ADDITIVE REGRESSION MODEL FOR CONTINUOUS TIME PROCESSES

Mohammed DEBBARH and Bertrand MAILLOT

Université Paris 6
175, Rue du Chevaleret, 75013 Paris.
debbarh@ccr.jussieu.fr.

Key Words: Additive regression; Continuous time processes; Curse of dimensionality; Marginal integration.

ABSTRACT

In the setting of additive regression model for continuous time process, we establish the optimal uniform convergence rates and optimal asymptotic quadratic error of additive regression. To build our estimate, we use the marginal integration method.

1 Introduction and motivations

The multivariate regression function estimation is an important problem which has been extensively treated for discrete time processes. It is well-known from (11) that the additive regression models bring out a solution to the problem of the curse of dimensionality in non-parametric multivariate regression estimation, which is characterized by a loss in the rate of convergence of the regression function estimator when the dimension of the covariates increases. Additive models allow to reach even univariate rate when these models fit well.
For continuous time processes, (2) obtained the optimal rate for the estimator of multivariate regression, which is the same as in the i.i.d. case. He even proved that, for processes with irregular paths, it is possible to reach the *parametric rate*. This one, called the superoptimal rate, does not depend on the dimension of the variables, but the needed conditions on the processes are very strong. That is the reason why it is relevant to study additive models to bring out a solution to the problem of the curse of dimensionality.

Let $\mathbf{Z}_t = (\mathbf{X}_t, Y_t), (t \in \mathbb{R})$ be a $\mathbb{R}^d \times \mathbb{R}$-valued measurable stochastic process defined on a probability space $(\Omega, \mathcal{A}, P)$. Denote by ψ a given real measurable function. We consider the additive regression function associated to $m_\psi(Y)$ defined by,

$$m_\psi(\mathbf{x}) \quad = \quad E\left(\psi(Y) \mid \mathbf{X} = \mathbf{x}\right), \ \forall \ \mathbf{x} = (x_1, ..., x_d) \in \mathbb{R}^d, \tag{1}$$

$$= \quad \mu + \sum_{l=1}^{d} m_l(x_l) := m_{\psi,add}(\mathbf{x}). \tag{2}$$

Let K_1, K_2, K_3 and K, be kernels respectively defined on $\mathbb{R}$, $\mathbb{R}^{d-1}$, $\mathbb{R}^d$ and $\mathbb{R}^d$. We denote by $\hat{f}_T$ the estimate of f, the density function of the covariable $\mathbf{X}$, (see (1)), that is,

$$\hat{f}_T(\mathbf{x}) = \frac{1}{Th_T^d}\int_0^T K\Big(\frac{\mathbf{x}-\mathbf{X}_s}{h_T}\Big)ds,$$

where (h_T) is a positive real function. In estimating the regression function defined in (1), we use the following two estimators (see for exemple (3) and (5))

$$\widetilde{m}_{\psi,T}(\mathbf{x}) = \int_0^T W_{T,t}(\mathbf{x})\psi(Y_t)dt \quad \text{with} \quad W_{T,t}(\mathbf{x}) = \frac{K_3(\frac{\mathbf{x}-\mathbf{X}_t}{h_{1,T}})}{Th_{1,T}^d \hat{f}_T(\mathbf{X}_t)}, \tag{3}$$

and

$$\widetilde{m}_{\psi,T,l}(\mathbf{x}) := \int_0^T W_{T,t}^l(\mathbf{x})\psi(Y_t)dt \quad \text{with} \quad W_{T,t}^l(\mathbf{x}) = \frac{K_1(\frac{x_l-X_{t,l}}{h_{1,T}})K_2(\frac{\mathbf{x}_{-l}-\mathbf{X}_{t,-l}}{h_{2,T}})}{Th_{1,T}h_{2,T}^{d-1}\hat{f}_T(\mathbf{X}_t)}, \tag{4}$$

where $(h_{j,T}), j = 1,2$ are positive real functions. Let $q_1, ..., q_d$ be d density functions defined in $\mathbb{R}$. Setting $q(\mathbf{x}) = \prod_{l=1}^d q_l(x_l)$ and $q_{-l}(\mathbf{x}_{-l}) = \prod_{j\neq l} q_j(x_j)$. To estimate the additive components of the regression function, we use the marginal integration method (see (6) and (8)). We obtain then

$$\eta_l(x_l) = \int_{\mathbb{R}^{d-1}} m_\psi(\mathbf{x})q_{-l}(\mathbf{x}_{-l})d\mathbf{x}_{-l} - \int_{\mathbb{R}^d} m_\psi(\mathbf{x})q(\mathbf{x})d\mathbf{x}, \quad l = 1, ..., d, \tag{5}$$

in such a way that the following two equalities hold,

$$\eta_l(x_l) = m_l(x_l) - \int_{\mathbb{R}} m_l(z)q_l(z)dz, \quad l = 1, ..., d, \tag{6}$$

$$m_\psi(\mathbf{x}) = \sum_{l=1}^d \eta_l(x_l) + \int_{\mathbb{R}^d} m_\psi(\mathbf{z})q(\mathbf{z})d\mathbf{z}. \tag{7}$$

In view of (6) and (7), we note thatη_l and m_l are equal up to an additional constant. Therefore, η_l is also an additive component, fulfilling a different identifiability condition. From (4) and (5), a natural estimate of this l-th component is given by

$$\widehat{\eta}_l(x_l) = \int_{\mathbb{R}^{d-1}} \widetilde{m}_{\psi,T,l}(\mathbf{x})q_{-l}(\mathbf{x}_{-l})d\mathbf{x}_{-l} - \int_{\mathbb{R}^d} \widetilde{m}_{\psi,T,l}(\mathbf{x})q(\mathbf{x})d\mathbf{x}, \; l = 1, ..., d, \tag{8}$$

from which we deduce the estimate $\widehat{m}_{\psi,T,add}$ of the additive regression function,

$$\widehat{m}_{\psi,T,add}(\mathbf{x}) \;=\; \sum_{l=1}^{d}\widehat{\eta}_l(x_l)+\int_{\mathbb{R}^d}\widetilde{m}_{\psi,T}(\mathbf{x})q(\mathbf{x})d\mathbf{x}. \tag{9}$$

Before stating our results, we introduce some additional notations and our assumptions. Let $\mathcal{C}_1,...,\ \mathcal{C}_d$, be d compact intervals of $\mathbb{R}$ and set $\mathcal{C}=\mathcal{C}_1\times...\times\mathcal{C}_d$. For every subset $\mathcal{E}$ of $\mathbb{R}^q$, $q\geq 1$, and any $\delta>0$, introduce the δ-neighborhood $\mathcal{E}^\delta$ of $\mathcal{E}$, namely, $\mathcal{E}^\delta=\{\mathbf{x}:\inf_{\mathbf{y}\in\mathcal{E}}\|\mathbf{x}-\mathbf{y}\|_{\mathbb{R}^q}<\delta\}$, with $\|\cdot\|_{\mathbb{R}^q}$ standing for the euclidian norm on $\mathbb{R}^q$.

(C.1) There exists a positive constant M such that $|\psi(y)|\leq M<\infty$.

(C.2) The function m_ψ is k-times continuously differentiable, $k\geq 1$, and

$$\sup_{\mathbf{x}}\left|\frac{\partial^k m_\psi}{\partial x_\ell^k}(\mathbf{x})\right|<\infty;\ \ell=1,...,d.$$

Denote by f_ℓ, $\ell=1,...,d$ the density functions of X_ℓ, $\ell=1,...,d$. The functions f and $f_\ell,\ell=1,...,d$, are supposed to be continuous, bounded and

$(F.1)\forall\mathbf{x}\in\mathcal{C}^\delta,\ f(\mathbf{x})>0$ and $f_\ell(x_\ell)>0,\ \ell=1,...,d.$

$(F.2)$ f is k'-times continuously differentiable on $\mathcal{C}^\delta, k'>kd$.

$(F.3)$ For all $0<\lambda\leq 1, \left|\frac{\partial f^{(k)}}{\partial x_1^{j_1}...\partial_d^{j_d}}(\mathbf{x}')-\frac{\partial f^{(k)}}{\partial x_1^{j_1}...\partial_d^{j_d}}(\mathbf{x})\right|\leq L\|\mathbf{x}'-\mathbf{x}\|^\lambda$ with $j_1+...+j_d=k'$. Where $\|.\|$ is a norm on $\mathbb{R}^d$ and L is a positive constant.

The kernels K_1, K_2, K_3 and K are assumed to fulfill the following conditions

(K.1) K_1, K_2, K_3 and K are continuous respectively on the compact supports $S_1\subset\mathcal{C}_1$, $S_2\subset\mathcal{C}_2\times...\times\mathcal{C}_d$, $S_3\subset\mathcal{C}$ and S,

(K.2) $\int K=1$ and $\int K_j=1,\quad j=1,2,3,$

(K.3) K_1, K_2 and K_3 are of order k,

(K.4) K is of order k'.

(K.5) K_1 is a Lipschitz function.

The density functions q_ℓ, $\ell = 1, ..., d$, satisfy the following assumption

$(Q.1)$ For any $1 \leq l \leq d$, q_ℓ has k continuous and bounded derivatives, with a compact support included in $\mathcal{C}_\ell$.

There exists a set $\Gamma \in \mathcal{B}_{\mathbb{R}^2}$ containing $D = \{(s,t) \in \mathbb{R}^2 : s = t\}$ such that

$(D.1)$ $f_{(\mathbf{X}_s,Y_s),(\mathbf{X}_t,Y_t)} - f_{(\mathbf{X}_s,Y_s)} \bigotimes f_{(\mathbf{X}_t,Y_t)}$ exists everywhere for $(s,t) \in \Gamma^C$,

$(D.2)$ $A_f(\Gamma) := \sup_{(s,t)\in\Gamma^C} \sup_{\mathbf{x},\mathbf{y}\in\mathcal{C}^\delta\times\mathcal{C}^\delta} \int_{u,v\in\mathbb{R}^2} |f_{(\mathbf{X}_s,Y_s),(\mathbf{X}_t,Y_t)}(\mathbf{x},u,\mathbf{y},v)$
$- f_{(\mathbf{X}_s,Y_s)}(\mathbf{x},u) f_{(\mathbf{X}_t,Y_t)}(\mathbf{y},t)| du dv < \infty$,

$(D.3)$ there exists $\ell_\Gamma < \infty$ and T_0 such that, $\forall T > T_0$, $\frac{1}{T}\int_{[0,T]^2\cap\Gamma} ds dt \leq \ell_\Gamma$.

We work under the following conditions upon the smoothing parameters h_T and $h_{j,T}$, $j = 1, 2$,

$(H.1)$ $h_T = c'\left(\frac{\log T}{T}\right)^{1/(2k'+d)}$, for a fixed $0 < c' < \infty$,

$(H.2)$ $h_{1,T} = c_1 T^{-1/(2k+1)}$ and $h_{2,T} = c_2 T^{-1/(2k+1)}$, for fixed $0 < c_1, c_2 < \infty$,

$(H.2)'$ $h_{1,T} = c_1(\frac{\log(T)}{T})^{1/(2k+1)}$ and $h_{2,T} = c_2(\frac{\log(T)}{T})^{1/(2k+1)}$, for fixed $0 < c_1, c_2 < \infty$.

Throughout this work, we use the α-mixing dependance structure where the associated coefficient is defined, for every σ-fields $\mathcal{A}$ and $\mathcal{B}$ by

$$\alpha(\mathcal{A}, \mathcal{B}) = \sup_{(A,B)\in(\mathcal{A},\mathcal{B})} |P(A \cap B) - P(A)P(B)|.$$

For all Borelian set I in $\mathbb{R}^+$ the σ-algebra defined by $(Z_t, t \in I)$ is denoted by $\sigma(Z_t, t \in I)$. Writing $\alpha(u) = \sup_{t\in\mathbb{R}_+} \alpha(\sigma(Z_v, v \leq t), \sigma(Z_v, v \geq t + u))$, we use the condition

$(A.1)$ $\alpha(t) = \mathcal{O}(t^{-b})$ with $b > 2d + 10 + \frac{6+4d}{k}$.

Theorem 1 *Under the conditions* $(A.1)$, $(C.1)-(C.2)$, $(F.1)-(F.3)$, $(K.1)-(K.4)$, $(Q.1)$, $(D.1)-(D.3)$ *and* $(H.1)-(H.2)$, *we have, for all* $\mathbf{x} \in \mathcal{C}^\delta$

$$E(\widehat{m}_{\psi,T,add}(\mathbf{x}) - m_\psi(\mathbf{x}))^2 = \mathcal{O}(T^{-2k/2k+1}).$$

Theorem 2 *Under the conditions* $(A.1)$, $(C.1)-(C.2)$, $(F.1)-(F.3)$, $(K.1)-(K.5)$, $(Q.1)$, $(D.1)-(D.3)$ *and* $(H.1)-(H.2)'$, *we have*

$$\sup_{\mathbf{x}\in\mathcal{C}} |\widehat{m}_{\psi,T,add}(\mathbf{x}) - m_\psi(\mathbf{x})| = \mathcal{O}\Big(\Big(\frac{\log T}{T}\Big)^{k/2k+1}\Big) \quad a.s.$$

2 Proofs

The proofs of our theorems are split into two steps. First, we consider the case where the density is assumed to be known. Subsequently, we treat the general case when f is unknown. Denote by $\hat{\hat{\eta}}$, $\widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x})$ and $\widetilde{\widetilde{m}}_{\psi,T,l}(\mathbf{x})$ the versions of $\hat{\eta}$, $\widetilde{m}_{\psi,T}(\mathbf{x})$ and $\widetilde{m}_{\psi,T,l}(\mathbf{x})$ associated to a known (formally, we replace $\hat{f}_T$ by f in the expressions (3), (4) and $\widetilde{m}_{\psi,T,l}(\mathbf{x})$ by $\widetilde{\widetilde{m}}_{\psi,T,l}(\mathbf{x})$ in (8).

Introduce now the following quantities (see, for the discrete case (4)), we establish the proof for the first component,

$$\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x}) = \sum_{l=1}^{d} \widehat{\widehat{\eta}}_l(x_l) + \int_{\mathbb{R}^d} \widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x}) q(\mathbf{x}) d\mathbf{x}. \tag{10}$$

$$\widetilde{Y}_{\psi,T,t} = \psi(Y_t) \int_{\mathbb{R}^{d-1}} \frac{1}{h_{2,T}^{d-1}} K_2\Big(\frac{\mathbf{x}_{-1} - \mathbf{X}_{t,-1}}{h_{2,T}}\Big) \frac{q_{-1}(\mathbf{x}_{-1})}{f(X_{t,-1}|X_{t,1})} d\mathbf{x}_{-1}, \tag{11}$$

$$\mathcal{G}(\mathbf{u}_{-1}) = \int_{\mathbb{R}^{d-1}} \frac{1}{h_{2,T}^{d-1}} K_2\Big(\frac{\mathbf{x}_{-1} - \mathbf{u}_{-1}}{h_{2,T}}\Big) q_{-1}(\mathbf{x}_{-1}) d\mathbf{x}_{-1}, \tag{12}$$

$$\widehat{\alpha}_1(x_1) = \frac{1}{Th_{1,T}} \int_0^T \frac{\widetilde{Y}_{\psi,T,t}}{f_1(X_{t,1})} K_1\Big(\frac{x_1 - X_{t,1}}{h_{1,T}}\Big) dt, \text{ for } x_1 \in \mathcal{C}_1, \tag{13}$$

$$\widetilde{m}_T(x_1) = E\Big(\widetilde{Y}_{\psi,T,t} \Big| X_{t,1} = x_1\Big), \tag{14}$$

$$C_T = \mu + \int_{\mathbb{R}^{d-1}} \sum_{j=2}^{d} m_j(u_j) \mathcal{G}(\mathbf{u}_{-1}) d\mathbf{u}_{-1}, \tag{15}$$

$$\widehat{C}_T = \int_{\mathbb{R}^d} \widetilde{\widetilde{m}}_{\psi,T,1}(\mathbf{x}) q(\mathbf{x}) d\mathbf{x}, \tag{16}$$

$$C = \int_{\mathbb{R}} m_1(x_1) q_1(x_1) dx_1. \tag{17}$$

The following Lemma is of particular interest to establish the result of theorem (1). Note that (19) is "only" be instrumental in the proof of (20).

Lemma 1 *Under the assumptions* $(C.1)-(C.2)$, $(F.1)-(F.2)$, $(K.1)$, $(Q.1)$ *and* $(H.2)$, *we have*

$$E(\hat{C}_T - C_T + C)^2 = \mathcal{O}\Big(T^{-2k/(2k+1)}\Big), \tag{18}$$

$$\mathrm{Var}(\hat{\alpha}_1(x_1)) = \mathcal{O}\Big(T^{-2k/(2k+1)}\Big), \tag{19}$$

$$E(\widehat{\widehat{\eta}}_1(x_1) - \eta_1(x_1))^2 = \mathcal{O}\Big(T^{-2k/(2k+1)}\Big). \tag{20}$$

Proof: According to Fubini's Theorem and under the additive model assumption, we have

$$\begin{aligned}
\mathbb{E}(\hat{C}_T - C_T) &= \mathbb{E}\Big\{\int_{\mathbb{R}^d} \widetilde{\widetilde{m}}_{\psi,T,1}(\mathbf{x})q(\mathbf{x})d\mathbf{x} - \mu - \int_{\mathbb{R}^{d-1}} \sum_{j=2}^{d} m_j(u_j)\mathcal{G}(\mathbf{u}_{-1})d\mathbf{u}_{-1}\Big\} \\
&= \int_{\mathbb{R}^d} E(\widetilde{\widetilde{m}}_{\psi,T,1}(\mathbf{x}))q(\mathbf{x})d\mathbf{x} - \mu - \int_{\mathbb{R}^{d-1}} \sum_{j=2}^{d} m_j(u_j)\mathcal{G}(\mathbf{u}_{-1})d\mathbf{u}_{-1} \\
&= \int_{\mathbb{R}^d} \frac{1}{h_{1,t}} m_\psi(\mathbf{u})\mathcal{G}(\mathbf{u}_{-1}) \int_{\mathbb{R}} K_1\Big(\frac{x_1-u_1}{h_{1,T}}\Big)q_1(x_1)dx_1 d\mathbf{u} \\
&\qquad - \mu - \int_{\mathbb{R}^{d-1}} \sum_{j=2}^{d} m_j(u_j)\mathcal{G}(\mathbf{u}_{-1})d\mathbf{u}_{-1} \\
&= \sum_{j=1}^{d} \int_{\mathbb{R}^d} \frac{1}{h_{1,T}} m_j(u_j)\mathcal{G}(\mathbf{u}_{-1}) \int_{\mathbb{R}} K_1\Big(\frac{x_1-u_1}{h_{1,T}}\Big)q_1(x_1)dx_1 d\mathbf{u} \\
&\qquad - \int_{\mathbb{R}^{d-1}} \sum_{j=2}^{d} m_j(u_j)\mathcal{G}(\mathbf{u}_{-1})d\mathbf{u}_{-1} \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{1}{h_{1,T}} m_1(u_1)K_1\Big(\frac{x_1-u_1}{h_{1,T}}\Big)q_1(x_1)dx_1 du_1.
\end{aligned}$$

Setting $v_1 h_{1,T} = x_1 - u_1$ and using a Taylor expansion, we get, by $(C.2)$ and $(K.1)-(K.3)$,

$$\begin{aligned}
\mathbb{E}(\hat{C}_T - C_T) - C &= \int_{\mathbb{R}}\int_{\mathbb{R}} q_1(x_1)m_1(x_1 - h_{1,T}v_1)K_1(v_1)dv_1dx_1 - C \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} q_1(x_1)[m_1(x_1 - h_{1,T}v_1) - m_1(x_1)]K_1(v_1)dv_1dx_1 \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} q_1(x_1)\Big[\frac{(-h_{1,T})^k v_1^k}{k!} m_1^{(k)}(x_1)\Big]K_1(v_1)dv_1dx_1 \\
&\quad + o(h_{1,T}^k).
\end{aligned} \tag{21}$$

Under $(H.2)$, it follows that,

$$\Big[\mathbb{E}\Big(\hat{C}_T - C_T - C\Big)\Big]^2 = \mathcal{O}(T^{-2k/(2k+1)}). \tag{22}$$

The Fubini's theorem gives us

$$\begin{aligned}
\mathrm{Var}(\hat{C}_T) &= \frac{1}{(Th_{1,T})^2}\mathrm{Var}\bigg(\int_0^T \frac{\psi(Y_t)}{f(\mathbf{X}_t)}\mathcal{G}(\mathbf{X}_{t,-1})\int_{\mathbb{R}} K_1\Big(\frac{x_1 - X_{t,1}}{h_{1,T}}\Big)q_1(x_1)dx_1dt\bigg) \\
&= \frac{1}{(Th_{1,T})^2}\int_{t,s\in[0;T]^2} \mathrm{Cov}\bigg(\frac{\psi(Y_t)}{f(\mathbf{X}_t)}\mathcal{G}(\mathbf{X}_{t,-1})\int_{\mathbb{R}} K_1\Big(\frac{x_1 - X_{t,1}}{h_{1,T}}\Big) \\
&\quad q_1(x_1)dx_1; \frac{\psi(Y_s)}{f(\mathbf{X}_s)}\mathcal{G}(\mathbf{X}_{s,-1})\int_{\mathbb{R}} K_1\Big(\frac{y_1 - X_{s,1}}{h_{1,T}}\Big)q_1(y_1)dy_1\bigg)dsdt.
\end{aligned} \tag{23}$$

Under $(C.1)$, $(F.1)$, $(K.1)-(K.2)$ and $(Q.1)$, there exists a finite constant M_3 such that, for T large enough,

$$\inf\bigg\{a \;:\; P\Big(\frac{\psi(Y_t)}{f(\mathbf{X}_t)}\mathcal{G}(\mathbf{X}_{t,-1})\int_{\mathbb{R}} K_1\Big(\frac{x_1 - X_{t,1}}{h_{1,T}}\Big)q_1(x_1)dx_1 > a\Big) = 0\bigg\} \le h_{1,T}M_3.$$

Thus, using the Billingsley's inequality and the condition $(A.1)$,

$$\mathrm{Var}(\hat{C}_T) \le \frac{8M_3^2}{T^2}\int_{s\in[0;T]}\int_{t\in[0,T-s]} \alpha_t dtds = \mathcal{O}\Big(\frac{1}{T}\Big). \tag{24}$$

Finally, by combining the statements (22) and (24), we obtain (18).

Proof of (19). Recalling (13), we have

$$\begin{aligned}
\mathrm{Var}(\hat{\alpha}_1(x_1)) &= \frac{1}{T^2}\Big[\int_{[0,T]^2\cap\Gamma} \mathrm{Cov}\Big(\frac{\widetilde{Y}_{\psi,T,t}}{f_1(X_{t,1})h_{1,T}}K_1\Big(\frac{x_1 - X_{t,1}}{h_{1,T}}\Big), \\
&\quad \frac{\widetilde{Y}_{\psi,T,s}}{f_1(X_{s,1})h_{1,T}}K_1\Big(\frac{x_1 - X_{s,1}}{h_{1,T}}\Big)\Big)dsdt \\
&\quad + \int_{[0,T]^2\cap\Gamma^c} \mathrm{Cov}\Big(\frac{\widetilde{Y}_{\psi,T,t}}{f_1(X_{i,1})h_{1,T}}K_1\Big(\frac{x_1 - X_{t,1}}{h_{1,T}}\Big), \\
&\quad \frac{\widetilde{Y}_{\psi,T,s}}{f_1(X_{s,1})h_{1,T}}K_1\Big(\frac{x_1 - X_{s,1}}{h_{1,T}}\Big)\Big)dsdt\Big] \\
&:= A + B.
\end{aligned} \tag{25}$$

For the first term, noting that, under $(C.1)$, $(F.1)$, $(K.1)-(K.2)$ and $(Q.1)$, there exists a finite constant M_4 such that, for T large enough,

$$M_4 \ge \inf\bigg\{a \;:\; P\bigg(\frac{\widetilde{Y}^2_{\psi,T,0}}{f_1(X_{0,1})^2}\Big|K_1\Big(\frac{x_1 - X_{0,1}}{h_{1,T}}\Big)\Big| > a\bigg) = 0\bigg\}.$$

Thus, we have

$$\begin{aligned} A &\le \frac{1}{T^2 h_{1,T}^2} \int_{[0,T]^2\cap\Gamma} \mathbb{E}\left(\frac{\widetilde{Y}_{\psi,T,0}}{f_1(X_{0,1})} K_1\left(\frac{x_1 - X_{0,1}}{h_{1,T}}\right)\right)^2 dsdt \\ &\le \frac{M_4}{T^2 h_{1,T}^2} \int_{[0,T]^2\cap\Gamma} \int_{\mathbb{R}} \left| K_1\left(\frac{x_1 - u}{h_{1,T}}\right)\right| |f_1(u)| du ds dt \\ &\le \frac{M_4 \|f\|_\infty l_\Gamma}{T h_{1,T}} \int_{\mathbb{R}} |K_1(v)| dv = \mathcal{O}\left(\frac{1}{T h_{1,T}}\right). \end{aligned} \tag{26}$$

To treat the second term, we introduce the set $S_{a(T)} = \{(s,t) \in \mathbb{R}^2; |t-s| \le a(T)\}$, where $a(T) = h_T^{-1}$, we have

$$\begin{aligned} B &= \frac{1}{T^2} \int_{[0,T]^2\cap\Gamma^c\cap S_{a(T)}} \text{Cov}\Big(\frac{\widetilde{Y}_{\psi,T,t}}{f_1(X_{t,1}) h_{1,T}} K_1\Big(\frac{x_1 - X_t}{h_{1,T}}\Big), \\ &\quad \frac{\widetilde{Y}_{\psi,T,s}}{f_1(X_{s,1}) h_{1,T}} K_1\Big(\frac{x_1 - X_s}{h_{1,T}}\Big)\Big) dsdt \\ &+ \frac{1}{T^2} \int_{[0,T]^2\cap\Gamma^c\cap S^c_{a(T)}} \text{Cov}\Big(\frac{\widetilde{Y}_{\psi,T,t}}{f_1(X_{t,1}) h_{1,T}} K_1\Big(\frac{x_1 - X_t}{h_{1,T}}\Big), \\ &\quad \frac{\widetilde{Y}_{T,s}}{f_1(X_{s,1}) h_{1,T}} K_1\Big(\frac{x_1 - X_s}{h_{1,T}}\Big)\Big) dsdt \\ &:= E + F. \end{aligned} \tag{27}$$

Under the conditions $(C.1)$, $(F.1)$, $(K.1)-(K.2)$ and $(Q.1)$, there exists a constant M_5 such that, for T large enough,

$$\sup_{z_1\in\mathbb{R}} \int_{(y,\mathbf{z_{-1}})\in\mathbb{R}\times\mathbb{R}^{d-1}} \left| \frac{\psi(y)\mathcal{G}(\mathbf{z}_{-1})}{f(\mathbf{z})} \mathbb{1}_{S_1}\left(\frac{x_1 - z_1}{h_{1,T}}\right)\right| dy d\mathbf{z_{-1}} \le M_5.$$

Consider now the term E, we have

$$\begin{aligned} E &= \frac{1}{T^2 h_{1,T}^2} \int_{(s,t)\in[0,T]^2\cap\Gamma^c\cap S_{a(T)}} \int_{(u,\mathbf{v})\in\mathbb{R}\times\mathbb{R}^d} \int_{(y,\mathbf{z})\in\mathbb{R}\times\mathbb{R}^d} \frac{\psi(y)\mathcal{G}(\mathbf{z}_{-1})}{f(\mathbf{z})} \\ &\quad K_1\Big(\frac{x_1 - z_1}{h_{1,T}}\Big) \frac{\psi(u)\mathcal{G}(\mathbf{v}_{-1})}{f(\mathbf{v})} K_1\Big(\frac{x_1 - v_1}{h_{1,T}}\Big)\Big(f_{Y_t,\mathbf{X}_t,Y_s,\mathbf{X}_s}(u,\mathbf{v},y,\mathbf{z}) \\ &- f_{Y_t,bfX_t}(u,\mathbf{v}) f_{Y_s,bfX_s}(y,\mathbf{z})\Big) dy d\mathbf{z} du d\mathbf{v} ds dt \\ &\le \frac{2a(T) M_5^2 \|K_1\|_{\mathbb{L}_1}^2 A_f(\Gamma)}{T}. \end{aligned} \tag{28}$$

Noting that, under the conditions $(C.1)$, $(F.1)$, $(K.1)-(K.2)$ and $(Q.1)$, there exists a finite constant M_6 such that, for T large enough,

$$M_6 \geq \inf\left\{a \ : \ P\Big(\frac{\widetilde{Y}_{\psi,T,0}}{f_1(X_{0,1})}\Big|K_1\left(\frac{x_1-X_{0,1}}{h_{1,T}}\right)\Big| > a\Big) = 0\right\}.$$

Using the Billingsley's inequality, it follows that

$$\begin{aligned} F &\leq \frac{2}{T^2h_{1,T}^2}\int_{[0,T]^2\cap\Gamma^c\cap S^c_{a(T)}\cap\{u>v\}} 4M_6^2\alpha(u-v)dudv \\ &\leq \frac{8M_6^2}{Th_{1,T}^2}\int_{\{t>a(T)\}}\alpha(t)dt \\ &\leq \frac{8M_6^2}{Th_{1,T}^2}La(T)^{-1}. \end{aligned} \tag{29}$$

Finally, combining the hypothesis $(H.2)$ and the statements (25)and (29), we obtain (19).
Proof of (20). We have

$$\begin{aligned} &\mathbb{E}(\hat{\alpha}_1(x_1)) - \widetilde{m}_T(x_1) \\ &= \int_{\mathbb{R}}\frac{1}{h_{1,T}}\widetilde{m}_T(u_1)K_1\Big(\frac{x_1-u_1}{h_{1,T}}\Big)du_1 - \widetilde{m}_T(x_1) \\ &= \int_{\mathbb{R}}[\widetilde{m}_T(x_1-v_1h_{1,T}) - \widetilde{m}_T(x_1)]K_1(v_1)dv_1 \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}^{d-1}}[m_\psi(x_1-v_1h_{1,T},\mathbf{u}_{-1}) - m_\psi(x_1,\mathbf{u}_{-1})]\mathcal{G}(\mathbf{u}_{-1})d\mathbf{u}_{-1}K_1(v_1)dv_1 \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}^{d-1}}\left[\frac{(-h_{1,T}v_1)^k}{k!}\frac{\partial^k m_\psi}{\partial v_1^k}(v_1,\mathbf{u}_{-1})\right]\mathcal{G}(\mathbf{u}_{-1})d\mathbf{u}_{-1}K_1(v_1)dv_1 + o(h_{1,T}^k). \end{aligned}$$

Under the condition $(H.2)$, we obtain

$$\Big[\mathbb{E}\Big(\hat{\alpha}_1(x_1) - \widetilde{m}_T(x_1)\Big)\Big]^2 = \mathcal{O}\Big(T^{-2k/(2k+1)}\Big). \tag{30}$$

Thus, we have

$$\mathbb{E}\Big(\hat{\alpha}_1(x_1) - \widetilde{m}_T(x_1)\Big)^2 = \Big[\mathbb{E}\Big(\hat{\alpha}_1(x_1) - \widetilde{m}_T(x_1)\Big)\Big]^2 + \mathrm{Var}(\hat{\alpha}_1(x_1)). \tag{31}$$

Consequently, by combining the following inequality

$$E(\widehat{\widetilde{\eta}}_1(x_1) - \eta_1(x_1))^2 \leq 2\mathbb{E}\Big(\hat{\alpha}_1(x_1) - \widetilde{m}_T(x_1)\Big)^2 + 2E(\hat{C}_T - C_T - C)^2, \tag{32}$$

and the statements (30), (31), (19) and (32), the proof of (20) is readily achieved.

2.1 Proof of Theorem 1

Using the classical inequality $(a+b)^2 \leq 2(a^2+b^2)$, if follows that, for all $\mathbf{x} \in \mathcal{C}$,

$$\begin{aligned} E(\widehat{m}_{\psi,T,add}(\mathbf{x}) - m_\psi(\mathbf{x}))^2 &\leq 2E(\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x}) - m_\psi(\mathbf{x}))^2 + 2E(\widehat{m}_{\psi,T,add}(\mathbf{x}) - \widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x}))^2 \\ &:= I_1(\mathbf{x}) + I_2(\mathbf{x}). \end{aligned} \tag{33}$$

First, consider the term I_1, we have

$$\begin{aligned} I_1(\mathbf{x}) &= 2E(\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x}) - m_\psi(\mathbf{x}))^2 \\ &\leq 4d \sum_{\ell=1}^{d} E(\widehat{\widehat{\eta}}_\ell(x_\ell) - \eta_\ell(x_\ell))^2 + 4E\Big[\int_{\mathbb{R}^d} (\widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x}) - m_{\psi,T}(\mathbf{x}))q(\mathbf{x})d\mathbf{x}\Big]^2. \end{aligned} \tag{34}$$

Arguing as in proof of Lemma (1), we obtain

$$\begin{aligned} \widehat{C}_T - C_T - C &= \int_{\mathbb{R}^d} \widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x})q(\mathbf{x})d\mathbf{x} - \mu - \int_{\mathbb{R}^{d-1}} \sum_{j=2}^{d} m_j(u_j)\mathcal{G}(\mathbf{u}_{-1})d\mathbf{u}_{-1} - \int_{\mathbb{R}} m_1(x_1)q_1(x_1)dx_1, \\ &= \int_{\mathbb{R}^d} \widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x})q(\mathbf{x})d\mathbf{x} - \mu - \int_{\mathbb{R}^{d-1}} \sum_{j=1}^{d} m_j(u_j)q(\mathbf{u})d\mathbf{u} + \mathcal{O}\Big(h_{1,T}^k\Big), \\ &= \int_{\mathbb{R}^d} (\widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x}) - m_{\psi,T}(\mathbf{x}))q(\mathbf{x})d\mathbf{x} + \mathcal{O}\Big(T^{-k/(2k+1)}\Big). \end{aligned}$$

It follows that,

$$E\Big[\int_{\mathbb{R}^d} (\widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x}) - m_{\psi,T}(\mathbf{x}))q(\mathbf{x})d\mathbf{x}\Big]^2 \leq 2E\Big(\widehat{C}_T - C_T - C\Big)^2 + \mathcal{O}\Big(T^{-2k/(2k+1)}\Big). \tag{35}$$

By combining (34), (35), (18), and (20), we conclude that, for all $\mathbf{x} \in \mathcal{C}$

$$I_1(\mathbf{x}) = \mathcal{O}\Big(T^{-2k/(2k+1)}\Big). \tag{36}$$

Turning our attention to $I_2(\mathbf{x})$, it holds that,

$$\begin{aligned} &E(\widehat{m}_{\psi,T,add}(\mathbf{x}) - \widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x}))^2 \\ &= E\left[\sum_{\ell=1}^{d}(\widehat{\widehat{\eta}}_\ell(x_\ell) - \widehat{\eta}_\ell(x_\ell)) + \int_{\mathbb{R}^d} \widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x})q(\mathbf{x})d\mathbf{x} - \int_{\mathbb{R}^d} \widetilde{m}_{\psi,T}(\mathbf{x})q(\mathbf{x})d\mathbf{x}\right]^2 \\ &\leq 4d \sum_{\ell=1}^{d} E\Big(\int_{\mathbb{R}^{d-1}} (\widetilde{m}_{\psi,T,\ell}(\mathbf{x}) - \widetilde{\widetilde{m}}_{\psi,T,\ell}(\mathbf{x}))q(\mathbf{x}_{-\ell})d\mathbf{x}_{-\ell}\Big)^2 \end{aligned}$$

$$
\begin{aligned}
&+4d\sum_{\ell=1}^{d} E\Big(\int_{\mathbb{R}^d}(\widetilde{m}_{\psi,T,\ell}(\mathbf{x})-\widetilde{\widetilde{m}}_{\psi,T,\ell}(\mathbf{x}))q(\mathbf{x})d\mathbf{x}\Big)^2\\
&+2E\left(\int_{\mathbb{R}^d}\widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x})q(\mathbf{x})d\mathbf{x}-\int_{\mathbb{R}^d}\widetilde{m}_{\psi,T}(\mathbf{x})q(\mathbf{x})d\mathbf{x}\right)^2\\
\leq\;& 4d\sum_{\ell=1}^{d} E\int_{\mathbb{R}^{d-1}}(\widetilde{m}_{\psi,T,\ell}(\mathbf{x})-\widetilde{\widetilde{m}}_{\psi,T,\ell}(\mathbf{x}))^2q^2(\mathbf{x}_{-\ell})d\mathbf{x}_{-\ell}\\
&+4d\sum_{\ell=1}^{d} E\int_{\mathbb{R}^d}(\widetilde{m}_{\psi,T,\ell}(\mathbf{x})-\widetilde{\widetilde{m}}_{\psi,T,\ell}(\mathbf{x}))^2q^2(\mathbf{x})d\mathbf{x}\\
&2E\int_{\mathbb{R}^d}(\widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x})-\widetilde{m}_{\psi,T}(\mathbf{x}))^2q^2(\mathbf{x})d\mathbf{x}\\
\leq\;& 4d\sum_{\ell=1}^{d}\int_{\mathbb{R}^{d-1}} E(\widetilde{m}_{\psi,T,\ell}(\mathbf{x})-\widetilde{\widetilde{m}}_{\psi,T,\ell}(\mathbf{x}))^2q^2(\mathbf{x}_{-\ell})d\mathbf{x}_{-\ell}\\
&+4d\sum_{\ell=1}^{d}\int_{\mathbb{R}^d} E(\widetilde{m}_{\psi,T,\ell}(\mathbf{x})-\widetilde{\widetilde{m}}_{\psi,T,\ell}(\mathbf{x}))^2q^2(\mathbf{x})d\mathbf{x}\\
&+2\int_{\mathbb{R}^d} E(\widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x})-\widetilde{m}_{\psi,T}(\mathbf{x}))^2q^2(\mathbf{x})d\mathbf{x}
\end{aligned}
$$

Using the decomposition $1/f = 1/\widehat{f}_T + (\widehat{f}_T - f)/(\widehat{f}_T f)$, it is easily shown that for some positive constant $M_1 < \infty$, we have, under $(Q.2)$, for all $\mathbf{x} \in \mathcal{C}$ and T large enough,

$$
\begin{aligned}
&E\Big(\widetilde{m}_{\psi,T,\ell}(\mathbf{x})-\widetilde{\widetilde{m}}_{\psi,T,\ell}(\mathbf{x})\Big)^2\\
\leq\;& M_1 E\left(\frac{1}{h_{1,T}h_{2,T}^{d-1}}\Big|K_1\Big(\frac{x_\ell - X_{t,\ell}}{h_{1,T}}\Big)K_2\Big(\frac{\mathbf{x}_{-\ell}-\mathbf{X}_{t,-\ell}}{h_{2,T}}\Big)\Big| \times \sup_{\mathbf{x}\in\mathcal{C}}|\widehat{f}_T(\mathbf{x})-f(\mathbf{x})|\right)^2.
\end{aligned}
$$

It's easily seen that under our assumptions, following the demonstration of Theorem 4.9. in (2) p.112 and replacing $\log_m$ by 1, we have,

$$
\sup_{\mathbf{x}\in\mathcal{C}}|\widehat{f}_T(\mathbf{x})-f(\mathbf{x})| = \mathcal{O}\Big(\Big(\frac{\log T}{T}\Big)^{k'/(2k'+d)}\Big) \text{ almost surely },
$$

We conclude that, for all $\mathbf{x} \in \mathcal{C}$,

$$
E\Big(\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x})-\widehat{m}_{\psi,T,add}(\mathbf{x})\Big)^2 = \mathcal{O}\Big(\Big(\frac{\log T}{T}\Big)^{2k'/(2k'+d)}\Big) = \mathcal{O}\Big(T^{-2k/(2k+1)}\Big). \qquad \sqcap \tag{37}
$$

2.2 Proof of Theorem 2

In the next lemma we evaluate the difference between the estimator of the additive regression function $\widehat{\widehat{m}}_{\psi,T,add}$, for continuous time process, and the estimator $\widehat{\widehat{m}}_{\psi,n,add}$ where $n \in \mathbb{N}$.

Lemma 2 *For $n \in \mathbb{N}$ large enough, there exists a deterministic constant C such that for all ω in Ω and for all T in $[n, n+1[$,*

$$|\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x})(\omega) - E\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x})(\omega) - \widehat{\widehat{m}}_{\psi,n,add}(\mathbf{x})(\omega) + E\widehat{\widehat{m}}_{\psi,n,add}(\mathbf{x})(\omega)| < C\Big(\frac{\log(T)}{T}\Big)^{\frac{k}{2k+1}}.$$

Proof: It is sufficient to prove that $\forall\ \omega \in \Omega,\ \forall T \in [n, n+1[$,

$$|\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x}(\omega) - \widehat{\widehat{m}}_{\psi,n,add}(\mathbf{x})(\omega)| < C'\Big(\frac{\log(T)}{T}\Big)^{\frac{k}{2k+1}},$$

the other part being a trivial consequence of this inequality. Moreover, in view of (8) and (10), we can establish the following inequalities

$$\int_{\mathbb{R}^d} \widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x})(\omega)q(\mathbf{x})d\mathbf{x} - \int_{\mathbb{R}^d} \widetilde{\widetilde{m}}_{\psi,n}(\mathbf{x})(\omega)q(\mathbf{x})d\mathbf{x} < C_1\Big(\frac{\log(T)}{T}\Big)^{\frac{k}{2k+1}}, \tag{38}$$

$$\int_{\mathbb{R}^{d-1}} \widetilde{\widetilde{m}}_{\psi,T,l}(\mathbf{x})(\omega)q_{-l}(\mathbf{x}_{-l})d\mathbf{x}_{-l} - \int_{\mathbb{R}^{d-1}} \widetilde{\widetilde{m}}_{\psi,n,l}(\mathbf{x})(\omega)q_{-l}(\mathbf{x}_{-l})d\mathbf{x}_{-l} < C_l'\Big(\frac{\log(T)}{T}\Big)^{\frac{k}{2k+1}}, \tag{39}$$

$$\int_{\mathbb{R}^d} \widetilde{\widetilde{m}}_{\psi,T,l}(\mathbf{x})(\omega)q(\mathbf{x})d\mathbf{x} - \int_{\mathbb{R}^d} \widetilde{\widetilde{m}}_{\psi,n,l}(\mathbf{x})(\omega)q(\mathbf{x})d\mathbf{x} < C_l''\Big(\frac{\log(T)}{T}\Big)^{\frac{k}{2k+1}}, \tag{40}$$

with $l = 1, ...d$. We just establish the first inequality, the techniques being the same for (39) and (40). For fixed ω in Ω and $\mathbf{x}$ in $\mathbb{R}^d$, we have, for n large enough,

$$y \notin \mathcal{C}^\delta \Rightarrow K_3\Big(\frac{\mathbf{x}-\mathbf{y}}{h_t}\Big)q(\mathbf{x}) = 0, \quad \forall t \geq n.$$

So, by Fubini's Theorem

$$\begin{aligned}
\int_{\mathbb{R}^d} \widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x})(\omega)q(\mathbf{x})d\mathbf{x} &= \frac{1}{Th_{1,T}^d}\int_{[0,T]}\int_{\mathbb{R}^d} \frac{\psi(Y_t(\omega))}{f(\mathbf{X}_t(\omega))}K_3(\frac{\mathbf{x}-\mathbf{X}_t(\omega)}{h_{1,T}})q(\mathbf{x})d\mathbf{x}dt \\
&= \frac{1}{Th_{1,T}^d}\int_{[0,n]}\int_{\mathbb{R}^d} \frac{\psi(Y_t(\omega))}{f(\mathbf{X}_t(\omega))}K_3(\frac{\mathbf{x}-\mathbf{X}_t(\omega)}{h_{1,T}})q(\mathbf{x})d\mathbf{x}dt + \mathcal{O}(\frac{1}{T}) \\
&= \frac{1}{nh_{1,T}^d}\int_{[0,n]}\int_{\mathbb{R}^d} \frac{\psi(Y_t(\omega))}{f(\mathbf{X}_t(\omega))}K_3(\frac{\mathbf{x}-\mathbf{X}_t(\omega)}{h_{1,T}})q(\mathbf{x})d\mathbf{x}dt + \mathcal{O}(\frac{1}{T}) \\
&= \frac{1}{n}\int_{[0,n]}\int_{\mathbb{R}^d} \frac{\psi(Y_t(\omega))}{f(\mathbf{X}_t(\omega))}K_3(\mathbf{u})q(\mathbf{X}_t(\omega)+\mathbf{u}h_{1,T})d\mathbf{u}dt + \mathcal{O}(\frac{1}{T}).
\end{aligned}$$

Denoting $M_7 := \sup_{\mathbf{x}\in\mathcal{C}^\delta, y\in\mathbb{R}} \frac{\psi(y)}{f(\mathbf{x})} < \infty$, we have

$$\begin{aligned}
&\Big|\int_{\mathbb{R}^d} \widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x})(\omega)q(\mathbf{x})d\mathbf{x} - \int_{\mathbb{R}^d} \widetilde{\widetilde{m}}_{\psi,n}(\mathbf{x})(\omega)q(\mathbf{x})d\mathbf{x}\Big| \\
&\quad = \frac{1}{n}\Big|\int_{t\in[0,n]}\int_{u\in\mathbb{R}^d} \frac{\psi(Y_t(\omega))}{f(X_t(\omega)} K_3(\mathbf{u})(q(\mathbf{X}_t + h_{1,T}\mathbf{u}) - q(\mathbf{X}_t + h_{1,n}\mathbf{u}))dudt\Big| + \mathcal{O}(\frac{1}{T}) \\
&\quad \le \frac{M_7}{n}\Big|\int_{[0,n]}\int_{\mathbb{R}^d} K_3(\mathbf{u})(q(\mathbf{X}_t(\omega) + h_{1,T}\mathbf{u}) - q(\mathbf{X}_t(\omega) + h_{1,n}\mathbf{u}))d\mathbf{u}dt\Big| + \mathcal{O}(\frac{1}{T}) \\
&\quad \le \frac{dM_7}{n}\int_{[0,n]}\int_{S_3} |K_3(\mathbf{u}) \max_{1\le l\le d} \|\frac{\partial q}{\partial u_l}\|_\infty \|\mathbf{u}\|(h_{1,T} - h_{1,n})|d\mathbf{u}dt + \mathcal{O}(\frac{1}{T}) \\
&\quad = \mathcal{O}(h_{1,T} - h_{1,n}) + \mathcal{O}(\frac{1}{T}) \qquad (41)
\end{aligned}$$

Which implies (38) by (K.1). This achieves the proof of Lemma 2.

Set $\epsilon^2(T) = C\Big(\frac{\log T}{T}\Big)^{\frac{2k}{2k+1}}$, where C is a finite constant. There exists a finite number $r(T) := (3/M'h_{1,T}^2\varepsilon(T))^d$ of balls B_p of center $\mathbf{x}_p$ and radius $h_{1,T}^2\varepsilon(T)$, such that $\mathcal{C} \subset \cup_{p=1}^{r(T)} B_p$, where M' is a constant. For each $\mathbf{x} \in B_p$ we denote $t(\mathbf{x}) = \mathbf{x}_p$.Write now,

$$\begin{aligned}
&\sup_{\mathbf{x}\in\mathcal{C}} |\widehat{m}_{\psi,T,add}(\mathbf{x}) - m_\psi(\mathbf{x})| \\
&\qquad \le \sup_{\mathbf{x}\in\mathcal{C}} |\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x}) - \widehat{m}_{\psi,T,add}(\mathbf{x})| + \sup_{\mathbf{x}\in\mathcal{C}} |E\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x}) - m_\psi(\mathbf{x})| \\
&\qquad\quad + \sup_{\mathbf{x}\in\mathcal{C}} |\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x}) - \widehat{\widehat{m}}_{\psi,T,add}(t(\mathbf{x}))| + \sup_{\mathbf{x}\in\mathcal{C}} |E\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x}) - E\widehat{\widehat{m}}_{\psi,T,add}(t(\mathbf{x}))| \\
&\qquad\quad + \sup_{\mathbf{x}\in\mathcal{C}} |\widehat{\widehat{m}}_{\psi,add}(t(\mathbf{x})) - E\widehat{\widehat{m}}_{\psi,add}(t(\mathbf{x}))|.
\end{aligned}$$

Thus, to prove the Theorem 2, it suffices to establish the following Lemma.

Lemma 3 *Under the same hypothesis as Theorem 2, we have*

$$\sup_{\mathbf{x}\in\mathcal{C}} |E\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x}) - m_\psi(\mathbf{x})| = \mathcal{O}(h_{1,T}^k), \qquad (42)$$

$$\sup_{\mathbf{x}\in\mathcal{C}} |\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x}) - \widehat{\widehat{m}}_{\psi,T,add}(t(\mathbf{x}))| = \mathcal{O}\Big(\varepsilon(T)\Big) \qquad (43)$$

$$\sup_{\mathbf{x}\in\mathcal{C}} |E\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x}) - E\widehat{\widehat{m}}_{\psi,T,add}(t(\mathbf{x}))|\mathcal{O}\Big(\varepsilon(T)\Big), \qquad (44)$$

$$\sup_{\mathbf{x}\in\mathcal{C}} |\widehat{\widehat{m}}_{\psi,T,add}(t(\mathbf{x})) - E\widehat{\widehat{m}}_{\psi,T,add}(t(\mathbf{x}))| = \mathcal{O}\bigg(\Big(\frac{\log T}{T}\Big)^{k/2k+1}\bigg) \quad a.s., \qquad (45)$$

$$\sup_{\mathbf{x}\in\mathcal{C}} |\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x}) - m_\psi(\mathbf{x})| = \mathcal{O}\bigg(\Big(\frac{\log T}{T}\Big)^{k/(2k+1)}\bigg) \quad a.s.. \qquad (46)$$

Proof of 42: We have

$$
\begin{aligned}
&E\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x}) - m_\psi(\mathbf{x}) \\
&\quad = \sum_{l=1}^{d}(E\widehat{\widehat{\eta}}_l(x_l) - \eta_l(x_l)) + E(\int_{\mathbb{R}^d}\widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x})q(\mathbf{x})d\mathbf{x}) - \int_{\mathbb{R}^d} m_\psi(\mathbf{x})q(\mathbf{x})d\mathbf{x}.
\end{aligned}
$$

By Fubini's theorem, we obtain

$$
\text{Bias}(\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x})) = \sum_{l=1}^{d}\text{Bias}(\widehat{\widehat{\eta}}_l(x_l)) + \int_{\mathbb{R}^d}\text{Bias}(\widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x}))q(\mathbf{x})d\mathbf{x}. \tag{47}
$$

We can write,

$$
\begin{aligned}
\text{Bias}(\widehat{\widehat{\eta}}_1(x_1)) &= E(\widehat{\widehat{\eta}}_1(x_1)) - \eta_1(x_1) \\
&= \{E(\widehat{\alpha}_1(x_1)) - \widetilde{m}_{\psi,T}(x_1)\} + E(\widehat{C}_T - C_T - C) \\
&:= (I) + (II). \qquad (48)
\end{aligned}
$$

First consider the term (I), we have

$$
\begin{aligned}
\widetilde{m}_{\psi,T}(x_1) &= \frac{1}{T}\int_0^T E\Big\{E(\psi(Y_t)|\mathbf{X}_t)\frac{\mathcal{G}(\mathbf{X}_t)}{f(X_{t,-1}|X_{t,1})}d\mathbf{x}_{-1}\Big| X_{t,1} = x_1\Big\}dt \\
&= \int_{\mathbb{R}^{d-1}} m_\psi(x_1,\mathbf{u}_{-1})\int_{\mathbb{R}^{d-1}}\frac{1}{h_{2,T}^{d-1}}K_2\left(\frac{\mathbf{x}_{-1}-\mathbf{u}_{-1}}{h_{2,T}}\right)q_{-1}(\mathbf{x}_{-1})d\mathbf{x}_{-1}\ d\mathbf{u}_{-1} \\
&= \int_{\mathbb{R}^{d-1}} m_\psi(x_1,\mathbf{u}_{-1})\mathcal{G}(\mathbf{u}_{-1})d\mathbf{u}_{-1}.
\end{aligned}
$$

It follows that, under the conditions $(C.2)$, $(K.2)$ and $(K.3)$

$$
\begin{aligned}
&E(\widehat{\alpha}_1(x_1)) - \widetilde{m}_{\psi,T}(x_1) = \int_{\mathbb{R}}\frac{1}{h_{1,T}}\widetilde{m}_{\psi,T}(u_1)K_1\Big(\frac{x_1-u_1}{h_{1,T}}\Big)du_1 - \widetilde{m}_{\psi,T}(x_1) \\
&= \int_{\mathbb{R}}[\widetilde{m}_{\psi,T}(x_1 - v_1h_{1,T}) - \widetilde{m}_{\psi,T}(x_1)]K_1(v_1)dv_1 \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}^{d-1}}\Big[\sum_{i=1}^{k-1}\frac{(-h_{1,T}v_1)^i}{i!}\frac{\partial^i m_\psi}{\partial x_1^i}(x_1,\mathbf{u}_{-1}) \\
&\quad + \frac{(-h_{1,T}v_1)^k}{k!}\frac{\partial^k m_\psi}{\partial x_1^k}(x_1 - \theta h_{1,T}v1,\mathbf{u}_{-1})\Big]\mathcal{G}(\mathbf{u}_{-1})d\mathbf{u}_{-1}K_1(v_1)dv_1 \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}^{d-1}}\left[\frac{(-h_{1,T}v_1)^k}{k!}\frac{\partial^k m_\psi}{\partial x_1^k}(x_1 - \theta h_{1,T}v_1,\mathbf{u}_{-1})\right]\mathcal{G}(\mathbf{u}_{-1})d\mathbf{u}_{-1}K_1(v_1)dv_1.
\end{aligned}
$$

Thus, we obtain

$$
\sup_{x_1}|E(\widehat{\alpha}_1(x_1)) - \widetilde{m}_{\psi,T}(x_1)| = \mathcal{O}\left(h_{1,T}^k\right). \tag{49}
$$

Next, turning our attention to (II), by (21) we have

$$E(\widehat{C}_T - C_T) - C \quad = \quad \mathcal{O}(h_{1,T}^k). \tag{50}$$

Combining (49) and (50), it follows that

$$\sup_{x_1} |E(\widehat{\widehat{\eta}}_1(x_1)) - \eta_1(x_1)| = \mathcal{O}\left(h_{1,T}^k\right). \tag{51}$$

On the other hand, we have, for all $0 < \theta < 1$,

$$\begin{aligned}
\text{Bias}(\widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x})) &:= E\widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x}) - m_\psi(\mathbf{x}) & (52)\\
&= \int_{\mathbb{R}^d} [m_\psi(\mathbf{x} + h_{1,T}\mathbf{v}) - m_\psi(\mathbf{x})] K_3(\mathbf{v}) d\mathbf{v}. \\
&= \int_{\mathbb{R}^d} \sum_{i_1+\ldots+i_d=k} \frac{h_{1,T}^k}{k!} \frac{\partial^{i_1+\ldots+i_d} m_\psi}{\partial x_1^{i_1} \ldots \partial x_d^{i_d}} (\mathbf{x} + h_1\theta\mathbf{v}) v_1^{i_1} \ldots v_d^{i_d} K_3(\mathbf{v}) d\mathbf{v} & (53)\\
&:= \mathcal{O}(h_{1,T}^k),
\end{aligned}$$

Combining the decomposition (47) and the statements (51) and (54), we deduce the result (42).

Proof of (43) Under the condition $(K.5)$, there exists a constant M such that

$$\frac{1}{T} \int_0^T |Z_t(\mathbf{x}) - Z_t(\mathbf{t}(\mathbf{x}))| dt \leq \sum_{l=1}^d \frac{M}{h_{1,T}^2} |x_l - t(\mathbf{x})_l|$$

Consequently, using the expression of $r(T)$, we obtain

$$\sup_{\mathbf{x} \in \mathcal{C}} |\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x}) - \widehat{\widehat{m}}_{\psi,T,add}(t(\mathbf{x}))| = \mathcal{O}(\epsilon(T)).$$

Proof of (44): Similarly as above, we may deduce (44).

Proof of (45): In view of Lemma 2, it is sufficient to prove discrete version of (45), that is

$$\sup_{\mathbf{x} \in \mathcal{C}} |\widehat{\widehat{m}}_{\psi,n,add}(t(\mathbf{x})) - E\widehat{\widehat{m}}_{\psi,n,add}(t(\mathbf{x}))| = \mathcal{O}\Big(\varepsilon(n)\Big) \quad a.s. \tag{54}$$

Set n in $\mathbb{N}$ and,introduce some notations. Set,

$$\widehat{\widehat{m}}_{\psi,n,add}(\mathbf{x}) - E\widehat{\widehat{m}}_{\psi,n,add}(\mathbf{x}) \quad =: \quad \frac{1}{n} \int_0^n \xi_t(\mathbf{x}) dt, \tag{55}$$

where

$$\xi_t = \xi_t(\mathbf{u}) := (Z_t(\mathbf{u}) - E(Z_t(\mathbf{u})))$$

and

$$\begin{aligned}
Z_t &= Z_t(\mathbf{u}) \\
&= \frac{\psi(Y_t)}{h_{1,n}h_{2,n}^{d-1}f(\mathbf{X}_t)}\sum_{l=1}^{d}\Bigg\{\int_{\mathbb{R}^{d-1}} K_1\left(\frac{u_l - X_{t,l}}{h_{1,n}}\right) K_2\left(\frac{\mathbf{x}_{-l} - \mathbf{X}_{t,-l}}{h_{2,n}}\right) q_{-l}(\mathbf{x}_{-l})d\mathbf{x}_{-l} \\
&\quad - \int_{\mathbb{R}^d} K_1\left(\frac{x_l - X_{t,l}}{h_{1,n}}\right) K_2\left(\frac{\mathbf{x}_{-l} - \mathbf{X}_{t,-l}}{h_{2,n}}\right) q(\mathbf{x})d\mathbf{x}\Bigg\} \\
&\quad + \frac{1}{h_{1,n}^d}\frac{\psi(Y_t)}{f(\mathbf{X}_t)}\int_{\mathbb{R}^{d-1}} K_3\left(\frac{\mathbf{x} - \mathbf{X}_t}{h_{1,n}}\right) q(\mathbf{x})d\mathbf{x}.
\end{aligned}$$

Finally, we use the notation

$$V_i^n(\mathbf{x}) := \frac{1}{n}\int_{(i-1)p}^{ip} \xi_t(\mathbf{x})dt \quad i = 1, ..., 2q' \text{ where } p := \frac{n}{2q'} := \epsilon(n)^{-1/2} \tag{56}$$

So we can write

$$\widehat{\widehat{m}}_{\psi,n,add}(\mathbf{x}) - E\widehat{\widehat{m}}_{\psi,n,add}(\mathbf{x}) = \sum_{i=1}^{2q'} V_i^n(\mathbf{x}). \tag{57}$$

We have to show that the following quantity is summable

$$\mathbf{P}(\sup_{\mathbf{x}\in\mathcal{C}} |\sum_{i=1}^{2q'} V_i(t(\mathbf{x}))| \geq \varepsilon(n)) \leq r(n) \sup_{p=1,\ldots,r(n)} \mathbf{P}(|\sum_{i=1}^{2q'} V_i(\mathbf{t}_p)| \geq \varepsilon(n)). \tag{58}$$

Let j be fixed in $[1, r(n)]$. We have

$$\mathbf{P}(|\sum_{i=1}^{2q'} V_i(\mathbf{t}_j)| \geq \varepsilon(n)) \leq \mathbf{P}(|\sum_{i=1}^{q'} V_{2i}(\mathbf{t}_j)| \geq \varepsilon(n)/2) + \mathbf{P}(|\sum_{i=1}^{q'} V_{2i-1}(\mathbf{t}_j)| \geq \varepsilon(n)/2).$$

Observing that for a given M'', $\xi_t(\mathbf{x})(\omega) < \frac{M''}{h_{1,n}}$, $\forall \omega \in \Omega$, we can use recursively Bradley's lemma and define the independent random variables $W_2(\mathbf{t}_j), ..., W_{2q'}(\mathbf{t}_j)$ such that, $\forall i \in [1, q']$, W_{2i} and V_{2i} have the same law and $\forall \nu > 0$

$$P\Big(|W_{2i}(\mathbf{t}_j) - V_{2i}(\mathbf{t}_j)| > \nu\Big) \leq 11\Big(\frac{\|V_{2i}(\mathbf{t}_j)\|_\infty}{\nu}\Big)^{\frac{1}{2}}\alpha(p) \leq 11\Big(\frac{pM''}{h_{1,n}\nu}\Big)^{\frac{1}{2}}\alpha(p). \tag{59}$$

We have, for all $0 < \lambda < \frac{\varepsilon(n)}{2}$

$$\begin{aligned}
\Big\{|\sum_{i=1}^{q'} V_{2i}(\mathbf{t}_j)| > \frac{\epsilon(n)}{2}\Big\} \subset &\Big\{\{|\sum_{i=1}^{q'} V_{2i}(\mathbf{t}_j)| > \frac{\epsilon(n)}{2}; |V_i(\mathbf{t}_j) - W_i(\mathbf{t}_j)| \leq \frac{\lambda}{q'} \quad 1 \leq i \leq q'\}\Big\} \\
&\bigcup \bigcup_{j=1}^{q'} \{|V_i(\mathbf{t}_j) - W_i(\mathbf{t}_j)| > \frac{\lambda}{q'}\}\Big\}.
\end{aligned}$$

The choice $\lambda = \dfrac{\epsilon(n)}{4}$ gives us

$$
\begin{aligned}
P\Big(|\sum_{i=1}^{q'} V_{2i}(\mathbf{t}_j)| > \frac{\epsilon(n)}{2}\Big) &\leq P\Big(|\sum_{i=1}^{q'} W_{2i}(\mathbf{t}_j)| > \frac{\epsilon(n)}{4}\Big) \\
&\quad + \sum_{i=1}^{q'} P\Big(|V_{2i}(\mathbf{t}_j) - W_{2i}(\mathbf{t}_j)| > \frac{\epsilon(n)}{4q'}\Big). \qquad (60)
\end{aligned}
$$

We treat separately the two terms of the last inequality. For the second one, the application of (59)under the condition $(A.1)$ drives us to

$$
\sum_{i=1}^{q'} P\Big(|V_{2i}(\mathbf{t}_j) - W_{2i}(\mathbf{t}_j)| > \frac{\epsilon(n)}{4q'}\Big) \leq 11q'\Big(\frac{4q'pM'}{h_{1,n}\epsilon(n)}\Big)^{1/2}\alpha(p) = \mathcal{O}(r(n)^{-1}n^{\mu}) \qquad (61)
$$

$$
where\ \mu < -1.
$$

In order to dominate $P\Big(|\sum_{i=1}^{q'} W_{2i}(\mathbf{t}_j)| > \frac{\epsilon(n)}{4}\Big)$, we must bound the variance of W_{2i} (which has the same law as V_{2i}) to use Bernstein's inequality

$$
\begin{aligned}
\mathrm{Var}(W_{2i}(\mathbf{t}_j)) &= E(V_{2i}(\mathbf{t}_j)^2) \\
&\leq \frac{1}{n^2}\int_{[(2j-1)p,2jp]} E(\xi_t^2)dt \qquad (62)
\end{aligned}
$$

The kernels are bounded, so we can easily see, after a change of variables, that there exists a constant M''' such that

$$
E(Z_t^2) \leq \frac{M'''}{h_{1,n}},
$$

witch implies

$$
\mathbb{E}(\xi_t^2) \leq \frac{M'''}{h_{1,n}} \text{ and } \mathrm{Var}(W_{2i}(\mathbf{t}_j)) \leq \frac{pM'''}{n^2 h_{1,n}}.
$$

Observe that, for a given S in $\mathbb{R}^{*+}$, $\xi_t(\omega) < \frac{S}{h_{1,n}}$, $\forall \omega \in \Omega$, we readily have

$$
E|W_i|^k \leq (p\frac{M'}{nh_{1,n}})^{k-2} p! E|W_i|^2, \forall i.
$$

This allows us to apply Bernstein's inequality

$$
\begin{aligned}
P\Big(|\sum_{i=1}^{q'} W_i(\mathbf{t}_j)| > \frac{\epsilon(n)}{4}\Big) &\leq 2\exp\Big(-\frac{\epsilon(n)^2}{16(\frac{4q'pM}{n^2h_{1,n}} + \frac{M'p\epsilon(n)}{2nh_{1,n}})}\Big) \\
&= 2\exp\Big(-\frac{\epsilon(n)^2 nh_{1,n}}{32M + 8M'p\epsilon(n))}\Big). \qquad (63)
\end{aligned}
$$

The expression of p and $\varepsilon(n)$ gives us $p\varepsilon(n) \to 0$ and the sequence $\sum_{n=1}^{N} r(n) P\Big(|\sum_{i=1}^{q'} W_i(\mathbf{t}_j)| > \frac{\epsilon(n)}{4}\Big)$ converges as N grows to infinity if we choose a large enough C in $\epsilon(n)$. In view of this last inequality and (61), we obtain (54) by Borel-Cantelli.

Proof of (46): By (4), we have

$$\sup_{\mathbf{x}\in\mathcal{C}} |\widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x}) - \widetilde{m}_{\psi,T}(\mathbf{x})| \leq M \frac{\sup_{\mathbf{x}\in\mathcal{C}} |f(\mathbf{x}) - \widehat{f}_T(\mathbf{x})|}{\inf_{\mathbf{x}\in\mathcal{C}} f^2(\mathbf{x}) + o(1)} \frac{1}{T h_{1,T}^d} \int_0^T \Big| K_3\Big(\frac{\mathbf{x} - \mathbf{X}_t}{h_{1,T}}\Big)\Big| dt.$$

Using the statements (5) and (7), and the Theorem on a density estimator due to (2), we obtain

$$\begin{aligned} &\sup_{\mathbf{x}\in\mathcal{C}} |\widehat{\widehat{m}}_{\psi,T,add}(\mathbf{x}) - \widehat{m}_{\psi,T,add}(\mathbf{x})| \\ &\leq 2d \max_{1\leq l\leq d} \sup_{\mathbf{x}\in\mathcal{C}} |\widetilde{\widetilde{m}}_{\psi,T,l}(\mathbf{x}) - \widetilde{m}_{\psi,T,l}(\mathbf{x})| + \sup_{\mathbf{x}\in\mathcal{C}} |\widetilde{\widetilde{m}}_{\psi,T}(\mathbf{x}) - \widetilde{m}_{\psi,T}(\mathbf{x})| \\ &= \mathcal{O}\left(\Big(\frac{\log T}{T}\Big)^{k/(2k+1)}\right) \quad \text{a.s..} \end{aligned}$$

References

[1] Banon, G. (1978). Errata: "Nonparametric identification for diffusion processes" (SIAM J. Control Optim. **16** (1978), no. 3, 380–395)

[2] Bosq, D. (1993). Vitesses optimales et superoptimales des estimateurs fonctionnels pour les processus à temps continu. *C. R. Acad. Sci. Paris Sér. I Math.*, **317**(10), 1075–1078.

[3] D. Bosq. *Nonparametric statistics for stochastic processes*, volume 110 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1996. Estimation and prediction.

[4] Camlong-Viot, C., Sarda, P., and Vieu, P. (2000). Additive time series: the kernel integration method. *MSpringer-Verlag*, New York.

[5] Jones, M., Davies, S., and Park, B. (1994). Versions of kernel-type regression estimators. *J. Am. Statist. Assoc.*, **89**, 825–832.

[6] Linton, O. and Nielsen, J. (1995). A kernel method of estimating structured nonparametric regression based on marginal integration . *Biometrika*, **82**, 93–100.

[7] Nadaraya, E. and Watson (1964). On estimating regression. *Theory of proba. and Appl.*, **9**, 141–142.

[8] Newey, W. K. (1994). Kernel estimation of partial means and a general variance estimator. *Econometric Theory*, **10**(2), 233–253.

[9] Parzen, E. (1962). On estimation of probability density function and mode. *Ann.Math.Stat.*, **33**, 1065–1076.

[10] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.*, **27**, 832–837.

[11] Stone, C. J. (1985). Additive regression and other nonparametric models. *Ann. Statist.*, **13**(2), 689–705.