



HAL
open science

Multivariate sensitivity analysis: Minimum variance unbiased estimators of the first-order and total-effect covariance matrices

Matieyendou Lamboni

► **To cite this version:**

Matieyendou Lamboni. Multivariate sensitivity analysis: Minimum variance unbiased estimators of the first-order and total-effect covariance matrices. *Reliability Engineering and System Safety*, 2019, 187, pp.67 - 92. 10.1016/j.ress.2018.06.004 . hal-03484941

HAL Id: hal-03484941

<https://hal.science/hal-03484941>

Submitted on 20 Dec 2021

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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Multivariate sensitivity analysis: minimum variance unbiased estimators of the first-order and total-effect covariance matrices

Matieyendou Lamboni^{a,b}

^aUniversity of Guyane, Department DFRST, 97346 Cayenne, French Guiana, France

^b228-UMR Espace-Dev, University of Guyane, University of Réunion, University of Montpellier, IRD.

Abstract

In uncertainty quantification, multivariate sensitivity analysis (MSA) extends variance-based sensitivity analysis to cope with the multivariate response, and it aims to apportion the variability of the multivariate response into input factors and their interactions. The first-order and total-effect covariance matrices from MSA, which assess the effects of input factors, provide useful information about interactions among input factors, the order of interactions, and the magnitude of interactions over all model outputs. **In this paper, first, we propose and study generalized sensitivity indices (GSIs)** using the first-order and total-effect covariance matrices. The new GSIs make use of matrix norms when partial orders such as the Loewner ordering on covariance matrices is not possible, and we obtain the classical GSIs using the Frobenius norm. Second, we propose minimum variance unbiased estimators (MVUEs) of the first-order and total-effect covariance matrices, and third, we provide an efficient estimator of the first-order and total (classical) GSIs. We also derive the consistency, the asymptotic normality, and the asymptotic confidence regions of these estimators. Our estimator allows for improving the GSIs estimates.

Keywords: Generalized sensitivity indices, Matrix norms, Multivariate sensitivity analysis, MVU estimators, U-statistics.

1. Introduction

Comprehensive, mathematical, and multivariate response models (such as dynamic models) are widely used as experimental tools for supporting decision making in natural or human-induced phenomena. They often include numerous uncertain input factors. These uncertainties can strongly affect the model output(s) and pose certain difficulties when building scenarios. Assessing the effects of input factors on

1 the whole model outputs is interesting for practitioners prior to scenario-building,
2 model-reducing, or model diagnostic activities.

3

4 Multivariate sensitivity analysis ([1; 2; 3; 4; 5; 6; 7]) extends variance-based sen-
5 sitivity analysis ([8; 9; 10]) to cope with the multivariate response. It provides the
6 generalized sensitivity indices, which allow for assessing interactions among input
7 factors, the order of interactions, and the magnitude of interactions over all model
8 outputs.

9 So far, we distinguish mainly two approaches of MSA. The first approach is based
10 on the classical, multivariate analysis of variance, and it is well suited for factorial
11 designs ([1; 4; 3; 2]). The second approach is based on the functional, multivariate
12 analysis of variance by sampling input values randomly or quasi-randomly ([2; 5; 7]).
13 For both MSA approaches, the GSIs (first-order and total indices) can be computed
14 using either the original model outputs or a transformation of the latter (e.g., princi-
15 pal components), thereby showing the existence of an invariant property of the GSIs
16 ([4; 5]). The transformation becomes interesting when quantifying the sensitivity in
17 a given direction, defined by eigenvectors or another orthogonal basis ([1]).

18

19 Recently, MSA has been adapted to spatial outputs ([11]), and it has been com-
20 pared to an extension of MSA using the Hilbert-Schmidt independence criterion
21 ([11]). The GSIs from MSA have also been compared to new types of GSIs for mul-
22 tivariate outputs based on a vector projection or direction of eigen space [12; 7]).
23 For complex systems or expensive models, a meta-modeling or emulator approach
24 is an attractive way to compute GSIs ([13; 12]). However, in the following text, we
25 focus on sampling-based methods for estimating GSIs.

26

27 **The estimator of Sobol' indices has been improved in the case of single response**
28 **models.** It is known in [14; 15] that the new estimator of the first-order and total

1 Sobol' indices is based on a minimum variance unbiased estimator of the first-order
2 and total-effect variances thanks to the theory of U-statistics. In the case of multi-
3 variate response models, the definitions of all types of GSIs make use of covariance
4 matrices (first-order and total-effect covariance matrices), including the definitions
5 proposed in [5], the aggregated Sobol' indices used in [2], and those from [12; 7].
6 While the former definitions aims at assessing the influence of input factors on the
7 trace of the covariance matrices (magnitudes), the latter accounts for both the mag-
8 nitudes and the directions simultaneously. Thus, it is worth interesting to use all
9 valuable information contained in the first-order and total-effect covariance matrices
10 to assess the influence of a given input factor, and to better estimate these covari-
11 ance matrices, including the off-diagonal elements.

12

13 **In this paper, first, we propose and study new types of GSI using the first-order**
14 **and** total-effect covariance matrices. The new GSIs make use of matrix norms when
15 partial orders such as the Loewner ordering on these covariance matrices does not
16 exist, and we obtain the classical GSIs (from papers [2; 5]) using the Frobenius norm.
17 Second, we propose a minimum variance unbiased estimator of the first-order and
18 total-effect covariance matrices, and third, we provide an efficient estimator of the
19 classical first-order and total GSIs. We also derive the consistency, the asymptotic
20 normality, and the asymptotic confidence regions of these estimators.

21

22 **The paper is organized as follows:** in Section 2, we recall the multivariate Hoeffd-
23 ing decomposition, and we propose the re-organized decomposition that allows for
24 assessing the overall contribution of input factors through the total-effect function.
25 While Section 3 proposes new definitions of GSIs using some matrix norms, Section
26 4 provides and studies the estimators of the first-order and total-effect covariance
27 matrices, including the joint estimator of the classical GSIs (from [2; 5]). The joint
28 estimator makes use of a function with two types of inputs (kernel of degree (p, q)).

1 In Section 5, we use a proxy measure for choosing the convenient degree of the kernel
 2 (p, q) , and we provide an algorithm for computing the two covariance matrices. We
 3 demonstrate the applicability of these results by means of numerical tests in Section
 4 6. Section 7 concludes this work.

5 Notation and preliminary

6 Throughout the paper, $\mathbf{f}(\mathbf{X})$ denotes a multivariate response model (i.e., $\mathbf{f}(\mathbf{X}) \in$
 7 \mathbb{R}^N) with d random input factors $\mathbf{X} = (X_1, \dots, X_d)$. We use u as a non-empty
 8 subset of $\{1, 2, \dots, d\}$, and we use $|u|$ for its cardinality (i.e., the number of ele-
 9 ments in u). Let $\mathbf{X}_u = \{X_j, j \in u\}$ be a subset of input factors and $\mathbf{X}_{\sim u}$ denote
 10 the vector containing all input factors except \mathbf{X}_u . We have the following partition:
 11 $\mathbf{X} = (\mathbf{X}_u, \mathbf{X}_{\sim u})$. We use $\mu(\mathbf{X}) = \mu(X_1, \dots, X_d)$ as the joint probability measure or
 12 distribution of input factors. We assume throughout the paper that input factors
 13 are independent (assumption A1) and all model outputs have finite second moments
 14 (assumption A2); that is, $\mathbb{E} [\|\mathbf{f}(\mathbf{X})\|_{L^2}^2] < +\infty$. We use $\mathbb{P}(\cdot)$ for the probability,
 15 $\mathbb{E}(\cdot)$ for the expectation, $\mathbb{V}(\cdot)$ for the variance, $\mathbb{T}_r(\cdot)$ for the trace, $\mathbb{C}_{ov}(\cdot)$ for the
 16 covariance, $\binom{n}{p}$ for the number of combinations for selecting p objects out of n , and
 17 \xrightarrow{D} and \xrightarrow{P} for the convergence in distribution and in probability respectively.

18

For a matrix Σ , the Frobenius norm of Σ is given by

$$\|\Sigma\|_F^2 = \sum_{i,j} |\Sigma_{ij}|^2 = \mathbb{T}_r(\Sigma\Sigma^T) .$$

The spectral norm of Σ is the square root of the largest eigenvalue of $\Sigma^T\Sigma$, that
 is,

$$\|\Sigma\|_S^2 = \max \{ \lambda, \det(\Sigma^T\Sigma - \lambda\mathcal{I}) = 0 \} .$$

The l_p norm is defined as follows:

$$\|\Sigma\|_{l_p} = \left(\sum_{i,j} |\Sigma_{ij}|^p \right)^{1/p},$$

1 and $\|\Sigma\|_{l_p} = \|\Sigma\|_F$ if $p = 2$.

2

Let Σ_1 and Σ_2 be two symmetric and positive semi-definite matrices (i.e., $\mathbf{v}^T \Sigma_i \mathbf{v} \geq 0 \forall \mathbf{v} \in \mathbb{R}^N$, $i = 1, 2$). The Loewner ordering on matrices Σ_1 and Σ_2 is defined as follows:

$$\Sigma_1 \preceq \Sigma_2 \quad \text{if } \Sigma_2 - \Sigma_1 \text{ is positive semi-definite.}$$

3 2. Multivariate Hoeffding decomposition

4 **A transformation may be applied to the model outputs** $\mathbf{f}(\mathbf{X})$, of the form $D_{\mathbf{w}}\mathbf{f}(\mathbf{X})$,
 5 where $D_{\mathbf{w}}$ is a $N \times N$ weighting matrix. For instance, in the case of the model outputs
 6 with different units, some transformations may be used to obtain unit-less outputs.

7 **Remark 1.** *A classical way to obtain unit-less outputs is to divide each output by its*
 8 *standard deviation, that is, $D_{\mathbf{w}}^{-1} = \text{diag}(\mathbf{w})$ is a diagonal matrix with \mathbf{w} representing*
 9 *the vector of the standard deviations of the outputs. One may use \mathbf{w} as a vector*
 10 *containing the expectation of the absolute value of each output ([12]). In the following*
 11 *text, we use $\mathbf{f}(\mathbf{X})$ as either the original model outputs or a given transformation of*
 12 *the latter.*

13 **Under the independence assumption** A1, the multivariate Hoeffding decomposi-
 14 tion ([16]; [17]) is given by

$$\begin{aligned} \mathbf{f}(\mathbf{X}) &= \mathbf{f}_\emptyset + \sum_j^d \mathbf{f}_j(X_j) + \sum_{j_1 < j_2}^d \mathbf{f}_{j_1 j_2}(X_{j_1}, X_{j_2}) + \dots + \mathbf{f}_{1\dots d}(X_1, \dots, X_d) \\ &= \mathbf{f}_\emptyset + \sum_{u \subseteq \{1, 2, \dots, d\}} \mathbf{f}_u(\mathbf{X}_u), \end{aligned} \quad (2.1)$$

15 where $\mathbf{f}_\emptyset = \mathbb{E}[\mathbf{f}(\mathbf{X})]$ is the expectation of the model output, $\mathbf{f}_j(X_j) = \mathbb{E}[\mathbf{f}(\mathbf{X})|X_j] - \mathbf{f}_\emptyset$,

16 and $\mathbf{f}_u(\mathbf{X}_u) = \mathbb{E}[\mathbf{f}(\mathbf{X})|\mathbf{X}_u] - \sum_{\substack{v \subseteq \{1, 2, \dots, d\} \\ v \subset u}} \mathbf{f}_v(\mathbf{X}_v)$ for a non-empty subset $u \subseteq \{1, 2, \dots, d\}$.

17

1 The function $\mathbf{f}_j(X_j) = \mathbb{E}[\mathbf{f}(\mathbf{X})|X_j] - \mathbf{f}_\emptyset$ in (2.1) contains the information pro-
 2 vided by X_j only to the model outputs. It is used to quantify the single con-
 3 tribution of X_j over model outputs. Moreover, we have the following property:
 4 $\mathbb{E}_{X_j}[\mathbf{f}_u(\mathbf{X}_u)] = \mathbf{0}, \forall j \in u$. In general, $\mathbb{E}[\mathbf{f}(\mathbf{X})|\mathbf{X}_u] - \mathbf{f}_\emptyset$ allows for quantifying the
 5 single contribution of the input factors \mathbf{X}_u over model outputs. We refer to this
 6 function as the first-order function.

7
 8 The Hoeffding decomposition in (2.1) can be organized as follows ([18; 19]):

$$\mathbf{f}(\mathbf{X}) = \mathbf{f}_\emptyset + \mathbf{g}_u(\mathbf{X}_u, \mathbf{X}_{\sim u}) + \mathbf{h}_{\sim u}(\mathbf{X}_{\sim u}), \quad (2.2)$$

9 with $\mathbf{g}_u(\mathbf{X}_u, \mathbf{X}_{\sim u}) = \sum_{\substack{v \subseteq \{1, 2, \dots, d\} \\ v \cap u \neq \emptyset}} \mathbf{f}_v(\mathbf{X}_v)$.

10

11 The decomposition in (2.2) is also unique under the independence assumption
 12 A1. By definition, the function $\mathbf{g}_u(\cdot)$ in (2.2) contains all information brought by
 13 the input factors \mathbf{X}_u to the model outputs, and it is sufficient to quantify the overall
 14 contribution of \mathbf{X}_u over model outputs. We refer to this function as the total-effect
 15 function (TEF).

16

17 Based on the re-organized Hoeffding decomposition in (2.2), we can show that

$$\mathbf{g}_u(\mathbf{X}_u, \mathbf{X}_{\sim u}) = \mathbf{f}(\mathbf{X}) - \mathbb{E}[\mathbf{f}(\mathbf{X})|\mathbf{X}_{\sim u}]. \quad (2.3)$$

18 **Indeed, since** $\mathbb{E}[\mathbf{g}_u(\mathbf{X}_u, \mathbf{X}_{\sim u})|\mathbf{X}_{\sim u}] = \mathbf{0}$, Equation (2.2) implies $\mathbb{E}[\mathbf{f}(\mathbf{X})|\mathbf{X}_{\sim u}] =$
 19 $\mathbf{f}_\emptyset + \mathbf{h}_{\sim u}(\mathbf{X}_{\sim u})$, and gives (2.3).

20

21 **Remark 2.** Equation (2.3) generalizes the total-effect function for the single re-
 22 sponse proposed in ([19; 14]) to the multivariate response.

23 Any definition of the first-order and total GSIs of \mathbf{X}_u should be based on the

1 random vectors $\mathbb{E}[\mathbf{f}(\mathbf{X})|X_u] - \mathbf{f}_\emptyset$ and $\mathbf{g}_u(\mathbf{X}_u, \mathbf{X}_{\sim u})$, respectively. We often measure
 2 the variability of a random vector by its variance. By taking the covariance of
 3 Equation (2.1) and bearing in mind the L^2 orthogonality, we have the following
 4 covariance decomposition assuming that assumptions A1 and A2 hold.

$$\Sigma = \sum_{v \subseteq \{1, 2, \dots, d\}} \mathbf{D}_v, \quad (2.4)$$

5 with $\Sigma = \mathbb{C}_{ov}[\mathbf{f}(\mathbf{X})]$ and $\mathbf{D}_v = \mathbb{C}_{ov}[\mathbf{f}_v(\mathbf{X}_v)]$.

6

7 **3. First-order and total-effect covariance matrices: new GSIs**

8 Using the variance measure, the definition of the sensitivity indices for the mul-
 9 tivariate response should be based on the following covariance matrices.

10

11 The first-order covariance matrix is given by

$$\Sigma_u = \mathbb{C}_{ov}[\mathbb{E}(\mathbf{f}(\mathbf{X})|X_u)]. \quad (3.5)$$

12 Further, the total-effect covariance matrix is given by

$$\Sigma_u^{tot} = \mathbb{C}_{ov}[\mathbf{g}_u(\mathbf{X}_u, \mathbf{X}_{\sim u})]. \quad (3.6)$$

13 Using Equations (2.2 , 2.3, 2.4, 3.6) and bearing in mind the L^2 orthogonality,
 14 we can see that the total-effect covariance matrix is also given by ([3; 2; 7])

$$\Sigma_u^{tot} = \sum_{\substack{v \subseteq \{1, 2, \dots, d\} \\ v \cap u \neq \emptyset}} \mathbf{D}_v. \quad (3.7)$$

15 Regarding the model outputs, the first-order and total-effect covariance matrices
 16 of \mathbf{X}_u provide all information about \mathbf{X}_u , including the parts of correlations among

1 model outputs. The covariance matrices Σ_u and Σ_u^{tot} are symmetric and positive
 2 semi-definite (i.e., $\mathbf{v}^T \Sigma_u \mathbf{v} \geq 0$ and $\mathbf{v}^T \Sigma_u^{tot} \mathbf{v} \geq 0 \forall \mathbf{v} \in \mathbb{R}^N$). Moreover, we have the
 3 following properties using the Loewner ordering

$$\Sigma_u \preceq \Sigma_u^{tot}. \quad (3.8)$$

4 **Indeed,** $\Sigma_u^{tot} - \Sigma_u = \sum_{\substack{v \subseteq \{1, 2, \dots, d\} \\ u \subset v}} \mathbf{D}_v$, with \mathbf{D}_v positive semi-definite. Thus, (3.8)
 5 generalizes the inequality between the first-order and total effects in the case of mul-
 6 tivariate response models. As Σ_u^{tot} accounts for interactions, it should be preferred for
 7 ranking input factors. In the same sense, when the Loewner ordering exists for ma-
 8 trices Σ_j^{tot} , $j = 1, 2, \dots, d$, that is, $\Sigma_{j_1}^{tot} \preceq \Sigma_{j_2}^{tot} \preceq \dots \preceq \Sigma_{j_d}^{tot}$ with $j_i \in \{1, 2, \dots, d\}$
 9 and $i = 1, 2, \dots, d$, the input factor X_{j_d} will be the most influential input factor,
 10 followed by $X_{j_{d-1}}$.

11

12 **The prioritization of input factors based on the covariance matrices is** straightforward
 13 for the single response ($N = 1$), as the covariance matrices are scalars. In the case of
 14 the multivariate response, we should apply a matrix norm on the covariance matrices
 15 $(\Sigma_u, \Sigma_u^{tot})$ prior to prioritize input factors when the Loewner ordering on matrices
 16 Σ_u and Σ_u^{tot} is not possible. Thus, new definitions of the sensitivity indices for the
 17 multivariate response become possible regarding the choice of the matrix norm. In
 18 the following text, we choose three matrix norms that are complementary.

19 3.1. **Classical generalized sensitivity indices and Frobenius norm**

20 A classical extension of the variance to a random vector is the expectation of the
 21 L^2 norm of the centered vector. In previous works ([3; 2; 5]), the definitions of the
 22 GSIs rely on this measure, and the harmonized definitions are given below.

23

24 **Definition 1.** Let $\mathbf{f}(\mathbf{X}) \in \mathbb{R}^N$ be the model outputs with d random inputs \mathbf{X} and Σ ,
 25 Σ_u , and Σ_u^{tot} be the covariance matrices of the model outputs, the first-order and the

1 total-effect functions, respectively.

2

3 The first-order GSI of \mathbf{X}_u is defined as follows:

$$GSI_u = \frac{\mathbb{T}_r(\Sigma_u)}{\mathbb{T}_r(\Sigma)}. \quad (3.9)$$

4 Further, the total GSI of \mathbf{X}_u is given by

$$GSI_{T_u} = \frac{\mathbb{T}_r(\Sigma_u^{tot})}{\mathbb{T}_r(\Sigma)}. \quad (3.10)$$

5 By using the Frobenius norm, the GSIs definitions in (3.9) and (3.10) can be
6 expressed as follows:

$$GSI_u^F = \frac{\left\| \left\| \Sigma_u^{1/2} \right\|_F^2 \right\|_F^2}{\left\| \left\| \Sigma^{1/2} \right\|_F^2 \right\|_F^2}, \quad (3.11)$$

7 and

$$GSI_{T_u}^F = \frac{\left\| \left\| (\Sigma_u^{tot})^{1/2} \right\|_F^2 \right\|_F^2}{\left\| \left\| \Sigma^{1/2} \right\|_F^2 \right\|_F^2}. \quad (3.12)$$

8 **Remark 3.** Let \mathcal{V} be an $N \times N$ orthogonal matrix (i.e., $\mathcal{V}^T \mathcal{V} = \mathcal{I}_k$), the first-order
9 and total GSIs of \mathbf{X}_u are invariant by left-composing $\mathbf{f}(\mathbf{X})$ by \mathcal{V} , that is,

$$GSI_u^F(\mathcal{V}\mathbf{f}(\mathbf{X})) = \frac{\mathbb{T}_r[\mathbf{C}_{ov}(\mathcal{V}\mathbf{E}[\mathbf{f}(\mathbf{X})|\mathbf{X}_u])]}{\mathbb{T}_r[\mathbf{C}_{ov}(\mathcal{V}\mathbf{f}(\mathbf{X}))]} = GSI_u, \quad (3.13)$$

10 and

$$GSI_{T_u}^F(\mathcal{V}\mathbf{f}(\mathbf{X})) = \frac{\mathbb{T}_r[\mathbf{C}_{ov}(\mathcal{V}[\mathbf{g}_u(\mathbf{X}_u, \mathbf{X}_{\sim u})])]}{\mathbb{T}_r[\mathbf{C}_{ov}(\mathcal{V}\mathbf{f}(\mathbf{X}))]} = GSI_{T_u}. \quad (3.14)$$

11 **When** \mathcal{V} contains the eigenvectors of Σ , the definitions in (3.13) and (3.14) come
12 down to an aggregation of the Sobol indices across all principal components ([4; 3;
13 2]).

14 The definitions in (3.11) and (3.12) include only the diagonal elements of the
15 covariance matrices. Thus, the GSIs provide the contributions of input factors to
16 the sum of the variance of the model outputs (called inertia).

17 3.2. New generalized sensitivity indices based on the spectral norm

18 **The spectral norm can bring some information about the correlations between**
19 **the** components of the total-effect function $\mathbf{g}_u(\mathbf{X})$ for instance. While strong correla-

1 tions between the components of $\mathbf{g}_u(\mathbf{X})$ will lead to a high value of the spectral norm
2 of $(\Sigma_u^{tot})^{1/2}$ (i.e., $\left\|(\Sigma_u^{tot})^{1/2}\right\|_S$), uncorrelated components will reduce $\left\|(\Sigma_u^{tot})^{1/2}\right\|_S^2$ to
3 the highest value of the variances of these components. Based on the spectral norm,
4 we define the new generalized sensitivity indices as follows.

5

6 **Definition 2.** Let $\mathbf{f}(\mathbf{X}) \in \mathbb{R}^N$ be the model outputs with d random inputs \mathbf{X} and
7 Σ , Σ_u , and Σ_u^{tot} be the covariance matrices of the model outputs, the first-order and
8 total-effect functions, respectively.

9

10 The first-order GSI of \mathbf{X}_u is defined as follows:

$$GSI_u^S = \frac{\left\|\Sigma_u^{1/2}\right\|_S^2}{\left\|\Sigma^{1/2}\right\|_F^2}. \quad (3.15)$$

11 Further, the total GSI of \mathbf{X}_u is given by

$$GSI_{T_u}^S = \frac{\left\|(\Sigma_u^{tot})^{1/2}\right\|_S^2}{\left\|\Sigma^{1/2}\right\|_F^2}. \quad (3.16)$$

12 In Definition 2, $\left\|\Sigma^{1/2}\right\|_F^2$ is used as a constant of normalization, and it allows
13 the derivation of the following properties of the new indices.

14

15 **Proposition 1.** Let $\mathbf{f}(\mathbf{X}) \in \mathbb{R}^N$ be the model outputs with d random inputs \mathbf{X} , \mathcal{V}
16 be an $N \times N$ orthogonal matrix. If assumptions A1 and A2 hold, then

17

18 i) the indices GSI_u^S and $GSI_{T_u}^S$ satisfy

$$0 \leq GSI_u^S \leq GSI_{T_u}^S \leq 1; \quad (3.17)$$

19 ii) the indices GSI_u^S and $GSI_{T_u}^S$ are invariant by left-composing $\mathbf{f}(\mathbf{X})$ by \mathcal{V} , that
20 is,

$$GSI_u^S(\mathcal{V}\mathbf{f}(\mathbf{X})) = GSI_u^S(\mathbf{f}(\mathbf{X})), \quad (3.18)$$

21 and

$$GSI_{T_u}^S(\mathcal{V}\mathbf{f}(\mathbf{X})) = GSI_{T_u}^S(\mathbf{f}(\mathbf{X})). \quad (3.19)$$

22 **Proof.** See Appendix A.

23

□

1 3.3. New generalized sensitivity indices based on the l_p -norm

2 **The** l_p -norm of Σ_u^{tot} ($\|\Sigma_u^{tot}\|_{l_p}$) accounts for all elements of Σ_u^{tot} , including the
3 off-diagonal elements. It includes the correlations between the components of the
4 total-effect function $\mathbf{g}_u(\mathbf{X})$ in the analysis. Thus, strong correlations between the
5 components of $\mathbf{g}_u(\mathbf{X})$ will increase the value of $\|\Sigma_u^{tot}\|_{l_p}$.

6
7 Based on the l_p -norm, we define the new generalized sensitivity indices below.

8
9 **Definition 3.** Let $\mathbf{f}(\mathbf{X}) \in \mathbb{R}^N$ be the model outputs with d random inputs \mathbf{X} and
10 Σ , Σ_u , and Σ_u^{tot} be the covariance matrices of the model outputs, the first-order and
11 total-effect functions, respectively.

12
13 The first-order GSI of \mathbf{X}_u is defined as follows:

$$GSI_u^{l_p} = \frac{\|\Sigma_u\|_{l_p}}{N \|\Sigma^{1/2}\|_F^2}. \quad (3.20)$$

14 Further, the total GSI of \mathbf{X}_u is given by

$$GSI_{T_u}^{l_p} = \frac{\|\Sigma_u^{tot}\|_{l_p}}{N \|\Sigma^{1/2}\|_F^2}. \quad (3.21)$$

15 In Definition 3, $N \|\Sigma^{1/2}\|_F^2$ is used as a constant of normalization, and it allows
16 the derivation of the following properties of the new indices.

17
18 **Proposition 2.** Let $\mathbf{f}(\mathbf{X}) \in \mathbb{R}^N$ be the model outputs with d random inputs \mathbf{X} . If
19 assumptions A1 and A2 hold, then the indices $GSI_u^{l_1}$ and $GSI_{T_u}^{l_1}$ satisfy

$$0 \leq GSI_u^{l_1} \leq 1 \quad \text{and} \quad 0 \leq GSI_{T_u}^{l_1} \leq 1. \quad (3.22)$$

20 **Proof.** See Appendix B.

21 □

22 3.4. Comparison of the three types of the generalized sensitivity indices

23 The three types of the GSIs provide different information about input factors,
24 and it is possible to come up with different ranking of input factors. However, we
25 have the following relationship between the three types of the GSIs.

$$GSI_u^S \leq GSI_u^F \leq NGSI_u^{l_1}. \quad (3.23)$$

1 Likewise, we have

$$GSI_{T_u}^S \leq GSI_{T_u}^F \leq NGSI_{T_u}^{l_1}. \quad (3.24)$$

2 It comes out that $GSI_{T_u}^F = NGSI_{T_u}^{l_1}$ in the presence of uncorrelated components
 3 of the total-effect function $\mathbf{g}_u(\mathbf{X})$, and $GSI_{T_u}^S = GSI_{T_u}^F$ for some correlations between
 4 the components of $\mathbf{g}_u(\mathbf{X})$.

5 **4. Estimators of the first-order and total-effect covariance matrices and** 6 **their traces**

7 The definitions of GSIs are based on the first-order and total-effect covariance
 8 matrices. In this section, we construct the estimators of the first-order and total-
 9 effect covariance matrices Σ_u and Σ_u^{tot} , and we investigate their performance using
 10 the theory of U-statistics.

11

12 **The theory of U-statistics allows for deriving the properties of an estimator** ([20],
 13 [21] [22], [16; 23]). However, it requires, first, to find a kernel (i.e., a function
 14 with expectation coinciding exactly with our parameter of interest), and second, to
 15 propose the estimator of our parameter using this kernel.

16 *4.1. Kernel functions for the first-order and total-effect covariance matrices*

17 For integers $p \geq 2, q \geq 2$, let $(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)})$ be p i.i.d copies of \mathbf{X}_u and
 18 $(\mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})$ be q i.i.d copies of $\mathbf{X}_{\sim u}$. We consider a function with two types
 19 of inputs $(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)})$ and $(\mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})$. This function is called a kernel of
 20 degree (p, q) in the theory of U-statistics; it has $q + p$ arguments.

21

22 To estimate the first-order and total-effect covariance matrices, we define the
 23 following kernels. The kernel $\mathbf{K}(\cdot)$ for the first-order covariance matrix is given by

$$\mathbf{K}(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)}) = \frac{1}{p^2(p-1)q(q-1)} \sum_{j=1}^p \sum_{l=1}^q \sum_{i \neq l}^q \times$$

$$\left(\sum_{\substack{j_1=1 \\ j_1 \neq j}}^p [\mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) - \mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)})] \right) \times \left(\sum_{\substack{j_2=1 \\ j_2 \neq j}}^p [\mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(i)}) - \mathbf{f}(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(i)})] \right)^T.$$

(4.25)

1 Further, the kernel $\mathbf{K}^{tot}(\cdot)$ for the total-effect covariance matrix is defined as
 2 follows:

$$\mathbf{K}^{tot}(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)}) = \frac{1}{p^2(p-1)q} \sum_{j=1}^p \sum_{l=1}^q \times$$

$$\left(\sum_{\substack{j_1=1 \\ j_1 \neq j}}^p [\mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) - \mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)})] \right) \times \left(\sum_{\substack{j_2=1 \\ j_2 \neq j}}^p [\mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) - f_s(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(l)})] \right)^T.$$

(4.26)

3 The kernels $\mathbf{K}(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})$ and $\mathbf{K}^{tot}(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})$
 4 are symmetric under independent permutations of the first arguments $(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)})$
 5 and the second arguments $(\mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})$. Indeed, the kernels values do not change
 6 if we permute the position of $\mathbf{X}_u^{(i_1)}$ and $\mathbf{X}_u^{(i_2)}$ in one hand, and the position of $\mathbf{X}_{\sim u}^{(j_1)}$
 7 and $\mathbf{X}_{\sim u}^{(j_2)}$ in the other hand, with $i_1, i_2 \in \{1, 2, \dots, p\}$ and $j_1, j_2 \in \{1, 2, \dots, q\}$.
 8 Theorem 1 gives other properties of these kernels.

9
 10 **Theorem 1.** Let $(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)})$ be p i.i.d copies of \mathbf{X}_u and $(\mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})$ be q
 11 i.i.d copies of $\mathbf{X}_{\sim u}$, with $p, q \geq 2$. If assumptions A1 (independent inputs) and A2
 12 $(\mathbb{E} [\|\mathbf{f}(\mathbf{X})\|_{L^2}^2] < +\infty)$ hold, then we have

$$\mathbb{E} [\mathbf{K}(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})] = \Sigma_u; \quad (4.27)$$

$$\mathbb{E} [\mathbf{K}^{tot} (\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})] = \Sigma_u^{tot}. \quad (4.28)$$

1 **Proof.** See Appendix C.

2

□

3 Theorem 1 shows that the kernels $\mathbf{K}(\cdot)$ (resp. $\mathbf{K}^{tot}(\cdot)$) is an unbiased estimator of Σ_u
 4 (resp. Σ_u^{tot}). Both kernels are symmetric, and we can use the theory of U-statistics
 5 of two samples to obtain the statistical properties of the estimators based on the
 6 kernels $\mathbf{K}(\cdot)$ and $\mathbf{K}^{tot}(\cdot)$.

7

8 **Remark 4.** In Theorem 1, q can take the value 1 ($q = 1$) in the case of the kernel
 9 $\mathbf{K}^{tot}(\cdot)$. However, for estimating both Σ_u and Σ_u^{tot} with the same number of model
 10 runs, we need to have $q \geq 2$.

11 4.2. Estimators of the first-order and total-effect covariance matrices

12 This section provides the MVU estimators of Σ_u and Σ_u^{tot} and the main theorems
 13 about the properties of these estimators. Theorems 2 and 3 deal with these issues.

14

15 **Theorem 2.** Let $\mathcal{X} = (\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)})$ be p i.i.d copies of \mathbf{X}_u , $\mathcal{Y} = (\mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})$
 16 be q i.i.d copies of $\mathbf{X}_{\sim u}$, $\mathcal{X}_i = (\mathbf{X}_{i,u}^{(1)}, \dots, \mathbf{X}_{i,u}^{(p)})$ and $\mathcal{Y}_i = (\mathbf{X}_{i,\sim u}^{(1)}, \dots, \mathbf{X}_{i,\sim u}^{(q)})$,
 17 $i = 1, 2, \dots, m$, be two independent samples of size m from \mathcal{X} and \mathcal{Y} respectively. If
 18 assumptions A1, A3 ($\mathbb{E} \|f(\mathbf{X})\|_{L^2}^4 < +\infty$), and A4 ($2 \leq p, 2 \leq q$) hold, then we have

19

20 i) the minimum variance unbiased estimator of Σ_u for a given (p, q) and m is
 21 given by

$$\widehat{\Sigma}_u = \frac{1}{mp^2(p-1)q(q-1)} \sum_{i=1}^m \sum_{j=1}^p \sum_{l=1}^q \sum_{k \neq l}^q \left(\sum_{\substack{j_1=1 \\ j_1 \neq j}}^p \left[\mathbf{f}(\mathbf{X}_{i,u}^{(j)}, \mathbf{X}_{i,\sim u}^{(l)}) - \mathbf{f}(\mathbf{X}_{i,u}^{(j_1)}, \mathbf{X}_{i,\sim u}^{(l)}) \right] \right) \\ \times \left(\sum_{\substack{j_2=1 \\ j_2 \neq j}}^p \left[\mathbf{f}(\mathbf{X}_{i,u}^{(j)}, \mathbf{X}_{i,\sim u}^{(k)}) - \mathbf{f}(\mathbf{X}_{i,u}^{(j_2)}, \mathbf{X}_{i,\sim u}^{(k)}) \right] \right)^T, \quad (4.29)$$

1 and we have

$$2 \quad \mathbb{E} \left(\widehat{\Sigma}_u \right) = \Sigma_u ; \quad (4.30)$$

3 ii) if $m \rightarrow +\infty$, $\widehat{\Sigma}_u$ is consistent, that is,

$$4 \quad \widehat{\Sigma}_u \xrightarrow{\mathcal{P}} \Sigma_u . \quad (4.31)$$

4 **Proof.** See Appendix D.

5 □

6 In Theorem 2, the estimator $\widehat{\Sigma}_u$ is an average of the kernel $\mathbf{K}(\cdot)$ over $i =$
 7 $1, 2, \dots, m$, with m the sample size. The kernel $\mathbf{K} \left(\mathbf{X}_{i,u}^{(1)}, \dots, \mathbf{X}_{i,u}^{(p)}, \mathbf{X}_{i,\sim u}^{(1)}, \dots, \mathbf{X}_{i,\sim u}^{(q)} \right)$
 8 requires p i.i.d copies of \mathbf{X}_u and q i.i.d copies of $\mathbf{X}_{\sim u}$. For different values of (p, q)
 9 and m , $\widehat{\Sigma}_u$ does not use the same information for computing Σ_u . Similar to Theorem
 10 2, Theorem 3 provides a MVUE of Σ_u^{tot} .

11
 12 **Theorem 3.** Let $\mathcal{X} = \left(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)} \right)$ be p i.i.d copies of \mathbf{X}_u , $\mathcal{Y} = \left(\mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)} \right)$
 13 be q i.i.d copies of $\mathbf{X}_{\sim u}$, $\mathcal{X}_i = \left(\mathbf{X}_{i,u}^{(1)}, \dots, \mathbf{X}_{i,u}^{(p)} \right)$ and $\mathcal{Y}_i = \left(\mathbf{X}_{i,\sim u}^{(1)}, \dots, \mathbf{X}_{i,\sim u}^{(q)} \right)$,
 14 $i = 1, 2, \dots, m$, be two independent samples of size m from \mathcal{X} and \mathcal{Y} respectively. If
 15 assumptions A1, A3 ($\mathbb{E} \|f(\mathbf{X})\|_{L^2}^4 < +\infty$), and A4 ($2 \leq p, 2 \leq q$) hold, then we have

16
 17 i) the minimum variance unbiased estimator of Σ_u^{tot} for a given (p, q) and m is
 18 given by

$$\begin{aligned} \widehat{\Sigma}_u^{tot} &= \frac{1}{mp^2(p-1)q} \sum_{i=1}^m \sum_{j=1}^p \sum_{l=1}^q \left(\sum_{\substack{j_1=1 \\ j_1 \neq j}}^p \left[\mathbf{f}(\mathbf{X}_{i,u}^{(j)}, \mathbf{X}_{i,\sim u}^{(l)}) - \mathbf{f}(\mathbf{X}_{i,u}^{(j_1)}, \mathbf{X}_{i,\sim u}^{(l)}) \right] \right) \\ &\quad \times \left(\sum_{\substack{j_2=1 \\ j_2 \neq j}}^p \left[\mathbf{f}(\mathbf{X}_{i,u}^{(j)}, \mathbf{X}_{i,\sim u}^{(l)}) - \mathbf{f}(\mathbf{X}_{i,u}^{(j_2)}, \mathbf{X}_{i,\sim u}^{(l)}) \right] \right)^T, \end{aligned} \quad (4.32)$$

19 and we have

$$20 \quad \mathbb{E} \left(\widehat{\Sigma}_u^{tot} \right) = \Sigma_u^{tot} ; \quad (4.33)$$

1 *ii) if $m \rightarrow +\infty$, $\widehat{\Sigma}_u$ is consistent, that is,*

$$\widehat{\Sigma}_u^{tot} \xrightarrow{\mathcal{P}} \Sigma_u^{tot}. \quad (4.34)$$

2 **Proof.** *See Appendix E.*

3 □

4 Theorems 2 and 3 provide interesting estimators of the first-order and total-effect
5 covariance matrices, and these estimators can be used to obtain the most influential
6 input factors when the Loewner ordering on these matrices exists. These estimators
7 are going to be used for estimating the three types of GSIs proposed in Section 3.

8 *4.3. Estimators of the classical generalized sensitivity indices*

9 In this section, we investigate the estimations of the classical GSIs defined in
10 [2; 5]. To make use of the theory of U-statistics in the case of the estimation of the
11 non-normalized first-order GSI, we define the following kernel.

$$K(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)}) = \mathbb{T}_r [\mathbf{K}(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})], \quad (4.35)$$

12 with $\mathbf{K}(\cdot)$ the kernel defined in (4.25).

13

14 It is obvious that the kernel $K(\cdot)$ is symmetric with respect to its first and sec-
15 ond arguments. As the kernel $K(\cdot)$ is the sum of the diagonal elements of $\mathbf{K}(\cdot)$ by
16 definition, it is equal to the following kernel.

17

$$\begin{aligned}
K(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)}) &= \frac{2}{p^2(p-1)q(q-1)} \sum_{j=1}^p \sum_{l=1}^{q-1} \sum_{i=l+1}^q \times \\
\mathbb{T}_r \left[\left(\sum_{\substack{j_1=1 \\ j_1 \neq j}}^p [\mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) - \mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)})] \right) \left(\sum_{\substack{j_2=1 \\ j_2 \neq j}}^p [\mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(i)}) - \mathbf{f}(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(i)})] \right) \right]^T.
\end{aligned} \tag{4.36}$$

1 Further, for the estimation of the non-normalized total GSI, we define the fol-
2 lowing kernel.

$$K^{tot}(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)}) = \mathbb{T}_r [\mathbf{K}^{tot}(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})], \tag{4.37}$$

4 with \mathbf{K}^{tot} the kernel defined in (4.26).

6 The kernel $K(\cdot)$ is symmetric with respect to its first and second arguments.
7 Other properties of the kernels K and K^{tot} are given in Corollary 1.

9 **Corollary 1.** *Let $(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)})$ be p i.i.d copies of \mathbf{X}_u and $(\mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})$ be
10 q i.i.d copies of $\mathbf{X}_{\sim u}$, with $p, q \geq 2$. If assumptions A1 (independent inputs) and
11 A2 ($\mathbb{E} [\|\mathbf{f}(\mathbf{X})\|_{L^2}^2] < +\infty$) hold, then we have*

$$\mathbb{E} [K(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})] = \mathbb{T}_r(\Sigma_u); \tag{4.38}$$

$$\mathbb{E} [K^{tot}(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})] = \mathbb{T}_r(\Sigma_u^{tot}). \tag{4.39}$$

13 **Proof.** *The proofs are straightforward using Equations (4.27-4.28), the linearity
14 of the expectation, and the linearity of the trace function ($\mathbb{T}_r(\cdot)$).*

15 □

1 Corollary 1 shows that the kernels $K(\cdot)$ (resp. $K^{tot}(\cdot)$) will lead to a general-
2 ized, unbiased estimator of $\mathbb{T}_r(\Sigma)_u$ (resp. $\mathbb{T}_r(\Sigma_u^{tot})$), which is the non-normalized
3 first-order (resp. total) GSI of \mathbf{X}_u . Both kernels are symmetric in their first and
4 second arguments, and this property allows for using the theory of U-statistics of
5 two samples to obtain the statistical properties of estimators based on the kernels
6 $K(\cdot)$ and $K^{tot}(\cdot)$ ([20], [21] [22], [16; 23]).

7 4.4. Joint estimator of the classical first-order and total generalized sensitivity in- 8 dices

9 This section aims to provide the joint estimator of the classical GSIs. The follow-
10 ing theorems give the minimum variance unbiased estimator of the non-normalized
11 GSIs as well as its variance and its asymptotic distribution for a given degree (p, q)
12 of the kernel. Based on the kernels $K(\cdot)$ and $K^{tot}(\cdot)$, we define $\mathbf{K}_t(\cdot)$ as follows:

$$\mathbf{K}_t(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)}) = \begin{bmatrix} K(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)}) \\ K^{tot}(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)}) \end{bmatrix}. \quad (4.40)$$

13 Of course, the multivariate kernel $\mathbf{K}_t(\cdot)$ is symmetric in its first and second
14 arguments. For $j, i \in \{0, 1, \dots, d\}$, we define $\Sigma^{j,l}$ as

$$\Sigma^{j,l} = \mathbb{V}(\mathbb{E}[\mathbf{K}(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)}) | \mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(l)}]).$$

15 Regarding MSA, $\Sigma^{j,l}$ is the first-order covariance matrix of $(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(l)})$
16 associated with the kernel $\mathbf{K}_t(\cdot)$. In particular, $\Sigma^{p,q}$ is the covariance matrix of the
17 kernel $\mathbf{K}_t(\cdot)$. Now, we have all elements to derive the joint estimator of GSIs. Corol-
18 lary 2 deals with this issue.

19

1 **Corollary 2.** Let $\mathcal{X} = (\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)})$ be p i.i.d copies of \mathbf{X}_u , $\mathcal{Y} = (\mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})$
2 be q i.i.d copies of $\mathbf{X}_{\sim u}$, $\mathcal{X}_i = (\mathbf{X}_{i,u}^{(1)}, \dots, \mathbf{X}_{i,u}^{(p)})$ and $\mathcal{Y}_i = (\mathbf{X}_{i,\sim u}^{(1)}, \dots, \mathbf{X}_{i,\sim u}^{(q)})$,
3 $i = 1, 2, \dots, m$, be two independent samples of size m from \mathcal{X} and \mathcal{Y} respectively. If
4 assumptions A1, A3 ($\mathbb{E} \|f(\mathbf{X})\|_{L^2}^4 < +\infty$), and A4 ($2 \leq p, 2 \leq q$) hold, then we have

5
6 i) the minimum variance unbiased estimator of $[\mathbb{T}_r(\Sigma_u) \mathbb{T}_r(\Sigma_u^{tot})]^T$ for a given
7 (p, q) and m is given by

$$\begin{aligned} \begin{bmatrix} \widehat{\mathbb{T}_r(\Sigma_u)} \\ \widehat{\mathbb{T}_r(\Sigma_u^{tot})} \end{bmatrix} &= \begin{bmatrix} \frac{2}{mp^2(p-1)q(q-1)} \sum_{i=1}^m \sum_{j=1}^p \sum_{l=1}^{q-1} \sum_{k>l}^q \\ \frac{2}{mp^2(p-1)q} \sum_{i=1}^m \sum_{j=1}^p \sum_{l=1}^q \end{bmatrix} \\ &\times \mathbb{T}_r \left[\left(\sum_{\substack{j_1=1 \\ j_1 \neq j}}^p \left[\mathbf{f}(\mathbf{X}_{i,u}^{(j)}, \mathbf{X}_{i,\sim u}^{(l)}) - \mathbf{f}(\mathbf{X}_{i,u}^{(j_1)}, \mathbf{X}_{i,\sim u}^{(l)}) \right] \right) \right] \\ &\times \mathbb{T}_r \left[\left(\sum_{\substack{j_1=1 \\ j_1 \neq j}}^p \left[\mathbf{f}(\mathbf{X}_{i,u}^{(j)}, \mathbf{X}_{i,\sim u}^{(l)}) - \mathbf{f}(\mathbf{X}_{i,u}^{(j_1)}, \mathbf{X}_{i,\sim u}^{(l)}) \right] \right) \right] \\ &\times \left(\sum_{\substack{j_2=1 \\ j_2 \neq j}}^p \left[\mathbf{f}(\mathbf{X}_{i,u}^{(j)}, \mathbf{X}_{i,\sim u}^{(k)}) - \mathbf{f}(\mathbf{X}_{i,u}^{(j_2)}, \mathbf{X}_{i,\sim u}^{(k)}) \right] \right)^T \\ &\times \left(\sum_{\substack{j_1=1 \\ j_2 \neq j}}^p \left[\mathbf{f}(\mathbf{X}_{i,u}^{(j)}, \mathbf{X}_{i,\sim u}^{(l)}) - \mathbf{f}(\mathbf{X}_{i,u}^{(j_2)}, \mathbf{X}_{i,\sim u}^{(l)}) \right] \right)^T \end{bmatrix}; \end{aligned} \quad (4.41)$$

9 ii) the mean square error of $[\mathbb{T}_r(\Sigma_u) \mathbb{T}_r(\Sigma_u^{tot})]^T$ is

$$m\mathbb{E} \left(\left\| \begin{bmatrix} \widehat{\mathbb{T}_r(\Sigma_u)} \\ \widehat{\mathbb{T}_r(\Sigma_u^{tot})} \end{bmatrix} - \begin{bmatrix} \mathbb{T}_r(\Sigma_u) \\ \mathbb{T}_r(\Sigma_u^{tot}) \end{bmatrix} \right\|^2 \right) = \mathbb{T}_r(\Sigma^{p,q}); \quad (4.42)$$

10 iii) if $m \rightarrow +\infty$, we have the asymptotic normality, that is,

$$\sqrt{m} \left(\begin{bmatrix} \widehat{\mathbb{T}_r(\Sigma_u)} \\ \widehat{\mathbb{T}_r(\Sigma_u^{tot})} \end{bmatrix} - \begin{bmatrix} \mathbb{T}_r(\Sigma_u) \\ \mathbb{T}_r(\Sigma_u^{tot}) \end{bmatrix} \right) \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}, \Sigma^{p,q}). \quad (4.43)$$

11 **Proof.** See Appendix F.

12 □

13 Corollary 2 gives an interesting (joint) estimator of the non-normalized GSIs of
14 \mathbf{X}_u , and the theorem below provides the joint estimator of GSIs of \mathbf{X}_u .

15

16 **Theorem 4.** Let $\mathcal{X} = (\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)})$ be p i.i.d copies of \mathbf{X}_u , $\mathcal{Y} = (\mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})$
17 be q i.i.d copies of $\mathbf{X}_{\sim u}$, $\mathcal{X}_i = (\mathbf{X}_{i,u}^{(1)}, \dots, \mathbf{X}_{i,u}^{(p)})$ and $\mathcal{Y}_i = (\mathbf{X}_{i,\sim u}^{(1)}, \dots, \mathbf{X}_{i,\sim u}^{(q)})$,

1 $i = 1, 2, \dots, m$, be two independent samples of size m from \mathcal{X} and \mathcal{Y} respectively. If
 2 assumptions A1, A3 ($\mathbb{E} \|f(\mathbf{X})\|_{L^2}^4 < +\infty$), and A4 ($2 \leq p, 2 \leq q$) hold, then we have

3

4 i) the joint estimator of GSIs of \mathbf{X}_u ($[GSI_u, GSI_{T_u}]^T$) for a given (p, q) and m
 5 is given by

$$\begin{aligned} \begin{bmatrix} \widehat{GSI}_u \\ \widehat{GSI}_{T_u} \end{bmatrix} &= \frac{1}{\widehat{\mathbb{T}}_r(\Sigma)} \times \left[\frac{2}{mp^2(p-1)q(q-1)} \sum_{i=1}^m \sum_{j=1}^p \sum_{l=1}^{q-1} \sum_{k>l}^q \right. \\ &\quad \times \mathbb{T}_r \left[\left(\sum_{\substack{j_1=1 \\ j_1 \neq j}}^p \left[\mathbf{f}(\mathbf{X}_{i,u}^{(j)}, \mathbf{X}_{i,\sim u}^{(l)}) - \mathbf{f}(\mathbf{X}_{i,u}^{(j_1)}, \mathbf{X}_{i,\sim u}^{(l)}) \right] \right) \right. \\ &\quad \times \mathbb{T}_r \left[\left(\sum_{\substack{j_1=1 \\ j_1 \neq j}}^p \left[\mathbf{f}(\mathbf{X}_{i,u}^{(j)}, \mathbf{X}_{i,\sim u}^{(l)}) - \mathbf{f}(\mathbf{X}_{i,u}^{(j_1)}, \mathbf{X}_{i,\sim u}^{(l)}) \right] \right) \right. \\ &\quad \times \left(\sum_{\substack{j_2=1 \\ j_2 \neq j}}^p \left[\mathbf{f}(\mathbf{X}_{i,u}^{(j)}, \mathbf{X}_{i,\sim u}^{(k)}) - \mathbf{f}(\mathbf{X}_{i,u}^{(j_2)}, \mathbf{X}_{i,\sim u}^{(k)}) \right] \right)^T \left. \right] \\ &\quad \times \left. \left(\sum_{\substack{j_1=1 \\ j_2 \neq j}}^p \left[\mathbf{f}(\mathbf{X}_{i,u}^{(j)}, \mathbf{X}_{i,\sim u}^{(l)}) - \mathbf{f}(\mathbf{X}_{i,u}^{(j_2)}, \mathbf{X}_{i,\sim u}^{(l)}) \right] \right)^T \right] \right], \quad (4.44) \end{aligned}$$

6 where $\widehat{\mathbb{T}}_r(\Sigma) = \mathbb{T}_r \left[\frac{1}{M-1} \sum_{i=1}^M (\mathbf{f}(\mathbf{X}_i) - \bar{\mathbf{f}}) (\mathbf{f}(\mathbf{X}_i) - \bar{\mathbf{f}})^T \right]$ is the estimator of the model
 7 inertia and $\bar{\mathbf{f}} = \frac{1}{M} \sum_{i=1}^M \mathbf{f}(\mathbf{X}_i)$ is the estimator of the model mean for M model runs;

8

9 ii) the estimator $[\widehat{GSI}_u \widehat{GSI}_{T_u}]^T$ is consistent, that is,

$$[\widehat{GSI}_u \widehat{GSI}_{T_u}]^T \xrightarrow{\mathcal{P}} [GSI_u GSI_{T_u}]^T; \quad (4.45)$$

10 iii) if $m \rightarrow +\infty$ and $m/M \rightarrow 0$, we have the asymptotic normality, that is,

$$\sqrt{m} \left([\widehat{GSI}_u \widehat{GSI}_{T_u}]^T - [GSI_u GSI_{T_u}]^T \right) \xrightarrow{\mathcal{D}} \mathcal{N} \left(\mathbf{0}, \frac{\Sigma^{p,q}}{[\mathbb{T}_r(\Sigma)]^2} \right); \quad (4.46)$$

11 iv) the $100 \times (1 - \alpha)\%$ asymptotic confidence regions for $[GSI_u GSI_{T_u}]^T$ is given
 12 by

$$\mathbb{P} \left[m \left(\begin{bmatrix} \widehat{GSI}_u \\ \widehat{GSI}_{T_u} \end{bmatrix} - \begin{bmatrix} GSI_u \\ GSI_{T_u} \end{bmatrix} \right)^T \widehat{\Gamma}^{-1} \left(\begin{bmatrix} \widehat{GSI}_u \\ \widehat{GSI}_{T_u} \end{bmatrix} - \begin{bmatrix} GSI_u \\ GSI_{T_u} \end{bmatrix} \right) \leq T^2(\alpha) \right] = 1 - \alpha, \quad (4.47)$$

13 with $\widehat{\Gamma} = \frac{\Sigma^{p,q}}{[\mathbb{T}_r(\Sigma)]^2}$, and $T^2(\alpha)$ the 100α percentile of the Hotelling T -square distribu-
 14 tion.

15 **Proof.** See Appendix G.

16

□

1 **Remark 5.** *In practice, we can estimate the covariance matrix Γ , and the confidence*
2 *regions are the ellipsoid with center $\left[\widehat{GSI}_u \widehat{GSI}_{T_u}\right]^T$ and axes $\pm\sqrt{\lambda_i T^2(\alpha)}\mathbf{v}_i$, where*
3 *λ_i and \mathbf{v}_i are the eigenvalues and eigenvectors of $\widehat{\Gamma}$ respectively.*

4 5. Choice of the degree of the kernel and algorithm

5 5.1. Proxy measure for the choice of the degree of a kernel

6 A classical way of choosing the degree (p, q) of a kernel consists in using a proxy
7 measure for the variance of this kernel, as the variance of the kernel involves fourth
8 moments, which are often unknown and hard to estimate ([24]). A proxy measure
9 is a part of the upper bound of the variance that is based on known coefficients
10 (see [24] for comprehensive details). It aims at overcoming the estimation of fourth
11 moments during the estimation of variances, and it is based on the minimum-norm
12 quadratic estimation (MINQE) or the MINQUE and MINQIE versions using unbi-
13 asedness or invariance as constraints (see [25]). In paper [24], the author generalized
14 the principle of MINQE to deal with the variances of the estimators of sensitivity
15 indices, and the author in [14; 15] used the MINQE approach to identify the degree
16 of the kernel for estimating both Sobol's first-order and total indices.

17

18 The MINQE approach used in [14; 15] can be adapted to the kernels $K(\cdot)$ and
19 $K^{tot}(\cdot)$ of this paper, and it leads to the following conclusions: first, the proxy
20 measure does not depend on the value of q . Thus, the value of $q = 2$ should be used
21 as the referenced value of the degree of the kernel, and we should expect to have
22 good results with $q \leq p$. And second, the value $p = 2$ serves as the referenced degree
23 for p , as the proxy measure increases with p . AppendixH deals with these issues.

24 5.2. Algorithm: design scheme and main steps for the estimations of the first-order 25 and total-effect covariance matrices

26 For a given degree (p, q) and sample size m , the following steps are used to com-
27 pute the d first-order and total-effect covariance matrices, including the GSIs.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34

Algorithm 1. *Main steps with $q \leq p$.*

i) Sample p input values (matrices) of type $m \times d$ ($\mathcal{X}_1, \dots, \mathcal{X}_p$).

ii) For each factor X_j , replace the j^{th} column of \mathcal{X}_1 with the j^{th} column of $\mathcal{X}_2, \dots, \mathcal{X}_p$ to obtain $p - 1$ new matrices ($\mathcal{X}_{2j}, \dots, \mathcal{X}_{pj}$).

iii) Run the model for \mathcal{X}_1 and for the $p - 1$ input values $\mathcal{X}_{2j}, \dots, \mathcal{X}_{pj}$ to obtain p outputs (matrices of type $m \times N$): $\mathcal{Y}_1, \mathcal{Y}_{2j}, \dots, \mathcal{Y}_{pj}$.

iv) Repeat steps ii) and iii) $q - 1$ times by replacing \mathcal{X}_1 with $\mathcal{X}_2, \mathcal{X}_3, \dots, \mathcal{X}_q$.

v) Use the above model evaluations and the estimators in (4.29), (4.32), and (4.44) to obtain the estimates of the first-order and total-effect covariance matrices in one hand, and the GSIs estimates in the other hand.

From the description in Algorithm 1, the number of model evaluations carried out in step iii) is $m + m \times (p - 1) \times d = m[(p - 1) \times d + 1]$, as \mathcal{Y}_1 is used for the computation of each index. Thus, the computational cost or the total number of model runs for the computation of the d indices is $mq[(p - 1)d + 1]$, as we repeat in total step ii-iii) q times. Algorithm 1 can be used to compute the GSIs of a subset of input factors \mathbf{X}_u , with $|u| > 1$. We should modify the step ii) of Algorithm 1 as follows: replace the columns of \mathcal{X}_1 , which indices are in u , with the same columns of $\mathcal{X}_2, \dots, \mathcal{X}_p$ to obtain $p - 1$ new matrices ($\mathcal{X}_{2j}, \dots, \mathcal{X}_{pj}$).

The computational cost does not explicitly include the dimensionality of the outputs. In classical statistics, we should require the sample size m to be greater than the dimensionality of the outputs N in order to have better estimations. However, the original outputs can be transformed into a few new outputs that preserve much or all information (e.g., principal components [4]). In this case, we can use the invariance property of some estimators to obtain the GSIs estimates. For very high dimensionality of the outputs, modern estimators of eigenvectors and principal components should be used ([26]).

1 6. Numerical tests

2 In this section, we perform some numerical tests to compare the effectiveness of
3 our estimations depending on the degree (p, q) of the symmetric kernel. We present
4 the following: i) the functions used to illustrate our approach, ii) the main choices
5 made to obtain the results, and iii) the numerical results.

6 6.1. Test functions

7 To illustrate our approach, we consider two types of functions as follows: func-
8 tions with a small number of inputs ($d = 2, d = 3$), and functions with a medium
9 number of inputs ($d = 10$). We also give the classical GSIs ([2; 5]) of these functions.

10 6.1.1. Multivariate exponential function ($d = 2$)

11 The multivariate exponential function includes two independent inputs following
12 a normal distribution $\mathcal{N}(0, 1)$. It is defined as follows:

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} \exp(x_1 + 2x_2) \\ x_1^4 x_2^2 \end{bmatrix}. \quad (6.48)$$

The classical GSIs for this function are listed in Table 1

j	1	2
GSI_j	0.016	0.359
GSI_{T_j}	0.641	0.984

Table 1: Generalized Sensitivity Indices (GSIs) of the multivariate exponential function

13

14 This function belongs to the class of functions with important interactions among
15 input factors.

16 6.1.2. Multivariate Ishigami's function ($d = 3$)

17 The multivariate Ishigami function includes three independent input factors fol-
18 lowing a uniform distribution on $[-\pi, \pi]$, and it provides three outputs. It is defined

1 as follows:

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} \sin(x_1) + 7 \sin^2(x_2) + 0.1 x_3^4 \sin(x_1) \\ \sin(x_1) + 5.896 \sin^2(x_2) + 0.1 x_3^4 \sin(x_1) \\ \sin(x_1) + 6.494 \sin^2(x_2) + 0.125 x_3^4 \sin(x_1) \end{bmatrix}. \quad (6.49)$$

The classical GSIs for this function are listed in Table 2.

j	1	2	3
GSI_j	0.345	0.372	0.000
GSI_{T_j}	0.628	0.372	0.284

Table 2: Generalized Sensitivity Indices (GSIs) of the multivariate Ishigami function

2

3 6.1.3. Multivariate Sobol's function ($d = 10$)

4 The multivariate Sobol function includes 10 independent input factors following
5 a uniform distribution on $[0, 1]$. It is defined as follows:

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} \prod_{j=1}^{d=10} \frac{|4x_j - 2| + \mathcal{A}[1,j]}{1 + \mathcal{A}[1,j]} \\ \prod_{j=1}^{d=10} \frac{|4x_j - 2| + \mathcal{A}[2,j]}{1 + \mathcal{A}[2,j]} \\ \prod_{j=1}^{d=10} \frac{|4x_j - 2| + \mathcal{A}[3,j]}{1 + \mathcal{A}[3,j]} \\ \prod_{j=1}^{d=10} \frac{|4x_j - 2| + \mathcal{A}[4,j]}{1 + \mathcal{A}[4,j]} \end{bmatrix}. \quad (6.50)$$

6 According to the values of \mathcal{A} (matrix of type $4 \times d$), this function has different
7 properties:

- if

$$\mathcal{A} = \begin{bmatrix} 0 & 0 & 6.52 & 6.52 & 6.52 & 6.52 & 6.52 & 6.52 & 6.52 & 6.52 \\ 0 & 1 & 4.5 & 9 & 99 & 99 & 99 & 99 & 99 & 99 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 50 & 50 & 50 & 50 & 50 & 50 & 50 & 50 & 50 & 50 \end{bmatrix},$$

8 the resulting values of the classical GSIs are those listed in Table 3.

j	1	2	3	4	5	6	7	8	9	10
GSI_j	0.491	0.297	0.025	0.015	0.010	0.008	0.007	0.007	0.006	0.006
GSI_{T_j}	0.605	0.406	0.034	0.021	0.014	0.013	0.011	0.011	0.010	0.009

Table 3: Generalized Sensitivity Indices (GSIs) of the multivariate Sobol function of type A

1 Thus, only a few inputs are important and the function has a low effective
2 dimension (function of type A);

• if

$$\mathcal{A} = \begin{bmatrix} 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 \\ 20 & 1 & 20 & 20 & 20 & 20 & 20 & 20 & 20 & 20 & 20 \\ 50 & 50 & 50 & 50 & 50 & 50 & 50 & 50 & 50 & 50 & 50 \\ 60 & 60 & 60 & 60 & 60 & 60 & 60 & 60 & 60 & 60 & 60 \end{bmatrix},$$

the resulting GSI values are those listed in Table 4. Thus, all input factors are

j	1	2	3	4	5	6	7	8	9	10
GSI_j	0.099	0.099	0.099	0.099	0.099	0.099	0.099	0.099	0.099	0.099
GSI_{T_j}	0.101	0.101	0.101	0.101	0.101	0.101	0.101	0.101	0.101	0.101

Table 4: Generalized Sensitivity Indices (GSIs) of the multivariate Sobol function of type B

3
4 important, but there is no interaction among these inputs. The function has
5 a high effective dimension (function of type B);

• If

$$\mathcal{A} = \begin{bmatrix} 0 & 0 & 6.52 & 6.52 & 6.52 & 6.52 & 6.52 & 6.52 & 6.52 & 6.52 & 6.52 \\ 0 & 1 & 4.5 & 9 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\ 50 & 50 & 50 & 50 & 50 & 50 & 50 & 50 & 50 & 50 & 50 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

6 the function belongs to the class of functions with important interactions
7 among input factors. The GSI values are listed in Table 5. Due to these
8 important interactions, it is clear that all input factors are important. The
9 function has a high effective dimension (function of type C).

j	1	2	3	4	5	6	7	8	9	10
GSI_j	0.055	0.041	0.019	0.019	0.019	0.019	0.019	0.019	0.019	0.019
GSI_{T_j}	0.291	0.277	0.247	0.246	0.246	0.246	0.246	0.246	0.246	0.246

Table 5: Generalized Sensitivity Indices (GSIs) of the multivariate Sobol function of type C

Remark 6. *To obtain the exact values of GSIs listed in Tables 1-5, we use the aggregated definition of GSIs ([4]), that is,*

$$GSI_j = \frac{\sum_{i=1}^d D_{i,j}}{\sum_{i=1}^d \mathbb{V}[f_i(\mathbf{X})]},$$

$$GSI_{T_j} = \frac{\sum_{i=1}^d D_{i,j}^{tot}}{\sum_{i=1}^d \mathbb{V}[f_i(\mathbf{X})]},$$

where $D_{i,j}$ (resp. $D_{i,j}^{tot}$) is the non-normalized first-order (resp. total) index of X_j associated with the i^{th} component of $\mathbf{f}(\mathbf{X})$ (i.e., $f_i(\mathbf{X})$).

The values of $\mathbb{V}[f_i(\mathbf{X})]$, $D_{i,j}$, and $D_{i,j}^{tot}$ for different single-response models, used in this paper, can be found in [14; 27; 28; 29].

6.2. Implementation issues

In this paper, we used Sobol's design from the R-package randtoolbox ([30]) to generate input values according to their probability measure. We considered six values of the degree, as follows: $(p = 3, q = 3)$, $(p = 4, q = 3)$, $(p = 4, q = 4)$, $(p = 5, q = 3)$, $(p = 5, q = 4)$, and $(p = 5, q = 5)$. We used $(p = 2, q = 2)$ as the referenced value of the degree, and we added the degree $(p, q = 2)$ to assess the numerical impact of q (i.e., fix $q = 2$ and consider various choices for p).

For a given value of the degree (p, q) , we increased the sample size (m) by 30 from 5 up to 500 depending on the convergence of estimations. For a fair comparison, we computed the indices for degrees (p, q) , $(p = 2, q = 2)$, and $(p, q = 2)$ using the same number of model evaluations, that is, the sample size (m_r) for the referenced estimator $(p = 2, q = 2)$ is $m_r = \text{round}(qm[(p - 1)d + 1]/(2d + 2))$ for a given (p, q) (see Section 5.2). We also added the up-to-date estimator of the classical GSIs (from papers [5; 29]) implemented in the R-package sensitivity ([31]).

1 We used the root mean square error (RMSE) to assess the accuracy of our estima-
 2 tions. For each sample size (m) and for each degree (p, q), we replicate the process
 3 of computing the indices $R = 30$ times (changing the seed randomly when sampling
 4 input values). The average RMSE of the d first-order indices is defined as follows:

$$RMSE_d = \frac{1}{d} \sum_{j=1}^d \sqrt{\frac{1}{R} \sum_{r=1}^R (\widehat{GSI}_{j,r} - GSI_j)^2}, \quad (6.51)$$

5
 6 where GSI_j and $\widehat{GSI}_{j,r}$ are the true and estimated values, respectively, of the first-
 7 order index of a given factor $X_j, j = 1, 2, \dots, d$. Furthermore, $\widehat{GSI}_{j,r}$ is the first-
 8 order estimate for a given replication r . We used the same expression of RMSE for
 9 the total indices GSI_{T_j} .

10 6.3. Numerical results and discussion

11 6.3.1. Estimates of the three types of the generalized sensitivity indices

12 Tables 6 and 7 show the estimates of the three types of the GSIs for the mul-
 13 tivariate exponential function and Ishigami function, respectively, when the sample
 14 size $m = 1000$ and the degree ($p = 2, q = 2$).

	First-order GSIs				Total GSIs			
	GSI^F	GSI^S	GSI^{l1}	GSI^{l2}	GSI_T^F	GSI_T^S	GSI_T^{l1}	GSI_T^{l2}
X_1	0.117	0.070	0.066	0.042	0.769	0.653	0.403	0.332
X_2	0.316	0.316	0.180	0.158	0.968	0.918	0.517	0.460

Table 6: Estimates of the three types of GSIs for the multivariate exponential function

	First-order GSIs				Total GSIs			
	GSI^F	GSI^S	GSI^{l1}	GSI^{l2}	GSI_T^F	GSI_T^S	GSI_T^{l1}	GSI_T^{l2}
X_1	0.349	0.349	0.347	0.116	0.626	0.626	0.621	0.209
X_2	0.376	0.376	0.375	0.125	0.376	0.376	0.375	0.125
X_3	0.002	0.002	0.002	0.001	0.280	0.280	0.277	0.093

Table 7: Estimates of the three types of GSIs for the multivariate Ishigami function

15 It comes out that the three types of GSIs give the same ranking of input fac-
 16 tors for both functions. **For more information about the estimations, we also added**

1 **the asymptotic** confident regions of the GSIs of both functions in AppendixI. The
2 asymptotic confident regions of the GSIs for the multivariate exponential function
3 show a high uncertainty on the total GSIs. This high uncertainty highlights the dif-
4 ficulty of estimating the GSIs of the multivariate exponential function (see Sections
5 6.3.2-6.3.3).

6 *6.3.2. Comparison of the estimates of the classical generalized first-order indices*

7 Figures 1, 2, 3, 4, and 5 show the RMSEs of the d first-order indices when the
8 total number of model evaluations increases for the functions in (6.48), (6.49), and
9 (6.50) with the three types (Sobol's function of types A, B, and C) respectively. The
10 figures show the trends of the average RMSEs for the six values of the degree (p, q)
11 compared to the RMSEs associated with the reference degree $(p = 2, q = 2)$ and
12 the degree $(p, q = 2)$. We also added the RMSEs associated with the up-to-date
13 estimator (from papers [5; 29]).

14

15 In Figures 1–5, the RMSEs of our estimators (this paper), decrease with the
16 number of model evaluations for different values of the degree, and we have con-
17 verging estimations. In these figures, our estimators outperform the up-to-date
18 estimator, with a significant difference observed in Figures 1 and 5. Moreover, Fig-
19 ure 1 shows that the trends of the RMSEs associated with the up-to-date estimator
20 do not decrease with the total number of model evaluations. Thus, this estimator
21 (i.e. estimator and sampling strategies used) fails to converge. Similar results were
22 obtained in [14; 28].

23

24 It can be seen in Figure 1 that the kernel of degree $(p = 5, q = 5)$ performs
25 better compared to the others (degrees $(p, q = 2)$ and $(p = 2, q = 2)$), while Figure
26 5 shows that the kernel of degree $(p, q = 2)$ is generally the best among the six
27 kernels, followed by the kernel of degree (p, q) . For the remaining models (Figures
28 2, 3, and 4), the referenced degree $(p = 2, q = 2)$ performs generally better compared

1 to the others.

2 6.3.3. Comparison of the estimates of the classical generalized total indices

3 Figures 6, 7, 8, 9, and 10 show the average RMSEs of the d total indices when
4 the total number of model evaluations increases for the functions in (6.48), (6.49),
5 and (6.50) with the three types (Sobol's function of types A, B, and C) respectively.
6 The figures show the trends of the average RMSEs for the six values of the degree
7 (p, q) compared to the RMSEs associated with the reference degree $(p = 2, q = 2)$
8 and the degree $(p, q = 2)$. We also added the RMSEs associated with the up-to-date
9 estimator.

10

11 We have converging estimations in Figures 6–10. Our estimators outperform the
12 up-to-date estimator in Figure 6. The kernel of degree (p, q) gives the best results
13 for the model in (6.48), followed by the kernel of degree $(2, 2)$. In the case of Figure
14 7, the kernel of degree $(2, 2)$ appears as the best kernel for estimating total indices.
15 Furthermore, when increasing (p, q) , all estimators become equivalent, including the
16 up-to-date estimator.

17

18 For the remaining models (Figures 9 and 10), we can find a degree that performs
19 slightly better and a degree that performs slightly worse than the up-to-date estima-
20 tor. Furthermore, the estimates from the degree $(2, 2)$ outperform the results from
21 other degrees and become equivalent to the up-to-date estimator when increasing
22 the sample size. For Figure 8, we can see that the up-to-date estimator outperforms
23 our estimator $(p = 3, q = 2)$ when the sample size is high.

24

25 The differences in the estimations of the total GSIs seem to come from the design
26 strategies. The better performance of the up-to-date estimator in the case of Figure 8
27 is probably due to our design scheme, which is well suited for the estimations of the
28 first-order GSI. Indeed, the estimations of the first-order GSI require $q \geq 2$ while the

1 estimations of the total GSI can be obtained with $q = 1$ (see Remark 4). Thus, our
2 estimates of the total GSI make use of the available model runs performed for better
3 estimating the first-order GSI. The worse performance of the up-to-date estimator in
4 the case of Figure 6 is due to the difficulty of this approach to capture the first-order
5 effect and important interactions among input factors in the presence of skewed or
6 heavy-tailed distributions of input factors.

7 7. Conclusion

8 **In this paper, we propose and study new GSIs using the first-order and to-**
9 **tal-effect** covariance matrices. The new GSIs make use of matrix norms when the
10 Loewner ordering on these covariance matrices does not exist, and we obtain the
11 classical GSIs from papers [2; 5] using the Frobenius norm. To estimate these in-
12 dices, first, we propose a minimum variance unbiased estimator of the first-order
13 and total-effect covariance matrices, and second, we provide an efficient (joint) esti-
14 mator of the first-order and total GSIs ([2; 5]). We also derive the consistency, the
15 asymptotic distribution, and the asymptotic confidence regions of these estimators.
16 As the joint estimator makes use of a kernel of degree (p, q) , we found that a kernel
17 of degree (p, q) with the smallest variance should be preferred for estimating the
18 classical GSIs. Although the proxy measure is in favor of the degree $(2, 2)$, the
19 degree (p, q) can be larger than $(2, 2)$ in practice.

20
21 The numerical tests confirmed the superiority of our estimators of the first-order
22 GSI compared to the estimator from papers [5; 29]. The superiority of our estimators
23 is also observed in the case of total indices, mainly when using challenging models.
24 In the case of total indices, our estimator associated with the degree $(p = 2, q)$ is
25 an average of the aggregated Jansen estimator ([32]) used in [29], and the difference
26 of results observed is due to the fact that both approaches use different information
27 (input values) to compute total indices.

1

2 **It also comes out from numerical tests that we can obtain an efficient kernel with**
3 degree larger than $(2, 2)$ in practice. Furthermore, there is no absolutely efficient
4 degree for all functions or models. Thus, some adaptive strategies are needed to
5 properly choose the degree of the kernel, for each input factor, prior to the estima-
6 tion of the indices. The strategy should be based on the variance of the kernel, but
7 it requires more investigations.

8

9 **The new indices based on the l_p norms provide complementary information about**
10 **the** influence of input factors, but more investigations are needed to i) identify the
11 most interesting value of p , ii) find the appropriate constants of normalization, and
12 iii) establish a link with the Hilbert-Schmidt norm, which does not require the
13 assumption of independence of input factors.

14 **References**

- 15 [1] K. Campbell, M. D. McKay, B. J. Williams, Sensitivity analysis when model
16 outputs are functions, *Reliability Engineering & System Safety* 91 (10-11)
17 (2006) 1468 – 1472.
- 18 [2] M. Lamboni, H. Monod, D. Makowski, Multivariate sensitivity analysis to mea-
19 sure global contribution of input factors in dynamic models, *Reliability Engi-
20 neering and System Safety* 96 (2011) 450–459.
- 21 [3] M. Lamboni, D. Makowski, S. Lehuger, B. Gabrielle, H. Monod, Multivariate
22 global sensitivity analysis for dynamic crop models, *Fields Crop Reasearch* 113
23 (2009) 312–320.
- 24 [4] M. Lamboni, D. Makowski, H. Monod, Multivariate global sensitivity analysis
25 for discrete-time models, *Rapport technique 2008-3*, INRA, UR341 Mathéma-
26 tiques et Informatique Appliquées, Jouy-en-Josas, France (2008).

- 1 [5] F. Gamboa, A. Janon, T. Klein, A. Lagnoux, Sensitivity indices for multivariate
2 outputs, *Comptes Rendus de l'Académie des Sciences* (2014) In press.
- 3 [6] H. Xiao, L. Li, Discussion of paper by Matieyendou Lamboni, Hervé Monod,
4 David Makowski: Multivariate sensitivity analysis to measure global contri-
5 bution of input factors in dynamic models, *Reliab. Eng. Syst. Saf.* 99 (2011)
6 450-459, *Reliability Engineering & System Safety* 147 (2016) 194 – 195.
- 7 [7] S. Xiao, Z. Lu, L. Xu, Multivariate sensitivity analysis based on the direction
8 of eigen space through principal component analysis, *Reliability Engineering &*
9 *System Safety* 165 (2017) 1 – 10.
- 10 [8] I. M. Sobol, Sensitivity analysis for non-linear mathematical models, *Mathe-*
11 *matical Modelling and Computational Experiments* 1 (1993) 407–414.
- 12 [9] A. Saltelli, K. Chan, E. Scott, *Variance-Based Methods, Probability and Statis-*
13 *tics*, John Wiley and Sons, 2000.
- 14 [10] R. Ghanem, D. Higdon, H. Owhadi, *Handbook of Uncertainty Quantification*,
15 Springer International Publishing, 2017.
- 16 [11] M. De Lozzo, A. Marrel, Sensitivity analysis with dependence and variance-
17 based measures for spatio-temporal numerical simulators, *Stochastic Environ-*
18 *mental Research and Risk Assessment* (2016) 1–17.
- 19 [12] L. Xu, Z. Lu, S. Xiao, The new importance measures based on vector projec-
20 tion for multivariate output: application on hydrological model, *Hydrology and*
21 *Earth System Sciences Discussions* 2016 (2016) 1–22.
- 22 [13] O. Garcia-Cabrejo, A. Valocchi, Global sensitivity analysis for multivariate out-
23 put using polynomial chaos expansion, *Reliability Engineering & System Safety*
24 126 (2014) 25 – 36.

- 1 [14] M. Lamboni, Global sensitivity analysis: a generalized, unbiased and
2 optimal estimator of total-effect variance, *Statistical Papers* (2016) 1–
3 26doi:10.1007/s00362-016-0768-5.
4 URL <http://dx.doi.org/10.1007/s00362-016-0768-5>
- 5 [15] M. Lamboni, Uncertainty quantification: A minimum variance unbiased (joint)
6 estimator of the non-normalized Sobol's indices, submitted in *Statistical Paper*
7 (2018) x–x.
- 8 [16] W. Hoeffding, A class of statistics with asymptotically normal distribution,
9 *Annals of Mathematical Statistics* 19 (1948) 293–325.
- 10 [17] B. Efron, C. Stein, The jackknife estimate of variance, *The Annals of Statistics*
11 9 (1981) 586–596.
- 12 [18] M. Lamboni, B. Iooss, A.-L. Popelin, F. Gamboa, Derivative-based global sen-
13 sitivity measures: General links with Sobol' indices and numerical tests, *Math-*
14 *ematics and Computers in Simulation* 87 (0) (2013) 45 – 54.
- 15 [19] M. Lamboni, Global sensitivity analysis: an efficient numerical method for
16 approximating the total sensitivity index, *International Journal for Uncertainty*
17 *Quantification* (accepted).
- 18 [20] E. L. Lehmann, Consistency and unbiasedness of certain nonparametric tests,
19 *Annals of Mathematical Statistics* 22 (1951) 165–179.
- 20 [21] T. S. Ferguson, *A Course in Large Sample Theory*, Chapman-Hall, New York,
21 1996.
- 22 [22] E. L. Lehmann, *Elements of Large Sample Theory*, Springer, 1999.
- 23 [23] W. Hoeffding, A non-parametric test for independence, *Annals of Mathematical*
24 *Statistics* 19 (1948) 546–557.

- 1 [24] A. B. Owen, Variance components and generalized Sobol' indices, SIAM/ASA
2 Journal on Uncertainty Quantification 1 (1) (2013) 19–41.
- 3 [25] C. R. Rao, J. Kleffe, Estimation of variance components and applications, North
4 - Holland, Amsterdam, 1988.
- 5 [26] X. Mestre, Improved estimation of eigenvalues and eigenvectors of covariance
6 matrices using their sample estimates, IEEE Transactions on Information The-
7 ory 54 (11) (2008) 5113–5129.
- 8 [27] E. Plischke, E. Borgonovo, C. L. Smith, Global sensitivity measures from given
9 data, European Journal of Operational Research 226 (3) (2013) 536 – 550.
- 10 [28] E. Borgonovo, S. Tarantola, E. Plischke, M. D. Morris, Transformations and
11 invariance in the sensitivity analysis of computer experiments, Journal of the
12 Royal Statistical Society Series B 76 (5) (2014) 925–947.
- 13 [29] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, S. Tarantola, Vari-
14 ance based sensitivity analysis of model output. Design and estimator for the
15 total sensitivity index, Computer Physics Communications 181 (2) (2010) 259–
16 270.
- 17 [30] C. Dutang, P. Savicky, randtoolbox: Generating and Testing Random Numbers,
18 R package version 1.13 (2013).
- 19 [31] G. Pujol, B. Iooss, A. Janon, sensitivity: Sensitivity Analysis, R package version
20 1.7 (2013).
- 21 [32] M. J. W. Jansen, Analysis of variance designs for model output, Computer
22 Physics Communications 117 (1999) 35–43.
- 23 [33] N. Sugiura, Multisample and multivariate nonparametric tests based on U-
24 statistics and their asymptotic efficiencies, Osaka J. Math. 2 (2) (1965) 385–426.
25 URL <http://projecteuclid.org/euclid.ojm/1200691466>

1 FIGURES

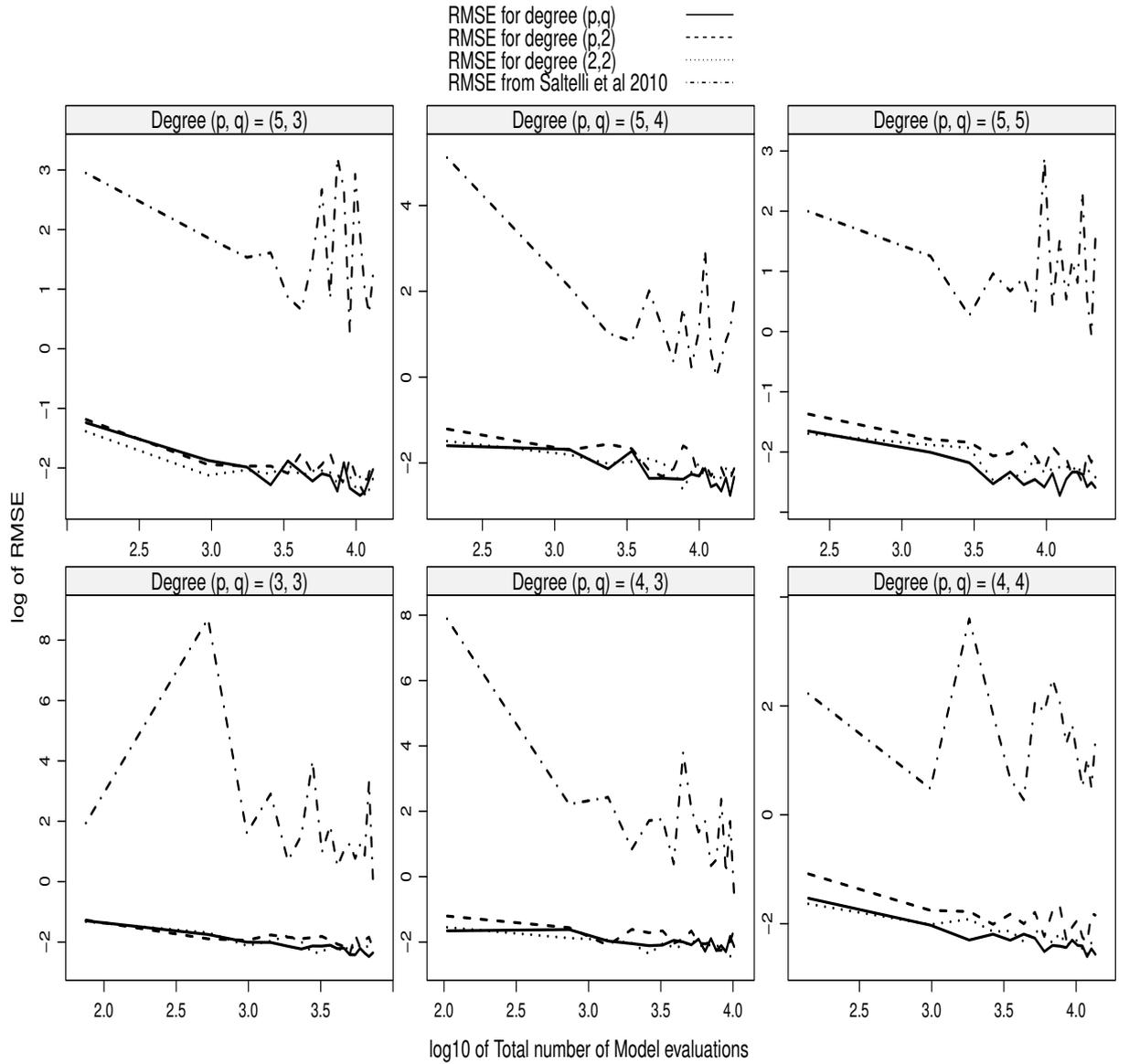


Figure 1: First-order indices: Log-RMSEs of the model in (6.48) against the total number of model evaluations (in \log_{10}) for six values of the degree ($p = 3, q = 3$), ($p = 4, q = 3$), ($p = 4, q = 4$), ($p = 5, q = 3$), ($p = 5, q = 4$), and ($p = 5, q = 5$). For each degree, we show the corresponding RMSE (solid line), the RMSE for degree ($p, q = 2$) (dashed line), the RMSE for degree (2, 2) (dotted line), and the RMSE from [29] (dash-dotted line).

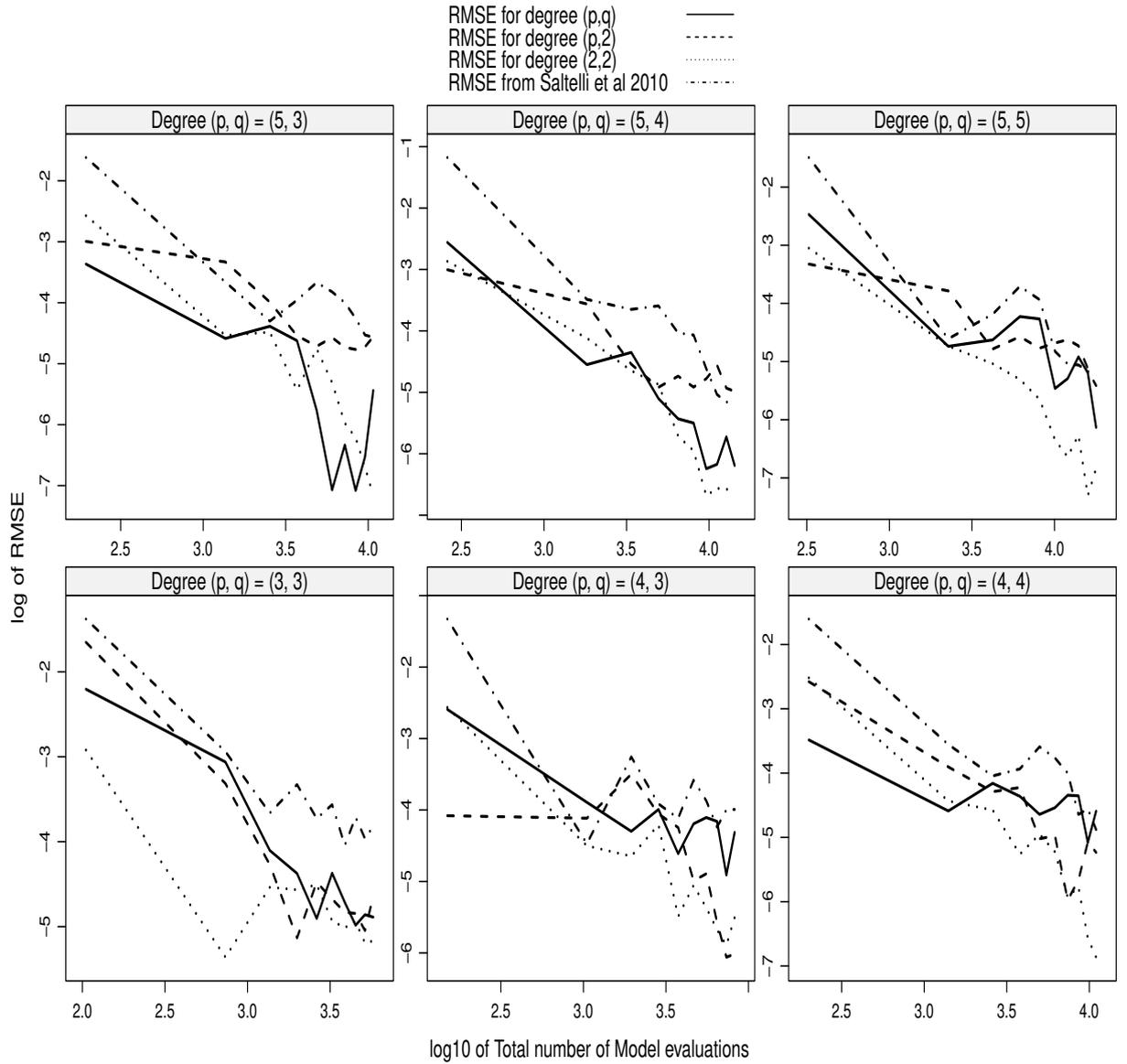


Figure 2: First-order indices: Log-RMSEs of the model in (6.49) against the total number of model evaluations (in \log_{10}) for six values of the degree ($p = 3, q = 3$), ($p = 4, q = 3$), ($p = 4, q = 4$), ($p = 5, q = 3$), ($p = 5, q = 4$), and ($p = 5, q = 5$). For each degree, we show the corresponding RMSE (solid line), the RMSE for degree ($p, q = 2$) (dashed line), the RMSE for degree ($2, 2$) (dotted line), and the RMSE from [29] (dash-dotted line).

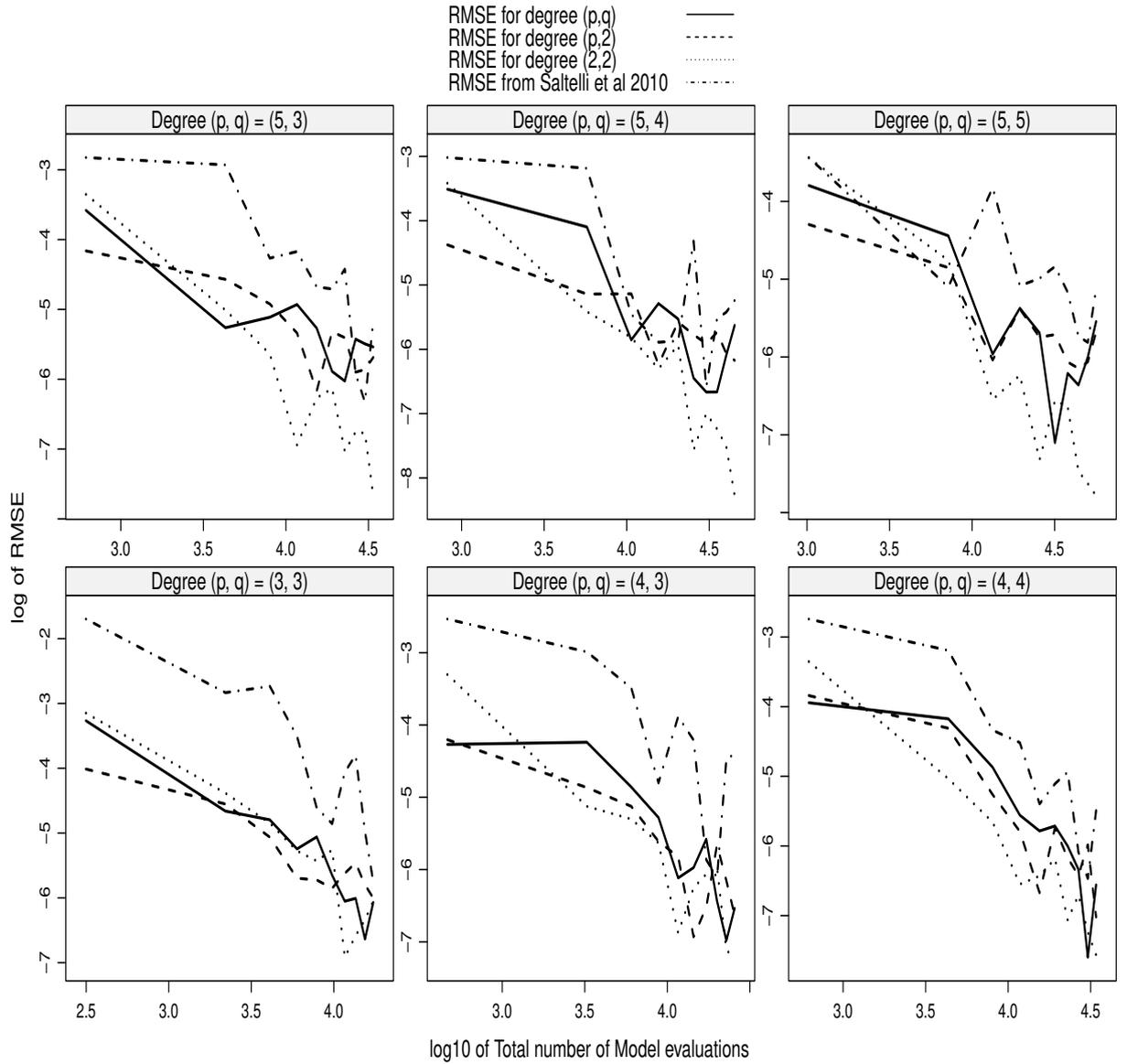


Figure 3: First-order indices: Log-RMSEs of the model in (6.50) of type A against the total number of model evaluations (in \log_{10}) for six values of the degree ($p = 3, q = 3$), ($p = 4, q = 3$), ($p = 4, q = 4$), ($p = 5, q = 3$), ($p = 5, q = 4$), and ($p = 5, q = 5$). For each degree, we show the corresponding RMSE (solid line), the RMSE for degree ($p, q = 2$) (dashed line), the RMSE for degree (2, 2) (dotted line), and the RMSE from [29] (dash-dotted line).

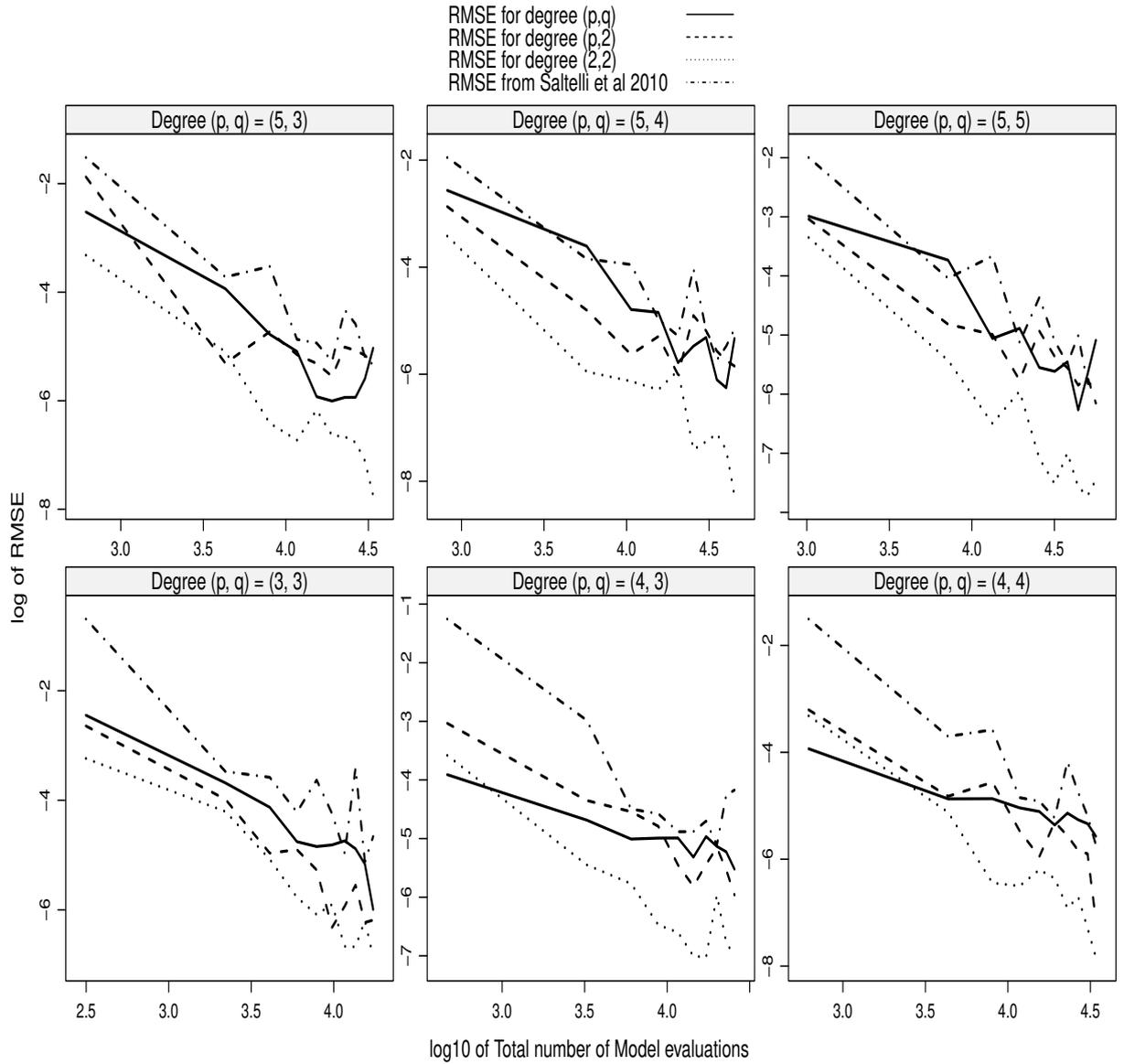


Figure 4: First-order indices: Log-RMSEs of the model in (6.50) of type B against the total number of model evaluations (in \log_{10}) for six values of the degree ($p = 3, q = 3$), ($p = 4, q = 3$), ($p = 4, q = 4$), ($p = 5, q = 3$), ($p = 5, q = 4$), and ($p = 5, q = 5$). For each degree, we show the corresponding RMSE (solid line), the RMSE for degree ($p, q = 2$) (dashed line), the RMSE for degree (2, 2) (dotted line), and the RMSE from [29] (dash-dotted line).

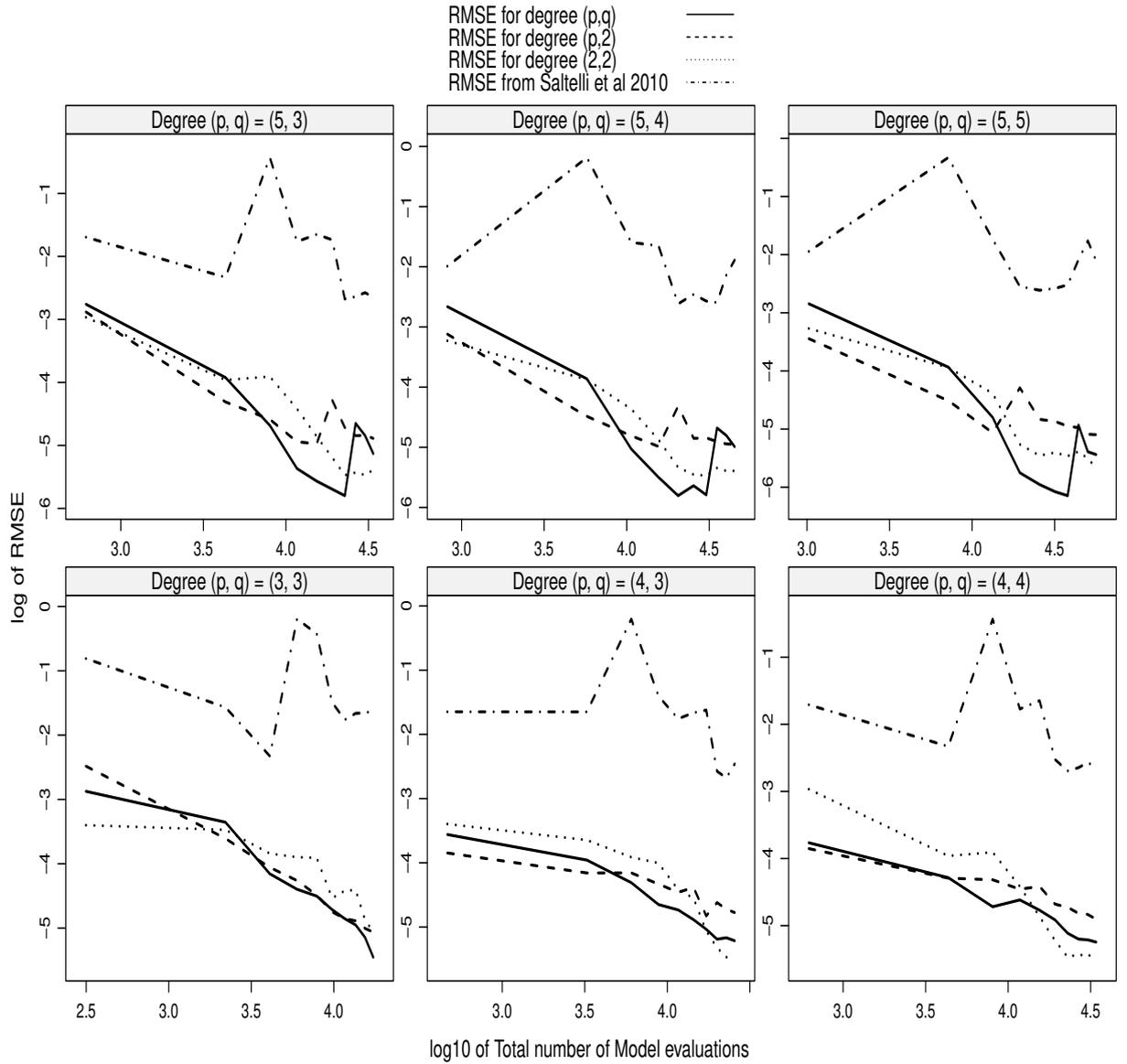


Figure 5: First-order indices: Log-RMSEs of the model in (6.50) of type C against the total number of model evaluations (in \log_{10}) for six values of the degree $(p = 3, q = 3)$, $(p = 4, q = 3)$, $(p = 4, q = 4)$, $(p = 5, q = 3)$, $(p = 5, q = 4)$, and $(p = 5, q = 5)$. For each degree, we show the corresponding RMSE (solid line), the RMSE for degree $(p, q = 2)$ (dashed line), the RMSE for degree $(2, 2)$ (dotted line), and the RMSE from [29] (dash-dotted line).

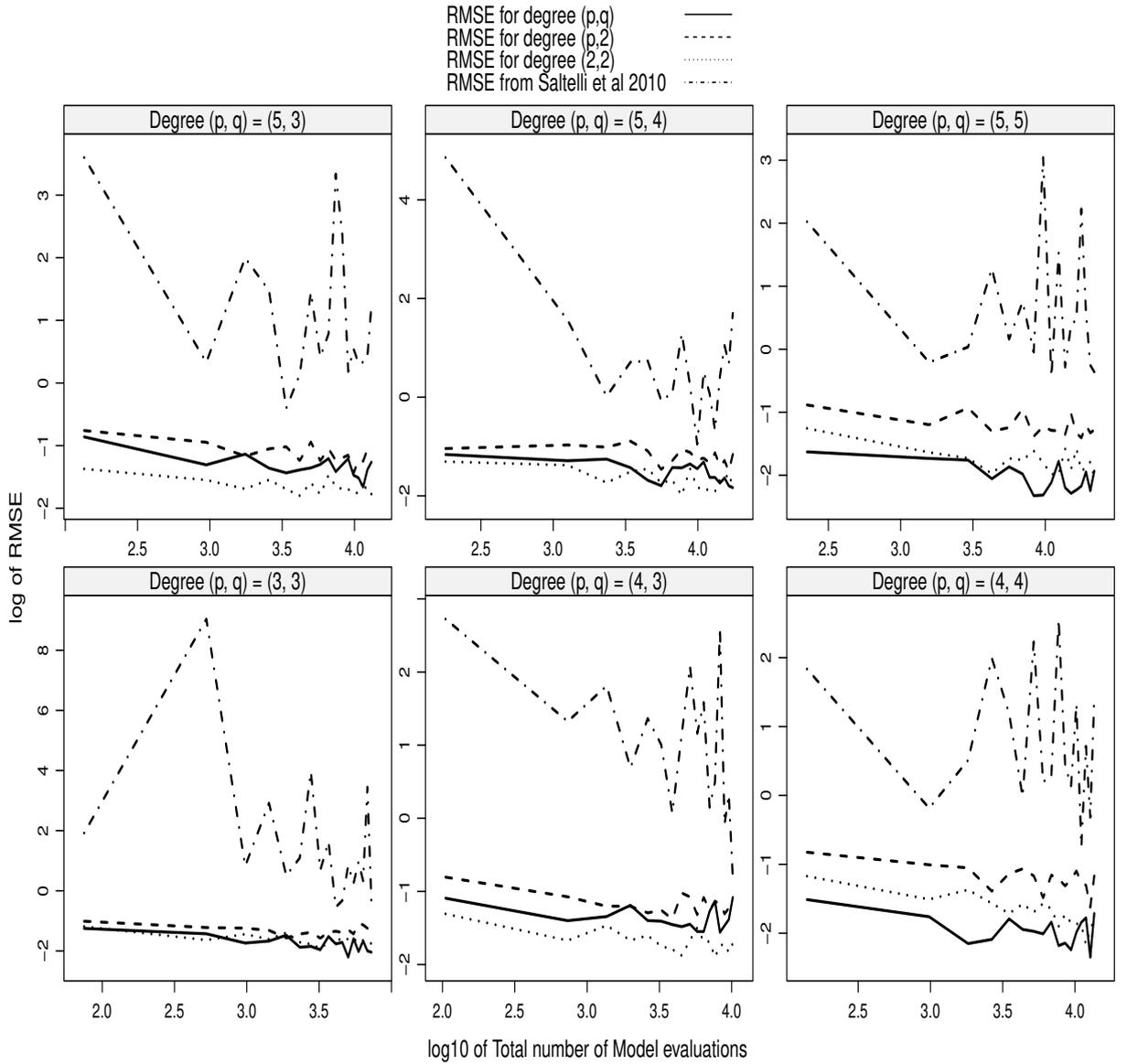


Figure 6: Total indices: Log-RMSEs of the model in (6.48) against the total number of model evaluations (in \log_{10}) for six values of the degree $(p = 3, q = 3)$, $(p = 4, q = 3)$, $(p = 4, q = 4)$, $(p = 5, q = 3)$, $(p = 5, q = 4)$, and $(p = 5, q = 5)$. For each degree, we show the corresponding RMSE (solid line), the RMSE for degree $(p, q = 2)$ (dashed line), the RMSE for degree $(2, 2)$ (dotted line), and the RMSE from [29] (dash-dotted line).

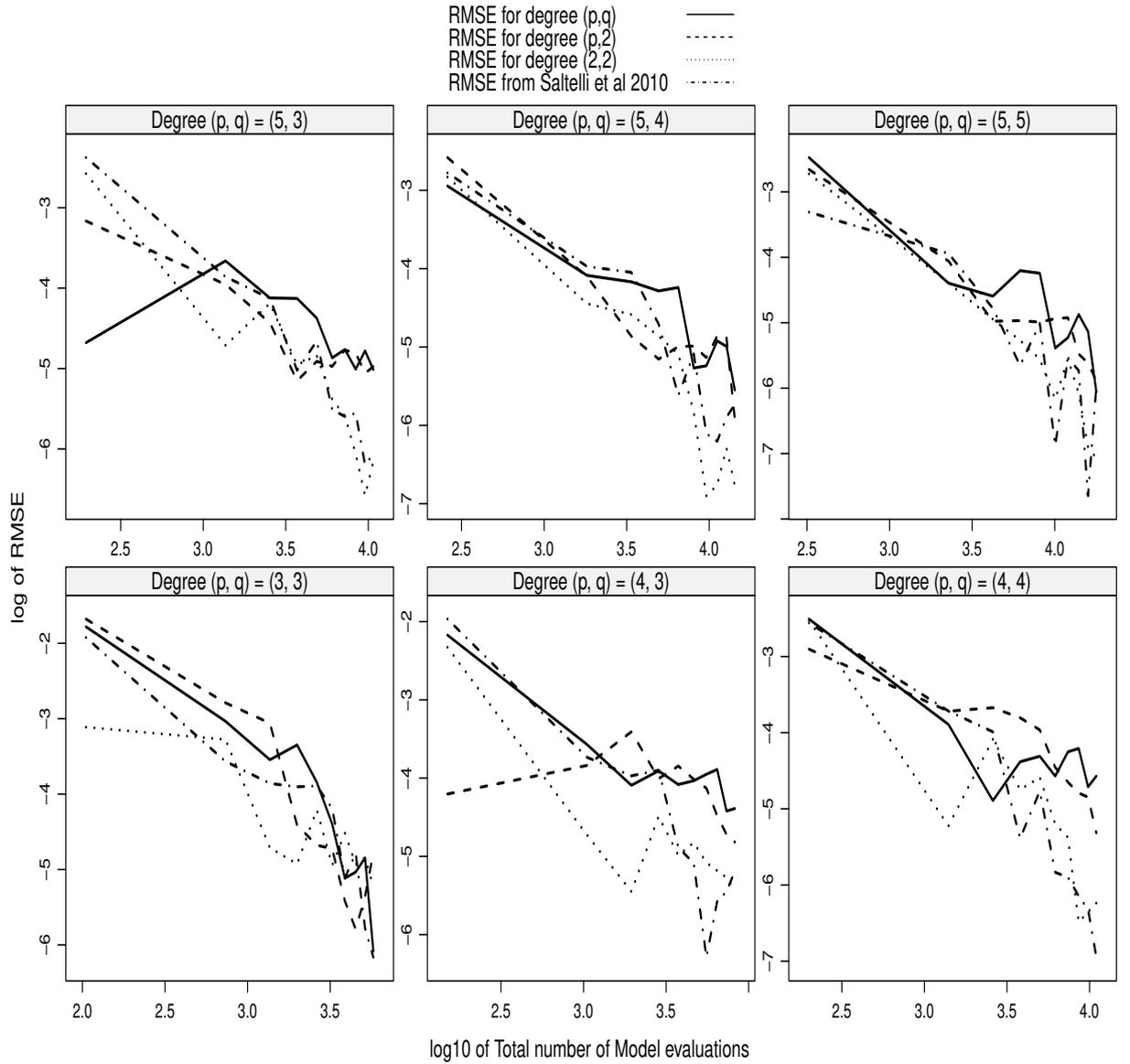


Figure 7: Total indices: Log-RMSEs of the model in (6.49) against the total number of model evaluations (in \log_{10}) for six values of the degree $(p = 3, q = 3)$, $(p = 4, q = 3)$, $(p = 4, q = 4)$, $(p = 5, q = 3)$, $(p = 5, q = 4)$, and $(p = 5, q = 5)$. For each degree, we show the corresponding RMSE (solid line), the RMSE for degree $(p, q = 2)$ (dashed line), the RMSE for degree $(2, 2)$ (dotted line), and the RMSE from [29] (dash-dotted line).

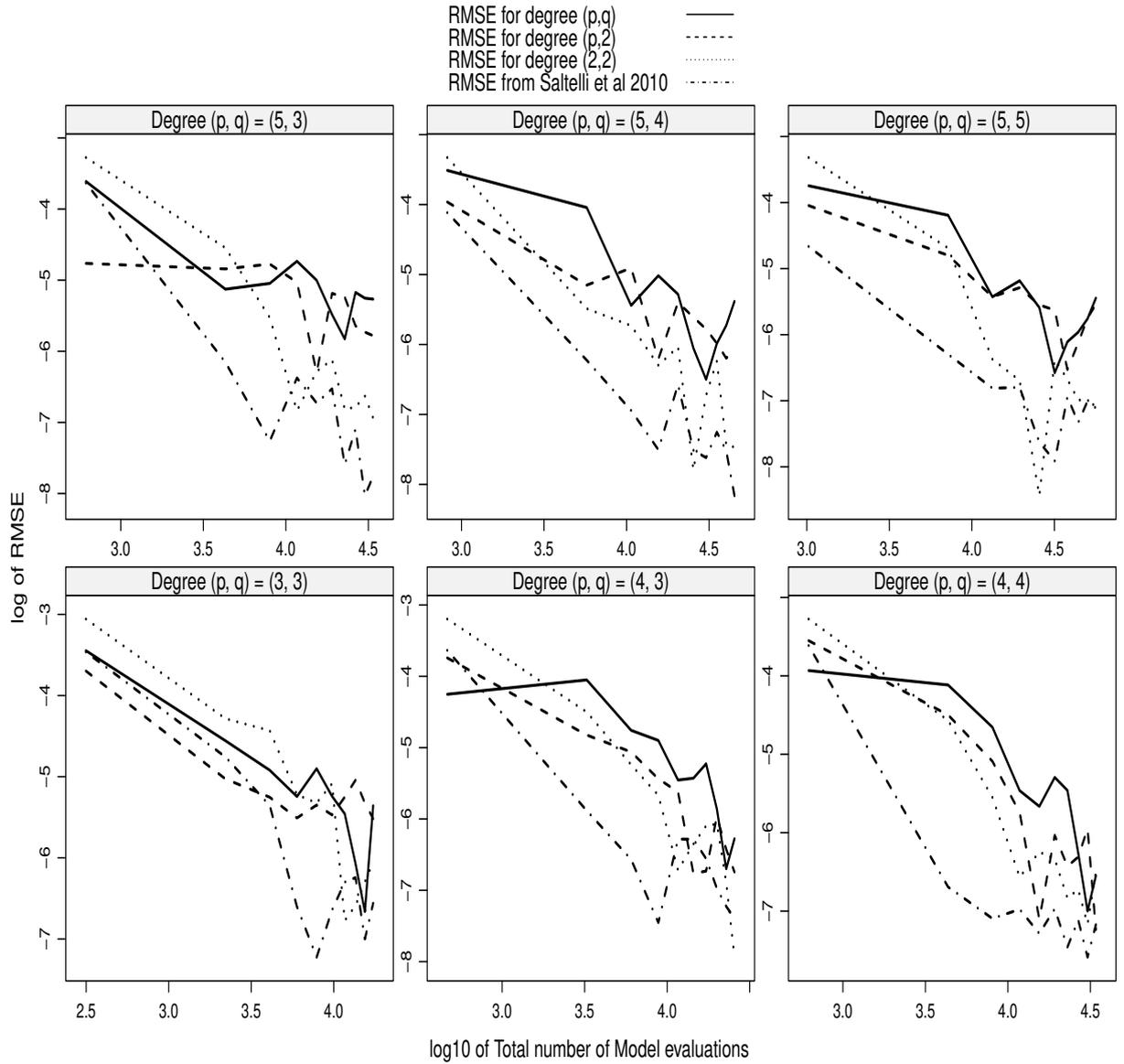


Figure 8: Total indices: Log-RMSEs of the model in (6.50) of type A against the total number of model evaluations (in \log_{10}) for six values of the degree ($p = 3, q = 3$), ($p = 4, q = 3$), ($p = 4, q = 4$), ($p = 5, q = 3$), ($p = 5, q = 4$), and ($p = 5, q = 5$). For each degree, we show the corresponding RMSE (solid line), the RMSE for degree ($p, q = 2$) (dashed line), the RMSE for degree (2, 2) (dotted line), and the RMSE from [29] (dash-dotted line).

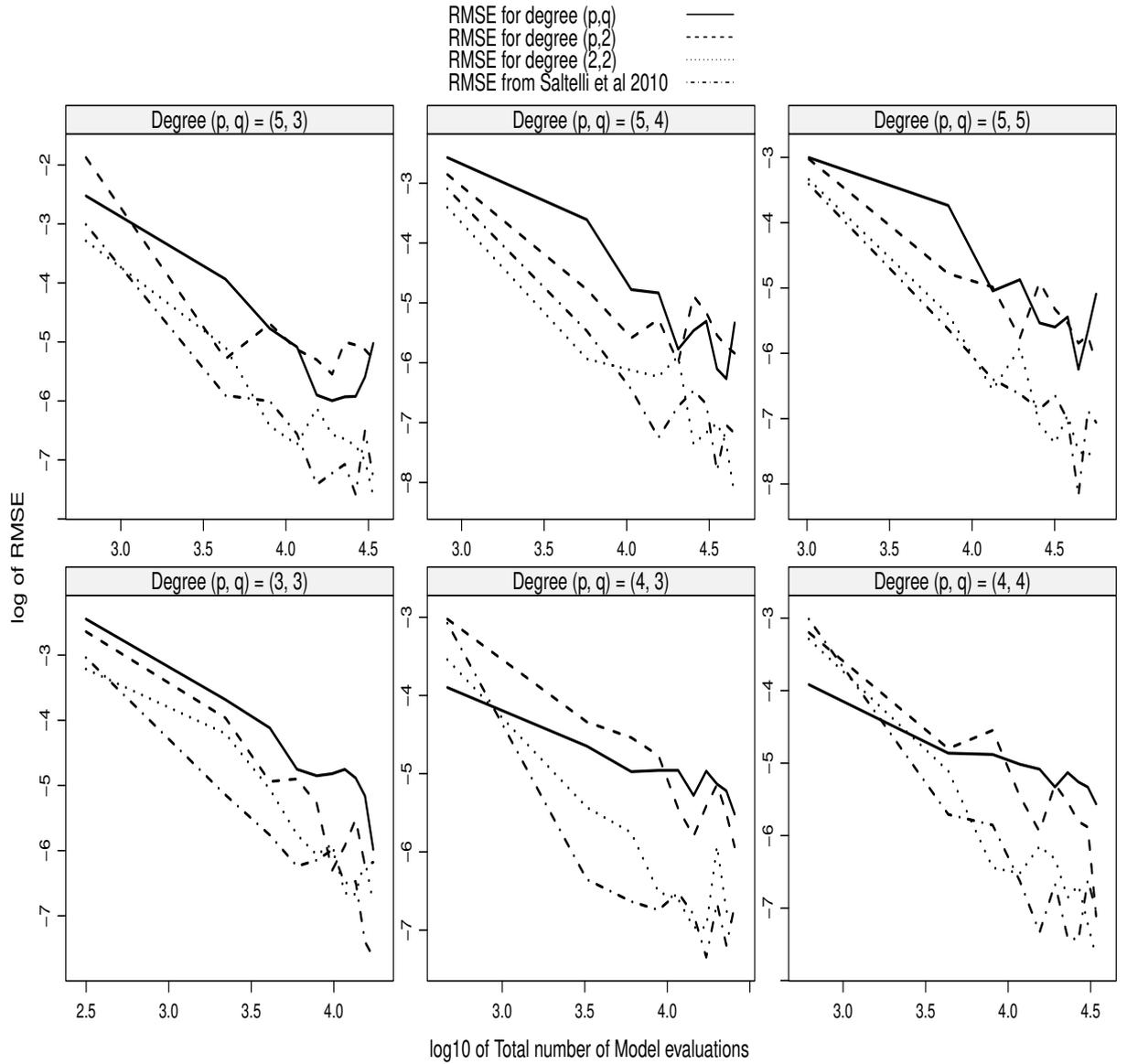


Figure 9: Total indices: Log-RMSEs of the model in (6.50) of type B against the total number of model evaluations (in \log_{10}) for six values of the degree $(p = 3, q = 3)$, $(p = 4, q = 3)$, $(p = 4, q = 4)$, $(p = 5, q = 3)$, $(p = 5, q = 4)$, and $(p = 5, q = 5)$. For each degree, we show the corresponding RMSE (solid line), the RMSE for degree $(p, q = 2)$ (dashed line), the RMSE for degree $(2, 2)$ (dotted line), and the RMSE from [29] (dash-dotted line).

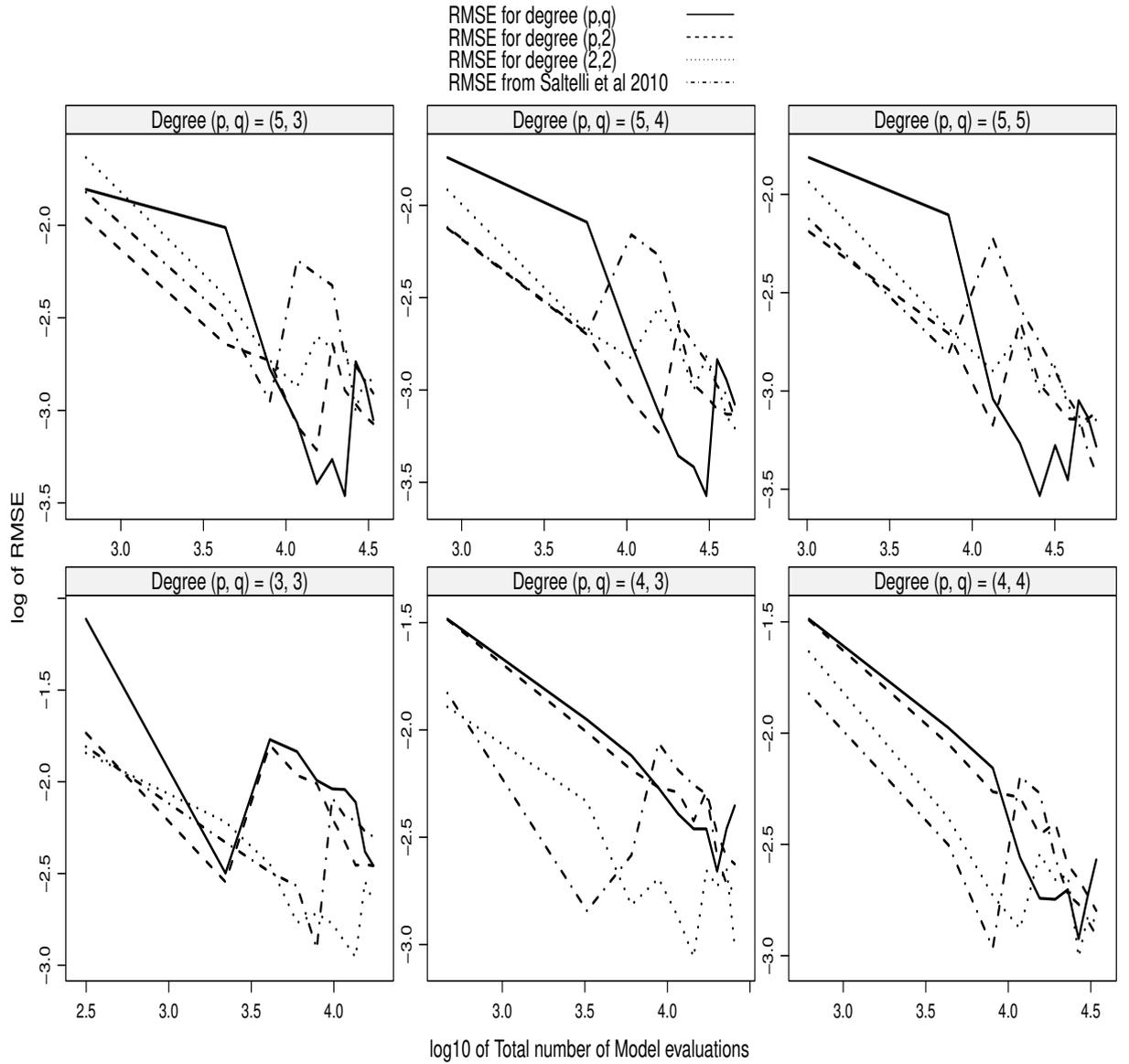


Figure 10: Total indices: Log-RMSEs of the model in (6.50) of type C against the total number of model evaluations (in \log_{10}) for six values of the degree $(p = 3, q = 3)$, $(p = 4, q = 3)$, $(p = 4, q = 4)$, $(p = 5, q = 3)$, $(p = 5, q = 4)$, and $(p = 5, q = 5)$. For each degree, we show the corresponding RMSE (solid line), the RMSE for degree $(p, q = 2)$ (dashed line), the RMSE for degree $(2, 2)$ (dotted line), and the RMSE from [29] (dash-dotted line).

1 **AppendixA. Proof of Proposition 1**

2 As $\left\| \mathbf{D}_v^{1/2} \right\|_S^2$ is the largest eigenvalue of \mathbf{D}_v , we have $\left\| \mathbf{D}_v^{1/2} \right\|_S^2 \leq \mathbb{T}_r(\mathbf{D}_v)$ for
 3 $v \subseteq \{1, \dots, d\}$, and $\mathbb{T}_r(\Sigma) = \sum_{v \subseteq \{1, \dots, d\}} \mathbb{T}_r(\mathbf{D}_v) \geq \sum_{v \subseteq \{1, \dots, d\}} \left\| \mathbf{D}_v^{1/2} \right\|_S^2$. Thus, we
 4 have $GSI_u^S \leq 1$ and $GSI_{T_u}^S \leq 1$.

5 Point i) is then straightforward bearing in mind $\Sigma_u \preceq \Sigma_u^{tot}$. Indeed, if $\Sigma_u \preceq \Sigma_u^{tot}$
 6 then $\left\| (\Sigma_u^{tot})^{1/2} \right\|_S^2 \geq \left\| \Sigma_u^{1/2} \right\|_S^2$.

7 For Point ii), it is sufficient to show that $\mathbb{C}_{ov}[\mathcal{V}\mathbf{f}(\mathbf{X})]$ and $\mathbb{C}_{ov}[\mathbf{f}(\mathbf{X})]$ have the same
 8 largest eigenvalue. It is obvious to see that the two matrices have the same eigen-
 9 values, and therefore, they have the same largest eigenvalue.

10 □

11 **AppendixB. Proof of Proposition 2**

12 As \mathbf{D}_v , $v \subseteq \{1, \dots, d\}$ are positive semi-definite matrices, we have $|(\mathbf{D}_v)_{ij}| \leq$
 13 $\frac{(\mathbf{D}_v)_{ii} + (\mathbf{D}_v)_{jj}}{2}$ and $\|\mathbf{D}_v\|_{l_1} = \sum_{i=1}^N |(\mathbf{D}_v)_{ij}| \leq \sum_{j=1}^N \frac{(\mathbf{D}_v)_{ii} + (\mathbf{D}_v)_{jj}}{2}$. By arranging the
 14 right-hand side, we obtain $\|\mathbf{D}_v\|_{l_1} \leq N\mathbb{T}_r(\mathbf{D}_v)$ and $\sum_{v \subseteq \{1, \dots, d\}} \|\mathbf{D}_v\|_{l_1} \leq N\mathbb{T}_r(\Sigma)$.

15 □

16 **AppendixC. Proof of Theorem 1**

17 First, we establish some useful equalities for the proofs, and second, we derive
 18 the two results of Theorem 1. Without loss of generality, we suppose that $\mathbf{f}(\mathbf{X})$ is
 19 centered, that is, $\mathbb{E}[\mathbf{f}(\mathbf{X})] = \mathbf{0}$.

20

21 Bearing in mind that $\mathbf{X}_u^{(j_1)}$ and $\mathbf{X}_{\sim u}^{(j_2)}$ are independent for $j_1 = 1, 2, \dots, p$ and
 22 $j_2 = 1, 2, \dots, q$, $\mathbf{X}_u^{(j_1)}$ and $\mathbf{X}_u^{(j_2)}$ (resp. $\mathbf{X}_{\sim u}^{(j_1)}$ and $\mathbf{X}_{\sim u}^{(j_2)}$) have the same distribution,
 23 we have the following equalities.

$$\mathbb{E}[\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)}) | \mathbf{X}_u^{(j_1)}] = \mathbb{E}_{\mathbf{X}_{\sim u}^{(j_1)}}[\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)})] = \mathbb{E}_{\mathbf{X}_{\sim u}^{(j_2)}}[\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_2)})], \quad (\text{C.1})$$

1 where $\mathbb{E}_{\mathbf{X}_{\sim u}^{(j_1)}} [\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)})]$ means that the expectation is taken with respect to
2 $\mathbf{X}_{\sim u}^{(j_1)}$.

3

4 Likewise, we have

$$\mathbb{E} [\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)}) | \mathbf{X}_{\sim u}^{(j_1)}] = \mathbb{E}_{\mathbf{X}_u^{(j_1)}} [\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)})] = \mathbb{E}_{\mathbf{X}_u^{(j_2)}} [\mathbf{f}(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(j_1)})] . \quad (\text{C.2})$$

5 According to Equation (3.5) and using Equation (C.1), the first-order covariance
6 matrix is also given by

$$\Sigma_u = \mathbb{E} [\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)}) \mathbf{f}(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(j_2)})^T] , \quad (\text{C.3})$$

7 with $j_1 \neq j_2$. Indeed,

$$\begin{aligned} \Sigma_u &= \mathbb{E} [\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)}) \mathbf{f}(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(j_2)})^T] \\ &= \mathbb{E} (\mathbb{E} [\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)}) \mathbf{f}(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(j_2)})^T | \mathbf{X}_u^{(j_1)}]) \\ &= \mathbb{E} (\mathbb{E} [\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)}) | \mathbf{X}_u^{(j_1)}] \mathbb{E} [\mathbf{f}(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(j_2)})^T | \mathbf{X}_u^{(j_1)}]) \\ &= \mathbb{E} \left(\mathbb{E}_{\mathbf{X}_{\sim u}^{(j_1)}} [\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)})] \mathbb{E}_{\mathbf{X}_{\sim u}^{(j_2)}} [\mathbf{f}(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(j_2)})^T] \right) \\ &= \mathbb{E} \left(\mathbb{E}_{\mathbf{X}_{\sim u}^{(j_1)}} [\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)})] \mathbb{E}_{\mathbf{X}_{\sim u}^{(j_1)}} [\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)})^T] \right) \\ &= \text{Cov} (\mathbb{E} [\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)}) | \mathbf{X}_u^{(j_1)}]) \end{aligned}$$

8 Now, we have all elements to start the proof of Equation (4.27) related to the
9 first-order covariance matrix $(\mathbf{K}(\cdot))$.

$$\begin{aligned}
& p^2(p-1)q(q-1)\mathbb{E} [\mathbf{K}(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})] \\
= & \sum_{j=1}^p \sum_{l=1}^q \sum_{\substack{i=1 \\ i \neq l}}^q \mathbb{E} \left[\left((p-1)\mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) - \sum_{\substack{j_1=1 \\ j_1 \neq j}}^p \mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)}) \right) \right. \\
& \times \left. \left((p-1)\mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(i)}) - \sum_{\substack{j_2=1 \\ j_2 \neq j}}^p \mathbf{f}(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(i)}) \right)^T \right] \\
= & \sum_{\substack{j=1 \\ l=1 \\ i=1 \\ i \neq l}}^{p,q,q} \mathbb{E} [(p-1)^2 \mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) \mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(i)})^T] \\
& - \sum_{\substack{j=1 \\ l=1 \\ i=1 \\ i \neq l}}^{p,q,q} \mathbb{E} \left[(p-1) \sum_{\substack{j_2=1 \\ j_2 \neq j}}^p \mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) \mathbf{f}(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(i)})^T \right] \\
& + \sum_{\substack{j=1 \\ l=1 \\ i=1 \\ i \neq l}}^{p,q,q} \mathbb{E} \left[(1-p) \sum_{\substack{j_1=1 \\ j_1 \neq j}}^p \mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)}) \mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(i)})^T \right] \\
& + \sum_{\substack{j=1 \\ l=1 \\ i=1 \\ i \neq l}}^{p,q,q} \mathbb{E} \left[\sum_{\substack{j_1=1 \\ j_2=1 \\ j_1 \neq j \\ j_2 \neq j}}^{p,p} \mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)}) \mathbf{f}(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(i)})^T \right].
\end{aligned}$$

As the functions $\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)})$ and $\mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(i)})$ are independent for $j_1 \neq j$ and $i \neq l$ and the function $\mathbf{f}(\cdot)$ is centered, we have

$$\mathbb{E} [\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)}) \mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(i)})^T] = \mathbb{E} [\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)})] \mathbb{E} [\mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(i)})^T] = \mathbf{0}.$$

Likewise, we have

$$\mathbb{E} [\mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)})\mathbf{f}(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(i)})^T] = \mathbf{0} \quad \text{if } j_2 \neq j, l \neq i.$$

1 Using the linearity of the expectation and the above equations, we can write

$$\begin{aligned} & p^2(p-1)q(q-1)\mathbb{E} [\mathbf{K}(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})] \\ = & \sum_{\substack{j=1 \\ l=1 \\ i=1 \\ i \neq l}}^{p,q,q} \left[(p-1)^2 \mathbb{E} [\mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)})\mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(i)})^T] + \sum_{\substack{j_1=1 \\ j_2=1 \\ j_1 \neq j \\ j_2 \neq j}}^{p,p} \mathbb{E} [\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)})\mathbf{f}(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(i)})^T] \right]. \end{aligned}$$

2 By using Equation (C.3), we have

$$\begin{aligned} & p^2(p-1)q(q-1)\mathbb{E} [\mathbf{K}(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})] \\ = & \sum_{j=1, l=1, i=1, i \neq l}^{p,q,q} \left[(p-1)^2 \Sigma_u + \sum_{\substack{j_1=1 \\ j_2=j_1 \\ j_1 \neq j}}^p \mathbb{E} [\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)})\mathbf{f}(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(i)})^T] \right] \\ = & \sum_{j=1, l=1, i=1, i \neq l}^{p,q,q} \left[(p-1)^2 \Sigma_u + \sum_{\substack{j_1=1 \\ j_1 \neq j}}^p \mathbb{E} [\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)})\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(i)})^T] \right] \\ = & \sum_{j=1, l=1, i=1, i \neq l}^{p,q,q} \left[(p-1)^2 \Sigma_u + \sum_{j_1=1, j_1 \neq j}^p \Sigma_u \right] \\ = & \sum_{j=1, l=1, i=1, i \neq l}^{p,q,q} [(p-1)^2 \Sigma_u + (p-1) \Sigma_u] \\ = & \sum_{j=1, l=1, i=1, i \neq l}^{p,q,q} [p(p-1) \Sigma_u] \\ = & p^2(p-1)q(q-1) \Sigma_u. \end{aligned}$$

1

□

2 The derivation for the expectation of $\mathbf{K}^{tot}(\cdot)$ in Equation (4.28) is similar to the
 3 proof of $\mathbf{K}(\cdot)$, and it is given below.

4

5 As $\mathbf{g}_u(\mathbf{X}_u, \mathbf{X}_{\sim u}) = \mathbf{f}(\mathbf{X}_u, \mathbf{X}_{\sim u}) - \mathbb{E}_{\mathbf{X}_u}[\mathbf{f}(\mathbf{X}_u, \mathbf{X}_{\sim u})]$ (see Equation (2.3)), the ex-
 6 pectation $\mathbb{E}[\mathbf{g}_u(\mathbf{X}_u, \mathbf{X}_{\sim u})] = \mathbb{E}[\mathbf{g}_u(\mathbf{X}_u, \mathbf{X}_{\sim u}) | \mathbf{X}_{\sim u}] = \mathbf{0}$. For $j_1 \neq j_2$, we have

$$\mathbb{E}[\mathbf{g}_u(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(j_1)})\mathbf{g}_u(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)})^T] = \mathbf{0}. \quad (\text{C.4})$$

7 Indeed,

$$\begin{aligned} A &= \mathbb{E}[\mathbf{g}_u(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(j_1)})\mathbf{g}_u(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)})^T] \\ &= \mathbb{E}[\mathbb{E}(\mathbf{g}_u(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(j_1)})\mathbf{g}_u(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)})^T | \mathbf{X}_{\sim u}^{(j_1)})] \\ &= \mathbb{E}[\mathbb{E}(\mathbf{g}_u(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(j_1)}) | \mathbf{X}_{\sim u}^{(j_1)}) \mathbb{E}(\mathbf{g}_u(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)})^T | \mathbf{X}_{\sim u}^{(j_1)})] \\ &= \mathbb{E}[\mathbf{0}\mathbf{0}^T] \end{aligned}$$

8 The expectation of \mathbf{K}^{tot} is given by

$$\begin{aligned} &qp^2(p-1)\mathbb{E}[\mathbf{K}^{tot}(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})] \\ &= \sum_{\substack{j=1 \\ l=1}}^{p,q} \mathbb{E} \left[\left(\sum_{\substack{j_1=1 \\ j_1 \neq j}}^p [\mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) - \mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)})] \right) \right. \\ &\quad \left. \times \left(\sum_{\substack{j_2=1 \\ j_2 \neq j}}^p [\mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) - \mathbf{f}(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(l)})] \right)^T \right]. \end{aligned}$$

1 Bearing in mind Equation (C.2), we have

$$\begin{aligned}
\mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) - \mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)}) &= \mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) - \mathbb{E}_{\mathbf{X}_u^{(j)}} [\mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)})] \\
&\quad + \mathbb{E}_{\mathbf{X}_u^{(j_1)}} [\mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)})] - \mathbf{f}(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)}) \\
&= \mathbf{g}_u(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) - \mathbf{g}_u(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)}).
\end{aligned}$$

Likewise, we have

$$\mathbf{f}(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) - \mathbf{f}(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(l)}) = \mathbf{g}_u(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) - \mathbf{g}_u(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(l)}).$$

2 Thus, the expectation of \mathbf{K}^{tot} becomes

$$\begin{aligned}
& qp^2(p-1)\mathbb{E}[\mathbf{K}^{tot}(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})] \\
&= \sum_{\substack{j=1 \\ l=1}}^{p,q} \mathbb{E} \left[\left(\sum_{\substack{j_1=1 \\ j_1 \neq j}}^p [\mathbf{g}_u(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) - \mathbf{g}_u(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)})] \right) \right. \\
&\quad \times \left. \left(\sum_{\substack{j_2=1 \\ j_2 \neq j}}^p [\mathbf{g}_u(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) - \mathbf{g}_u(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(l)})] \right)^T \right] \\
&= \sum_{\substack{j=1 \\ l=1}}^{p,q} \mathbb{E} \left[\left((p-1)\mathbf{g}_u(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) - \sum_{\substack{j_1=1 \\ j_1 \neq j}}^p \mathbf{g}_u(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)}) \right) \right. \\
&\quad \times \left. \left((p-1)\mathbf{g}_u(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) - \sum_{\substack{j_2=1 \\ j_2 \neq j}}^p \mathbf{g}_u(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(l)}) \right)^T \right] \\
&= \sum_{\substack{j=1 \\ l=1}}^{p,q} \mathbb{E} [(p-1)^2 \mathbf{g}_u(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) \mathbf{g}_u(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)})^T] \\
&\quad - \sum_{\substack{j=1 \\ l=1}}^{p,q} \mathbb{E} \left[(p-1) \sum_{\substack{j_1=1 \\ j_1 \neq j}}^p \mathbf{g}_u(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)}) \mathbf{g}_u(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)})^T \right] \\
&\quad - \sum_{\substack{j=1 \\ l=1}}^{p,q} \mathbb{E} \left[(p-1) \sum_{\substack{j_2=1 \\ j_2 \neq j}}^p \mathbf{g}_u(\mathbf{X}_u^{(j)}, \mathbf{X}_{\sim u}^{(l)}) \mathbf{g}_u(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(l)})^T \right] \\
&\quad + \sum_{\substack{j=1 \\ l=1}}^{p,q} \mathbb{E} \left[\sum_{\substack{j_1=1 \\ j_1 \neq j \\ j_2=1 \\ j_2 \neq j}}^p \mathbf{g}_u(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)}) \mathbf{g}_u(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(l)})^T \right].
\end{aligned}$$

- 1 Using the linearity of the expectation, Equation (C.4), and the fact that $\Sigma_u^{tot} =$
- 2 $\mathbb{C}_{ov}[\mathbf{g}_u(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)})] = \mathbb{E}[\mathbf{g}_u(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)}) \mathbf{g}_u(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(j_1)})^T]$, we obtain the result,

1 that is,

$$\begin{aligned}
& qp^2(p-1)\mathbb{E}[\mathbf{K}^{tot}(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)})] \\
&= \sum_{\substack{j=1 \\ l=1}}^{p,q} \left[(p-1)^2 \Sigma_u^{tot} + \sum_{\substack{j_1=1 \\ j_1 \neq j \\ j_2=1 \\ j_2 \neq j}}^p \mathbb{E}[\mathbf{g}_u(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)}) \mathbf{g}_u(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(l)})^T] \right] \\
&= \sum_{\substack{j=1 \\ l=1}}^{p,q} \left[(p-1)^2 \Sigma_u^{tot} + \sum_{\substack{j_1=1 \\ j_1 \neq j \\ j_2=j_1}}^p \mathbb{E}[\mathbf{g}_u(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)}) \mathbf{g}_u(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(l)})^T] \right] \\
&= \sum_{\substack{j=1 \\ l=1}}^{p,q} \left[(p-1)^2 \Sigma_u^{tot} + \sum_{\substack{j_1=1 \\ j_1 \neq j}}^p \mathbb{E}[\mathbf{g}_u(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)}) \mathbf{g}_u(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)})^T] \right] \\
&= \sum_{\substack{j=1 \\ l=1}}^{p,q} \left[(p-1)^2 \Sigma_u^{tot} + \sum_{\substack{j_1=1 \\ j_1 \neq j}}^p \Sigma_u^{tot} \right] \\
&= \sum_{\substack{j=1 \\ l=1}}^{p,q} [(p-1)^2 \Sigma_u^{tot} + (p-1) \Sigma_u^{tot}] \\
&= \sum_{\substack{j=1 \\ l=1}}^{p,q} [p(p-1) \Sigma_u^{tot}] \\
&= qp^2(p-1) \Sigma_u^{tot}.
\end{aligned}$$

2

□

3 AppendixD. Proof of Theorem 2

Consider $\mathcal{Z}_i = (\mathcal{X}_i, \mathcal{Y}_i)$, $i = 1, 2, \dots, m$, an independent sample of size m .

Theorem 1 shows that

$$\mathbb{E}[\mathbf{K}(\mathcal{Z}_1)] = \Sigma_u.$$

1 The corresponding U-statistic of Σ_u is

$$\begin{aligned}\widehat{\Sigma}_u &= \frac{1}{m} \sum_{i=1}^m \mathbf{K}(\mathcal{Z}_i) \\ &= \frac{1}{m} \sum_{i=1}^m \mathbf{K}(\mathbf{X}_{i,u}^{(1)}, \dots, \mathbf{X}_{i,u}^{(p)}, \mathbf{X}_{i,\sim u}^{(1)}, \dots, \mathbf{X}_{i,\sim u}^{(q)}).\end{aligned}$$

2 The U-statistic $\widehat{\Sigma}_u$ is an average of the kernel $\mathbf{K}(\mathcal{Z}_i)$ over $i = 1, 2, \dots, m$, and
3 it is symmetric, as the kernel $\mathbf{K}(\cdot)$ is symmetric.

4

5 **Now, let** $\text{vec}(\widehat{\Sigma}_u)$ be the vectorization of $\widehat{\Sigma}_u$, that is, a vector containing the
6 first column of $\widehat{\Sigma}_u$, followed by the second column, and so on. It is obvious that
7 $\text{vec}(\widehat{\Sigma}_u)$ is an unbiased estimator of $\text{vec}(\Sigma_u)$, and it is symmetric w.r.t the two type
8 of inputs. Then, it follows from the theory of U-statistics ([21; 16; 23; 33]) that
9 $\text{vec}(\widehat{\Sigma}_u)$ is the unique, uniformly minimum variance unbiased estimator of $\text{vec}(\Sigma_u)$
10 for the class of functions having finite 4th moment, and we have the following results:

11

(i) unbiasedness

$$\mathbb{E} \left[\text{vec}(\widehat{\Sigma}_u) \right] = \text{vec}(\Sigma_u);$$

(ii) consistency

$$\text{vec}(\widehat{\Sigma}_u) \xrightarrow{\mathcal{P}} \text{vec}(\Sigma_u);$$

(iii) asymptotic normality

$$\sqrt{m} \left(\text{vec}(\widehat{\Sigma}_u) - \text{vec}(\Sigma_u) \right) \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}, \mathbb{V}[\text{vec}(\mathbf{K})]).$$

Points (i) and (ii) are equivalent to

$$\mathbb{E} \left[\widehat{\Sigma}_u \right] = \Sigma_u \text{ and } \widehat{\Sigma}_u \xrightarrow{\mathcal{P}} \Sigma_u.$$

1

□

2 **AppendixE. Proof of Theorem 3**

3

The proof is similar to the proof derived in AppendixD by replacing \mathbf{K} with \mathbf{K}^{tot} .

4

□

5 **AppendixF. Proof of Corollary 2**

6

For Point i), the proof is similar to the proofs in Sections AppendixD and AppendixE.

7

8

Points ii) and iii) are the properties of the multivariate U-statistic associated with the kernel $\mathbf{K}_t(\cdot)$. It can be found in [21; 16; 23; 33].

9

10

□

11 **AppendixG. Proof of Theorem 4**

12

Point i) is obvious using the definition of GSIs.

13

Results ii) and iii) are obtained by combining Corollary 2 and Slutsky's theorem.

14

Comprehensive details of the proof can be found in [14; 15].

15

Point iv) represents the classical confidence regions for a mean vector associated with a multivariate normal distribution.

16

17

□

18 **AppendixH. Proxy measure for the GSI variances**

19

We use the MINQE approach to obtain the best degree of the kernel $\mathbf{K}_t(\cdot)$.

20

First, the idea of MINQE consists in expressing the estimator of a sensitivity index

21

as $\mathbb{T}_r(\boldsymbol{\Omega}^T \hat{\boldsymbol{\Theta}})$ for given coefficients $\boldsymbol{\Omega}$. The components of the matrix $\hat{\boldsymbol{\Theta}}$ are the

22

product of functions $f(\mathbf{X}_u, \mathbf{X}_{\sim u})$ and $f(\mathbf{X}_v, \mathbf{X}_{\sim v})$, with $u, v \subseteq \{1, 2, \dots, d\}$ (see

1 [24; 14; 15] for a comprehensive treatment). Second, a proxy measure for the variance
 2 of the estimator is given by ([24])

$$V(\boldsymbol{\Omega}) = \left(\sum_{u \subseteq \{1, 2, \dots, d\}} \sum_{v \subseteq \{1, 2, \dots, d\}} |\boldsymbol{\Omega}_{u,v}| \right)^2. \quad (\text{H.1})$$

3 **From Equation** (4.42), it comes out that the mean square error of the estimator
 4 of the non-normalized GSIs is the trace of the covariance of the kernel $\mathbf{K}_t(\cdot)$ up to a
 5 constant. The trace of the covariance of $\mathbf{K}_t(\cdot)$ is the sum of the variances of $K^{tot}(\cdot)$
 6 and $K(\cdot)$. To identify the best degree of $\mathbf{K}_t(\cdot)$, we start with the methodology used
 7 in [14; 15].

8

For integers j_1, j_2, k, l , and i , let $(\hat{\boldsymbol{\Theta}}_{j_1, l}^k)$ be a matrix of type $(p \times q - l)$, with

$$(\hat{\boldsymbol{\Theta}}_{j_1, l}^k)_{j_2, i} = f(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)}) f(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(i)}),$$

9 $j_2 = 1, \dots, p$, and $i = l + 1, \dots, q$.

10

11 We use $\mathbf{r}_{\bullet j_1}^{(k)} = [r_1^{(k)} r_{j_1}^{(k)}, \dots, r_p^{(k)} r_{j_1}^{(k)}]^T$ as a vector of size p , with $r_i^{(k)} = -1$ if
 12 $i \neq k$ and $r_k^{(k)} = p - 1$ otherwise, $\hat{\boldsymbol{\Theta}}^k = \text{diag}(\hat{\boldsymbol{\Theta}}_{1,1}^k, \dots, \hat{\boldsymbol{\Theta}}_{p, q-1}^k)$ as a block diago-
 13 nal matrix of type $p^2 \times q(q-1)/2$, and $\mathbf{r}^{(k)} = [\mathbf{r}_{\bullet 1}^{(k)T}, \dots, \mathbf{r}_{\bullet p}^{(k)T}]^T$ as a vector of size p^2 .

14

15 Finally, we use $\hat{\boldsymbol{\Theta}} = \text{diag}(\hat{\boldsymbol{\Theta}}^1, \dots, \hat{\boldsymbol{\Theta}}^p)$ as a block diagonal matrix of type
 16 $p^3 \times pq(q-1)/2$, $\mathbf{r} = [\mathbf{r}^{(1)T}, \dots, \mathbf{r}^{(p)T}]^T$ as a vector of size p^3 , and $\mathbf{s} = [1, \dots, 1]^T$ as
 17 a vector of size $qp(q-1)/2$.

18

19 Now, consider $\mathbf{f}(\cdot) = [f_1(\cdot) f_2(\cdot) \dots f_N(\cdot)]^T$. As the trace of a matrix is the sum
 20 of the diagonal elements, the kernel $K(\cdot)$ in (4.36) can be written as follows:

21

$$\begin{aligned}
K(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)}) &= \frac{2}{p^2(p-1)q(q-1)} \sum_{n=1}^N \sum_{k=1}^p \sum_{l=1}^{q-1} \sum_{i=l+1}^q \\
&\sum_{j_1=1}^p \sum_{j_2=1}^p r_{j_1}^{(k)} r_{j_2}^{(k)} f_n(\mathbf{X}_u^{(j_1)}, \mathbf{X}_{\sim u}^{(l)}) f_n(\mathbf{X}_u^{(j_2)}, \mathbf{X}_{\sim u}^{(i)}),
\end{aligned} \tag{H.2}$$

1 or as a bi-linear form, that is,

$$\begin{aligned}
K(\mathbf{X}_u^{(1)}, \dots, \mathbf{X}_u^{(p)}, \mathbf{X}_{\sim u}^{(1)}, \dots, \mathbf{X}_{\sim u}^{(q)}) &= \frac{2}{p^2(p-1)q(q-1)} \sum_{n=1}^N \mathbf{r}^T \hat{\Theta}_n \mathbf{s} \\
&= \frac{2}{p^2(p-1)q(q-1)} \mathbb{T}_r \left(\sum_{n=1}^N \mathbf{s} \mathbf{r}^T \hat{\Theta}_n \right).
\end{aligned}$$

2 Using the definition of the proxy measure in (H.1), the proxy measure for the
3 variance of the kernel $K(\cdot)$ is given by

$$\begin{aligned}
V(\Omega) &= V \left(\frac{2}{p^2(p-1)q(q-1)} \sum_{n=1}^N \mathbf{s} \mathbf{r}^T \right) \\
&= \frac{4}{p^4(p-1)^2 q^2 (q-1)^2} V \left(\sum_{n=1}^N \mathbf{s} \mathbf{r}^T \right) \\
&= \frac{4}{p^4(p-1)^2 q^2 (q-1)^2} \left(\frac{qp(q-1)N}{2} \sum_{i=1}^{p^3} |r_i| \right)^2 \\
&= \frac{N^2}{p^2(p-1)^2} \left(\sum_{i=1}^{p^3} |r_i| \right)^2,
\end{aligned} \tag{H.3}$$

4 with r_i the i^{th} coordinate of the vector \mathbf{r} .

5

6 It comes out that the proxy measure does not depend on the value of q . Thus, for
7 a given value of p , any value of q will yield to the same value of the proxy measure
8 of the variance. The value of $q = 2$ should be used as the referenced value, and we

1 should expect to have good results with $q \leq p$.

2

3 We can see that the value of $V(\mathbf{\Omega})$ increases with p ($V[\mathbf{\Omega}] = 16N^2$ if $p = 2$
4 and $V[\mathbf{\Omega}] = 64N^2$ if $p = 3$), and we should use small values of p . We use 2 as the
5 referenced value of p .

6

7 The proxy measure for the variance of the kernel $K(\cdot)$ is sufficient for choosing
8 the degree of the Kernel $\mathbf{K}_t(\cdot)$, as the model runs performed for estimating the
9 first-order GSI are sufficient for estimating the total GSI.

10 AppendixI. **Asymptotic confident regions of classical GSIs**

11 **Figures** I.11-I.12 show the asymptotic 90% confident regions of the classical GSIs
12 for the multivariate exponential and Ishigami functions respectively. In Figures I.11-
13 I.12, the confident regions were obtained when $m = 2000$ and $(p = 2, q = 2)$ by using
14 the R-package car.

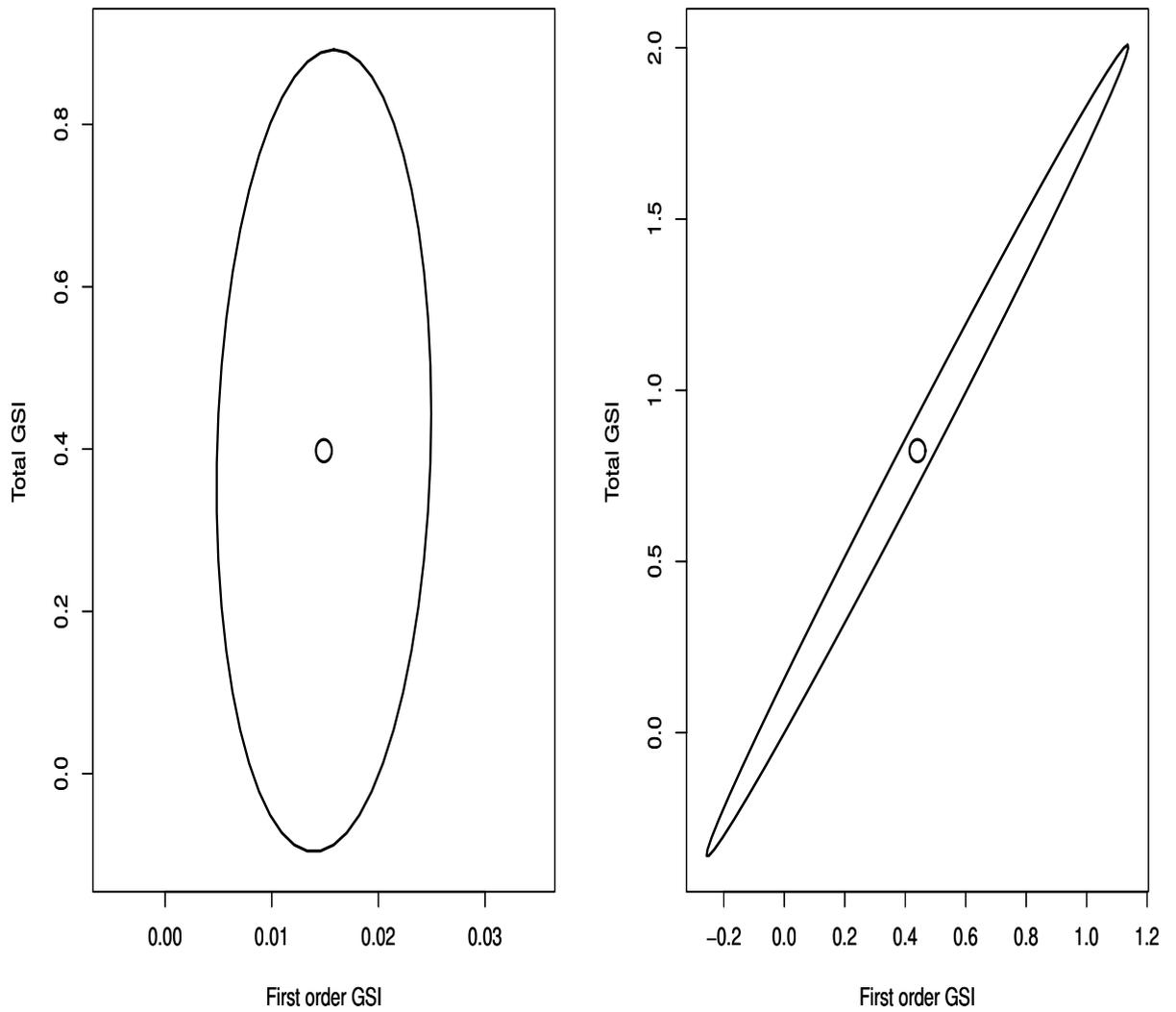


Figure I.11: Asymptotic 90% confident regions of the classical GSIs for the multivariate exponential function. The first ellipsoid is the confident regions of the GSIs of X_1 and the second one is the confident regions of the GSIs of X_2 .

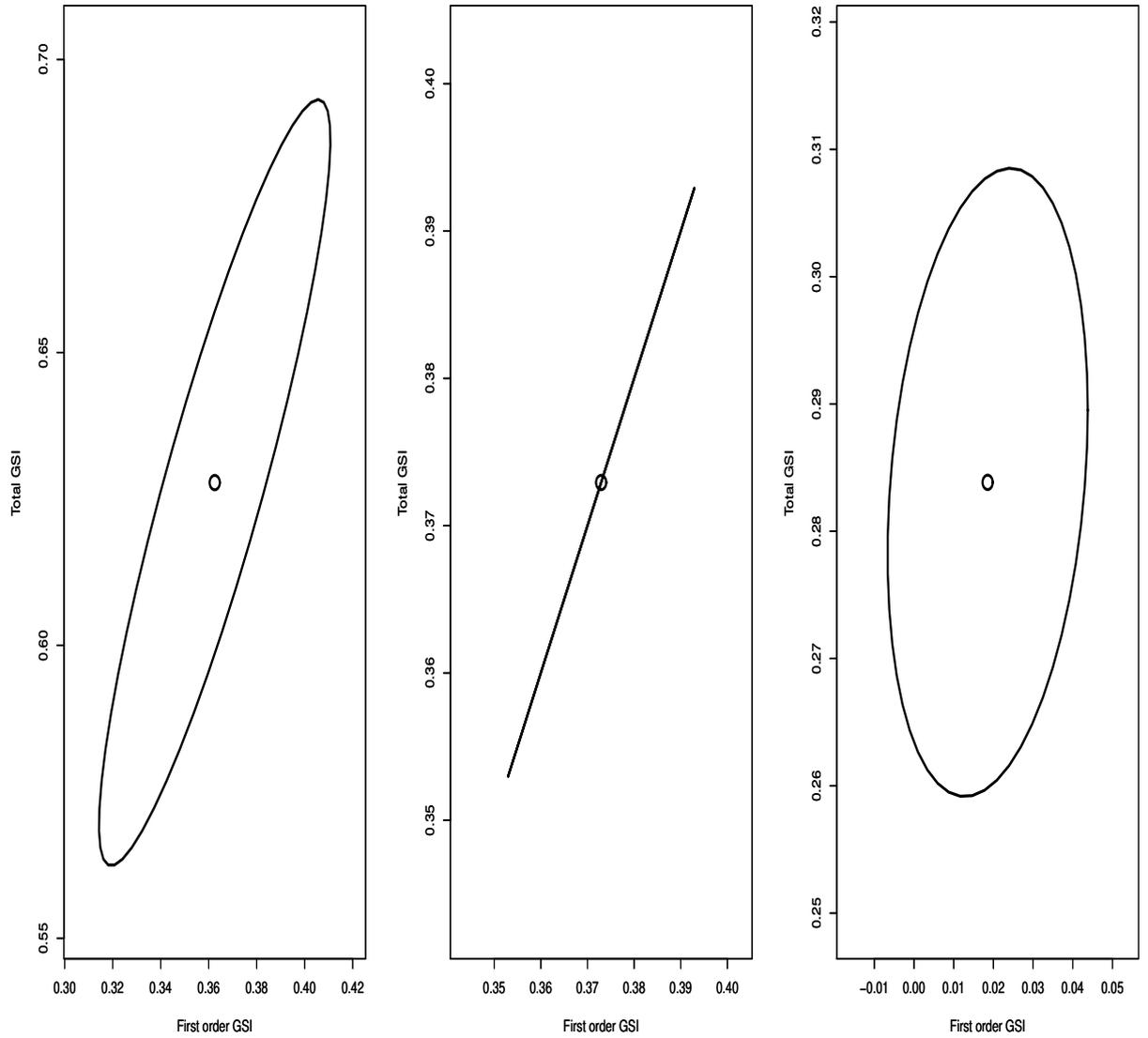


Figure I.12: Asymptotic 90% confident regions of the classical GSIs for the multivariate Ishigami function. The first ellipsoid is the confident regions of the GSIs of X_1 , the second one is the confident regions of the GSIs of X_2 , and the last one is the confident regions of the GSIs of X_3 .