Large and moderate deviations principles for kernel estimators of the multivariate regression

Abdelkader Mokkadem, Mariane Pelletier, Baba Thiam

Large and moderate deviations principles for kernel estimators of the multivariate regression

Abdelkader Mokkadem Mariane Pelletier Baba Thiam

(mokkadem, pelletier, thiam)@math.uvsq.fr

Université de Versailles-Saint-Quentin
Département de Mathématiques
45, Avenue des Etats-Unis
78035 Versailles Cedex
France

Abstract : In this paper, we prove large deviations principle for the Nadaraya-Watson estimator and for the semi-recursive kernel estimator of the regression in the multidimensional case. Under suitable conditions, we show that the rate function is a good rate function. We thus generalize the results already obtained in the unidimensional case for the Nadaraya-Watson estimator. Moreover, we give a moderate deviations principle for these two estimators. It turns out that the rate function obtained in the moderate deviations principle for the semi-recursive estimator is larger than the one obtained for the Nadaraya-Watson estimator.

AMS Subj. Classification: 62G08, 60F10
Key words and phrases :
Nadaraya-Watson estimator ; Recursive kernel estimator ; Large deviations principle ; Moderate deviations principle

1 Introduction

Let (X,Y), $(X_1,Y_1),\ldots,(X_n,Y_n)$ be a sequence of independent and identically distributed $\mathbb{R}^d\times\mathbb{R}^q$-valued random variables with probability density $f(x,y)$ with $\mathbb{E}|Y|<\infty$. Moreover, let $g(x)$ be the marginal density of X and $r(x)=\mathbb{E}\left(Y|X=x\right)=m(x)/g(x)$ the regression of Y on X. The purpose of this paper is to establish large and moderate deviations principles for the Nadaraya-Watson estimator and for the semi-recursive kernel estimator of the regression.

Let us first recall the concept of large and moderate deviations. A speed is a sequence (ν_n) of positive numbers going to infinity. A good rate function on $\mathbb{R}^m$ is a lower semicontinuous function $I:\mathbb{R}^m\to[0,\infty]$ such that, for each $\alpha<\infty$, the level set $\{x\in\mathbb{R}^m,\ \ I(x)\leq\alpha\}$ is a compact set. If the level sets of I are only closed, then I is said to be a rate function. A sequence $(Z_n)_{n\geq1}$ of $\mathbb{R}^m$-valued random variables is said to satisfy a large deviations principle (LDP) with speed (ν_n) and rate function I if:

$$\begin{aligned}\liminf_{n\to\infty}\nu_n^{-1}\log\mathbb{P}\left[Z_n\in U\right] &\geq -\inf_{x\in U}I(x)\ \text{ for every open subset U of } \mathbb{R}^m,\\ \limsup_{n\to\infty}\nu_n^{-1}\log\mathbb{P}\left[Z_n\in V\right] &\leq -\inf_{x\in V}I(x)\ \text{ for every closed subset V of } \mathbb{R}^m.\end{aligned}$$

Moreover, let (v_n) be a nonrandom sequence that goes to infinity; if (v_nZ_n) satisfies a LDP, then (Z_n) is said to satisfy a moderate deviations principle (MDP).

The Nadaraya-Watson estimator ([15], [20]) of the regression function $r(x)$ is defined by

$$r_n(x)=\begin{cases}\dfrac{m_n(x)}{g_n(x)} & \text{if } g_n(x)\neq0\\ 0 & \text{otherwise,}\end{cases}\tag{1}$$

with

$$m_n(x)=\frac{1}{nh_n^d}\sum_{i=1}^nY_iK\left(\frac{x-X_i}{h_n}\right)\quad\text{and}\quad g_n(x)=\frac{1}{nh_n^d}\sum_{i=1}^nK\left(\frac{x-X_i}{h_n}\right),$$

where the bandwidth (h_n) is a positive sequence such that

$$\lim_{n\to\infty}h_n=0\quad\text{and}\quad\lim_{n\to\infty}nh_n^d=\infty,\tag{2}$$

and the kernel K a continuous function such that $\lim_{\|x\|\to\infty}K(x)=0$ and $\int_{\mathbb{R}^d}K(x)dx=1$. The weak and strong consistency of r_n has been widely discussed by many authors; let us cite, among many others, Collomb [4], Collomb and Härdle [5], Devroye [7], Mack and Silverman [12] and Senoussi [19]. For other works on the consistency of r_n, the reader is refered to the monographs of Bosq [3] and Prakasa Rao [16]. The large deviations behaviour of r_n has been studied at first by Louani [11], and then by Joutard [10] in the univariate framework. Moderate deviations principles have been obtained by Worms [21] in the particular case $Y=r(X)+\varepsilon$ with ε and X independent. The first aim of this paper is to generalize these large and moderate deviations results.

The approach used by Louani [11] and Joutard [10] to study the large deviations behaviour of r_n is to note that, if $d=q=1$ and if the kernel is positive, then, for all $\delta>0$,

$$\mathbb{P}\left[r_n(x)-r(x)\geq\delta\right]\ =\ \mathbb{P}\left[\frac{1}{nh_n}\sum_{j=1}^n\left[Y_j-r(x)-\delta\right]K\left(\frac{x-X_j}{h_n}\right)\geq0\right].$$

Obviously, their approach can not be extended to the multivariate framework. Thus, to study the large deviations behaviour of r_n, our approach is totally different. We first establish a large deviations principle for the sequence $(m_n(x), g_n(x))$, and then show how the large deviations behaviour of r_n can be deduced. More precisely, for $x \in \mathbb{R}^d$, let Ψ_x be the function defined for any $(u, v) \in \mathbb{R}^q \times \mathbb{R}$ by

$$\Psi_x(u,v) \quad = \quad \int_{\mathbb{R}^d \times \mathbb{R}^q} \left(e^{(\langle u,y\rangle + v)K(z)} - 1 \right) f(x,y) dz dy,$$

(where $\langle u, y\rangle$ denotes the scalar product of u and y) and let I_x be the Fenchel-Legendre transform of Ψ_x. We give conditions ensuring that the sequence $(r_n(x))$ satisfies a LDP with speed (nh_n^d) and good rate function J defined, for any $s \in \mathbb{R}^q$, by

$$J(s) = \inf_{t \in \mathbb{R}} I_x(st, t).$$

Concerning the moderate deviations behaviour of the Nadaraya-Watson estimator, we prove that, for any positive sequence (v_n) such that

$$\lim_{n\to\infty} v_n = \infty, \quad \lim_{n\to\infty} \frac{v_n^2}{nh_n^d} = 0, \quad \text{and} \quad \lim_{n\to\infty} v_n h_n^p = 0, \tag{3}$$

(where p denotes the order of the kernel K) the sequence $(v_n [r_n(x) - r(x)])$ satisfies a LDP with speed $\left(nh_n^d / v_n^2\right)$ and good rate function G_x defined for all $v \in \mathbb{R}^q$ by

$$G_x(v) \quad = \quad \frac{g(x)}{2\int_{\mathbb{R}^d} K^2(z) dz} v^T \Sigma_x^{-1} v, \tag{4}$$

where Σ_x denotes the $q \times q$ covariance matrix $V(Y|X = x)$. Let us note that, in the case the model $Y = r(X) + \varepsilon$ (with X and ε independent) is considered, the matrix Σ_x is the covariance matrix of ε and does depend on x; we then find the MDP proved in Worms [21] again.

A semi-recursive version of the Nadaraya-Watson estimator (1) is defined as

$$\tilde{r}_n(x) = \begin{cases} \dfrac{\tilde{m}_n(x)}{\tilde{g}_n(x)} & \text{if } \tilde{g}_n(x) \neq 0 \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

where

$$\tilde{m}_n(x) = \frac{1}{n} \sum_{i=1}^n \frac{Y_i}{h_i^d} K\left(\frac{x - X_i}{h_i}\right) \quad \text{and} \quad \tilde{g}_n(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_i^d} K\left(\frac{x - X_i}{h_i}\right).$$

Weak conditions for various forms of consistency of $\tilde{r}_n$ have been obtained by Ahmad and Lin [1] and Devroye and Wagner [8]. Roussas [18] studied its almost sure convergence rate. The second aim of this paper is to establish the large and moderate deviations behaviour of $\tilde{r}_n$.

It turns out that the rate function that appears in the LDP is much more complex to explicit in the case the semi-recursive kernel regression estimator is considered than in the case the Nadaraya-Watson estimator is used. That is the reason why we only consider bandwidths defined as $(h_n) = (cn^{-a})$ with $c > 0$ and $a \in]0, 1/d[$ (instead of bandwidths satisfying (2)). For $x \in \mathbb{R}^d$, let $\tilde{\Psi}_{a,x}$ be the function defined for all $(u, v) \in \mathbb{R}^q \times \mathbb{R}$ by

$$\tilde{\Psi}_{a,x}(u,v) \quad = \quad \int_{[0,1]\times\mathbb{R}^d\times\mathbb{R}^q} s^{-ad} \left(e^{s^{ad}(\langle u,y\rangle + v)K(z)} - 1 \right) f(x,y) ds dz dy,$$

and let $\tilde{I}_{a,x}$ be the Fenchel-Legendre transform of $\tilde{\Psi}_{a,x}$. We give conditions ensuring that the sequence $(\tilde{r}_n(x))$ satisfies a LDP with speed (nh_n^d) and good rate function $\tilde{J}_a$ defined, for any $s \in \mathbb{R}^q$, by

$$\tilde{J}_a(s) = \inf_{t \in \mathbb{R}} \tilde{I}_{a,x}(st, t).$$

To establish the moderate deviations behaviour of $\tilde{r}_n$, we consider bandwidths (h_n) which vary regularly with exponent $(-a)$, $a \in]0, 1/d[$. We prove that, for any positive sequence (v_n) satisfying (3), the sequence $(v_n [\tilde{r}_n(x) - r(x)])$ satisfies a LDP with speed $\left(nh_n^d / v_n^2\right)$ and good rate function defined for all $v \in \mathbb{R}^q$ by

$$\tilde{G}_{a,x}(v) \quad = \quad \frac{(1+ad)g(x)}{2\int_{\mathbb{R}^d} K^2(z)dz} v^T \Sigma_x^{-1} v. \tag{6}$$

Let us underline that, because of the factor $(1 + ad)$ which is present in (6) but not in (4), the rate function obtained in the MDP in the case the semi-recursive estimator is used is larger than the one which appears in the case the Nadaraya-Watson kernel estimator is considered; this means that the semi-recursive estimator $\tilde{r}_n(x)$ is more concentrated around $r(x)$ than the Nadaraya-Watson estimator.

Our main results are stated in Section 2, whereas Section 3 is devoted to the proofs.

2 Assumptions and Main Results

We shall use the following notations.

- $\mathcal{D}(\mathcal{F}) = \{x,\ \mathcal{F}(x) < \infty\}$ denotes the domain of a function $\mathcal{F}$ and $\overset{\circ}{\mathcal{D}}(\mathcal{F})$ is the interior domain of $\mathcal{F}$.
- $\|x\|$ is the euclidean norm of x.
- λ is the Lebesgue measure.
- $a \wedge b = \min\{a, b\}$.
- $\vec{0} = (0, \ldots, 0) \in \mathbb{R}^q$.

The large and moderate deviations behaviours of the Nadaraya-Watson estimator r_n are given in Section 2.1, whereas the ones of the semi-recursive kernel estimator $\tilde{r}_n$ are stated in Section 2.2.

2.1 Large and moderate deviations principles for the Nadaraya-Watson estimator

The assumptions required for the LDP of the Nadaraya-Watson estimator are the following.

(A1) $K : \mathbb{R}^d \to \mathbb{R}$ is a bounded and integrable function, $\int_{\mathbb{R}^d} K(z)dz = 1$ and $\lim_{\|z\| \to \infty} K(z) = 0$.

(A2) For any $u \in \mathbb{R}^q$, $t \mapsto \displaystyle\int_{\mathbb{R}^q} e^{\langle u, y \rangle} f(t, y)dy$ is continuous at x and bounded.

Comments

- Notice that (A2) implies that the density g is continuous at x and bounded.
- In the model $Y = r(X) + \varepsilon$ with ε and X independent, let h be the probability density of ε. Then

$$\begin{aligned} f(t,y) &= g(t)h\left(y - r(t)\right) \\ \int_{\mathbb{R}^q} \|y\| f(t,y)dy &= g(t)\int_{\mathbb{R}^q} \|y + r(t)\| h(y)dy \\ \int_{\mathbb{R}^q} e^{\langle u,y\rangle} f(t,y)dy &= g(t)e^{\langle u,r(t)\rangle}\int_{\mathbb{R}^q} e^{\langle u,y\rangle} h(y)dy. \end{aligned}$$

  Thus, (A2) can be translated as assumptions on g and r and on the moments of ε.
- As it can be seen from the proofs, the boundness assumption in (A2) is useless if K has a compact support.
- The boundness of the function $t \mapsto \int_{\mathbb{R}^q} e^{\langle u,y\rangle} f(t,y)dy$ for any $u \in \mathbb{R}^q$ implies that

$$\forall m \geq 0, \forall \rho \geq 0 \ \text{ the function } \ t \mapsto \int_{\mathbb{R}^q} \|y\|^m e^{\rho\|y\|} f(t,y)dy \quad \text{is bounded.} \tag{7}$$

Proof It suffices to prove that the function $t \mapsto \int_{\mathbb{R}^q} e^{\rho\|y\|} f(t,y)dy$ is bounded for any $\rho > 0$. Set $y = (y_1, \ldots, y_q)$, we first note that

$$\begin{aligned} \int_{\mathbb{R}^q} e^{q\rho|y_j|} f(t,y)dy &\leq \int_{\{y_j \geq 0\}} e^{q\rho y_j} f(t,y)dy + \int_{\{y_j<0\}} e^{-q\rho y_j} f(t,y)dy \\ &\leq \int_{\mathbb{R}^q} e^{q\rho y_j} f(t,y)dy + \int_{\mathbb{R}^q} e^{-q\rho y_j} f(t,y)dy. \end{aligned}$$

Now, we have

$$\begin{aligned} &\int_{\mathbb{R}^q} e^{\rho\|y\|} f(t,y)dy \\ &\leq \int_{\mathbb{R}^q} e^{\rho|y_1| + \cdots + \rho|y_q|} f(t,y)dy \\ &\leq \left(\int_{\mathbb{R}^q} e^{q\rho|y_1|} f(t,y)dy \ldots \int_{\mathbb{R}^q} e^{q\rho|y_q|} f(t,y)dy\right)^{\frac{1}{q}} \quad \text{by the generalized Hölder inequality.} \\ &\leq \left(\left(\int_{\mathbb{R}^q} e^{q\rho y_1} f(t,y)dy + \int_{\mathbb{R}^q} e^{-q\rho y_1} f(t,y)dy\right) \ldots \left(\int_{\mathbb{R}^q} e^{q\rho y_q} f(t,y)dy + \int_{\mathbb{R}^q} e^{-q\rho y_q} f(t,y)dy\right)\right)^{\frac{1}{q}} \end{aligned}$$

which is bounded.

Before stating our results, we need to introduce the rate function for the LDP of the Nadaraya-Watson estimator. Let $\Psi_x : \mathbb{R}^q \times \mathbb{R} \to \mathbb{R}$ and $I_x, \hat{I}_x : \mathbb{R}^q \times \mathbb{R} \to \mathbb{R}$ be the functions defined as follows:

$$\Psi_x(u,v) = \int_{\mathbb{R}^d \times \mathbb{R}^q} \left(e^{(\langle u,y\rangle + v)K(z)} - 1\right) f(x,y)dzdy, \tag{8}$$

$$I_x(t_1,t_2) = \sup_{(u,v)\in\mathbb{R}^q\times\mathbb{R}} \left\{\langle u,t_1\rangle + vt_2 - \Psi_x(u,v)\right\}, \tag{9}$$

$$\hat{I}_x(s,t) = I_x(st,t). \tag{10}$$

Moreover, for any $s \in \mathbb{R}^q$, set

$$\begin{aligned} J^*(s) &= \inf_{t\in\mathbb{R}^*} I_x(st,t) \\ &= \inf_{t\in\mathbb{R}^*} \hat{I}_x(s,t), \\ J(s) &= J^*(s) \wedge I_x(\vec{0},0) \\ &= \inf_{t\in\mathbb{R}} \hat{I}_x(s,t). \end{aligned}$$

To prove that J is a rate function, we need to assume that the following condition (C) is fulfilled.

(C) $\inf_{s\in\mathbb{R}^q} I_x(s,0) = I_x(\vec{0},0)$.

Before stating the properties of the function J, let us give some cases when Condition (C) is satisfied (under Assumptions (A1) and (A2)).

Example 1: Nonnegative kernel
Condition (C) is satisfied when K is nonnegative since, in this case, $I_x(s,0) = +\infty$ for any $s \neq \vec{0}$, (this is stated in Proposition 3 of Section 3).

Example 2: Model with symmetry
Condition (C) holds when f is symmetric in each coordinate of the second variable $y \in \mathbb{R}^q$. As a matter of fact, for a diagonal $q \times q$ matrix A such that $A_{ii} = \pm 1$, observe that

$$\begin{aligned} \Psi_x(Au,v) &= \int_{\mathbb{R}^d\times\mathbb{R}^q} \left(e^{(\langle Au,y\rangle+v)K(z)} - 1\right) f(x,y)dzdy \\ &= \int_{\mathbb{R}^d\times\mathbb{R}^q} \left(e^{(\langle u,Ay\rangle+v)K(z)} - 1\right) f(x,y)dzdy \\ &= \int_{\mathbb{R}^d\times\mathbb{R}^q} \left(e^{(\langle u,y'\rangle+v)K(z)} - 1\right) f(x,A^{-1}y')dzdy' \\ &= \Psi_x(u,v). \end{aligned}$$

For any given $s \in \mathbb{R}^q$, set

$$\mathcal{U}_s = \{u \in \mathbb{R}^q,\ \langle u,s\rangle \geq 0\}.$$

We have,

$$\sup_{u,v}(-\Psi_x(u,v)) = \sup_{u\in\mathcal{U}_s,\ v\in\mathbb{R}} (-\Psi_x(u,v)).$$

Now, for any $u \in \mathcal{U}_s$ and $v \in \mathbb{R}$,

$$\langle u,s\rangle - \Psi_x(u,v) \geq -\Psi_x(u,v),$$

so that

$$\sup_{u,v}\{\langle u,s\rangle - \Psi_x(u,v)\} \geq \sup_{u\in\mathcal{U}_s,\ v\in\mathbb{R}} (-\Psi_x(u,v)),$$

and thus,

$$I_x(\vec{0},0) \leq I_x(s,0) \quad \forall s \in \mathbb{R}^q,$$

so that Condition (C) follows.

Example 3: A negative kernel without symmetry assumption on f, and for $d = q = 1$
If the kernel K can be written as $K = \mathbb{1}_D - \mathbb{1}_{D'}$ where D and D' are two subsets of $\mathbb{R}$ such that $D \cap D' = \emptyset$ and $\lambda(D) - \lambda(D') = 1$, then Condition (C) holds. As a matter of fact, we then have

$$\begin{aligned}
\Psi_x(u,v) &= \int_{\mathbb{R}\times\mathbb{R}} \left(e^{(uy+v)K(z)} - 1\right) f(x,y)dzdy \\
&= \int_{D\times\mathbb{R}} \left(e^{uy+v} - 1\right) f(x,y)dzdy + \int_{D'\times\mathbb{R}} \left(e^{-uy-v} - 1\right) f(x,y)dzdy \\
&= e^v\lambda(D)\int_{\mathbb{R}} e^{uy} f(x,y)dy - \left[\lambda(D) + \lambda(D')\right] g(x) + e^{-v}\lambda(D')\int_{\mathbb{R}} e^{-uy} f(x,y)dy.
\end{aligned}$$

Now, let M_x denote the Laplace transform of $f(x,\cdot)$, then

$$\Psi_x(u,v) = e^v\lambda(D)M_x(u) + e^{-v}\lambda(D')M_x(-u) - \left[\lambda(D) + \lambda(D')\right] g(x).$$

For any given u, it can easily be seen that the infimum of $\Psi_x(u,\cdot)$ is reached at

$$v_0 = \log\sqrt{\frac{\lambda(D')M_x(-u)}{\lambda(D)M_x(u)}},$$

and

$$\Psi_x(u,v_0) = 2\sqrt{\lambda(D)\lambda(D')}\sqrt{M_x(u)M_x(-u)} - \left[\lambda(D) + \lambda(D')\right] g(x).$$

Observe that

$$\Psi_x(u,v_0) = \Psi_x(-u,v_0),$$

and thus

$$\begin{aligned}
\sup_u\left(-\Psi_x(u,v_0)\right) &= \sup_{u\geq 0}\left(-\Psi_x(u,v_0)\right) \\
&= \sup_{u\leq 0}\left(-\Psi_x(u,v_0)\right) \\
&= I_x(0,0).
\end{aligned}$$

Now, if $s \geq 0$, we have for any $u \geq 0$

$$us - \Psi_x(u,v_0) \geq -\Psi_x(u,v_0),$$

and thus

$$I_x(s,0) \geq I_x(0,0)\ \forall s \geq 0.$$

Proceeding in the same way for $s < 0$, we obtain Condition (C).
Such an example of a four order kernel is $K = \mathbb{1}_{[-a,a]} - \mathbb{1}_{[-b,-a[\cup]a,b]}$, with

$$\begin{aligned}
a &= \frac{1}{6}\sqrt[3]{2} + \frac{1}{12}\left(\sqrt[3]{2}\right)^2 + \frac{1}{3} \\
b &= \frac{1}{3}\sqrt[3]{2} + \frac{1}{6}\left(\sqrt[3]{2}\right)^2 + \frac{1}{6}.
\end{aligned}$$

Let us now give the properties of the function J.

Proposition 1 *Assume that (A1), (A2) and (C) hold. Then,*

(i) J is a rate function on $\mathbb{R}^q$. More precisely, for $\alpha \in \mathbb{R}$,

- *if $\alpha < I_x(\vec{0}, 0)$, then $\{J(s) \leq \alpha\}$ is compact.*
- *if $\alpha \geq I_x(\vec{0}, 0)$, then $\{J(s) \leq \alpha\} = \mathbb{R}^q$.*

(ii) If $I_x(\vec{0}, 0) = \infty$, then J is a good rate function on $\mathbb{R}^q$ and $J = J^*$.

(iii) If $J^*(s) < \infty$, then $J(s) = J^*(s)$.

(iv) If $\alpha < I_x(\vec{0}, 0)$, then $\{J^*(s) \leq \alpha\} = \{J(s) \leq \alpha\}$.

Remark 1 *In view of the definition of J and J^*, and of Proposition 1 (iii), we have:*

$$J(s) \quad = \quad \begin{cases} J^*(s) & \text{if } J^*(s) < \infty \\ I_x(\vec{0}, 0) & \text{if } J^*(s) = \infty. \end{cases}$$

Let us now state the LDP for the Nadaraya-Watson estimator.

Theorem 1 ***(Pointwise LDP for the Nadaraya-Watson estimator)***
Assume that (A1), (A2) and (C) hold, and that (h_n) satisfies the conditions in (2). Then, for any open subset U of $\mathbb{R}^q$,

$$\liminf_{n\to\infty} \frac{1}{nh_n^d} \log \mathbb{P}\left[r_n(x) \in U\right] \geq -\inf_{s\in U} J^*(s),$$

and for any closed subset V of $\mathbb{R}^q$,

$$\limsup_{n\to\infty} \frac{1}{nh_n^d} \log \mathbb{P}\left[r_n(x) \in V\right] \leq -\inf_{s\in V} J(s).$$

Comments.

1) Set $E = \{J^*(s) < \infty\}$. For any open subset U of $\mathbb{R}^q$ such that $U \cap E \neq \emptyset$, we have

$$\liminf_{n\to\infty} \frac{1}{nh_n^d} \log \mathbb{P}\left[r_n(x) \in U\right] \geq -\inf_{s\in U} J(s).$$

2) If I_x is finite in a neighbourhood of $(\vec{0}, 0)$, then J^* is finite everywhere and by Proposition 1 (iii), $J(s) = J^*(s) < \infty \ \ \forall s$; thus (r_n) satisfies a LDP with speed (nh_n^d) and rate function J. Of course, this does not hold for nonnegative kernel since in this case $I_x(s, 0) = +\infty$ for any s (see Proposition 3 in Section 3). However, it can hold for kernels which take negative values. For example, consider the previous Example 3, and assume $f(x, y)$ is symmetric in y ; in this case $M_x(u) = M_x(-u)$. The equation

$$\frac{\partial \Psi_x}{\partial v}(u, v) = \left[\lambda(D)e^v - \lambda(D')e^{-v}\right] M_x(u) = 0$$

has solution $v_0 = \log \sqrt{\dfrac{\lambda(D')}{\lambda(D)}}$ independent from u. Moreover, M' is continuous and has range $\mathbb{R}$, thus, there exists u_0 such that $M'(u_0) = 0$. This implies that the equation

$$\frac{\partial \Psi_x}{\partial u}(u, v) = \left[\lambda(D)e^v + \lambda(D')e^{-v}\right] M_x'(u) = 0$$

has a solution u_0 independent from v. Thus $(0, 0)$ is in the range of $\nabla \Psi_x$. It follows from Proposition 3 Section 3 that I_x is finite in a neighbourhood of $(0, 0)$.

3) When $I_x(\vec{0},0)=\infty$, it follows from Proposition 1 and Theorem 1 that (r_n) satisfies a LDP with speed (nh_n^d) and good rate function J.

In the case K is a nonnegative kernel whose support has an infinity measure, we will show in Proposition 3 that $I_x(\vec{0},0)=\infty$. We have thus the following corollary.

Corollary 1 *Let the assumptions of Theorem 1 hold. If K is a nonnegative kernel such that $\lambda\left(\left\{x\in\mathbb{R}^d, K(x)>0\right\}\right)=\infty$, then the sequence (r_n) satisfies a LDP with speed (nh_n^d) and good rate function J.*

This corollary is an extension of the results of Louani [11] and Joutard [10] to the multivariate framework (and to the case the kernel K may vanish). Moreover, it proves that the rate function that appears in their large deviations results is in fact a good rate function.

To establish pointwise MDP for the Nadaraya-Watson estimator, we need the following additionnal assumptions.

(A3) For any $u\in\mathbb{R}^q$, $t\mapsto\int_{\mathbb{R}^q}\langle u,y\rangle^2 f(t,y)dy$ and $t\mapsto\int_{\mathbb{R}^q}\langle u,y\rangle f(t,y)dy$ are continuous at x and $g(x)\neq 0$.

(A4) $\lim_{n\to\infty} v_n=\infty$ and $\lim_{n\to\infty}\dfrac{nh_n^d}{v_n^2}=\infty$.

(A5) i) There exists an integer $p\geq 2$ such that $\forall s\in\{1,\dots,p-1\},\forall j\in\{1,\dots,d\}$, $\int_{\mathbb{R}^d} y_j^s K(y)dy_j=0$, and $\int_{\mathbb{R}^d}\left|y_j^p K(y)\right|dy<\infty$.
ii) $\lim_{n\to\infty} v_n h_n^p=0$.
iii) m and g are p-times differentiable on $\mathbb{R}^d$, and their differentials of order p are bounded and continuous at x.

We can now state the MDP for the Nadaraya-Watson estimator.

Theorem 2 ***(Pointwise MDP for the Nadaraya-Watson kernel estimator of the regression)*** *Assume that (A1)-(A5) hold. Then, the sequence $(v_n(r_n(x)-r(x)))$ satisfies a LDP with speed $\left(\dfrac{nh_n^d}{v_n^2}\right)$ and good rate function G_x defined in (4).*

2.2 Large and moderate deviations principles for the semi-recursive estimator

For $a\in]0,1/d[$, let $\tilde{\Psi}_{a,x}:\mathbb{R}^q\times\mathbb{R}\to\mathbb{R}$ and $\tilde{I}_{a,x}:\mathbb{R}^q\times\mathbb{R}\to\mathbb{R}$ be the functions defined as follows:

$$\tilde{\Psi}_{a,x}(u,v) = \int_{[0,1]\times\mathbb{R}^d\times\mathbb{R}^q} s^{-ad}\left(e^{s^{ad}(\langle u,y\rangle+v)K(z)}-1\right)f(x,y)dsdzdy, \tag{11}$$

$$\tilde{I}_{a,x}(t_1,t_2) = \sup_{(u,v)\in\mathbb{R}^q\times\mathbb{R}}\left\{\langle u,t_1\rangle+vt_2-\tilde{\Psi}_{a,x}(u,v)\right\}. \tag{12}$$

Moreover, let $\tilde{J}_a$ and $\tilde{J}_a^*$ be defined as follows: for any $s\in\mathbb{R}^q$,

$$\tilde{J}_a^*(s) = \inf_{t\in\mathbb{R}^*}\tilde{I}_{a,x}(st,t) \tag{13}$$

$$\tilde{J}_a(s) = \tilde{J}_a^*(s)\wedge\tilde{I}_{a,x}(\vec{0},0). \tag{14}$$

Let us give the following additionnal hypotheses.

(A'1) For any $u \in \mathbb{R}^q$, $t \mapsto \int_{\mathbb{R}^q} e^{\alpha\langle u,y\rangle} f(t,y)dy$ is continuous at x uniformly with respect to $\alpha \in [0,1]$.

Condition (C) above is substituted by the following one,

(C') $\inf_{s\in\mathbb{R}^q} \tilde{I}_{a,x}(s,0) = \tilde{I}_{a,x}(\vec{0},0)$.

Examples for which Condition (C') holds are Examples 1 and 2 given for (C). The following proposition gives the properties of the function $\tilde{J}_a$.

Proposition 2 *Assume that (A1), (A2), (A'1) and (C') hold. Then,*

(i) $\tilde{J}_a$ is a rate function on $\mathbb{R}^q$. More precisely, for $\alpha \in \mathbb{R}$,

- *if $\alpha < \tilde{I}_{a,x}(\vec{0},0)$, then $\left\{\tilde{J}_a(s) \le \alpha\right\}$ is compact.*
- *if $\alpha \ge \tilde{I}_{a,x}(\vec{0},0)$, then $\left\{\tilde{J}_a(s) \le \alpha\right\} = \mathbb{R}^q$.*

(ii) If $\tilde{I}_{a,x}(\vec{0},0) = \infty$, then $\tilde{J}_a$ is a good rate function on $\mathbb{R}^q$ and $\tilde{J}_a = \tilde{J}_a^*$.

(iii) If $\tilde{J}_a^*(s) < \infty$, then $\tilde{J}_a(s) = \tilde{J}_a^*(s)$.

(iv) If $\alpha < \tilde{I}_{a,x}(\vec{0},0)$, then $\left\{\tilde{J}_a^*(s) \le \alpha\right\} = \left\{\tilde{J}_a(s) \le \alpha\right\}$.

Notice that, like for J and J^*, we have

$$\tilde{J}_a(s) \quad = \quad \begin{cases} \tilde{J}_a^*(s) & \text{if } \tilde{J}_a^*(s) < \infty \\ \tilde{I}_{a,x}(\vec{0},0) & \text{if } \tilde{J}_a^*(s) = \infty. \end{cases}$$

We can now state the LDP for the semi-recursive kernel estimator of the regression.

Theorem 3 ***(Pointwise LDP for the semi-recursive estimator of the regression)*** *Set $(h_n) = (cn^{-a})$ with $c > 0$ and $0 < a < 1/d$, and let (A1), (A2), (A'1) and (C') hold. Then, for any open subset U of $\mathbb{R}^q$,*

$$\liminf_{n\to\infty} \frac{1}{nh_n^d} \log \mathbb{P}\left[\tilde{r}_n(x) \in U\right] \ge -\inf_{s\in U} \tilde{J}_a^*(s),$$

and for any closed subset V of $\mathbb{R}^q$,

$$\limsup_{n\to\infty} \frac{1}{nh_n^d} \log \mathbb{P}\left[\tilde{r}_n(x) \in V\right] \le -\inf_{s\in V} \tilde{J}_a(s).$$

The comments made for Theorem 1 are valid for Theorem 3. In particular, we have the following corollary.

Corollary 2 *Let the assumptions of Theorem 3 hold. If K is a nonnegative kernel such that $\lambda\left(\left\{x \in \mathbb{R}^d, K(x) > 0\right\}\right) = \infty$, then the sequence $(\tilde{r}_n)$ satisfies a LDP with speed (nh_n^d) and good rate function $\tilde{J}_a$.*

Before stating pointwise MDP for the semi-recursive estimator of the regression, let us recall that a sequence (u_n) is said to vary regularly with exponent α if there exists a function u which varies regularly with exponent α and such that $u_n = u(n)$ for all n (see, for example, Feller [9] page

275). We will use in the sequel the following property (see Bingham et al. [2] page 26). If (h_n) varies regularly with exponent $(-a)$ and if $\beta a < 1$, then

$$\lim_{n\to\infty} \frac{1}{nh_n^\beta} \sum_{i=1}^{n} h_i^\beta = \frac{1}{1-a\beta}. \tag{15}$$

We also consider the following condition.

$$\sup_{i\leq n} \frac{h_n}{h_i} < \infty. \tag{16}$$

(For example, this condition holds when h_n is nonincreasing).

Theorem 4 ***(Pointwise MDP for the semi-recursive kernel estimator of the regression)*** *Assume that (h_n) varies regularly with exponent $(-a)$ with $a \in]0, 1/d[$, and satisfies (16). Let (A1)-(A5) hold. Then, the sequence $(v_n(\tilde{r}_n(x) - r(x)))$ satisfies a LDP with speed $\left(\frac{nh_n^d}{v_n^2}\right)$ and good rate function $\tilde{G}_{a,x}$ defined in (6).*

3 Proofs

The proofs of the results for the Nadaraya-Watson kernel estimator are in many cases similar to those of the semi-recursive kernel estimator of the regression, so we omit some details of the proofs for this last one.

First, let us state the following propositions which give the properties of the functions Ψ_x, $\tilde{\Psi}_{a,x}$, I_x and $\tilde{I}_{a,x}$. Set

$$S_+ = \left\{x \in \mathbb{R}^d,\ K(x) > 0\right\} \quad \text{and} \quad S_- = \left\{x \in \mathbb{R}^d,\ K(x) < 0\right\}.$$

Proposition 3 ***(Properties of Ψ_x and I_x)***
Let Assumptions (A1) and (A2) hold. Then,

i) Ψ_x is strictly convex, continuously differentiable on $\mathbb{R}^q \times \mathbb{R}$, and I_x is a good rate function on $\mathbb{R}^q \times \mathbb{R}$.

ii) $\nabla\Psi_x$ is an open map and the range of Ψ_x is $\overset{\circ}{\mathcal{D}}(I_x)$. I_x is strictly convex on $\overset{\circ}{\mathcal{D}}(I_x)$ and for any $t \in \overset{\circ}{\mathcal{D}}(I_x) \subset \mathbb{R}^q \times \mathbb{R}$,

$$I_x(t) = \langle (\nabla\Psi_x)^{-1}(t), t\rangle - \Psi_x\left((\nabla\Psi_x)^{-1}(t)\right). \tag{17}$$

iii) If $\lambda(S_-) = 0$, then $I_x(\vec{0}, 0) = g(x)\lambda(S_+)$, and for any $t_1 \neq \vec{0}$, $I_x(t_1, 0) = +\infty$.

Proposition 4 ***(Properties of $\tilde{\Psi}_{a,x}$ and $\tilde{I}_{a,x}$)***
Let Assumptions (A1) and (A2) hold. Then,

i) $\tilde{\Psi}_{a,x}$ is strictly convex, continuously differentiable on $\mathbb{R}^q \times \mathbb{R}$, and $\tilde{I}_{a,x}$ is a good rate function on $\mathbb{R}^q \times \mathbb{R}$.

ii) $\nabla\tilde{\Psi}_{a,x}$ *is an open map and the range of* $\tilde{\Psi}_{a,x}$ *is* $\overset{\circ}{\mathcal{D}}(\tilde{I}_{a,x})$. $\tilde{I}_{a,x}$ *is strictly convex on* $\overset{\circ}{\mathcal{D}}(\tilde{I}_{a,x})$, *and for any* $t \in \overset{\circ}{\mathcal{D}}(\tilde{I}_{a,x}) \subset \mathbb{R}^q \times \mathbb{R}$,

$$\tilde{I}_{a,x}(t) = \langle (\nabla\tilde{\Psi}_{a,x})^{-1}(t), t\rangle - \tilde{\Psi}_{a,x}\left((\nabla\tilde{\Psi}_{a,x})^{-1}(t)\right). \tag{18}$$

iii) If $\lambda(S_-) = 0$, *then* $\tilde{I}_{a,x}(\vec{0}, 0) = g(x)\lambda(S_+)/(1-ad)$, *and for any* $t_1 \neq \vec{0}$, $\tilde{I}_{a,x}(t_1, 0) = +\infty$.

The two following lemmas are used for the proofs of Theorems 1 and 3.

Lemma 1 ***(Pointwise LDP for the sequence*** $(m_n(x), g_n(x))$***)***
Let Assumptions (A1) and (A2) hold. Then, the sequence $(m_n(x), g_n(x))$ *satisfies a LDP with speed* $\left(nh_n^d\right)$ *and rate function* I_x *defined in* (9).

Lemma 2 ***(Pointwise LDP for the sequence*** $(\tilde{m}_n(x), \tilde{g}_n(x))$***)***
Set $h_n = cn^{-a}$ *with* $c > 0$ *and* $a \in]0, 1/d[$, *and let Assumptions (A1), (A2) and (A'1) hold. Then, the sequence* $(\tilde{m}_n(x), \tilde{g}_n(x))$ *satisfies a LDP with speed* $\left(nh_n^d\right)$ *and rate function* $\tilde{I}_{a,x}$ *defined in* (12).

Our proofs are now organized as follows. Lemmas 1 and 2 are proved in Section 3.1, Theorems 1 and 3 in Section 3.2, Theorem 2 in Section 3.3, Theorem 4 is proved in Section 3.4. Section 3.5 is devoted to the proof of Propositions 3 and 4 on the rate functions I_x and $I_{a,x}$. Propositions 1 and 2 are proved in Section 3.6.

3.1 Proof of Lemmas 1 and 2

3.1.1 Proof of Lemma 1

For any $w = (u, v) \in \mathbb{R}^q \times \mathbb{R}$, set

$$\begin{aligned} \Psi_n(x) &= (m_n(x), g_n(x)), \\ \Lambda_{n,x}(w) &= \frac{1}{nh_n^d} \log \mathbb{E}\left[\exp\left(nh_n^d \langle w, \Psi_n(x)\rangle\right)\right]. \end{aligned}$$

Let us at first assume that the following lemma holds.

Lemma 3 ***(Convergence of*** $\Lambda_{n,x}$***)***
Assume that (A1) and (A2) hold, then

$$\lim_{n\to\infty} \Lambda_{n,x}(u, v) = \Psi_x(u, v), \tag{19}$$

where Ψ_x *is defined in* (8).

To prove Lemma 1, we apply Proposition 3, Lemma 3 and the Gärtner-Ellis Theorem (see Dembo and Zeitouni [6]). Proposition 3 ensures that Ψ_x is essentially smooth, lower semicontinuous function so that Lemma 1 follows from the Gärtner-Ellis Theorem.
Let us now prove Lemma 3. Set

$$Z_i = [\langle u, Y_i\rangle + v]\, K\left(\frac{x - X_i}{h_n}\right).$$

For any $(u, v) \in \mathbb{R}^q \times \mathbb{R}$, we have

$$\Lambda_{n,x}(u, v) = \frac{1}{nh_n^d} \log \mathbb{E}\left[\exp\left(\sum_{i=1}^n Z_i\right)\right],$$

and, since the random vectors (X_i, Y_i), $i = 1, \ldots, n$ are independent and identically distributed, we get

$$\Lambda_{n,x}(u,v) = \frac{1}{h_n^d} \log \mathbb{E}\left[e^{Z_n}\right].$$

A Taylor's expansion implies that there exists c_n between 1 and $\mathbb{E}\left[e^{Z_n}\right]$ such that

$$\begin{aligned}
\Lambda_{n,x}(u,v) &= \frac{1}{h_n^d}\mathbb{E}\left[e^{Z_n}-1\right] - \frac{1}{2c_n^2 h_n^d}\left(\mathbb{E}\left[e^{Z_n}-1\right]\right)^2 \\
&= \frac{1}{h_n^d}\int_{\mathbb{R}^d\times\mathbb{R}^q}\left[e^{(\langle u,y\rangle+v)K\left(\frac{x-s}{h_n}\right)}-1\right] f(s,y)dsdy - R_{n,x}^{(1)}(u,v) \\
&= \Psi_x(u,v) - R_{n,x}^{(1)}(u,v) + R_{n,x}^{(2)}(u,v),
\end{aligned}$$

with

$$\begin{aligned}
R_{n,x}^{(1)}(u,v) &= \frac{1}{2c_n^2 h_n^d}\left(\mathbb{E}\left[e^{Z_n}-1\right]\right)^2, \\
R_{n,x}^{(2)}(u,v) &= \int_{\mathbb{R}^d\times\mathbb{R}^q}\left[e^{(\langle u,y\rangle+v)K(z)}-1\right]\left[f(x-h_n z,y)-f(x,y)\right]dzdy.
\end{aligned}$$

Let us prove that

$$\lim_{n\to\infty} R_{n,x}^{(2)}(u,v) = 0. \tag{20}$$

Set $A > 0$ and $\epsilon > 0$; we then have

$$\begin{aligned}
R_{n,x}^{(2)}(u,v) &= \int_{\{\|z\|\le A\}\times\mathbb{R}^q}\left[e^{(\langle u,y\rangle+v)K(z)}-1\right]\left[f(x-h_n z,y)-f(x,y)\right]dzdy \\
&\quad + \int_{\{\|z\|>A\}\times\mathbb{R}^q}\left[e^{(\langle u,y\rangle+v)K(z)}-1\right]\left[f(x-h_n z,y)-f(x,y)\right]dzdy. \qquad (21)
\end{aligned}$$

Next, since for any $t \in \mathbb{R}$, $\left|e^t - 1\right| \le |t|e^{|t|}$, we have

$$\begin{aligned}
&\int_{\{\|z\|>A\}\times\mathbb{R}^q}\left|e^{(\langle u,y\rangle+v)K(z)}-1\right|\left|f(x-h_n z,y)-f(x,y)\right|dzdy \\
&\le \int_{\{\|z\|>A\}\times\mathbb{R}^q}|\langle u,y\rangle+v|\,|K(z)|\,e^{|\langle u,y\rangle+v||K(z)|}\left|f(x-h_n z,y)-f(x,y)\right|dzdy \\
&\le \int_{\{\|z\|>A\}\times\mathbb{R}^q}|\langle u,y\rangle+v|\,|K(z)|\,e^{|\langle u,y\rangle+v||K(z)|}f(x-h_n z,y)dzdy \\
&\quad + \int_{\{\|z\|>A\}\times\mathbb{R}^q}|\langle u,y\rangle+v|\,|K(z)|\,e^{|\langle u,y\rangle+v||K(z)|}f(x,y)dzdy \\
&\le e^{|v|\|K\|_\infty}\|u\|\int_{\{\|z\|>A\}}|K(z)|\left[\int_{\mathbb{R}^q}\|y\|e^{\|K\|_\infty\|u\|\|y\|}f(x-h_n z,y)dy\right]dz \\
&\quad + e^{|v|\|K\|_\infty}|v|\int_{\{\|z\|>A\}}|K(z)|\left[\int_{\mathbb{R}^q}e^{\|K\|_\infty\|u\|\|y\|}f(x-h_n z,y)dy\right]dz \\
&\quad + e^{|v|\|K\|_\infty}\|u\|\int_{\{\|z\|>A\}}|K(z)|\,dz\int_{\mathbb{R}^q}\|y\|e^{\|K\|_\infty\|u\|\|y\|}f(x,y)dy \\
&\quad + e^{|v|\|K\|_\infty}|v|\int_{\{\|z\|>A\}}|K(z)|\,dz\int_{\mathbb{R}^q}e^{\|K\|_\infty\|u\|\|y\|}f(x,y)dy \\
&\le B\int_{\{\|z\|>A\}}|K(z)|\,dz, \qquad (22)
\end{aligned}$$

where B is a constant ; this last inequality follows from (7) and from the fact that K is bounded. Now, since K is integrable, we can choose A such that

$$\int_{\{\|z\|>A\}\times\mathbb{R}^q} \left|e^{(\langle u,y\rangle+v)K(z)}-1\right| |f(x-h_nz,y)-f(x,y)|\,dzdy \leq \frac{\epsilon}{2}. \tag{23}$$

Now, observe that

$$\begin{aligned}
&\int_{\{\|z\|\leq A\}\times\mathbb{R}^q} \left[e^{(\langle u,y\rangle+v)K(z)}-1\right][f(x-h_nz,y)-f(x,y)]\,dzdy \\
&= \int_{\{\|z\|\leq A\}} e^{vK(z)}\left[\int_{\mathbb{R}^q} e^{\langle u,y\rangle K(z)}\left(f(x-h_nz,y)-f(x,y)\right)dy\right]dz \qquad (24)\\
&\quad -\int_{\{\|z\|\leq A\}}\left[\int_{\mathbb{R}^q}\left(f(x-h_nz,y)-f(x,y)\right)dy\right]dz. \qquad (25)
\end{aligned}$$

Assumption (A2) together with (7), and the dominated convergence theorem ensure that both integrals in (24) and (25) converge to 0. We deduce that for n large enough,

$$\left|\int_{\{\|z\|\leq A\}\times\mathbb{R}^q}\left[e^{(\langle u,y\rangle+v)K(z)}-1\right][f(x-h_nz,y)-f(x,y)]\,dzdy\right| \leq \frac{\epsilon}{2}, \tag{26}$$

so that (20) follows from (23) and (26).
Let us now consider $R^{(1)}_{n,x}$; since c_n is between 1 and $\mathbb{E}\left[e^{Z_n}\right]$, we get

$$\frac{1}{c_n} \leq \max\left\{1, \frac{1}{\mathbb{E}\left(e^{Z_n}\right)}\right\}.$$

By Jensen's inequality, we obtain

$$\frac{1}{\mathbb{E}\left[e^{Z_n}\right]} \leq \frac{1}{e^{\mathbb{E}[Z_n]}}.$$

Observe that

$$\begin{aligned}
|\mathbb{E}\left(Z_n\right)| &= \left|\mathbb{E}\left[\left(\langle u,Y_n\rangle+v\right)K\left(\frac{x-X_n}{h_n}\right)\right]\right| \\
&\leq \int_{\mathbb{R}^d\times\mathbb{R}^q}\|u\|\|y\|\left|K\left(\frac{x-s}{h_n}\right)\right| f(s,y)dsdy + |v|\int_{\mathbb{R}^d\times\mathbb{R}^q}\left|K\left(\frac{x-s}{h_n}\right)\right| f(s,y)dsdy \\
&\leq h_n^d\left(\|u\|\int_{\mathbb{R}^d}|K(z)|\left[\int_{\mathbb{R}^q}\|y\|f(x-h_nz,y)dy\right]dz + |v|\int_{\mathbb{R}^d}|K(z)|\left[\int_{\mathbb{R}^q}f(x-h_nz,y)dy\right]dz\right),
\end{aligned}$$

which goes to 0 in view of (7) and since $\lim_{n\to\infty}h_n=0$. We deduce that there exists $c\in\mathbb{R}^*_+$ such that

$$\frac{1}{c_n^2}\leq c.$$

Noting that by (7),

$$\begin{aligned}
\mathbb{E}\left|e^{Z_n}-1\right| &\leq h_n^d\int_{\mathbb{R}^d\times\mathbb{R}^q}|\langle u,y\rangle+v|\,|K(z)|e^{|\langle u,y\rangle+v||K(z)|}f(x-h_nz,y)dydz \\
&\leq Bh_n^d,
\end{aligned}$$

where B is a constant. It follows that

$$\lim_{n\to\infty}R^{(1)}_{n,x}(u,v)=0,$$

which proves Lemma 3. ■

3.1.2 Proof of Lemma 2

Similarly as the proof of Lemma 1, for any $w = (u, v) \in \mathbb{R}^q \times \mathbb{R}$, set

$$\begin{aligned} \tilde{\Psi}_n(x) &= (\tilde{m}_n(x), \tilde{g}_n(x)), \\ \tilde{\Lambda}_{n,x}(w) &= \frac{1}{nh_n^d} \log \mathbb{E}\left[\exp\left(nh_n^d \langle w, \tilde{\Psi}_n(x)\rangle\right)\right]. \end{aligned}$$

When $h_n = cn^{-a}$, $c > 0$ and $0 < a < 1/d$, assume for the moment that

$$\lim_{n\to\infty} \tilde{\Lambda}_{n,x}(u, w) = \tilde{\Psi}_{a,x}(u, v), \tag{27}$$

where $\tilde{\Psi}_{a,x}$ is defined in (11). The conclusion of Lemma 2 follows from Proposition 4 and again the Gärtner-Ellis Theorem.

Let us now prove (27). Set

$$M_i = [\langle u, Y_i\rangle + v]\, K\left(\frac{x - X_i}{h_i}\right),$$

then, for $(u, v) \in \mathbb{R}^q \times \mathbb{R}$,

$$\begin{aligned} \tilde{\Lambda}_{n,x}(u, v) &= \frac{1}{nh_n^d} \log \mathbb{E}\left[\exp\left(\sum_{i=1}^n M_i \frac{h_n^d}{h_i^d}\right)\right] \\ &= \frac{1}{nh_n^d} \sum_{i=1}^n \log \mathbb{E}\left[\exp\left(M_i \frac{h_n^d}{h_i^d}\right)\right]. \end{aligned}$$

By Taylor expansion, there exists $b_{i,n}$ between 1 and $\mathbb{E}\left[\exp\left(M_i \frac{h_n^d}{h_i^d}\right)\right]$ such that

$$\log \mathbb{E}\left[\exp\left(M_i \frac{h_n^d}{h_i^d}\right)\right] = \mathbb{E}\left[\exp\left(M_i \frac{h_n^d}{h_i^d}\right) - 1\right] - \frac{1}{2b_{i,n}^2}\left(\mathbb{E}\left[\exp\left(M_i \frac{h_n^d}{h_i^d}\right) - 1\right]\right)^2.$$

Noting that $h_n = cn^{-a}$ with $c > 0$ and $a \in]0, 1/d[$, $\tilde{\Lambda}_{n,x}$ can be rewritten as

$$\begin{aligned} \tilde{\Lambda}_{n,x}(u, v) &= \frac{1}{h_n^d} \sum_{i=1}^n \mathbb{E}\left[\exp\left(M_i \frac{h_n^d}{h_i^d}\right) - 1\right] - \frac{1}{2nh_n^d} \sum_{i=1}^n \frac{1}{b_{i,n}^2}\left(\mathbb{E}\left[\exp\left(M_i \frac{h_n^d}{h_i^d}\right) - 1\right]\right)^2 \\ &= \frac{1}{n} \sum_{i=1}^n \left(\frac{i}{n}\right)^{-ad} \int_{\mathbb{R}^d \times \mathbb{R}^q} \left[e^{\left(\frac{i}{n}\right)^{ad}(\langle u,y\rangle + v)K(z)} - 1\right] f(x, y)dzdy - R_{n,x}^{(1)}(u, v) + R_{n,x}^{(2)}(u, v), \end{aligned}$$

with

$$\begin{aligned} \tilde{R}_{n,x}^{(1)}(u, v) &= \frac{1}{2nh_n^d} \sum_{i=1}^n \frac{1}{b_{i,n}^2}\left(\mathbb{E}\left[\exp\left(M_i \frac{h_n^d}{h_i^d}\right) - 1\right]\right)^2 \\ \tilde{R}_{n,x}^{(2)}(u, v) &= \frac{1}{nh_n^d} \sum_{i=1}^n h_i^d \int_{\mathbb{R}^d \times \mathbb{R}^q} \left[e^{\frac{h_n^d}{h_i^d}(\langle u,y\rangle + v)K(z)} - 1\right] [f(x - h_i z, y) - f(x, y)]\, dzdy. \end{aligned}$$

Since $b_{i,n}$ is between 1 and $\mathbb{E}\left[\exp\left(M_i \frac{h_n^d}{h_i^d}\right)\right]$, we have

$$\frac{1}{b_{i,n}} \leq \max\left\{1, \frac{1}{\mathbb{E}\left(e^{M_i \frac{h_n^d}{h_i^d}}\right)}\right\}.$$

By Jensen's inequality, we obtain

$$\frac{1}{\mathbb{E}\left[e^{M_i\frac{h_n^d}{h_i^d}}\right]} \leq \frac{1}{e^{\mathbb{E}\left[M_i\frac{h_n^d}{h_i^d}\right]}}.$$

Observe that

$$\begin{aligned}
&\left|\mathbb{E}\left(M_i\frac{h_n^d}{h_i^d}\right)\right| \\
&\leq \frac{h_n^d}{h_i^d}\int_{\mathbb{R}^d}|\langle u,y\rangle+v|\left|K\left(\frac{x-z}{h_i}\right)\right|f(z,y)dzdy \\
&\leq h_n^d\left(\|u\|\int_{\mathbb{R}^d\times\mathbb{R}^q}|K(z)|\left[\int_{\mathbb{R}^q}\|y\|f(x-h_iz,y)dy\right]dz+|v|\int_{\mathbb{R}^d}|K(z)|\left[\int_{\mathbb{R}^q}f(x-h_iz,y)dy\right]dz\right),
\end{aligned}$$

which goes to 0 in view of (7) and since $\lim_{n\to\infty}h_n=0$. We deduce that the sequence $\left(\mathbb{E}\left[\exp\left(M_i\frac{h_n^d}{h_i^d}\right)\right]\right)$ is bounded, so that there exists $c>0$ such that

$$\frac{1}{b_{i,n}^2}\leq c,$$

and thus

$$\tilde{R}_{n,x}^{(1)}(u,v) \leq \frac{c}{2nh_n^d}\sum_{i=1}^n\left(\mathbb{E}\left[e^{M_i\frac{h_n^d}{h_i^d}}-1\right]\right)^2.$$

Now, in view of (7), and since K is bounded integrable, we have

$$\begin{aligned}
\mathbb{E}\left|e^{M_i\frac{h_n^d}{h_i^d}}-1\right| &\leq \mathbb{E}\left[\left|M_i\frac{h_n^d}{h_i^d}\right|e^{\left|M_i\frac{h_n^d}{h_i^d}\right|}\right] \\
&\leq \frac{h_n^d}{h_i^d}\int_{\mathbb{R}^d}|\langle u,y\rangle+v|\left|K\left(\frac{x-s}{h_i}\right)\right|e^{|\langle u,y\rangle+v|\left|K\left(\frac{x-s}{h_i}\right)\right|}f(s,y)dsdy \\
&\leq h_n^d\int_{\mathbb{R}^d\times\mathbb{R}^q}|\langle u,y\rangle+v|\,|K(z)|\,e^{|\langle u,y\rangle+v||K(z)|}f(x-h_iz,y)dzdy \\
&\leq Bh_n^d,
\end{aligned}$$

where B is a constant. Thus

$$\left|\tilde{R}_{n,x}^{(1)}(u,v)\right| \leq \frac{cB^2}{2}h_n^d,$$

and

$$\lim_{n\to\infty}\left|\tilde{R}_{n,x}^{(1)}(u,v)\right|=0.$$

Let us now consider $R_{n,x}^{(2)}$. Set $A>0$ and $\epsilon>0$; we then have

$$\begin{aligned}
&\tilde{R}_{n,x}^{(2)}(u,v) \\
&= \frac{1}{nh_n^d}\sum_{i=1}^n h_i^d\int_{\{\|z\|\leq A\}\times\mathbb{R}^q}\left[e^{\frac{h_n^d}{h_i^d}(\langle u,y\rangle+v)K(z)}-1\right][f(x-h_iz,y)-f(x,y)]\,dzdy \\
&\quad+\frac{1}{nh_n^d}\sum_{i=1}^n h_i^d\int_{\{\|z\|>A\}\times\mathbb{R}^q}\left[e^{\frac{h_n^d}{h_i^d}(\langle u,y\rangle+v)K(z)}-1\right][f(x-h_iz,y)-f(x,y)]\,dzdy \\
&= I+II.
\end{aligned}$$

Since $|e^t - 1| \leq |t|e^{|t|}$, it follows that

$$|II| \leq \frac{1}{n}\sum_{i=1}^{n}\int_{\{\|z\|>A\}\times\mathbb{R}^q} |\langle u,y\rangle + v|\,|K(z)|\,e^{|\langle u,y\rangle+v||K(z)|}\,|f(x-h_iz,y)-f(x,y)|\,dzdy.$$

Using the same argument as in (22), it holds that

$$|II| \leq \frac{\epsilon}{2}.$$

Now, for I, we write

$$\begin{aligned} I &= \frac{1}{nh_n^d}\sum_{i=1}^{n} h_i^d \int_{\{\|z\|\leq A\}\times\mathbb{R}^q} e^{\frac{h_n^d}{h_i^d}(\langle u,y\rangle+v)K(z)}\,[f(x-h_iz,y)-f(x,y)]\,dzdy \\ &\quad -\frac{1}{nh_n^d}\sum_{i=1}^{n} h_i^d \int_{\{\|z\|\leq A\}\times\mathbb{R}^q} [f(x-h_iz,y)-f(x,y)]\,dzdy. \end{aligned}$$

On the one hand, Assumption (A2) with $u=0$ ensures that

$$\lim_{i\to\infty}\int_{\{\|z\|\leq A\}\times\mathbb{R}^q} [f(x-h_iz,y)-f(x,y)]\,dzdy = 0.$$

Moreover, since $ad<1$, (15) ensures that

$$\lim_{n\to\infty}\frac{1}{nh_n^d}\sum_{i=1}^{n} h_i^d \int_{\{\|z\|\leq A\}\times\mathbb{R}^q} [f(x-h_iz,y)-f(x,y)]\,dzdy = 0,$$

so that for n large enough,

$$\left|\frac{1}{nh_n^d}\sum_{i=1}^{n} h_i^d \int_{\{\|z\|\leq A\}\times\mathbb{R}^q} [f(x-h_iz,y)-f(x,y)]\,dzdy\right| \leq \frac{\epsilon}{4}. \tag{28}$$

On the other hand, since for $i\leq n$, $0\leq \frac{h_n^d}{h_i^d}\leq 1$, by Assumption (A'1), there exists $n_0\in\mathbb{N}$ such that for any $i>n_0$,

$$\left|\int_{\mathbb{R}^q} e^{\frac{h_n^d}{h_i^d}(\langle u,y\rangle+v)K(z)}\,[f(x-h_iz,y)-f(x,y)]\,dy\right| \leq \frac{\epsilon}{8(2-ad)\int_{\|z\|\leq A}dz} \quad \forall n>i.$$

Noting that by (7), for any $\alpha\in[0,1]$,

$$\sup_t\left|\int_{\mathbb{R}^q} e^{\alpha\langle u,y\rangle} f(t,y)dy\right| \leq \sup_t\int_{\mathbb{R}^q} e^{\|u\|\|y\|} f(t,y)dy < \infty.$$

Since $ad<1$, by (15), we get for n sufficiently large

$$\left|\frac{1}{nh_n^d}\sum_{i=n_0+1}^{n} h_i^d \int_{\{\|z\|\leq A\}\times\mathbb{R}^q} e^{\frac{h_n^d}{h_i^d}(\langle u,y\rangle+v)K(z)}\,[f(x-h_iz,y)-f(x,y)]\,dydz\right| \leq \frac{\epsilon}{8}.$$

Now, for n large enough, in view of (7),

$$\begin{aligned} &\left|\frac{1}{nh_n^d}\sum_{i=1}^{n_0} h_i^d \int_{\{\|z\|\leq A\}\times\mathbb{R}^q} e^{\frac{h_n^d}{h_i^d}(\langle u,y\rangle+v)K(z)}\,[f(x-h_iz,y)-f(x,y)]\,dydz\right| \\ &\leq \frac{1}{nh_n^d}\sum_{i=1}^{n_0} h_i^d \int_{\{\|z\|\leq A\}\times\mathbb{R}^q} e^{|\langle u,y\rangle+v||K(z)|}\,|f(x-h_iz,y)-f(x,y)|\,dydz \\ &\leq \frac{\epsilon}{8}. \end{aligned}$$

It follows that for n large enough,

$$\left|\frac{1}{nh_n^d}\sum_{i=1}^{n} h_i^d \int_{\{\|z\|\leq A\}\times\mathbb{R}^q} e^{\frac{h_n^d}{h_i^d}(\langle u,y\rangle+v)K(z)}\left[f(x-h_iz,y)-f(x,y)\right]dydz\right| \leq \frac{\epsilon}{4}. \tag{29}$$

The combination of (28) and (29) ensures that $|I|\leq\frac{\epsilon}{2}$, which ensures that

$$\lim_{n\to\infty}\left|\tilde{R}_{n,x}^{(2)}(u)\right| = 0.$$

Hence, (27) follows from analysis considerations. ■

3.2 Proof of Theorems 1 and 3

Let us consider the following functions defined as:

$$\begin{aligned} H_1 : \mathbb{R}^q\times\mathbb{R}^* &\to \mathbb{R}^q \\ (\alpha,\beta) &\mapsto \frac{\alpha}{\beta}, \end{aligned}$$

and

$$\begin{aligned} H_2 : \mathbb{R}^q\times\mathbb{R} &\to \mathbb{R}^q \\ (\alpha,\beta) &\mapsto \begin{cases} \frac{\alpha}{\beta} & \text{if } \beta\neq 0 \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

3.2.1 Proof of Theorem 1

i) Let U be an open subset of $\mathbb{R}^q$, we have

$$\frac{1}{nh_n^d}\log\mathbb{P}\left[r_n(x)\in U\right] = \frac{1}{nh_n^d}\log\mathbb{P}\left[(m_n(x),g_n(x))\in H_2^{-1}(U)\right]. \tag{30}$$

Observe that $H_1^{-1}(U)\subset H_2^{-1}(U)$ and $H_1^{-1}(U)$ is an open subset on $\mathbb{R}^q\times\mathbb{R}^*$ which is open, it follows that $H_1^{-1}(U)$ is an open subset on $\mathbb{R}^q\times\mathbb{R}$. We deduce from (30) that

$$\frac{1}{nh_n^d}\log\mathbb{P}\left[r_n(x)\in U\right] \geq \frac{1}{nh_n^d}\log\mathbb{P}\left[(m_n(x),g_n(x))\in H_1^{-1}(U)\right].$$

The application of Lemma 1 ensures that

$$\begin{aligned} \liminf_{n\to\infty}\frac{1}{nh_n^d}\log\mathbb{P}\left[r_n(x)\in U\right] &\geq \liminf_{n\to\infty}\frac{1}{nh_n^d}\log\mathbb{P}\left[(m_n(x),g_n(x))\in H_1^{-1}(U)\right] \\ &\geq -\inf_{(t_1,t_2)\in H_1^{-1}(U)} I_x(t_1,t_2) = -\inf_{s\in U} J^*(s), \end{aligned}$$

and the first part of Theorem 1 is proved.

ii) Let V be a closed subset of $\mathbb{R}^q$, we have

$$\begin{aligned} \frac{1}{nh_n^d}\log\mathbb{P}\left[r_n(x)\in V\right] &= \frac{1}{nh_n^d}\log\mathbb{P}\left[(m_n(x),g_n(x))\in H_2^{-1}(V)\right] \\ &\leq \frac{1}{nh_n^d}\log\mathbb{P}\left[(m_n(x),g_n(x))\in \overline{H_2^{-1}(V)}\right]. \end{aligned}$$

Now, observe that $\overline{H_2^{-1}(V)} = H_1^{-1}(V) \cup A$ where $A \subset \mathbb{R}^q \times \{0\}$ and $\left(\vec{0}, 0\right) \in A$ (since for any $s \in \mathbb{R}^q$, $(\vec{0}, 0) \in \overline{H_2^{-1}(s)}$). The application of Lemma 1 again ensures that

$$\begin{aligned}
\limsup_{n\to\infty} \frac{1}{nh_n^d} \log \mathbb{P}\left[r_n(x) \in V\right] &\leq -\inf_{(s,t)\in H_1^{-1}(V)\cup A} I_x(s,t) \\
&\leq -\inf_{(s,t)\in H_1^{-1}(V)\cup\{(\vec{0},0)\}} I_x(s,t) \\
&\leq -\inf_{s\in V,\ t\in\mathbb{R}} I_x(st,t) \\
&\leq -\inf_{s\in V,\ t\in\mathbb{R}} \hat{I}_x(s,t) \\
&\leq -\inf_{s\in V} J(s),
\end{aligned}$$

where the second inequality comes from Condition (C); this concludes the proof of Theorem 1. ■

3.2.2 Proof of Theorem 3

Applying Lemma 2, Theorem 3 is proved by following the same approach as for the proof of Theorem 1 with replacing m_n, g_n, J^* and J by $\tilde{m}_n$, $\tilde{g}_n$, $\tilde{J}_a^*$ and $\tilde{J}_a$ respectively. ■

3.3 Proof of Theorem 2

Set

$$B_n(x) = \frac{1}{g(x)}\left(m_n(x) - m(x)\right) - \frac{r(x)}{g(x)}\left(g_n(x) - g(x)\right).$$

Let us at first state the two following lemmas.

Lemma 4 *Under the assumptions of Theorem 2, the sequence* $\left(v_n\left(B_n(x) - \mathbb{E}\left(B_n(x)\right)\right)\right)$ *satisfies a LDP with speed* $\left(\frac{nh_n^d}{v_n^2}\right)$ *and good rate function* G_x.

Lemma 5 *Under the assumptions of Theorem 2,*

$$\lim_{n\to\infty} v_n \mathbb{E}(B_n(x)) = 0. \tag{31}$$

We first show that how Theorem 2 can be deduced from the application of Lemmas 4 and 5, and then prove Lemmas 4 and 5 successively.

3.3.1 Proof of Theorem 2

Lemmas 4 and 5 imply that the sequence $(v_n B_n(x))$ satisfies a LDP with speed $\left(\dfrac{nh_n^d}{v_n^2}\right)$ and good rate function G_x. To prove Theorem 2, we show that $(v_n(r_n - r))$ and $(v_n B_n)$ are exponentially contiguous.

Let us first note that, for x such that $g_n(x) \neq 0$, we have:

$$\begin{aligned}
r_n(x) - r(x) &= \frac{m_n(x)}{g_n(x)} - \frac{m(x)}{g(x)} \\
&= \frac{\left(m_n(x) - m(x)\right) g(x) + \left(g(x) - g_n(x)\right) m(x)}{g_n(x) g(x)} \\
&= B_n(x) \frac{g(x)}{g_n(x)}.
\end{aligned}$$

It follows that, for any $\delta > 0$, we have

$$\begin{aligned}
&\mathbb{P}\Big[v_n\|\left(r_n(x)-r(x)\right)-B_n(x)\| > \delta\Big] \\
&\quad\leq\ \mathbb{P}\Big[v_n\|B_n(x)\left(\frac{g(x)}{g_n(x)}-1\right)\| > \delta \ \text{ and } \ g_n(x)\neq 0\Big] + \mathbb{P}\left[g_n(x)=0\right] \\
&\quad\leq\ \mathbb{P}\left[\sqrt{v_n}\|B_n(x)\| > \delta\right] + \mathbb{P}\left[\sqrt{v_n}\,|g(x)-g_n(x)| > \delta\,|g_n(x)|\right] + \mathbb{P}\left[|g(x)-g_n(x)| > \frac{g(x)}{2}\right] \\
&\quad\leq\ \mathbb{P}\left[\sqrt{v_n}\|B_n(x)\| > \delta\right] + \mathbb{P}\left[\sqrt{v_n}\,|g(x)-g_n(x)| > \delta\,|g_n(x)| \ \text{ and } \ \frac{g_n(x)}{g(x)} > \frac{1}{2}\right] \\
&\qquad +\mathbb{P}\left[\frac{g_n(x)}{g(x)} \leq \frac{1}{2}\right] + \mathbb{P}\left[|g_n(x)-g(x)| > \frac{g(x)}{2}\right] \\
&\quad\leq\ \mathbb{P}\left[\sqrt{v_n}\|B_n(x)\| > \delta\right] + \mathbb{P}\left[\sqrt{v_n}\,|g(x)-g_n(x)| > \delta\frac{g(x)}{2}\right] + \mathbb{P}\left[g(x)-g_n(x) \geq \frac{g(x)}{2}\right] \\
&\qquad +\mathbb{P}\left[|g_n(x)-g(x)| > \frac{g(x)}{2}\right].
\end{aligned}$$

Since $\lim_{n\to\infty} v_n = \infty$, it follows that, for n large enough,

$$\begin{aligned}
&\mathbb{P}\Big[v_n\|\left(r_n(x)-r(x)\right)-B_n(x)\| > \delta\Big] \\
&\quad\leq\ 4\max\left\{\mathbb{P}\left[\sqrt{v_n}\|B_n(x)\| > \delta\right]\ ;\ \mathbb{P}\left[\sqrt{v_n}\,|g(x)-g_n(x)| > \delta\frac{g(x)}{2}\right]\right\},
\end{aligned}$$

and thus

$$\begin{aligned}
&\frac{v_n^2}{nh_n^d}\log\mathbb{P}\Big[v_n\|\left(r_n(x)-r(x)\right)-B_n(x)\| > \delta\Big] \\
&\quad\leq\ \frac{v_n^2}{nh_n^d}\log 4 + \max\left\{\frac{v_n^2}{nh_n^d}\log\mathbb{P}\left[\sqrt{v_n}\|B_n(x)\| > \delta\right]\ ;\ \frac{v_n^2}{nh_n^d}\log\mathbb{P}\left[\sqrt{v_n}\,|g(x)-g_n(x)| > \delta\frac{g(x)}{2}\right]\right\}.
\end{aligned}$$

Now, since the sequence $(v_nB_n(x))$ satisfies a LDP with speed $\left(\frac{nh_n^d}{v_n^2}\right)$ and good rate function G_x, there exists $c_1 > 0$ such that

$$\limsup_{n\to\infty}\frac{v_n}{nh_n^d}\log\mathbb{P}\left[\sqrt{v_n}\|B_n(x)\| > \delta\right] \ \leq\ -c_1.$$

Moreover, the application of Theorem 1 in Mokkadem et al. [14] guarantees the existence of $c_2 > 0$ such that

$$\limsup_{n\to\infty}\frac{v_n}{nh_n^d}\log\mathbb{P}\left[\sqrt{v_n}\,|g(x)-g_n(x)| > \delta\frac{g(x)}{2}\right] < -c_2.$$

We thus deduce that

$$\lim_{n\to\infty}\frac{v_n^2}{nh_n^d}\log\mathbb{P}\Big[v_n\|\left(r_n(x)-r(x)\right)-B_n(x)\| > \delta\Big] = -\infty,$$

which means that the sequences $(v_n(r_n(x)-r(x)))$ and $(v_nB_n(x))$ are exponentially contiguous. Theorem 2 thus follows. ■

3.3.2 Proof of Lemma 4

For any $u \in \mathbb{R}^q$, set

$$\begin{aligned}
\Gamma_{n,x}(u) &= \frac{v_n^2}{nh_n^d} \log \mathbb{E}\left[\exp\left(\frac{nh_n^d}{v_n}\langle u, B_n(x) - \mathbb{E}\left(B_n(x)\right)\rangle\right)\right], \\
\Phi_x(u) &= \frac{1}{2g^2(x)} \int_{\mathbb{R}^d \times \mathbb{R}^q} \langle u, y - r(x)\rangle^2 K^2(z) f(x,y) dz dy \\
&= \frac{u^T \Sigma_x u}{2g(x)} \int_{\mathbb{R}^d} K^2(z) dz.
\end{aligned}$$

To prove Lemma 4, it suffices to show that, for all $u \in \mathbb{R}^q$,

$$\lim_{n\to\infty} \Gamma_{n,x}(u) = \Phi_x(u).$$

As a matter of fact, since Φ_x is a quadratic function, Lemma 4 then follows from the application of the Gärtner-Ellis Theorem. For $u \in \mathbb{R}^q$, set

$$\hat{Z}_i = \langle u, Y_i - r(x)\rangle K\left(\frac{x - X_i}{h_n}\right),$$

and note that

$$\Gamma_{n,x}(u) = \frac{v_n^2}{nh_n^d} \log \mathbb{E}\left[\exp\left(\frac{1}{v_n g(x)} \sum_{i=1}^n \left[\hat{Z}_i - \mathbb{E}(\hat{Z}_i)\right]\right)\right].$$

Since (X_i, Y_i), $i = 1, \ldots, n$ are independent and identically distributed, it holds that

$$\Gamma_{n,x}(u) = \frac{v_n^2}{h_n^d} \log \mathbb{E}\left[e^{\frac{\hat{Z}_n}{v_n g(x)}}\right] - \frac{v_n}{h_n^d g(x)} \mathbb{E}(\hat{Z}_n).$$

Now, we follow the same lines as in the proof of Lemma 3. A Taylor's expansion ensures that there exists $\hat{c}_n$ between 1 and $\mathbb{E}\left[e^{\frac{\hat{Z}_n}{v_n g(x)}}\right]$ such that

$$\begin{aligned}
\Gamma_{n,x}(u) &= \frac{v_n^2}{h_n^d} \mathbb{E}\left[e^{\frac{\hat{Z}_n}{v_n g(x)}} - 1 - \frac{\hat{Z}_n}{v_n g(x)}\right] - \hat{R}_{n,x}^{(1)}(u) \\
&= v_n^2 \int_{\mathbb{R}^d \times \mathbb{R}^q} \left[e^{\frac{1}{v_n g(x)}\langle u, y - r(x)\rangle K(z)} - 1 - \frac{1}{v_n g(x)}\langle u, y - r(x)\rangle K(z)\right] f(x,y) dz dy \\
&\quad - \hat{R}_{n,x}^{(1)}(u) + \hat{R}_{n,x}^{(2)}(u),
\end{aligned}$$

with

$$\begin{aligned}
&\hat{R}_{n,x}^{(1)}(u) \\
&= \frac{v_n^2}{2\hat{c}_n^2 h_n^d} \left(\mathbb{E}\left[e^{\frac{\hat{Z}_n}{v_n g(x)}} - 1\right]\right)^2 \\
\hat{R}_{n,x}^{(2)}(u) &= v_n^2 \int_{\mathbb{R}^d \times \mathbb{R}^q} \left[e^{\frac{1}{v_n g(x)}\langle u, y - r(x)\rangle K(z)} - 1 - \frac{1}{v_n g(x)}\langle u, y - r(x)\rangle K(z)\right] \left[f(x - h_n z, y) - f(x,y)\right] dz dy,
\end{aligned}$$

and

$$\frac{1}{\hat{c}_n} \leq \max\left\{1, \frac{1}{\mathbb{E}\left(e^{\frac{\hat{Z}_n}{v_n g(x)}}\right)}\right\}.$$

Noting that

$$\left|\mathbb{E}\left(\frac{\hat{Z}_n}{v_n g(x)}\right)\right| \leq \frac{h_n^d\|u\|}{v_n g(x)}\int_{\mathbb{R}^d}|K(z)|\left[\int_{\mathbb{R}^q}\|y\|f(x-h_n z,y)dy\right]dz + \frac{h_n^d\|u\|\|r(x)\|}{v_n g(x)}\int_{\mathbb{R}^d}|K(z)|\left[\int_{\mathbb{R}^q}f(x-h_n z,y)dy\right]dz. \tag{32}$$

It follows from (7) that,

$$\mathbb{E}\left(\frac{\hat{Z}_n}{v_n g(x)}\right) \to 0.$$

We deduce that there exists $c' \in \mathbb{R}_+^*$ such that

$$\frac{1}{\hat{c}_n^2} \leq c',$$

and thus, in view of (7),

$$\begin{aligned}
\hat{R}_{n,x}^{(1)}(u) &\leq \frac{c'}{2}\frac{v_n^2}{h_n^d}\left(\int_{\mathbb{R}^d\times\mathbb{R}^q}\left[e^{\frac{1}{v_n g(x)}\langle u,y-r(x)\rangle K\left(\frac{x-s}{h_n}\right)}-1\right]f(s,y)dsdy\right)^2 \\
&\leq \frac{c'}{2g^2(x)}h_n^d\left(\int_{\mathbb{R}^d\times\mathbb{R}^q}|\langle u,y-r(x)\rangle K(z)|\,e^{\left|\frac{1}{g(x)}\langle u,y-r(x)\rangle K(z)\right|}f(x-h_n z,y)dzdy\right)^2 \\
&\leq \frac{c'e^{\frac{2\|K\|_\infty}{g(x)}\|u\|\|r(x)\|}}{2g^2(x)}h_n^d\left(\int_{\mathbb{R}^d\times\mathbb{R}^q}|\langle u,y-r(x)\rangle K(z)|\,e^{\frac{\|K\|_\infty}{g(x)}\|u\|\|y\|}f(x-h_n z,y)dzdy\right)^2 \\
&\leq Bh_n^d,
\end{aligned}$$

where B is a constant, so that

$$\lim_{n\to\infty}\left|\hat{R}_{n,x}^{(1)}(u)\right| = 0.$$

On the other hand, since $\forall x\in\mathbb{R}$, $e^x-1-x = \frac{x^2}{2}+\frac{x^3}{6}d(x)$, with $d(x)\leq e^{|x|}$, we get

$$\begin{aligned}
&\hat{R}_{n,x}^{(2)}(u) \\
&= \frac{1}{2g^2(x)}\int_{\mathbb{R}^d\times\mathbb{R}^q}\langle u,y-r(x)\rangle^2K^2(z)\left[f(x-h_n z,y)-f(x,y)\right]dzdy + \mathcal{R}_{n,x}(u),
\end{aligned} \tag{33}$$

with

$$|\mathcal{R}_{n,x}(u)| \leq \frac{1}{6v_n g^3(x)}\int_{\mathbb{R}^d\times\mathbb{R}^q}\left|\langle u,y-r(x)\rangle^3K^3(z)\right|e^{\frac{1}{g(x)}|\langle u,y-r(x)\rangle K(z)|}\left|f(x-h_n z,y)-f(x,y)\right|dzdy.$$

It follows from (7) that $\mathcal{R}_{n,x}$ converges to 0. Applying then (A2) and (A3), we find

$$\lim_{n\to\infty}\left|\hat{R}_{n,x}^{(2)}(u)\right| = 0.$$

Finally, we have

$$\begin{aligned}
&\Gamma_{n,x}(u) \\
&= v_n^2\int_{\mathbb{R}^d\times\mathbb{R}^q}\left[e^{\frac{1}{v_n g(x)}\langle u,y-r(x)\rangle K(z)}-1-\frac{1}{v_n g(x)}\langle u,y-r(x)\rangle K(z)\right]f(x,y)dzdy - \hat{R}_{n,x}^{(1)}(u) \\
&\quad +\hat{R}_{n,x}^{(2)}(u) \\
&= \Phi_x(u) - \hat{R}_{n,x}^{(1)}(u) + \hat{R}_{n,x}^{(2)}(u) + \hat{R}_{n,x}^{(3)}(u),
\end{aligned}$$

with

$$\hat{R}_{n,x}^{(3)}(u)$$
$$= v_n^2 \int_{\mathbb{R}^d\times\mathbb{R}^q} \left[e^{\frac{1}{v_n g(x)}\langle u, y-r(x)\rangle K(z)} - 1 - \frac{\langle u, y-r(x)\rangle K(z)}{v_n g(x)} - \frac{\langle u, y-r(x)\rangle^2 K^2(z)}{2v_n^2 g^2(x)} \right] f(x,y)dzdy.$$

By the majoration $\left| e^x - 1 - x - \frac{x^2}{2} \right| \leq \left| \frac{x^3}{6} d(x) \right|$, we get

$$\left| \hat{R}_{n,x}^{(3)}(u) \right|$$
$$\leq \frac{1}{6v_n g^3(x)} \int_{\mathbb{R}^d\times\mathbb{R}^q} \left| \langle u, y-r(x)\rangle^3 K^3(z) \right| e^{\frac{1}{v_n g(x)}|\langle u, y-r(x)\rangle K(z)|} f(x,y)dzdy,$$

and (7) ensures that

$$\lim_{n\to\infty} \left| \hat{R}_{n,x}^{(3)}(u) \right| = 0,$$

which concludes the proof of Lemma 4. ■

3.3.3 Proof of Lemma 5

Observe that

$$\mathbb{E}(B_n(x)) = \frac{1}{g(x)} \left[\mathbb{E}(m_n(x)) - m(x) \right] - \frac{r(x)}{g(x)} \left[\mathbb{E}(g_n(x)) - g(x) \right]. \tag{34}$$

Since

$$\begin{aligned}
\mathbb{E}(m_n(x)) - m(x) &= \frac{1}{h_n^d} \mathbb{E}\left(Y_1 K\left(\frac{x - X_1}{h_n} \right) \right) - m(x) \\
&= \frac{1}{h_n^d} \int_{\mathbb{R}^d\times\mathbb{R}^q} yK\left(\frac{x-z}{h_n} \right) f(z,y)dzdy - m(x) \\
&= \frac{1}{h_n^d} \int_{\mathbb{R}^d} m(z)K\left(\frac{x-z}{h_n} \right) dz - m(x) \\
&= \int_{\mathbb{R}^d} K(y) \left[m(x - h_n y) - m(x) \right] dy,
\end{aligned}$$

Assumptions (A5)i), (A5)iii) and a Taylor's expansion of m of order p ensure that

$$\mathbb{E}(m_n(x)) - m(x) = O(h_n^p). \tag{35}$$

Similarly, we have

$$\mathbb{E}(g_n(x)) - g(x) = O(h_n^p). \tag{36}$$

We deduce from (34), (35), and (36) that

$$\mathbb{E}(B_n(x)) = O(h_n^p),$$

and thus Lemma 5 follows from Assumption (A5)ii). ■

3.4 Proof of Theorem 4

Set

$$\tilde{B}_n(x) = \frac{1}{g(x)}\left(\tilde{m}_n(x) - m(x)\right) - \frac{r(x)}{g(x)}\left(\tilde{g}_n(x) - g(x)\right),$$

and, for any $u \in \mathbb{R}^q$,

$$\begin{aligned}
\tilde{\Gamma}_{n,x}(u) &= \frac{v_n^2}{nh_n^d}\log \mathbb{E}\left[\exp\left(\frac{nh_n^d}{v_n}\langle u, \tilde{B}_n(x) - \mathbb{E}\left(\tilde{B}_n(x)\right)\rangle\right)\right], \\
\tilde{\Phi}_{a,x}(u) &= \frac{1}{2(1+ad)g^2(x)}\int_{\mathbb{R}^d\times\mathbb{R}^q}\langle u, y - r(x)\rangle^2 K^2(z) f(x,y)dzdy \\
&= \frac{1}{1+ad}\frac{u^T\Sigma_x u}{2g(x)}\int_{\mathbb{R}^d} K^2(z)dz.
\end{aligned}$$

By following the steps of the proof of Lemma 4 and by using the property (15), we prove that

$$\lim_{n\to\infty}\tilde{\Gamma}_{n,x}(u) = \tilde{\Phi}_{a,x}(u). \tag{37}$$

We first show how (37) implies Theorem 4. The function $\tilde{\Phi}_{a,x}$ being quadratic, the application of the Gärtner-Ellis Theorem then ensures that

$$\begin{aligned}
&\text{the sequence } \left(v_n\left(\tilde{B}_n(x) - \mathbb{E}\left(\tilde{B}_n(x)\right)\right)\right) \text{ satisfies a LDP} \\
&\text{with speed } \left(\frac{nh_n^d}{v_n^2}\right) \text{ and good rate function } \tilde{G}_{a,x}.
\end{aligned} \tag{38}$$

Now, following the proof of Lemma 5, we have

$$\begin{aligned}
\mathbb{E}\left(\tilde{m}_n(x)\right) - m(x) &= \frac{1}{n}\sum_{i=1}^n \frac{1}{h_i^d}\int_{\mathbb{R}^d\times\mathbb{R}^q} yK\left(\frac{x-z}{h_i}\right) f(z,y)dzdy - m(x) \\
&= \frac{1}{n}\sum_{i=1}^n \int_{\mathbb{R}^d} K(y)\left[m(x - h_i y) - m(x)\right]dy.
\end{aligned}$$

Here again, Assumptions (A5)i), (A5)iii) and a Taylor's expansion of m of order p ensure that

$$\mathbb{E}\left(\tilde{m}_n(x)\right) - m(x) = O\left(\frac{1}{n}\sum_{i=1}^n h_i^p\right).$$

and similarly,

$$\mathbb{E}\left(\tilde{g}_n(x)\right) - g(x) = O\left(\frac{1}{n}\sum_{i=1}^n h_i^p\right),$$

thus

$$v_n\mathbb{E}\left(\tilde{B}_n(x)\right) = O\left(\frac{v_n}{n}\sum_{i=1}^n h_i^p\right).$$

- If $ap < 1$, since (h_n) varies regularly with exponent $(-a)$, we have, in view of (15) and Assumption (A5)ii),

$$\frac{v_n}{n}\sum_{i=1}^{n} h_i^p = O\left(\frac{v_n}{n}\left[nh_n^p\right]\right) = o(1).$$

- If $ap > 1$, we have $\sum_i h_i^p < \infty$ and thus, since $v_n = o(nh_n^d)$, we get

$$\frac{v_n}{n}\sum_{i=1}^{n} h_i^p = O\left(\frac{v_n}{n}\right) = o(1).$$

- In the case $ap = 1$, let $\mathcal{L}$ be the slowly varying function such that $h_n = n^{-a}\mathcal{L}(n)$, and set $\varepsilon > 0$ small enough. Since $a(p-\varepsilon) < 1$, we have $h_n^p = o(h_n^{p-\varepsilon})$, and in view of (15) and (A4),

$$\begin{aligned}\frac{v_n}{n}\sum_{i=1}^{n} h_i^p &= o\left(v_n h_n^{p-\varepsilon}\right) = o\left(nh_n^{d+p-\varepsilon}\right)\\ &= o\left(n^{1-a(d+p-\varepsilon)}\left[\mathcal{L}(n)\right]^{d+p-\varepsilon}\right)\\ &= o\left(n^{-a(d-\varepsilon)}\left[\mathcal{L}(n)\right]^{d+p-\varepsilon}\right) = o(1).\end{aligned}$$

We thus deduce that

$$\lim_{n\to\infty} v_n \mathbb{E}(\tilde{B}_n(x)) = 0. \tag{39}$$

To conclude the proof of Theorem 4, we follow the same lines as for the proof of Theorem 2 (see Section 3.3), except that we apply (38) instead of Lemma 4, (39) instead of Lemma 5, and Theorem 1 in Mokkadem et al. [13] instead of Theorem 1 in Mokkadem et al. [14]. ■
Let us now prove (37). For $u \in \mathbb{R}^q$, set

$$T_i = \langle u, Y_i - r(x)\rangle K\left(\frac{x - X_i}{h_i}\right),$$

and note that

$$\tilde{\Gamma}_{n,x}(u) = \frac{v_n^2}{nh_n^d}\log\mathbb{E}\left[\exp\left(\frac{h_n^d}{v_n g(x)}\sum_{i=1}^{n}\frac{1}{h_i^d}\left[T_i - \mathbb{E}(T_i)\right]\right)\right].$$

Since (X_i, Y_i), $i = 1, \ldots, n$ are independent and identically distributed, it holds that

$$\tilde{\Gamma}_{n,x}(u) = \frac{v_n^2}{h_n^d}\sum_{i=1}^{n}\log\mathbb{E}\left[e^{\frac{h_n^d T_i}{v_n g(x) h_i^d}}\right] - \frac{v_n}{ng(x)}\sum_{i=1}^{n}\frac{1}{h_i^d}\mathbb{E}(T_i).$$

By Taylor expansion, there exists $c_{i,n}$ between 1 and $\mathbb{E}\left[e^{\frac{h_n^d T_i}{v_n g(x)}h_i^d}\right]$ such that

$$\log\mathbb{E}\left[e^{\frac{h_n^d T_i}{v_n g(x)}h_i^d}\right] = \mathbb{E}\left[e^{\frac{h_n^d T_i}{v_n g(x)}h_i^d} - 1\right] - \frac{1}{2c_{i,n}^2}\left(\mathbb{E}\left[e^{\frac{h_n^d T_i}{v_n g(x)}h_i^d} - 1\right]\right)^2,$$

and $\tilde{\Gamma}_{n,x}$ can be rewritten as

$$\tilde{\Gamma}_{n,x}(u) = \frac{v_n^2}{nh_n^d}\sum_{i=1}^n \mathbb{E}\left[e^{\frac{T_i h_n^d}{v_n g(x) h_i^d}} - 1\right] - \frac{v_n^2}{2nh_n^d}\sum_{i=1}^n \frac{1}{c_{i,n}^2}\left(\mathbb{E}\left[e^{\frac{T_i h_n^d}{v_n g(x) h_i^d}} - 1\right]\right)^2 - \frac{v_n}{ng(x)}\sum_{i=1}^n \frac{1}{h_i^d}\mathbb{E}(T_i).$$

A Taylor expansion implies again that there exists $c'_{i,n}$ between 0 and $\frac{T_i h_n^d}{v_n g(x) h_i^d}$ such that

$$\begin{aligned}
&\mathbb{E}\left[e^{\frac{T_i h_n^d}{v_n g(x) h_i^d}} - 1\right] \\
&= \frac{h_n^d}{v_n g(x) h_i^d}\mathbb{E}(T_i) + \frac{1}{2}\left(\frac{h_n^d}{v_n g(x) h_i^d}\right)^2 \mathbb{E}(T_i^2) + \frac{1}{6}\left(\frac{h_n^d}{v_n g(x) h_i^d}\right)^3 \mathbb{E}(e^{c'_{i,n}} T_i^3).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\tilde{\Gamma}_{n,x}(u) \\
&= \frac{1}{2g^2(x)}\frac{1}{nh_n^{-d}}\sum_{i=1}^n \frac{1}{h_i^d}\int_{\mathbb{R}^d\times\mathbb{R}^q} \langle u, y - r(x)\rangle^2 K^2(z) f(x,y)dzdy + \ddot{R}_{n,x}^{(1)}(u) + \ddot{R}_{n,x}^{(2)}(u), \quad (40)
\end{aligned}$$

with

$$\begin{aligned}
\ddot{R}_{n,x}^{(1)}(u) &= \frac{1}{6}\frac{h_n^{2d}}{v_n g^3(x)}\frac{1}{n}\sum_{i=1}^n \frac{1}{h_i^{2d}}\mathbb{E}(e^{c'_{i,n}} T_i^3) - \frac{v_n^2}{2nh_n^d}\sum_{i=1}^n \frac{1}{c_{i,n}^2}\left(\mathbb{E}\left[e^{\frac{T_i h_n^d}{v_n g(x) h_i^d}} - 1\right]\right)^2, \\
\ddot{R}_{n,x}^{(2)}(u) &= \frac{h_n^d}{2g^2(x)}\frac{1}{n}\sum_{i=1}^n \frac{1}{h_i^d}\int_{\mathbb{R}^d\times\mathbb{R}^q} \langle u, y - r(x)\rangle^2 K^2(z)\left[f(x - h_i z, y) - f(x,y)\right] dzdy.
\end{aligned}$$

In view of (15), the first term in the right-hand-side of (40) converges to $\tilde{\Phi}_{a,x}$.
It remains to prove that $\ddot{R}_{n,x}^{(1)}$ and $\ddot{R}_{n,x}^{(2)}$ converge to 0. We have

$$\begin{aligned}
\left|\mathbb{E}\left[\frac{h_n^d T_i}{v_n g(x) h_i^d}\right]\right| &\leq \frac{h_n^d}{v_n g(x) h_i^d}\int_{\mathbb{R}^d\times\mathbb{R}^q}\left|\langle u, y - r(x)\rangle K\left(\frac{x-s}{h_i}\right)\right| f(s,y)dsdy \\
&\leq \frac{h_n^d}{v_n g(x)}\int_{\mathbb{R}^d\times\mathbb{R}^q} |\langle u, y - r(x)\rangle K(z)| f(x - h_i z, y)dzdy.
\end{aligned}$$

In view of (7), the integral is bounded, thus

$$\lim_{n\to\infty}\sup_{i\leq n}\mathbb{E}\left[\frac{h_n^d T_i}{v_n g(x) h_i^d}\right] = 0,$$

so that, there exists $c > 0$ such that

$$\frac{1}{c_{i,n}^2} \leq c.$$

Now, on the one hand, since $|e^t - 1| \leq |t|e^{|t|}$, and in view of (7) and (16), we have

$$\begin{aligned}
\mathbb{E}\left|e^{\frac{h_n^d T_i}{v_n g(x) h_i^d}} - 1\right| &\leq \frac{h_n^d}{v_n g(x)}\int_{\mathbb{R}^d\times\mathbb{R}^q} |\langle u, y - r(x)\rangle K(z)| e^{c\left|\frac{1}{g(x)}\langle u, y - r(x)\rangle K(z)\right|} f(x - h_i z, y)dzdy \\
&\leq B_1\frac{h_n^d}{v_n g(x)},
\end{aligned}$$

where B_1 and c are constants. We deduce that

$$\lim_{n\to\infty} \frac{v_n^2}{2nh_n^d} \sum_{i=1}^n \frac{1}{c_{i,n}^2} \left(\mathbb{E}\left[e^{\frac{T_i h_n^d}{v_n g(x) h_i^d}} - 1 \right] \right)^2 = 0.$$

On the other hand,

$$\begin{aligned} \mathbb{E}\left[T_i^3 e^{c'_{i,n}}\right] &\leq \mathbb{E}\left[|T_i|^3 e^{|c'_{i,n}|}\right] \\ &\leq h_i^d \int_{\mathbb{R}^d\times\mathbb{R}^q} |\langle u, y - r(x)\rangle K(z)|^3 \, e^{\frac{c}{g(x)}|\langle u, y-r(x)\rangle K(z)|} f(x - h_i z, y) dz dy \\ &\leq B_2 h_i^d, \end{aligned}$$

where B_2 is a constant. Thus,

$$\left| \frac{h_n^{2d}}{6nv_n g^3(x)} \sum_{i=1}^n \frac{1}{h_i^{2d}} \mathbb{E}(e^{c'_{i,n}} T_i^3) \right| \leq \frac{h_n^d}{6v_n g^3(x)} \frac{B_2}{nh_n^{-d}} \sum_{i=1}^n h_i^{-d}.$$

Since $\lim_{n\to\infty} \dfrac{h_n^d}{v_n} = 0$, (15) ensures that

$$\lim_{n\to\infty} \frac{h_n^{2d}}{6v_n g^3(x)} \frac{1}{n} \sum_{i=1}^n \frac{1}{h_i^{2d}} \mathbb{E}(e^{c'_{i,n}} T_i^3) = 0,$$

which proves that

$$\lim_{n\to\infty} \left| \ddot{R}_{n,x}^{(1)}(u) \right| = 0.$$

Finally, using (15), (A2) and (A3), we have

$$\begin{aligned} &\lim_{n\to\infty} \ddot{R}_{n,x}^{(2)}(u) \\ &= \lim_{n\to\infty} \frac{1}{2g^2(x)} \frac{\sum_{i=1}^n h_i^{-d}}{nh_n^{-d}} \frac{1}{\sum_{i=1}^n h_i^{-d}} \sum_{i=1}^n h_i^{-d} \int_{\mathbb{R}^d\times\mathbb{R}^q} \langle u, y - r(x)\rangle^2 K^2(z) \left[f(x - h_i z, y) - f(x, y) \right] dz dy \\ &= 0, \end{aligned}$$

which proves (37). ■

3.5 Proof of Propositions 3 and 4

3.5.1 Proof of Proposition 3

- The strict convexity of Ψ_x follows from its definition, since for any $\gamma \in]0, 1[$, and $(u, v) \neq (u', v')$,

$$\begin{aligned} \Psi_x\left(\gamma(u, v) + (1-\gamma)(u', v')\right) &= \Psi_x\left((\gamma u + (1-\gamma)u', \gamma v + (1-\gamma)v')\right) \\ &= \int_{\mathbb{R}^d\times\mathbb{R}^q} \left(e^{[\langle \gamma u + (1-\gamma)u', y\rangle + \gamma v + (1-\gamma)v']K(z)} - 1 \right) f(x, y) dz dy \\ &< \gamma \int_{\mathbb{R}^d\times\mathbb{R}^q} \left(e^{(\langle u, y\rangle + v)K(z)} - 1 \right) f(x, y) dz dy \\ &\quad + (1-\gamma) \int_{\mathbb{R}^d\times\mathbb{R}^q} \left(e^{(\langle u', y\rangle + v')K(z)} - 1 \right) f(x, y) dz dy, \end{aligned}$$

where the last inequality follows from the fact that $x \mapsto e^x$ is strictly convex.
Since $|e^t - 1| \leq |t|e^{|t|}$ $\forall t \in \mathbb{R}$ and K is bounded and integrable, (7) imply that

$$\begin{aligned}
&\int_{\mathbb{R}^d \times \mathbb{R}^q} \left| \left(e^{(\langle u,y\rangle + v)K(z)} - 1 \right) f(x,y) \right| dzdy \\
&\leq \int_{\mathbb{R}^d \times \mathbb{R}^q} |(\langle u,y\rangle + v)\, K(z)|\, e^{|(\langle u,y\rangle+v)K(z)|} f(x,y)dzdy \\
&\leq e^{|v|\|K\|_\infty} \|u\| \int_{\mathbb{R}^d} |K(z)|\, dz \int_{\mathbb{R}^q} \|y\| e^{\|u\|\|y\|\|K\|_\infty} f(x,y)dy \\
&\quad + e^{|v|\|K\|_\infty} |v| \int_{\mathbb{R}^d} |K(z)|\, dz \int_{\mathbb{R}^q} e^{\|u\|\|y\|\|K\|_\infty} f(x,y)dy < \infty,
\end{aligned}$$

which ensures the existence of Ψ_x.
Next, set

$$h_x(u,v,y,z) = \left[e^{(\langle u,y\rangle+v)K(z)} - 1 \right] f(x,y).$$

Since h_x is differentiable with respect to (u,v) and

$$\nabla h_x(u,v,y,z) = \begin{pmatrix} y e^{(\langle u,y\rangle+v)K(z)} K(z) f(x,y) \\ e^{(\langle u,y\rangle+v)K(z)} K(z) f(x,y) \end{pmatrix},$$

using Assumption (A1) and (7), it can be seen that Ψ_x is differentiable on $\mathbb{R}^q \times \mathbb{R}$. Since Ψ_x is a smooth convex on $\mathbb{R}^q \times \mathbb{R}$, it follows that Ψ_x is essentially smooth so that I_x is a good rate function on $\mathbb{R}^q \times \mathbb{R}$ (see Dembo and Zeitouni [6]), which proves the first part of Proposition 3.
Now, observe that $\overset{\circ}{\mathcal{D}}(\Psi_x) = \mathbb{R}^q \times \mathbb{R}$, and since Ψ_x is strictly convex, it holds that the pair $\left(\overset{\circ}{\mathcal{D}}(\Psi_x), \Psi_x \right)$ is a convex function of Legendre type. It follows that $\left(\overset{\circ}{\mathcal{D}}(I_x), I_x \right)$ is a convex function of Legendre type (See Rockafellar [17]). Thus, Part 2 of Proposition 3 follows from Theorem 26.5 of Rockafellar [17].

- Let us now assume that $\lambda(S_-) = 0$. Thus

$$\Psi_x(u,v) = \int_{\mathbb{R}^d \times \mathbb{R}^q} \left(e^{u^T y K(z)} e^{vK(z)} - 1 \right) \mathbb{1}_{S_+}(z) f(x,y) dzdy.$$

For each $u \in \mathbb{R}^q$, the function $v \mapsto \left(e^{u^T y K(z)} e^{vK(z)} - 1 \right) \mathbb{1}_{S_+}(z) f(x,y)$ is increasing in v and goes to $-f(x,y)$ when $v \to -\infty$. Thus $\lim_{v\to-\infty} \Psi_x(u,v) = -g(x)\lambda(S_+)$ and $I_x(\vec{0},0) = g(x)\lambda(S_+)$. Now, when $t_1 \neq \vec{0}$, let us show that

$$I_x(t_1,0) = +\infty.$$

Let $M > 0$, $\epsilon > 0$ and set $u = (M+\epsilon)t_1/\|t_1\|^2$. Let $v \in \mathbb{R}$ such that

$$\begin{cases} -\Psi_x(u,v) > g(x)\lambda(S_+) - \epsilon & \text{if } \lambda(S_+) < \infty \\ -\Psi_x(u,v) > M & \text{if } \lambda(S_+) = \infty. \end{cases}$$

Then, on the one hand, when $\lambda(S_+) < \infty$, we have

$$u^T t_1 - \Psi_x(u,v) \geq M + \epsilon + g(x)\lambda(S_+) - \epsilon > M.$$

On the other hand, when $\lambda(S_+) = \infty$, we get

$$u^T t_1 - \Psi_x(u,v) \geq M + \epsilon + M > M.$$

It follows that $\sup_{u,v} \left(u^T t_1 - \Psi_x(u,v) \right) = +\infty$. ■

3.5.2 Proof of Proposition 4

Following the same lines of the proof of Proposition 3, we prove Proposition 4. When $\lambda(S_-) = 0$, for each $u \in \mathbb{R}^q$ and $s \in]0,1]$, the map $v \mapsto s^{-ad}\left(e^{s^{ad}u^T yK(z)}e^{vK(z)} - 1\right)\mathbb{1}_{S_+}(z)f(x,y)$ is increasing in v and goes to $-s^{-ad}f(x,y)$ when $v \to -\infty$. We deduce that $\lim_{v\to-\infty}\tilde{\Psi}_{a,x}(u,v) = -g(x)\lambda(S_+)\int_0^1 s^{-ad} = -g(x)\lambda(S_+)/(1-ad)$ and $I_x(\vec{0},0) = g(x)\lambda(S_+)/(1-ad)$. ■

3.6 Proof of Propositions 1 and 2

3.6.1 Proof of Proposition 1

(i) Let us prove the first part of Proposition 1.

- If $\alpha < I_x(\vec{0},0)$, set

$$G = \{(a,b) \in \mathbb{R}^q \times \mathbb{R},\ \ I_x(a,b) \le \alpha\} \ \text{ and } \ \hat{G} = \{(s,t) \in \mathbb{R}^q \times \mathbb{R},\ \ \hat{I}_x(s,t) \le \alpha\}.$$

  We first show that $\hat{G}$ is a compact subset of $\mathbb{R}^q \times \mathbb{R}$.
  First, observe that since I_x is a good rate function, G is a compact subset of $\mathbb{R}^q \times \mathbb{R}$. Let us define the following function

$$\begin{array}{rcl} F : \mathbb{R}^q \times \mathbb{R} & \to & \mathbb{R}^q \times \mathbb{R} \\ (s,t) & \mapsto & (st,t). \end{array}$$

  Observe that F is continuous and $\hat{G} = F^{-1}(G)$. We deduce that $\hat{G}$ is a closed subset of $\mathbb{R}^q \times \mathbb{R}$.
  Now, let (s_n,t_n) be a sequence of real numbers of $\hat{G}$, there exists $(x_n,y_n) \in G$ such that $(x_n,y_n) = F(s_n,t_n) = (s_nt_n,t_n) \in G$.
  The compactness of G on $\mathbb{R}^q \times \mathbb{R}$ ensures that there exists a sequence of real numbers $(x_{n_k},y_{n_k}) \in G$ such that $(x_{n_k},y_{n_k}) \to (x_0,y_0)$ as $k \to \infty$, where $(x_0,y_0) \in G$. Therefore, $(s_{n_k}t_{n_k},t_{n_k}) \to (x_0,y_0)$ as $k \to \infty$.
  Noting that Condition (C) ensures that $\forall s \in \mathbb{R}^q$, $I_x(s,0) \ge I_x(\vec{0},0) > \alpha$ so that $(s,0) \notin G$.
  It follows that $y_0 \neq 0$, and thus $t_{n_k} \to y_0$ and $s_{n_k} \to s_0$ as $k \to \infty$, where $s_0 = x_0/y_0$. We deduce that $(s_{n_k},t_{n_k}) \to (s_0,y_0)$ as $k \to \infty$, so that $(s_0,y_0) \in \hat{G}$. Thus $\hat{G}$ is a compact set. Now we claim that the set $A = \{s,\ J(s) \le \alpha\}$ is the image of $\hat{G}$ by the continuous map $\pi : (s,t) \mapsto s$, and thus it is a compact.
  Indeed, clearly $\pi(\hat{G}) \subset A$. For the opposite inclusion, consider $\alpha < \alpha' < I_x\left(\vec{0},0\right)$; the set $\hat{G}' = \left\{\hat{I}_x(s,t) \le \alpha'\right\}$ is compact. Let $s_0 \in A$, since $J(s_0) \le \alpha$, we have $J(s_0) = \inf_{(s_0,t)\in\hat{G}'}\hat{I}_x(s_0,t)$; by compacity, there exists t_0 such that $J(s_0) = \hat{I}_x(s_0,t_0)$; $(s_0,t_0) \in \hat{G}$ and $\pi(s_0,t_0) = s_0$, thus $A \subset \pi(\hat{G})$.
- If $\alpha \ge I_x(\vec{0},0)$, let $s \in \mathbb{R}^q$, we have

$$\begin{array}{rcl} J(s) & \le & I_x(st,t) \ \ \forall t \\ & \le & I_x(\vec{0},0) \\ & \le & \alpha. \end{array}$$

  We deduce that $\mathbb{R}^q \subseteq \{J(s) \le \alpha\}$ and the second part of Proposition 1 (i) follows.

(ii) It is an obvious consequence of (i) and the definitions of J and J^*.

(iii) Assume that $J^*(s) < \infty$.

- If $I_x(\vec{0}, 0) > \inf_t I_x(st, t)$, then

$$\inf_t I_x(st, t) \quad = \inf_{t \neq 0} I_x(st, t),$$

so that $J(s) = J^*(s)$.

- If $I_x(\vec{0}, 0) = \inf_t I_x(st, t)$, since $J^*(s) < \infty$, there exists $t_0 \neq 0$ such that $I_x(st_0, t_0) < \infty$. By the convexity of I_x, we have for any $\nu \in]0, 1]$, $I_x(st_0\nu, t_0\nu) < \infty$ and

$$I_x(st_0\nu, t_0\nu) \leq \nu I_x(st_0, t_0) + (1 - \nu) I_x(\vec{0}, 0).$$

We deduce that

$$0 \leq I_x(st_0\nu, t_0\nu) - I_x(\vec{0}, 0) \leq \nu \left(I_x(st_0, t_0) - I_x(\vec{0}, 0) \right),$$

and if we take $\nu \to 0$, the third part of Proposition 1 follows.

(iv) Let us suppose that $\alpha < I_x(\vec{0}, 0)$ and let $s \in \{J^*(s) \leq \alpha\}$, then we have $J^*(s) < \infty$. We deduce from (iii) that $J(s) = J^*(s)$. It follows that $J(s) \leq \alpha$, which ensures that $s \in \{J(s) \leq \alpha\}$. Conversely, if $s \in \{J(s) \leq \alpha\}$, then

$$I_x(\vec{0}, 0) > \inf_t I_x(st, t),$$

so that

$$\inf_t I_x(st, t) = \inf_{t \neq 0} I_x(st, t).$$

That is $J(s) = J^*(s)$. Therefore, $J^*(s) \leq \alpha$, which ensures that $s \in \{J^*(s) \leq \alpha\}$, and thus Proposition 1 is proved. ■

3.6.2 Proof of Proposition 2

Proposition 2 is proved by following the same approach as for the proof of Proposition 1 with replacing I_x, J and J^* by $\tilde{I}_{a,x}$, $\tilde{J}_a$ and $\tilde{J}_a^*$ respectively. ■

References

[1] Ahmad, I.A. and Lin, P. (1976), Nonparametric sequential estimation of a multiple regression function. *Bull. Math. Statist.*, **17**, pp. 63-75.

[2] Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987), Regular variation. *Cambridge University Press.*

[3] Bosq, D. (1985), Nonparametric statistics for stochastic processes. *Lectures notes in Control and Inform. Sc. Springer.*

[4] Collomb, G. (1984), Proprietés de convergence presque-complète du prédicateur à noyau. *Z. Wahrsch. verw. Gebiete*, **66**, pp. 441-460.

[5] Collomb, G. and Härdle W. (1986), Strong uniform convergence rates in robust nonparametric time series analysis and prediction: kernel regression estimation from dependent observations. *Stoch. Proc. and their Appl.* **23**, pp. 77-89.

[6] Dembo, A. and Zeitouni, O. (1998), Large deviations techniques and applications. Springer, Applications of mathematics, New-York.

[7] Devroye, L. (1979), The uniform convergence of the Nadaraya-Watson regression function estimate. *Can. J. Stat.* **6**, pp. 179-191.

[8] Devroye, L. and Wagner, T.J. (1980), On the L_1 convergence of kernel estimators of regression function with applications in discrimination. *Wahrsch. verw. Gebiete*, **51**, pp. 15-25.

[9] Feller, W. (1970), An introduction to probability theory and its applications, Second edition Volume II, Wiley.

[10] Joutard, C. (2006), Sharp large deviations in nonparametric estimation. *J. of Nonparam. Stat.*, **18**, pp. 293-306.

[11] Louani, D. (1999), Some large deviations limit theorems in conditionnal nonparametric statistics. *Statistics*, **33**, pp. 171-196.

[12] Mack, Y. P. and Silverman, B.W. (1982), Weak and strong uniform consistency of kernel regression estimates. *Z. Wahrsch. verw. Gebiete*, **61**, pp. 405-415.

[13] Mokkadem, A., Pelletier, M. and Thiam, B. (2005), Large and moderate deviations principles for recursive kernel estimators of a multivariate density and its partial derivatives. arXiv:math.ST/0601429.

[14] Mokkadem, A., Pelletier, M. and Worms, J. (2005), Large and moderate deviations principles for kernel estimation of a multivariate density and its partial derivatives. *Austral. J. Statist.*, **4**, pp. 489-502.

[15] Nadaraya, E. A. (1964), On estimating regression. *Theory Probab. Appl.* **10**, pp. 186-190.

[16] Prakasa Rao, B. L. S. (1983), Nonparametric functionnal estimation. *Academic presse New York.*

[17] Rockafellar, R. T. (1970), Convex analysis. Princeton University Press.

[18] Roussas, G. (1992), Exact rates of almost sure convergence of a recursive kernel estimate of a probability density function: Application to regression and hazard rate estimate. *J. of Nonparam. Statist.* **3**, pp. 171-195.

[19] Senoussi, R. (1991), Loi du log itéré et identification, Thèse. *Université Paris-Sud.*

[20] Watson, G. S. (1964), Smooth regression analysis. *Sankhya Ser. A.* **26**, pp. 359-372.

[21] Worms, J. (2000), Principes de déviations modérées pour des martingales et applications statistiques, Thèse. *Université Marne-la-vallée.*