

Detailed proofs of Propositions 3.1 and 3.2

1 Proof of Proposition 3.1

NB: integration sets are generally omitted for integrals defined over a unit hypercube $[0, 1]^s$ with any $s \leq d$.

We first give four lemmas. The proof of Proposition 3.1 is presented in Section 1.2.

1.1 Technical lemmas

Let \mathbf{X}^1 and \mathbf{X}^2 two distinct points of a Latin hypercube of size n in $[0, 1]^d$. For any function f defined on $[0, 1]^d$, consider $Y^{1,LHS} = f(\mathbf{X}^1)$ and $Y^{2,LHS} = f(\mathbf{X}^2)$. In Theorem 1 in Stein (1987), Stein gives the following result

Theorem 1. *If f is a square integrable function then as n tends to $+\infty$, we have*

$$\text{Cov}(Y^{1,LHS}, Y^{2,LHS}) = -\frac{1}{n} \sum_{i=1}^d \sigma_i^2 + o(n^{-1}) \quad (1)$$

with $\sigma_i^2 = \text{Var}[\mathbb{E}[Y|X_i]]$, $i = 1, \dots, d$, $Y = f(X)$, X uniformly distributed on $[0, 1]^d$.

In this subsection, we prove an analogous result with more general settings and without the asymptotic assumption on n (see Lemma 4 below).

Notation and definitions For s and n in \mathbb{N}^* , define the partition of $[0, 1]^s$ in elementary hypercubes of side $1/n$,

$$\mathcal{Q}_s(n) = \left\{ Q \subseteq [0, 1]^s \mid Q = \prod_{i=1}^s [\alpha_i, \beta_i), \alpha_i \in \left\{ 0, \frac{1}{n}, \dots, \frac{n-1}{n} \right\}, \beta_i = \alpha_i + \frac{1}{n} \right\}. \quad (2)$$

For any square integrable function g defined on $[0, 1]^s$, $s \leq d$, define the sequence with general term

$$u_n(g) = n^s \sum_{Q \in \mathcal{Q}_s(n)} \left(\int_Q g(\mathbf{x}) d\mathbf{x} \right)^2, \quad n \in \mathbb{N}. \quad (3)$$

Outline The first lemma is the analogous result for Lebesgue integrability of a result given in Equation (A.4) in Stein (1987) for Riemann integrability. The second lemma gives an important inequality which allows to work without asymptotic assumption on n . The third one consists in simplifying integrals under LHS using the ANOVA decomposition. Lemma 4 provides the expected inequalities.

Lemma 1. *Let $s \in \mathbb{N}^*$. If g is a continuous function on $[0, 1]^s$, the sequence $(u_n(g))$ converges to $\int g^2(\mathbf{x}) d\mathbf{x}$ as n tends to $+\infty$.*

Proof. Noting that

$$u_n(g) = \int g_n^2(\mathbf{x}) d\mathbf{x} \quad (4)$$

where

$$\forall \mathbf{x} \in [0, 1]^s, g_n^2(\mathbf{x}) = \sum_{Q \in \mathcal{Q}_s(n)} \left(n^s \int_Q g(\mathbf{y}) d\mathbf{y} \right)^2 \mathbf{1}_Q(\mathbf{x}), \quad (5)$$

We can rewrite g_n as $g_n^2(\mathbf{x}) = \left(n^s \int_{Q_{\mathbf{x}}} g(\mathbf{y}) d\mathbf{y} \right)^2$, where $Q_{\mathbf{x}}$ is the set Q in $\mathcal{Q}_s(n)$ containing \mathbf{x} .

We now prove that $\|g_n - g\|_{\infty} \xrightarrow{n \rightarrow +\infty} 0$. Indeed, as g is continuous on a compact, it is uniformly continuous on this compact. Thus $g_n(x) - g(x) = n^s \int (g(y) - g(x)) \mathbf{1}_Q(\mathbf{x}) dy$ converges uniformly to zero. From this latter convergence, we also deduce that for n large enough, $g_n^2 \leq g^2 + 1$. We then conclude by applying the dominated convergence theorem, as $\int_{[0,1]^s} (g^2(x) + 1) dx < \infty$. \square

Lemma 2. *The sequence $(u_n(g))$ is dominated by $\int g^2(\mathbf{x}) d\mathbf{x}$.*

Proof. Let $n \in \mathbb{N}^*$, the result is proved by showing that the sequence of general term $v_k(g) = u_{2^k n}(g)$ is increasing. In this case, by Lemma 1, we have $\lim v_k(g) = \int g^2(\mathbf{x}) d\mathbf{x}$, and since v_k is increasing, all the terms of this sequence are dominated by $\int g^2(\mathbf{x}) d\mathbf{x}$, hence $v_0(g) = u_n(g) \leq \int g^2(\mathbf{x}) d\mathbf{x}$. To prove that the sequence $(v_k(g))$ is increasing, note that

$$\begin{aligned} v_{k+1}(g) &= (2^{k+1}n)^s \sum_{Q \in \mathcal{Q}_s(2^{k+1}n)} \left(\int_Q g(\mathbf{x}) d\mathbf{x} \right)^2 \\ &= (2^k n)^s \sum_{Q \in \mathcal{Q}_s(2^k n)} \left(2^s \sum_{P \in \mathcal{P}(Q, 2^{k+1}n)} \left(\int_P g(\mathbf{x}) d\mathbf{x} \right)^2 \right) \end{aligned} \quad (6)$$

where $\mathcal{P}(Q, 2^{k+1}n) = \mathcal{Q}(2^{k+1}n) \cap Q$. Then by Jensen inequality, we have

$$2^s \sum_{P \in \mathcal{P}(Q, 2^{k+1}n)} \left(\int_P g(\mathbf{x}) d\mathbf{x} \right)^2 \geq \left(\int_Q g(\mathbf{x}) d\mathbf{x} \right)^2 \quad (7)$$

and we conclude that $v_{k+1}(g) \geq v_k(g)$. \square

For $0 \leq x_1, x_2 \leq 1$ define

$$r_n(x_1, x_2) = \begin{cases} 1 & \text{if } \lfloor nx_1 \rfloor = \lfloor nx_2 \rfloor \\ 0 & \text{otherwise,} \end{cases} \quad (8)$$

where $\lfloor \cdot \rfloor$ is the floor function. We now end with the following result

Lemma 3. *Let \mathbf{v} be a subset of $\{1, \dots, d\}$, we have*

$$\int f(\mathbf{x}_1) f(\mathbf{x}_2) \prod_{i \in \mathbf{v}} r_n(x_{1i}, x_{2i}) d\mathbf{x}_1 d\mathbf{x}_2 = \int \sum_{\mathbf{w} \subseteq \mathbf{v}} f_{\mathbf{w}}(\mathbf{x}_{1\mathbf{w}}) f_{\mathbf{w}}(\mathbf{x}_{2\mathbf{w}}) \prod_{i \in \mathbf{v}} r_n(x_{1i}, x_{2i}) d\mathbf{x}_1 d\mathbf{x}_2. \quad (9)$$

Proof. By the ANOVA decomposition, we have

$$\int f(\mathbf{x}_1)f(\mathbf{x}_2)\prod_{i\in\mathbf{v}}r_n(x_{1i},x_{2i})d\mathbf{x}_1d\mathbf{x}_2=\int\sum_{\mathfrak{w}_1\subseteq\{1,\dots,d\}}\sum_{\mathfrak{w}_2\subseteq\{1,\dots,d\}}f_{\mathfrak{w}_1}(\mathbf{x}_{1\mathfrak{w}_1})f_{\mathfrak{w}_2}(\mathbf{x}_{2\mathfrak{w}_2})\prod_{i\in\mathbf{v}}r_n(x_{1i},x_{2i})d\mathbf{x}_1d\mathbf{x}_2. \quad (10)$$

Then note that a certain number of terms in the member on the right-hand side vanishes. If $(\mathfrak{w}_1\cap\mathbf{v}^c)\cup(\mathfrak{w}_2\cap\mathbf{v}^c)\neq\emptyset$ then suppose without any loss of generality that there exists $k\in\mathfrak{w}_1\setminus\mathbf{v}$. We have

$$\int f_{\mathfrak{w}_1}(\mathbf{x}_{1\mathfrak{w}_1})f_{\mathfrak{w}_2}(\mathbf{x}_{2\mathfrak{w}_2})\prod_{i\in\mathbf{v}}r_n(x_{1i},x_{2i})d\mathbf{x}_1d\mathbf{x}_2=\int\underbrace{\left(\int f_{\mathfrak{w}_1}(\mathbf{x}_{1\mathfrak{w}_1})dx_{1k}\right)}_{I_1}f_{\mathfrak{w}_2}(\mathbf{x}_{2\mathfrak{w}_2})\prod_{i\in\mathbf{v}}r_n(x_{1i},x_{2i})d\mathbf{x}_{1\setminus\{k\}}d\mathbf{x}_2 \quad (11)$$

and note that, by a basic property of the ANOVA decomposition, $I_1=0$. If $(\mathfrak{w}_1\cap\mathbf{v}^c)\cup(\mathfrak{w}_2\cap\mathbf{v}^c)=\emptyset$ and $\mathfrak{w}_1\neq\mathfrak{w}_2$, then suppose without any loss of generality that there exists $k\in\mathfrak{w}_1\setminus\mathfrak{w}_2$. In this case, we have

$$\begin{aligned} &\int f_{\mathfrak{w}_1}(\mathbf{x}_{1\mathfrak{w}_1})f_{\mathfrak{w}_2}(\mathbf{x}_{2\mathfrak{w}_2})\prod_{i\in\mathbf{v}}r_n(x_{1i},x_{2i})d\mathbf{x}_1d\mathbf{x}_2 \\ &= \int\underbrace{\left(\int f_{\mathfrak{w}_1}(\mathbf{x}_{1\mathfrak{w}_1})r_n(x_{1k},x_{2k})dx_{1k}dx_{2k}\right)}_{I_2}f_{\mathfrak{w}_2}(\mathbf{x}_{2\mathfrak{w}_2})\left(\prod_{i\in\mathbf{v}\setminus\{k\}}r_n(x_{1i},x_{2i})\right)d\mathbf{x}_{1\setminus\{k\}}d\mathbf{x}_{2\setminus\{k\}} \quad (12) \end{aligned}$$

and note that by the definition of r_n , we have

$$\int f_{\mathfrak{w}_1}(\mathbf{x}_{1\mathfrak{w}_1})r_n(x_{1k},x_{2k})dx_{1k}dx_{2k}=\int f_{\mathfrak{w}_1}(\mathbf{x}_{1\mathfrak{w}_1})dx_{1k} \quad (13)$$

and thus $I_2=0$. The conclusion of the lemma follows. \square

Let \mathbf{u} be a non-empty subset of $\{1,\dots,d\}$ and consider $(\mathbf{X}^j)_j\sim\mathcal{LH}(n,d)$ and $(\mathbf{Z}^j)_j\sim\mathcal{LH}(n,d)$. For any function f defined on $[0,1]^d$, consider

$$Y^{1,LHS}=f(\mathbf{X}^1)\text{ and }Y_{\mathbf{u}}^{2,LHS}=f(\mathbf{X}_{\mathbf{u}}^2,\mathbf{Z}_{\mathbf{u}^c}^2). \quad (14)$$

We have the following result

Lemma 4. *If f is a square integrable function then we have*

$$-\sum_{\substack{\emptyset\neq\mathfrak{w}\subseteq\mathbf{u} \\ |\mathfrak{w}|\text{ odd}}}\frac{\sigma_{\mathfrak{w}}^2}{(n-1)^{|\mathfrak{w}|}}\leq\text{Cov}(Y^{1,LHS},Y_{\mathbf{u}}^{2,LHS})\leq\sum_{\substack{\emptyset\neq\mathfrak{w}\subseteq\mathbf{u} \\ |\mathfrak{w}|\text{ even}}}\frac{\sigma_{\mathfrak{w}}^2}{(n-1)^{|\mathfrak{w}|}}. \quad (15)$$

Proof. Recall that for $0\leq x_1, x_2\leq 1$,

$$r_n(x_1,x_2)=\begin{cases} 1 & \text{if } \lfloor nx_1 \rfloor = \lfloor nx_2 \rfloor \\ 0 & \text{otherwise,} \end{cases} \quad (16)$$

where $\lfloor\cdot\rfloor$ is the floor function. For $\mathbf{x}_1=(x_{11},\dots,x_{1d})$ in $[0,1]^d$, define $\mathbf{x}_{1\mathbf{v}}=(x_{1i_1},\dots,x_{1i_{|\mathbf{v}|}})$ where $\mathbf{v}=\{i_1,\dots,i_{|\mathbf{v}|}\}$. Due to the joint density of $(\mathbf{X}_{\mathbf{u}}^1,\mathbf{X}_{\mathbf{u}^c}^2)$ under LHS — see McKay et al. (1979) or Stein (1987) — and by Lemma 3, we have

$$\text{Cov}(Y^{1,LHS},Y_{\mathbf{u}}^{2,LHS})+\left(\int f(\mathbf{x})d\mathbf{x}\right)^2=\int f(\mathbf{x}_1)f(\mathbf{x}_2)\left(\frac{n}{n-1}\right)^{|\mathbf{u}|}\prod_{i\in\mathbf{u}}(1-r_n(x_{1i},x_{2i}))d\mathbf{x}_1d\mathbf{x}_2$$

$$\begin{aligned}
&= \left(\frac{n}{n-1}\right)^{|\mathbf{u}|} \sum_{\mathbf{v} \subseteq \mathbf{u}} (-1)^{|\mathbf{v}|} \int f(\mathbf{x}_1) f(\mathbf{x}_2) \prod_{i \in \mathbf{v}} r_n(x_{1i}, x_{2i}) d\mathbf{x}_1 d\mathbf{x}_2 \\
&= \left(\frac{n}{n-1}\right)^{|\mathbf{u}|} \sum_{\mathbf{v} \subseteq \mathbf{u}} (-1)^{|\mathbf{v}|} \int \sum_{\mathbf{w} \subseteq \mathbf{v}} f_{\mathbf{w}}(\mathbf{x}_{1\mathbf{w}}) f_{\mathbf{w}}(\mathbf{x}_{2\mathbf{w}}) \prod_{i \in \mathbf{v}} r_n(x_{1i}, x_{2i}) d\mathbf{x}_1 d\mathbf{x}_2 \\
&= \left(\frac{n}{n-1}\right)^{|\mathbf{u}|} \sum_{\mathbf{v} \subseteq \mathbf{u}} (-1)^{|\mathbf{v}|} \sum_{\mathbf{w} \subseteq \mathbf{v}} \left(\frac{1}{n}\right)^{|\mathbf{v}|-|\mathbf{w}|} \int f_{\mathbf{w}}(\mathbf{x}_{1\mathbf{w}}) f_{\mathbf{w}}(\mathbf{x}_{2\mathbf{w}}) \prod_{i \in \mathbf{w}} r_n(x_{1i}, x_{2i}) d\mathbf{x}_{1\mathbf{w}} d\mathbf{x}_{2\mathbf{w}}. \quad (17)
\end{aligned}$$

Then note that for any function of \mathbf{w} denoted by $A(\mathbf{w})$, we have

$$\begin{aligned}
\sum_{\mathbf{v} \subseteq \mathbf{u}} (-1)^{|\mathbf{v}|} \sum_{\mathbf{w} \subseteq \mathbf{v}} \left(\frac{1}{n}\right)^{|\mathbf{v}|-|\mathbf{w}|} A(\mathbf{w}) &= \sum_{\mathbf{v} \subseteq \mathbf{u}} \left(-\frac{1}{n}\right)^{|\mathbf{v}|} \sum_{\mathbf{w} \subseteq \mathbf{v}} \left(\frac{1}{n}\right)^{-|\mathbf{w}|} A(\mathbf{w}) \\
&= \sum_{\mathbf{w} \subseteq \mathbf{u}} \left(\sum_{k=0}^{|\mathbf{u}|-|\mathbf{w}|} \binom{|\mathbf{u}|-|\mathbf{w}|}{k} \left(-\frac{1}{n}\right)^{k+|\mathbf{w}|} \right) \left(\frac{1}{n}\right)^{-|\mathbf{w}|} A(\mathbf{w}) \\
&= \sum_{\mathbf{w} \subseteq \mathbf{u}} \left(\frac{n-1}{n}\right)^{|\mathbf{u}|-|\mathbf{w}|} (-1)^{|\mathbf{w}|} A(\mathbf{w}) \quad (18)
\end{aligned}$$

Hence, we deduce that

$$\text{Cov}(Y^{1,LHS}, Y_u^{2,LHS}) = \sum_{\substack{\mathbf{w} \subseteq \mathbf{u} \\ \mathbf{w} \neq \emptyset}} \left(\frac{n}{n-1}\right)^{|\mathbf{w}|} (-1)^{|\mathbf{w}|} \int f_{\mathbf{w}}(\mathbf{x}_{1\mathbf{w}}) f_{\mathbf{w}}(\mathbf{x}_{2\mathbf{w}}) \prod_{i \in \mathbf{w}} r_n(x_{1i}, x_{2i}) d\mathbf{x}_{1\mathbf{w}} d\mathbf{x}_{2\mathbf{w}}. \quad (19)$$

Finally by the definition of r_n , we have

$$0 \leq \int f_{\mathbf{w}}(\mathbf{x}_{1\mathbf{w}}) f_{\mathbf{w}}(\mathbf{x}_{2\mathbf{w}}) \prod_{i \in \mathbf{w}} r_n(x_{1i}, x_{2i}) d\mathbf{x}_{1\mathbf{w}} d\mathbf{x}_{2\mathbf{w}} \leq \sum_{Q \in \mathcal{Q}_{|\mathbf{w}|}(n)} \left(\int_Q f_{\mathbf{w}}(\mathbf{x}_{1\mathbf{w}}) d\mathbf{x}_{1\mathbf{w}} \right)^2 \quad (20)$$

and by Lemma 2, this gives

$$0 \leq \int f_{\mathbf{w}}(\mathbf{x}_{1\mathbf{w}}) f_{\mathbf{w}}(\mathbf{x}_{2\mathbf{w}}) \prod_{i \in \mathbf{w}} r_n(x_{1i}, x_{2i}) d\mathbf{x}_{1\mathbf{w}} d\mathbf{x}_{2\mathbf{w}} \leq \frac{\sigma_{\mathbf{w}}^2}{n^{|\mathbf{w}|}}. \quad (21)$$

The latter inequalities and (19) lead to Lemma 4. \square

Remark 1. Note that if $\mathbf{u} = \{1, \dots, d\}$, the resulting inequalities are

$$-\sum_{\substack{\mathbf{w} \neq \emptyset \\ |\mathbf{w}| \text{ odd}}} \frac{\sigma_{\mathbf{w}}^2}{(n-1)^{|\mathbf{w}|}} \leq \text{Cov}(Y^{1,LHS}, Y^{2,LHS}) \leq \sum_{\substack{\mathbf{w} \neq \emptyset \\ |\mathbf{w}| \text{ even}}} \frac{\sigma_{\mathbf{w}}^2}{(n-1)^{|\mathbf{w}|}}. \quad (22)$$

This consists of a non-asymptotic equivalent of the theorem due to Stein presented at the beginning of this section.

1.2 Proof of Proposition 3.1

(i) This is a consequence of the strong law of large numbers for LHS given in Theorem 3 in Loh (1996).

(ii) The proof consists in translating the original proof given for simple random sampling — see Proposition 2.2 in Janon et al. (2013) — for LHS. Concerning $\tilde{\underline{S}}_{u,n}^{LHS}$, it is easy to show that

$$\tilde{\underline{S}}_{u,n}^{LHS} = \Phi(\bar{\mathbf{V}}_n) \quad (23)$$

where $\bar{\mathbf{V}}_n = \frac{1}{n} \sum_{j=1}^n \mathbf{V}_j$ and \mathbf{V}_j equal to

$$\left((Y^{j,LHS} - \mathbb{E}[Y]) (Y_u^{j,LHS} - \mathbb{E}[Y]), Y^{j,LHS} - \mathbb{E}[Y], Y_u^{j,LHS} - \mathbb{E}[Y], (Y^{j,LHS} - \mathbb{E}[Y])^2 \right)^T$$

and with $\Phi(x, y, z, t) = \frac{x-yz}{t-y^2}$. Then we deduce from Theorem 2 in Loh (1996) that

$$\sqrt{n}(\bar{\mathbf{V}}_n - \mu) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_4(0, \Gamma) \quad (24)$$

where $\mu = (\tau_u^2, 0, 0, \sigma^2)^T$ and Γ is the covariance matrix of $\mathbf{R}_1 = \mathbf{V}_1 - \mathbf{A}_1$ — see details in Eq. (3) in Loh (1996) — defined by

$$\forall i \in \{1, \dots, 4\}, A_{1i} \text{ is the additive part — see (2) in the main document — of } V_{1i}. \quad (25)$$

Thus the Delta method — see Theorem 3.1 in Van der Vaart (1998) — gives

$$\sqrt{n}(\tilde{\underline{S}}_{u,n}^{LHS} - \underline{S}_u) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_1(0, g^T \Gamma g) \quad (26)$$

where $g = \nabla \Phi(\mu)$. Developing the term $g^T \Gamma g$ does not seem to provide any useful information. However, denoting σ_{LHS}^2 this term, and σ_{IID}^2 the analogous quantity in the CLT for simple random sampling, we can show that $\sigma_{LHS}^2 \leq \sigma_{IID}^2$. Indeed we first note that, for simple random sampling, the variance given in Janon et al. (2013) reads

$$\sigma_{IID}^2 = \frac{\text{Var}[V_{11} - \underline{S}_u V_{14}]}{\sigma^2} \quad (27)$$

and for LHS, it is easy to show that

$$\sigma_{LHS}^2 = \frac{\text{Var}[R_{11} - \underline{S}_u R_{14}]}{\sigma^2}. \quad (28)$$

Hence

$$\sigma_{IID}^2 = \sigma_{LHS}^2 + \frac{\text{Var}[A_{11} - \underline{S}_u A_{14}]}{\sigma^2} \quad (29)$$

and the conclusion of (ii) for $\tilde{\underline{S}}_{u,n}^{LHS}$ follows. Concerning $\hat{\underline{S}}_{u,n}^{LHS}$, the proof follows the same lines — see Proof of (10) in Janon et al. (2013) for details.

(iii) First we have

$$\begin{aligned} \mathbb{E}[\tilde{\tau}_{u,n}^{2,LHS}] &= \frac{n-1}{n^2} \sum_{j=1}^n \mathbb{E}[Y^{j,LHS} Y_u^{j,LHS}] - \frac{1}{n^2} \sum_{j=1}^n \sum_{\substack{l=1 \\ l \neq j}}^n \mathbb{E}[Y^{j,LHS} Y_u^{l,LHS}] \\ &= \frac{n-1}{n} \left(\mathbb{E}[Y]^2 + \tau_u^2 \right) - \frac{n-1}{n} \left(\text{Cov}(Y^{1,LHS}, Y_u^{2,LHS}) + \mathbb{E}[Y]^2 \right) \end{aligned} \quad (30)$$

and thanks to Lemma 4, this gives

$$\mathbb{E}[\widehat{\tau}_{u,n}^{2,LHS}] = \tau_u^2 + B_{n,1} \quad (31)$$

with

$$-\frac{1}{n-1}\tau_u^2 \leq B_{n,1} \leq 0. \quad (32)$$

Concerning $\widehat{\sigma}_n^{2,LHS}$, we have

$$\begin{aligned} \mathbb{E}[\widehat{\sigma}_{\bar{u},n}^{2,LHS}] &= \frac{n-1}{n^2} \sum_{j=1}^n \mathbb{E}[(Y^{j,LHS})^2] - \frac{1}{n^2} \sum_{j=1}^n \sum_{\substack{l=1 \\ l \neq j}}^n \mathbb{E}[Y^{j,LHS} Y^{l,LHS}] \\ &= \frac{n-1}{n} \mathbb{E}[Y^2] - \frac{n-1}{n} \left(\text{Cov}(Y^{1,LHS}, Y^{2,LHS}) + \mathbb{E}[Y]^2 \right) \end{aligned} \quad (33)$$

and noting that

$$\text{Cov}(Y^{1,LHS}, Y^{2,LHS}) = \text{Cov}(Y^{1,LHS}, Y_{\{1,\dots,d\}}^{2,LHS}) \quad (34)$$

we conclude that

$$\mathbb{E}[\widehat{\sigma}_n^{2,LHS}] = \sigma^2 + B_{n,2} \quad (35)$$

with

$$-\frac{1}{n-1}\sigma^2 \leq B_{n,2} \leq 0. \quad (36)$$

As for $\widehat{\tau}_u^{2,LHS}$ and $\widehat{\sigma}_n^{2,LHS}$, we have

$$\begin{aligned} &\mathbb{E} \left[\left(\frac{1}{n} \sum_{j=1}^n \frac{Y^{j,LHS} + Y_u^{j,LHS}}{2} \right)^2 \right] \\ &= \frac{1}{4n} \mathbb{E}[(Y^{1,LHS} + Y_u^{1,LHS})^2] + \frac{1}{4n^2} \sum_{j=1}^n \sum_{\substack{l=1 \\ l \neq j}}^n \mathbb{E}[(Y^{j,LHS} + Y_u^{j,LHS})(Y^{l,LHS} + Y_u^{l,LHS})] \\ &= \frac{1}{2n} \left(\mathbb{E}[Y^2] + \tau_u^2 + \mathbb{E}[Y]^2 \right) + \frac{n-1}{n} \mathbb{E}[Y]^2 + \frac{n-1}{2n} \left(\text{Cov}(Y^{1,LHS}, Y^{2,LHS}) + \text{Cov}(Y^{1,LHS}, Y_u^{2,LHS}) \right) \\ &= \frac{1}{2n} \left(\sigma^2 + \tau_u^2 \right) + \mathbb{E}[Y]^2 + \frac{n-1}{2n} \left(\text{Cov}(Y^{1,LHS}, Y^{2,LHS}) + \text{Cov}(Y^{1,LHS}, Y_u^{2,LHS}) \right). \end{aligned} \quad (37)$$

Then it is easy to conclude that

$$\mathbb{E}[\widehat{\tau}_{u,n}^{2,LHS}] = \tau_u^2 + B_{n,3} \quad (38)$$

$$\mathbb{E}[\widehat{\sigma}_n^{2,LHS}] = \sigma^2 + B_{n,3} \quad (39)$$

with

$$-\frac{1}{2(n-1)}(\sigma^2 + \tau_u^2) \leq B_{n,3} \leq 0. \quad (40)$$

2 Proof of Proposition 3.2

We first give three lemmas. The proof of Proposition 3.2 is presented in Section 2.2.

2.1 Technical lemmas

Lemma 5. *Let $d \in \mathbb{N}^*$, if $n \geq \frac{d^2}{2}$ then*

$$\left(1 + \frac{1}{n}\right)^d - 1 \leq \frac{d+1}{n}. \quad (41)$$

Proof. If $d = 1$, the result is obvious. Otherwise, for any $x > 0$, consider the function g_d defined by

$$g_d(x) = \left(1 + \frac{1}{x}\right)^d - 1 - \frac{d+1}{x}. \quad (42)$$

We show that

(a) if there exists $x_0 > 0$ such that $g_d(x_0) \leq 0$ then for all $x \geq x_0$, $g_d(x) \leq 0$

(b) $g_d(d^2/2) \leq 0$

and the conclusion follows. Concerning (a) note that

$$g_d(x) = 1 + \frac{d}{x} + O(x^{-2}) - 1 - \frac{d}{x} - \frac{1}{x} = -\frac{1}{x} + O(x^{-2}) \quad (43)$$

and then that g_d is negative as x tends to $+\infty$. Moreover for any $d > 1$, g_d is first decreasing and then increasing. Indeed, we have

$$g'_d(x) = -\frac{d}{x^2} \left(1 + \frac{1}{x}\right)^{d-1} + \frac{d+1}{x^2} \quad (44)$$

and we deduce that $g'_d(x_0) = 0$ with

$$x_0 = \frac{1}{\left(\frac{d+1}{d}\right)^{1/(d-1)} - 1} > 0 \quad (45)$$

and is negative on the left-hand side and positive on the right-hand side. The conclusion of (a) follows. Concerning (b), it is easy to check that it is true for $d = 1$ and 2, and for $d \geq 3$ we have

$$\begin{aligned} g_d\left(\frac{d^2}{2}\right) &= \sum_{k=0}^d \binom{d}{k} \left(\frac{2}{d^2}\right)^k - 1 - \frac{2}{d} - \frac{2}{d^2} \\ &= -\frac{2}{d^3} + \sum_{k=3}^d \binom{d}{k} \left(\frac{2}{d^2}\right)^k \\ &\leq -\frac{2}{d^3} + \sum_{k=3}^d \frac{1}{k!} \left(\frac{2}{d}\right)^k \\ &\leq -\frac{2}{d^3} + \frac{1}{d^3} + \frac{1}{3d^3} + \frac{2}{3} \sum_{k=4}^d \frac{1}{d^k} \\ &\leq -\frac{2}{d^3} + \sum_{k=3}^d \frac{1}{d^k} + \frac{1}{3d^3} \left(1 - \sum_{k=1}^{d-3} \frac{1}{d^k}\right) \\ &\leq -\frac{2}{d^3} + \sum_{k=3}^d \frac{1}{d^k} \\ &\leq -\frac{2}{d^3} + \frac{2}{d^3} \end{aligned} \quad (46)$$

and the conclusion follows. \square

Let \mathbf{u} be a non-empty subset of $\{1, \dots, d\}$ and consider $(\mathbf{X}^j)_j \sim \mathcal{LH}(n, d)$. We have the following result

Lemma 6. *If f is a square integrable function, we have*

$$\mathbb{E}[f(\mathbf{X}^1)f(\mathbf{X}_{\mathbf{u}}^1, \mathbf{X}_{\mathbf{u}^c}^2)] = \mathbb{E}[Y]^2 + \tau_{\mathbf{u}}^2 + B_{\mathbf{u},n} \quad (47)$$

where

$$-\mathbb{E}[Y^2] \sum_{\substack{\emptyset \neq \mathbf{v} \subseteq \mathbf{u}^c \\ |\mathbf{v}| \text{ odd}}} \frac{1}{(n-1)^{|\mathbf{v}|}} \leq B_{\mathbf{u},n} \leq \mathbb{E}[Y^2] \sum_{\substack{\emptyset \neq \mathbf{v} \subseteq \mathbf{u}^c \\ |\mathbf{v}| \text{ even}}} \frac{1}{(n-1)^{|\mathbf{v}|}} . \quad (48)$$

Proof. First due to the joint density of $(\mathbf{X}_{\mathbf{u}}^1, \mathbf{X}_{\mathbf{u}^c}^2)$ under LHS — see McKay et al. (1979) or Stein (1987) — we have

$$\begin{aligned} \mathbb{E}[f(\mathbf{X}^1)f(\mathbf{X}_{\mathbf{u}}^1, \mathbf{X}_{\mathbf{u}^c}^2)] &= \int f(\mathbf{x}, \mathbf{x}_1)f(\mathbf{x}, \mathbf{x}_2) \left(\frac{n}{n-1}\right)^{d-|\mathbf{u}|} \prod_{i \in \mathbf{u}^c} (1 - r_n(x_{1i}, x_{2i})) d\mathbf{x} d\mathbf{x}_1 d\mathbf{x}_2 \\ &= \left(\frac{n}{n-1}\right)^{d-|\mathbf{u}|} \int \underbrace{\left(\sum_{\mathbf{v} \subseteq \mathbf{u}^c} (-1)^{|\mathbf{v}|} \int f(\mathbf{x}, \mathbf{x}_1)f(\mathbf{x}, \mathbf{x}_2) \prod_{i \in \mathbf{v}} r_n(x_{1i}, x_{2i}) d\mathbf{x}_1 d\mathbf{x}_2 \right)}_{I(\mathbf{x})} d\mathbf{x}. \end{aligned} \quad (49)$$

We now denote $f_{\mathbf{x}} : \mathbf{y} \mapsto f(\mathbf{x}, \mathbf{y})$ and then by (17) and (18) we have

$$\begin{aligned} I(\mathbf{x}) &= \sum_{\mathbf{v} \subseteq \mathbf{u}^c} (-1)^{|\mathbf{v}|} \int f_{\mathbf{x}}(\mathbf{x}_1)f_{\mathbf{x}}(\mathbf{x}_2) \prod_{i \in \mathbf{v}} r_n(x_{1i}, x_{2i}) d\mathbf{x}_1 d\mathbf{x}_2 \\ &= \sum_{\mathbf{v} \subseteq \mathbf{u}^c} (-1)^{|\mathbf{v}|} \sum_{\mathbf{w} \subseteq \mathbf{v}} \left(\frac{1}{n}\right)^{|\mathbf{v}|-|\mathbf{w}|} \int f_{\mathbf{x},\mathbf{w}}(\mathbf{x}_{1\mathbf{w}})f_{\mathbf{x},\mathbf{w}}(\mathbf{x}_{2\mathbf{w}}) \prod_{i \in \mathbf{w}} r_n(x_{1i}, x_{2i}) d\mathbf{x}_{1\mathbf{w}} d\mathbf{x}_{2\mathbf{w}} \\ &= \sum_{\mathbf{w} \subseteq \mathbf{u}^c} (-1)^{|\mathbf{w}|} \left(\frac{n-1}{n}\right)^{d-|\mathbf{u}|-|\mathbf{w}|} \int f_{\mathbf{x},\mathbf{w}}(\mathbf{x}_{1\mathbf{w}})f_{\mathbf{x},\mathbf{w}}(\mathbf{x}_{2\mathbf{w}}) \prod_{i \in \mathbf{w}} r_n(x_{1i}, x_{2i}) d\mathbf{x}_{1\mathbf{w}} d\mathbf{x}_{2\mathbf{w}} . \end{aligned} \quad (50)$$

Hence by (21) we have for all $\mathbf{w} \neq \emptyset$,

$$0 \leq \int f_{\mathbf{x},\mathbf{w}}(\mathbf{x}_{1\mathbf{w}})f_{\mathbf{x},\mathbf{w}}(\mathbf{x}_{2\mathbf{w}}) \prod_{i \in \mathbf{w}} r_n(x_{1i}, x_{2i}) d\mathbf{x}_{1\mathbf{w}} d\mathbf{x}_{2\mathbf{w}} \leq \frac{\int f_{\mathbf{x},\mathbf{w}}^2(\mathbf{x}_{1\mathbf{w}}) d\mathbf{x}_{1\mathbf{w}}}{n^{|\mathbf{w}|}} \leq \frac{\int f_{\mathbf{x}}^2(\mathbf{x}_1) d\mathbf{x}_1}{n^{|\mathbf{w}|}} \quad (51)$$

and note that

$$\int f_{\mathbf{x},\emptyset}(\mathbf{x}_{1\emptyset})f_{\mathbf{x},\emptyset}(\mathbf{x}_{2\emptyset}) \prod_{i \in \emptyset} r_n(x_{1i}, x_{2i}) d\mathbf{x}_{1\emptyset} d\mathbf{x}_{2\emptyset} = f_{\mathbf{x},\emptyset}^2 . \quad (52)$$

Finally, note that

$$\int f_{\mathbf{x},\emptyset}^2 d\mathbf{x} = \tau_{\mathbf{u}}^2 + \mathbb{E}[Y]^2 \quad (53)$$

and

$$\int \int f_{\mathbf{x}}^2(\mathbf{x}_1) d\mathbf{x}_1 d\mathbf{x} = \mathbb{E}[Y^2] \quad (54)$$

and conclude that

$$\mathbb{E}[f(\mathbf{X}^1)f(\mathbf{X}_u^1, \mathbf{X}_{u^c}^2)] = \left(\frac{n}{n-1}\right)^{d-|\mathbf{u}|} \int I(\mathbf{x})d\mathbf{x} = \tau_u^2 + \mathbb{E}[Y]^2 + B_{u,n} \quad (55)$$

with

$$-\mathbb{E}[Y^2] \sum_{\substack{\emptyset \neq \mathbf{v} \subseteq u^c \\ |\mathbf{v}| \text{ odd}}} \frac{1}{(n-1)^{|\mathbf{v}|}} \leq B_{u,n} \leq \mathbb{E}[Y^2] \sum_{\substack{\emptyset \neq \mathbf{v} \subseteq u^c \\ |\mathbf{v}| \text{ even}}} \frac{1}{(n-1)^{|\mathbf{v}|}} . \quad (56)$$

□

Lemma 7. *The inequalities in Equation (48) imply that*

$$\left| \sum_{\mathbf{w} \subseteq u^c} \left(\frac{1}{n}\right)^{|\mathbf{w}|} B_{u \cup \mathbf{w}, n} \right| \leq \left(\frac{d-|\mathbf{u}|+1}{n} + 1\right) \left(\frac{d-|\mathbf{u}|+1}{n-1}\right) \mathbb{E}[Y^2] . \quad (57)$$

Proof. By (48), we have

$$\begin{aligned} \sum_{\mathbf{w} \subseteq u^c} \left(\frac{1}{n}\right)^{|\mathbf{w}|} B_{u \cup \mathbf{w}, n} &\leq \mathbb{E}[Y^2] \sum_{\mathbf{w} \subseteq u^c} \left(\frac{1}{n}\right)^{|\mathbf{w}|} \sum_{\emptyset \neq \mathbf{v} \subseteq (u \cup \mathbf{w})^c} \frac{1}{(n-1)^{|\mathbf{v}|}} \\ &\leq \mathbb{E}[Y^2] \sum_{\mathbf{w} \subseteq u^c} \left(\frac{1}{n}\right)^{|\mathbf{w}|} \left(\left(1 + \frac{1}{n-1}\right)^{d-|\mathbf{u}|-|\mathbf{w}|} - 1 \right) \\ &\leq \mathbb{E}[Y^2] \left[\left(1 + \frac{1}{n-1}\right)^{d-|\mathbf{u}|} \sum_{\mathbf{w} \subseteq u^c} \left(\frac{1}{n}\right)^{|\mathbf{w}|} - \sum_{\mathbf{w} \subseteq u^c} \left(\frac{1}{n}\right)^{|\mathbf{w}|} \right] \\ &\leq \mathbb{E}[Y^2] \left[\left(1 + \frac{1}{n-1}\right)^{d-|\mathbf{u}|} \left(1 + \frac{1}{n}\right)^{d-|\mathbf{u}|} - \left(1 + \frac{1}{n}\right)^{d-|\mathbf{u}|} \right] \end{aligned} \quad (58)$$

and the conclusion follows by applying twice Lemma 5. □

2.2 Proof of Proposition 3.2

(i) The proof is divided into two parts. In the first one, we only consider continuous functions, and in the second one, we extend the result to the larger class of functions such that f^4 is integrable.

First part: Consistency is obvious as in Proposition 3.1, except for the term

$$\frac{1}{n} \sum_{j=1}^n Y^{j,RLHS} Y_u^{j,RLHS} . \quad (59)$$

So denote $\bar{\mathbf{Z}}^j$ the Latin hypercube defined by

$$\bar{\mathbf{Z}}^j = \frac{\lfloor n\mathbf{Z}^j \rfloor + \mathbf{U}^j}{n} \quad (60)$$

where the \mathbf{U}^j s are independent random vectors uniformly distributed in $[0, 1]^d$ independent from all the permutations and shifts in the definition of $(\mathbf{Z}^j)_j$, and $\lfloor \cdot \rfloor$ is the floor function. We can write $\frac{1}{n} \sum_{j=1}^n Y^{j,RLHS} Y_{\mathbf{u}}^{j,RLHS} = \frac{1}{n} \sum_{j=1}^n f(\mathbf{X}_{\mathbf{u}}^j, \mathbf{Z}_{\mathbf{u}^c}^j) f(\mathbf{X}_{\mathbf{u}}^j, \mathbf{Z}_{\mathbf{u}^c}^{\prime j})$ as

$$\frac{1}{n} \sum_{j=1}^n f(\mathbf{X}_{\mathbf{u}}^j, \mathbf{Z}_{\mathbf{u}^c}^j) f(\mathbf{X}_{\mathbf{u}}^j, \bar{\mathbf{Z}}_{\mathbf{u}^c}^j) + \frac{1}{n} \sum_{j=1}^n f(\mathbf{X}_{\mathbf{u}}^j, \mathbf{Z}_{\mathbf{u}^c}^j) \left(f(\mathbf{X}_{\mathbf{u}}^j, \mathbf{Z}_{\mathbf{u}^c}^{\prime j}) - f(\mathbf{X}_{\mathbf{u}}^j, \bar{\mathbf{Z}}_{\mathbf{u}^c}^j) \right). \quad (61)$$

The first term on the right-hand side is an estimator as described in Section 3.1. in the main document since by the construction proposed in Eq. (60), $(\mathbf{Z}^j)_{j=1..n}$ and $(\bar{\mathbf{Z}}^j)_{j=1..n}$ are two independent LHS; so Proposition 3.1 states it converges to $\mathbb{E}[Y]^2 + \underline{\tau}_{\mathbf{u}}^2$ almost surely. The second term on the right-hand side converges to 0 since, as f is bounded, — by continuity on a compact — it is bounded by

$$\frac{\sup |f|}{n} \sum_{j=1}^n \left| f(\mathbf{X}_{\mathbf{u}}^j, \mathbf{Z}_{\mathbf{u}^c}^{\prime j}) - f(\mathbf{X}_{\mathbf{u}}^j, \bar{\mathbf{Z}}_{\mathbf{u}^c}^j) \right| \quad (62)$$

and by uniform continuity of f — due to Heine-Cantor theorem — this quantity tends to 0 as n tends to $+\infty$. Thus the sum in the right-hand side, i.e. $\frac{1}{n} \sum_{j=1}^n Y^{j,RLHS} Y_{\mathbf{u}}^{j,RLHS}$, converges to $\mathbb{E}[Y]^2 + \underline{\tau}_{\mathbf{u}}^2$ almost surely.

Second part: Since the space of continuous functions on $[0, 1]^d$ — denoted $\mathcal{C}([0, 1]^d)$ — is dense in $L^4([0, 1]^d)$, let $(f_m)_{m \in \mathbb{N}^*}$ be a sequence in $\mathcal{C}([0, 1]^d)$ such that $\mathbb{E}[|f_m(\mathbf{X}) - f(\mathbf{X})|^4]$ converges to 0 as m tends to $+\infty$, where \mathbf{X} is uniformly distributed on $[0, 1]^d$.

Now let $\varepsilon > 0$ and $M = M(\varepsilon) \in \mathbb{N}^*$ such that

$$\mathbb{E} \left[(f_M(\mathbf{X}) - f(\mathbf{X}))^2 \right] < \frac{\varepsilon^2}{65 \mathbb{E}[f^2(\mathbf{X})]}. \quad (63)$$

We can write $\frac{1}{n} \sum_{j=1}^n Y^{j,RLHS} Y_{\mathbf{u}}^{j,RLHS}$ as

$$\begin{aligned} & \frac{1}{n} \sum_{j=1}^n f(\mathbf{X}_{\mathbf{u}}^j, \mathbf{Z}_{\mathbf{u}^c}^j) f(\mathbf{X}_{\mathbf{u}}^j, \bar{\mathbf{Z}}_{\mathbf{u}^c}^j) + \frac{1}{n} \sum_{j=1}^n f(\mathbf{X}_{\mathbf{u}}^j, \mathbf{Z}_{\mathbf{u}^c}^j) \left(f_M(\mathbf{X}_{\mathbf{u}}^j, \mathbf{Z}_{\mathbf{u}^c}^{\prime j}) - f_M(\mathbf{X}_{\mathbf{u}}^j, \bar{\mathbf{Z}}_{\mathbf{u}^c}^j) \right) \\ & + \frac{1}{n} \sum_{j=1}^n f(\mathbf{X}_{\mathbf{u}}^j, \mathbf{Z}_{\mathbf{u}^c}^j) \left(f(\mathbf{X}_{\mathbf{u}}^j, \mathbf{Z}_{\mathbf{u}^c}^{\prime j}) - f_M(\mathbf{X}_{\mathbf{u}}^j, \mathbf{Z}_{\mathbf{u}^c}^{\prime j}) \right) + \frac{1}{n} \sum_{j=1}^n f(\mathbf{X}_{\mathbf{u}}^j, \mathbf{Z}_{\mathbf{u}^c}^j) \left(f_M(\mathbf{X}_{\mathbf{u}}^j, \bar{\mathbf{Z}}_{\mathbf{u}^c}^j) - f(\mathbf{X}_{\mathbf{u}}^j, \bar{\mathbf{Z}}_{\mathbf{u}^c}^j) \right). \end{aligned} \quad (64)$$

As noted in the proof of (i) in Proposition 3.1, the first term on the right-hand side of (64) converges to $\underline{\tau}_{\mathbf{u}}^2 + \mathbb{E}[Y]^2$ almost surely as n tends to $+\infty$ i.e.

$$\mathbb{P} \left(\forall \varepsilon > 0, \exists N_1 \in \mathbb{N}^*, \forall n > N_1, \left| \frac{1}{n} \sum_{j=1}^n f(\mathbf{X}_{\mathbf{u}}^j, \mathbf{Z}_{\mathbf{u}^c}^j) f(\mathbf{X}_{\mathbf{u}}^j, \bar{\mathbf{Z}}_{\mathbf{u}^c}^j) - \underline{\tau}_{\mathbf{u}}^2 - \mathbb{E}[Y]^2 \right| < \frac{\varepsilon}{4} \right) = 1. \quad (65)$$

Since f_M is uniformly continuous on $[0, 1]^d$, we have that

$$A_n = \sup_{1 \leq j \leq n} \left| f_M(\mathbf{X}_{\mathbf{u}}^j, \mathbf{Z}_{\mathbf{u}^c}^{\prime j}) - f_M(\mathbf{X}_{\mathbf{u}}^j, \bar{\mathbf{Z}}_{\mathbf{u}^c}^j) \right| \quad (66)$$

converges almost surely to 0 as n tends to $+\infty$. Moreover, since f is integrable, we have that $\frac{1}{n} \sum_{j=1}^n |f(\mathbf{X}_u^j)|$ converges to $\mathbb{E}[|Y|]$ as n tends to $+\infty$. Hence

$$\begin{aligned} & \mathbb{P}\left(\forall \varepsilon > 0, \exists N_1 \in \mathbb{N}^*, \forall n > N_1, \left| \frac{1}{n} \sum_{j=1}^n f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) \left(f_M(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) - f_M(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j) \right) \right| < \frac{\varepsilon}{4}\right) \\ & \geq \mathbb{P}\left(\forall \varepsilon > 0, \exists N_2 \in \mathbb{N}^*, \forall n > N_2, A_n \frac{1}{n} \sum_{j=1}^n |f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j)| < \frac{\varepsilon}{4}\right) \\ & = 1. \end{aligned} \tag{67}$$

For the third and the fourth terms on the right-hand side of (64), we apply twice the same proof. First the Cauchy-Schwartz inequality gives

$$\begin{aligned} & \mathbb{P}\left(\forall \varepsilon > 0, \exists N_3 \in \mathbb{N}^*, \forall n > N_3, \left| \frac{1}{n} \sum_{j=1}^n f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) \left(f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) - f_M(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) \right) \right| < \frac{\varepsilon}{4}\right) \\ & \geq \mathbb{P}\left(\forall \varepsilon > 0, \exists N_3 \in \mathbb{N}^*, \forall n > N_3, \left(\frac{1}{n} \sum_{j=1}^n f^2(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) \right)^{1/2} \left(\frac{1}{n} \sum_{j=1}^n (f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) - f_M(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j))^2 \right)^{1/2} < \frac{\varepsilon}{4}\right). \end{aligned} \tag{68}$$

Then note that $\frac{1}{n} \sum_{j=1}^n f^2(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j)$ and $\frac{1}{n} \sum_{j=1}^n (f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) - f_M(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j))^2$ converge almost surely to $\mathbb{E}[Y^2]$ and $\mathbb{E}[(f_M(\mathbf{X}) - f(\mathbf{X}))^2]$ — where \mathbf{X} is uniformly distributed on $[0, 1]^d$ — respectively. And deduce that there exists $N_4 \in \mathbb{N}^*$ such that for all $n > N_4$, we have $\frac{1}{n} \sum_{j=1}^n f^2(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) < 2 \mathbb{E}[Y^2]$ and $\frac{1}{n} \sum_{j=1}^n (f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) - f_M(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j))^2 < 2 \mathbb{E}[(f_M(\mathbf{X}) - f(\mathbf{X}))^2]$ almost surely. As a consequence, deduce from Eq. (63) that

$$\begin{aligned} & \mathbb{P}\left(\forall \varepsilon > 0, \exists N_3 \in \mathbb{N}^*, \forall n > N_3, \left| \frac{1}{n} \sum_{j=1}^n f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) \left(f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) - f_M(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) \right) \right| < \frac{\varepsilon}{4}\right) \\ & \geq \mathbb{P}\left(\forall \varepsilon > 0, \exists N_3 > N_4, \forall n > N_3, \varepsilon \sqrt{\frac{4}{65}} < \frac{\varepsilon}{4}\right) \\ & = 1 \end{aligned} \tag{69}$$

Finally, Eqs. (65–69) gives

$$\mathbb{P}\left(\forall \varepsilon > 0, \exists N \in \mathbb{N}^*, \forall n > N, \left| \frac{1}{n} \sum_{j=1}^n Y^{j,RLHS} Y_u^{j,RLHS} \right| < \varepsilon\right) = 1 \tag{70}$$

and we have the conclusion.

(ii) As in (i), the only term to treat is

$$\frac{1}{n} \sum_{j=1}^n Y^{j,RLHS} Y_u^{j,RLHS}, \tag{71}$$

so asymptotic normality is shown in the same way by using the decomposition in (61). We always obtain the sum of a term already considered in Section 3 in the main document which converges in law to a normal distribution and a term which converges to 0 in probability, and the conclusion follows from Slutsky's lemma. We only detail the proof for $\tilde{S}_{u,n}^{RLHS}$, it is exactly

the same for $\widehat{\underline{S}}_{u,n}^{RLHS}$. So note that following the proof of (ii) in Proposition 3.1 and the notation above, it is sufficient to show that

$$\sqrt{n} \left(\frac{1}{n} \sum_{j=1}^n (f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) - \mathbb{E}[Y]) (f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^{\prime j}) - f(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j)) \right) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0 \quad (72)$$

to prove the asymptotic normality of $\widetilde{\underline{S}}_{u,n}^{RLHS}$.

So consider $\varepsilon, \eta > 0$ and prove that there exists $N \in \mathbb{N}^*$ such that for all $n > N$, the quantity

$$P = \mathbb{P} \left(\left| \frac{1}{n} \sum_{j=1}^n (f(\mathbf{X}_u^j) - \mathbb{E}[Y]) (f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^{\prime j}) - f(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j)) \right| > \varepsilon \right) \quad (73)$$

is less than η . First as f^6 is integrable, there exists a constant $K > 0$ such that $\mathbb{P}(|f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j)| > K) < \eta/4$. Hence

$$\begin{aligned} P &\leq \mathbb{P} \left(\left(\left| \frac{1}{\sqrt{n}} \sum_{j=1}^n (f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) - \mathbb{E}[Y]) (f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^{\prime j}) - f(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j)) \right| > \varepsilon \right) \cap (|f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j)| \leq K) \right) \\ &\quad + \mathbb{P} \left(\left(\left| \frac{1}{\sqrt{n}} \sum_{j=1}^n (f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) - \mathbb{E}[Y]) (f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^{\prime j}) - f(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j)) \right| > \varepsilon \right) \cap (|f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j)| > K) \right) \\ &< \mathbb{P} \left(\frac{K + |\mathbb{E}[Y]|}{\sqrt{n}} \sum_{j=1}^n |f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^{\prime j}) - f(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j)| > \varepsilon \right) + \frac{\eta}{4}. \end{aligned} \quad (74)$$

Now note that the space of continuous functions on $[0, 1]^d$, denoted by $\mathcal{C}([0, 1]^d)$, is dense in $L^6([0, 1]^d)$ and let $(f_m)_{m \in \mathbb{N}^*}$ be a sequence in $\mathcal{C}([0, 1]^d)$ such that $\mathbb{E}[|f_m(\mathbf{X}) - f(\mathbf{X})|^6]$ converges to 0 as m tends to $+\infty$ where \mathbf{X} is uniformly distributed on $[0, 1]^d$. It is easy to note that there exists $M = M(n)$ such that $\mathbb{P}(|f_M(\mathbf{X}) - f(\mathbf{X})| > 1/n) < \eta/4$. Thus we get from Eq. (74) that

$$\begin{aligned} P &< \sum_{i=1}^4 \mathbb{P} \left(\left(\frac{K + |\mathbb{E}[Y]|}{\sqrt{n}} \sum_{j=1}^n (|f_M(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^{\prime j}) - f_M(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j)| + |f_M(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^{\prime j}) - f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^{\prime j})| \right. \right. \\ &\quad \left. \left. + |f_M(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j) - f(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j)| \right) \cap A_i \right) + \frac{\eta}{4} \end{aligned} \quad (75)$$

where

$$A_1 = (|f_M(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j) - f(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j)| > \frac{1}{n}) \cap (|f_M(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^{\prime j}) - f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^{\prime j})| > \frac{1}{n}) \quad (76)$$

$$A_2 = (|f_M(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j) - f(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j)| > \frac{1}{n}) \cap (|f_M(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^{\prime j}) - f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^{\prime j})| < \frac{1}{n}) \quad (77)$$

and A_3 and A_4 are the complementary events of A_1 and A_2 , respectively. So we deduce

$$\begin{aligned} P &< \mathbb{P} \left(\left(\frac{K + |\mathbb{E}[Y]|}{\sqrt{n}} \sum_{j=1}^n (|f_M(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^{\prime j}) - f_M(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j)| + |f_M(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^{\prime j}) - f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^{\prime j})| \right. \right. \\ &\quad \left. \left. + |f_M(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^{\prime j}) - f(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j)| \right) \cap A_3 \right) + \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_4) + \frac{\eta}{4} \\ &< \mathbb{P} \left(\frac{K + |\mathbb{E}[Y]|}{\sqrt{n}} \left(2 + \sum_{j=1}^n |f_M(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^{\prime j}) - f_M(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j)| \right) > \varepsilon \right) + \eta. \end{aligned} \quad (78)$$

Now by an other density argument, note that there exists a sequence of Lipschitz continuous functions with constant 1, denoted $(f_{M,q})_{q \in \mathbb{N}^*}$, such that $\sup_{[0,1]^d} |f_{M,q}(\mathbf{x}) - f_M(\mathbf{x})|$ converges to 0 as q tends to $+\infty$. Then there exists $Q = Q(n) \in \mathbb{N}^*$ such that $\sup_{[0,1]^d} |f_{M,Q}(\mathbf{x}) - f_M(\mathbf{x})| < 1/n$ and deduce that

$$\begin{aligned} P &< \mathbb{P}\left(\frac{K + |\mathbb{E}[Y]|}{\sqrt{n}} \left(2 + \sum_{j=1}^n (|f_{M,Q}(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) - f_{M,Q}(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j)| + |f_{M,Q}(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) - f_M(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j)| \right. \right. \\ &\quad \left. \left. + |f_{M,Q}(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j) - f_M(\mathbf{X}_u^j, \bar{\mathbf{Z}}_{u^c}^j)|\right) > \varepsilon\right) + \eta \\ &< \mathbb{P}\left(\frac{5(K + |\mathbb{E}[Y]|)}{\sqrt{n}} > \varepsilon\right) + \eta. \end{aligned} \quad (79)$$

and the conclusion follows.

(iii) First note that since $(\mathbf{X}^j)_j \sim \mathcal{LH}(n, d)$ and $(\mathbf{Z}^j, \mathbf{Z}'^j)_j \sim \mathcal{RLH}(n, d)$, we have

$$Y^{j,RLHS} = f(\mathbf{X}^j, \mathbf{Z}_{u^c}^j) \quad \text{and} \quad Y_u^{j,RLHS} = f(\mathbf{X}_u^j, \mathbf{Z}_{u^c}^j) \quad (80)$$

with

$$\mathbf{X}^j = \left(\frac{\pi_1''(j) - U'_{1,\pi_1''(j)}}{n}, \dots, \frac{\pi_d''(j) - U'_{d,\pi_d''(j)}}{n} \right) \quad (81)$$

$$\mathbf{Z}^j = \left(\frac{\pi_1(j) - U_{1,\pi_1(j)}}{n}, \dots, \frac{\pi_d(j) - U_{d,\pi_d(j)}}{n} \right) \quad (82)$$

and

$$\mathbf{Z}'^j = \left(\frac{\pi_1'(j) - U_{1,\pi_1'(j)}}{n}, \dots, \frac{\pi_d'(j) - U_{d,\pi_d'(j)}}{n} \right) \quad (83)$$

where the π_i s, the π_i' s, the π_i'' s, the $U_{i,j}$ s and the $U'_{i,j}$ s are independent random variables uniformly distributed on Π_n — see Definition 1 in the main document —, Π_n , Π_n , $[0, 1]$ and $[0, 1]$, respectively. Moreover note that if for an index i in $\{1, \dots, d\}$, we have $\pi_i(j) = \pi_i'(j)$ then $Z_i^j = Z_i'^j$; and if $\pi_i(j) \neq \pi_i'(j)$ then $U_{i,\pi_i(j)}$ and $U_{i,\pi_i'(j)}$ are independent and therefore Z_i^j and $Z_i'^j$ are two distinct points of a Latin hypercube of size n in $[0, 1]$. For $j \in \{1, \dots, n\}$, denote by $\mathfrak{e}(j)$ the set of integers $i \in \mathbf{u}^c$ such that $\pi_i(j) = \pi_i'(j)$.

Thus we have $\mathbb{E}[Y^{j,RLHS} Y_u^{j,RLHS}]$ equal to

$$\begin{aligned} &\frac{1}{n^{2d-|\mathbf{u}|}} \sum_{\mathbf{w} \subseteq \mathbf{u}^c} \sum_{\substack{\pi_u''(j) \in \\ \{1, \dots, n\}^{|\mathbf{u}|}}} \sum_{\substack{\pi_{u^c}(j) \in \\ \{1, \dots, n\}^{d-|\mathbf{u}|}}} \sum_{\substack{\pi_{u^c}'(j) \in \\ \{1, \dots, n\}^{d-|\mathbf{u}|}}} \mathbf{1}_{\{\mathfrak{e}(j)=\mathbf{w}\}} \int f\left(\dots, \underbrace{\frac{\pi_i''(j) - u_{1i}}{n}}_{i \in \mathbf{u}}, \dots, \underbrace{\frac{\pi_k(j) - u_{1k}}{n}}_{k \in \mathbf{u}^c}, \dots\right) \dots \\ &\quad f\left(\dots, \underbrace{\frac{\pi_i''(j) - u_{1i}}{n}}_{i \in \mathbf{u}}, \dots, \underbrace{\frac{\pi_k(j) - u_{1k}}{n}}_{k \in \mathbf{u}^c \cap \mathbf{w}}, \dots, \underbrace{\frac{\pi_l'(j) - u_{2l}}{n}}_{l \in \mathbf{u}^c \cap \mathbf{w}^c}, \dots\right) d\mathbf{u}_1 d\mathbf{u}_2(\mathbf{u} \cup \mathbf{w})^c \end{aligned} \quad (84)$$

where for all $i \in \mathbf{u} \cup \mathfrak{e}(j)$, $\pi_i(j) = \pi_i'(j)$ and $u_{1i}(j) = u_{2i}(j)$. And noting that

$$\frac{1}{(n-1)^{d-|\mathbf{u}|-|\mathbf{w}|} n^d} \sum_{\substack{\pi_u''(j) \in \\ \{1, \dots, n\}^{|\mathbf{u}|}}} \sum_{\substack{\pi_{u^c}(j) \in \\ \{1, \dots, n\}^{d-|\mathbf{u}|}}} \sum_{\substack{\pi_{u^c}'(j) \in \\ \{1, \dots, n\}^{d-|\mathbf{u}|}}} \mathbf{1}_{\{\mathfrak{e}(j)=\mathbf{w}\}} \int f\left(\dots, \underbrace{\frac{\pi_i''(j) - u_{1i}}{n}}_{i \in \mathbf{u}}, \dots, \underbrace{\frac{\pi_k(j) - u_{1k}}{n}}_{k \in \mathbf{u}^c}, \dots\right) \dots$$

$$f\left(\dots, \underbrace{\frac{\pi_i''(j) - u_{1i}}{n}}_{i \in \mathbf{u}}, \dots, \underbrace{\frac{\pi_k(j) - u_{1k}}{n}}_{k \in \mathbf{u}^c \cap \mathbf{w}}, \dots, \underbrace{\frac{\pi_l'(j) - u_{2l}}{n}}_{l \in \mathbf{u}^c \cap \mathbf{w}^c}, \dots\right) d\mathbf{u}_1 d\mathbf{u}_{2(\mathbf{u} \cup \mathbf{w})^c} \quad (85)$$

is equal to $\mathbb{E}[f(\mathbf{X}^1)f(\mathbf{X}_{\mathbf{u} \cup \mathbf{w}}^1, \mathbf{X}_{(\mathbf{u} \cup \mathbf{w})^c}^2)]$, Lemma 6 gives

$$\mathbb{E}[Y^{j,RLHS}Y_{\mathbf{u}}^{j,RLHS}] = \sum_{\mathbf{w} \subseteq \mathbf{u}^c} \left(\frac{1}{n}\right)^{|\mathbf{w}|} \left(\mathbb{E}[Y]^2 + \tau_{\mathbf{u} \cup \mathbf{w}}^2 + B_{\mathbf{u} \cup \mathbf{w}, n}\right). \quad (86)$$

By Lemmas 5 and 7, and noting that $\mathbb{E}[Y]^2 + \tau_{\mathbf{u} \cup \mathbf{w}}^2 \leq \mathbb{E}[Y^2]$, we obtain

$$\mathbb{E}[Y^{j,RLHS}Y_{\mathbf{u}}^{j,RLHS}] = \mathbb{E}[Y]^2 + \tau_{\mathbf{u}}^2 + B_{|\mathbf{u}|, n} \quad (87)$$

where

$$|B_{|\mathbf{u}|, n}| \leq \left(\frac{d - |\mathbf{u}| + 1}{n} + 2\right) \left(\frac{d - |\mathbf{u}| + 1}{n - 1}\right) \mathbb{E}[Y^2]. \quad (88)$$

Following the same proof, it is easy to show that for $j \neq l$, we have

$$\mathbb{E}[Y^{j,RLHS}Y_{\mathbf{u}}^{l,RLHS}] = \mathbb{E}[Y]^2 + B_{n,1} \quad (89)$$

where

$$|B_{n,1}| \leq \left(\frac{d+1}{n} + 2\right) \left(\frac{d+1}{n-1}\right) \mathbb{E}[Y^2]. \quad (90)$$

Thus noting that

$$\mathbb{E}[\tau_{\mathbf{u}, n}^{2,RLHS}] = \frac{n-1}{n} \mathbb{E}[Y^{1,RLHS}Y_{\mathbf{u}}^{1,RLHS}] - \frac{n-1}{n} \mathbb{E}[Y^{1,RLHS}Y_{\mathbf{u}}^{2,RLHS}] \quad (91)$$

we conclude that

$$\mathbb{E}[\tilde{\tau}_{\mathbf{u}, n}^{2,RLHS}] = \tau_{\mathbf{u}}^2 - \frac{1}{n} \tau_{\mathbf{u}}^2 + \frac{n-1}{n} (B_{n,1} + B_{|\mathbf{u}|, n}) \quad (92)$$

where the biases are $O(n^{-1})$ as specified in Proposition 3.2. Concerning $\tilde{\sigma}_n^{2,RLHS}$, note that $\tilde{\sigma}_n^{2,RLHS} = \tilde{\sigma}_n^{2,LHS}$ and the conclusion follows from (iii) in Proposition 3.1. Concerning $\hat{\tau}_{\mathbf{u}, n}^{2,RLHS}$

and $\hat{\sigma}_n^{2,RLHS}$, we have $\mathbb{E}\left[\left(\frac{1}{n} \sum_{j=1}^n \frac{Y^{j,RLHS} + Y_{\mathbf{u}}^{j,RLHS}}{2}\right)^2\right]$ equal to

$$\begin{aligned} & \frac{1}{4n} \mathbb{E}[(Y^{1,RLHS} + Y_{\mathbf{u}}^{1,RLHS})^2] + \frac{1}{4n^2} \sum_{j=1}^n \sum_{\substack{l=1 \\ l \neq j}}^n \mathbb{E}[(Y^{j,RLHS} + Y_{\mathbf{u}}^{j,RLHS})(Y^{l,RLHS} + Y_{\mathbf{u}}^{l,RLHS})] \\ &= \frac{1}{2n} \left(\mathbb{E}[(Y^{1,RLHS})^2] + \mathbb{E}[Y^{1,RLHS}Y_{\mathbf{u}}^{1,RLHS}]\right) + \frac{n-1}{2n} \left(\mathbb{E}[Y^{1,RLHS}Y^{2,RLHS}] + \mathbb{E}[Y^{1,RLHS}Y_{\mathbf{u}}^{2,RLHS}]\right). \end{aligned}$$

Then using notation introduced in Section 3 in the main document, note that

$$\mathbb{E}[Y^{1,RLHS}Y^{2,RLHS}] = \mathbb{E}[Y^{1,LHS}Y^{2,LHS}] = \text{Cov}(Y^{1,LHS}, Y^{2,LHS}) + \mathbb{E}[Y]^2 = \text{Cov}(Y^{1,LHS}, Y_{\{1, \dots, d\}}^{2,LHS}) + \mathbb{E}[Y]^2$$

and by (88), (90) and Lemma 4, we deduce

$$\mathbb{E}\left[\left(\frac{1}{n} \sum_{j=1}^n \frac{Y^{j,RLHS} + Y_{\mathbf{u}}^{j,RLHS}}{2}\right)^2\right] = \frac{1}{2n} \tau_{\mathbf{u}}^2 + \mathbb{E}[Y]^2 + B_{n,1} + B_{n,2}$$

where

$$|B_{n,2}| \leq \frac{\sigma^2}{2n} \quad (93)$$

and $B_{n,1}$ is specified in (90). Then it is easy to conclude that

$$\mathbb{E}[\widehat{\tau}_{u,n}^{2,RLHS}] = \tau_u^2 - \frac{1}{2n}\tau_u^2 + B_{n,1} + B_{n,2} + \frac{n-1}{n}B_{|u|,n} \quad (94)$$

$$\mathbb{E}[\widehat{\sigma}_n^{2,RLHS}] = \sigma^2 - \frac{1}{2n}\tau_u^2 + B_{n,1} + B_{n,2} \quad (95)$$

where the biases are $O(n^{-1})$ as specified in Proposition 3.2.

References

- Janon, A., Klein, T., Lagnoux, A., Nodet, M., and Prieur, C. (2013). Asymptotic normality and efficiency of two sobol index estimators. *ESAIM: Probability and Statistics (to appear)*.
- Loh, W. L. (1996). On latin hypercube sampling. *The Annals of Statistics*, 24(5):2058–2080.
- McKay, M. D., Conover, W. J., and Beckman, R. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245.
- Stein, M. (1987). Large sample properties of simulations using latin hypercube sampling. *Technometrics*, 29(2):143–151.
- Van der Vaart, A. W. (1998). *Asymptotics Statistics*. Cambridge University Press.