Mathematics and Statistics 10(5): 956-970, 2022
DOI: 10.13189/ms.2022.100507 http://www.hrpub.org

Asymptotically Minimax Goodness-of-fit Testing for Single-index Models

Jean-Philippe Tchiekre[1], **Christophe Pouet**[2], **Armel Fabrice E. Yodé**[1,*]

[1]Department of Mathematics and Computer Science, Félix Houphouët-Boigny University, 22 BP 582 Abidjan 22, Ivory Coast

[2] Marseille Institute of Mathematics, Marseille Central School, 38 Rue Fr´ed´eric Joliot-Curie, 13013 Marseille, France

Received March 31, 2022; Revised August 10, 2022; Accepted August 20, 2022

Cite This Paper in the following Citation Styles

***(a)**: [1] Jean-Philippe Tchiekre, Christophe Pouet, Armel Fabrice E. Yodé, "Asymptotically Minimax Goodness-of-fit Testing for Single-index Models," Mathematics and Statistics, Vol.10, No.5, pp. 956-970, 2022. DOI: 10.13189/ms.2022.100507*

***(b)**: Jean-Philippe Tchiekre, Christophe Pouet, Armel Fabrice E. Yodé, (2022). Asymptotically Minimax Goodness-of-fit Testing for Single-index Models. Mathematics and Statistics, 10(5), 956-970. DOI: 10.13189/ms.2022.100507*

Abstract In the context of non parametric multivariate regression model, we are interested in goodness-of-fit testing for the single-index models. These models are dimension reduction models and are therefore useful in multidimensional nonparametric statistics because of the well-known phenomenon called the curse of dimensionality. Fan and Li [5] have proposed the first consistent test for goodness-of-fit testing of the single-index by using nonparametric kernel estimation method and a central limit theorem for degenerate U-statistics of order higher than two. Since then, the minimax properties of this test have not been investigated. Following this work, we use here the asymptotic minimax approach. We are interested in finding the asymptotic minimax rate of testing $\varphi(\alpha_n)$ which gives the minimal distance between the null and alternative hypotheses such that a successful testing is possible. We propose a test procedure of level α_n which can tend to zero when the sample size tends to infinity. We have established the minimax asymptotic properties of our test procedure by showing that it reaches the asymptotic minimax rate $\varphi_n(\alpha_n)$ for the dimension $d = 3$ and there is no test of level α_n reaching this rate for $d \geq 3$. Because of its minimax asymptotic properties, our test is able to distinguish the null hypothesis of the closest possible alternative. The results obtained were possible thanks to a large deviation result that we established for a degenerate U-statistic of order two appearing in our decision variable.

Keywords Single-index Model, Asymptotically Minimax Hypothesis Testing, Asymptotics of Errors Probabilities, Large Deviation, U-statistics

1 Introduction

The main difficulty arising in multivariate nonparametric estimation is the so-called curse of dimensionality: the larger dimension and the worse quality of estimation. In particular, the dimension affects the optimal rate of estimation. According to Stone [19], under appropriate regularity conditions, the optimal rate of convergence for the nonparametric estimation of an unknown d-dimensional β-times differentiable regression function or density function is $n^{-\frac{\beta}{2\beta+d}}$, where β is an integer and n is the sample size. As we see, the higher d is, the lower the optimal rate is. A way to overcome the curse of dimensionality consists in imposing additional structural assumptions on the function to be estimated. In this paper, we are interested in the single-index model. In this case, from Gaïffa and Lecué [6], the minimax rate of convergence is given by $n^{-\frac{\beta}{2\beta+1}}$ i.e. the quality of estimation does not depend on the dimension parameter d. Therefore, it is important and relevant to develop appropriate goodness-of-fit tests for this model. Fan and Li [5] have proposed the first consistent test for the single-index model. This test is constructed by invoking the Central Limit Theorem for degenerate U-statistics of order higher than two. Other authors such as Aït-Sahalia, Bickel and Stoker [1], Härdle, Mammen and Proença [9], Xia, Li, Tong, and Zhang [26], Stute and Zhu [20], Chen and Van Keilegom [3], Lin and Kulakesara [14] and Maistre and Patilea [17] have also presented goodness-of-fit testing for the single-index model. The test statistics proposed by Maistre and Patilea [17] are quite similar to those of Fan and Li [5]. However, the minimax properties of the above tests are not investigated. In this paper, we construct a test for which we establish minimax asymptotic properties. Moreover, the error of the first kind of our test depends on the sample size n and it can decrease to zero when n tends to infinity. We show also how this property affects the error of the second kind. This condition has already been used in Yodé [22, 24] and Chiabrando [2]. Such kind of testing problems appear in the concept of random normalizing factor initiated by Lepski [16] and extended by Hoffmann and Lepski [11], Hoffman [10], Yodé [23], and Chiabrando [2].

Suppose that the observations (X_j, Y_j) $j = 1, \ldots, n$ are generated by the classical regression model

$$Y_j = g(X_j) + \varepsilon_j \tag{1.1}$$

where Y_j are scalar response variables, the covariates X_j are d-dimensional vectors, ε_j are independent errors with zero mean and known variance $\sigma^2 > 0$ and g is an unknown real valued function with compact support $S \subseteq [0,1]^d$. We allow for fixed covariate designs in this article.
We shall make the following technical asssumption :

($\mathcal{H}_1$) The function g is a function belonging to $\Sigma_d(\beta, L) \cap L^2([0,1]^d)$, $\beta = m + \tau$, $m \in \mathbb{N}$, $\tau \in (0,1)$, $L > 0$ where $\Sigma_d(\beta, L)$ is the d-dimensional isotropic Hölder space. We say that $g \in \Sigma_d(\beta, L)$ if

$$\sup_{(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_d)\in[0,1]^{d-1}} \left| \frac{\partial^m g}{\partial x_i^m}(x^{(1)}) - \frac{\partial^m g}{\partial x_i^m}(x^{(2)}) \right| \leq L|x_i^{(1)} - x_i^{(2)}|^\tau$$

where $x^{(l)} = (x_1, \ldots, x_{i-1}, x_i^{(l)}, x_{i+1}, \ldots, x_d)'$, $l = 1, 2$, for all $i \in \{1, \ldots, d\}$ and $x_i^{(1)}$, $x_i^{(2)} \in [0,1]$. Here and later, the prime denotes the transpose.

We say that the model (1.1) is a linear single-index model if

$$g(x) = f_0(\theta' x), \tag{1.2}$$

where f_0 is an unknown univariate smooth function and θ is an unknown index vector. Let the set

$$\Theta = \{\theta \in \mathbb{R}^d, \|\theta\| = 1, \text{ with first nonzero component positive}\}.$$

According to Kulasekera and Lin [15], if g is continuous and non-constant and $\theta \in \Theta$ then the representation (1.2) is unique. The estimation of f_0 and θ has been the subject of several papers: Powell, Stock and Stoker [18], Duan and Li [4], Ichimura [13], Härdle, Hall and Ichimura [8], Xia, Tong and Li [25], Hristache, Juditsky and Spokoiny [12].

This paper is organized as follows. In Section 2, we present the hypothesis testing problem and the basic tools that we need. In Section 3, we present our main results. Proofs of results are found in Section 4.

2 Hypothesis testing problem

2.1 Formulation of the problem

In this paper, we develop a test for null hypothesis

$$(H_0) : g \in \Delta_0$$

against alternative hypothesis

$$(H_1) : g \in \Delta_n(C\varphi_n)$$

where

$$\Delta_0 = \big\{g : \exists\theta_0 \in \Theta, \exists f_0 \in L^2([0,1]) \cap \Sigma_1(\beta, L) / g(x) = f_0(\theta_0' x),\ \forall x \in S\big\}$$

and for $n \geq 1$,

$$\Delta_n(C\varphi_n) = \big\{g : \inf_{\theta\in\Theta} \inf_{f\in\Sigma_1(\beta,L)\cap L^2([0,1])} \|g - f(\theta'\cdot)\|_2 \geq C\varphi_n\big\}, \tag{2.1}$$

$(\varphi_n)_{n\geq1}$ is a sequence of strictly positive real numbers with $\|h\|_2 = (\int_S h^2(x)dx)^{1/2}$. The null hypothesis (H_0) means that the model (1.1) is single-index.
Now, we determine the closest function to g in L^2-norm in the direction of θ. For all $\theta \in \Theta$, note that

$$f_\theta = \arg \inf_{f\in\Sigma_1(\beta,L)\cap L^2([0,1])} \int_S (g(x) - f(\theta' x))^2 \, dx.$$

Let A be an orthogonal matrix with first row θ'. Let us put $y = Ax$. This transformation yields

$$\int_S (g(x) - f(\theta' x))^2 dx = \int_{AS} (g(A'y) - f(y_1))^2 dy = \int_0^1 \left[\int_{S(y_1)} (g(A'y) - f(y_1))^2 dy_2 \ldots dy_d \right] dy_1$$

where $AS = \{Ax | x \in S\}$ and $S(t) = \{y \in AS | y_1 = t\}$. Hence, it suffices to minimize the inner integral for every y_1 which gives

$$f_\theta(y_1) = \frac{\int_{S(y_1)} g(A'y) dy_2 \ldots dy_d}{\int_{S(y_1)} dy_2 \ldots dy_d}.$$

Therefore, the alternative hypothesis (2.1) becomes

$$(H_1) : g \in \Delta_n(C\varphi_n) = \left\{ g : \inf_{\theta\in\Theta} \|g - f_\theta(\theta'\cdot)\|_2 \geq C\varphi_n \right\}. \tag{2.2}$$

2.2 Asymptotic minimax framework

Let ϕ_n be a test i.e. a measurable function depending only on observation $(X_1, Y_1), \ldots, (X_n, Y_n)$ with values in the two-point set $\{0, 1\}$. The value $\phi_n = 0$ means that (H_0) is accepted and $\phi_n = 1$ means that (H_0) is rejected. Introduce the first-type error

$$R_1(\phi_n) = \sup_{g\in\Delta_0} \mathbb{P}_g^n \{\phi_n = 1\},$$

and the second-type error

$$R_2(C, \varphi_n, \phi_n) = \sup_{g\in\Delta_n(C\varphi_n)} \mathbb{P}_g^n \{\phi_n = 0\},$$

where $\mathbb{P}_g^n$ is the probability measure of the $(Y_1, \ldots, Y_n)$. The properties of the test ϕ_n are characterized by both types of errors.

Definition 2.1. *Let α_n be a positive sequence in $(0,1)$. We call ϕ_n a test of asymptotical level α_n if*

$$\limsup_{n\to+\infty} \alpha_n^{-1} \sup_{g\in\Delta_0} \mathbb{P}_g^n\{\phi_n = 1\} \le 1. \tag{2.3}$$

Let $\Gamma_n(\alpha_n)$ be the set of asymptotical α_n-level tests.

Definition 2.2. *Let α_n and γ_n be two positive sequences in $(0,1)$. The positive sequence $\varphi_n(\alpha_n)$ is called minimax rate of testing if*

1. *there exists a constant C_* such that for any $C < C_*$, we have*

$$\liminf_{n\to+\infty} \gamma_n^{-1} \inf_{\phi_n\in\Gamma_n(\alpha_n)} \sup_{g\in\Delta_n(C\varphi_n(\alpha_n))} \mathbb{P}_g^n\{\phi_n = 0\} \ge 1; \tag{2.4}$$

2. *there exists a constant $C^* > 0$ and a test $\phi_n^* \in \Gamma_n(\alpha_n)$ such that for any $C > C^*$*

$$\limsup_{n\to+\infty} \gamma_n^{-1} \sup_{g\in\Delta_n(C\varphi_n(\alpha_n))} \mathbb{P}_g^n\{\phi_n^* = 0\} \le 1. \tag{2.5}$$

ϕ_n^* is called asymptotically optimal test.

According to Definition 2.2, if the alternative is too close to the null hypothesis set of functions then (2.4) ensures that no asymptotically α_n-level test procedure can asymptotically achieve a second kind error lower than γ_n. Though, (2.5) states that it is possible to construct a test that detects (H_0) against an alternative separated away from the null hypothesis by a distance asymptotically equal to $\varphi_n(\alpha_n)$.

3 Test procedure

In this section we give explaination of the proposed test procedure. According to (2.2), an estimate of $\inf_{\theta\in\Theta} \|g - f_\theta(\theta'\cdot)\|_2^2$ can detect whether the single-index hypothesis is accepted or not.

For all $i = (i_1,\ldots,i_d) \in \mathbb{N}^d$, $u = (u_1,\ldots,u_d) \in \mathbb{R}^d$, $h_n > 0$, let us introduce the following notations

$$|i| = \sum_{k=1}^d i_k \qquad i! = i_1!\cdots i_d! \qquad D^i g(x) = \frac{\partial^{|i|} g(x)}{\partial x_1^{i_1}\ldots\partial x_d^{i_d}}$$

$$u^i = u_1^{i_1}\ldots u_d^{i_d} \qquad \sum_{0\le|i|\le m} = \sum_{s=0}^m \Big(\sum_{i_1=0}^s \cdots \sum_{i_d=0}^s\Big).$$

For X_j in a neighborhood of x, by using Taylor's expansion, we obtain

$$g(X_j) \approx \sum_{0\le|i|\le m} \frac{1}{i!}(X_j - x)^i D^i g(x).$$

Let us put

$$\Upsilon_l(x) = \big(\frac{1}{i!}x^i : |i| = l\big)' \qquad \beta_l(x) = \big(h^{|i|} D^i g(x) : |i| = l\big)'$$

$$\Upsilon(x) = \big(1, \Upsilon_1'(x), \ldots, \Upsilon_m'(x)\big)$$
$$\beta(x) = \big(g(x), \beta_1'(x), \ldots, \beta_m'(x)\big).$$

Here and later, $\beta(x)$ and $\Upsilon(x)$ are $\mathcal{N}_m \times 1$ column vectors, where

$$\mathcal{N}_m = \sum_{l=0}^m \binom{l+d-1}{d-1} = \frac{(d+m)!}{d!m!}.$$

Let us fix a kernel function K, that is, a real function defined on $\mathbb{R}^d$ verifying $\int_{\mathbb{R}^d} K(u)du = 1$ and a sequence of real numbers $h_n > 0$. The local polynomial estimator, which estimates β, is the minimizer of the following weighted least squares problem:

$$\min_{\beta\in\mathbb{R}^{\mathcal{N}_m}} \sum_{j=1}^n \Big(Y_j - \beta'\Upsilon\Big(\frac{X_j - x}{h_n}\Big)\Big)^2 K\Big(\frac{X_j - x}{h_n}\Big).$$

Let us denote

$$a_{nx} = \frac{1}{nh_n^d}\sum_{j=1}^n \Upsilon\Big(\frac{X_j - x}{h_n}\Big) K\Big(\frac{X_j - x}{h_n}\Big) Y_j$$

$$\mathcal{B}_{nx} = \frac{1}{nh_n^d}\sum_{j=1}^n \Upsilon\Big(\frac{X_j - x}{h_n}\Big)\Upsilon'\Big(\frac{X_j - x}{h_n}\Big) K\Big(\frac{X_j - x}{h_n}\Big).$$

We suppose that $\mathcal{B}_{nx}$ is a symmetric positive definite matrix. Then,

$$\hat{g}_n(x) = \sum_{j=1}^n W_{nj}(x) Y_j$$

where

$$W_{nj}(x) = \frac{1}{nh_n^d}\Upsilon'(0)\mathcal{B}_{nx}^{-1}\Upsilon\Big(\frac{X_j - x}{h_n}\Big) K\Big(\frac{X_j - x}{h_n}\Big) \tag{3.1}$$

is locally polynomial estimator of order $\mathcal{N}_m$ of g.

Now, we consider the following technical assumptions:

$(\mathcal{H}_2)$ *There exists $\lambda_0 > 0$ and an integer n_0 such that the smallest eigenvalue $\lambda_{min}(\mathcal{B}_{nx})$ of $\mathcal{B}_{nx}$ satisfies*

$$\lambda_{min}(\mathcal{B}_{nx}) \ge \lambda_0$$

for all $n \ge n_0$ and for $x \in [0,1]^d$.

$(\mathcal{H}_3)$ *There exists $a_0 > 0$ such that, for all interval $A \subseteq [0,1]^d$ and for all $n \ge 1$*

$$\frac{1}{n}\sum_{j=1}^n 1_{\{X_j\in A\}} \le a_0 \max(Leb(A), 1/n),$$

where $Leb(A)$ is the Lebesgue measure of A.

$(\mathcal{H}_4)$ *The kernel K is compactly supported on $[-1,1]^d$ and there exists $K_{max} < +\infty$ such that $|K(u)| \le K_{max}$, $\forall u \in \mathbb{R}^d$.*

The following proposition is a multidimensional version of Lemma 1.3 of Tsybakov[21].The proof is omitted.

Proposition 3.1. *Under Assumptions $(\mathcal{H}_2)$, $(\mathcal{H}_3)$ and $(\mathcal{H}_4)$, for all $n \ge n_0$, $h_n \ge \frac{1}{2n^{1/d}}$ and $x \in [0,1]^d$, the weights $W_{nj}(x)$ are such that*

1. $\sup_{j,x}|W_{nj}(x)| \leq \frac{c^*}{nh_n^d}$;
2. $\sum_{j=1}^{n}|W_{nj}(x)| \leq c^*$;
3. $W_{nj}(x) = 0$ *if* $\|X_j - x\| > h_n$,

where $\|\cdot\|$ *is Euclidian norm and the constant* c^* *depends only on* λ_0, a_0, d *and* K_{max}.

For fixed $\theta \in \Theta$, the function $f_\theta(\theta'\cdot)$ can be estimated by

$$\hat{f}_\theta(\theta' x) = \sum_{j=1}^{n} \widetilde{W}_{nj}(\theta' x) Y_j \tag{3.2}$$

where

$$\widetilde{W}_{nj}(\theta' x) = \frac{1}{nb_n}\widetilde{\Upsilon}'(0)\widetilde{\mathcal{B}}_{nx}^{-1}\widetilde{\Upsilon}\left(\frac{\theta' X_j - \theta' x}{b_n}\right)\widetilde{K}\left(\frac{\theta' X_j - \theta' x}{b_n}\right),$$

with $\widetilde{K}$ is an univariate kernel function, $b_n > 0$, $\widetilde{\Upsilon}(\widetilde{u}) = \left(1, \widetilde{u}, \widetilde{u}^2/2!, \cdots, \widetilde{u}^m/m!\right)'$ and

$$\widetilde{\mathcal{B}}_{nx} = \frac{1}{nb_n}\sum_{j=1}^{n}\widetilde{\Upsilon}\left(\frac{\theta' X_j - \theta' x}{b_n}\right)\widetilde{\Upsilon}'\left(\frac{\theta' X_j - \theta' x}{b_n}\right)\widetilde{K}\left(\frac{\theta' X_j - \theta' x}{b_n}\right).$$

The assumptions on K are used for $\widetilde{K}$ in the one-dimensional case. Moreover, Proposition 3.1 is valid for $\widetilde{W}_{nj}$ in the univariate case and $b_n = n^{-\frac{1}{2\beta+1}}$. Therefore, $\inf_{\theta\in\Theta}\|g - f_\theta(\theta'\cdot)\|_2^2$ is estimated by $\inf_{\theta\in\Theta} T_n(\theta)$ with

$$T_n(\theta) = \left\|\hat{g}_n - \hat{f}_\theta(\theta'\cdot)\right\|_2^2 - \sigma^2\sum_{j=1}^{n}\int_S W_{nj}^2(x)dx. \tag{3.3}$$

Here $T_n(\theta)$ is an unbiased estimator of $\|g - f_\theta(\theta'\cdot)\|_2^2$.

4 Main results

Let the following assumption:

($\mathcal{H}_5$) *Assume that there exists* $a, \nu > 0$ *positive constants and a scale sequence* $(\omega_n)_{n\in\mathbb{N}^*}$ *such that :*

(a) $\forall\, m \in \mathbb{N}^*, \quad \mathbb{E}(|\varepsilon|^m) \leq a^m m^{\nu m}$.

(b) $\exists\, c < 1/2, \quad \omega_n = O(n^c)$.

(c) $\forall\, m, n \in \mathbb{N}^*$, $\mathbb{E}\Big(\max_{i=1,\dots,n}|\varepsilon_i|^m\Big) \leq \omega_n^m\ m^{\nu m}$.

This assumption is not restrictive and is satisfied for standard noise distributions. For example, if ε_i has the gaussian distribution then Assumption ($\mathcal{H}_5$) is satisfied with $\omega_n = \sqrt{\ln(n)}$ and $\nu = \frac{1}{2}$.
Our test statistic is defined as follows

$$\phi_{n,\alpha_n} = 1_{\left\{\inf_{\theta\in\Theta} T_n(\theta) \geq (\lambda\varphi_n(\alpha_n))^2\right\}}, \tag{4.1}$$

where $T_n(\theta)$ is (3.3), $\varphi_n(\alpha_n) = \left(n^{-1}\sqrt{\log\left(\frac{e^4}{\alpha_n}\right)}\right)^{\frac{2\beta}{4\beta+d}}$ and λ is a positive constant such that $\lambda > \left(4e^2K_1^2\Gamma_1\right)^{1/4}$ with

$$\Gamma_1 = \frac{\left(d + 1 + \binom{d}{2}\right)^2 9^m((d+m)!)^2 K_{max}^4 a^4 2^{4\nu}}{(m!)^4(d!)^2\lambda_0^4}.$$

K_1 is universal constant.

Theorem 4.1. *Let a positive sequence* $\alpha_n \in (0,1)$. *Assume that* $d > 2$, $\beta > \frac{d}{4}$, *Assumptions* ($\mathcal{H}_1$), ($\mathcal{H}_2$), ($\mathcal{H}_3$), ($\mathcal{H}_4$) *and* ($\mathcal{H}_5$) *are satisfied. Then, we have*

$$\limsup_{n\to+\infty}\alpha_n^{-1}\sup_{g\in\Delta_0}\mathbb{P}_g^n\Big(\phi_{n,\alpha_n} = 1\Big) \leq 1.$$

In practice the variance σ^2 is often unknown. In this case, we can replace it in (3.3) with a consistent estimator $\widehat{\sigma}^2$. For example, if

$$\widehat{\sigma}^2 = \frac{1}{2(n-2)}\sum_{i=2}^{n}(Y_i - Y_{i-1})^2$$

and $0 < \beta < 1$ then $\widehat{\sigma}^2$ converges in probability to σ^2. Moreover, if $\frac{3}{4} \leq \beta < 1$ then the test ϕ_{n,α_n} is of asymptotic level α_n.

Remark 4.1. - *The condition* $\beta > \frac{d}{4}$ *implies that* $nh_n^d \to +\infty$, *where* $h_n = \left(n^{-1}\sqrt{\log\left(\frac{e^4}{\alpha_n}\right)}\right)^{\frac{2}{4\beta+d}}$. *This condition is essential for the convergence of the locally polynomial estimator.*

- *The condition* $d > 2$ *ensures that the one-dimensional optimal rate of estimation* $n^{\frac{-2\beta}{2\beta+1}}$ *is negligible compared to* $\frac{1}{nh_n^{d/2}}$ *which is the order of the main term of the decision variable.*

We make the following assumption:

($\mathcal{H}_6$) $\mathbb{P}(|\varepsilon_i| \geq y) = O(y^{-(3+\delta)})$, $\quad\forall i \in \{1, ..., n\}$ and for $\delta > 0$.

This hypothesis is also used for the control of $\sup_i|\varepsilon_i|$ as it was the case in Gayraud and Pouet [7].

Theorem 4.2. *Assume that* $\frac{3}{4} \leq \beta < 1$ *and that Assumptions* ($\mathcal{H}_1$), ($\mathcal{H}_2$), ($\mathcal{H}_3$), ($\mathcal{H}_4$), ($\mathcal{H}_5$), *and* ($\mathcal{H}_6$), *are satisfied. Then, for all* $C > C^* = \frac{\sqrt{2}}{2}\lambda$, *we have*

$$\limsup_{n\to+\infty}\sup_{g\in\Delta_n(C\varphi_n(\alpha_n))}\mathbb{P}_g^n(\phi_{n,\alpha_n} = 0) = 0.$$

Let us introduce the following assumption:

($\mathcal{H}_7$) The random variables $\varepsilon_1, \dots, \varepsilon_n$ are i.i.d. centered Gaussian with known variance $\sigma^2 > 0$.

We note that Assumption ($\mathcal{H}_7$) implies ($\mathcal{H}_6$). Let us put $C_* = \left(a_*^{-\frac{1}{4\beta+d-4}}\|\psi\|_\infty^{-\frac{4d}{4\beta+d-4}}\right)^\beta$ with $a_* > 0$, such that for any $x \in \mathbb{R}$, $\cosh(x) \leq \exp(a_*\,x^2)$.

Theorem 4.3. *Assume that $\beta > \frac{d}{4}$ and that Assumptions $(\mathcal{H}_1)$, $(\mathcal{H}_2)$, $(\mathcal{H}_3)$, $(\mathcal{H}_4)$, $(\mathcal{H}_5)$, and $(\mathcal{H}_7)$ are satisfied. Then, for any $C < C_*$ we have*

$$\liminf_{n\to+\infty} \inf_{\phi_n\in\Gamma_n(\alpha_n)} \sup_{g\in\Delta_n(C\varphi_n(\alpha_n))} \mathbb{P}_g^n(\phi_n = 0) = 1.$$

Remark 4.2.

- *Theorems 4.1, 4.2, and 4.3 show that the sequence $\varphi_n(\alpha_n)$ is the asymptotically optimal rate of the test and the procedure ϕ_{n,α_n} is asymptotically optimal for $d = 3$ and $3/4 \le \beta < 1$.*
- *Since the Theorem 4.2 could not be established for $\beta \ge 1$ and $d \ge 4$, the optimality of the test procedure (4.1) in the family of α_n level tests remains an open problem. However, no α_n level test can reach the $\varphi_n(\alpha_n)$ rate.*

5 Preliminaries results

Let X_n be a sequence of random variables and $\psi_n > 0$ a sequence of constants. We write

- $X_n = O_p(\psi_n)$ if, for each $\varepsilon > 0$, there is an M (depending on ε but not n) such that $\mathbb{P}(|X_n| > M\psi_n) < \varepsilon$;
- $X_n = o_p(\psi_n)$ if $\dfrac{X_n}{\psi_n}$ converges in probability to zero.

5.1 Behavior of the test statistic under null hypothesis

Let us fix θ_0 in Θ. Therefore, we have

$$T_n(\theta_0) = I_1 + I_2(\theta_0) + I_3(\theta_0)$$

where

$$I_1 = \|\hat{g}_n - g\|_2^2 - \sigma^2 \sum_{j=1}^n \int_S W_{nj}^2(x)dx,$$
$$I_2(\theta_0) = \left\|g - \hat{f}_{\theta_0}(\theta_0'\cdot)\right\|_2^2,$$
$$I_3(\theta_0) = 2\int_S (\hat{g}_n(x) - g(x))\left(g(x) - \hat{f}_{\theta_0}(\theta_0' x)\right)dx.$$

The following decomposition holds

$$I_1 = \left\|\hat{g}_n - \mathbb{E}_g^n(\hat{g}_n)\right\|_2^2 - \sigma^2\sum_{j=1}^n \int_S W_{nj}^2(x)dx + \left\|\mathbb{E}_g^n(\hat{g}_n) - g\right\|_2^2 + 2\int_S \left(\hat{g}_n(x) - \mathbb{E}_g^n(\hat{g}_n(x))\right)\left(\mathbb{E}_g^n(\hat{g}_n(x)) - g(x)\right)dx.$$

Let us put

$$U_n = \sum_{j\neq k} \varepsilon_j\varepsilon_k \int_S W_{nj}(x)W_{nk}(x)dx. \tag{5.1}$$

Proposition 5.1. *Asume that Assumptions $(\mathcal{H}_1)$, $(\mathcal{H}_2)$, $(\mathcal{H}_3)$ and $(\mathcal{H}_4)$ are satisfied. Then, we have*

1. $U_n = O_p\left(\dfrac{1}{nh_n^{d/2}}\right)$
2. $$\left\|\hat{g}_n - \mathbb{E}_g^n(\hat{g}_n)\right\|_2^2 - \sigma^2\sum_{j=1}^n\int_S W_{nj}^2(x)dx = U_n + o_p\left(\frac{1}{nh_n^{d/2}}\right).$$
3. $\left\|\mathbb{E}_g^n(\hat{g}_n) - g\right\|_2^2 \le \left(\dfrac{Ld^{2/(2-\tau)}c^*}{m!}\right)^2 h_n^{2\beta}.$
4. $$\int_S \left(\hat{g}_n(x) - \mathbb{E}_g^n(\hat{g}_n(x))\right)\left(\mathbb{E}_g^n(\hat{g}_n(x)) - g(x)\right)dx = o_p\left(\frac{1}{nh_n^{d/2}}\right).$$

The following proposition shows that under (H_0), $I_1(\theta_0)$ and $I_2(\theta_0)$ are negligible in probability with respect to $\frac{1}{nh_n^{d/2}}$.

Proposition 5.2. *Under the null hypothesis (H_0), we have*

$$I_2(\theta_0) = o_p\left(\frac{1}{nh_n^{d/2}}\right) \tag{5.2}$$
$$I_3(\theta_0) = o_p\left(\frac{1}{nh_n^{d/2}}\right) \tag{5.3}$$

5.2 Behavior of the test under alternative hypothesis

We have the following decomposition

$$\inf_{\theta\in\Theta} T_n(\theta) = I_1 + \inf_{\theta\in\Theta}\{I_2(\theta) + I_3(\theta)\}$$

where

$$I_1 = \|\hat{g}_n - g\|_2^2 - \sigma^2\int_S W_{nj}^2(x)dx$$
$$I_2(\theta) = \|g - \hat{f}_\theta(\theta'\cdot)\|_2^2$$
$$I_3(\theta) = 2\int_S (\hat{g}_n(x) - g(x))(g(x) - \hat{f}_\theta(\theta' x))dx.$$

The following result is useful for controlling the decision variable under alternative hypothesis.

Proposition 5.3. *Suppose that $3/4 \le \beta < 1$ and $d \ge 3$ and Θ consist of d points $\theta_1 = (1,0,\ldots,0)$, $\theta_2 = (0,1,0,\ldots,0),\ldots,$ $\theta_d = (0,0,\ldots,0,1)$.*

$$\left|\int_{S(\theta' x)} g(A'y)dy_2\ldots dy_d - \sum_{j=1}^n \widetilde{W}_{nj}(\theta' x)g(X_j)\right| \le L\left(\sqrt{b_n^2 + (d-1)\left(\frac{1}{2n^{1/d}}\right)^2}\right)^\beta$$

Proposition 5.4. *Using Assumptions of Proposition 5.3 and supposing that Assumptions $(\mathcal{H}_1)$, $(\mathcal{H}_2)$, $(\mathcal{H}_3)$, $(\mathcal{H}_4)$, and $(\mathcal{H}_5)$ are satisfied, we have:*

$$\begin{aligned}\inf_{\theta\in\Theta}\{I_2(\theta)+I_3(\theta)\} &\geq (C\varphi_n(\alpha_n))^2\\ &-\left(2C_n^*+\alpha\sup_j|\varepsilon_j|+2C_n^{**}\right)\inf_{\theta\in\Theta}\{\|g-f_\theta(\theta'\cdot)\|_2\}\\ &-2C_n^*C_n^{**}-(C_n^*)^2-\left((c_*)^2+2c^*c_*\right)(\sup_j|\varepsilon_j|)^2\\ &-\left(2c^*C_n^{**}+\alpha C_n^*\right)\sup_j|\varepsilon_j|,\end{aligned}$$

with

$$C_n^*=L\left(\sqrt{b_n^2+(d-1)\Big(\frac{1}{2n^{1/d}}\Big)^2}\right)^\beta,\quad C_n^{**}=\Big(\frac{Ld^{2/(2-\tau)}C}{m!}\Big)h_n^\beta$$

and C,c^*,c_* are positive constants with $\alpha=2(c^*+c_*)$.

Proposition 5.5. *There exists a positive sequence $(B_n)_{n\in\mathbb{N}}$ such that*

$$\mathbb{P}_g^n(\sup_j|\varepsilon_j|>B_n)\longrightarrow 0\quad \text{as } n\longrightarrow+\infty$$

where $\lim_{n\to+\infty}B_n=+\infty$, $\lim_{n\to+\infty}B_nC_n^*=\lim_{n\to+\infty}B_nC_n^{**}=0$.

5.3 Large deviation for U-statistic

To prove our results, we need to give sharp exponential inequality for U-statistic (5.1). This inequality is deduced from the following result.

Proposition 5.6 (Chiabrando [2])**.** *Assume that there exist $u_{s,n},\ k_{s,n}$ and $c_s>0,\ s=1,2,3$ such that for any $p\in 4\mathbb{N}$.*

(i) $\mathbb{E}h_n^p(X_1,X_2)\leq k_{1,n}p^{c_1p}u_{1,n}^p$

(ii) $\max\left\{\mathbb{E}\max_{i\neq j}|\underline{h}_n^p(X_i,X_j)|;\mathbb{E}\underline{h}_{(0)}^p\right\}\leq k_{2,n}p^{c_2p}u_{2,n}^p$

(iii) $\mathbb{E}\max_i\left(\mathbb{E}\Big[\underline{h}_n^2(X_i,X)|X_i\Big]\right)^p\leq k_{3,n}p^{c_3p}u_{3,n}^p$

Then for any $x>0$,

$$\mathbb{P}\left(\frac{1}{n(n-1)}\left|\sum_{i\neq j}h_n(X_i,X_j)\right|\geq x\right)\leq c_n(x)e^{4-m_n(x)}.$$

The mixture term is defined by

$$m_n(x)=\inf[m_{i,n}(x),1\leq i\leq n]$$

where if $K_1,...,K_5$ are universal constants , we set

$$\begin{cases}m_{1,n}(x)=\left(\frac{nx}{eK_1\sigma_n}\right)^2, & m_{2,n}(x)=\frac{nx}{eK_2\sqrt{\underline{\sigma}_n}}\\ m_{3,n}(x)=\left(\frac{n\sqrt{n}x}{eK_3\sigma_n}\right)^{\frac{2}{3}}, & m_{4,n}(x)=\left(\frac{n^{\frac54}x}{eK_4\sqrt{\underline{\sigma}_{(0)}}}\right)^{\frac43}\\ m_{5,n}(x)=\left(\frac{n^2x}{eK_2u_{1,n}}\right)^{\delta(c_1)}, & \\ m_{6,n}(x)=\left(\frac{n\sqrt{n}x}{e(K_2\vee K_5)\sqrt{u_{2,n}}}\right)^{\delta(c_2)}, & \\ m_{7,n}(x)=\left(\frac{n^{\frac54}x}{eK_2u_{3,n}^{\frac14}}\right)^{\delta(c_3)} & \end{cases}$$

where

$$\sigma_n^2=\mathbb{E}(h_n^2(X_1,X_2)),\ \underline{\sigma}_n^2=\mathbb{E}(\underline{h}_n^2(X_1,X_2)),$$
$$\sigma_{(0)}^2=\mathbb{E}(\underline{h}_{(0)}^2(X)).$$

and

$$\delta(c_1)=\frac{1}{2+c_1},\ \delta(c_2)=\frac{2}{3+c_2},\delta(c_3)=\frac{4}{5+c_3}.$$

Moreover,

$$c_n(x)=\max\{c_{j,n}:j\mid m_n(x)=m_{j,n}(x)\}$$

where

$$\begin{cases}c_{1,n}=c_{2,n}=c_{4,n}=1,\\ c_{3,n}=n,\\ c_{5,n}=2nk_{1,n},\\ c_{6,n}=n^2k_{2,n},\\ c_{7,n}=k_{3,n}.\end{cases}$$

Corollary 5.1. *For all $y>0$, we have*

$$\mathbb{P}_g^n\Big(|U_n|>y\Big)\leq e^4\exp\Big(-\frac{n^2h_n^dy^2}{e^2K_1^2\Gamma_1}\Big).$$

6 Proofs of main results

6.1 Proof of Theorem 4.1

Using Proposition 5.1 and Corollary 5.1, for all $g\in\Delta_0$, we obtain

$$\begin{aligned}\mathbb{P}_g^n\Big(\phi_{n,\alpha_n}=1\Big)&=\mathbb{P}_g^n\Big(\inf_{\theta\in\Theta}T_n(\theta)>(\lambda\varphi_n(\alpha_n))^2\Big)\\ &\leq\mathbb{P}_g^n\Big(T_n(\theta_0)>(\lambda\varphi_n(\alpha_n))^2\Big)\\ &\leq\mathbb{P}_g^n\Bigg(U_n>(\lambda\varphi_n(\alpha_n))^2-\\ &\qquad\Big(\frac{Ld^{2/(2-\tau)}}{m!}\Big)^2h_n^{2\beta}+o_p\Big(\frac{1}{nh_n^{d/2}}\Big)\Bigg)\end{aligned}$$

$$\begin{aligned}\mathbb{P}_g^n\Big(\phi_{n,\alpha_n}=1\Big)&\leq\mathbb{P}_g^n\Big(U_n>\frac{(\lambda\varphi_n(\alpha_n))^2}{2}\Big)\\ &\leq e^4\exp\Big(-\frac{n^2h_n^d(\lambda\varphi_n(\alpha_n))^4}{4e^2K_1^2\Gamma_1}\Big)\end{aligned}$$

Choosing $\lambda>\left(4e^2K_1^2\Gamma_1\right)^{1/4}$, we obtain

$$\limsup_{n\to+\infty}\alpha_n^{-1}\sup_{g\in\Delta_0}\mathbb{P}_g^n\Big(\phi_{n,\alpha_n}=1\Big)\leq 1.$$

6.2 Proof of Theorem 4.2

Using Propositions 5.4 and 5.5, for all $g \in \Delta_n(C\varphi_n)$, we obtain :

$$\mathbb{P}_g^n\big(\phi_{n,\alpha_n}=0\big)=\mathbb{P}_g^n\big(I_1+\inf_{\theta\in\Theta}\{I_2(\theta)+I_3(\theta)\}<\lambda_n\big)$$
$$\le \mathbb{P}_g^n\Bigg(I_1+(C\varphi_n(\alpha_n))^2- \Big(2C_n^*+\alpha B_n+2C_n^{**}\Big)\inf_{\theta\in\Theta}\{\|g-f_\theta(\theta'\cdot)\|_2\}- 2C_n^*C_n^{**}-(C_n^*)^2-\Big((c_*)^2+2c^*c_*\Big)(B_n)^2- \Big(2c^*C_n^{**}+\alpha\, C_n^*\Big)B_n<\lambda_n\Bigg)$$

$$\mathbb{P}_g^n\big(\phi_{n,\alpha_n}=0\big)\le \mathbb{P}_g^n\Bigg(U_n<\lambda_n+\frac{1}{2}\lambda^2\varphi_n^2(\alpha_n)+ o_p\Big(\frac{1}{nh_n^{d/2}}\Big)-(C\varphi_n(\alpha_n))^2+ \Big(2C_n^*+\alpha B_n+2C_n^{**}\Big)\inf_{\theta\in\Theta}\{\|g-f_\theta(\theta'\cdot)\|_2\}+ 2C_n^*C_n^{**}+(C_n^*)^2+\Big((c_*)^2+2c^*c_*\Big)(B_n)^2+ \Big(2c^*C_n^{**}+\alpha\, C_n^*\Big)B_n\Bigg)$$

Thus, for n large enough, we have

$$\mathbb{P}_g^n\left(I_1+\inf_{\theta\in\Theta}\{I_2(\theta)+I_3(\theta)\}<\lambda_n\right)\le$$
$$\mathbb{P}_g^n\Bigg(U_n<-\Bigg[\lambda^2\varphi_n^2(\alpha_n)\Big(\frac{C^2}{\lambda^2}-\frac{1}{2}\Big)-\alpha B_n\inf_{\theta\in\Theta}\{\|g-f_\theta(\theta'\cdot)\|_2\}- \Big((c_*)^2+2c^*c_*\Big)B_n^2\Bigg]\Bigg)$$

Thus, we obtain

$$\mathbb{P}_g^n\left(I_1+\inf_{\theta\in\Theta}\{I_2(\theta)+I_3(\theta)\}<\lambda_n\right)\le \mathbb{P}_g^n(U_n<-y_n)$$
$$\le e^4\exp\Big(-\frac{1}{e^2K_1^2\Gamma_1}(nh_n^{d/2}y_n)^2\Big)$$

with

$$y_n=\lambda^2\varphi_n^2(\alpha_n)\Big(\frac{C^2}{\lambda^2}-\frac{1}{2}\Big)-\alpha B_n\inf_{\theta\in\Theta}\{\|g-f_\theta(\theta'\cdot)\|_2\}- \Big((c_*)^2+2c^*c_*\Big)B_n^2$$

Next, we obtain

$$\lim_{n\to+\infty}\sup_{g\in\Delta_n(C\varphi_n(\alpha_n))}\mathbb{P}_g^n(\phi_{n,\alpha_n}=0)\le$$
$$\lim_{n\to+\infty}e^4\exp\Big(-\frac{1}{e^2K_1^2\Gamma_1}\Big[\Big(\frac{C^2}{\lambda^2}-\frac{1}{2}\Big)-\alpha nh_n^{d/2}B_n\inf_{\theta\in\Theta}\{\|g-f_\theta(\theta'\cdot)\|_2\} -\Big((c_*)^2+2c^*c_*\Big)nh_n^{d/2}\,B_n^2\Big]^2\Big)$$

We recall that $\lim_{n\to+\infty}nh_n^{d/2}=+\infty$, $\lim_{n\to+\infty}B_n=+\infty$ and we choose B_n such that $\lim_{n\to+\infty}B_nC_n^*=\lim_{n\to+\infty}B_nC_n^{**}=0$. Hence, if $C>\frac{\sqrt{2}}{2}\lambda$, we have

$$\limsup_{n\to+\infty}\sup_{g\in\Delta_n(C\varphi_n(\alpha_n))}\mathbb{P}_g^n(\phi_{n,\alpha_n}=0)=0.$$

□

6.3 Proof of lower bound

The proof of this theorem is partly inspired by the lower bound established in Yodé[22] and Yodé[24].

Let us fix $\sigma>0$ and put $\delta_n=\sigma h_n$ and $M_n=\delta_n^{-1}$. The value of σ can be determined later. Suppose that M_n is an integer. Otherwise, one take its integer part. Let $\{u_1,...,u_{M_n},1\}$ be a regular subdivision of $[0,1]$ and put $A_{n,l}=[u_l,u_{l+1}[,\ l=1,...,M_n-1,\quad A_{n,M_n}=[u_{M_n},1]$ with $u_l=\frac{l-1}{M_n}$. For a multi-index $s=(s_1,...,s_d)\in\mathcal{M}_n\triangleq\{1,...,M_n\}^d$, define $A_{n,s}=A_{n,s_1}\times...\times A_{n,s_d}$. Thus, the family $\{A_{n,s},s\in\mathcal{M}_n\}$ is a partition of $[0,1]^d$.

Let ψ be a bounded and infinitely differentiable function with support $[0,1]$ such that:

$$\int_{[0,1]}\psi(x)dx=0\ ,\int_{[0,1]}\psi^2(x)dx=1.$$

For any $s=(s_1,...,s_d)\in\mathcal{M}_n$, consider the function :

$$\psi_{n,s}:x=(x_1,...,x_d)\in\mathbb{R}^d\longrightarrow\prod_{r=1}^d\psi\Big(\frac{x_r-u_{s_r}}{\delta_n}\Big)$$

such that

$$\delta_n^{\beta+d/2}\sum_{s\in\mathcal{M}_n}\sup_{(x_1,\dots,x_{i-1},x_{i+1},\dots,x_d)\in[0,1]^{d-1}}\left|\frac{\partial^m\psi_{n,s}}{\partial x_i^m}(x^{(1)})-\frac{\partial^m\psi_{n,s}}{\partial x_i^m}(x^{(2)})\right|\le\frac{L}{2}|x_i^{(1)}-x_i^{(2)}|^\tau$$

where $x^{(l)}=(x_1,\dots,x_{i-1},x_i^{(l)},x_{i+1},\dots,x_d)$, $l=1,2$, for all $i\in\{1,\dots,d\}$ and $x_i^{(1)},x_i^{(2)}\in[0,1]$. The function $\psi_{n,s}$ is compactly supported in $A_{n,s}$ and we have

$$\int_{A_{n,s}}\psi_{n,s}(x)dx=0\ ,\int_{A_{n,s}}\psi_{n,s}^2(x)dx=1.$$

Let θ_0 be the real value of the parameter. We assume that the function $f_{\theta_0}(\theta_0'.)$ is such that $f_{\theta_0}(\theta_0'.)\in\Sigma_d(\beta,\frac{L}{2})$. Next consider the collection of function

$$\mathcal{F}_n=\{g_k,\ k=1,...,2^{M_n^d}\}$$

where

$$g_k(.)\triangleq f_{\theta_0}(\theta_0'.)+\delta_n^{\beta+d/2}\sum_{s\in\mathcal{M}_n}a_{k,s}\psi_{n,s}(.),$$

where $a_{k,s}$, $s \in \mathcal{M}_n$ are i.i.d random variables taking values 1 and -1 with probability $\frac{1}{2}$. If $\sigma^\beta \geq C$ then $g_k \in \Delta_n(C\varphi_n(\alpha_n))$. Moreover, we obtain that $g_k \in \Sigma_d(\beta, L)$.

Proof of Theorem 4.3. Let us put

$$Z_n = \frac{1}{2^{M_n^d}} \sum_{k=1}^{2^{M_n^d}} \frac{\mathrm{d}\mathbb{P}^n_{g_k}}{\mathrm{d}\mathbb{P}^n_{f_{\theta_0}}}.$$

Then we obtain

$$Z_n = \prod_{s\in\mathcal{M}_n} \exp\Big(-\frac{\delta_n^{2\beta+d}}{2\sigma^2}\sum_{i=1}^n \psi^2_{n,s}(X_i)\Big)\times \cosh\left(\frac{\delta_n^{\beta+d/2}}{\sigma^2}\sum_{i=1}^n \big(Y_i - f_{\theta_0}(\theta_0' X_i)\big)\psi_{n,s}(X_i)\right).$$

Let ϕ_n a test of asymptotical level α_n. So we have $\mathbb{P}^n_{f_{\theta_0}}\{\phi_n = 1\} \leq \alpha_n$ and

$$\sup_{g\in\Delta_n(C\varphi_n(\alpha_n))} \mathbb{P}^n_g\{\phi_n = 0\} \geq 1 - c - c^{-1}\alpha_n \mathbb{E}^n_{f_{\theta_0}}(Z_n^2). \tag{6.1}$$

Moreover, we obtain

$$\mathbb{E}^n_{f_{\theta_0}}(Z_n^2) = \prod_{s\in\mathcal{M}_n} \exp\Big(-\frac{\delta_n^{2\beta+d}}{\sigma^2}\sum_{i=1}^n \psi^2_{n,s}(X_i)\Big)\times \mathbb{E}^n_{f_{\theta_0}}\left[\cosh^2\left(Z_s\frac{\delta_n^{\beta+d/2}}{\sigma}\sqrt{\sum_{i=1}^n \psi^2_{n,s}(X_i)}\right)\right]$$

Next, using Assumption $(\mathcal{H}_7)$ and applying the following equality

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \cosh^2(ux)\exp\Big(\frac{-x^2}{2}\Big)dx = \frac{1}{2}\Big(1 + \exp(2u^2)\Big),$$

$u \in \mathbb{R}$ and the fact that there exists $a_* > 0$, such that for any $x \in \mathbb{R}$, $\cosh(x) \leq \exp(a_* x^2)$, we obtain :

$$\mathbb{E}^n_{f_{\theta_0}}(Z_n^2) \leq \exp\left(a_*\frac{\delta_n^{4\beta+2d}}{\sigma^4}\sum_{s\in\mathcal{M}_n}\Big(\sum_{i=1}^n \psi^2_{n,s}(X_i)\Big)^2\right) \leq e^{4a_*\,\sigma^{4\beta+d-4}\|\psi\|_\infty^{4d}}\cdot \alpha_n^{-a_*\,\sigma^{4\beta+d-4}\|\psi\|_\infty^{4d}}$$

Thus, for any $C \leq \sigma^\beta$, and using (6.1) we obtain

$$\inf_{\phi_n\in\Gamma_n(\alpha_n)}\sup_{g\in\Delta_n(C\varphi_n(\alpha_n))} \mathbb{P}^n_g\{\phi_n = 0\} \geq 1 - c - c^{-1}\alpha_n^{1-a_*\,\sigma^{4\beta+d-4}\|\psi\|_\infty^{4d}}\, e^{4a_*\,\sigma^{4\beta+d-4}\|\psi\|_\infty^{4d}}$$

For all $c > 0$, for n large enough, if $\sigma \leq a_*^{-\frac{1}{4\beta+d-4}}\|\psi\|_\infty^{-\frac{4d}{4\beta+d-4}}$, then we have the result. □

7 Proofs of preliminaries results

7.1 Proof of Proposition 5.1

1. To prove this result, we show that $\mathbb{E}^n_g(U_n) = 0$ and there exists $C_{**} > 0$ such that $var(U_n) \leq \frac{C_{**}}{n^2 h_n^d}$. Since errors ε_j are independent centered random variables with $var(\varepsilon_i) = \sigma^2$, we deduce that $\mathbb{E}^n_g(U_n) = 0$ and

$$var(U_n) \leq \sigma^4 \sum_{j\neq k}\Big(\int_S \sqrt{|W_{nj}(x)|}\sqrt{|W_{nj}(x)|}|W_{nk}(x)|dx\Big)^2.$$

Using Cauchy-Schwarz inequality, Assumption $(\mathcal{H}_3)$ and Proposition 3.1, we obtain

$$var(U_n) \leq \sigma^4\sum_{j\neq k}\int_S |W_{nj}(x)|dx \times \int_S |W_{nj}(x)||W_{nk}(x)|^2 1_{\{\|X_k - x\|\leq h_n\}}dx \leq \frac{(c^*)^2\sigma^4}{n^2h_n^{2d}}\times \sum_{j=1}^n\int_S |W_{nj}(x)|dx\int_S |W_{nj}(x)|\Big(\sum_{k=1}^n 1_{\{\|X_k - x\|\leq h_n\}}\Big)dx.$$

$$var(U_n) \leq \frac{(c^*)^4\sigma^4 a_0}{n^2h_n^d}.$$

2. We have the following decomposition

$$\big\|\hat{g}_n - \mathbb{E}^n_g(\hat{g}_n)\big\|^2 - \sigma^2\sum_{j=1}^n\int_S W^2_{nj}(x)dx = \sum_{j=1}^n(\varepsilon_j^2 - \sigma^2)\int_S W^2_{nj}(x)dx + U_n.$$

Thus, we obtain

$$\mathbb{E}^n_g\Big(\sum_{j=1}^n(\varepsilon_j^2 - \sigma^2)\int_S W^2_{nj}(x)dx + U_n\Big) = 0$$

and

$$var\left(\sum_{j=1}^n(\varepsilon_j^2 - \sigma^2)\int_S W^2_{nj}(x)dx\right) = O\left(\frac{1}{n^3h_n^{2d}}\right).$$

We deduce that

$$\sum_{j=1}^n(\varepsilon_j^2 - \sigma^2)\int_S W^2_{nj}(x)dx = o_p\left(\frac{1}{nh_n^{d/2}}\right).$$

3. Using Taylor's formula, we obtain

$$\mathbb{E}^n_g(\hat{g}_n(x)) - g(x) = \sum_{j=1}^n W_{nj}(x)(g(X_j) - g(x))$$
$$= \sum_{1\leq|i|\leq m-1}\frac{1}{i!}D^i g(x)\sum_{j=1}^n W_{nj}(x)(X_j - x)^i + \frac{1}{m!}\sum_{s=1}^d\sum_{j=1}^n W_{nj}(x)(X_{j,s} - x_s)^m\times \Big(\frac{\partial^m g}{\partial x_s^m}(X_j) - \frac{\partial^m g}{\partial x_s^m}(\tilde{X}_j)\Big)$$

where

$$X_j = (X_{j,1}, \ldots, X_{j,s-1}, X_{j,s}, X_{j,s+1}, \ldots, X_{j,d})'$$
$$\tilde{X}_j = (X_{j,1}, \ldots, X_{j,s-1}, \tilde{u}_s, X_{j,s+1}, \ldots, X_{j,d})'$$

According to $(\mathcal{H}_1)$ and Proposition 3.1, we obtain

$$\left|\mathbb{E}_g^n(\hat{g}_n(x)) - g(x)\right| \leq \frac{L}{m!}\sum_{j=1}^n |W_{nj}(x)| \left(\sum_{s=1}^d |X_{j,s} - x_s|^\beta\right) 1_{\{\|X_j - x\| \leq h_n\}}$$

$$\left|\mathbb{E}_g^n(\hat{g}_n(x)) - g(x)\right| \leq \frac{L}{m!}\sum_{j=1}^n |W_{nj}(x)|\, d^{2/(2-\tau)} \|X_j - x\|^\beta 1_{\{\|X_j - x\| \leq h_n\}}.$$

$$\left\|\mathbb{E}_g^n(\hat{g}_n) - g\right\|_2^2 \leq \left(\frac{L d^{2/(2-\tau)} c^*}{m!}\right)^2 h_n^{2\beta}.$$

4. We have

$$\int_S \left(\hat{g}_n(x) - \mathbb{E}_g^n(\hat{g}_n(x))\right)\left(\mathbb{E}_g^n(\hat{g}_n(x)) - g(x)\right) dx = \sum_{j=1}^n \varepsilon_j \int_S W_{nj}(x)\left(\mathbb{E}_g^n(\hat{g}_n(x)) - g(x)\right) dx.$$

Therefore, we obtain

$$\mathbb{E}_g^n\left[\int_S \left(\hat{g}_n(x) - \mathbb{E}_g^n(\hat{g}_n(x))\right)\left(\mathbb{E}(\hat{g}_n(x)) - g(x)\right) dx\right] = 0.$$

Moreover, by using the Cauchy-Schwartz inequality and Proposition 3.1, we obtain

$$var\left[\sum_{j=1}^n \varepsilon_j \int_S W_{nj}(x)\left(\mathbb{E}_g^n(\hat{g}_n(x)) - g(x)\right) dx\right] = O\left(\frac{h_n^{2\beta}}{n}\right).$$

This implies the result since $\frac{h_n^{2\beta}}{n} = o\left(\frac{1}{n^2 h_n^d}\right)$.

7.2 Proof of Proposition 5.2

We have the following decomposition

$$I_2(\theta_0) = \sum_{s=1}^6 I_{2,s}(\theta_0)$$

where

$$I_{2,1}(\theta_0) = \|g - f_{\theta_0}(\theta_0'\cdot)\|_2^2$$
$$I_{2,2}(\theta_0) = \left\|f_{\theta_0}(\theta_0'\cdot) - \mathbb{E}_g^n\left(\hat{f}_{\theta_0}(\theta_0'\cdot)\right)\right\|_2^2$$
$$I_{2,3}(\theta_0) = \left\|\hat{f}_{\theta_0}(\theta_0'\cdot) - \mathbb{E}_g^n\left(\hat{f}_{\theta_0}(\theta_0'\cdot)\right)\right\|_2^2$$
$$I_{2,4}(\theta_0) = 2\int_S \left(g(x) - f_{\theta_0}(\theta_0'x)\right)\left(f_{\theta_0}(\theta_0'x) - \mathbb{E}_g^n\left(\hat{f}_{\theta_0}(\theta_0'x)\right)\right) dx$$
$$I_{2,5}(\theta_0) = 2\int_S \left(g(x) - f_{\theta_0}(\theta_0'x)\right)\left(\mathbb{E}_g^n\left(\hat{f}_{\theta_0}(\theta_0'x)\right) - \hat{f}_{\theta_0}(\theta_0'x)\right) dx$$
$$I_{2,6}(\theta_0) = 2\int_S \left(f_{\theta_0}(\theta_0'x) - \mathbb{E}_g^n\left(\hat{f}_{\theta_0}(\theta_0'x)\right)\right) \times \left(\mathbb{E}_g^n\left(\hat{f}_{\theta_0}(\theta_0'x)\right) - \hat{f}_{\theta_0}(\theta_0'x)\right) dx$$

To proove our Proposition, we need the following result

Lemma 7.1. *Under the null hypothesis, we have*

$$\left|\mathbb{E}_g^n(\hat{f}_{\theta_0}(\theta_0'x)) - f_{\theta_0}(\theta_0'x)\right| \leq \frac{Lc_* b_n^\beta}{m!}.$$

Proof. We have

$$\mathbb{E}_g^n(\hat{f}_{\theta_0}(\theta_0'x)) - f_{\theta_0}(\theta_0'x) = \sum_{j=1}^n \widetilde{W}_{nj}(\theta_0'x)\Big(g(X_j) - f_{\theta_0}(\theta_0'x)\Big).$$

Using Taylor's formula, we have

$$\mathbb{E}_g^n(\hat{f}_{\theta_0}(\theta_0'x)) - f_{\theta_0}(\theta_0'x) = \sum_{l=1}^{m-1} \frac{f_{\theta_0}^{(l)}(\theta_0'x)}{l!}\left[\sum_{j=1}^n \widetilde{W}_{nj}(\theta_0'x)(\theta_0'X_j - \theta_0'x)^l\right] + \frac{1}{m!}\sum_{j=1}^n \widetilde{W}_{nj}(\theta_0'x)(\theta_0'X_j - \theta_0'x)^m \left[f_{\theta_0}^{(m)}(\tilde{u}) - f_{\theta_0}^{(m)}(\theta_0'x)\right]$$

where $|\theta_0'X_j - \tilde{u}| \leq |\theta_0'X_j - \theta_0'x|$. Since

1. $\sum_{j=1}^n \widetilde{W}_{nj}(\theta_0'x)(\theta_0'X_j - \theta_0'x)^l = 0$ for $l = 1, \ldots, m$
2. $\sum_{j=1}^n \left|\widetilde{W}_{nj}(\theta_0'x)\right| \leq c_* < +\infty,$
3. $\widetilde{W}_{nj}(\theta_0'x) = 0$ if $|\theta_0'X_j - \theta_0'x| > b_n,$

then, we obtain

$$\left|\frac{1}{m!}\sum_{j=1}^n \widetilde{W}_{nj}(\theta_0'x)(\theta_0'X_j - \theta_0'x)^m \left[f_{\theta_0}^{(m)}(\tilde{u}) - f_{\theta_0}^{(m)}(\theta_0'x)\right]\right| \leq \frac{Lc_* b_n^\beta}{m!}.$$

□

Proof of (5.2)**.** Under the null hypothesis (H_0), we have $I_{2,1}(\theta_0) = I_{2,4}(\theta_0) = I_{2,5}(\theta_0) = 0$. From Lemma 7.1, we deduce that $I_{2,2} = O_p\left(b_n^{2\beta}\right)$. Thus, we obtain

$$I_{2,2}(\theta_0) = o_p\left(\frac{1}{nh_n^{d/2}}\right). \tag{7.1}$$

We can write

$$I_{2,6}(\theta_0) = -2\sum_{j=1}^n \left[\varepsilon_j \int_S \left(f_{\theta_0}(\theta_0'x) - \mathbb{E}_g^n\left(\hat{f}_{\theta_0}(\theta_0'x)\right)\right) \times \widetilde{W}_{nj}(\theta_0'x) dx\right].$$

Therefore, we obtain $\mathbb{E}_g^n(I_{2,6}(\theta_0)) = 0$. Since $\frac{b_n^{2\beta}}{nb_n} \asymp b_n^{4\beta}$, using Cauchy-Schwartz inequality and Lemma 7.1, we have

$$var(I_{2,6}(\theta_0)) = 4\sigma^2 \sum_{j=1}^{n} \left(\int_S \left(f_{\theta_0}(\theta_0' x) - \mathbb{E}_g^n \left(\hat{f}_{\theta_0}(\theta_0' x) \right) \right) \widetilde{W}_{nj}(\theta_0' x) dx \right)^2$$

$$var(I_{2,6}(\theta_0)) \leq \frac{4\sigma^2 L^2 c_*^3}{(m!)^2} b_n^{4\beta}$$

Therefore, we deduce that $I_{2,6}(\theta_0) = O_p\left(b_n^{2\beta}\right)$. This implies

$$I_{2,6}(\theta_0) = o_p\left(\frac{1}{nh_n^{d/2}}\right). \tag{7.2}$$

For $I_{2,3}(\theta_0)$, we have

$$I_{2,3}(\theta_0) = \int_S \left(\sum_{j=1}^{n} \widetilde{W}_{nj}(\theta_0' x)\varepsilon_j \right)^2 dx = I_{2,3}^1(\theta_0) + I_{2,3}^2(\theta_0)$$

where

$$I_{2,3}^1(\theta_0) = \sum_{j=1}^{n} \varepsilon_j^2 \int_S \widetilde{W}_{nj}^2(\theta_0' x) dx$$

$$I_{2,3}^2(\theta_0) = \sum_{j \neq k} \varepsilon_j \varepsilon_k \int_S \widetilde{W}_{nj}(\theta_0' x)\widetilde{W}_{nk}(\theta_0' x) dx.$$

We have the following results

(i)

$$\mathbb{E}_g^n\left(I_{2,3}^1(\theta_0)\right) = \mu_2 \sum_{j=1}^{n} \int_S \widetilde{W}_{nj}^2(\theta_0' x) dx \leq \mu_2 \frac{c_*^2}{nb_n}$$

with $\mu_2 = \mathbb{E}_g^n(\varepsilon_j^2)$

$$var\left(I_{2,3}^1(\theta_0)\right) = \sum_{j=1}^{n} var(\varepsilon_j^2) \left(\int_S \widetilde{W}_{nj}^2(\theta_0' x) dx \right)^2$$

$$var\left(I_{2,3}^1(\theta_0)\right) \leq \frac{\mu_4 c_*^4}{n^3 b_n^3} \quad \text{with } \mu_4 = var(\varepsilon_j^2)$$

Then, we deduce $I_{2,3}^1(\theta_0) = O_p\left(\frac{1}{nb_n}\right)$.

(ii) We have

$$\mathbb{E}_g^n(I_{2,3}^2(\theta_0)) = 0$$

$$var(I_{2,3}^2(\theta_0)) = (\mathbb{E}_g^n(\varepsilon^2))^2 \sum_{j \neq k} \left(\int_S \widetilde{W}_{nj}(\theta' x)\widetilde{W}_{nk}(\theta_0' x) dx \right)^2 \leq \frac{\mu_2^2 c_*^4}{n^2 b_n^2}$$

Then, we have $I_{2,3}^2(\theta_0) = O_p\left(\frac{1}{nb_n}\right)$.

From (i) and (ii), we deduce that

$$I_{2,3}(\theta_0) = o_p\left(\frac{1}{nh_n^{d/2}}\right) \tag{7.3}$$

Under the null hypothesis (H_0), according to (7.1), (7.2) and (7.3), we deduce that

$$I_2(\theta_0) = o_p\left(\frac{1}{nh_n^{d/2}}\right). \tag{7.4}$$

Proof of (5.3)**.** We have the following decomposition

$$I_3(\theta_0) = \sum_{s=1}^{6} I_{3,s}(\theta_0)$$

where

$$I_{3,1}(\theta_0) = 2\int_S \left(\hat{g}_n(x) - \mathbb{E}_g^n(\hat{g}_n(x))\right)\left(g(x) - f_{\theta_0}(\theta_0' x)\right) dx$$

$$I_{3,2}(\theta_0) = 2\int_S \left(\hat{g}_n(x) - \mathbb{E}_g^n(\hat{g}_n(x))\right) \times \left(f_{\theta_0}(\theta_0' x) - \mathbb{E}_g^n\left(\hat{f}_{\theta_0}(\theta_0' x)\right)\right) dx$$

$$I_{3,3}(\theta_0) = 2\int_S \left(\hat{g}_n(x) - \mathbb{E}_g^n(\hat{g}_n(x))\right) \times \left(\mathbb{E}_g^n\left(\hat{f}_{\theta_0}(\theta_0' x)\right) - \hat{f}_{\theta_0}(\theta_0' x)\right) dx$$

$$I_{3,4}(\theta_0) = 2\int_S \left(\mathbb{E}_g^n(\hat{g}_n(x)) - g(x)\right)\left(g(x) - f_{\theta_0}(\theta_0' x)\right) dx$$

$$I_{3,5}(\theta_0) = 2\int_S \left(\mathbb{E}_g^n(\hat{g}_n(x)) - g(x)\right) \times \left(f_{\theta_0}(\theta_0' x) - \mathbb{E}_g^n\left(\hat{f}_{\theta_0}(\theta_0' x)\right)\right) dx$$

$$I_{3,6}(\theta_0) = 2\int_S \left(\mathbb{E}_g^n(\hat{g}_n(x)) - g(x)\right) \times \left(\mathbb{E}_g^n\left(\hat{f}_{\theta_0}(\theta_0' x)\right) - \hat{f}_{\theta_0}(\theta_0' x)\right) dx$$

1. Under (H_0), we have $I_{3,1}(\theta_0) = I_{3,4}(\theta_0) = 0$.
2. $\mathbb{E}_g^n(I_{3,2}(\theta_0)) = 0$. Next, we can write

$$var(I_{3,2}(\theta_0)) = 4\sigma^2 \times \sum_{j=1}^{n} \left(\int_{\|X_j - x\| \leq h_n} W_{nj}(x) \left(f_{\theta_0}(\theta_0' x) - \mathbb{E}_g^n(\hat{f}_{\theta_0}(\theta_0' x)) \right) dx \right)^2 \leq \frac{4\sigma^2 C b_n^{2\beta} h_n^d}{nh_n^d} = \frac{4\sigma^2 C b_n^{2\beta}}{n}.$$

We have $I_{3,2}(\theta_0) = o_p\left(\frac{1}{nh_n^{d/2}}\right)$.

3. We have the following decomposition

$$I_{3,3}(\theta_0) = -2\int_S \left(\sum_{j=1}^n \varepsilon_j W_{nj}(x)\right)\left(\sum_{k=1}^n \varepsilon_k \widetilde{W}_{nk}(\theta_0' x)\right)dx = I_{3,3}^1(\theta_0) + I_{3,3}^2(\theta_0)$$

where

$$I_{3,3}^1(\theta_0) = -2\sum_{j=1}^n \varepsilon_j^2 \int_S W_{nj}(x)\widetilde{W}_{nj}(\theta_0' x)dx$$
$$I_{3,3}^2(\theta_0) = -2\sum_{j\neq k} \varepsilon_j\varepsilon_k \int_S W_{nj}(x)\widetilde{W}_{nk}(\theta_0' x)dx.$$

$$\left|\mathbb{E}_g^n(I_{3,3}^1(\theta_0))\right| = 2\sigma^2\left|\sum_{j=1}^n \int_S W_{nj}(x)\widetilde{W}_{nj}(\theta_0' x)dx\right| \leq \frac{2C\sigma^2}{nb_n}\sum_{j=1}^n \int_S |W_{nj}(x)|dx.$$

Then, we deduce that

$$\mathbb{E}_g^n(I_{3,3}^1(\theta_0)) = O\left(\frac{1}{nb_n}\right) \Rightarrow \mathbb{E}(I_{3,3}^1(\theta_0)) = o\left(\frac{1}{nh_n^{d/2}}\right). \tag{7.5}$$

Moreover, Cauchy-Schwartz inequality implies

$$var(I_{3,3}^1(\theta_0)) = 4\sum_{j=1}^n var(\varepsilon_j^2)\left(\int_S W_{nj}(x)\widetilde{W}_{nj}(\theta_0' x)dx\right)^2.$$

$$var(I_{3,3}^1(\theta_0)) \leq \frac{4C\mu_4}{n^3 b_n^2}.$$

Therefore, we get

$$\sqrt{var(I_{3,3}^1(\theta_0))} = O\left(\frac{1}{n^{3/2}b_n}\right) \Rightarrow \sqrt{var(I_{3,3}^1(\theta_0))} = o\left(\frac{1}{nh_n^{d/2}}\right) \tag{7.6}$$

From (7.5) and (7.6), we deduce that

$$I_{3,3}^1(\theta_0) = o_p\left(\frac{1}{nh_n^{d/2}}\right) \tag{7.7}$$

We have $\mathbb{E}(I_{3,3}^2(\theta_0)) = 0$ and

$$\mathbb{E}\left[\left(I_{3,3}^2(\theta_0)\right)^2\right] = 4\sum_{j\neq k}\mathbb{E}(\varepsilon_j^2\varepsilon_k^2)\left(\int_S W_{nj}(x)\widetilde{W}_{nk}(\theta_0' x)dx\right)^2 \leq$$
$$4\sigma^4\sum_{j=1}^n \int_S W_{nj}^2(x)dx\int_{\|X_j - x\|\leq h}\left(\sum_{k=1}^n \widetilde{W}_{nk}^2(\theta_0' x)\right)dx.$$

we get

$$\mathbb{E}\left[\left(I_{3,3}^2(\theta_0)\right)^2\right] = O\left(\frac{1}{n^2 b_n}\right).$$

Thus, according to $\frac{1}{n\sqrt{b_n}} = o\left(\frac{1}{nh_n^{d/2}}\right)$, we have

$$I_{3,3}^2(\theta_0) = o_p\left(\frac{1}{nh_n^{d/2}}\right). \tag{7.8}$$

From (7.7) and (7.8), we deduce that

$$I_{3,3}(\theta_0) = o_p\left(\frac{1}{nh_n^{d/2}}\right). \tag{7.9}$$

4. Using Cauchy-Schwartz inequality and Proposition 7.1, we have

$$|I_{3,5}(\theta_0)| \leq \left\|\mathbb{E}_g^n(\hat{g}_n) - g\right\|_2 \left\|f_{\theta_0}(\theta'\cdot) - \mathbb{E}_g^n\left(\hat{f}_{\theta_0}(\theta_0'\cdot)\right)\right\|_2.$$

Since

$$\left\|\mathbb{E}_g^n(\hat{g}_n) - g\right\|_2 = O(h_n^\beta)$$

and $\left\|f_{\theta_0}(\theta_0'\cdot) - \mathbb{E}_g^n\left(\hat{f}_{\theta_0}(\theta_0'\cdot)\right)\right\|_2 = O(b_n^\beta)$, we have $I_{3,5}(\theta_0) = O(h_n^\beta b_n^\beta)$. According to $d > 2$, we deduce that

$$I_{3,5}(\theta_0) = o\left(\frac{1}{nh_n^{d/2}}\right).$$

5. Using Cauchy-Schwartz inequality, we have

$$|I_{3,6}(\theta_0)| \leq \left\|\mathbb{E}_g^n(\hat{g}_n) - g\right\|_2 \times \left\|\mathbb{E}_g^n\left(\hat{f}_{\theta_0}(\theta_0'\cdot)\right) - \hat{f}_{\theta_0}(\theta_0'\cdot)\right\|_2.$$

Since

$$\left\|\mathbb{E}_g^n(\hat{g}_n) - g\right\|_2 = O(h^\beta)_n$$

and $\left\|\mathbb{E}_g^n\left(\hat{f}_{\theta_0}(\theta_0'\cdot)\right) - \hat{f}_{\theta_0}(\theta_0'\cdot)\right\|_2 = O\left(\frac{1}{\sqrt{nb_n}}\right)$, we have

$$I_{3,6}(\theta_0) = O\left(\frac{h_n^\beta}{\sqrt{nb_n}}\right).$$

Therefore, we have

$$I_{3,6}(\theta_0) = o\left(\frac{1}{nh_n^{d/2}}\right).$$

Thus, we deduce that

$$I_3(\theta_0) = o_p\left(\frac{1}{nh_n^{d/2}}\right).$$

7.3 Proof Corollary 5.1

We consider the U-statistic

$$U_n = \sum_{1\leq j\neq k\leq n} L_n(\varepsilon_j, \varepsilon_k)$$

and

$$L_n(\varepsilon_j,\varepsilon_k)=\varepsilon_j\varepsilon_k\int_S W_{nj}(x)W_{nk}(x)dx.$$

We have

$$\mathbb{E}_g^n\Big(L_n^p(\varepsilon_1,\varepsilon_2)\Big)\le \frac{a^{2p}p^{2\nu p}(3^m(d+m)!)^p\Big(d+1+\binom{d}{2}\Big)^p K_{max}^{2p}\ h_n^{-pd}h_n^d}{(d!)^p(m!)^{2p}\ n^{2p}\lambda_0^{2p}},$$

$$\mathbb{E}_g^n\Big(\max_{i\neq j}|\underline{L}_n^p(\varepsilon_i,\varepsilon_j)|\Big)\le \frac{(c^*)^{2p}\Big(d+1+\binom{d}{2}\Big)^{2p}\sigma^{2p}K_{max}^{2p}}{\lambda_0^{2p}n^{2p}h_n^{pd}}\omega_n^{2p}\ 2^{2p\nu}p^{2p\nu},$$

$$\mathbb{E}_g^n\big(\underline{L}_0^p(\varepsilon)\big)\le \frac{\Big(d+1+\binom{d}{2}\Big)^{2p}K_{max}^{4p}}{\lambda_0^{4p}n^{2p}h_n^{pd}}\sigma^{2p}a^{2p}2^{2p\nu+p}p^{2p\nu},$$

$$\max\Big\{\mathbb{E}_g^n\max_{i\neq j}|\underline{L}_n^p(\varepsilon_i,\varepsilon_j)|;\mathbb{E}_g^n\underline{L}_{(0)}^p(\varepsilon)\Big\}\le \frac{(c^*)^{2p}\Big(d+1+\binom{d}{2}\Big)^{2p}K_{max}^{2p}\sigma^{2p}2^{2p\nu}p^{2p\nu}\omega_n^{2p}}{\lambda_0^{2p}n^{2p}h_n^{pd}},$$

$$\mathbb{E}_g^n\Big\{\max_i\Big(\mathbb{E}\Big[\underline{L}_n^2(\varepsilon_i,\varepsilon)|\varepsilon_i\Big]\Big)^p\Big\}\le \frac{(c^*)^{2p}\Big(d+1+\binom{d}{2}\Big)^{4p}K_{max}^{6p}\sigma^{6p}}{\lambda_0^{6p}n^{2p}h_n^{pd}}\omega_n^{2p}2^{2p\nu}p^{2p\nu}.$$

Applying the Proposition 5.1 to the kernel L

$$\mathbb{P}_g^n(|U_n|\ge x)\le c_n(x)e^{4-m_n(x)}$$

where $m_n(x)$ is the mixture defined as the minimum of $m_{i,n}(x)$, $i=\overline{1,7}$:

$$\begin{cases} m_{1,n}(x)=\Big(\frac{nh_n^{d/2}x}{eK_1\Gamma_1^{1/2}}\Big)^2,\\ m_{2,n}(x)=\frac{nh_n^{d/4}x}{eK_2\Gamma_2^{1/4}}\\ m_{3,n}(x)=\Big(\frac{n\sqrt{n}h_n^{d/2}x}{eK_3\Gamma_1^{1/2}}\Big)^{\frac{2}{3}},\\ m_{4,n}(x)=\Big(\frac{n^{\frac{5}{4}}h_n^{d/2}x}{eK_4\Gamma_3^{1/4}}\Big)^{\frac{4}{3}}\\ m_{5,n}(x)=\Big(\frac{n^2h_n^dx}{eK_2a^2K_{max}^2}\Big)^{\frac{1}{2+c_1}},\\ m_{6,n}(x)=\Big(\frac{n\sqrt{n}h_n^{d/2}x}{e(K_2\vee K_5)\sigma\ \omega_nK_{max}}\Big)^{\frac{2}{3+c_2}},\\ m_{7,n}(x)=\Big(\frac{n^{\frac{5}{4}}h_n^{d/4}x}{eK_2\ 2^{\nu/2}\omega_n^{1/2}\sigma^{3/2}K_{max}^{3/2}}\Big)^{\frac{4}{5+c_3}}\end{cases}$$

with

$$u_{n,1}=a^2K_{max}^2h_n^{-d},\quad u_{n,2}=\sigma^2\omega_n^2h_n^{-d}K_{max}^2,$$

$$u_{n,3}=2^{2\nu}\omega_n^2h_n^{-d}\sigma^6K_{max}^6$$

$$c_1=c_2=c_3=2\nu,$$

$$k_{1,n}=\frac{h_n^d(3^m(d+m)!)^p\Big(d+1+\binom{d}{2}\Big)^p}{n^{2p}\lambda_0^{2p}(d!)^p(m!)^{2p}},$$

$$k_{2,n}=\frac{(c^*)^{2p}\Big(d+1+\binom{d}{2}\Big)^{2p}\ 2^{2p\nu}}{\lambda_0^{2p}n^{2p}},$$

$$k_{3,n}=\frac{\Big(d+1+\binom{d}{2}\Big)^{4p}(c^*)^{2p}}{\lambda_0^{6p}n^{2p}}.$$

7.4 Proof of Lemma 5.3

Let I_t be such that

$$I_t=\Big\{j:\theta'X_j=\frac{t}{n^{1/d}}\Big\}\ \text{ with }\ \sharp I_t=n^{(d-1)/d}.$$

$\sharp I_t$ depends on θ. It exactly $n^{(d-1)/d}$ if $\theta\in\Theta$.
We know that

$$f_\theta(u)=\frac{\int_{S(u)}g(A'y)dy_2\dots dy_d}{\int_{S(u)}dy_2\dots dy_d}$$

with $S(u)=\{(y_2,\dots,y_d):y_1=u,A'y\in[0,1]^d\}$.
As $\theta\in\Theta$, we have

$$\int_{S(\theta'x)}dy_2\dots dy_d=1.$$

If $\theta\notin\Theta$, then the integral above must depends on θ.
The estimator under the null hypothesis (H_0) is

$$\hat{f}_\theta(\theta'x)=\sum_{j=1}^n\widetilde{W}_{nj}(\theta'x)Y_j$$

with the following two conditions:

$$\sum_{j=1}^n\widetilde{W}_{nj}(\theta'x)=1,$$

$$\forall i,j\in I_t:\widetilde{W}_{ni}(\theta'x)=\widetilde{W}_{ni}(\theta'x)=\widetilde{W}_{nI_t}(\theta'x).$$

Therefore, we have

$$\sum_{t=1}^{n^{1/d}}n^{(d-1)/d}\widetilde{W}_{nI_t}(\theta'x)=1.$$

Let $a=(a_1,\dots,a_d)$ and let us denote $S_a(\theta'x)$ the subset

$$S_a(u)=\Big\{(y_2,\dots,y_d):y_1=u,\frac{a_k}{n^{1/d}}-\frac{1}{2n^{1/d}}\le y_k\le \frac{a_k}{n^{1/d}}+\frac{1}{2n^{1/d}},k=2,\dots,d,A'y\in[0,1]^d\Big\}$$

We have

$$\int_{S_a(\theta' x)} dy_2 \dots dy_d = n^{-(d-1)/d},$$

and

$$\sum_{a_k=0,k=2,\dots,d} \int_{S_a(\theta' x)} g(A'y) dy_2 \dots dy_d = \int_{S(\theta' x)} g(A'y) dy_2 \dots dy_d.$$

If we restrict to all points X_j that have the same orthogonal projection on the hyperplane $\{y : y = \theta' x\}$ i.e $X_j^{(k)} = \frac{a_k}{n^{1/d}}$ (we call this set J_a), we have

$$\int_{S_a(\theta' x)} g(A'y) dy_2 \dots dy_d - \sum_{k=0,j\in J_a\cap I_k} \widetilde{W}_{nI_k}(\theta' x) g(X_j)$$
$$= \int_{S_a(\theta' x)} g(A'y) dy_2 \dots dy_d - n^{-(d-1)/d} \sum_{k=0,j\in J_a\cap I_k} n^{(d-1)/d} \widetilde{W}_{nI_k}(\theta' x) g(X_j)$$
$$= \int_{S_a(\theta' x)} \sum_{k=0,j\in J_a\cap I_k} n^{(d-1)/d} \widetilde{W}_{nI_k}(\theta' x) \times (g(A'y) - g(X_j)) dy_2 \dots dy_d.$$

We know that

$$\left|\widetilde{W}_{nI_k}(\theta' x)\right| |g(A'y) - g(X_j)| \le L \|A'y - X_j\|^\beta \le L \left(\sqrt{b^2 + (d-1)\left(\frac{1}{2n^{1/d}}\right)^2} \right)^\beta$$

because of the property

$$\widetilde{W}_{nI_k}(\theta' x) = 0 \text{ as soon as } |\theta'(x - X_j)| > b_n.$$

Moreover, the following result holds

$$\sum_{j=1}^n |\widetilde{W}_{nj}(\theta' x)| \le c_*.$$

Therefore, this entails the result.

7.5 Proof of Corollary 5.4

We consider $\inf_{\theta\in\Theta} \{I_2(\theta) + I_3(\theta)\}$. First, we look at

$$I_2(\theta) = \sum_{s=1}^6 I_{2,s}(\theta)$$

We know that $I_{2,1}(\theta) = \|g - f_\theta(\theta' \cdot)\|_2^2$ and

$$(C\varphi_n(\alpha_n))^2 \le I_{2,1}(\theta)$$

under (H_1).
We consider $I_{2,2}(\theta) = \left\| f_\theta(\theta' \cdot) - \mathbb{E}_g^n \left(\hat{f}_\theta(\theta' \cdot)\right)\right\|_2^2$.

Using lemma 5.5, we put

$$C_n^* = L \left(\sqrt{b^2 + (d-1)\left(\frac{1}{2n^{1/d}}\right)^2} \right)^\beta.$$

So

$$-(C_n^*)^2 \le I_{2,2}(\theta)$$

We consider $I_{2,3}(\theta) = \left\| \hat{f}_\theta(\theta' \cdot) - \mathbb{E}_g^n \left(\hat{f}_\theta(\theta' \cdot)\right)\right\|_2^2$
We obtain,

$$-(c_*)^2 \, (\sup_j |\varepsilon_j|)^2 \le I_{2,3}(\theta)$$

We consider now

$$I_{2,4}(\theta) = 2 \int_S (g(x) - f_\theta(\theta' x)) \left(f_\theta(\theta' x) - \mathbb{E}_g^n \left(\hat{f}_\theta(\theta' x)\right)\right) dx.$$

So

$$-2C_n^* \times \|g - f_\theta(\theta' \cdot)\|_2 \le I_{2,4}(\theta).$$

We consider

$$I_{2,5}(\theta) = 2 \int_S (g(x) - f_\theta(\theta' x)) \left(\mathbb{E}_g^n \left(\hat{f}_\theta(\theta' x)\right) - \hat{f}_\theta(\theta' x)\right) dx$$

Using lemmas we obtain,

$$-2c_* \sup_j |\varepsilon_j| \times \|g - f_\theta(\theta' \cdot)\|_2 \le I_{2,5}(\theta)$$

We consider now,

$$I_{2,6}(\theta) = 2 \int_S \left(f_\theta(\theta' x) - \mathbb{E}_g^n \left(\hat{f}_\theta(\theta' x)\right)\right) \times \left(\mathbb{E}_g^n \left(\hat{f}_\theta(\theta' x)\right) - \hat{f}_\theta(\theta' x)\right) dx$$

Using lemmas we obtain,

$$-2c_* C_n^* \sup_j |\varepsilon_j| \le I_{2,6}(\theta)$$

Now we look at $I_3(\theta)$.
We consider

$$I_{3,1}(\theta) = 2 \int_S \left(\hat{g}_n(x) - \mathbb{E}_g^n(\hat{g}_n(x))\right) (g(x) - f_\theta(\theta' x))\, dx$$

Then we have,

$$-2c^* \sup_j |\varepsilon_j| \times \|g - f_\theta(\theta' \cdot)\|_2 \le I_{3,1}(\theta)$$

We consider now

$$I_{3,2}(\theta) = 2 \int_S \left(\hat{g}_n(x) - \mathbb{E}_g^n(\hat{g}_n(x))\right) \times \left(f_\theta(\theta' x) - \mathbb{E}_g^n \left(\hat{f}_\theta(\theta' x)\right)\right) dx$$

We obtain,

$$-2c^*\, C_n^* \times \sup_j |\varepsilon_j| \leq I_{3,2}(\theta)$$

We consider

$$I_{3,3}(\theta) = 2\int_S \left(\hat{g}_n(x) - \mathbb{E}_g^n(\hat{g}_n(x))\right) \times \left(\mathbb{E}_g^n\left(\hat{f}_\theta(\theta' x)\right) - \hat{f}_\theta(\theta' x)\right) dx$$

$$-2c^*\, c_*\, (\sup_j |\varepsilon_j|)^2 \leq I_{3,3}(\theta)$$

We consider

$$I_{3,4}(\theta) = 2\int_S \left(\mathbb{E}_g^n(\hat{g}_n(x)) - g(x)\right)(g(x) - f_\theta(\theta' x))\, dx.$$

$$|I_{3,4}(\theta)| \leq 2\int_S \left|\mathbb{E}_g^n(\hat{g}_n(x)) - g(x)\right| |g(x) - f_\theta(\theta' x)|\, dx$$

$$|\mathbb{E}_g^n(\hat{g}_n(x)) - g(x)| \leq C_n^{**} = \left(\frac{Ld^{2/(2-\tau)}C}{m!}\right) h_n^\beta,$$

according to Proposition 5.1 point 3.

$$-2C_n^{**} \times \|g - f_\theta(\theta' \cdot)\|_2 \leq I_{3,4}(\theta)$$

We consider

$$I_{3,5}(\theta) = 2\int_S \left(\mathbb{E}_g^n(\hat{g}_n(x)) - g(x)\right) \times \left(f_\theta(\theta' x) - \mathbb{E}_g^n\left(\hat{f}_\theta(\theta' x)\right)\right) dx$$

and we obtain

$$-2C_n^* C_n^{**} \leq I_{3,5}(\theta)$$

We consider

$$I_{3,6}(\theta) = 2\int_S \left(\mathbb{E}_g^n(\hat{g}_n(x)) - g(x)\right) \times \left(\mathbb{E}_g^n\left(\hat{f}_\theta(\theta' x)\right) - \hat{f}_\theta(\theta' x)\right) dx$$

$$-2c_* C_n^{**} \sup_j |\varepsilon_j| \leq I_{3,6}(\theta)$$

We obtain the result

$$(C\varphi_n(\alpha_n))^2 - \left(2C_n^* + \alpha \sup_j |\varepsilon_j| + 2C_n^{**}\right) \inf_{\theta\in\Theta}\{\|g - f_\theta(\theta' \cdot)\|_2\} - 2C_n^* C_n^{**} - (C_n^*)^2 - \left((c_*)^2 + 2c^* c_*\right)(\sup_j |\varepsilon_j|)^2 - \left(2c^* C_n^{**} + \alpha\, C_n^*\right)\sup_j |\varepsilon_j| \leq \inf_{\theta\in\Theta}\{I_2(\theta) + I_3(\theta)\},$$

with $\alpha = 2(c^* + c_*)$.

8 Conclusion

In this paper, we have studied the problem of goodness-of-fit testing for the single-index models when the unknown regression function belongs to the d-dimensional isotropic Hölder space $\Sigma_d(\beta, L)$. We have indeed, constructed an asymptotically optimal test of level α_n reaching the minimax rate $\varphi_n(\alpha_n)$ for dimension $d = 3$ and $\frac{3}{4} \leq \beta < 1$. The minimax properties of our test allow it to distinguish the null hypothesis from the closest altenative. For $d \geq 4$ and $\beta \geq 1$, we have shown that no α_n-level test can achieve a rate better than $\varphi_n(\alpha_n)$. However the upper bound remains an open problem in this case. The results obtained have been established thanks to a large deviation inequality for degenerate U-statistics of order two.

REFERENCES

[1] Aït-Sahalia Y., Bickel P. J. and Stoker T. M., "Goodness-of-fit tests for kernel regression with an application to option implied volatilities," Journal of Econometrics, vol. 105, no. 2, pp. 363-412, 2001, DOI: 10.1016/S0304-4076(01)00091-4.

[2] Chiabrando F., "Risque avec normalisation aléatoire et test adaptatif dans le modèle additif." [PhD thesis, Aix-Marseille University], Retrieved from https://tel.archives-ouvertes.fr/tel-00348271, 2008.

[3] Chen, S. X. and Van Keilegom, I., "A goodness-of-fit test for parametric and semiparametric models in multiresponse regression.", Bernoulli, vol. 15, no. 4, pp. 955-976, 2009, DOI: 10.3150/09-BEJ208.

[4] Duan N. and Li K-C., "Slicing Regression: A Link-Free Regression Method," The Annals of Statistics, vol. 19, no. 2, pp. 505-530, 1991, https://www.jstor.org/stable/2242072.

[5] Fan Y. and Li Q., "Consistent Model Specification Tests: Omitted Variables and Semiparametric Functional Forms," Econometrica, vol. 64, no. 4, pp. 865-890, 1996, DOI: 10.2307/2171848.

[6] Gaïffa S. and Lecué G., "Optimal rates and adaptation in the single-index model using aggregation," Electronic Journal of Statistics , vol. 1, pp. 538-573, 2007, DOI: 10.1214/07-EJS077.

[7] Gayraud G. and Pouet C., "Minimax Testing composite null hypotheses in the discrete regression scheme," Mathematical Methods of Statistics, vol. 10, no. 4, pp. 375-394, 2001.

[8] Härdle W., Hall P. and Ichimura H., "Optimal Smoothing in Single-Index Models," The Annals of Statistics, vol. 21, no. 1, pp. 157-178, 1993, https://www.jstor.org/stable/3035585.

[9] Härdle W., Mammen E. and Proença I., "A bootstrap test for single index models," Statistics: a journal of theoretical and applied statistics, vol. 35, no. 4, pp. 427-451, 2001, DOI: 10.1080/02331880108802746

[10] Hoffmann M., "On estimating the diffusion coefficient: parametric versus nonparametric," Annales de l'Institut Henri Poincare (B) Probability and Statistics, vol. 37, no. 3, pp. 339-372, 2001, DOI: 10.1016/S0246-0203(00)01070-0.

[11] Hoffmann M., Lepski O. V., "Random rates in anisotropic regression (with a discussion and a rejoinder by the authors)," The Annals of Statistics, vol. 30, no. 2, pp. 325-396, 2002, DOI: 10.1214/aos/1021379858

[12] Hristache M., Judistsky A. and Spokoiny V. "Direct estimation of the index coefficient in a single index model," The Annals of Statistics, vol. 29, no. 3, pp. 595-623, 2001, DOI: 10.1214/aos/1009210682.

[13] Ichimura H., "Semiparametric least squares (SLS) and weighted SLS estimation of single-index models," Journal of Econometrics, vol. 58, no. 1-2, pp. 71-120, 1993, DOI: 10.1016/0304-4076(93)90114-K.

[14] Kulasekera K. B. and Lin W., "Error variance estimation for the single-index model," Australian & New Zealand Journal of Statistics, vol. 52, no. 2, pp. 201-219, 2010, DOI: 10.1111/j.1467-842X.2010.00575.x.

[15] Kulasekera K. B. and Lin W., "Identifiability of single-index models and additive-index models," Biometrika, vol. 94, no. 2, pp. 496-501, 2007, DOI: 10.1093/biomet/asm029.

[16] Lepski O. V., "How to improve the accuracy of estimation," Mathematical Methods of Statistics, vol. 8, pp. 441-486, 1999.

[17] Maistre S. and Patilea V., "Nonparametric model checks of single-index assumptions.", Statistica Sinica, vol. 29, no. 1, pp. 113-138, 2019, DOI: //doi.org/10.5705/ss.202015.0337.

[18] Powell J. L., Stock J. and Stoker T. M., "Semiparametric Estimation of Index Coefficients," Econometrica: Journal of the Econometric Society, vol. 57, no. 6, 1989, pp. 1403-30, DOI: 10.2307/1913713.

[19] Stone C.J, "Optimal Global Rates of Convergence for Nonparametric Regression," The Annals of Statistics, vol. 10, no. 4, pp. 1040-1053, 1982, DOI: 10.1214/aos/1176345969

[20] Stute W. and Zhu L.X. "Nonparametric checks for single-index models," Annals of Statistics, vol.33, no. 3, pp. 1048-1083, 2005, DOI: 10.1214/009053605000000020.

[21] Tsybakov A. B., "Nonparametric estimators," in Introduction to Nonparametric Estimation, Springer, 2009, pp. 1-72.

[22] Yodé A. F., "Asymptotically minimax test of independance," Mathematical Methods of Statistics, vol. 13, no. 2, 201-234, 2004, http://link.springer.com/journal/12004.

[23] Yodé A. F., "Multidimensional Nonparametric Density Estimates: Minimax Risk with Random Normalizing Factor," African Diaspora Journal of Mathematics. New Series, vol. 10, no. 2, pp. 27-57, 2010, https://projecteuclid.org:443/euclid.adjm/1291058599.

[24] Yodé A. F., "Adaptative minimax test of independance," Mathematical Methods of Statistics, vol. 20, no. 3, pp.246-368, 2011, DOI: 10.3103/S1066530711030069.

[25] Xia Y., Tong H.and Li W.K. "On extended partially linear single-index models," Biometrika, vol. 86, no. 4, pp. 831-842, 1999, DOI: 10.1093/biomet/86.4.831.

[26] Xia, Y., LI, W. K., Tong, H. and Zhang, D., "A goodness-of-fit test for single-index models (with discussion)," Statistica Sinica, vol. 14, no. 1, pp. 1-39, 2004, http://hdl.handle.net/10722/45362.