



Robust estimation with latin hypercube sampling: a central limit theorem for Z-estimators

Faouzi Hakimi

► To cite this version:

Faouzi Hakimi. Robust estimation with latin hypercube sampling: a central limit theorem for Z-estimators. 2026. <hal-04462415v4>

HAL Id: hal-04462415

<https://hal.science/hal-04462415v4>

Preprint submitted on 8 Jan 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Robust Estimation With Latin Hypercube Sampling: A Central Limit Theorem for Z-estimators

Faouzi Hakimi^{1,2}

¹Aix-Marseille University, 13284 Marseille, France

²Institut Mathématique de Toulouse, 31062 Toulouse, France

Correspondence

Faouzi Hakimi, Marseille, France
Email: faouzi.hakimi@univ-amu.fr

Funding information

Latin Hypercube Sampling (LHS) is a widely used stratified sampling method in computer experiments. In this work, we extend existing convergence results for the sample mean under LHS to the broader class of Z -estimators – estimators defined as the zeros of a sample mean function. We derive the asymptotic variance of these estimators and demonstrate that it is smaller when using LHS compared to traditional independent and identically distributed sampling. Furthermore, we establish a Central Limit Theorem for Z -estimators under LHS, providing a theoretical foundation for its improved efficiency.

KEYWORDS

Design of Experiments, Statistical Computing, Robust Estimation, Variance Reduction, Latin Hypercube Sampling

1 | INTRODUCTION

Latin Hypercube Sampling (LHS), introduced in [1], is a compelling alternative to *independent and identically distributed (i.i.d.)* random sampling for exploring the behavior of complex systems (often treated as black-boxes) through computer experiments [2, 3, 4]. A significant number of subsequent sampling methods are founded upon LHS (see [5], [6], [7] or [8] for instance), underscoring its importance in the field. Consequently, the comprehension and theoretical development of LHS is fundamental in contemporary research.

To generate an LHS sample of size n , the range of each variable is divided into n equally probable intervals. In the case of two variables, the n sample points are then positioned such that there is exactly one sample in each row and each column. Figure 1 illustrates schematic examples of LHS designs with dimension $d = 2$ and size $n = 4$. The process generalizes naturally to higher dimensions.

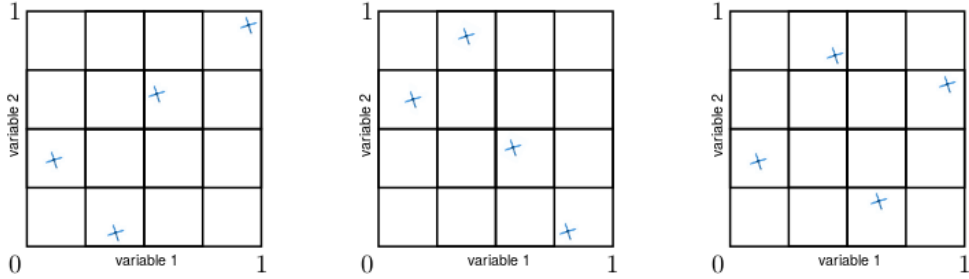


Figure 1 Three schematic examples of LHS designs with dimension $d = 2$ and size $n = 4$.

Several theoretical results have been established regarding the convergence of estimators under LHS. Most of these results focus on the empirical mean of a measurable function with a finite second-order moment. For instance, it was shown in [9] that the asymptotic variance of the sample mean is smaller under LHS compared to classic *i.i.d.* random sampling for such statistics. Additionally, a Central Limit Theorem (CLT) for the empirical mean of bounded functions was proven in [10] and later extended to functions with finite third-order moments in [11].

Subsequent research has extended these results on Latin Hypercube Sampling (LHS) to various methods derived from LHS, such as those accounting for dependencies (e.g., [12]) and other derivative approaches (e.g., [6]). However, the majority of these findings remain focused on the empirical mean.

This paper aims to broaden the scope of the convergence results for empirical mean estimators under classic LHS to encompass the more general class of Z -estimators. This class, closely related to the well-established class of M -estimators, includes all estimators that can be expressed as the zeros of an empirical mean function. This work therefore provides novel insights into convergence properties using Latin Hypercube Sampling, which is particularly valuable in computational experiments where variance reduction and efficiency are important to consider.

Most results discussed in this paper are also presented in Chapter 2 of the thesis manuscript [13]. The paper is organized as follows: Section 2 provides a formal definition of Latin Hypercube Sampling along with its key convergence properties. Section 3 introduces the definitions and relevant properties of Z -estimators. Original results concerning the asymptotic normality of Z -estimators under LHS are presented in Section 4. Finally, an application example is discussed in Section 5.

2 | GENERAL NOTATIONS, DEFINITIONS, AND PROPERTIES OF LHS

Let us introduce some useful notations regarding this work. We first denote by $\mathbf{X} = (X_1, \dots, X_d)$ the vector of d (with $d \in \mathbb{N}^* = \mathbb{N} \setminus \{0\}$) independent random variables evolving in $\mathcal{X} \subset \mathbb{R}^d$. For simplicity and without loss of generality, we assume that the d inputs vary uniformly in $[0, 1]$ so we have that, for j in $\llbracket 1, d \rrbracket = \{1, 2, \dots, d\}$ and $X_j \sim U_{[0,1]}$. Indeed, one can always work under uniformity and then use the inverse transformation method [14] to place the support back on the original scale and retrieve the original distribution, as long as the sampling distribution of interest is a product measure (see for instance [10] p543 for details).

Consider a size n ($n \in \mathbb{N}^*$) sample of \mathbf{X} generated using either *i.i.d.* random sampling or Latin Hypercube Sampling (LHS). For either of these methods (denoted here as *METHOD*), the sample is represented as follows:

- The matrix of samples is denoted by $\mathbf{X}^{METHOD} = \left(x^{(1)}, \dots, x^{(n)} \right)^T \in M_{n,d}([0, 1])$. Here, $M_{n,d}([0, 1])$ denotes the space of matrices of size $n \times d$ with coefficients in $[0, 1]$.
- The j th column of \mathbf{X}^{METHOD} , denoted by $x_j^{METHOD} = \left(x_j^{(1)}, \dots, x_j^{(n)} \right)^T$ with $j \in \llbracket 1, d \rrbracket$, represents the effective generated sample of the input X_j .

Specifically, a sample generated by classic *i.i.d.* random sampling will be denoted by \mathbf{X}^{IID} , and a sample generated by Latin Hypercube sampling will be denoted by \mathbf{X}^{LHS} . Similar notations will be used for any quantities estimated with either of these two sampling methods. If no sampling method is mentioned, it means that the results presented do not depend on the sampling method.

We also define the measurable function $g : \mathcal{X} \rightarrow \mathbb{R}^q$ with $q \in \mathbb{N}^*$. This function represents in practice the studied simulation code. We denote by $g(\mathbf{X}^{METHOD}) = (g(x^{(1)}), \dots, g(x^{(n)}))^T$ the matrix of output samples corresponding to \mathbf{X}^{METHOD} .

For $\alpha = (a_1, \dots, a_q) \in \mathbb{R}^q$ with $q \in \mathbb{N}^*$, we denote by $\|\alpha\|$ the Euclidean norm of α such that $\|\alpha\|^2 = \sum_{i=1}^q a_i^2$. Similarly, for any matrix \mathbf{A} in $M_{q,q}(\mathbb{R})$, we denote by $\|\mathbf{A}\|$ the *pseudo* Euclidean norm (Frobenius norm) such that $\|\mathbf{A}\|^2 = \sum_{1 \leq i,j \leq q} A_{i,j}^2$. Here, $A_{i,j}$ with $i, j \in \llbracket 1, q \rrbracket$ are the components of the matrix \mathbf{A} . A *pseudo* Euclidean norm $\|\cdot\|$ is finally associated with the tensor space $T_{q,q,q}(\mathbb{R})$. This norm is defined, for all \mathbf{T} in $T_{q,q,q}(\mathbb{R})$, by $\|\mathbf{T}\|^2 = \sum_{1 \leq i,j,k \leq q} T_{i,j,k}^2$. Here, $T_{i,j,k}$ with $i, j, k \in \llbracket 1, q \rrbracket$ are the components of the tensor \mathbf{T} .

We denote by $o(1)$ ("small oh-one") a deterministic sequence that converges to 0 and $O(1)$ ("big oh-one") a deterministic sequence that is bounded. We denote by $o_p(1)$ ("small oh-P-one") a sequence of random variables that converges in probability to 0. The expression $O_p(1)$ ("big oh-P-one") denotes a sequence of random variables that is bounded in probability. We recall that a sequence of random variables $(\mathbf{W}_n)_{n \in \mathbb{N}}$ is bounded in probability if, for any scalar $\epsilon > 0$, there exist M and N such that, for all $n > N$, $\mathbb{P}(\|\mathbf{W}_n\| > M) < \epsilon$ (note that this definition holds in the general case where the norm $\|\cdot\|$ is not Euclidean).

Finally, a multivariate normal distribution of dimension q ($q \in \mathbb{N}^*$) with a mean equal to $\mu \in \mathbb{R}^q$ and a covariance matrix equal to Σ in $M_{q,q}(\mathbb{R})$ is denoted $\mathcal{N}_q(\mu, \Sigma)$.

As previously stated, Latin Hypercube Sampling is a statistical method used to generate a near-random sample of parameter values from a multidimensional distribution. To define it formally, we denote, for $d, n \in \mathbb{N}^*$:

1. $\pi_j = (\pi_j(1) \dots \pi_j(n))^T, j \in \llbracket 1, d \rrbracket$ as a random permutation of $\llbracket 1, n \rrbracket$, according to the uniform distribution on the set of all possible permutations of $\llbracket 1, n \rrbracket$. Random permutations $(\pi_j)_{j \in \llbracket 1, d \rrbracket}$ are assumed to be independent.
2. $u_j = (u_j^{(1)}, \dots, u_j^{(n)})^T, j \in \llbracket 1, d \rrbracket$ as an i.i.d. sample of the uniform distribution $U_{[0,1]}$. The samples $(u_j)_{j \in \llbracket 1, d \rrbracket}$ are assumed to be independent.

Random permutations $(\pi_j)_{j \in \llbracket 1, d \rrbracket}$ and samples $(u_j)_{j \in \llbracket 1, d \rrbracket}$ are also assumed to be independent. The n -sized sampling x_j^{LHS} of the input $X_j, j \in \llbracket 1, d \rrbracket$, is then defined as follows:

$$x_j^{LHS} = (x_j^{(1)}, \dots, x_j^{(n)})^T = \left(\frac{1}{n}(\pi_j(1) - u_j^{(1)}), \dots, \frac{1}{n}(\pi_j(n) - u_j^{(n)}) \right)^T. \quad (1)$$

The corresponding LHS design of dimension d and size n is then $\mathbf{X}^{LHS} = (x_1^{LHS}, \dots, x_d^{LHS})$.

As a result of its stratified nature, the realizations of the LHS design are not i.i.d.. However, several results have been indeed established for the convergence of estimators under LHS. Most of them concern the sample mean (first order U-statistics) of measurable functions. For instance, it has been shown in [1] that, for any measurable function, this estimator is unbiased:

Proposition 1 Let $g : [0, 1]^d \rightarrow \mathbb{R}^q$ with $d, q \in \mathbb{N}^*$ be a measurable function such that $\mathbb{E}(\|g(\mathbf{X})\|) < +\infty$. Denote

$$G_n^{LHS} = \frac{1}{n} \sum_{i=1}^n g(x^{(i)}),$$

where $x^{(i)}, i \in \llbracket 1, n \rrbracket$ is such that $\mathbf{X}^{LHS} = (x^{(1)}, \dots, x^{(n)})^T$ with \mathbf{X}^{LHS} being defined using Equation (1). Then, G_n^{LHS} is an unbiased estimator of $G = \mathbb{E}(g(\mathbf{X}))$.

Similarly to Proposition 1, we denote by G_n^{IID} the classic sample mean of an i.i.d. design: $G_n^{IID} = \frac{1}{n} \sum_{i=1}^n g(x^{(i)})$, with $x^{(i)}, i \in \llbracket 1, n \rrbracket$ being a sample of an i.i.d. design $\mathbf{X}^{IID} \in M_{n,d}([0, 1])$.

A second interesting characteristic of mean value estimators under LHS is their variance. Indeed, Stein [9] showed that if g is a real-valued function such that $\mathbb{E}(g^2(\mathbf{X})) < +\infty$, then $\text{Var}(G_n^{LHS})$ is always asymptotically smaller than $\text{Var}(G_n^{IID})$. This result is generalized to multidimensional functions by Loh in [11]. Proposition 2 summarizes the main results regarding the covariance matrix of G_n^{LHS} :

Proposition 2 Let $g : [0, 1]^d \rightarrow \mathbb{R}^q$ ($d, q \in \mathbb{N}^*$) be a measurable function with $\mathbb{E}(\|g(\mathbf{X})\|^2) < +\infty$. Let $\Sigma_{G_n^{IID}}, \Sigma_{G_n^{LHS}} \in M_{q,q}(\mathbb{R})$ be the covariance matrices of G_n^{IID} and G_n^{LHS} respectively, with $\Sigma_{G_n^{IID}} = \frac{1}{n} \mathbb{E} \left((g(\mathbf{X}) - G)(g(\mathbf{X}) - G)^T \right)$.

We also define, for $x = (x_1, \dots, x_d) \in [0, 1]^d$:

- $g_{-j}(x_j) = \int_{[0,1]^{d-1}} [g(x) - G] \prod_{1 \leq k \leq d, k \neq j} dx_k = \mathbb{E}(g(\mathbf{X}) - G | X_j)$ with $j \in \llbracket 1, d \rrbracket$.
- $g_{rem}(x) = g(x) - G - \sum_{j=1}^d g_{-j}(x_j)$.
- $R_g = \int_{[0,1]^d} g_{rem}(x) g_{rem}(x)^T dx$.

Then we have:

- $\Sigma_{G_n^{LHS}} = \frac{1}{n} R_g + \frac{1}{n} o(1)$.
- $\Sigma_{G_n^{IID}} = \frac{1}{n} R_g + \frac{1}{n} \sum_{j=1}^d \int_{[0,1]} g_{-j}(x_j) g_{-j}(x_j)^T dx_j$.

We therefore have that $\Sigma_{G_n^{IID}} - \Sigma_{G_n^{LHS}}$ is asymptotically positive semi-definite, that is,

$$\forall \xi \in \mathbb{R}^d, \quad \lim_{n \rightarrow +\infty} n \xi^T (\Sigma_{G_n^{IID}} - \Sigma_{G_n^{LHS}}) \xi \geq \sum_{j=1}^d \int_{[0,1]} \xi^T g_{-j}(x_j) g_{-j}(x_j)^T \xi dx_j \geq 0.$$

Since G_n^{IID} converges in quadratic mean to G and that G_n^{LHS} is an unbiased estimator of G (as established in Proposition 1), we can conclude that G_n^{LHS} also converges in quadratic mean to G : $\lim_{n \rightarrow +\infty} \mathbb{E}(\|G_n^{LHS} - G\|^2) = 0$. Consequently, G_n^{LHS} converges in probability to G .

In addition, Owen [10] showed a Central Limit Theorem (CLT) for this class of estimators under LHS when the model function g is bounded. This was generalized to any function with finite third moment in [11]:

Theorem 1 *In the framework of Proposition 2, let $g : [0, 1]^d \rightarrow \mathbb{R}^q$ ($d, q \in \mathbb{N}^*$) be a measurable function with $\mathbb{E}(\|g(\mathbf{X})\|^3) < +\infty$. Then, assuming that R_g is non-singular, we have that $\sqrt{n}(G_n^{LHS} - G)$ tends in distribution to $N_q(0, R_g)$ as $n \rightarrow +\infty$.*

3 | DEFINITIONS AND PROPERTIES ON Z-ESTIMATORS

The primary objective of this work is to extend the convergence results under Latin Hypercube Sampling (LHS) to the class of Z -estimators. The Z -estimator class is intimately related to the well-established class of M -estimators, yet it offers a distinct formulation. Specifically, Z -estimators are defined as solutions to a set of estimating equations, which can be viewed as a generalization of the optimization problem associated with M -estimators. This formulation provides a flexible framework for parameter estimation, encompassing a wide range of statistical models and inference procedures. For a comprehensive discussion on these topics, we refer the reader to [15] and [16]. Significant theoretical work has been conducted on this class of estimators, which remains a major research topic today (see, for example, [17] or [18]).

More formally, let $\mathbf{X} = (x^{(1)}, \dots, x^{(n)})^T$ be the vector of n realizations of a random vector X evolving in $\mathcal{X} \subset \mathbb{R}^d$, with $n, d \in \mathbb{N}^*$. Its law is parameterized by a vector $\theta \in \Theta \subset \mathbb{R}^q$, $q \in \mathbb{N}^*$.

For $x \in \mathcal{X}$, $\theta \in \Theta$, let $(x, \theta) \rightarrow \psi_\theta(x) \in \mathbb{R}^q$ be a known measurable function such that $\psi_\theta(x) = (\psi_{\theta_1}(x), \dots, \psi_{\theta_q}(x))^T$. We also define the empirical mean of this function $(X, \theta) \rightarrow \Psi_n(\theta) \in \mathbb{R}^q$ such that $\Psi_n(\theta) = \frac{1}{n} \sum_{i=1}^n \psi_\theta(x^{(i)})$.

The Z -estimator $\hat{\theta}_n = \hat{\theta}_n(x^{(1)}, \dots, x^{(n)}) \in \Theta$ associated with $\psi_\theta(x)$ corresponds to the solution of the following vectorial equation:

$$\Psi_n(\theta) = 0. \quad (2)$$

Many known estimators can be defined as Z -estimators. For instance, let X have a distribution function f_θ with a continuous first derivative in $\theta \in \Theta$ and a separable log-likelihood function. In this case, the maximum likelihood estimator of θ can be written as a Z -estimator as defined by 2 with, for $x \in \mathbb{R}^d$, $d \in \mathbb{N}^*$, $\psi_\theta(x) = (\frac{\partial \log(f_\theta(x))}{\partial \theta_1}, \dots, \frac{\partial \log(f_\theta(x))}{\partial \theta_q})^T$.

The Z -estimator framework also encompasses Generalized Method of Moments (GMM) estimators [19], where $\psi_\theta(x)$ comprises moment conditions $g(x, \theta)$ equating sample and theoretical moments. Quantile regression also fits within this framework, with $\psi_\theta(x, y) = [I(y \leq x^T \theta) - \tau]x$, where τ is the quantile of interest, and $I(\cdot)$ is the indicator function. This estimates conditional quantiles rather than the mean. These examples highlight the versatility of the Z -estimator framework, where the key is specifying an appropriate ψ function reflecting the estimation objectives.

The first useful properties regarding Z -estimators concern the link between the consistency of $\Psi_n(\theta)$ and the consistency of $\hat{\theta}_n$. For instance, in [20], one can find assumptions for which the consistency of $\hat{\theta}_n$ is ensured:

Proposition 3 *Let Θ be a compact subset of \mathbb{R}^q with $q \in \mathbb{N}^*$. Let also assume that the following hypotheses are true, for any $\theta \in \Theta$ and $n \in \mathbb{N}^*$:*

- the functions $\theta \rightarrow \Psi_n(\theta)$ and $\theta \rightarrow \Psi(\theta)$ are continuous measurable functions of $\theta \in \Theta$ evolving in \mathbb{R}^q ;
- each function $\Psi_n(\theta)$ has exactly one zero $\hat{\theta}_n \in \Theta$;
- $\Psi_n(\theta)$ converges to $\Psi(\theta)$ in probability;
- $\Psi(\theta)$ vanishes only at θ_0 with $\theta_0 \in \Theta$;
- denoting, for $\eta \geq 0$, $w_n(\eta) = \sup\{|\Psi_n(\theta_1) - \Psi_n(\theta_2)|; |\theta_1 - \theta_2| \leq \eta, \theta_1, \theta_2 \in \Theta\}$; there exists two sequences (η_k) and (ϵ_k) both decreasing to 0 such that, for all $k \in \mathbb{N}$, $\mathbb{P}(w_n(\eta_k) > \epsilon_k) \xrightarrow[n \rightarrow +\infty]{} 0$.

Then $\hat{\theta}_n$ is a consistent estimator of θ_0 , that is $\hat{\theta}_n \xrightarrow[n \rightarrow +\infty]{P} \theta_0$.

The assumption on $w_n(\eta)$ seems difficult to grasp at first glance. However, as mentioned in [20], if we find a function ϕ from \mathbb{R}_+ to \mathbb{R} such that $\lim_{\eta \rightarrow 0^+} \phi(\eta) = 0$, this assumption on w_n can be obtained through: $\mathbb{P}(w_n(\eta) \geq 2\phi(\eta)) \xrightarrow[n \rightarrow +\infty]{} 0$ for each $\eta \geq 0$. For instance, $w_n(\eta) \xrightarrow[n \rightarrow +\infty]{} \phi(\eta)$, or $\lim_{n \rightarrow +\infty} w_n(\eta) \leq \phi(\eta)$ give both sufficient conditions. Note that Proposition 3 is general and does not mention any sampling scheme.

In addition to these convergence properties, several Central Limit Theorems for Z -estimators have been proved. Here we give one of them, proposed in [16]. Theorem 2 relies on the so-called *classic conditions*, formulated to mathematically tighten the informal derivation of the asymptotic normality of maximum likelihood proposed by [21]. These conditions are stringent, but they are simple. They lead to a simple proof of the Central Limit Theorem. This simplicity will allow us to adapt this theorem to the LHS case.

In particular, a needed assumption for the application of this theorem concerns the existence of a first and a second order derivatives in θ for ψ_θ . Let us introduce these terms.

For any $\theta \in \Theta$ and for any $x \in \mathcal{X}$, let $(x, \theta) \rightarrow \dot{\psi}_\theta(x)$ be the first order partial derivative of $\psi_\theta \in \mathbb{R}^q$, assuming it exists. This first order partial derivative is evolving in $M_{q,q}(\mathbb{R})$. Its components are $\dot{\psi}_{\theta_{j,k}} = \frac{\partial \psi_{\theta_j}}{\partial \theta_k}$ with $j, k \in \llbracket 1, q \rrbracket$.

Similarly, for any $\theta \in \Theta$ and for any $x \in \mathcal{X}$, let $(x, \theta) \rightarrow \ddot{\psi}_\theta(x)$ be the second order partial derivative of $\psi_\theta \in \mathbb{R}^q$, assuming it exists. This second order partial derivative is evolving in $T_{q,q,q}(\mathbb{R})$. Its components are $\ddot{\psi}_{\theta_{j,k,l}} = \frac{\partial^2 \psi_{\theta_j}}{\partial \theta_k \partial \theta_l}$, with $j, k, l \in \llbracket 1, q \rrbracket$.

Theorem 2 *Let Θ be an open subset of an Euclidean space of dimension q , $q \in \mathbb{N}^*$ and let \mathcal{X} be a subspace of \mathbb{R}^d , $d \in \mathbb{N}^*$. Assume that, for all θ in Θ and for all x in \mathcal{X} , the function $(x, \theta) \rightarrow \psi_\theta(x)$ evolving in \mathbb{R}^q is twice continuously differentiable*

in θ .

Let $(x, \theta) \rightarrow \psi_\theta(x) \in M_{q,q}(\mathbb{R})$ and $(x, \theta) \rightarrow \dot{\psi}_\theta(x) \in T_{q,q,q}(\mathbb{R})$ denote the first and second-order derivatives of ψ_θ , respectively.

Let $X^{IID} = (x^{(1)}, \dots, x^{(n)})^T$ be the vector of i.i.d. realizations of a random variable $X = (X_1, \dots, X_d)$ evolving in \mathcal{X} .

Suppose also that the following assumptions are fulfilled:

1. $\Psi_n^{IID}(\hat{\theta}_n^{IID}) = \frac{1}{n} \sum_{i=1}^n \psi_{\hat{\theta}_n^{IID}}(x^{(i)}) = 0, \forall n \in \mathbb{N}^*$;
2. there exists a unique θ_0 in Θ such that $\mathbb{E}(\psi_{\theta_0}(X)) = \Psi(\theta_0) = 0$ with θ_0 in Θ ;
3. $\mathbb{E}(\|\psi_{\theta_0}(X)\|^2) < +\infty$;
4. $\mathbb{E}(\dot{\psi}_{\theta_0}(X))$ exists and is non-singular;
5. For any $x \in \mathcal{X}$ and for any θ in the neighborhood of θ_0 , the function $(x, \theta) \rightarrow \dot{\psi}_\theta(x) \in T_{q,q,q}(\mathbb{R})$ is dominated, in norm, by a fixed integrable function $x \rightarrow \tilde{\psi}(x) \in T_{q,q,q}(\mathbb{R})$.

Then, if $\hat{\theta}_n^{IID}$ is a consistent estimator of θ_0 , we have:

$$(\hat{\theta}_n^{IID} - \theta_0) = -[\mathbb{E}(\dot{\psi}_{\theta_0}(X))]^{-1} \frac{1}{n} \sum_{i=1}^n \psi_{\theta_0}(x^{(i)}) + \frac{1}{\sqrt{n}} o_p(1). \quad (3)$$

Moreover, we have that the sequence $\sqrt{n}(\hat{\theta}_n^{IID} - \theta_0)$ tends in distribution to $\mathcal{N}_q(0, [\mathbb{E}(\dot{\psi}_{\theta_0}(X))]^{-1} \mathbb{E}(\psi_{\theta_0}(X) \psi_{\theta_0}(X)^T) [\mathbb{E}(\dot{\psi}_{\theta_0}(X))]^{-T})$ as $n \rightarrow +\infty$.

For the following, it is important to note that we have $\Sigma_{\Psi_n^{IID}(\theta_0)} = \frac{1}{n} \mathbb{E}(\psi_{\theta_0}(X) \psi_{\theta_0}(X)^T)$ with $\Sigma_{\Psi_n^{IID}(\theta_0)} \in M_{q,q}(\mathbb{R})$ being the covariance matrix of $\Psi_n^{IID}(\theta_0)$. It is also important to remark that among the results presented in this section, only Theorem 2 requires the specific use of an i.i.d. sample, since its proof relies on the classical Central Limit Theorem (CLT) [22].

While Theorem 2 assumes Θ is open, one can modify this assumption to consider Θ as the interior of a compact set. This allows us to maintain the differentiability conditions required for asymptotic normality while preserving the compactness needed for Proposition 3 concerning consistency.

4 | Z-ESTIMATORS UNDER LHS

In this section, we extend the convergence properties of Z -estimators to LHS designs. The idea is to combine all the above properties. Indeed, one can first notice that the Z -function $\Psi_n(\theta)$ is the empirical mean of ψ_θ . Now, as mentioned in Section 2, the convergence of this type of statistic under LHS holds. We use that here to show a Central Limit Theorem for Z -estimators under LHS.

As in Section 2, let, for any $\theta \in \Theta$ and $X^{LHS}, X^{IID} \in M_{n,d}([0,1])$, $\Sigma_{\Psi_n^{IID}(\theta)}, \Sigma_{\Psi_n^{LHS}(\theta)} \in M_{q,q}(\mathbb{R})$ be the covariance matrices of $\Psi_n^{IID}(\theta)$ and $\Psi_n^{LHS}(\theta)$ respectively. Let us now give some noteworthy convergence properties on $\Psi_n^{LHS}(\theta)$.

Proposition 4 Let Θ be a compact subset of \mathbb{R}^q and $X = [0, 1]^d$ (q, d in \mathbb{N}^*). Let $X^{LHS} = (x^{(1)}, \dots, x^{(n)})^T$ be the vector of LHS realizations of a random variable $X = (X_1, \dots, X_d)$ evolving in \mathcal{X} such that $X \sim U_{[0,1]^d}$. Assume also that, for all $\theta \in \Theta$ and $x \in \mathcal{X}$, the function $(x, \theta) \rightarrow \psi_\theta(x)$ is measurable regarding x . We then have the following properties on $\Psi_n^{LHS}(\theta) = \frac{1}{n} \sum_{i=1}^n \psi_\theta(x^{(i)})$:

1. If, for all $\theta \in \Theta$, $\mathbb{E}(\|\psi_\theta(X)\|) < +\infty$, $\Psi_n^{LHS}(\theta)$ is an unbiased estimator of $\Psi(\theta) = \mathbb{E}(\psi_\theta(X))$.
2. If, for all $\theta \in \Theta$, $\mathbb{E}(\|\psi_\theta(X)\|^2) < +\infty$, we also have:
 $\Sigma_{\Psi_n^{LHS}(\theta)} = \frac{1}{n} \int_{[0,1]^d} \psi_{\theta_{rem}}(x) \psi_{\theta_{rem}}(x)^T dx + \frac{1}{n} o(1)$, with $\psi_{\theta_{rem}}$ being defined as in Proposition 2.
 Moreover, we have that $\Sigma_{\Psi_n^{LHS}(\theta)} - \Sigma_{\Psi_n^{IID}(\theta)}$ is asymptotically positive semi-definite and that $\Psi_n^{LHS}(\theta)$ converges in quadratic mean to $\Psi(\theta)$. In other words, we have $\lim_{n \rightarrow +\infty} \mathbb{E}(\|\Psi_n^{LHS}(\theta) - \Psi(\theta)\|^2) = 0$.
3. If, for all $\theta \in \Theta$, $\mathbb{E}(\|\psi_\theta(X)\|^3) < +\infty$ and if $R_{\psi_\theta} = \int_{[0,1]^d} \psi_{\theta_{rem}}(x) \psi_{\theta_{rem}}(x)^T dx$ is non-singular, we have that $\sqrt{n}(\Psi_n^{LHS} - \Psi(\theta))$ tends in distribution to $N_q(0, R_{\psi_\theta})$ as $n \rightarrow +\infty$.

Proof. Let us show these properties one by one:

1. Since, for all $\theta \in \Theta$ and $x \in \mathcal{X}$, the function $(x, \theta) \rightarrow \psi_\theta(x)$ is measurable regarding $x \in \mathcal{X}$ and $\mathbb{E}(\|\psi_\theta(X)\|) < +\infty$, $\Psi_n^{LHS}(\theta)$ is an unbiased estimator of $\Psi(\theta)$ by Proposition 1.
2. This is a direct consequence of Proposition 2.
3. This is a direct consequence of Theorem 1. ■

All these properties on $\Psi_n^{LHS}(\theta)$ allow to show that $\hat{\theta}_n^{LHS}$ is a consistent estimator of θ_0 . Indeed, the assertion 2 of Proposition 4 ensures the convergence in probability of $\Psi_n^{LHS}(\theta)$ to $\Psi(\theta)$. As mentioned before, Proposition 3 does not impose any other conditions on the sampling scheme. We therefore have, under the conditions of application of this proposition, that $\hat{\theta}_n^{LHS}$ converges to θ_0 in probability. Let us now establish a Central Limit Theorem for Z-estimators under LHS.

Theorem 3 Let Θ be the interior of a compact subset of \mathbb{R}^q , $q \in \mathbb{N}^*$, and $\mathcal{X} = [0, 1]^d$, $d \in \mathbb{N}^*$. For all $\theta \in \Theta$ and $x \in \mathcal{X}$, assume $(x, \theta) \rightarrow \psi_\theta(x)$, where $\psi_\theta = (\psi_{\theta_1}(x), \dots, \psi_{\theta_q}(x))^T \in \mathbb{R}^q$, is twice continuously differentiable in θ .

Let $(x, \theta) \rightarrow \dot{\psi}_\theta(x) \in M_{q,q}(\mathbb{R})$ and $(x, \theta) \rightarrow \ddot{\psi}_\theta(x) \in T_{q,q,q}(\mathbb{R})$ denote the first and second-order derivatives of ψ_θ , respectively.

For any $n \in \mathbb{N}^*$, let $X^{LHS} = (x^{(1)}, \dots, x^{(n)})^T$ be LHS realizations of $X \sim U_{[0,1]^d}$ of size n . Suppose also that the following hypotheses are fulfilled:

1. For any $n \in \mathbb{N}^*$, $\Psi_n(\hat{\theta}_n^{LHS}) = \frac{1}{n} \sum_{i=1}^n \psi_{\hat{\theta}_n^{LHS}}(x^{(i)}) = 0$;
2. There is $\theta_0 \in \Theta$ such that $\mathbb{E}(\psi_{\theta_0}(X)) = \Psi(\theta_0) = 0$;
3. $\mathbb{E}(\|\psi_{\theta_0}(X)\|^2) < +\infty$;
4. $\mathbb{E}(\dot{\psi}_{\theta_0}(X))$ is non-singular and $\mathbb{E}(\|\dot{\psi}_{\theta_0}(X)\|^2) < +\infty$;
5. There is an integrable function $x \rightarrow \ddot{\psi}(x) \in T_{q,q,q}(\mathbb{R})$, $x \in \mathcal{X}$, such that $\|\ddot{\psi}_\theta(x)\| \leq \|\ddot{\psi}(x)\|$ and $\mathbb{E}(\|\ddot{\psi}(X)\|^2) < +\infty$ for all θ in the neighborhood of θ_0 .

If $\hat{\theta}_n^{LHS}$ is a consistent estimator of θ_0 , then:

$$\hat{\theta}_n^{LHS} - \theta_0 = -[\mathbb{E}(\dot{\psi}_{\theta_0}(X))]^{-1} \frac{1}{n} \sum_{i=1}^n \psi_{\theta_0}(x^{(i)}) + \frac{1}{\sqrt{n}} o_p(1). \quad (4)$$

Moreover, let $\Sigma_{\hat{\theta}_n^{LHS}} \in M_{q,q}(\mathbb{R})$ be the covariance matrix of $\hat{\theta}_n^{LHS}$. Then:

$$\Sigma_{\hat{\theta}_n^{LHS}} = [\mathbb{E}(\dot{\psi}_{\theta_0}(X))]^{-1} \Sigma_{\psi_{\theta_0}^{LHS}(\theta_0)} [\mathbb{E}(\dot{\psi}_{\theta_0}(X))]^{-T} + \frac{1}{n} o(1), \quad (5)$$

where

$$\Sigma_{\psi_{\theta_0}^{LHS}(\theta_0)} = \frac{1}{n} R_{\psi_{\theta_0}} + \frac{1}{n} o(1). \quad (6)$$

with $R_{\psi_{\theta_0}} = \int_{[0,1]^d} \psi_{\theta_{0,rem}}(x) \psi_{\theta_{0,rem}}(x)^T dx$. Furthermore, $\Sigma_{\hat{\theta}_n^{LHS}} - \Sigma_{\hat{\theta}_n^{IID}}$ is asymptotically positive semi-definite ($\Sigma_{\hat{\theta}_n^{IID}} \in M_{q,q}(\mathbb{R})$) corresponding to the covariance matrix of $\hat{\theta}_n^{IID}$.

Finally, if $\mathbb{E}(\|\psi_{\theta_0}(X)\|^3) < +\infty$ and $R_{\psi_{\theta_0}}$ is non-singular, then $\sqrt{n}(\hat{\theta}_n^{LHS} - \theta_0)$ tends in distribution to $N(0, [\mathbb{E}(\dot{\psi}_{\theta_0}(X))]^{-1} R_{\psi_{\theta_0}} [\mathbb{E}(\dot{\psi}_{\theta_0}(X))]^{-T})$ as $n \rightarrow +\infty$.

Proof. The proof follows the reasoning in [16] for Theorem 2.

By Taylor's Theorem, as $\Psi_n(\cdot)$ is continuous and twice differentiable in θ , $\exists \tilde{\theta}_n^{LHS}$ between θ_0 and $\hat{\theta}_n^{LHS}$ such that:

$$\begin{aligned} \Psi_n^{LHS}(\hat{\theta}_n^{LHS}) - \Psi_n^{LHS}(\theta_0) &= \dot{\Psi}_n^{LHS}(\theta_0)(\hat{\theta}_n^{LHS} - \theta_0) \\ &\quad + \frac{1}{2}(\hat{\theta}_n^{LHS} - \theta_0)^T \ddot{\Psi}_n^{LHS}(\tilde{\theta}_n^{LHS})(\hat{\theta}_n^{LHS} - \theta_0). \end{aligned} \quad (7)$$

Since $\mathbb{E}(\|\psi_{\theta_0}(X)\|^2) < +\infty$, Proposition 4 implies:

$$\Psi_n^{LHS}(\theta_0) = \frac{1}{n} \sum_{i=1}^n \psi_{\theta_0}(x^{(i)}) \xrightarrow[n \rightarrow +\infty]{P} \mathbb{E}(\psi_{\theta_0}(X)) = 0. \quad (8)$$

Now, let $\dot{\Psi}_n^{LHS}(\theta) = \frac{1}{n} \sum_{i=1}^n \dot{\psi}_{\theta}(x^{(i)})$ be the empirical mean over x of the matrix function $(x, \theta) \rightarrow \dot{\psi}_{\theta}(x)$, with $\dot{\psi}_{\theta}(x) \in M_{q,q}(\mathbb{R})$. Similarly, $\ddot{\Psi}_n^{LHS}(\theta) \xrightarrow[n \rightarrow +\infty]{P} \mathbb{E}(\ddot{\psi}_{\theta_0}(X))$, which is non-singular by assumption.

Let also $\ddot{\Psi}_n^{LHS}(\theta) = \frac{1}{n} \sum_{i=1}^n \ddot{\psi}_{\theta}(x^{(i)})$ be the empirical mean over x of the tensor function $(x, \theta) \rightarrow \ddot{\psi}_{\theta}(x)$, with $\ddot{\psi}_{\theta}(x) \in T_{q,q,q}(\mathbb{R})$.

For $\ddot{\Psi}_n^{LHS}(\theta)$, let \mathcal{B} be a ball around θ_0 where $\|\ddot{\psi}_{\theta}\| \leq \|\ddot{\psi}\|$ with $x \rightarrow \ddot{\psi}(x) \in T_{q,q,q}(\mathbb{R})$ being an integrable function and $\mathbb{E}(\|\ddot{\psi}(X)\|^2) < +\infty$ (this ball exists by assumption). Since $\hat{\theta}_n^{LHS} \xrightarrow[n \rightarrow +\infty]{P} \theta_0$, we have $\mathbb{P}(\hat{\theta}_n^{LHS} \in \mathcal{B}) \rightarrow 1$. For $\hat{\theta}_n^{LHS} \in \mathcal{B}$:

$$\|\ddot{\Psi}_n^{LHS}(\hat{\theta}_n^{LHS})\| \leq \frac{1}{n} \sum_{i=1}^n \|\ddot{\psi}(x^{(i)})\|. \quad (9)$$

The right-hand side converges to a finite value by Proposition 2, implying the same for the left-hand side.

Rewriting the Taylor expansion:

$$-\Psi_n^{LHS}(\theta_0) = \left(\mathbb{E}(\dot{\psi}_{\theta_0}(X)) + o_p(1) + \frac{1}{2}(\hat{\theta}_n^{LHS} - \theta_0)^T O_p(1) \right) (\hat{\theta}_n^{LHS} - \theta_0). \quad (10)$$

As $\hat{\theta}_n^{LHS} \xrightarrow[n \rightarrow +\infty]{P} \theta_0$, we have:

$$-\Psi_n^{LHS}(\theta_0) = \left(\mathbb{E}(\dot{\psi}_{\theta_0}(X)) + o_p(1) \right) (\hat{\theta}_n^{LHS} - \theta_0). \quad (11)$$

This yields equation (4), since $\mathbb{E}(\dot{\psi}_{\theta_0})$ is non-singular and $\Psi_n^{LHS}(\theta_0) = \frac{1}{\sqrt{n}} O_p(1)$ asymptotically.

We also have $\Sigma_{\hat{\theta}_n^{LHS}} = [\mathbb{E}(\dot{\psi}_{\theta_0}(X))]^{-1} \Sigma_{\Psi_n^{LHS}(\theta_0)} [\mathbb{E}(\dot{\psi}_{\theta_0}(X))]^{-T} + \frac{1}{n} o(1)$.

By Proposition 4, $\Sigma_{\hat{\theta}_n^{IID}(\theta_0)} - \Sigma_{\hat{\theta}_n^{LHS}(\theta_0)}$ is asymptotically positive semi-definite, which implies the same property for $\Sigma_{\hat{\theta}_n^{IID}} - \Sigma_{\hat{\theta}_n^{LHS}}$.

Finally, if $\mathbb{E}(\|\psi_{\theta_0}(X)\|^3) < +\infty$ and $R_{\psi_{\theta_0}}$ is non-singular, the asymptotic normality follows from assertion 3 of Proposition 4. ■

These results give an asymptotic convergence for $\hat{\theta}_n^{LHS}$ with, in the univariate case, a lower asymptotic variance of estimation than $\hat{\theta}_n^{IID}$ (corresponding to $\Sigma_{\hat{\theta}_n^{IID}} - \Sigma_{\hat{\theta}_n^{LHS}}$ being asymptotically positive semi-definite in the multivariate case). Moreover, it gives a Central Limit Theorem for Z -estimators under LHS. Although strong regularity conditions on ψ_{θ} are needed for these results to be valid, it remains very useful in many practical cases (e.g., for estimation by maximum likelihood). In the next section, we give an example of application.

5 | APPLICATION: PARAMETERS ESTIMATION OF GENERALIZED LINEAR MODELS (GLM)

When performing statistical analysis on a computational code, it is common to approximate its outputs using a regression or classification model, also known as a metamodel. If the estimation of the modeling parameters can be expressed as a Z -estimator and the other conditions of use are satisfied, Theorem 3 ensures that the estimation variance of these parameters is asymptotically lower under LHS than under IID sampling. It also provides a Central Limit Theorem under LHS.

Consider for instance the case of Generalized Linear Models (GLM), proposed in [23]. They were formulated as a way of unifying various statistical models, including linear regression, logistic regression and Poisson regression. To estimate the parameters of a GLM, one generally uses a Maximum Likelihood Estimator (MLE). It is therefore a special case of Z -estimation supposing that the likelihood can be differentiated. Thus, the results presented above can be applied to parameters estimation of a GLM.

5.1 | Definitions and main properties on GLM

Before entering in more details, let us first define GLM more formally. For simplicity and without loss of generality, we focus here on the canonical case. Let Z be a random variable on $\mathcal{Z} \subset \mathbb{R}$ and $X = (X_1, \dots, X_d)$ a vector of covariables

on $\mathcal{X} \subset \mathbb{R}^d$, $d \in \mathbb{N}^*$. A GLM is characterized by:

1. A probability distribution: Z follows an exponential family distribution with density:

$$f(z, \alpha, \phi) = a(\alpha) b(z) \exp\left(z \frac{\alpha}{\phi}\right), \quad z \in \mathcal{Z}, \alpha \in \mathbb{R}, \phi > 0, \quad (12)$$

where ϕ is the known dispersion parameter, $a(\alpha) = \exp(-v(\alpha)/\phi)$, $v : \mathbb{R} \rightarrow \mathbb{R}$ is twice continuously differentiable, and $b(z) = \exp(w(z, \phi))$ with $w : \mathcal{Z} \times \mathbb{R}^+ \rightarrow \mathbb{R}$ being also twice continuously differentiable in z .

2. A linear predictor: For $\theta = (\theta_1, \dots, \theta_d)^T \in \Theta \subset \mathbb{R}^d$, Θ being open and bounded:

$$\eta : \mathcal{X} \times \Theta \rightarrow \mathbb{R}, \quad \eta(x, \theta) = x^T \theta = \sum_{j=1}^d x_j \theta_j. \quad (13)$$

3. A link function: Let $h : \mathcal{H} \subset \mathbb{R} \rightarrow \mathbb{R}$ be a monotone, differentiable function and $\alpha = (\alpha^{(1)}, \dots, \alpha^{(n)})^T \in \mathbb{R}^n$

$$h(\mu(x)) = \eta(x, \theta), \quad (14)$$

where $\mu(x) = \mathbb{E}[Z|X = x]$. Note that this hypothesis on h implies the existence of the inverse function $h^{-1} : \mathbb{R} \rightarrow \mathcal{H}$ so that, for any $a \in \mathcal{H}$, $h^{-1} \circ h(a) = a$.

Given n independent realizations $\{(x^{(i)}, z^{(i)})\}_{i=1}^n$, with $\mathbf{X} = (x^{(1)}, \dots, x^{(n)})^T$ and $\mathbf{Z} = (z^{(1)}, \dots, z^{(n)})^T$, we also have $\alpha = (\alpha^{(1)}, \dots, \alpha^{(n)})^T \in \mathbb{R}^n$ such that:

$$\alpha^{(i)} = h^{-1}(x^{(i)T} \theta), \quad i \in \llbracket 1, n \rrbracket. \quad (15)$$

In this framework, the log-likelihood of each observation is:

$$l(z^{(i)}, x^{(i)}, \theta, \phi) = \frac{1}{\phi} [z^{(i)} \alpha^{(i)} - v(\alpha^{(i)})] + w(z^{(i)}, \phi). \quad (16)$$

The maximum likelihood estimator $\hat{\theta}_n$ can be obtained by maximizing the log-likelihood function over the parameter space Θ . Under regularity conditions, including the continuity and differentiability of the log-likelihood function over Θ , the maximum likelihood estimator satisfies the first-order optimality conditions and therefore the following vectorial equation:

$$\sum_{i=1}^n \nabla_{\theta} l(z^{(i)}, x^{(i)}, \theta, \phi) = \mathbf{0}. \quad (17)$$

This defines $\hat{\theta}_n$ as a Z -estimator with $\psi_{\theta}(x^{(i)}) = \nabla_{\theta} l(z^{(i)}, x^{(i)}, \theta, \phi)$.

For the canonical case, the components of $\psi_\theta = (\psi_{\theta_1}(x), \dots, \psi_{\theta_d}(x))^T$ are:

$$\psi_{\theta_j}(x) = \frac{z - h^{-1}(x\theta)}{\phi} x_j, \quad j \in \llbracket 1, d \rrbracket, x \in \mathcal{X}, z \in \mathcal{Z}, \theta \in \Theta. \quad (18)$$

We can see that the estimation of the parameters of a GLM by maximum likelihood fits into the framework of Z -estimation. Thus, let us suppose that the observations of \mathbf{X} are obtained by a LHS. We consider the Z -estimator defined by the equation 17, even though in this case the realizations are no longer *i.i.d*. Let us discuss the convergence of this estimator under LHS.

5.2 | Z -estimation of GLM parameters under LHS

Let $X^{LHS} = (x^{(1)}, \dots, x^{(n)})^T$ be the realizations of \mathbf{X} generated by a LHS. As before, for simplicity, we assume that we have $\mathcal{X} = [0, 1]^d$ and $\mathbf{X} \sim U_{[0,1]^d}$. We also assume that Θ is the interior of a compact subset of \mathbb{R}^d . Moreover, we suppose that h is defined and is twice continuously derivable on \mathcal{H} . We also suppose that h and its first derivative $\dot{h} : \mathcal{H} \rightarrow \mathbb{R}$ have no zero on \mathcal{H} . Since h is monotone by construction, note that h^{-1} is also defined and twice continuously derivable for any $\theta \in \Theta$ and $x \in \mathcal{X}$ thanks to the inverse function theorem (see, for instance, [24] for more details).

Since we suppose that \mathbf{X} and θ are bounded, we have that $\Psi_n^{LHS}(\theta) = \frac{1}{n} \sum_{i=1}^n \psi_\theta(x^{(i)})$ converges in probability to $\mathbb{E}(\psi_\theta(\mathbf{X})) = \Psi(\theta)$. As we have seen, the other conditions concerning the convergence of $\hat{\theta}_n$ to θ_0 are not specific to the sampling scheme. The conditions of application of Proposition 3 are verified both in the case of an IID or a LHS design. We can thus conclude that $\hat{\theta}_n$ converges in probability in θ_0 .

Let us now verify that the conditions of application of Theorem 3 are fulfilled. First, we see that $(X^{LHS}, \theta) \rightarrow \Psi_n^{LHS}(\theta)$ is continuous and twice continuously differentiable in θ . Plus, $\mathbb{E}(\psi_{\theta_0}(\mathbf{X})) = \Psi(\theta_0) = 0$ by construction.

We also have, for $j, k \in \llbracket 1, d \rrbracket$, $x = (x_1, \dots, x_d)^T \in \mathcal{X}$ and $\theta \in \Theta$:

$$\frac{\partial \psi_{\theta_j}(x)}{\partial \theta_k} = \frac{-1}{\phi \dot{h}(h^{-1}(x\theta))} x_j x_k. \quad (19)$$

Thus, we have that the matrix of partial derivatives $x \rightarrow \dot{\psi}_{\theta_0}(x)$ is such that $\mathbb{E}(\dot{\psi}_{\theta_0}(\mathbf{X}))$ is defined and non-singular since h has no zero on \mathcal{H} . Since the values of \mathbf{X} and θ are bounded in norm, the function $(x, \theta) \rightarrow \psi_\theta(x)$ is bounded and thus, for any $\theta \in \Theta$, $\mathbb{E}(\|\psi_\theta(\mathbf{X})\|^3) < +\infty$ (and especially for $\theta = \theta_0$).

Finally, we have that the elements of the tensor $\ddot{\psi}_\theta(x)$ are, for $j, k, l \in \llbracket 1, d \rrbracket$ and $x = (x_1, \dots, x_d)^T \in \mathcal{X}$, as follows:

$$\frac{\partial^2 \psi_{\theta_j}(x)}{\partial \theta_k \partial \theta_l} = \frac{\ddot{h}(h^{-1}(x\theta))}{\phi (\dot{h}(h^{-1}(x\theta)))^3} x_j x_k x_l. \quad (20)$$

Here $\ddot{h} : \mathcal{H} \rightarrow \mathbb{R}$ is the second order derivative of h .

Thus, $(x, \theta) \rightarrow \|\ddot{\psi}_\theta(x)\|$ can be bounded by an integrable function with finite second order moment in the neighbor-

hood of θ_0 since we assume that the values of θ and X are bounded.

All of these statements allow us to apply Theorem 3. We therefore have that the covariance matrix of estimation of $\hat{\theta}_n^{LHS}$ is equal to $\Sigma_{\hat{\theta}_n^{LHS}} = [\mathbb{E}(\dot{\psi}_{\theta_0}(X))]^{-1} \Sigma_{\psi_{\theta_0}^{LHS}(\theta_0)} [\mathbb{E}(\dot{\psi}_{\theta_0}(X))]^{-T} + \frac{1}{n} \mathcal{O}(1)$ and that $\Sigma_{\hat{\theta}_n^{LHS}} - \Sigma_{\hat{\theta}_n^{LHS}}$ is asymptotically positive semi-definite. Note that we have, with the previously introduced notations, $\Sigma_{\psi_{\theta_0}^{LHS}(\theta_0)} = \frac{1}{n} R_{\psi_{\theta_0}} + \frac{1}{n} \mathcal{O}(1)$ asymptotically, with $R_{\psi_{\theta_0}} = \int_{[0,1]^d} \psi_{\theta_0,rem}(x) \psi_{\theta_0,rem}(x)^T dx$.

Finally, since we have $\mathbb{E}(\|\psi_{\theta_0}(X)\|^3) < +\infty$, we have that $\sqrt{n}(\hat{\theta}_n^{LHS} - \theta_0)$ is asymptotically normal with mean zero and a covariance matrix equal to $[\mathbb{E}(\dot{\psi}_{\theta_0}(X))]^{-1} R_{\psi_{\theta_0}} [\mathbb{E}(\dot{\psi}_{\theta_0}(X))]^{-1}$, assuming that $R_{\psi_{\theta_0}}$ is non-singular.

5.3 | Numerical example: a Poisson regression under LHS

To illustrate this result, let us consider a numerical example with a count random variable Z and a vector of covariables $X = (X_1, \dots, X_9)^T$, $X \sim U_{[0,1]^9}$. In an industrial context, Z could represent for instance the number of operating problems evaluated by a simulation code of an industrial facility. In this example, we define Z by the following Poisson density function, for $z \in \mathbb{N}^*$:

$$f(z, \lambda_0) = \exp(-\lambda_0) \frac{1}{z!} \exp(z \log(\lambda_0)), \quad (21)$$

with $\log(\lambda_0) = x\theta_0$, $x \in [0, 1]^9$ and $\theta_0 = (\theta_{0,1}, \dots, \theta_{0,9})^T = (7, -\sqrt{2}, 1/2, -1/3, \sqrt{5}, -7, \sqrt{2}, -1/2, -\sqrt{5})^T$.

One can notice Z fits in the framework of Equations 12, 13 and 14.

Let us compare numerically the performances of the maximum likelihood estimation of θ_0 regarding the sampling method (IID or LHS) in this example. To do so, we first compare the estimation variance of each parameter $(\theta_{0,1}, \dots, \theta_{0,9})^T$ with respect to the sampling scheme and size. We also verify that there is no significant difference concerning the square bias of estimation $[\mathbb{E}(\hat{\theta}_j) - \theta_{0,j}]^2, j \in \llbracket 1, 9 \rrbracket$. Additionally, we display the Mean Squared Error (MSE) for each parameter, defined as:

$$\text{MSE}(\hat{\theta}_j) = \mathbb{E}[(\hat{\theta}_j - \theta_{0,j})^2] = \text{Var}(\hat{\theta}_j) + [\mathbb{E}(\hat{\theta}_j) - \theta_{0,j}]^2, \quad (22)$$

where $\hat{\theta}_j$ is the estimator of $\theta_{0,j}$ and $j \in \llbracket 1, 9 \rrbracket$.

For each sampling method, the average values of these three metrics (variance, squared bias, and MSE) are computed over $L = 1000$ independent LHS and IID designs with sample sizes n ranging from 40 to 100 (in increments of 10).

Figure 2, 3 and 4 show respectively the evolution of the variance, the square bias of estimation and the MSE of the nine estimated parameters $(\theta_{0,1}, \dots, \theta_{0,9})^T$. As expected, we observe that for the nine estimated parameters, the average variance of estimation is overall lower for the classic LHS design compared to IID. No significant differences between LHS and IID designs are observed in terms of the square bias of estimation. The MSE is also significantly

lower. As shown previously, classic LHS designs allow better estimation performances than IID ones, regardless of the theoretical value of the estimated parameters.

To further validate our theoretical results, we conduct an additional numerical experiment comparing the normalized empirical estimation variances ($n \times \text{Var}(\hat{\theta}_j)$, $j \in \llbracket 1, 9 \rrbracket$) with their theoretical asymptotic values. The covariance matrix formula, and hence the variance of each parameter (extracted from its diagonal), is derived from Theorem 3, using the application results given in 5.2 and Proposition 2. The asymptotic variances evaluation is performed using a classic Monte Carlo method with a very large sample size ($n_{\text{MonteCarlo}} = 10^6$), as the analytical evaluation involves complex integrals and expectations that may not have closed-form solutions. The resulting theoretical asymptotic variances, rounded to four decimal places, are (0.6195, 0.3763, 0.3388, 0.3395, 0.3963, 1.4805, 0.3575, 0.3423, 0.4328).

Figure 5 shows the convergence of the normalized empirical variances with their theoretical counterparts as the sample size increases from $n = 50$ to $n = 1000$ in increments of 50. Each of the nine subplots demonstrates that the empirical normalized variances converge to the asymptotic values as n grows, confirming the accuracy of our theoretical predictions. The average values are computed over $L = 1000$ independent LHS designs, reinforcing the reliability of the observed trends.

Furthermore, to explore the asymptotic properties of the estimators, we conduct a third experiment examining the Q-Q plots (Quantile-Quantile plots) of the normalized parameter estimates compared to the standard normal distribution $\mathcal{N}(0, 1)$. These plots compare the quantiles of the empirical distribution to the quantiles of a theoretical distribution, providing a visual assessment of how closely the data follows the specified distribution. Specifically, we generate these Q-Q plots for sample sizes of $n = 50$, $n = 250$, and $n = 1000$. Each plot is obtained with $L = 1000$ independent LHS designs. For clarity, we present the Q-Q plots for the first three parameters in Figure 6, and similar plots for the remaining six parameters are provided in Figures 7 and 8. The Q-Q plots reveal that the empirical distribution exhibits heavy tails for $n = 50$, which improve with $n = 250$ and align even more closely with the theoretical distribution for $n = 1000$. This comparison allows us to visually assess the convergence of the estimators to normality, as predicted by our theoretical results.

In summary, our numerical experiments confirm that Latin Hypercube Sampling (LHS) consistently outperforms Independent and Identically Distributed (IID) sampling in terms of estimation variance and Mean Squared Error (MSE). The normalized empirical variances converge to their theoretical values as the sample size increases, validating our theoretical work. Additionally, the Q-Q plots visually confirm the asymptotic normality of the estimators.

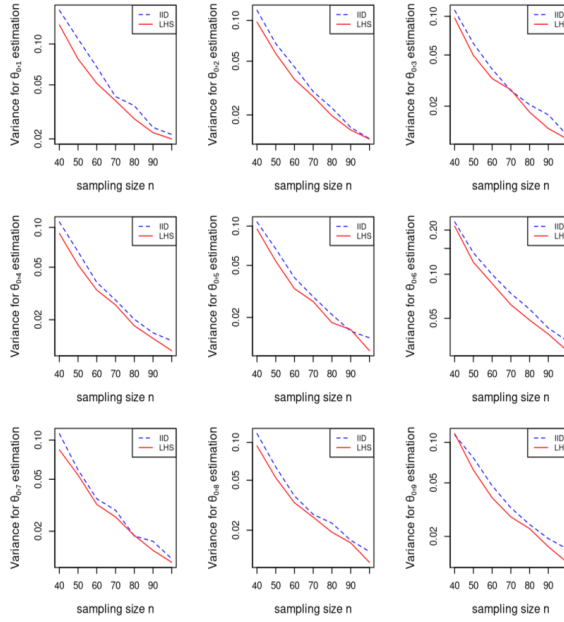


Figure 2 Average estimation variances of $(\theta_{0,1}, \dots, \theta_{0,9})^T$ according to the sampling size n for IID and LHS designs (decimal logarithmic scale).

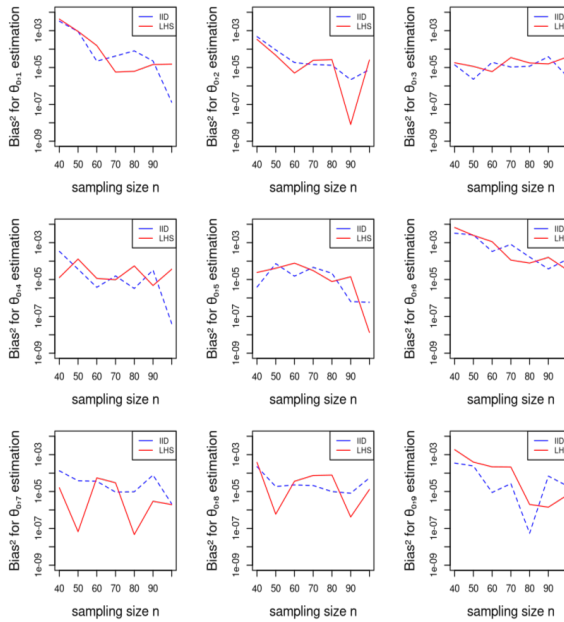


Figure 3 Average estimation square bias of $(\theta_{0,1}, \dots, \theta_{0,9})^T$ according to the sampling size n for IID and LHS designs (decimal logarithmic scale).

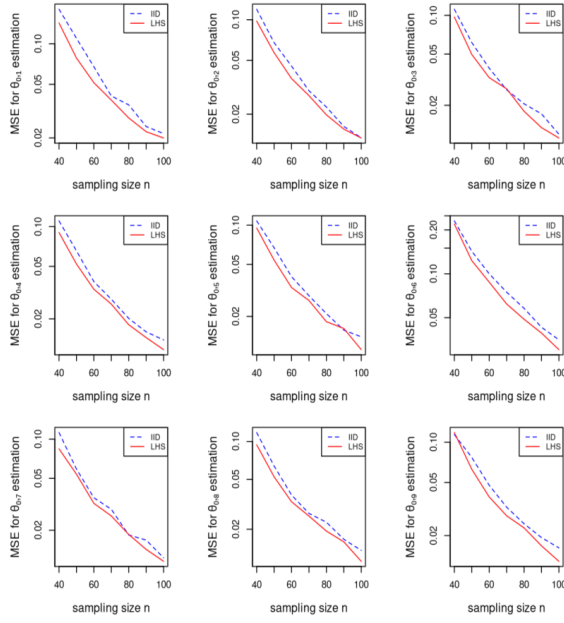


Figure 4 Average estimation MSE of $(\theta_{0,1}, \dots, \theta_{0,9})^T$ according to the sampling size n for IID and LHS designs (decimal logarithmic scale).

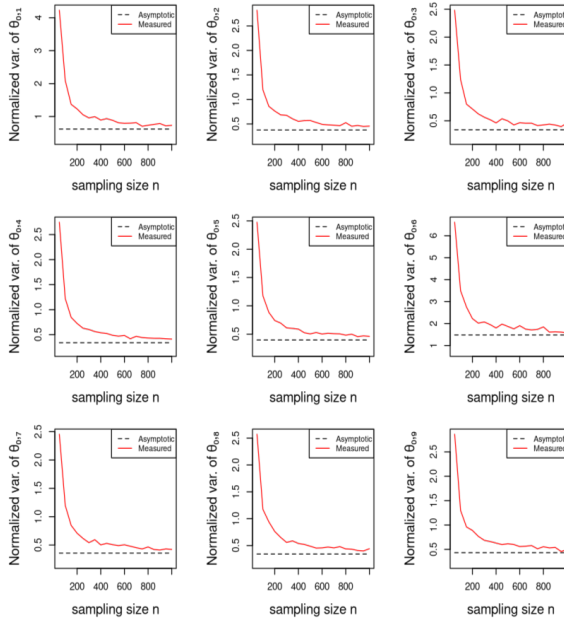


Figure 5 Evolution of measured normalized variances of $\theta_{0,1}, \dots, \theta_{0,9}$ according to the sampling size n for IID and LHS designs (decimal logarithmic scale).

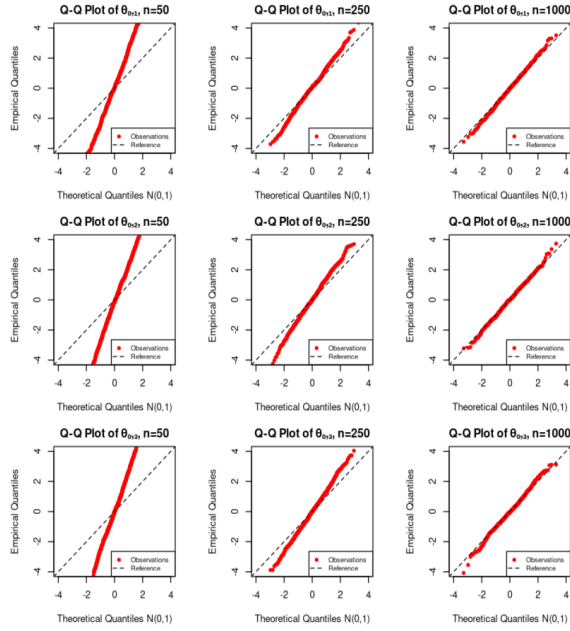


Figure 6 Q-Q plots of the first three parameter estimates $\theta_{0,1}$, $\theta_{0,2}$, and $\theta_{0,3}$ (red dots) compared to the standard normal distribution for sample sizes of 50, 250, and 1000.

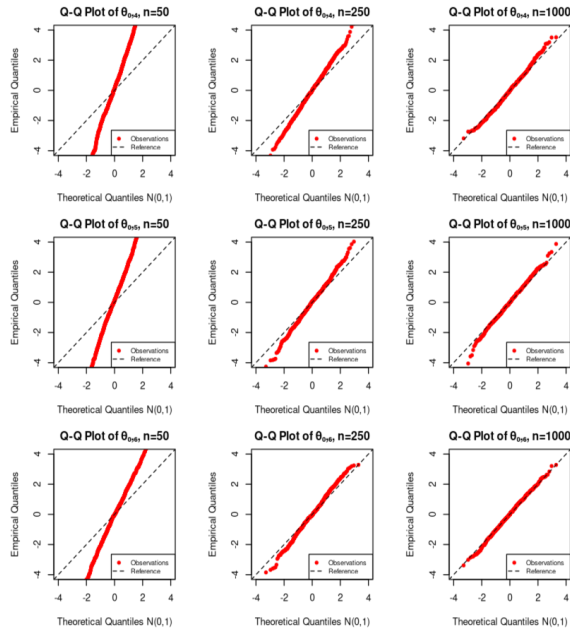


Figure 7 Q-Q plots of the parameter estimates $\theta_{0,4}$, $\theta_{0,5}$, and $\theta_{0,6}$ compared to the standard normal distribution for sample sizes of 50, 250, and 1000.

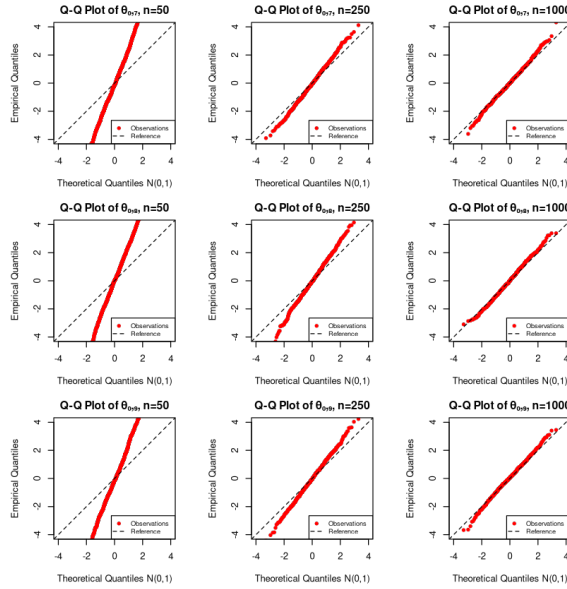


Figure 8 Q-Q plots of the parameter estimates $\theta_{0,7}$, $\theta_{0,8}$, and $\theta_{0,9}$ compared to the standard normal distribution for sample sizes of 50, 250, and 1000.

6 | CONCLUSION AND PROSPECTS

In conclusion, Latin Hypercube Sampling (LHS) has demonstrated its robustness as a powerful method for conducting computer experiments, particularly in the analysis of complex black-box functions. This paper has advanced the understanding of the asymptotic convergence of estimators using this sampling method. Specifically, we have extended the convergence results under LHS previously established in [9], [10], and [11] for the empirical mean to the broader class of Z -estimators. A key contribution of this work is the introduction of a Central Limit Theorem (CLT) for Z -estimators under this sampling method with a reduced asymptotic variance compared to traditional independent and identically distributed (i.i.d.) random sampling. Furthermore, we have illustrated the practical relevance of these theoretical findings through an application to parameter estimation in Generalized Linear Models (GLMs) under LHS. However, it is worth noting that certain restrictive regularity conditions were necessary to establish these convergence results.

A promising direction for future research is to relax some of the regularity assumptions, particularly the requirement for the second derivative of the Z -function. Alternative formulations of the CLT for Z -estimators that do not depend on the existence of a second derivative have been proposed, as discussed in [16]. Exploring these approaches for the convergence of Z -estimators under LHS could lead to more generalized results that extend beyond those presented in this work. Furthermore, generalizing the given convergence results under LHS to the class of M -estimators presents another valuable perspective for future research.

Another valuable extension of this research would be to adapt our theoretical framework to models with complex dependency structures, such as Gaussian Process (GP) regression [25, 26, 27], a key application of LHS. While GP re-

gression parameter estimation does not align with the Z -estimator framework due to its non-separable log-likelihood function, developing methodologies to address these dependencies could substantially expand the scope and applicability of our findings.

Ultimately, this work underscores the significant value of LHS in industrial applications, particularly for analyzing simulation codes that are computationally intensive and involve numerous input parameters. The versatility of LHS enables the efficient implementation of various statistical techniques—including variable selection, sensitivity analysis, and metamodeling—within a single numerical design of experiments. For further exploration of practical industrial applications of LHS, we refer readers to [13, 28, 29, 30, 31] as a few examples among many available in the literature.

Acknowledgments

Support from my supervisors (Claude Brayer, Fabrice Gamboa, Benoît Habert and Amandine Marrel) during my PhD thesis, of which this work is a continuation, is gratefully acknowledged. I would also like to thank Anouar Meynaoui for his help in writing this paper.

references

- [1] Mckay M, Beckman R, Conover W. A Comparison of Three Methods for Selecting Vales of Input Variables in the Analysis of Output From a Computer Code 21. *Technometrics* 1979;p. 239–245.
- [2] Helton JC, Davis FJ. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety* 81 2003;p. 23–69.
- [3] Helton JC, Davis FJ, Johnson JD. A comparison of uncertainty and sensitivity analysis results obtained with random and Latin hypercube sampling. *Reliability Engineering & System Safety* 89 2005;p. 305–330.
- [4] Viana FA. A tutorial on Latin hypercube design of experiments. *Quality and reliability engineering international* 2016;32(5):1975–1985.
- [5] Deutsch JL, Deutsch CV. Latin hypercube sampling with multidimensional uniformity. *Journal of Statistical Planning and Inference* 2012;142(3):763–772.
- [6] Shields MD, Zhang J. The generalization of Latin hypercube sampling. *Reliability Engineering & System Safety* 2016;148:96–108.
- [7] Packham N, Schmidt WM. Latin hypercube sampling with dependence and applications in finance. Available at SSRN 1269633 2008;.
- [8] Mondal A, Mandal A. Stratified random sampling for dependent inputs in Monte Carlo simulations from computer experiments. *Journal of Statistical Planning and Inference* 2020;205:269–282.
- [9] Stein M. Large Sample Properties of Simulations Using Latin Hypercube Sampling. *Technometrics* 2 1987;p. 143–151.
- [10] Owen AB. A Central Limit Theorem for Latin Hypercube Sampling. *Journal of the Royal Statistical Society* 54: Series B (Methodological) 1992;p. 541–551.
- [11] Loh WL. On latin hypercube sampling. *Annals of Statistics* 24 1996;p. 2058–2080.
- [12] Aistleitner C, Hofer M, Tichy RF. A central limit theorem for Latin hypercube sampling with dependence and application to exotic basket option pricing. *International Journal of Theoretical and Applied Finance* 2013;15(07).

- [13] Hakimi F. High-dimensional sensitivity analysis methods for computationally expensive simulators modeling a severe nuclear accident. PhD thesis, Paul Sabatier University, Toulouse, France; 2023.
- [14] Devroye L. 2: General principles in random variate generation. In: *Non-Uniform Random Variate Generation*(originally published with Springer-Verlag; 1986. p. 27–39.
- [15] Rousseeuw PJ, Hampel FR, Ronchetti EM, Stahel WA. *Robust Statistics: The Approach Based on Influence Functions*. Wiley; 1986.
- [16] Van der Vaart AW. *Asymptotic statistics*. Cambridge University Press; 1998.
- [17] Bouzebda S, Chaouch M. Uniform limit theorems for a class of conditional Z-estimators when covariates are functions. *Journal of Multivariate Analysis* 2022;189:104872.
- [18] Zhan Y. Central limit theorems for functional Z-estimators. *Statistica Sinica* 2002;12:609–634.
- [19] Hall AR. Generalized method of moments. *A companion to theoretical econometrics* 2003;p. 230–255.
- [20] Dacunha-Castelle D, Duflo M. *Probability and Statistics Volume II*. Springer-Verlag; 1986.
- [21] Fisher RA. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London, Series A* 222;p. 309–368.
- [22] Fischer H. *A History of the Central Limit Theorem: From Classical to Modern Probability Theory*. New York: Springer; 2011.
- [23] Nelder J, Wedderburn R. Generalized Linear Models. *Journal of the Royal Statistical Society Series A (General)* Blackwell Publishing 135;p. 370–384.
- [24] Allendoerfer CB. *Calculus of Several Variables and Differentiable Manifolds*. New York: Macmillan; 1974.
- [25] Williams C, Rasmussen C. Gaussian processes for regression. *Advances in neural information processing systems* 1995;8.
- [26] Kang F, Han S, Salgado R, Li J. System probabilistic stability analysis of soil slopes using Gaussian process regression with Latin hypercube sampling. *Computers and geotechnics* 2015;63:13–25.
- [27] Marrel A, Iooss B. Probabilistic surrogate modeling by Gaussian process: A review on recent insights in estimation and validation. *Reliability Engineering & System Safety* 2024;247:110094.
- [28] Karolczuk A, Kurek M. Fatigue life uncertainty prediction using the Monte Carlo and Latin hypercube sampling techniques under uniaxial and multiaxial cyclic loading. *International Journal of Fatigue* 2022;160:106867.
- [29] Bourcet J, Kubilay A, Derome D, Carmeliet J. Representative meteorological data for long-term wind-driven rain obtained from Latin Hypercube Sampling – Application to impact analysis of climate change. *Building and Environment* 2023;228:109875.
- [30] Ferrari R, Froio D, Rizzi E, Gentile C, Chatzi EN. Model updating of a historic concrete bridge by sensitivity- and global optimization-based Latin Hypercube Sampling. *Engineering Structures* 2019;179:139–160.
- [31] Marrel A, Iooss B, Chabridon V. The icscream methodology: Identification of penalizing configurations in computer experiments using screening and metamodel—applications in thermal hydraulics. *Nuclear Science and Engineering* 2022;196(3):301–321.