



**HAL**  
open science

# An optimal tradeoff between explorations and repetitions in global sensitivity analysis for stochastic computer models

Gildas Mazo

► **To cite this version:**

Gildas Mazo. An optimal tradeoff between explorations and repetitions in global sensitivity analysis for stochastic computer models. 2020. hal-02113448v3

**HAL Id: hal-02113448**

**<https://hal.science/hal-02113448v3>**

Preprint submitted on 22 May 2020 (v3), last revised 7 Jun 2021 (v7)

**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.

1 **AN OPTIMAL TRADEOFF BETWEEN EXPLORATIONS AND**  
2 **REPETITIONS IN GLOBAL SENSITIVITY ANALYSIS FOR**  
3 **STOCHASTIC MODELS\***

4 GILDAS MAZO<sup>†</sup>

5 **Abstract.** Global sensitivity analysis often accompanies computer modeling to understand what  
6 are the important factors of a model of interest. In particular, Sobol indices, naturally estimated  
7 by Monte-Carlo methods, permit to quantify the contribution of the inputs to the variability of the  
8 output. However, stochastic computer models raise difficulties. There is no unique definition of Sobol  
9 indices and their estimation is difficult because a good balance between repetitions of the computer  
10 code and explorations of the input space must be found. The problem of finding an optimal tradeoff  
11 between explorations and repetitions is addressed. Two kinds of Sobol-like indices are considered.  
12 Their estimators are built and their asymptotic properties are established. To find an optimal tradeoff  
13 between repetitions and explorations, an error criterion that penalizes bad rankings of the inputs  
14 is considered. A bound is found and minimized under a fixed computing budget. Estimators that  
15 asymptotically achieve the minimal bound are built. Numerical tests are performed.

16 **Key words.** Explorations, repetitions, Sobol, estimation, sensitivity, stochastic, random, model.

17 **AMS subject classifications.** 62G20, 60H99, 65C05

18 **1. Introduction.** Sensitivity analysis often accompanies computer modeling to  
19 understand what are the important factors of a model of interest [17, 18]. In partic-  
20 ular, Sobol indices [20, 21] permit to quantify the contribution of the inputs to the  
21 variability of the output. The estimation of Sobol indices is naturally performed by  
22 Monte-Carlo methods [6, 14, 20, 21], which permit to build estimators with statisti-  
23 cal guarantees [4, 10]. Sobol indices for multivariate, functional outputs [3, 11] or  
24 functional inputs [9] have been proposed as well.

25 Computer models employed to simulate physical systems/natural phenomena are  
26 increasingly stochastic. That is, two runs of the computer with the same input may  
27 lead to two different outputs. Examples can be found in epidemiology [1, 2, 15, 19]  
28 or ecology [22].

29 It is still unclear how sensitivity analysis should be performed when the models are  
30 stochastic. First, there is no unique definition of Sobol indices [5]. Second, it is unclear  
31 how to account for noise in the inference. Monte-Carlo sampling with repetitions is  
32 natural, but what is a good balance between the number of repetitions of the model  
33 and the number of explorations of the input space [22]? Having efficient estimators  
34 would permit to achieve the same level of precision but with less computations, an  
35 important practical issue. An approach based on meta-models has been proposed [12],  
36 but it is difficult to control the induced bias and the construction of the meta-model  
37 itself can be challenging.

38 The problem of finding an optimal Monte-Carlo design to estimate Sobol indices  
39 in stochastic computer models is addressed. Two definitions of Sobol indices are  
40 given. Their estimators, based on Monte-Carlo sampling with repetitions, are built  
41 and their asymptotic properties are established. To find an optimal tradeoff between  
42 repetitions and explorations, an error criterion that penalizes bad rankings of the  
43 inputs is considered. A bound is found and minimized under a fixed computing  
44 budget. To get estimators that asymptotically achieve the minimal bound, a two-

---

\*Submitted to the editors 2019/03/07.

<sup>†</sup>MaIAGE, INRA, Université Paris-Saclay, 78350, Jouy-en-Josas, France ([gildas.mazo@inra.fr](mailto:gildas.mazo@inra.fr),  
<http://genome.jouy.inra.fr/~gmazo/>).

45 stage procedure is proposed. The first stage estimates the combination of repetitions  
 46 and explorations that achieves the minimal bound and the second stage uses that  
 47 combination to optimize the design of the Monte-Carlo data generation. The resulting  
 48 estimators are shown to achieve the minimal bound asymptotically.

49 This paper is organized as follows. The sensitivity indices and their estimators are  
 50 defined and discussed in Section 2. The two-stage procedure to optimize the Monte-  
 51 Carlo design is given in Section 3. Some theoretical guarantees are given. Section 4  
 52 contains asymptotic results for the sensitivity indices estimators of Section 2. Section 4  
 53 and Section 3 are not related to each other and can be read independently. Numerical  
 54 experiments are provided in Section 5 to test and illustrate the theory. A discussion  
 55 closes the paper. The proofs are given in Appendix A.

## 56 2. Sensitivity analysis for stochastic models.

57 **2.1. Representation of a stochastic model.** A model is a mechanism that  
 58 takes an input  $X$  and returns an output  $Y$ . A *stochastic* model has the following  
 59 property: two runs of the model with the same input may return two different outputs.  
 60 To account for this property, we assume that there exist a function  $f$  and a hidden  
 61 random variable  $Z$  independent of  $X$  such that

$$62 \quad (2.1) \quad Y = f(X_1, \dots, X_p, Z),$$

64 where  $X_1, \dots, X_p$  are the components of  $X$ , which are assumed to be independent.  
 65 The variable  $Z$  is seen as an unobserved and uncontrollable noise variable that rep-  
 66 resents the intrinsic stochasticity of the model. Even if  $X$  were to be fixed to some  
 67 arbitrary value, say  $x$ , the output would remain a random variable, the distribution  
 68 of which would be that of  $f(x, Z)$ .

69 Note that no pairs  $(W_i, Y_i)$  with  $W_i = (X_i, Z_i)$  can be observed because  $Z$  is  
 70 not observable. Thus we have no access to the function  $f$ . The specification of  $Z$  is  
 71 unnecessary: it can be a random variable, a random vector or something else.

72 The following assumption is needed to derive some results in Section 3 and in  
 73 Section 4: there exists some function  $F$  with  $E F(X)^8 < \infty$  such that, for all  $x$  and  $z$   
 74 in the domain of definition of  $f$ ,

$$75 \quad (2.2) \quad |f(x, z)| \leq F(x).$$

76 This assumption, needed to apply various versions of central limit theorems, appears  
 77 to be mild. In particular every model with bounded outputs fulfill the condition.

78 **2.2. Sobol's decomposition.** Sobol showed that every integrable multidimen-  
 79 sional function  $h$  decomposes uniquely into a sum of lower dimensional functions [20].  
 80 If  $q$  is a natural integer and  $w = (w_1, \dots, w_q)$  lies in the euclidean space  $\mathbf{R}^q$ , then

$$81 \quad (2.3) \quad h(w) = h_0 + h_1(w_1) + \dots + h_q(w_q) \\
 82 \quad \quad \quad + h_{1,2}(w_1, w_2) + \dots + h_{q-1,q}(w_{q-1}, w_q) \\
 83 \quad \quad \quad + \dots \\
 84 \quad \quad \quad + h_{1,\dots,q}(w_1, \dots, w_q),$$

86 where  $h_0$  is a constant and

$$87 \quad \int_0^1 h_{i_1, \dots, i_k}(w_{i_1}, \dots, w_{i_k}) dw_{i_j}, \quad 1 \leq j \leq k \\
 88$$

89 for any  $k = 1, \dots, q$ . We call the lower-dimensional functions component functions.  
 90 It follows that the integral of every nonconstant component function is null and the  
 91 integral of the product of any two component functions is null as well. Integration  
 92 can be taken with respect to any product probability measure: the above properties  
 93 are not changed.

94 The decomposition (2.3) has been widely used to partition the variance of the  
 95 output of a given mathematical model. Let  $W$  be a random vector. From (2.3), it  
 96 follows that

$$\begin{aligned}
 97 \quad (2.4) \quad \text{Var } h(W) &= \text{Var } h_1(W_1) + \dots + \text{Var } h_q(W_q) \\
 98 &\quad + \text{Var } h_{1,2}(W_1, W_2) + \dots + \text{Var } h_{q-1,q}(W_{q-1}, W_q) \\
 100 &\quad + \dots + \text{Var } h_{1,\dots,q}(W_1, \dots, W_q).
 \end{aligned}$$

101 The variance of  $h(W)$  is decomposed into “individual” effects and “interaction” effects  
 102 of the components of  $W$ . The  $j$ th Sobol index, denoted by  $S_j$ , is defined as the fraction  
 103 of variance that is attributed to  $W_j$  alone in the decomposition:

$$104 \quad (2.5) \quad S_j = \frac{\text{Var } h_j(W_j)}{\text{Var } h(W)} = \frac{\text{Var } \text{E}(h(W)|W_j)}{\text{Var } h(W)}.$$

106 The last equality is true because  $h_j(W_j) = \text{E}(h(W)|W_j) - h_0$ .

107 The  $j$ th total sensitivity index [6], denoted by  $S_{Tj}$ , is the fraction of variance  
 108 attributed to  $W_j$  and its interactions with the other components of  $W$ . A convenient  
 109 formula for  $S_{Tj}$  can be found as follows. In (2.3), group all the component functions  
 110 (except the constant) that do not depend on  $w_j$  and denote the sum by  $h_{c_j}(w_{c_j})$ , where  
 111  $w_{c_j}$  stands for the vector complementary to  $w_j$ , that is, the vector whose components  
 112 are those of  $w$  with  $w_j$  removed. Likewise, group all the interactions between  $w_j$   
 113 and the other components of  $w$  and denote the sum by  $h_{j,c_j}(w_j, w_{c_j})$ . Then Sobol’s  
 114 decomposition rewrites

$$115 \quad h(w) = h_0 + h_j(w_j) + h_{c_j}(w_{c_j}) + h_{j,c_j}(w_j, w_{c_j})$$

117 and hence the  $j$ th total sensitivity index is given by

$$118 \quad (2.6) \quad S_{Tj} = 1 - \frac{\text{Var } h_{c_j}(W_{c_j})}{\text{Var } h(W)} = 1 - \frac{\text{Var } \text{E}(h(W)|W_{c_j})}{\text{Var } h(W)}.$$

120 The sensitivity indices defined above have been widely used to carry out sensitivity  
 121 analyses of computer implementations of mathematical models. These indices are  
 122 interpreted as a measure of “importance” or “influence” of the inputs for a given  
 123 model. The number  $h(w)$  is the output of the computer program and the input is  $w$ .  
 124 Monte-Carlo methods permit to estimate the sensitivity indices [20, 21, 4, 10] and to  
 125 get insight into what may be the “important” inputs of the mathematical model. The  
 126 goal in the next section is to find ways to apply Sobol’s decomposition to the special  
 127 case of stochastic models.

128 **2.3. Definition of the sensitivity indices.** To define sensitivity indices,  
 129 Sobol’s decomposition (2.3) can be exploited in two natural ways. The first approach  
 130 consists in applying Sobol’s decomposition directly to the hidden function  $f$  in (2.1).  
 131 Putting  $h = f$  and  $W = (X, Z)$  in (2.4) yields the indices in Definition 2.1. We call  
 132 them the indices of the first kind.

133 DEFINITION 2.1 (Sobol indices of the first kind). *The Sobol indices of the first*  
 134 *kind are defined as*

$$135 \quad S'_j = \frac{\text{Var E}(f(X, Z)|X_j)}{\text{Var } f(X, Z)}, \quad j = 1, \dots, p.$$

136 Interestingly, the inaccessibility of the function  $f$ , due to the lack of control over  
 137 the noise variable, does not prevent computing the indices of the first kind. Indeed,  
 138 the  $j$ th index depends on the conditional law of the output given  $X_j$  only and  $X_j$   
 139 is controllable: the output of  $f$  can be generated with  $X_j$  fixed to some value. Not  
 140 so much can be said with  $Z$ . As a result, in principle the  $j$ th total sensitivity index  
 141 can be defined as in (2.6) with  $W_{cj} = (X_1, \dots, X_{j-1}, X_{j+1}, \dots, X_p, Z)$  but it is not  
 142 estimable.

143 The second approach consists in turning the stochastic model (2.1) into a com-  
 144 pletely controllable deterministic one. To do this, one takes the conditional expecta-  
 145 tion of the output  $Y$  given the input  $X$ , given by

$$146 \quad g(X) = \text{E}(f(X, Z)|X).$$

147 The new function  $g$  is then subjected to Sobol's decomposition. In (2.3), taking  
 148  $W = X$  and  $h = g$  yields the indices in Definition 2.2.

149 DEFINITION 2.2 (Sobol indices of the second kind). *The Sobol indices of the*  
 150 *second kind are defined as*

$$151 \quad S''_j = \frac{\text{Var E}(\text{E}[f(X, Z)|X]|X_j)}{\text{Var E}[f(X, Z)|X]}, \quad j = 1, \dots, p.$$

152 Here, since  $g$  is accessible, it makes sense to define total sensitivity indices. The  
 153  $j$ th total sensitivity index is given by

$$154 \quad (2.7) \quad S''_{Tj} = 1 - \frac{\text{Var E}(g(X)|X_1, \dots, X_{j-1}, X_{j+1}, \dots, X_p)}{\text{Var } g(X)}.$$

156 The noise variable  $Z$  does not appear in the conditioning variables because the com-  
 157plementary vector  $X_{cj}$  is to be understood with respect to  $X$ , not  $(X, Z)$ .

158 However, the operation of taking the conditional expectation leads to a loss of  
 159 information. This is illustrated in Example 1.

160 EXAMPLE 1. *Let  $Y = aX_1 + cX_2\phi(Z)$ , where  $X_1, X_2, Z$  are independent standard*  
 161 *normal variables,  $a, c$  are real coefficients and  $\phi$  is a function such that  $\text{E}\phi(Z) = 0$ .*  
 162 *Then*

$$163 \quad S'_1 = \frac{a^2}{a^2 + c^2 \text{E}\phi(Z)^2}, \quad S'_2 = 0, \quad S''_1 = 1 \quad \text{and} \quad S''_2 = 0.$$

165 The information loss in Example 1 is severe: the sensitivity indices of the second  
 166 kind seem to indicate that only  $X_1$  is influential. This is because the part involving  
 167  $X_2$  has been “removed” along with the noise. In this example it may be argued that  
 168 the indices of the first kind better reflect the “importance” of the inputs.

169 In sum, the two kinds of sensitivity indices defined above seem to be complemen-  
 170 tary. Information about interaction effects will be missing with the indices of the first  
 171 kind but no first-order information is lost. The reverse is true for the indices of the

172 second kind. Another difference will be shown in Section 4: it is more difficult to  
 173 estimate the indices of the second kind than the indices of the first kind.

174 For estimation purposes, it is convenient to rewrite the indices as

$$175 \quad (2.8) \quad S'_j = \frac{\mathbb{E}\mathbb{E}[f(X, Z)|X] \mathbb{E}[f(\tilde{X}_{-j}, Z)|\tilde{X}_{-j}] - (\mathbb{E}\mathbb{E}[f(X, Z)|X])^2}{\mathbb{E}\mathbb{E}[f(X, Z)^2|X] - (\mathbb{E}\mathbb{E}[f(X, Z)|X])^2}$$

176 and

$$177 \quad (2.9) \quad S''_j = \frac{\mathbb{E}\mathbb{E}[f(X, Z)|X] \mathbb{E}[f(\tilde{X}_{-j}, Z)|\tilde{X}_{-j}] - (\mathbb{E}\mathbb{E}[f(X, Z)|X])^2}{\mathbb{E}\mathbb{E}[f(X, Z)|X]^2 - (\mathbb{E}\mathbb{E}[f(X, Z)|X])^2},$$

178 where  $\tilde{X} = (\tilde{X}_1, \dots, \tilde{X}_p)$  is an independent copy of  $X$  and

$$179 \quad \tilde{X}_{-j} = (\tilde{X}_1, \dots, \tilde{X}_{j-1}, X_j, \tilde{X}_{j+1}, \dots, \tilde{X}_p),$$

180 for  $j = 1, \dots, p$ . Note that  $S'_j$  and  $S''_j$  differ only by the lower left term. In particular,  
 181 the upper left term is the same in both formula. It is the only term that depends on  
 182  $j$ , and hence the only term that permits to discriminate between any two indices of  
 183 the same kind. For this reason, it is called the discriminator and is denoted by  $D_j$ .  
 184 Notice that  $S'_j \leq S''_j$ .

185 **2.4. Estimation of the sensitivity indices.** The sensitivity indices are es-  
 186 timated by Monte-Carlo sampling. Outputs of the stochastic model are produced  
 187 through Algorithm 2.1. The input space is explored  $n$  times and, for each explora-  
 188 tion, the computer is run  $m$  times to smooth out the noise. Thus, the total number  
 189 of calls to the computer is proportional to  $mn$ . The integer  $n$  is called the number of  
 190 explorations and the integer  $m$  is called the number of repetitions. The couple  $(n, m)$   
 191 is called the *design* of the Monte-Carlo sampling scheme.

---

**Algorithm 2.1** Generate a Monte-Carlo sample

---

```

for  $i = 1$  to  $n$  do
    draw two independent copies  $X^{(i)}, \tilde{X}^{(i)}$ 
    for  $j = 0, 1, \dots, p$  do
        for  $k = 1$  to  $m$  do
            run the computer model at  $\tilde{X}_{-j}^{(i)}$  to get an output  $Y_j^{(i,k)}$ 
        end for
    end for
end for
    
```

---

192 The data generated by the algorithm are

$$193 \quad (Y_j^{(i,k)}, \tilde{X}_{-j}^{(i)}),$$

194 for  $j = 0, 1, \dots, p$ ,  $i = 1, \dots, n$  and  $k = 1, \dots, m$ , with the convention  $\tilde{X}_{-0}^{(i)} = X^{(i)}$ .  
 195 By assumption, there are independent random elements  $(Z_j^{(i,k)})$  such that

$$196 \quad (2.10) \quad Y_j^{(i,k)} = f(\tilde{X}_{-j}^{(i)}, Z_j^{(i,k)}).$$

197 The estimators of the sensitivity indices are built by substituting empirical averages  
 198 for expectations in (2.8) and (2.9), that is,

$$199 \quad (2.11) \quad \widehat{S}'_{j;n,m} = \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{m} \sum_{k=1}^m Y_0^{(i,k)} \frac{1}{m} \sum_{k'=1}^m Y_j^{(i,k')} - \left( \frac{1}{n} \sum_{i=1}^n \frac{1}{m} \sum_{k=1}^m Y_0^{(i,k)} \right)^2}{\frac{1}{n} \sum_{i=1}^n \frac{1}{m} \sum_{k=1}^m Y_0^{(i,k)2} - \left( \frac{1}{n} \sum_{i=1}^n \frac{1}{m} \sum_{k=1}^m Y_0^{(i,k)} \right)^2}$$

200 and

$$201 \quad (2.12) \quad \widehat{S}''_{j;n,m} = \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{m} \sum_{k=1}^m Y_0^{(i,k)} \frac{1}{m} \sum_{k'=1}^m Y_j^{(i,k')} - \left( \frac{1}{n} \sum_{i=1}^n \frac{1}{m} \sum_{k=1}^m Y_0^{(i,k)} \right)^2}{\frac{1}{n} \sum_{i=1}^n \left( \frac{1}{m} \sum_{k=1}^m Y_0^{(i,k)} \right)^2 - \left( \frac{1}{n} \sum_{i=1}^n \frac{1}{m} \sum_{k=1}^m Y_0^{(i,k)} \right)^2}.$$

202 The estimation of total sensitivity indices of the second kind is not considered in this  
 203 paper. However, the formula (2.7) suggests that estimators can be built similarly.

204 To our knowledge (personal communication), when faced with stochastic com-  
 205 puter models, practitioners tend to use softwares for deterministic sensitivity analysis  
 206 in which an average over repetitions is given to the program as a substitute for the  
 207 value of the output. Thus, the second estimator is used in practice, albeit implicitly.  
 208 The first estimator, to the best of our knowledge, was not formally defined. The  
 209 second estimator appeared in [7, 8], where it was studied only in the case  $m = n$  (to  
 210 the best of our understanding).

211 In (2.11) and (2.12), if  $m = 1$  and the function  $f$  does in fact not depend on  
 212  $Z$ , then the estimators reduce to Sobol estimators [20, 21] for deterministic models.  
 213 These are sometimes called pick-freeze estimators [4].

214 **3. Choosing between Monte-Carlo designs.** The estimators in Section 2  
 215 depend on the design  $(n, m)$  of the Monte-Carlo sampling scheme. To estimate the  
 216 sensitivity indices, the stochastic model has to be called  $(p + 1)mn$  times.

217 It is reasonable to think of a sensitivity analysis as done the following way. The  
 218 total number of calls is set to a limit, say  $T$ . Then  $n$  and  $m$  are chosen so that  
 219  $T = (p + 1)mn$ . For instance, suppose that one cannot afford more than 150 calls to  
 220 a model with two inputs. Then  $T = 150$ ,  $p = 2$  and one can choose either one of the  
 221 columns in the following table

$$222 \quad \begin{array}{cccccc} n & 50 & 25 & 10 & 5 & 2 & 1 \\ m & 1 & 2 & 5 & 10 & 25 & 50. \end{array}$$

223 Denote by  $\text{div}_p(T)$  the set of all divisors of  $T/(p + 1)$  between 1 and  $T/(p + 1)$ .  
 224 In the example above,  $\text{div}_2(150) = \{1, 2, 5, 10, 25, 50\}$ . There are as many designs as  
 225 there are elements in the set  $\text{div}_p(T)$ . Each one of those elements corresponds to a  
 226 possible combination for  $n$  and  $m$  which Algorithm 2.1 can be run with. The resulting  
 227 estimators require the same number of calls but do not perform equally well. The  
 228 goal of this section is to find the “best” way to estimate the sensitivity indices.

229 **3.1. Introducing the miss-ranking error and its bound.** To compare the  
 230 estimators, a measure of performance has to be defined. We shall consider the miss-  
 231 ranking error (MRE), defined by

$$232 \quad \text{MRE} = \mathbb{E} \sum_{j=1}^p |\widehat{R}_{j;n,m} - R_j|,$$

233 where  $R_j$  is the rank of  $D_j$  among  $D_1, \dots, D_p$ , that is,  $R_j = \sum_{i=1}^p \mathbf{1}(D_i \leq D_j)$ ,  
 234 and  $\widehat{R}_{j;n,m}$  is an estimator of  $R_j$ . Recall that  $D_1, \dots, D_p$  are the upper-left terms  
 235 in (2.8) and (2.9). They determine the ranks of the sensitivity indices. Recall that  
 236 the ranks of the sensitivity indices of the first kind coincide with the ranks of the  
 237 sensitivity indices of the second kind. Thus, the MRE permits to find a unique  
 238 solution for both kinds of sensitivity indices. The MRE is small when one succeeds  
 239 in ranking the inputs from the most to the least important, a task which is called  
 240 “factors prioritization” in [18, p. 52].

241 The MRE has a bound with interesting mathematical properties. Denote by  
 242  $\text{MRE}(T, m)$  the MRE based on  $T$  number of calls and  $m$  repetitions, so that the  
 243 number of explorations is  $T/(p+1)/m$ . Remember the notation of Section 2: denote  
 244  $(X^{(1)}, \widetilde{X}^{(1)}) = \mathbf{X}$ ,  $f(X^{(1)}, Z_0^{(1,1)}) = Y_0$  and  $f(\widetilde{X}_{-j}^{(1)}, Z_j^{(1,1)}) = Y_j$ .

245 **PROPOSITION 3.1.** *Let  $\widehat{D}_{j;n,m}$ ,  $j = 1, \dots, p$ , be the upper-left term in (2.11)*  
 246 *(or (2.12)) and put  $\widehat{R}_{j;n,m} = \sum_{i=1}^p \mathbf{1}(\widehat{D}_{i;n,m} \leq \widehat{D}_{j;n,m})$ . Then*

$$\begin{aligned}
 247 \quad \text{MRE}(T, m) &\leq \frac{L}{nm} \left( m \sum_{j=1}^p \text{Var}(\mathbb{E}[Y_0 Y_j | \mathbf{X}]) \right. \\
 &\quad + \sum_{j=1}^p \mathbb{E}(\text{Var}[Y_0 Y_j | \mathbf{X}] - \text{Var}[Y_0 | \mathbf{X}] \text{Var}[Y_j | \mathbf{X}]) \\
 248 &\quad \left. + \frac{1}{m} \sum_{j=1}^p \mathbb{E}(\text{Var}[Y_0 | \mathbf{X}] \text{Var}[Y_j | \mathbf{X}]) \right), \\
 249 & \\
 250 &
 \end{aligned}$$

251 where

$$252 \quad L = \frac{4(p-1)}{\min_{j < j'} (|D_j - D_{j'}|^2)}.$$

254 The constant  $L$  tells us that the bound is smaller when the indices are well  
 255 separated. The bound goes to zero when the number of explorations goes to infinity.  
 256 This is true even if the number of repetitions is fixed. Most interestingly, the bound  
 257 separates  $T$  and  $m$ :

$$258 \quad (3.1) \quad \text{MRE}(T, m) \leq \frac{1}{T} v(m), \quad m \in \text{div}_p(T),$$

260 where the function  $v$  is implicitly defined through Proposition 3.1. Denote by  $m_T^\dagger$   
 261 the element  $m$  in  $\text{div}_p(T)$  that minimizes  $v(m)$ . Taking  $m = m_T^\dagger$  in (3.1), we get the  
 262 bound

$$263 \quad \text{MRE}(T, m_T^\dagger) \leq \frac{v(m_T^\dagger)}{T} \leq \frac{v(m)}{T}, \quad \text{for all } m \in \text{div}_p(T).$$

264 Thus the best guarantee coincides with  $m = m_T^\dagger$  and  $n = T/(p+1)/m_T^\dagger$  in Algo-  
 265 rithm 2.1. However  $m_T^\dagger$  is unknown.

266 *Remark 3.2.* The choice of  $T$ , through the specification of  $\text{div}_p(T)$ , will influence  
 267 the quality of the bound. It is clear that choosing  $T/(p+1)$  a prime number may  
 268 not be a good idea because  $v(m_T^\dagger)$  will be either one of  $v(1)$  or  $v(T/(p+1))$ . On the  
 269 opposite, choosing  $T/(p+1)$  a factorial number ensures many more choices (in fact,  
 270 all).



271 **3.2. A two-stage procedure to estimate the sensitivity indices.** The results in Section 3.1 suggest a two-stage procedure to estimate the sensitivity indices.  
 272 The procedure is given in Algorithm 3.1. The computational budget is split into two  
 273 parts  $K$  and  $T - K$ . The first  $K$  calls to the model are used to estimate  $m_{T-K}^\dagger$ . The  
 274 last  $T - K$  calls to the model are used to estimate the sensitivity indices.  
 275

---

**Algorithm 3.1** Estimate the sensitivity indices by a two-stage procedure

---

**Stage 1.** Choose an integer  $K$  such that  $K/(p+1)$  and  $(T-K)/(p+1)$  are integers also. Choose integers  $m_0$  and  $n_0$  such that  $K = m_0 n_0 (p+1)$ . Run Algorithm 2.1 with  $m = m_0$  and  $n = n_0$ . Estimate  $m_{T-K}^\dagger$  by an estimator  $\widehat{m}_{T-K}^\dagger$  in  $\text{div}_p(T-K)$ .

**Stage 2.** Run Algorithm 2.1 with  $m = \widehat{m}_{T-K}^\dagger$  and

$$n = \frac{T-K}{(p+1)\widehat{m}_{T-K}^\dagger}.$$

Compute the sensitivity indices estimators (2.11) and (2.12).

---

276 The estimator of  $m_{T-K}^\dagger$  is built as follows. Let  $m^*$  be the minimizer of  $v$  seen as  
 277 a function on the positive reals. Since  $v$  is convex, the minimizer is unique. It follows  
 278 from (3.1) and Proposition 3.1 that

$$279 \quad (3.2) \quad m^* := \sqrt{\frac{\sum_{j=1}^p \text{E Var}[Y_0|\mathbf{X}] \text{Var}[Y_j|\mathbf{X}]}{\sum_{j=1}^p \text{Var E}[Y_0 Y_j|\mathbf{X}]}} = \sqrt{\frac{\sum_{j=1}^p \zeta_{3,j}}{\sum_{j=1}^p \zeta_{1,j}}},$$

281 where  $\zeta_{3,j} = \text{E Var}[Y_0|\mathbf{X}] \text{Var}[Y_j|\mathbf{X}]$  and  $\zeta_{1,j} = \text{Var E}[Y_0 Y_j|\mathbf{X}]$ ,  $j = 1, \dots, p$ .

282 Let  $\varphi_T : (0, \infty) \rightarrow \text{div}_p(T)$ , be the function defined by  $\varphi_T(x) = 1$  if  $0 < x < 1$ ,  
 283  $\varphi_T(x) = T/(p+1)$  if  $x > T/(p+1)$ , and

$$284 \quad \varphi_T(x) = \begin{cases} \lfloor x \rfloor_T & \text{if } \sqrt{\lfloor x \rfloor_T \lceil x \rceil_T} > x \geq 1 \\ \lceil x \rceil_T & \text{if } \sqrt{\lfloor x \rfloor_T \lceil x \rceil_T} \leq x \leq \frac{T}{p+1} \end{cases}$$

285 where

$$286 \quad \lfloor x^* \rfloor_T = \max\{m \in \text{div}_p(T), m \leq x\}, \quad \lceil x^* \rceil_T = \min\{m \in \text{div}_p(T), m \geq x\}.$$

287 The function  $\varphi_T$  is piecewise constant with discontinuity points at  $\sqrt{i \cdot j}$ , where  $i$  and  
 288  $j$  are two consecutive elements of  $\text{div}_p(T)$ .

291 **PROPOSITION 3.3.** *If  $m^* > 0$  then  $m_{T-K}^\dagger = \varphi_{T-K}(m^*)$ . If, moreover,*  
 292  *$\lfloor m^* \rfloor_{T-K} \lceil m^* \rceil_{T-K}$  is not equal to  $m^{*2}$  then the minimizer of  $v(m)$ ,  $m \in \text{div}_p(T-K)$ ,*  
 293 *is unique.*

294 Proposition 3.3 suggests that  $m_{T-K}^\dagger$  can be estimated by applying the function  
 295  $\varphi_{T-K}$  to an estimate of  $m^*$ . Remember that  $K = m_0 n_0 (p+1)$  and put

$$296 \quad (3.3) \quad \widehat{m}_K^* := \sqrt{\frac{\sum_{j=1}^p \widehat{\zeta}_{3,j}}{\sum_{j=1}^p \widehat{\zeta}_{1,j}}},$$

297 where

$$\begin{aligned}
298 \quad & \hat{\zeta}_{3,j} = \\
299 \quad (3.4) \quad & \frac{1}{n_0} \sum_{i=1}^n \frac{1}{m_0} \sum_{k_1=1}^{m_0} f(X^{(i)}, Z_0^{(i,k_1)})^2 \frac{1}{m_0} \sum_{k_2=1}^{m_0} f(\tilde{X}_{-j}^{(i)}, Z_j^{(i,k_2)})^2 \\
300 \quad (3.5) \quad & + \frac{1}{n_0} \sum_{i=1}^n \left( \frac{1}{m_0} \sum_{k_1=1}^{m_0} f(X^{(i)}, Z_0^{(i,k_1)}) \right)^2 \left( \frac{1}{m_0} \sum_{k_2=1}^{m_0} f(\tilde{X}_{-j}^{(i)}, Z_j^{(i,k_2)}) \right)^2 \\
301 \quad (3.6) \quad & - \frac{1}{n_0} \sum_{i=1}^n \left( \frac{1}{m_0} \sum_{k_1=1}^{m_0} f(X^{(i)}, Z_0^{(i,k_1)}) \right)^2 \frac{1}{m_0} \sum_{k_2=1}^{m_0} f(\tilde{X}_{-j}^{(i)}, Z_j^{(i,k_2)})^2 \\
302 \quad (3.7) \quad & - \frac{1}{n_0} \sum_{i=1}^n \frac{1}{m_0} \sum_{k_1=1}^{m_0} f(X^{(i)}, Z_0^{(i,k_1)})^2 \left( \frac{1}{m_0} \sum_{k_2=1}^{m_0} f(\tilde{X}_{-j}^{(i)}, Z_j^{(i,k_2)}) \right)^2, \\
303
\end{aligned}$$

304 and

$$\begin{aligned}
305 \quad & \hat{\zeta}_{1,j} = \\
306 \quad (3.8) \quad & \frac{1}{n_0} \sum_{i=1}^n \left( \frac{1}{m_0} \sum_{k=1}^{m_0} f(X^{(i)}, Z_0^{(i,k)}) f(\tilde{X}_{-j}^{(i)}, Z_j^{(i,k)}) \right)^2 \\
307 \quad (3.9) \quad & - \left( \frac{1}{n_0} \sum_{i=1}^n \frac{1}{m_0} \sum_{k=1}^{m_0} f(X^{(i)}, Z_0^{(i,k)}) f(\tilde{X}_{-j}^{(i)}, Z_j^{(i,k)}) \right)^2. \\
308
\end{aligned}$$

309 Notice that  $\hat{\zeta}_{1,j} \geq 0$  and  $\hat{\zeta}_{3,j} \geq 0$  so that  $\hat{m}_K^* \geq 0$ . If  $m_0 = 1$  then  $\hat{\zeta}_{3,j} = 0$  and hence  
310  $\hat{m}_K^* = 0$ .

311 The estimator  $\hat{m}_K^*$  is consistent and asymptotically normal on some conditions  
312 on the rates of  $n_0$  and  $m_0$ .

313 **THEOREM 3.4.** *Assume (2.2) holds. Let  $n_0 \rightarrow \infty$ . If  $m_0$  is fixed then*

$$314 \quad \sqrt{n_0} \left( \hat{m}_K^* - \left[ m^* + \frac{C}{m_0} + \epsilon_{m_0} \right] \right) \xrightarrow{d} N(0, \sigma_{m_0}^2),$$

315 for some constant  $C$ , real  $\epsilon_{m_0}$  depending on  $m_0$  and variance  $\sigma_{m_0}^2$  depending on  
316  $m_0$ . If  $m_0 \rightarrow \infty$  then the above display with  $\epsilon_{m_0} = o(1/m_0)$  and  $\sigma_{m_0}$  replaced by  
317  $\lim_{m_0 \rightarrow \infty} \sigma_{m_0}$  is true.

318 Theorem 3.4 shows that  $\hat{m}_K^*$  is asymptotically biased. The bias is polynomial in  
319  $1/m_0$ . Corollary 3.5 shows that letting  $m_0 \rightarrow \infty$  suffices to get the consistency of  $\hat{m}_K^*$   
320 but to get a central limit theorem centered around  $m^*$ , it is furthermore needed that  
321  $\sqrt{n_0}/m_0 \rightarrow 0$ .

322 **COROLLARY 3.5.** *Assume (2.2) holds. Let  $n_0 \rightarrow \infty$  and  $m_0 \rightarrow \infty$ . Then  $\hat{m}_K^* \xrightarrow{P}$   
323  $m^*$ . If, moreover,  $\sqrt{n_0}/m_0 \rightarrow 0$ , then*

$$324 \quad \sqrt{n_0}(\hat{m}_K^* - m^*) \xrightarrow{d} N(0, \lim_{m_0 \rightarrow \infty} \sigma_{m_0}^2).$$

325 To estimate  $m_{T-K}^\dagger$ , put  $\hat{m}_{T-K}^\dagger = \varphi_{T-K}(\hat{m}_K^*)$ . Proposition 3.6 states that  $\hat{m}_{T-K}^\dagger$   
326 and  $m_{T-K}^\dagger$  are equal with probability going to one.

327 PROPOSITION 3.6. Assume (2.2) holds. Let  $n_0 \rightarrow \infty$  and  $m_0 \rightarrow \infty$ . Then

$$328 \quad P\left(\widehat{m}_{T-K}^\dagger = m_{T-K}^\dagger\right) \rightarrow 1.$$

330 All the details of Algorithm 3.1 have been given.

331 **3.3. Performance.** To get some insight into the performance of the procedure  
 332 given in Algorithm 3.1, we look at the performance of the sensitivity indices estimators  
 333 produced in Stage 2. Since they are built with  $T - K$  calls to the model with  $\widehat{m}_{T-K}^\dagger$   
 334 repetitions, they satisfy

$$335 \quad (3.10) \quad \text{MRE}(T - K, \widehat{m}_{T-K}^\dagger) \leq \frac{1}{T - K} v(\widehat{m}_{T-K}^\dagger),$$

337 where the left-hand side is the conditional expectation of the MRE, given the outputs  
 338 produced in Stage 1. The estimator  $\widehat{m}_{T-K}^\dagger$  is computed with  $K$  calls only.

339 It is difficult to compare the guarantee above with that which got by choosing  
 340 an arbitrary number of repetitions, say  $m$ . In the later case  $K = 0$  and hence the  
 341 guarantee is (3.1). The denominator in (3.10) is smaller but we expect that the  
 342 numerator  $v(\widehat{m}_{T-K}^\dagger)$  will be less than  $v(m)$  for many values of  $m$ . Indeed, the numer-  
 343 ator should be close to  $v(m_{T-K}^\dagger)$ . If  $T - K$  is well chosen then  $v(m_{T-K}^\dagger)$  and  $v(m_T^\dagger)$   
 344 should be close and since  $v(m) \leq v(m_T^\dagger)$  for all  $m$  in  $\text{div}_p(T)$ , the numerator  $v(\widehat{m}_{T-K}^\dagger)$   
 345 should be an approximate minimizer. For instance if  $K$  and  $T$  are large enough and  
 346  $\text{div}_p(T - K) = \{1, 2, \dots, (T - K)/(p + 1)\}$  and  $\text{div}_p(T) = \{1, 2, \dots, T/(p + 1)\}$  hold  
 347 then  $v(m_{T-K}^\dagger)$  and  $v(m_T^\dagger)$  are equal. Note that the numerator and the denominator  
 348 in (3.10) cannot be good at the same time and  $K$  determines the balance.

349 THEOREM 3.7. Assume that the conditions of Proposition 3.6 are fulfilled. Sup-  
 350 pose furthermore that  $K \rightarrow \infty$  such that  $K/T \rightarrow 0$ . Then

$$351 \quad \frac{1}{T - K} v(\widehat{m}_{T-K}^\dagger) = \frac{1}{T} v(m_{T-K}^\dagger) (1 + o_P(1)).$$

353 The bound in Theorem 3.7 is the best possible guarantee inflated by a factor not  
 354 much larger than one. This result is valid if  $K$  is large but not too large with regard  
 355 to  $T$ . In particular, if  $\text{div}_p(T) \cap \text{div}_p(T - K) = \text{div}_p(T - K)$  then for every fixed  
 356  $m \neq m_T^\dagger$ , it holds that  $P(T^{-1} v(m_{T-K}^\dagger) (1 + o_P(1)) \leq T^{-1} v(m)) \rightarrow 1$ . In other words  
 357 it is always better, in terms of obtainable guarantees, to use the procedure rather than  
 358 to choose the number of repetitions arbitrarily, except for the lucky case  $m = m_T^\dagger$ .

359 **4. Asymptotic normality of the sensitivity indices estimators.** The sen-  
 360 sitivity indices estimators of Section 2.4 depend on both  $m$  and  $n$ . It is clear that  
 361  $n$  should go to infinity to get central limit theorems. It may be less clear, however,  
 362 whether or not  $m$  should go to infinity as well. The answer depends on the kind of  
 363 the sensitivity index we are looking at.

364 Two frameworks are considered:

- 365 •  $n \rightarrow \infty$  and  $m$  is fixed;
- 366 •  $n \rightarrow \infty$  and  $m \rightarrow \infty$ .

367 In the second framework  $m_n$  is a sequence indexed by  $n$  that goes to infinity as  $n$  goes  
 368 to infinity. Denote by  $\mathbf{S}'$  (resp.  $\mathbf{S}''$ ) the (column) vector with coordinates  $S'_j$  (resp.  
 369  $S''_j$ ),  $j = 1, \dots, p$ , and denote by  $\widehat{\mathbf{S}}'_{n,m}$  (resp.  $\widehat{\mathbf{S}}''_{n,m}$ ) the vector with coordinates  $\widehat{S}'_{j;n,m}$   
 370 given in (2.11) (resp.  $\widehat{S}''_{j;n,m}$  given in (2.12)).

371 THEOREM 4.1. Assume (2.2) holds. Let  $n \rightarrow \infty$ . If  $m$  is fixed then

$$372 \quad \sqrt{n} \left( \widehat{\mathbf{S}}''_{n,m} - \mathbf{S}'' \left[ 1 - \frac{\widehat{\mathbf{S}}'_{n,m} - \mathbf{S}'}{\frac{\text{E Var}[f(X,Z)|X]}{\text{E Var}[f(X,Z)|X] + m \text{ Var E}[f(X,Z)|X]}} \right] \right) \xrightarrow{d} N(0, \Xi_m),$$

373 for some nonnegative matrix  $\Xi_m$  of size  $2p \times 2p$ . If  $m \rightarrow \infty$  then, elementwise,  
 374  $\lim_{m \rightarrow \infty} \Xi_m$  exists and the above display with  $\Xi_m$  replaced by  $\lim_{m \rightarrow \infty} \Xi_m$  is true.

375 Theorem 4.1 predicts the behavior of the joint vector  $(\widehat{\mathbf{S}}'_{n,m}, \widehat{\mathbf{S}}''_{n,m})$ . However the  
 376 behaviors of  $\widehat{\mathbf{S}}'_{n,m}$  and  $\widehat{\mathbf{S}}''_{n,m}$  are different. The estimator  $\widehat{\mathbf{S}}'_{n,m}$  is asymptotically normal  
 377 around  $\mathbf{S}'$ , even if  $m$  is kept fixed. The estimator  $\widehat{\mathbf{S}}''_{n,m}$  is also asymptotically normal,  
 378 but not around  $\mathbf{S}''$ .

379 The estimator  $\widehat{\mathbf{S}}''_{n,m}$  under-estimates  $\mathbf{S}''$ . The bias, given by

$$380 \quad \mathbf{S}'' \frac{\text{E Var}[f(X,Z)|X]}{\text{E Var}[f(X,Z)|X] + m \text{ Var E}[f(X,Z)|X]},$$

382 is null whenever  $f$  actually does not depend on  $Z$ , and large whenever the computer  
 383 model is highly stochastic. As Theorem 4.1 shows, the bias is still present even if  $m$   
 384 goes to infinity. Corollary 4.2 shows that  $m$  must go to infinity fast enough to avoid  
 385 the estimator to be tightly concentrated around the wrong target.

386 COROLLARY 4.2. Assume (2.2) holds. Let  $n \rightarrow \infty$ . If  $m \rightarrow \infty$  such that  $\sqrt{n}/m \rightarrow$   
 387 0 then

$$388 \quad \sqrt{n} \left( \widehat{\mathbf{S}}''_{n,m} - \mathbf{S}'' \right) \xrightarrow{d} N(0, \Xi_{22}),$$

389 where  $\Xi_{22}$  is the lower-right block of the matrix  $\lim_{m \rightarrow \infty} \Xi_m$  given in Theorem 4.1.

390 The difference between  $\widehat{\mathbf{S}}'_{n,m}$  and  $\widehat{\mathbf{S}}''_{n,m}$  is due to the difference between the lower-  
 391 left terms in (2.11) and (2.12). While the lower-left term in (2.11) is unbiased for all  $n$   
 392 and  $m$ , the lower-left term in (2.12) has a bias depending on  $m$  which propagates to the  
 393 estimator of the sensitivity indices. (The calculations are carried out in Appendix C.)

394 From a statistical perspective, it is more difficult to estimate the sensitivity indices  
 395 of the second kind than to estimate the sensitivity indices of the first kind. To estimate  
 396 the former, one needs to repeat the model many times. To estimate the later, this is  
 397 not necessary.

398 **5. Numerical tests.** Section 5.1 illustrates how the MRE responds to a change  
 399 in the Monte-Carlo design. In Section 5.1 the total budget  $T$  is kept fixed. Section 5.2  
 400 illustrates how the sensitivity indices estimators behave asymptotically. In Section 5.2  
 401 the total budget  $T$  increases.

402 **5.1. Comparison of Monte-Carlo designs.** The effect of the number of rep-  
 403 etitions on the sensitivity indices estimators and the effect of the calibration in the  
 404 two-stage procedure are examined in two kinds of experiments: the “direct” experi-  
 405 ments and the “calibration” experiments.

406 In the direct experiments, the sensitivity indices are estimated directly with the  
 407 given number of repetitions. Increasing numbers of repetitions  $m$  are tested. (Since  
 408 the budget is fixed, this goes with decreasing numbers of explorations.) For each  $m$ ,  
 409 the mean squared errors (MSEs), given by  $\text{E} \sum_{j=1}^p (\widehat{S}'_{j;n,m} - S'_j)^2$  and  $\text{E} \sum_{j=1}^p (\widehat{S}''_{j;n,m} -$   
 410  $S''_j)^2$ , are estimated with replications. They are also split into the sum of the squared

411 biases and the sum of the variances to get further insight about the behavior of the  
 412 estimators. The MREs are estimated as well. A normalized version is considered:  
 413 it is the MRE divided by the number of variables. For models with two inputs, the  
 414 normalized MRE is interpreted directly as the probability that the two inputs are  
 415 ranked incorrectly.

416 In the calibration experiments, the sensitivity indices are estimated with the two-  
 417 stage procedure, the results of which depend on the calibration parameters  $K$  and  
 418  $m_0$ . Various calibration parameters are tested to see their effect on the MRE. The  
 419 budgets for the direct experiments and the calibration experiments are the same so  
 420 that the numbers can be compared. In particular, the direct experiments correspond  
 421 to the case  $K = 0$  in the calibration experiments.

422 A linear model of the form  $Y = X_1 + \beta X_2 + \sigma Z$ , where  $X_1, X_2, Z$ , are standard  
 423 normal random variables and  $\beta, \sigma$  are real coefficients, has been considered because  
 424 the sensitivity indices are explicit and hence the performance of the estimators can  
 425 be evaluated easily. The quantity  $m^*$  is explicit: the formula is given in Appendix D.

426 **5.1.1. High noise context.** The coefficients are  $\beta = 1.2$  and  $\sigma = 4$ . The  
 427 sensitivity indices are  $S'_1 = 0.05$ ,  $S'_2 = 0.08$ ,  $S''_1 = 0.41$  and  $S''_2 = 0.59$ . The real  
 428  $m^*$  is about 5.8. The total budget is  $T = 3 \times 500 = 1500$  and hence  $\text{div}_2(1500) =$   
 429  $\{1, 2, 4, 5, 10, 20, 25, 50, 100, 125, 250, 500\}$ . The integer  $m_{1500}^\dagger$  is equal to  $\varphi_{1500}(m^*) =$   
 430 5. Since the budget is kept fixed, the numbers of explorations are, respectively,  
 431 500, 250, 125, 100, 50, 25, 20, 10, 5, 4, 2, 1. The number of replications is 1500.

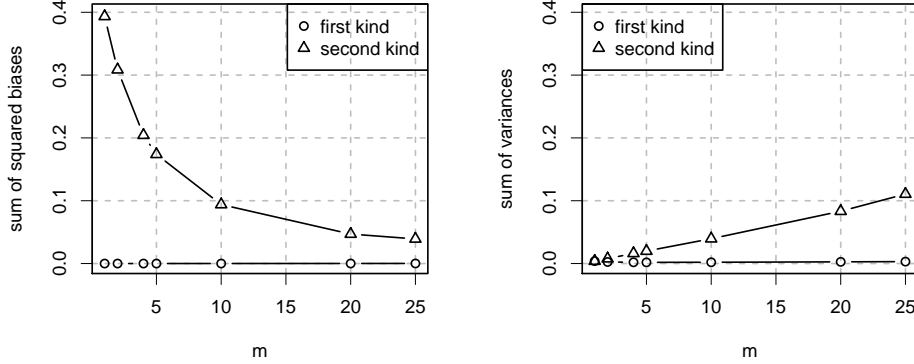
432 The results of the direct experiment are given in Figure 1 for  $m = 1, 2, 4, 5, 10,$   
 433 20, 25. The MSE of first kind does not vary with the number of repetitions and is  
 434 much lower than the MSE of second kind, see (c). The estimators of the second kind  
 435 are highly biased for small numbers of repetitions (a) and they have a higher variance  
 436 for larger numbers of repetitions (b). The fact that the bias is high for small numbers  
 437 of repetitions agrees with the theory, according to which the bias should vanish as  $m$   
 438 goes to infinity. Overall, the sensitivity indices of the second kind seem to be much  
 439 harder to estimate than the indices of the first kind, the estimators of which have a  
 440 negligible bias and a very small variance whatever the number of repetitions.

441 According to Figure 1(c), the normalized MRE curve has a banana shape with a  
 442 minimum of about slightly less than 30% reached around  $m \in \{5, 10\}$  and endpoints  
 443 with a value of about 35%. A value of 30% means that the probability of ranking  
 444 the inputs correctly is about 70%. The region of observed optimal performance  $m \in$   
 445  $\{5, 10\}$  coincides with  $m_{1500}^\dagger = 5$ , the point at which the bound is minimal.

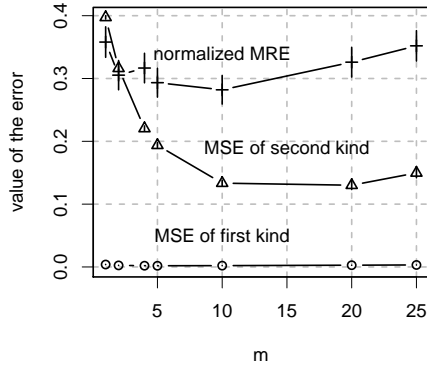
446 The results of the calibration experiment is given in Table 1 for the normalized  
 447 MRE. The lowest MREs are reached at the bottom right of the table, with values  
 448 corresponding to  $2 \leq m \leq 10$  in Figure 1 (c). Optimal performance is reached with  
 449 very few explorations in the first stage of the two-stage procedure. In this case, the  
 450 estimator  $\hat{m}_K^*$  has a small bias but a high variance. It seems to be better than an  
 451 estimator with a small variance but a large bias. This might be explained by the low  
 452 curvature of the MRE curve.

453 **5.1.2. Low noise context.** The coefficients are  $\beta = 1.2$  and  $\sigma = 0.9$ . The  
 454 sensitivity indices are  $S'_1 = 0.31$ ,  $S'_2 = 0.44$ ,  $S''_1 = 0.41$  and  $S''_2 = 0.59$ . The real  
 455  $m^*$  is about 0.30 and hence the integer  $m_{1500}^\dagger$  is equal to 1. As expected, these  
 456 numbers are smaller than the ones found in the high noise context. The total budget  
 457 is  $T = 3 \times 500 = 1500$ . The number of replications is 500.

458 The results for the direct experiment are given in Figure 2. The MSE of first  
 459 kind increases with the number of repetitions, see (c): this is due to the increase



(a) Squared bias (b) Variance



(c) Errors

Fig. 1: Sum of squared biases (a), sum of variances (b) and errors (c) of the sensitivity indices estimators for the linear model in the high noise setting. Confidence intervals of level 95% are added in (c).

$K/3$	$m_0$				$n_0$			
	2	5	10	20	20	10	5	2
400	0.43	0.42	0.42	-	-	0.42	0.39	0.40
200	0.38	0.39	0.37	-	-	0.35	0.35	0.34
100	0.36	0.37	-	-	-	-	0.32	0.30
50	0.39	0.33	-	-	-	-	0.33	0.31

Table 1: Normalized MRE in the linear model with high noise for various calibrations:  $K/(p + 1) = 50, 100, 200, 400$  and  $m_0 = 2, 5, 10, 20, \dots$ . For instance, for  $K/(p + 1) = 200 = m_0 n_0$ , the normalized MRE is available for  $m_0 = 2, 5, 10, 20, 40, 100$ .

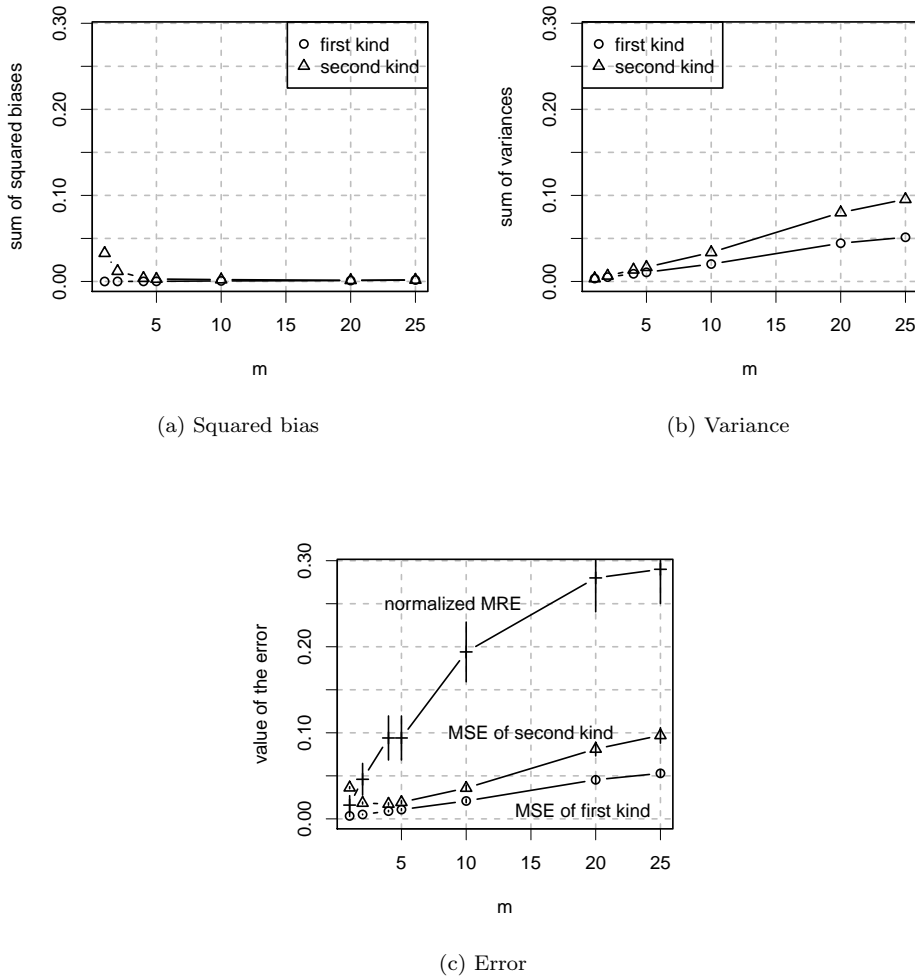


Fig. 2: Sum of squared biases (a), sum of variances (b) and errors (c) of the sensitivity indices estimators for the linear model in the low noise context. Confidence intervals of level 95% are added in (c).

460 of the variance (b), while the bias is negligible (a). As in the high noise context,  
 461 the estimators of the second kind have a decreasing bias and an increasing variance,  
 462 although the decrease of the bias is of much less magnitude. This agrees with the  
 463 theory, where we have seen that, for the sensitivity indices of the second kind, the  
 464 biases of the estimators are small when the noise of the model is low.

465 In Figure 2 (c), the normalized MRE varies a lot. It increases from about 2% at  
 466  $m = 1$  to 30% at  $m = 25$ . Thus, unlike in the high noise setting, choosing a good  
 467 number of repetitions is important. The best performance is achieved at  $m = 1$ , which  
 468 coincides with the minimizer  $m_{1500}^{\dagger} = 1$  of the bound.

469 The results of the calibration experiment for the normalized MRE is given in

$K/3$	$m_0$				$n_0$			
	2	5	10	20	20	10	5	2
400	0.18	0.15	0.17	-	-	0.16	0.18	0.20
200	0.05	0.04	0.04	-	-	0.06	0.05	0.07
100	0.02	0.04	-	-	-	-	0.04	0.04
50	0.03	0.02	-	-	-	-	0.02	0.04

Table 2: Normalized MRE in the linear model with low noise for various calibrations:  $K/(p + 1) = 50, 100, 200, 400$  and  $m_0 = 2, 5, 10, 20, \dots$ . For instance, for  $K/(p + 1) = 200 = m_0 n_0$ , the normalized MRE is available for  $m_0 = 2, 5, 10, 20, 40, 100$ .

470 Table 2. The best performance is reached at the bottom left of the table with numbers  
 471 that correspond to the optimal performance in Figure 2 (c). Moreover, notice that a  
 472 large spectrum of calibration parameters ( $K, m_0$ ) yield low errors.

473 **5.2. Asymptotic behavior of the sensitivity indices estimators.** To illus-  
 474 trate the asymptotic behavior of the sensitivity indices estimators, Sobol’s g-function,  
 475 a benchmark in sensitivity analysis [16, 13], is considered. Sobol’s g-function is given  
 476 by

$$477 \quad g(U_1, \dots, U_{p+1}) = \prod_{j=1}^{p+1} \frac{|4U_j - 2| + a_j}{1 + a_j},$$

479 where the  $a_j$  are nonnegative and the  $U_j$  are independent standard uniform random  
 480 variables. The less  $a_j$  the more  $U_j$  is important. Elementary calculations show that  
 481 the first-order Sobol index (2.5), associated with  $U_j$ , is given by

$$482 \quad S_j^{(a_1, \dots, a_{p+1})} = \frac{1}{3(1 + a_j)^2} \left( -1 + \prod_{j=1}^{p+1} \frac{(4/3 + a_j^2 + 2a_j)}{(1 + a_j)^2} \right)^{-1}.$$

484 To build a stochastic model out of Sobol’s g-function, we let one of the  $U_j$  play  
 485 the role of  $Z$ . For instance if  $U_i, 1 \leq i \leq p + 1$ , were to play this role, then the  
 486 stochastic model would be

$$487 \quad (5.1) \quad Y = f(X_1, \dots, X_p, Z) = g(X_1, \dots, X_{i-1}, Z, X_i, \dots, X_p).$$

489 Of course  $Y$  and  $f$  above depend on  $i$ . In the rest of this section we choose arbitrarily  
 490  $i = 2$  and  $p = 4$ .

491 The Sobol indices of the first and of the second kind (in the sense of Definition 2.1  
 492 and 2.2) are then easily seen to be

$$493 \quad S'_j = \begin{cases} S_j^{(a_1, \dots, a_{p+1})} & \text{if } 1 \leq j \leq i - 1 \\ S_{j+1}^{(a_1, \dots, a_{p+1})} & \text{if } i \leq j \leq p \end{cases}$$

495 and  $S''_j = S_j^{(b_{i1}, \dots, b_{ip})}$ , where

$$496 \quad b_{ij} = \begin{cases} a_j & \text{if } 1 \leq j \leq i - 1, \\ a_{j+1} & \text{if } i \leq j \leq p. \end{cases}$$



498 For each kind of Sobol index, we produced 500 estimates of the  $p$  Sobol indices  
 499 and computed the values of the mean squared error (MSE) by averaging over the  
 500 500 replications and summing over the  $p$  indices. We tested  $n = 100, 500, 2500$  and  
 501  $m = 1, 10, 100$ .

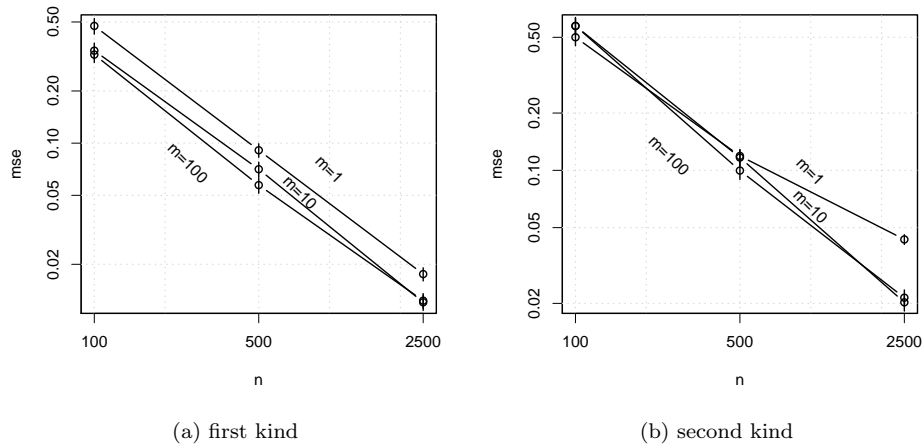


Fig. 3: MSEs for the Sobol index estimators of the first and second kind (logarithmic scale).

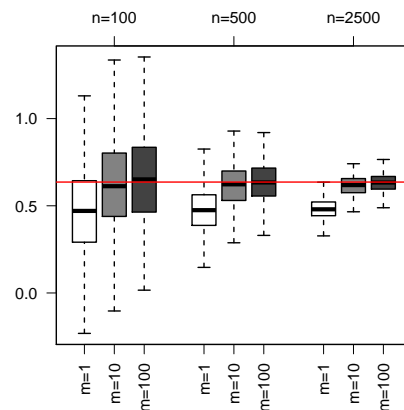


Fig. 4: Boxplots of the estimates for the Sobol index of the second kind associated with  $X_1$ . The red horizontal line is the truth.

502 The MSEs are shown in Figure 3. Let us look at 3a. As  $n$  increases, the decrease  
 503 is linear for each  $m$ . This indicates that the MSEs go to zero at a polynomial rate,  
 504 even if  $m$  is fixed (look at the line  $m = 1$ ). This agrees with the theoretical results

505 of Section 4. The picture is different for the estimator of Sobol indices of the second  
 506 kind. In 3b, the curve for  $m = 1$  is not a straight line, indicating that the MSE may  
 507 not go to zero. Indeed, the MSE for  $m$  fixed is not expected to go to zero because  
 508 of the bias depending on  $m$ . To make the MSE go to zero, one has to force  $m$  go to  
 509 infinity.

510 Figure 4, which shows the distribution of the estimates for the index associated  
 511 to  $X_1$ , better explains this phenomenon. Here the bias is apparent for  $m = 1$  and  
 512 vanishes as  $m$  goes to infinity. The bias for the indices associated with the other  
 513 inputs is not as large (not shown here).

514 **6. Discussion.** We have considered two kinds of sensitivity indices for stochastic  
 515 models. Asymptotic normality of the estimators, which depend both on the number of  
 516 explorations and the number of repetitions, has been established, and it was noticed  
 517 that the second kind, that which arises from smoothing out the computer model,  
 518 suffers from a bias term which vanishes only when the number of repetitions goes  
 519 to infinity. Assuming a fixed computing budget, the performance of the sensitivity  
 520 indices estimators, measured by the MRE, depends on the design of the Monte-Carlo  
 521 sampling scheme. The optimal design corresponds to the minimal MRE. A bound  
 522 on the MRE has been minimized and a two-stage procedure has been built to get  
 523 estimators that asymptotically achieve the minimal bound. To test the procedure,  
 524 simulation experiments were conducted, where the bias of the sensitivity estimator  
 525 of the second kind was confirmed. Optimal compromises between repetitions and  
 526 explorations have been identified and compared with the output of the two-stage  
 527 procedure for different values of the tuning parameters.

528 This work opens many research directions. First, the sensitivity estimators of the  
 529 two stages could be aggregated to build estimators with a lower variance. Second,  
 530 other methods might be developed to optimize the Monte-Carlo sampling scheme. For  
 531 instance the MSE might be approximated or asymptotic variance-covariance matrices  
 532 might be minimized. Third, multilevel Monte-Carlo sampling schemes might be con-  
 533 sidered to alleviate the bias issue. Fourth, a finite-sample analysis could be conducted  
 534 to get insight into the tradeoff  $K$  is subjected to. Fifth, since the bias is known, it  
 535 could be estimated to build bias-corrected sensitivity indices estimators. Sixth, the  
 536 problem of choosing a number of calls with many divisors must be addressed. It may  
 537 be worth to call the model a bit less if this permits to have a better set  $\text{div}_p(T)$ .  
 538 Seventh, the assumption that  $X$  and  $Z$  are independent might be relaxed.

539 **Appendix A. Proofs.**

540 **Proof of Proposition 3.1.** Assume without loss of generality that  $D_1 < \dots <$   
 541  $D_p$ . We first prove the following Lemma. For convenience, the subscripts  $n$  and  $m$   
 542 are left out.

543 LEMMA A.1. *Let  $i < j$ . Then*

544 
$$P(\widehat{D}_i - \widehat{D}_j \geq 0) \leq \frac{\text{Var } \widehat{D}_i + \text{Var } \widehat{D}_j}{\frac{1}{2}|D_i - D_j|^2}.$$

545

546 *Proof. We have*

547 
$$P(\widehat{D}_i - \widehat{D}_j \geq 0) \leq P(|\widehat{D}_i - D_i| + |\widehat{D}_j - D_j| \geq D_j - D_i)$$

548 
$$\leq P(|\widehat{D}_i - D_i|^2 + |\widehat{D}_j - D_j|^2 \geq \frac{1}{2}|D_j - D_i|^2)$$

549

550

551 *and the claim follows from Markov's inequality.*

552 We now prove Proposition 3.1. Recall that  $D_1 < \dots < D_p$ . We have

$$\begin{aligned}
553 \quad \sum_{i=1}^p \mathbb{E} |\widehat{R}_i - R_i| &\leq \sum_{i=1}^p \sum_{j=1}^p \mathbb{E} |\mathbf{1}(\widehat{D}_j \leq \widehat{D}_i) - \mathbf{1}(D_j \leq D_i)| \\
554 \quad &\leq \sum_{i=1}^p \sum_{j \neq i} \frac{\text{Var } \widehat{D}_i + \text{Var } \widehat{D}_j}{\frac{1}{2}|D_i - D_j|^2} \\
555 \quad &\leq \frac{4(p-1)}{\min_{j < j'} |D_j - D_{j'}|^2} \sum_{i=1}^p \text{Var } \widehat{D}_i, \\
556 \quad &
\end{aligned}$$

557 where the second inequality holds by Lemma A.1 and because

$$558 \quad \mathbb{E} |\mathbf{1}(\widehat{D}_j \leq \widehat{D}_i) - \mathbf{1}(D_j \leq D_i)| = \begin{cases} \mathbb{E} |\mathbf{1}(\widehat{D}_j > \widehat{D}_i)| & \text{if } j < i, \\ 0 & \text{if } j = i, \\ \mathbb{E} |\mathbf{1}(\widehat{D}_j \leq \widehat{D}_i)| & \text{if } j > i. \end{cases}$$

560 It remains to calculate the variances. But this is done in Lemma C.3 in Appendix C,  
561 where it is found that

$$\begin{aligned}
562 \quad \text{Var } \widehat{D}_j &= \frac{1}{n} \{ \text{Var } \mathbb{E}[Y_0 Y_j | \mathbf{X}] + \frac{1}{m} (\mathbb{E} \text{Var}[Y_0 Y_j | \mathbf{X}] - \text{Var}[Y_0 | \mathbf{X}] \text{Var}[Y_j | \mathbf{X}]) \\
563 \quad &+ \frac{1}{m^2} \mathbb{E} \text{Var}[Y_0 | \mathbf{X}] \text{Var}[Y_j | \mathbf{X}] \}. \\
564 \quad &
\end{aligned}$$

565 **Proof of Proposition 3.3.** We distinguish between three cases:  $0 < m^* < 1$ ,  
566  $m^* > (T-K)/(p+1)$  and  $1 \leq m^* \leq (T-K)/(p+1)$ . Recall that  $m_{T-K}^\dagger$  is the  
567 minimizer of  $v(m)$ ,  $m$  in  $\text{div}_p(T-K)$ .

568 If  $0 < m^* < 1$  then by definition  $\varphi_{T-K}(m^*) = 1$  and by convexity  $v(m^*) \leq$   
569  $v(1) \leq v(m)$  for all  $m$  in  $\text{div}_p(T-K)$ . Therefore  $m_{T-K}^\dagger = 1$ .

570 If  $m^* > (T-K)/(p+1)$  then by definition  $\varphi_{T-K}(m^*) = (T-K)/(p+1)$  and by  
571 convexity  $v(m^*) \leq v((T-K)/(p+1)) \leq v(m)$  for all  $m$  in  $\text{div}_p(T-K)$ . Therefore  
572  $m_{T-K}^\dagger = (T-K)/(p+1)$ .

573 If  $1 \leq m^* \leq (T-K)/(p+1)$  then by definition

$$\begin{aligned}
574 \quad \varphi_{T-K}(m^*) &= \begin{cases} \lfloor m^* \rfloor_{T-K} & \text{if } \sqrt{\lfloor m^* \rfloor_{T-K} \lceil m^* \rceil_{T-K}} > m^* \\ \lceil m^* \rceil_{T-K} & \text{if } \sqrt{\lfloor m^* \rfloor_{T-K} \lceil m^* \rceil_{T-K}} \leq m^*. \end{cases} \\
575 \quad &
\end{aligned}$$

576 By convexity  $m_{T-K}^\dagger$  must be  $\lfloor m^* \rfloor_{T-K}$  or  $\lceil m^* \rceil_{T-K}$ . If  $\lfloor m^* \rfloor_{T-K} = \lceil m^* \rceil_{T-K}$  then  
577  $m_{T-K}^\dagger = \lceil m^* \rceil_{T-K} = \varphi_{T-K}(m^*)$ . Otherwise, since  $v(x) = \zeta_1 x + \zeta_2 + \zeta_3/x$ ,  $x > 0$ ,  
578 for some constants  $\zeta_1, \zeta_2$  and  $\zeta_3$  such that  $\zeta_3/\zeta_1 = m^*$ , we have

$$\begin{aligned}
579 \quad v(\lfloor m^* \rfloor_{T-K}) &< v(\lceil m^* \rceil_{T-K}) \text{ iff } \sqrt{\lfloor m^* \rfloor_{T-K} \lceil m^* \rceil_{T-K}} > \frac{\zeta_3}{\zeta_1} = m^*. \\
580 \quad &
\end{aligned}$$

581 Therefore  $\varphi_{T-K}(m^*) = m_{T-K}^\dagger$ .

582 Let us prove that the minimizer of  $v(m)$ ,  $m \in \text{div}_p(T-K)$ , is unique if  $m^* \neq$   
583  $\sqrt{\lfloor m^* \rfloor_{T-K} \lceil m^* \rceil_{T-K}}$ . If it were not, then we would have  $v(\lfloor m^* \rfloor_{T-K})$   
584  $= v(\lceil m^* \rceil_{T-K})$ . But this implies  $m^* = \sqrt{\lfloor m^* \rfloor_{T-K} \lceil m^* \rceil_{T-K}}$ , which is a contra-  
585 diction.

586 **Proof of Theorem 3.4.** In this proof  $m_0$  and  $n_0$  are denoted by  $m$  and  $n$ ,  
 587 respectively. In view of (3.3) and (3.4)–(3.9), we have

$$588 \quad \widehat{m}_K^* = \sqrt{\frac{\sum_{j=1}^p \widehat{\zeta}_{3,j}}{\sum_{j=1}^p \widehat{\zeta}_{1,j}}} = \sqrt{\frac{\sum_{j=1}^p \frac{1}{n} \sum_{i=1}^n \xi_{j;m,i}^{(3.4)} + \xi_{j;m,i}^{(3.5)} - \xi_{j;m,i}^{(3.6)} - \xi_{j;m,i}^{(3.7)}}{\sum_{j=1}^p \frac{1}{n} \sum_{i=1}^n \xi_{j;m,i}^{(3.8)} - \left(\frac{1}{n} \sum_{i=1}^n \xi_{j;m,i}^{(3.9)}\right)^2}},$$

590 where the  $\xi_{j;m,i}^{(e)}$ ,  $i = 1, \dots, n$ ,  $j = 1, \dots, p$ ,  $e = 3.4, \dots, 3.9$ , are implicitly defined  
 591 through (3.4)–(3.9). Let

$$592 \quad \bar{\boldsymbol{\xi}} = \frac{1}{n} \sum_{i=1}^n \boldsymbol{\xi}_{m,i},$$

$$593 \quad \boldsymbol{\xi}_{m,i} = (\boldsymbol{\xi}_{1;m,i}^\top, \dots, \boldsymbol{\xi}_{p;m,i}^\top)^\top, \quad i = 1, \dots, n,$$

$$594 \quad \boldsymbol{\xi}_{j;m,i} = (\xi_{j;m,i}^{(3.4)}, \dots, \xi_{j;m,i}^{(3.9)})^\top, \quad j = 1, \dots, p, \quad i = 1, \dots, n.$$

596 Let  $s$  be the function defined by

$$597 \quad s(\mathbf{x}) = \sqrt{\frac{\sum_{j=1}^p x_j^{(3.4)} + x_j^{(3.5)} - x_j^{(3.6)} - x_j^{(3.7)}}{\sum_{j=1}^p x_j^{(3.8)} - x_j^{(3.9)2}},$$

599 where  $\mathbf{x} = (\mathbf{x}_1^\top, \dots, \mathbf{x}_p^\top)^\top$ ,  $\mathbf{x}_j = (x_j^{(3.4)}, \dots, x_j^{(3.9)})^\top$ ,  $j = 1, \dots, p$ . With the above  
 600 notation we have  $\widehat{m}_K^* = s(\bar{\boldsymbol{\xi}})$ . Moreover, elementary calculations show that

$$601 \quad (\text{A.1}) \quad \mathbf{E} \boldsymbol{\xi}_{m,1} = \boldsymbol{\theta} + \sum_{\nu=1}^4 \frac{\mathbf{C}_\nu}{m^\nu},$$

603 where the  $\mathbf{C}_\nu$  are vectors of constants,  $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \dots, \boldsymbol{\theta}_p^\top)^\top$  and

$$604 \quad \boldsymbol{\theta}_j = \mathbf{E} \begin{pmatrix} Y_0^{(1,1)2} Y_j^{(1,1)2} \\ Y_0^{(1,1)} Y_0^{(1,2)} Y_j^{(1,1)} Y_j^{(1,2)} \\ Y_0^{(1,1)} Y_0^{(1,2)} Y_j^{(1,1)2} \\ Y_j^{(1,1)} Y_j^{(1,2)} Y_0^{(1,1)2} \\ Y_0^{(1,1)} Y_0^{(1,2)} Y_j^{(1,1)} Y_j^{(1,2)} \\ Y_j^{(1,1)} Y_0^{(1,1)} \end{pmatrix}.$$

606 Check that  $m^* = s(\boldsymbol{\theta})$ . A concatenation of two Taylor expansions yield

$$607 \quad \sqrt{n}(\bar{\boldsymbol{\xi}} - \mathbf{E} \boldsymbol{\xi}_{m,1})^\top \dot{s}(\mathbf{E} \boldsymbol{\xi}_{m,1}) + \frac{1}{2}(\bar{\boldsymbol{\xi}} - \mathbf{E} \boldsymbol{\xi}_{m,1})^\top \ddot{s}_{n,m}(\bar{\boldsymbol{\xi}} - \mathbf{E} \boldsymbol{\xi}_{m,1})$$

$$608 \quad (\text{A.2}) \quad = \sqrt{n}(s(\bar{\boldsymbol{\xi}}) - s(\mathbf{E} \boldsymbol{\xi}_{m,1}))$$

$$609 \quad = \sqrt{n}(s(\bar{\boldsymbol{\xi}}) - s(\boldsymbol{\theta}) - (\mathbf{E} \boldsymbol{\xi}_{m,1} - \boldsymbol{\theta})^\top \dot{s}(\boldsymbol{\theta}) - \frac{1}{2}(\mathbf{E} \boldsymbol{\xi}_{m,1} - \boldsymbol{\theta})^\top \ddot{s}_m(\mathbf{E} \boldsymbol{\xi}_{m,1} - \boldsymbol{\theta})),$$

611 where  $\dot{s}$  is the gradient of  $s$ ,  $\ddot{s}_{n,m}$  is the Hessian matrix of  $s$  at a point between  $\bar{\boldsymbol{\xi}}$   
 612 and  $\boldsymbol{\theta}_m$ , and,  $\ddot{s}_m$  is the Hessian matrix of  $s$  at a point between  $\mathbf{E} \boldsymbol{\xi}_{m,1}$  and  $\boldsymbol{\theta}$ . It

613 follows from (A.1) that  $(\mathbf{E}\boldsymbol{\xi}_{m,1} - \boldsymbol{\theta})^\top \dot{s}(\boldsymbol{\theta})$  is clearly of the form  $\sum_{\nu=1}^4 C_\nu/m^\nu$  for  
 614 some constants  $C_\nu$ . Putting

$$615 \quad \epsilon_m = \frac{1}{2}(\mathbf{E}\boldsymbol{\xi}_{m,1} - \boldsymbol{\theta})^\top \ddot{s}_m(\mathbf{E}\boldsymbol{\xi}_{m,1} - \boldsymbol{\theta}) + \sum_{\nu=2}^4 \frac{C_\nu}{m^\nu},$$

617 it follows from (A.2) that

$$618 \quad (\text{A.3}) \quad \sqrt{n}(\bar{\boldsymbol{\xi}} - \mathbf{E}\boldsymbol{\xi}_{m,1})^\top \dot{s}(\mathbf{E}\boldsymbol{\xi}_{m,1}) + \frac{1}{2}(\bar{\boldsymbol{\xi}} - \mathbf{E}\boldsymbol{\xi}_{m,1})^\top \ddot{s}_{n,m}(\bar{\boldsymbol{\xi}} - \mathbf{E}\boldsymbol{\xi}_{m,1})$$

$$619 \quad \quad \quad = \sqrt{n}(\widehat{m}_K^* - m^* - \frac{C_1}{m} - \epsilon_m).$$

622 If  $m$  is fixed then Lemma B.2 in Appendix B yields

$$623 \quad \sqrt{n}(\bar{\boldsymbol{\xi}} - \mathbf{E}\boldsymbol{\xi}_{m,1}) \rightarrow N(0, \Sigma_m),$$

625 for some variance-covariance matrix  $\Sigma_m$  of size  $6p \times 6p$ . Moreover, the second term in  
 626 the left-hand side of (A.3) is  $o_P(1)$  by Cauchy-Schwartz's inequality and the continuity  
 627 of the second derivatives of  $s$ . The first term goes to  $N(0, \dot{s}(\mathbf{E}\boldsymbol{\xi}_{m,1})^\top \Sigma_m \dot{s}(\mathbf{E}\boldsymbol{\xi}_{m,1}))$   
 628 and hence the claim follows with  $\sigma_m^2 = \dot{s}(\mathbf{E}\boldsymbol{\xi}_{m,1})^\top \Sigma_m \dot{s}(\mathbf{E}\boldsymbol{\xi}_{m,1})$  and  $C = C_1$ .

629 If  $m \rightarrow \infty$  then again Lemma B.2 in Appendix B applies: we have

$$630 \quad \sqrt{n}(\bar{\boldsymbol{\xi}} - \mathbf{E}\boldsymbol{\xi}_{m,1}) \rightarrow N(0, \lim_{m \rightarrow \infty} \Sigma_m).$$

632 Since  $\epsilon_m - \sum_{\nu=2}^4 C_\nu/m^\nu = o(m^{-1})$ ,  $\dot{s}$  is continuous and  $\mathbf{E}\boldsymbol{\xi}_{m,1} \rightarrow \boldsymbol{\theta}$ , the claim follows.  
 633 The proof is complete.

634 **Proof of Proposition 3.6.** By definition,  $\widehat{m}_{T-K}^\dagger = \varphi_{T-K}(\widehat{m}_K^*)$  and  $m_{T-K}^\dagger =$   
 635  $\varphi_{T-K}(m^*)$ . The function  $\varphi_{T-K}$  is piecewise constant and has  $|\text{div}_p(T-K)| - 1$  points  
 636 of discontinuity of the form  $\sqrt{ij}$ , where  $i$  and  $j$  are two consecutive members of

$$637 \quad \text{div}_p(T-K) \setminus \left\{ 1, \frac{T-K}{p+1} \right\}.$$

639 Denote the set of discontinuity points by  $\mathcal{D}_{T-K}$ . Clearly,

$$640 \quad \mathcal{D}_{T-K} \subset \{ \sqrt{ij} : i \text{ and } j \text{ are two consecutive integers} \} = \mathcal{E}.$$

642 There exists an open interval that contains  $m^*$  but does not contain any points of  
 643  $\mathcal{E}$  and hence does not contain any points of  $\mathcal{D}_{T-K}$ , whatever  $T$  and  $K$ . If  $\widehat{m}_K^*$  is in  
 644 this interval then there are no discontinuity points between  $m^*$  and  $\widehat{m}_K^*$  and hence  
 645  $\widehat{m}_{T-K}^\dagger = \varphi_{T-K}(\widehat{m}_K^*) = \varphi_{T-K}(m^*) = m_{T-K}^\dagger$ . By Corollary 3.5, the probability of  
 646 that event goes to one as  $m_0$  and  $n_0$  go to infinity.

647 **Proof of Theorem 3.7.** Let  $\varepsilon > 0$ . An obvious algebraic manipulation and  
 648 Taylor's expansion yield

$$649 \quad P \left( \left| \frac{\frac{1}{T-K} v(\widehat{m}_{T-K}^\dagger) - \frac{1}{T} v(m_{T-K}^\dagger)}{\frac{1}{T} v(m_{T-K}^\dagger)} > \varepsilon \right| \right)$$

$$650 \quad \leq P \left( \left| \frac{T}{T-K} (\widehat{m}_{T-K}^\dagger - m_{T-K}^\dagger) v'(\tilde{m}) + \frac{K}{T-K} v(m_{T-K}^\dagger) \right| > v(m_{T-K}^\dagger) \varepsilon \right),$$

651

652 where  $\tilde{m}$  denotes a real between  $\hat{m}_{T-K}^\dagger$  and  $m_{T-K}^\dagger$ . A decomposition of the probability  
 653 above according to whether  $\hat{m}_{T-K}^\dagger - m_{T-K}^\dagger \neq 0$  or  $\hat{m}_{T-K}^\dagger - m_{T-K}^\dagger = 0$  yields the  
 654 bound

$$655 \quad P\left(\hat{m}_{T-K}^\dagger - m_{T-K}^\dagger \neq 0\right) + P\left(\frac{K}{T-K} > \varepsilon\right).$$

657 The first term goes to zero by Proposition 3.6. The second term goes to zero because  
 658  $K/T \rightarrow 0$ .

659 **Proof of Theorem 4.1.** The proof is based on the results in Appendix B. The  
 660 Sobol estimators in (2.11) and (2.12) are of the form

$$661 \quad \hat{S}_{j;n,m}' = \frac{\frac{1}{n} \sum_{i=1}^n \xi_{j;m,i}^{\text{UL}} - \left(\frac{1}{n} \sum_{i=1}^n \xi_{m,i}^{\text{UR}}\right)^2}{\frac{1}{n} \sum_{i=1}^n \xi_{m,i}^{\text{LL}} - \left(\frac{1}{n} \sum_{i=1}^n \xi_{m,i}^{\text{UR}}\right)^2}, \quad j = 1, \dots, p,$$

663 and

$$664 \quad \hat{S}_{j;n,m}'' = \frac{\frac{1}{n} \sum_{i=1}^n \xi_{j;m,i}^{\text{UL}} - \left(\frac{1}{n} \sum_{i=1}^n \xi_{m,i}^{\text{UR}}\right)^2}{\frac{1}{n} \sum_{i=1}^n \xi_{m,i}^{\text{LL}} - \left(\frac{1}{n} \sum_{i=1}^n \xi_{m,i}^{\text{UR}}\right)^2}, \quad j = 1, \dots, p,$$

666 where the notation is obvious. Denote  $\boldsymbol{\xi}_{m,i} := (\xi_{1;m,i}^{\text{UL}}, \dots, \xi_{p;m,i}^{\text{UL}}, \xi_{m,i}^{\text{UR}}, \xi_{m,i}^{\text{LL}}, \xi_{m,i}^{\text{LL}})^\top$ .  
 667 Elementary but burdensome calculations show that

$$668 \quad \mathbb{E} \boldsymbol{\xi}_{m,1} = \begin{pmatrix} \mathbb{E} \mathbb{E}[f(X, Z)|X] \mathbb{E}[f(\tilde{X}_{-1}, Z)|\tilde{X}_{-1}] \\ \vdots \\ \mathbb{E} \mathbb{E}[f(X, Z)|X] \mathbb{E}[f(\tilde{X}_{-p}, Z)|\tilde{X}_{-p}] \\ \mathbb{E} f(X, Z) \\ \mathbb{E} f(X, Z)^2 \\ \mathbb{E} \mathbb{E}[f(X, Z)|X]^2 + \frac{\mathbb{E} \text{Var}[f(X, Z)|X]}{m} \end{pmatrix}.$$

669 (Some calculations are carried out in Appendix C.) Define the function

$$670 \quad s(x_1, \dots, x_p, x_{p+1}, x_{p+2}, x_{p+3})$$

$$671 \quad = \left( \frac{x_1 - x_{p+1}^2}{x_{p+2} - x_{p+1}^2}, \dots, \frac{x_p - x_{p+1}^2}{x_{p+2} - x_{p+1}^2}, \frac{x_1 - x_{p+1}^2}{x_{p+3} - x_{p+1}^2}, \dots, \frac{x_p - x_{p+1}^2}{x_{p+3} - x_{p+1}^2} \right).$$

674 Clearly, we have

$$675 \quad s\left(\frac{1}{n} \sum_{i=1}^n \boldsymbol{\xi}_{m,i}\right) = \begin{pmatrix} \hat{\mathbf{S}}_{n,m}' \\ \hat{\mathbf{S}}_{n,m}'' \end{pmatrix}$$

676 and

$$677 \quad s(\mathbb{E} \boldsymbol{\xi}_{m,1}) = \begin{pmatrix} \mathbf{S}' \\ \mathbf{S}'' \left[ 1 - \frac{\mathbb{E} \text{Var}[f(X, Z)|X]}{\mathbb{E} \text{Var}[f(X, Z)|X] + m \text{Var} \mathbb{E}[f(X, Z)|X]} \right] \end{pmatrix}.$$

678 If  $m$  is fixed then Lemma B.2 in Appendix B yields

$$679 \quad \sqrt{n} \left( \frac{1}{n} \sum_{i=1}^n \boldsymbol{\xi}_{m,i} - \mathbb{E} \boldsymbol{\xi}_{m,1} \right) \xrightarrow{d} N(0, \Sigma_m),$$

680 for some nonnegative matrix  $\Sigma_m$  of size  $(p+3) \times (p+3)$  and the result follows by the  
681 delta-method.

682 If  $m \rightarrow \infty$ , Lemma B.2 still holds with the variance-covariance matrix replaced  
683 by its limit. Taylor's expansion yields

$$\begin{aligned}
684 \quad & \sqrt{n} \left( s \left( \frac{1}{n} \sum_{i=1}^n \boldsymbol{\xi}_{m,i} \right) - s(\mathbb{E} \boldsymbol{\xi}_{m,1}) \right) \\
685 \quad & = \sqrt{n} \left( \left( \frac{1}{n} \sum_{i=1}^n \boldsymbol{\xi}_{m,i} - \mathbb{E} \boldsymbol{\xi}_{m,1} \right) \dot{s}_m \right. \\
686 \quad & \quad \left. + \frac{1}{2} \left( \frac{1}{n} \sum_{i=1}^n \boldsymbol{\xi}_{m,i} - \mathbb{E} \boldsymbol{\xi}_{m,1} \right)^\top \ddot{s}_{n,m} \left( \frac{1}{n} \sum_{i=1}^n \boldsymbol{\xi}_{m,i} - \mathbb{E} \boldsymbol{\xi}_{m,1} \right) \right), \\
687
\end{aligned}$$

688 where  $\dot{s}_m$  is the gradient of  $s$  at  $\mathbb{E} \boldsymbol{\xi}_{m,1}$  and  $\ddot{s}_{n,m}$  is the Hessian matrix of  $s$  at a  
689 point between  $n^{-1} \sum_i \boldsymbol{\xi}_{m,i}$  and  $\mathbb{E} \boldsymbol{\xi}_{m,1}$ . Since that point goes to a constant and  $s$  has  
690 continuous second derivatives, it holds that  $\ddot{s}_{n,m}$  goes to a constant as well. So does  
691  $\dot{s}_m$  and the claim follows by Slutsky's lemma.

692 **Appendix B. A unified treatment of the asymptotics.** All estimators in  
693 this paper have a common form, given by

$$694 \quad (\text{B.1}) \quad \frac{1}{n} \sum_{i=1}^n \boldsymbol{\xi}_{m,i},$$

695 with

$$696 \quad (\text{B.2}) \quad \boldsymbol{\xi}_{m,i} = \prod_{l=1}^L \frac{1}{m} \sum_{k=1}^m \prod_{j=0}^p Y_j^{(i,k) b_{j;l}},$$

698 where  $Y_0^{(i,k)} = Y^{(i,k)} = f(X^{(i)}, Z_0^{(i,k)})$ ,  $Y_j^{(i,k)} = f(\tilde{X}_{-j}^{(i)}, Z_j^{(i,k)})$  for  $j = 1, \dots, p$ , and  
699  $b_{j;l}$ ,  $j = 0, \dots, p$ ,  $l = 1, \dots, L$ , are nonnegative coefficients. The coefficients are  
700 arranged in a matrix  $(b_{j;l})$  with  $L$  rows and  $p+1$  columns, where  $b_{j;l}$  is the element in  
701 the  $l$ th row and  $(j+1)$ th column. This way, all estimators of the form (B.1) and (B.2),  
702 or, equivalently, all summands (B.2), can be represented by a matrix. We sometimes  
703 write  $\boldsymbol{\xi}_{m,i} \simeq (b_{j;l})$ , where  $(b_{j;l})$  is the matrix of size  $L \times (p+1)$  with coefficients  $b_{j;l}$ ,  
704  $j = 0, \dots, p$ ,  $l = 1, \dots, L$ .

705 **B.1. Examples.** The estimator

$$706 \quad \frac{1}{n} \sum_{i=1}^n \frac{1}{m} \sum_{k=1}^m Y_0^{(i,k)} \frac{1}{m} \sum_{k'=1}^m Y_j^{(i,k')}$$

708 is of the form (B.1) and (B.2) with  $L = 2$  and coefficients

$$709 \quad \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix},$$

710

711 where the non-null columns are the first and the  $(j + 1)$ th ones. The estimators

$$712 \quad \frac{1}{n} \sum_{i=1}^n \frac{1}{m} \sum_{k=1}^m Y_0^{(i,k)}, \quad \frac{1}{n} \sum_{i=1}^n \frac{1}{m} \sum_{k=1}^m Y_0^{(i,k)2},$$

$$713 \quad \frac{1}{n} \sum_{i=1}^n \left( \frac{1}{m} \sum_{k=1}^m Y_0^{(i,k)} \right)^2$$

714

715 are of the form (B.1) and (B.2) with  $L = 2$  and coefficients

$$716 \quad \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix},$$

$$717 \quad \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix},$$

718

719 respectively.

720 **The estimators of Section 3.** In view of (3.4)–(3.9), the estimators  $\widehat{\zeta}_{3,j}$  and  
721  $\widehat{\zeta}_{1,j}$  can be expressed in terms of estimators of the form (B.1) and (B.2): we have

$$722 \quad \widehat{\zeta}_{3,j} = \frac{1}{n} \sum_{i=1}^n \xi_{j;m,i}^{(3.4)} + \xi_{j;m,i}^{(3.5)} - \xi_{j;m,i}^{(3.6)} - \xi_{j;m,i}^{(3.7)} \quad \text{and,}$$

$$723 \quad \widehat{\zeta}_{1,j} = \frac{1}{n} \sum_{i=1}^n \xi_{j;m,i}^{(3.8)} - \left( \frac{1}{n} \sum_{i=1}^n \xi_{j;m,i}^{(3.9)} \right)^2,$$

724

725 where

$$726 \quad \xi_{j;m,i}^{(3.4)}, \quad \xi_{j;m,i}^{(3.5)}$$

$$727 \quad \xi_{j;m,i}^{(3.6)}, \quad \xi_{j;m,i}^{(3.7)}$$

$$728 \quad \xi_{j;m,i}^{(3.8)}, \quad \xi_{j;m,i}^{(3.9)}$$

729

730 are all of the form (B.2) with  $L = 4$  and coefficients

$$731 \quad \begin{pmatrix} 2 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 2 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \end{pmatrix},$$

$$732 \quad \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 2 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix},$$

$$733 \quad \begin{pmatrix} 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix},$$

734

735 respectively. In the matrices above, the first and  $j + 1$ th columns are nonnull.



736 **The estimators of Section 4.** The Sobol estimators in (2.11) and (2.12) are  
 737 of the form (B.1) and (B.2) with  $L = 2$  and coefficients

$$738 \quad \xi_{1;m,i}^{\text{UL}} \simeq \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \end{pmatrix}, \dots, \xi_{p;m,i}^{\text{UL}} \simeq \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

739 for the upper left (UL) terms,

$$740 \quad \xi_{m,i}^{\text{UR}} \simeq \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

741 for the upper right (UR) term,

$$742 \quad \xi_{m,i}^{\text{LL}} \simeq \begin{pmatrix} 2 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

743 for the lower left (LL) term of  $\widehat{S}'_{j;n,m}$  and

$$744 \quad \xi_{m,i}^{\text{LL}} \simeq \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \end{pmatrix}$$

745 for the lower left (LL) term of  $\widehat{S}''_{j;n,m}$ .

746 **B.2. A central limit theorem.** For each  $n$ , the random variables  $\xi_{m,1}, \dots, \xi_{m,n}$   
 747 are independent and identically distributed. Denote by  $\mathcal{E}_{m,i}(L)$  the set of all sum-  
 748 mands (B.2). In other words,  $\mathcal{E}_{m,i}(L)$  is the set of all nonnegative matrices of size  
 749  $L \times (p+1)$ . This set has useful properties, gathered in Proposition B.1 for subsequent  
 750 use.

751 **PROPOSITION B.1.** *Let  $\xi$  be an element of  $\mathcal{E}_{m,i}(L)$  with coefficients  $(b_{j;l})$ . The*  
 752 *following statements are true.*

753 (i) *If  $\xi'$  is an element of  $\mathcal{E}_{m,i}(L)$  with coefficients  $(b'_{j;l})$  then  $\xi\xi'$  is an element of*  
 754  *$\mathcal{E}_{m,i}(2L)$  with coefficients*

$$755 \quad \begin{pmatrix} b_{0;1} & \cdots & b_{p;1} \\ \vdots & & \vdots \\ b_{0;L} & \cdots & b_{p;L} \\ b'_{0;1} & \cdots & b'_{p;1} \\ \vdots & & \vdots \\ b'_{0;L} & \cdots & b'_{p;L} \end{pmatrix}.$$

757 (ii) *The limit of  $\mathbb{E}\xi$  exists as  $m \rightarrow \infty$ .*

758 (iii) *If there exists some function  $F$  such that  $|f(x, z)| \leq F(x)$  for all  $x$  and  $z$  in the*  
 759 *domain of definition of  $f$  then*

$$760 \quad |\xi| \leq \left( \bigvee_{j=0}^p F_j(\mathbf{X}^{(i)}) \right)^{\sum_{j=0}^p \sum_{l=1}^L b_{j;l}},$$

761 *where  $F_j(\mathbf{X}^{(i)})$  is  $F(X^{(i)}) \vee 1$  if  $j = 0$  and  $F(\widetilde{X}_{-j}^{(i)}) \vee 1$  if  $j \geq 1$ .*

762 *Proof.* The proof of (i) is trivial. Let us prove (ii). We have

$$\begin{aligned}
 763 \quad \mathbb{E} \xi &= \frac{1}{m^L} \sum_{(k_1, \dots, k_L) \in \{1, \dots, m\}^L} \mathbb{E} \prod_{l=1}^L \prod_{j=0}^p Y_j^{(1, k_l) b_{j;l}} \\
 764 &= \frac{1}{m^L} \sum_{(k_1, \dots, k_L) \in \{1, \dots, m\}^L} \mathbb{E} \mathbb{E} \left( \prod_{l=1}^L \prod_{j=0}^p Y_j^{(1, k_l) b_{j;l}} \middle| \mathbf{X}^{(1)} \right) \\
 765 \quad (\text{B.3}) &= \frac{1}{m^L} \sum_{(k_1, \dots, k_L) \in \{1, \dots, m\}^L} \mathbb{E} \prod_{j=0}^p \mathbb{E} \left( \prod_{l=1}^L Y_j^{(1, k_l) b_{j;l}} \middle| \mathbf{X}^{(1)} \right). \\
 766
 \end{aligned}$$

767 Since (i)  $\mathbf{X}^{(1)}$  and  $\{\mathbf{Z}^{(1, k)}, k = 1, \dots, m\}$  are independent and (ii) the law of

$$768 \quad (\mathbf{Z}^{(1, k_1)}, \dots, \mathbf{Z}^{(1, k_L)})$$

769 is invariant through any permutation of distinct  $k_1, \dots, k_L$ , all the inner expectations  
770 in (B.3) are equal to some others. For if  $k_1, \dots, k_L$  are distinct then

$$771 \quad \mathbb{E} \left( \prod_{l=1}^L Y_j^{(1, k_l) b_{j;l}} \middle| \mathbf{X}^{(1)} \right) = \mathbb{E} \left( \prod_{l=1}^L Y_j^{(1, l) b_{j;l}} \middle| \mathbf{X}^{(1)} \right)$$

772  
773 for all  $j = 0, \dots, p$ . The number of inner expectations equal to the one above is  
774  $m(m-1) \cdots (m-L+1)$ , a polynomial in  $m$  with degree  $L$ . If some components of  
775 the tuple  $(k_1, \dots, k_L)$  are equal, then we can always write

$$776 \quad \mathbb{E} \left( \prod_{l=1}^L Y_j^{(1, k_l) b_{j;l}} \middle| \mathbf{X}^{(1)} \right) = \mathbb{E} \left( \prod_{l=1}^{L'} Y_j^{(1, l) \beta_{j;l}} \middle| \mathbf{X}^{(1)} \right)$$

777 for some  $L' \leq L$  and coefficients  $\beta_{j;l}$ . It is easy to see that the number of inner expecta-  
778 tions equal to the one above is a polynomial in  $m$  with degree at most  $L$ . (Looking  
779 at examples helps to see this; see e.g. the proof of Lemma C.2 in Appendix C.)  
780 Therefore, the sum in (B.3) is also a polynomial in  $m$  with degree at most  $L$  and the  
781 claim follows ( $\mathbb{E} \xi$  can be zero). To prove (iii), simply remember that, by assumption,  
782  $|Y^{(1, k)}| \leq F(X^{(1)})$  and  $|Y_j^{(1, k)}| \leq F(\tilde{X}_{-j}^{(1)})$  for all  $k$  and all  $j$ .  $\square$   
783

784 Two frameworks are considered:

- 785 •  $n \rightarrow \infty$  and  $m$  is fixed;
- 786 •  $n \rightarrow \infty$  and  $m \rightarrow \infty$ .

787 In the second framework  $m_n$  is a sequence indexed by  $n$  that goes to infinity as  $n$  goes  
788 to infinity.

789 LEMMA B.2. Let  $\xi_{m,i}^{(I)}$ ,  $I = 1, \dots, N$ , be elements of  $\mathcal{E}_{m,i}(L)$  with coefficients  
790  $(b_{j;l}^{(I)})$ . Assume

$$791 \quad \mathbb{E} F(X^{(1)})^2 \sum_{j=0}^p \sum_{l=1}^L b_{j;l}^{(I)} < \infty$$

792 for all  $I = 1, \dots, N$ . Let  $n \rightarrow \infty$ . If  $m$  is fixed then

$$793 \quad \sqrt{n} \left[ \frac{1}{n} \sum_{i=1}^n \xi_{m,i}^{(1)} - \mathbb{E} \xi_{m,1}^{(1)}, \dots, \frac{1}{n} \sum_{i=1}^n \xi_{m,i}^{(N)} - \mathbb{E} \xi_{m,1}^{(N)} \right]^\top \xrightarrow{d} N(0, \Sigma_m),$$

794

795 where  $\Sigma_m$  is the variance-covariance matrix of  $\boldsymbol{\xi}_{m,i} = (\xi_{m,i}^{(1)}, \dots, \xi_{m,i}^{(N)})^\top$ . If  $m \rightarrow$   
 796  $\infty$  then  $\lim_{m \rightarrow \infty} \Sigma_m$  exists elementwise and the above display with  $\Sigma_m$  replaced by  
 797  $\lim_{m \rightarrow \infty} \Sigma_m$  is true.

798 *Proof.* Let  $m$  be fixed. By Proposition B.1 (i),  $\xi_{m,i}^{(I)2}$ ,  $I = 1, \dots, N$ , belongs to  
 799  $\mathcal{E}_{m,i}(2L)$  and has coefficients

$$800 \quad \xi_{m,i}^{(I)2} \simeq \begin{pmatrix} b_{0;1}^{(I)} & \cdots & b_{p;1}^{(I)} \\ \vdots & & \vdots \\ b_{0;L}^{(I)} & \cdots & b_{p;L}^{(I)} \\ b_{0;1}^{(I)} & \cdots & b_{p;1}^{(I)} \\ \vdots & & \vdots \\ b_{0;L}^{(I)} & \cdots & b_{p;L}^{(I)} \end{pmatrix}.$$

802 Thus, denoting  $\sum_{j=0}^p \sum_{l=1}^L b_{j;l}^{(I)}$  by  $\beta$ , Proposition B.1 (iii) yields

$$803 \quad (\text{B.4}) \quad \xi_{m,i}^{(I)2} \leq \bigvee_{j=0}^p F_j(\mathbf{X}^{(i)})^{2\beta}$$

805 and hence

$$806 \quad \mathbb{E} \xi_{m,i}^{(I)2} \leq \mathbb{E} \bigvee_{j=0}^p F_j(\mathbf{X}^{(1)})^{2\beta} \leq (p+1) \mathbb{E} \left( 1 \vee F(X^{(1)}) \right)^{2\beta} < \infty.$$

808 Therefore we can apply the central limit theorem to finish the proof for  $m$  fixed.

809 Let  $m \rightarrow \infty$ . According to Lindeberg-Feller's central limit theorem (see e.g. [23]),  
 810 it suffices to show

811 (i) for all  $\epsilon > 0$ ,

$$812 \quad \sum_{i=1}^n \mathbb{E} \left\| \frac{1}{\sqrt{n}} \boldsymbol{\xi}_{m,i} \right\|^2 \mathbf{1} \left\{ \left\| \frac{1}{\sqrt{n}} \boldsymbol{\xi}_{m,i} \right\| > \epsilon \right\} \rightarrow 0,$$

814 and

815 (ii) the limit  $\sum_{i=1}^n \text{Cov}(\boldsymbol{\xi}_{m,i}/\sqrt{n})$  exists and is finite.

816 Let us show (i). Denoting  $\mathbf{X} = (X^{(1)}, \tilde{X}^{(1)})$ , we have

$$817 \quad \sum_{i=1}^n \mathbb{E} \left\| \frac{\boldsymbol{\xi}_{m,i}}{\sqrt{n}} \right\|^2 \mathbf{1} \{ \|\boldsymbol{\xi}_{m,i}\| > \sqrt{n}\epsilon \} = \mathbb{E} \|\boldsymbol{\xi}_{m,1}\|^2 \mathbf{1} \{ \|\boldsymbol{\xi}_{m,1}\| > \sqrt{n}\epsilon \}$$

$$818 \quad = \mathbb{E} \sum_{I=1}^N \xi_{m,1}^{(I)2} \mathbf{1} \{ \|\boldsymbol{\xi}_{m,1}\| > \sqrt{n}\epsilon \}$$

$$819 \quad = \sum_{I=1}^N \mathbb{E} \left[ \mathbb{E} \left( \xi_{m,1}^{(I)2} \mathbf{1} \{ \|\boldsymbol{\xi}_{m,1}\| > \sqrt{n}\epsilon \} \mid \mathbf{X} \right) \right].$$

820

821 By (B.4), we have

$$\begin{aligned}
 822 \quad \mathbb{E} \left( \xi_{m,1}^{(I)2} \mathbf{1} \{ \|\xi_{m,1}\| > \sqrt{n}\epsilon \} | \mathbf{X} \right) &\leq \prod_{j=0}^p F_j(\mathbf{X}^{(1)})^{2\beta} P(\|\xi_{m,1}\| > \sqrt{n}\epsilon | \mathbf{X}) \\
 823 \quad &\leq \prod_{j=0}^p F_j(\mathbf{X}^{(1)})^{2\beta} \frac{\sum_{I=1}^N \mathbb{E} \left( \xi_{m,1}^{(I)2} | \mathbf{X} \right)}{n\epsilon^2} \\
 824 \quad &\leq \frac{N \prod_{j=0}^p F_j(\mathbf{X}^{(1)})^{4\beta}}{n\epsilon^2}, \\
 825
 \end{aligned}$$

826 where the last inequality holds by using (B.4) once more. The upper bound goes to  
 827 zero and is dominated by an integrable function. Thus, we can apply the dominated  
 828 convergence theorem to complete the proof.

829 Let us show that (ii) holds. We have  $\sum_{i=1}^n \text{Cov}(\xi_{m,i}/\sqrt{n}) = \text{Cov}(\xi_{m,1})$ . The  
 830 element  $(I, J)$  in this matrix is given by  $\mathbb{E} \xi_{m,1}^{(I)} \xi_{m,1}^{(J)} - \mathbb{E} \xi_{m,1}^{(I)} \mathbb{E} \xi_{m,1}^{(J)}$ . Remember that  
 831  $\mathbb{E} \xi_{m,1}^{(I)2} < \infty$ ,  $I = 1, \dots, N$ , and hence  $\mathbb{E} \xi_{m,1}^{(I)} \xi_{m,1}^{(J)} \leq \mathbb{E} \xi_{m,1}^{(I)2}/2 + \mathbb{E} \xi_{m,1}^{(J)2}/2 < \infty$ . Therefore  
 832 the limit of  $\text{Cov} \xi_{m,1}$  exists and is finite. The proof is complete.  $\square$

833 **Appendix C. Explicit moment calculations.** Explicit moment calculations  
 834 are given for the summands in the proof of Theorem 4.1. In this section,  $\mathbb{E} f(X, Z)$   
 835 and  $\mathbb{E} \mathbb{E}[f(X, Z) | X]^2$  are denoted by  $\mu$  and  $D$ , respectively. Recall that the upper-left  
 836 term in (2.8) and (2.9) is denoted by  $D_j$ . The moments are given in Lemma C.1  
 837 and Lemma C.2. The variances and covariances are given in Lemma C.3. Let  $\mathbf{X} =$   
 838  $(X^{(1)}, \tilde{X}^{(1)})$ . Whenever there is a superscript  $\mathbf{X}$  added to the expectation symbol  $\mathbb{E}$   
 839 or the variance symbol  $\text{Var}$ , this means that these operators are to be understood  
 840 conditionally on  $\mathbf{X}$ . An integral with respect to  $P(d\mathbf{x})$  means that we integrate with  
 841 respect to the law of  $\mathbf{X}$ .

842 LEMMA C.1 (Moments of order 1). *The moments of order 1 are given by*

$$\begin{aligned}
 843 \quad \mathbb{E} \xi_{m1}^{UL} &= D_j, \\
 844 \quad \mathbb{E} \xi_{m1}^{UR} &= \mu, \\
 845 \quad \mathbb{E} \xi_{m1}^{LL} &= \frac{1}{m} \mathbb{E} \text{Var}^X f(X^{(1)}, Z^{(1,1)}) + D. \\
 846
 \end{aligned}$$

847

848 *Proof.* One has

$$\begin{aligned}
 849 \quad \mathbb{E} \xi_{m1}^{UL} &= \frac{1}{m^2} \sum_{k,k'} \mathbb{E} f(X^{(1)}, Z^{(1,k)}) f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,k')}) \\
 850 \quad &= \frac{1}{m^2} \sum_{k,k'} \int \mathbb{E} f(x, Z^{(1,k)}) f(\tilde{x}_{-j}, Z_j^{(1,k')}) P(d\mathbf{x}) \\
 851 \quad &= \mathbb{E} f(X^{(1)}, Z^{(1,1)}) f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)}) \\
 852 \quad &= D_j, \\
 853
 \end{aligned}$$

854 where the integral is taken with respect to the law of  $\mathbf{x} = (x, \tilde{x})$ , and,

$$\begin{aligned}
855 \quad \mathbb{E} \xi_{m1}^{\prime\prime LL} &= \frac{1}{m^2} \sum_{k, k'} \mathbb{E} f(X^{(1)}, Z^{(1,k)}) f(X^{(1)}, Z^{(1,k')}) \\
856 &= \frac{1}{m} \mathbb{E} \text{Var}^X f(X, Z) + \mathbb{E} (\mathbb{E}^X f(X, Z))^2 \\
857 &= \frac{1}{m} \mathbb{E} \text{Var}^X f(X, Z) + D. \\
858
\end{aligned}$$

859 The proof for  $\xi_{m1}^{\text{UR}}$  is similar. □

860 LEMMA C.2 (Moments of order 2). *The moments of order 2 are given by*

$$\begin{aligned}
861 \quad \mathbb{E} \xi_{m1}^{(UL)2} &= \text{Var} \mathbb{E}^{\mathbf{X}} f(X^{(1)}, Z^{(1,1)}) f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)}) + D_j^2 \\
862 &\quad + \frac{1}{m} [\mathbb{E} \text{Var}^{\mathbf{X}} f(X^{(1)}, Z^{(1,1)}) f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)}) \\
863 &\quad - \text{Var}^{\mathbf{X}} f(X^{(1)}, Z^{(1,1)}) \text{Var}^{\mathbf{X}} f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)})] \\
864 &\quad + \frac{1}{m^2} \mathbb{E} \text{Var}^{\mathbf{X}} f(X^{(1)}, Z^{(1,1)}) \text{Var}^{\mathbf{X}} f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)}), \\
865 \quad \mathbb{E} \xi_{m1}^{(UR)2} &= \frac{1}{m} \mathbb{E} \text{Var}^{\mathbf{X}} f(X^{(1)}, Z^{(1,1)}) + \mathbb{E} (\mathbb{E}^{\mathbf{X}} f(X^{(1)}, Z^{(1,1)}))^2, \\
866 \quad \mathbb{E} \xi_{m1}^{\prime\prime LL} &= \frac{m(m-1)(m-2)(m-3)}{m^4} \\
867 &\quad \mathbb{E} f(X^{(1)}, Z^{(1,1)}) f(X^{(1)}, Z^{(1,2)}) f(X^{(1)}, Z^{(1,3)}) f(X^{(1)}, Z^{(1,4)}) \\
868 &\quad + \frac{\binom{4}{2} m(m-1)(m-2)}{m^4} \mathbb{E} f(X^{(1)}, Z^{(1,1)})^2 f(X^{(1)}, Z^{(1,2)}) f(X^{(1)}, Z^{(1,3)}) \\
869 &\quad + \frac{\binom{4}{3} m(m-1)}{m^4} \mathbb{E} f(X^{(1)}, Z^{(1,1)})^3 f(X^{(1)}, Z^{(1,2)}) \\
870 &\quad + \frac{m}{m^4} \mathbb{E} f(X^{(1)}, Z^{(1,1)})^4 \\
871 &\quad + \frac{\binom{4}{2} m(m-1)/2}{m^4} \mathbb{E} f(X^{(1)}, Z^{(1,1)})^2 f(X^{(1)}, Z^{(1,2)})^2 \\
872
\end{aligned}$$

874 *Proof.* Let us first deal with  $\xi_{m1}^{\text{UL}}$ . We have

$$\begin{aligned}
875 \quad \mathbb{E} \xi_{m1}^{(UL)2} &= \frac{1}{m^4} \sum_{k_1, k_2, k_3, k_4} \mathbb{E} f(X^{(1)}, Z^{(1, k_1)}) f(X^{(1)}, Z^{(1, k_2)}) \\
876 &\quad f(\tilde{X}_{-j}^{(1)}, Z_j^{(1, k_3)}) f(\tilde{X}_{-j}^{(1)}, Z_j^{(1, k_4)}) \\
877
\end{aligned}$$

879 where, in the sum, the indices run over  $1, \dots, m$ . We split the sum into four parts.  
880 The first contains the  $m^2(m-1)^2$  terms that satisfy  $k_1 \neq k_3$  and  $k_2 \neq k_4$ . In this  
881 part, all the terms are equal to

$$\begin{aligned}
882 \quad (\text{term 1}) &\quad \mathbb{E} \left( \mathbb{E}^{\mathbf{X}} f(X^{(1)}, Z^{(1,1)}) f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)}) \right)^2. \\
883
\end{aligned}$$

884 The second part contains the  $m^2(m-1)$  terms that satisfy  $k_1 \neq k_3$  and  $k_2 = k_4$  and  
885 that are equal to

$$\begin{aligned}
886 \quad (\text{term 2}) &\quad \mathbb{E} f(X^{(1)}, Z^{(1,1)}) f(X^{(1)}, Z^{(1,2)}) f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)})^2. \\
887
\end{aligned}$$

888 The third part contains the  $m^2(m-1)$  terms that satisfy  $k_1 = k_3$  and  $k_2 \neq k_4$  and  
 889 that are equal to

$$890 \quad (\text{term 3}) \quad \mathbb{E} f(X^{(1)}, Z^{(1,1)})^2 f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)}) f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,2)}).$$

892 Finally, the fourth part contains the  $m^2$  terms that satisfy  $k_1 = k_4$  and  $k_2 = k_3$  and  
 893 are equal to

$$894 \quad (\text{term 4}) \quad \mathbb{E} f(X^{(1)}, Z^{(1,1)})^2 f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)})^2.$$

896 (One can see that the number of terms is  $m^4$ .) Thus,

$$\begin{aligned} 897 \quad \mathbb{E} \xi_{m1}^{(\text{UL})2} &= (\text{term 1}) \\ 898 \quad &+ \frac{1}{m} [(\text{term 2}) + (\text{term 3}) - 2(\text{term 1})] \\ 899 \quad &+ \frac{1}{m^2} [(\text{term 1}) - (\text{term 2}) - (\text{term 3}) + (\text{term 4})]. \end{aligned}$$

901 Furthermore,  $[(\text{term 1}) - (\text{term 2}) - (\text{term 3}) + (\text{term 4})]$  is equal to

$$\begin{aligned} 902 \quad &\int \left( \mathbb{E}^{\mathbf{X}} f(x, Z) f(\tilde{x}_{-j}, Z_j) \right)^2 \\ 903 \quad &- \mathbb{E}^{\mathbf{X}} f(x, Z^{(1,1)}) f(x, Z^{(1,2)}) f(\tilde{x}_{-j}, Z_j^{(1,1)})^2 \\ 904 \quad &- \mathbb{E}^{\mathbf{X}} f(x, Z^{(1,1)})^2 f(\tilde{x}_{-j}, Z_j^{(1,1)}) f(\tilde{x}_{-j}, Z_j^{(1,2)}) \\ 905 \quad &+ \mathbb{E}^{\mathbf{X}} f(x, Z^{(1,1)})^2 f(\tilde{x}_{-j}, Z_j^{(1,1)})^2 dP(\mathbf{x}) \\ 906 \quad &= \int \left( \mathbb{E}^{\mathbf{X}} f(x, Z) \right)^2 \left( \mathbb{E}^{\mathbf{X}} f(\tilde{x}_{-j}, Z_j) \right)^2 \\ 907 \quad &- \left( \mathbb{E}^{\mathbf{X}} f(x, Z) \right)^2 \mathbb{E}^{\mathbf{X}} f(\tilde{x}_{-j}, Z_j)^2 \\ 908 \quad &- \mathbb{E}^{\mathbf{X}} f(x, Z)^2 \left( \mathbb{E}^{\mathbf{X}} f(\tilde{x}_{-j}, Z_j) \right)^2 \\ 909 \quad &+ \mathbb{E}^{\mathbf{X}} f(x, Z)^2 \mathbb{E}^{\mathbf{X}} f(\tilde{x}_{-j}, Z_j)^2 dP(\mathbf{x}) \\ 910 \quad &= \int \text{Var}^{\mathbf{X}} f(X, Z) \text{Var}^{\mathbf{X}} f(\tilde{X}_{-j}, Z_j) dP(\mathbf{x}). \\ 911 \end{aligned}$$

912 Likewise, we find that  $[(\text{term 2}) + (\text{term 3}) - 2(\text{term 1})]$  is equal to

$$913 \quad \mathbb{E} \text{Var}^{\mathbf{X}} f(X, Z) f(\tilde{X}_{-j}, Z_j) - \text{Var}^{\mathbf{X}} f(X, Z) \text{Var}^{\mathbf{X}} f(\tilde{X}_{-j}, Z_j),$$

915 and term 1 is  $\text{Var} \mathbb{E}^{\mathbf{X}} f(X, Z) f(\tilde{X}_{-j}, \tilde{Z}) + D_j^2$ .

916 We now deal with  $\xi_{m1}^{\text{LL}}$ . We have

$$\begin{aligned} 917 \quad \mathbb{E} \xi_{m1}^{\text{LL}2} &= \frac{1}{m^4} \sum_{k_1, k_2, k_3, k_4} \mathbb{E} f(X^{(1)}, Z^{(1, k_1)}) f(X^{(1)}, Z^{(1, k_2)}) \\ 918 \quad &f(X^{(1)}, Z^{(1, k_3)}) f(X^{(1)}, Z^{(1, k_4)}). \end{aligned}$$

921 The sum is split into five parts. The first part consists of the  $m(m-1)(m-2)(m-3)$   
 922 terms with different indices; those terms are equal to

$$923 \quad \mathbb{E} f(X^{(1)}, Z^{(1,1)}) f(X^{(1)}, Z^{(1,2)}) f(X^{(1)}, Z^{(1,3)}) f(X^{(1)}, Z^{(1,4)}).$$

924 The second part consists of the  $\binom{4}{2}m(m-1)(m-2)$  terms with exactly two equal  
 925 indices; those terms are equal to

$$926 \quad \mathbb{E} f(X^{(1)}, Z^{(1,1)})^2 f(X^{(1)}, Z^{(1,2)}) f(X^{(1)}, Z^{(1,3)}).$$

927 The third part consists of the  $\binom{4}{3}m(m-1)$  terms with exactly three equal indices;  
 928 those terms are equal to

$$929 \quad \mathbb{E} f(X^{(1)}, Z^{(1,1)})^3 f(X^{(1)}, Z^{(1,2)}).$$

930 The fourth part consists of the  $m$  terms with exactly four equal indices; those terms  
 931 are equal to

$$932 \quad \mathbb{E} f(X^{(1)}, Z^{(1,1)})^4.$$

933 The fifth and last part consists of the  $\binom{4}{2}m(m-1)/2$  terms with exactly two pairs of  
 934 equal indices; those terms are equal to

$$935 \quad \mathbb{E} f(X^{(1)}, Z^{(1,1)})^2 f(X^{(1)}, Z^{(1,2)})^2.$$

936 (One can check that the total number of terms is  $m^4$ .) □

937 LEMMA C.3 (Variances and covariances).

$$\begin{aligned}
 938 \quad (i) \quad & \text{Var } \xi_{m1}^{UL} = \text{Var } \mathbf{E}^{\mathbf{X}} f(X^{(1)}, Z^{(1,1)}) f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)}) \\
 939 & \quad + \frac{1}{m} [\mathbf{E} \text{Var}^{\mathbf{X}} f(X^{(1)}, Z^{(1,1)}) f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)}) \\
 940 & \quad - \text{Var}^{\mathbf{X}} f(X^{(1)}, Z^{(1,1)}) \text{Var}^{\mathbf{X}} f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)})] \\
 941 & \quad + \frac{1}{m^2} \mathbf{E} \text{Var}^{\mathbf{X}} f(X^{(1)}, Z^{(1,1)}) \text{Var}^{\mathbf{X}} f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)}), \\
 942 \quad (ii) \quad & \text{Cov}(\xi_{m1}^{UL}, \xi_{m1}^{UR}) = \frac{m-1}{m} \mathbf{E} f(X^{(1)}, Z^{(1,1)}) f(X^{(1)}, Z^{(1,2)}) f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)}) \\
 943 & \quad + \frac{1}{m} \mathbf{E} f(X^{(1)}, Z^{(1,1)})^2 f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)}) - D_j \mu \\
 944 \quad (iii) \quad & \text{Cov}(\xi_{m1}^{UL}, f(X, Z)^2) = \frac{1}{m} \mathbf{E} f(X^{(1)}, Z^{(1,1)})^3 f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)}) \\
 945 \quad (iii) \quad & \quad + \frac{m-1}{m} \mathbf{E} f(X^{(1)}, Z^{(1,1)})^2 f(X^{(1)}, Z^{(1,2)}) f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)}) - D_j \kappa \\
 946 \quad (iv) \quad & \text{Var } \xi_{m1}^{UR} = \frac{1}{m} \text{Var } f(X, Z) \\
 947 \quad (v) \quad & \text{Cov}(\xi_{m1}^{UR}, f(X, Z)^2) = \frac{1}{m} f(X, Z)^3 \\
 948 & \quad + \frac{m-1}{m} \mathbf{E} f(X^{(1)}, Z^{(1,1)})^2 f(X^{(1)}, Z^{(1,2)}) - \mu \kappa \\
 949 \quad (vi) \quad & \text{Cov}(\xi_{m_{n1}}^{UL}, \xi_{m_{n1}}^{ULL}) = \frac{m}{m^3} \mathbf{E} f(X^{(1)}, Z^{(1,1)})^3 f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)}) \\
 950 & \quad + \frac{3m(m-1)}{m^3} \mathbf{E} f(X^{(1)}, Z^{(1,1)})^2 f(X^{(1)}, Z^{(1,2)}) f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)}) \\
 951 & \quad + \frac{m(m-1)(m-2)}{m^3} \mathbf{E} f(X^{(1)}, Z^{(1,1)}) f(X^{(1)}, Z^{(1,2)}) \\
 952 & \quad f(X^{(1)}, Z^{(1,3)}) f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)}) \\
 953 & \quad - \mathbf{E} f(X^{(1)}, Z^{(1,1)}) f(\tilde{X}_{-j}^{(1)}, Z_j^{(1,1)}) \\
 954 & \quad \left\{ \frac{1}{m} \mathbf{E} f(X^{(1)}, Z^{(1,1)})^2 + \frac{m-1}{m} \mathbf{E} f(X^{(1)}, Z^{(1,1)}) f(X^{(1)}, Z^{(1,2)}) \right\} \\
 955 & \\
 956 &
 \end{aligned}$$

957 *Proof.* The proof follows from direct calculations.  $\square$

#### 958 Appendix D. Calculations for the linear model.

959 LEMMA D.1. Suppose that  $f(X, Z) = \beta_0 + \beta_{p+1}Z + \sum_{j=1}^p \beta_j X_j$  where  $X =$   
 960  $(X_1, \dots, X_p)$ ,  $Z_k, \tilde{Z}_{ik}$  are independent,  $\mathbf{E} X_j = \mathbf{E} Z = 0$ ,  $\mathbf{E} X_j^2 = \mathbf{E} Z^2 = 1$ ,  $\mathbf{E} X_j^3 = 0$ ,  
 961  $\mathbf{E} X_j^4 = 3$ . Then the squared optimal number of repetitions is given by

$$962 \quad (m_i^*)^2 = \frac{\beta_{p+1}^4}{(\beta_0 + \beta_i)^2 - 2\beta_0^4 + (\sum_{j=0}^p \beta_j^2)^2}$$

963 and the discriminator (the upper-left term in (2.8) and (2.9)) is

$$964 \quad \beta_0^2 + \beta_i^2.$$

965



966 *Proof.* We have

$$967 \quad m_i^* = \frac{A_i + B_i + C_i + D_i}{E_i},$$

968 with

$$\begin{aligned} 969 \quad A_i &= \mathbf{E} f(X, Z_1)^2 f(\tilde{X}_{-i}, \tilde{Z}_{i1})^2 \\ 970 \quad B_i &= \mathbf{E} f(X, Z_1) f(\tilde{X}_{-i}, \tilde{Z}_{i1}) f(X, Z_2) f(\tilde{X}_{-i}, \tilde{Z}_{i2}) \\ 971 \quad C_i &= -\mathbf{E} f(X, Z_1)^2 f(\tilde{X}_{-i}, \tilde{Z}_{i1}) f(\tilde{X}_{-i}, \tilde{Z}_{i2}) \\ 972 \quad D_i &= -\mathbf{E} f(\tilde{X}_{-i}, \tilde{Z}_{i1})^2 f(X, Z_1) f(X, Z_2) \\ 973 \quad E_i &= B - [\mathbf{E} f(X, Z_1) f(\tilde{X}_{-i}, \tilde{Z}_{i1})]^2 \end{aligned}$$

975 where  $X = (X_1, \dots, X_p)$ ,  $Z_k, \tilde{Z}_{ik}$  are independent,  $\mathbf{E} X_j = \mathbf{E} Z = 0$ ,  $\mathbf{E} X_j^2 = \mathbf{E} Z^2 = 1$ ,  
976  $\mathbf{E} X_j^3 = 0$ ,  $\mathbf{E} X_j^4 = 3$ . We deal with the case

$$977 \quad f(X, Z) = \beta_0 + \beta_{p+1} Z + \sum_{j=1}^p \beta_j X_j.$$

978 We calculate the terms one by one as follows. We have

$$\begin{aligned} 979 \quad A_j &= \mathbf{E} \left( \beta_0 + \sum_{j=1}^p \beta_j X_j \right)^2 \left( \beta_0 + \beta_i X_i + \sum_{j:1 \leq j \neq i} \beta_j \tilde{X}_j \right)^2 \\ 980 \quad &+ \left( \beta_0 + \sum_{j=1}^p \beta_j X_j \right)^2 \beta_{p+1}^2 \tilde{Z}_{i1}^2 + \beta_{p+1}^4 Z_1^2 \tilde{Z}_{i1}^2 \\ 981 \quad &+ \beta_{p+1}^2 Z_1^2 \left( \beta_0 + \beta_i X_i + \sum_{j:1 \leq j \neq i} \beta_j \tilde{X}_j \right)^2 \\ 982 \quad &= A_{j1} + A_{j2} + A_{j3}, \end{aligned}$$

984 where  $\mathbf{E}(A2) = \beta_{p+1}^4 + \beta_{p+1}^2 \sum_{j=0}^p \beta_j^2$ ,  $\mathbf{E}(A3) = \beta_{p+1}^2 \sum_{j=0}^p \beta_j^2$ . Elementary but some-  
985 what tedious calculations yield

$$986 \quad \mathbf{E}(A1) = \beta_0^4 + 3\beta_i^4 + 6\beta_0^2 \beta_i^2 + 2(\beta_0^2 + \beta_i^2) \sum_{j:1 \leq j \neq i} \beta_j^2 + \left( \sum_{j:1 \leq j \neq i} \beta_j^2 \right)^2.$$

988 Similar calculations show that  $B_j = A_{j1}$ ,  $C_j = -A_{j1} - A_{j3}$ ,  $D_j = -A_{j1} - A_{j3}$ ,  
989  $E_j = A_{j1} - (\beta_0^2 + \beta_i^2)^2$ . Thus,

$$990 \quad (m_i^*)^2 = \frac{\beta_{p+1}^4}{(\beta_0 + \beta_i)^2 - 2\beta_0^4 + (\sum_{j=0}^p \beta_j^2)^2}. \quad \square$$

991 **Acknowledgments.** The author thanks two anonymous referees who made use-  
992 ful suggestions. The author thanks Hervé Monod and Elisabeta Vergu for stimulating  
993 discussions from which this work arose and Bertrand Iooss for the references [5, 7].

994

## REFERENCES

- 995 [1] A. COURCOUL, H. MONOD, M. NIELEN, D. KLINKENBERG, L. HOGERWERF, F. BEAUDEAU, AND  
 996 E. VERGU, *Modelling the effect of heterogeneity of shedding on the within herd coxiella*  
 997 *burnetii spread and identification of key parameters by sensitivity analysis*, Journal of  
 998 Theoretical Biology, 284 (2011), pp. 130–141.
- 999 [2] D. J. DALEY AND J. GANI, *Epidemic Modelling*, Cambridge, 1999.
- 1000 [3] F. GAMBOA, A. JANON, T. KLEIN, AND A. LAGNOUX, *Sensitivity analysis for multidimensional*  
 1001 *and functional outputs*, Electron. J. Stat., 8 (2014), pp. 575–603, [https://doi.org/10.1214/](https://doi.org/10.1214/14-EJS895)  
 1002 [14-EJS895](https://doi.org/10.1214/14-EJS895).
- 1003 [4] F. GAMBOA, A. JANON, T. KLEIN, A. LAGNOUX, AND C. PRIEUR, *Statistical inference for Sobol*  
 1004 *pick-freeze Monte Carlo method*, Statistics, 50 (2016), pp. 881–902.
- 1005 [5] J. L. HART, A. ALEXANDERIAN, AND P. A. GREMAUD, *Efficient computation of Sobol’indices*  
 1006 *for stochastic models*, SIAM Journal on Scientific Computing, 39 (2017), pp. A1514–A1530.
- 1007 [6] T. HOMMA AND A. SALTELLI, *Importance measures in global sensitivity analysis of non-*  
 1008 *linear models*, Reliability Engineering & System Safety, 52 (1996), pp. 1–17, [https://doi.org/https://doi.org/10.1016/0951-8320\(96\)00002-6](https://doi.org/https://doi.org/10.1016/0951-8320(96)00002-6), <http://www.sciencedirect.com/science/article/pii/0951832096000026>.
- 1009 [7] B. IOOSS, L. L. GRATIET, A. LAGNOUX, AND T. KLEIN, *Sensitivity analysis for stochastic*  
 1010 *computer codes: Theory and estimation methods*, tech. report, EDF R&D, 2014.
- 1011 [8] B. IOOSS, T. KLEIN, AND A. LAGNOUX, *Sobol sensitivity analysis for stochastic numerical codes*,  
 1012 in 8th International Conference Sensitivity Analysis of Model Output, 2016, pp. 48–49.
- 1013 [9] B. IOOSS AND M. RIBATET, *Global sensitivity analysis of computer models with functional in-*  
 1014 *puts*, Reliability Engineering & System Safety, 94 (2009), pp. 1194 – 1204, <https://doi.org/https://doi.org/10.1016/j.res.2008.09.010>, <http://www.sciencedirect.com/science/article/pii/S0951832008002299>.
- 1015 [10] A. JANON, T. KLEIN, A. LAGNOUX, M. NODET, AND C. PRIEUR, *Asymptotic normality and*  
 1016 *efficiency of two Sobol index estimators*, ESAIM: Probability and Statistics, 18 (2014),  
 1017 pp. 342–364, <https://doi.org/https://doi.org/10.1051/ps/2013040>.
- 1018 [11] M. LAMBONI, H. MONOD, AND D. MAKOWSKI, *Multivariate sensitivity analysis to measure*  
 1019 *global contribution of input factors in dynamic models*, Reliability Engineering and System  
 1020 Safety, 96 (2010), pp. 450–459.
- 1021 [12] A. MARREL, B. IOOSS, S. DA VEIGA, AND M. RIBATET, *Global sensitivity analysis of stochastic*  
 1022 *computer models with joint metamodels*, Statistics and Computing, 22 (2012), pp. 833–847,  
 1023 <https://doi.org/10.1007/s11222-011-9274-8>, <https://doi.org/10.1007/s11222-011-9274-8>.
- 1024 [13] A. MARREL, B. IOOSS, B. LAURENT, AND O. ROUSTANT, *Calculations of sobol indices for*  
 1025 *the gaussian process metamodel*, Reliability Engineering & System Safety, 94 (2009),  
 1026 pp. 742 – 751, <https://doi.org/https://doi.org/10.1016/j.res.2008.07.008>, <http://www.sciencedirect.com/science/article/pii/S0951832008001981>.
- 1027 [14] H. MONOD, C. NAUD, AND D. MAKOWSKI, *Working with dynamic crop models: Evaluation,*  
 1028 *analysis, parameterization, and applications*, in Uncertainty and sensitivity analysis for  
 1029 crop models, Elsevier, 2006, pp. 55–100.
- 1030 [15] L. RIMBAUD, C. BRUCHOU, S. DALLOT, D. R. J. PLEYDELL, E. JACQUOT, S. SOUBEYRAND,  
 1031 AND G. THÉBAUD, *Using sensitivity analysis to identify key factors for the propaga-*  
 1032 *tion of a plant epidemic*, Open Science, 5 (2018), <https://doi.org/10.1098/rsos.171435>,  
 1033 <http://rsos.royalsocietypublishing.org/content/5/1/171435>, [https://arxiv.org/abs/http://](https://arxiv.org/abs/http://rsos.royalsocietypublishing.org/content/5/1/171435.full.pdf)  
 1034 [rsos.royalsocietypublishing.org/content/5/1/171435.full.pdf](https://arxiv.org/abs/http://rsos.royalsocietypublishing.org/content/5/1/171435.full.pdf).
- 1035 [16] A. SALTELLI AND I. M. SOBOL’, *About the use of rank transformation in sensitivity analysis*  
 1036 *of model output*, Reliability Engineering & System Safety, 50 (1995), pp. 225 – 239, [https://doi.org/https://doi.org/10.1016/0951-8320\(95\)00099-2](https://doi.org/https://doi.org/10.1016/0951-8320(95)00099-2), <http://www.sciencedirect.com/science/article/pii/0951832095000992>.
- 1037 [17] A. SALTELLI, S. TARANTOLA, AND F. CAMPOLONGO, *Sensitivity analysis as an ingredient of*  
 1038 *modeling*, Statistical Science, 15 (2000), pp. 377–395.
- 1039 [18] A. SALTELLI, S. TARANTOLA, F. CAMPOLONGO, AND M. RATTO, *Sensitivity analysis in practice*,  
 1040 Wiley, 2004.
- 1041 [19] J. F. SAVALL, C. BIDOT, M. LEBLANC-MARIDOR, C. BELLOC, AND S. TOUZEAU, *Modelling*  
 1042 *salmonella transmission among pigs from farm to slaughterhouse: interplay between man-*  
 1043 *agement variability and epidemiological uncertainty*, International Journal of Food Micro-  
 1044 biology, 229 (2016), pp. 33–43.
- 1045 [20] I. M. SOBOL, *Sensitivity estimates for nonlinear mathematical models*, Mathematical modelling  
 1046 and computational experiments, 1 (1993), pp. 407–414.
- 1047 [21] I. M. SOBOL, *Global sensitivity indices for nonlinear mathematical models and their Monte*

- 1055 *Carlo estimates*, Mathematics and computers in simulation, 55 (2001), pp. 271–280.
- 1056 [22] M. SPENCE, *Statistical issues in ecological simulation models*, PhD thesis, University of  
1057 Sheffield, 2015. <http://etheses.whiterose.ac.uk/10517>.
- 1058 [23] A. W. VAN DER VAART, *Asymptotic Statistics*, Cambridge University Press, 1998.