

# Variance-based sensitivity indices of computer models with dependent inputs: The Fourier Amplitude Sensitivity Test

Stefano Tarantola, Thierry A. Mara

► **To cite this version:**

Stefano Tarantola, Thierry A. Mara. Variance-based sensitivity indices of computer models with dependent inputs: The Fourier Amplitude Sensitivity Test. Accepted for publication in the International Journal of Uncertainty Quantification (BegellHouse). 2017. <hal-01568006>

**HAL Id: hal-01568006**

**<https://hal.archives-ouvertes.fr/hal-01568006>**

Submitted on 24 Jul 2017

**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.

# VARIANCE-BASED SENSITIVITY INDICES OF COMPUTER MODELS WITH DEPENDENT INPUTS: THE FOURIER AMPLITUDE SENSITIVITY TEST

*S. Tarantola<sup>1</sup> & T.A. Mara<sup>2,3,\*</sup>*

<sup>1</sup>*Directorate for Energy, Transport and Climate, Joint Research Centre, European Commission, Ispra (VA), Italy, 21027*

<sup>2</sup>*PIMENT, EA 4518, Université de La Réunion, FST, 15 Avenue René Cassin, 97715 Saint-Denis, Réunion*

<sup>3</sup>*Directorate for Modelling, Indicators and Impact Evaluation, Joint Research Centre, European Commission, Ispra (VA), Italy, 21027*

\*Address all correspondence to: T.A. Mara, PIMENT, EA 4518, Université de La Réunion, FST, 15 Avenue René Cassin, 97715 Saint-Denis, Réunion, E-mail: thierry.mara@univ-reunion.fr

*Several methods are proposed in the literature to perform the global sensitivity analysis of computer models with independent inputs. Only a few allow for treating the case of dependent inputs. In the present work, we investigate how to compute variance-based sensitivity indices with the Fourier amplitude sensitivity test. This can be achieved with the help of the inverse Rosenblatt transformation or the inverse Nataf transformation. We illustrate so on two distinct benchmarks. As compared to the recent Monte Carlo based approaches recently proposed by the same authors in [1], the new approaches allow to divide by two the computational effort to assess the entire set of first-order and total-order variance-based sensitivity indices.*

**KEY WORDS:** *Fourier amplitude sensitivity test; inverse Rosenblatt transformation; inverse Nataf transformation; variance-based sensitivity indices; dependent contributions; independent contributions*

## 2 1. INTRODUCTION

3 Good practice in computer model simulations requires that the uncertainties in the process under study be acknowl-  
4 edged. This is achieved by treating the model scalar inputs like random variables and functional inputs (temporally or

5 spatially dependent) like stochastic fields [2,3]. Subsequently, the model responses are also random and their uncer-  
6 tainties can be assessed, among others, via Monte Carlo simulations. Global Sensitivity Analysis (GSA) of computer  
7 model responses usually accompanies the uncertainty assessment. It aims to point out the set of input factors that  
8 mainly contributes to the model responses uncertainty. Such an information is essential to indicate the modellers  
9 what is the subset of inputs upon which they should concentrate their effort in future works in order to obtain more  
10 accurate and relevant predictions.

11 Several methods exist to perform GSA of model outputs when the inputs are independent [4]. The choice of the  
12 method to employ depends on the importance measure to be estimated. Two types of quantitative sensitivity mea-  
13 sures are of particular interest: the variance-based sensitivity measures [5] and the moment-independent sensitivity  
14 measures [6,7]. Numerical methods to assess these sensitivity measures can be classified as: non-parametric Monte  
15 Carlo approaches (among others, [8,9]), parametric spectral methods (likewise the Fourier amplitude sensitivity test  
16 [10,11]) or emulator-based approaches (e.g. [12,13]). Amongst the variance-based sensitivity measures, practitioners  
17 mostly focus on the first-order sensitivity index (also called correlation ratio [14,15]), that measures the marginal  
18 effect of one single input, and the total sensitivity index that accounts for the both marginal effects and interaction  
19 effects of the input of interest with the other ones [16]. The interested readers are referred to [17] for more details  
20 about the usefulness of these sensitivity measures in some GSA settings.

21 Performing sensitivity analysis of computer models with dependent inputs is more challenging. Just like models  
22 can be of different natures (linear, non-linear, additive and non-additive), the dependence structure amongst the inputs  
23 can be of different natures too (linear, non-linear, pairwise and non-pairwise). In case of linear dependence structure,  
24 the inputs are said correlated. But in many cases, the dependence structure can be more complex. The latter is  
25 embedded in the joint probability distribution of the inputs and in their copula.

26 The authors in [18] propose to distinguish the influence of an input due to its correlation with the other variables,  
27 and the influence that is not due to its possible correlations. The idea is that if an input is not influential on its own  
28 (without accounting for its correlations with the other inputs) then it can be concluded that it is a spurious input, only  
29 influential because of its correlations. This concept was later on extended to the variance-based sensitivity measures  
30 in [1,19] and to the density-based sensitivity measure in [20].

31 Kucherenko et al. [21] generalize the first-order and total Sobol' indices for the case of computer models with  
32 dependent inputs. A non-parametric Monte Carlo approach is proposed to evaluate them based on the theory of  
33 Gaussian copula. Simultaneously, in [19] four new variance-based sensitivity measures are defined: two that account  
34 for the dependencies of an individual input with the other ones (the authors called them the first-order and total full

35 sensitivity indices respectively) and two that do not account for the dependencies (called uncorrelated/independent  
36 first-order and total sensitivity indices respectively). Moreover, the authors show that the first-order sensitivity index  
37 generalized in [21] is the same as their *full first-order sensitivity index* (that accounts for correlations) while the  
38 generalized total sensitivity index of [21] is the same as their *independent total sensitivity index* (that does not account  
39 for correlations). Furthermore, two new variance-based indices namely the *full total sensitivity index* and *independent*  
40 *first-order sensitivity index* are also introduced in [19]. These four different sensitivity measures are discussed in the  
41 next Section.

42 In [1], the new variance-based sensitivity measures are formally defined (see Eqs.(1-4)) and two non-parametric  
43 methods are proposed to estimate them. The first method uses sampling strategies in conjunction with the inverse  
44 Rosenblatt transformation [22] while the second method employs the inverse Nataf transformation [23]. The latter  
45 corresponds to the procedure of Iman and Conover when the target dependence structure is the correlation matrix  
46 instead of the rank correlation matrix as in their original paper [24] and the Gaussian-copula-based approach em-  
47 ployed in [21] for generating correlated samples. In [19], the polynomial chaos expansion method is employed to  
48 evaluate the four variance-based sensitivity measures. The latter is very efficient because it only requires one single  
49 input/output random sample. But it also requires a procedure to make the input sample at hand independent. For  
50 this purpose, the authors derive a specific procedure only valid for some specific correlation structures (like pairwise  
51 linear and non-linear correlations).

52 Another idea is proposed in [25] that consists of distinguishing the *correlative* effect of a given input onto the  
53 computer model response from its *structural* effect. This can be achieved, for instance, by first identifying the model  
54 structure via an ANalysis Of VAriance decomposition in the Sobol' sense (for ANOVA see [8]). For this purpose  
55 independence of the inputs is mandatory. Then, the structural and correlative effects can be inferred by analyzing  
56 the covariance structure of the ANOVA decomposition stemming from the correlation structure amongst the inputs.  
57 Caniou names this approach the ANCOVA (acronym for ANalysis of COVAriance) decomposition and the author  
58 uses the polynomial chaos expansion for prior ANOVA decomposition [26]. There are other approaches proposed in  
59 the literature that are not discussed here (the interested reader can refer to [27,28], among others).

60 So far, no numerical approach has been proposed in the literature to compute the four variance-based sensitivity  
61 indices introduced in [19] with the Fourier Amplitude Sensitivity Test (FAST). One can cite [29] in which the author  
62 derives a cheap FAST-based approach to evaluate the independent first-order sensitivity index and the full first-order  
63 sensitivity index. One can also mention the early work of Xu and Gertner [30] that allows to assess the full first-order  
64 sensitivity index. Therefore, the proposed approaches are unable to account for interactions in computer models

65 with dependent inputs. The present work aims to fill this gap. To this end, we show that the extended FAST [11]  
 66 in conjunction with the inverse Rosenblatt transformation [22] or the inverse Nataf transformation [23] allows to  
 67 compute the four sensitivity indices.

68 The paper is organized as follows: we start by recalling the definitions of the variance-based sensitivity indices in  
 69 the case of dependent inputs in Section 2. The link with the law of total variance is made in Section 3. In Section 4, we  
 70 discuss the two transformations, namely, the inverse Rosenblatt transformation and the inverse Nataf transformation.  
 71 Section 5 recalls the classical FAST method for models with independent inputs and then its extension to the case of  
 72 dependent inputs is described. In Section 6, the new approaches are tested before concluding.

## 73 2. DEFINITION OF THE SENSITIVITY INDICES

74 Let  $f(\mathbf{x})$  be a square integrable function over an  $n$ -dimensional space where  $\mathbf{x} = \{x_1, \dots, x_n\}$  a continuous random  
 75 vector defined by a joint probability density function  $p(\mathbf{x})$ . The scalar  $f(\mathbf{x})$  can be regarded, without loss of gener-  
 76 ality, as the scalar response of a computer model to the input set. In the sequel, we set  $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)$  with  $\mathbf{x}_1$  and  
 77  $\mathbf{x}_2$  two non-empty subsets of  $\mathbf{x}$ . The importance of  $\mathbf{x}_1$  for  $f(\mathbf{x})$  can be measured with the variance-based sensitivity  
 78 indices (also called Sobol' indices,[5]). The Sobol' indices can either measure the amount of variance of  $f(\mathbf{x})$  due  
 79 to  $\mathbf{x}_1$  alone, or measure the amount of variance that also includes its interactions with  $\mathbf{x}_2$ . Besides, when  $\mathbf{x}_1$  and  $\mathbf{x}_2$   
 80 are dependent, it is possible to distinguish the two possible types of contribution of  $\mathbf{x}_1$  to the variance of  $f(\mathbf{x})$ : i) the  
 81 independent contribution that does not account for the dependence of  $\mathbf{x}_1$  with  $\mathbf{x}_2$  and, ii) the full contribution that  
 82 accounts for the dependence between  $\mathbf{x}_1$  and  $\mathbf{x}_2$ . To the authors' best knowledge, this concept was first introduced  
 83 in [18] although the partial correlation coefficient of [31] is related to this concept. The variance-based sensitivity  
 84 measures were defined in [19] for correlated variables and recently generalized in [1]. They are defined as follows:

$$S_{\mathbf{x}_1} = \frac{\mathbb{V}[\mathbb{E}[f(\mathbf{x})|\mathbf{x}_1]]}{\mathbb{V}[f(\mathbf{x})]}, \quad (1)$$

$$ST_{\mathbf{x}_1}^{ind} = \frac{\mathbb{E}[\mathbb{V}[f(\mathbf{x})|\mathbf{x}_2]]}{\mathbb{V}[f(\mathbf{x})]}, \quad (2)$$

$$S_{\mathbf{x}_1}^{ind} = \frac{\mathbb{V}[\mathbb{E}[f(\mathbf{x})|\bar{\mathbf{x}}_1]]}{\mathbb{V}[f(\mathbf{x})]}, \quad (3)$$

$$ST_{\mathbf{x}_1} = \frac{\mathbb{E}[\mathbb{V}[[f(\mathbf{x})|\bar{\mathbf{x}}_2]]]}{\mathbb{V}[f(\mathbf{x})]}, \quad (4)$$

85 where  $\mathbb{V}[\cdot]$  is the variance operator,  $\mathbb{E}[\cdot]$  is the mathematical expectation while  $\mathbb{V}[\cdot|\cdot]$  and  $\mathbb{E}[\cdot|\cdot]$  are the conditional  
 86 variance and expectation respectively.

87 The variables with an overbar are conditional variables, therefore:  $\bar{x}_1 \sim p_{x_1|x_2}(x_1|x_2)$  and  $\bar{x}_2 \sim p_{x_2|x_1}(x_2|x_1)$ .  
 88 While the first two sensitivity indices Eqs.(1-2) are the classical definitions of the Sobol' indices [21], the last two  
 89 are only defined for dependent input variables. All these indices are scaled within  $[0, 1]$  and we have  $S_{x_1} \leq ST_{x_1}$ ,  
 90  $S_{x_1}^{ind} \leq ST_{x_1}^{ind}$ .

91 The full first-order sensitivity index  $S_{x_1}$  measures the amount of variance of  $f(\mathbf{x})$  due to  $x_1$  and its dependence  
 92 with  $x_2$  but does not include the interactions of  $x_1$  with  $x_2$ . The full total sensitivity index  $ST_{x_1}$  does account for these  
 93 two types of contributions (dependence and interaction). The independent first-order sensitivity index  $S_{x_1}^{ind}$  measures  
 94 the contribution of  $x_1$  by ignoring its correlations and interactions with  $x_2$  while  $ST_{x_1}^{ind}$  accounts for interactions and  
 95 ignores correlations. An input  $x_i$  can contribute to the model response variance only because of its strong correlations  
 96 with the other inputs. In this case, we shall find  $ST_{x_i} \geq 0$  and  $ST_{x_i}^{ind} = 0$ .

97 The authors in [1] propose two non-parametric methods to evaluate these sensitivity indices. The first approach  
 98 consists of generating random samples of the dependent variables from the inverse Rosenblatt transformation while  
 99 the second one uses the sampling technique of Iman and Conover [24]. The first approach is advisable when the  
 100 conditional densities  $p_{x_1|x_2}$  and  $p_{x_2|x_1}$  are known while the procedure of Iman and Conover (IC) is to be preferred  
 101 when the marginal densities and the rank correlation structure of the input variables are known. We note that, when the  
 102 target dependence structure is the correlation matrix, the IC procedure is equivalent to the Gaussian copula approach  
 103 used in [21] and the inverse Nataf transformation described in the present work.

### 104 3. LINK WITH THE LAW OF TOTAL VARIANCE

105 It is usual to define the Sobol' indices from the law of total variance, namely

$$\mathbb{V}[f(\mathbf{x})] = \mathbb{V}[\mathbb{E}[f(\mathbf{x})|\mathbf{x}_1]] + \mathbb{E}[\mathbb{V}[f(\mathbf{x})|\mathbf{x}_1]]. \quad (5)$$

106 Dividing this equation by the left-hand side term yields,

$$1 = S_{x_1} + ST_{x_2}^{ind}. \quad (6)$$

107 In principal, Eq. (5) should be written as follows,

$$\mathbb{V}_{\mathbf{x}}[f(\mathbf{x})] = \mathbb{V}_{x_1}[\mathbb{E}_{\bar{x}_2}[f(\mathbf{x})|\mathbf{x}_1]] + \mathbb{E}_{x_1}[\mathbb{V}_{\bar{x}_2}[f(\mathbf{x})|\mathbf{x}_1]] \quad (7)$$

108 with the variables over which the conditional operators are applied indicated in subscript. But in the case of indepen-  
 109 dent variables, it is not necessary to indicate so and Eq. (5) (without subscript) is adopted for the sake of simplicity.  
 110 However, such a precision is necessary when the variables are dependent because of the axiom of conditional proba-  
 111 bilities:  $p_{\mathbf{x}}(\mathbf{x}) = p_{\mathbf{x}_1}(\mathbf{x}_1)p_{\mathbf{x}_2|\mathbf{x}_1}(\mathbf{x}_2|\mathbf{x}_1) = p_{\mathbf{x}_2}(\mathbf{x}_2)p_{\mathbf{x}_1|\mathbf{x}_2}(\mathbf{x}_1|\mathbf{x}_2)$ . In effect, setting  $\bar{\mathbf{x}}_1 = \mathbf{x}_1|\mathbf{x}_2$  and  $\bar{\mathbf{x}}_2 = \mathbf{x}_2|\mathbf{x}_1$ , one  
 112 can also write the law of total variance as follows,

$$\mathbb{V}_{\mathbf{x}} [f(\mathbf{x})] = \mathbb{V}_{\bar{\mathbf{x}}_1} [\mathbb{E}_{\mathbf{x}_2} [f(\mathbf{x})|\bar{\mathbf{x}}_1]] + \mathbb{E}_{\bar{\mathbf{x}}_1} [\mathbb{V}_{\mathbf{x}_2} [f(\mathbf{x})|\bar{\mathbf{x}}_1]]. \quad (8)$$

113 Normalizing the latter equation, yields,

$$1 = S_{\mathbf{x}_1}^{ind} + ST_{\mathbf{x}_2}. \quad (9)$$

114 The two different versions of the law of total variance hold because according to the axiom of conditional probabilit-  
 115 ities,  $\mathbf{x}_1$  and  $\bar{\mathbf{x}}_2$  (resp.  $\bar{\mathbf{x}}_1$  and  $\mathbf{x}_2$ ) are independent random vectors (i.e.  $p(\mathbf{x}) = p_{\mathbf{x}_1}(\mathbf{x}_1)p_{\bar{\mathbf{x}}_2}(\bar{\mathbf{x}}_2)$ ). Therefore, for  
 116 the sake of clarity, the definitions of the Sobol' indices in the case of dependent input variables should be written as  
 117 follows,

$$S_{\mathbf{x}_1} = \frac{\mathbb{V}_{\mathbf{x}_1} [\mathbb{E}_{\bar{\mathbf{x}}_2} [f(\mathbf{x})|\mathbf{x}_1]]}{\mathbb{V} [f(\mathbf{x})]}, \quad (10)$$

$$ST_{\mathbf{x}_1}^{ind} = \frac{\mathbb{E}_{\mathbf{x}_2} [\mathbb{V}_{\bar{\mathbf{x}}_1} [f(\mathbf{x})|\mathbf{x}_2]]}{\mathbb{V} [f(\mathbf{x})]}, \quad (11)$$

$$S_{\mathbf{x}_1}^{ind} = \frac{\mathbb{V}_{\bar{\mathbf{x}}_1} [\mathbb{E}_{\mathbf{x}_2} [f(\mathbf{x})|\bar{\mathbf{x}}_1]]}{\mathbb{V} [f(\mathbf{x})]}, \quad (12)$$

$$ST_{\mathbf{x}_1} = \frac{\mathbb{E}_{\bar{\mathbf{x}}_2} [\mathbb{V}_{\mathbf{x}_1} [f(\mathbf{x})|\bar{\mathbf{x}}_2]]}{\mathbb{V} [f(\mathbf{x})]}. \quad (13)$$

## 118 4. TWO PROBABILISTIC TRANSFORMATIONS

### 119 4.1 The inverse Rosenblatt transformation

120 It is shown in [1] that the Rosenblatt transformation [22] is the key for estimating the four sensitivity indices defined  
 121 in Eqs.(1-4). Indeed, the Inverse Rosenblatt Transformation (IRT) provides a set of dependent variables  $(\mathbf{x}_1, \mathbf{x}_2)$  from  
 122 a set of independent random vectors  $(\mathbf{u}_1, \mathbf{u}_2)$  uniformly distributed over the unit hypercube  $\mathbb{K}^n = [0, 1]^n$ . Assuming

123 that  $(\mathbf{x}_1, \mathbf{x}_2)$  is a vector of continuous random variables, the inverse Rosenblatt transformation writes,

$$\begin{cases} \mathbf{x}_2 = F_{\mathbf{x}_2}^{-1}(\mathbf{u}_2) \\ \bar{\mathbf{x}}_1 = F_{\mathbf{x}_1|\mathbf{x}_2}^{-1}(\mathbf{u}_1|\mathbf{u}_2) \end{cases} \quad (14)$$

124 where,  $F_{\mathbf{x}_2}^{-1}$  is the inverse cumulative density function of  $\mathbf{x}_2$  (that is,  $p_{\mathbf{x}_2} = dF_{\mathbf{x}_2}/d\mathbf{x}_2$ ) and  $F_{\mathbf{x}_1|\mathbf{x}_2}^{-1}$  is the one  
 125 of  $\bar{\mathbf{x}}_1$  (that is,  $p_{\mathbf{x}_1|\mathbf{x}_2} = \partial F_{\mathbf{x}_1|\mathbf{x}_2}/\partial \mathbf{x}_1$ ). IRT simply exploits the axiom of conditional probabilities:  $p(\mathbf{x}_1, \mathbf{x}_2) =$   
 126  $p_{\mathbf{x}_1|\mathbf{x}_2}(\mathbf{x}_1|\mathbf{x}_2)p_{\mathbf{x}_2}(\mathbf{x}_2)$  but is not unique. Indeed, as already aforementioned, one can also write  $p(\mathbf{x}_1, \mathbf{x}_2) = p_{\mathbf{x}_2|\mathbf{x}_1}(\mathbf{x}_2|\mathbf{x}_1)$   
 127  $p_{\mathbf{x}_1}(\mathbf{x}_1)$ , which yields the following transformation,

$$\begin{cases} \mathbf{x}_1 = F_{\mathbf{x}_1}^{-1}(\mathbf{u}_1) \\ \bar{\mathbf{x}}_2 = F_{\mathbf{x}_2|\mathbf{x}_1}^{-1}(\mathbf{u}_2|\mathbf{u}_1) \end{cases} \quad (15)$$

128 To generate the sets  $(\bar{\mathbf{x}}_1, \mathbf{x}_2)$  and  $(\mathbf{x}_1, \bar{\mathbf{x}}_2)$  from these two transformations, an independent set  $(\mathbf{u}_1, \mathbf{u}_2)$  uniformly  
 129 distributed over the unit hypercube is required. This is efficiently performed with, for instance the  $LP_\tau$  sequences of  
 130 [32]. Using  $(\bar{\mathbf{x}}_1, \mathbf{x}_2)$  and  $(\mathbf{x}_1, \bar{\mathbf{x}}_2)$ , one can estimate the sensitivity measures defined in Eqs.(1-4) as shown in [1].

#### 131 4.2 The inverse Nataf transformation

132 The application of IRT requires the knowledge of the conditional probability densities. In many situations, only the  
 133 individual cumulative densities (i.e.  $F_{x_j}(x_j), \forall j \in \llbracket 1, n \rrbracket$ ) and the correlation matrix  $\mathbf{R}_x$  are known. In this case, the  
 134 inverse Nataf transformation (INT) is more suitable than IRT to generate samples with correlation approaching the  
 135 desired matrix  $\mathbf{R}_x$ . It is worth mentioning that the procedure of Iman and Conover [24] and the Gaussian copula-  
 136 based approach employed in [21] are tantamount to INT. The latter uses a set of correlated standard normal variables  
 137  $\mathbf{z}^c = (z_1^c, \dots, z_n^c)$  with correlation matrix  $\mathbf{R}_z$  and generates the desired correlated random variables as follows,

$$x_j = F_{x_j}^{-1}(\Phi(z_j^c)), \forall j = 1, \dots, n \quad (16)$$

138 where  $\Phi$  is the cumulative density function of the standard normal variable. We note that transformation (16) implies  
 139 that  $x_j$  and  $z_j^c$  have identical ranking. Therefore, INT is equivalent to the procedure of [24]. However, INT and the  
 140 original IC procedure differ in the fact that the former generates samples of  $\mathbf{x}$  w.r.t. the correlation matrix  $\mathbf{R}_x$  whereas  
 141 the latter generates samples w.r.t. the rank correlation matrix of  $\mathbf{x}$ . Consequently, INT is a bit more complicated than



142 the original IC procedure. We note that Eq. (16) is also employed in the theory of Gaussian copula (see [21]).

143 Sampling  $\boldsymbol{x}$  with INT requires to generate  $\boldsymbol{z}^c$  with the desired correlation matrix  $\mathbf{R}_z$ . This is achieved with the  
 144 Cholesky transformation. Let us denote by  $\mathbf{L}$  the lower triangular matrix such that  $\mathbf{R}_z = \mathbf{L}\mathbf{L}^T$ , the superscript  $T$   
 145 stands for the transpose operator. This decomposition is possible because  $\mathbf{R}_z$  is positive-semidefinite. Then, from an  
 146 independent standard normal vector  $\boldsymbol{z}$ , the correlated standard normal vector  $\boldsymbol{z}^c$  is obtained as follows,

$$\boldsymbol{z}^c = \boldsymbol{z}\mathbf{L}^T. \quad (17)$$

147 The issue with this approach is to find  $\mathbf{R}_z$  such that  $\boldsymbol{x}$  has the desired correlation matrix  $\mathbf{R}_x$ . This can be achieved  
 148 with an optimization scheme in which  $\mathbf{R}_z$  is iteratively adjusted until the correlation matrix of  $\boldsymbol{x}$  satisfactorily matches  
 149  $\mathbf{R}_x$  [33]. The relationship between  $\mathbf{R}_z$  and  $\mathbf{R}_x$  is discussed for some densities in [34].

## 150 5. THE FOURIER AMPLITUDE SENSITIVITY TEST

### 151 5.1 The classical FAST for independent variables

152 FAST was introduced in [10] to compute the individual first-order sensitivity index for models with independent  
 153 inputs (Eq. (1) with  $x_1 = x_i$ ). In FAST input values are sampled over a periodic curve that explores the input space.  
 154 Each input is associated with a distinct integer frequency. The periodic sampled values are propagated through the  
 155 model. Then, the Fourier transform of the model output is computed. The Parseval-Plancherel theorem allows to  
 156 compute the variance-based sensitivity indices via the Fourier coefficients evaluated at specific frequencies. The  
 157 individual first-order sensitivity index of a given input uses the Fourier coefficients of the associated frequency and  
 158 its higher harmonics (see Algorithm 1).

159 The set  $\boldsymbol{\omega} = \{\omega_1, \dots, \omega_n\}$  of integer frequencies must be chosen in order to avoid interferences between  
 160 higher harmonics. This is possible up to a given interference order  $M$ . Cukier et al. [35] provides an algorithm to  
 161 generate such a set of frequencies for prescribed  $M$  and  $n$ . In practice,  $N$  draws of the input values are generated  
 162 by discretizing  $s$  as follows:  $s_k = \frac{2k\pi}{N}$ ,  $k = 1, \dots, N$ . With classical FAST, all first-order sensitivity indices can be  
 163 obtained with only one set of model runs  $N$ . The Nyquist criterion imposes that  $N \geq 2M \times \max(\boldsymbol{\omega}) + 1$ . Therefore,  
 164 the dimension of the model  $n$  and the choice of the interference factor  $M$ , considerably impact the number of model  
 165 runs  $N$  and also complicate the choice of an interference-free set of frequencies. To circumvent this problem, the  
 166 random balance design trick of [36] was extended to FAST in [37]. Algorithm 1 describes the steps to perform the  
 167 classical FAST approach.

**Algorithm 1:** The classical FAST procedure

1. Set  $u_j(s) = \frac{1}{2} + \frac{1}{\pi} \arcsin(\sin(\omega_j s + \varphi_j)) \forall j \in \llbracket 1, n \rrbracket$ , with  $s$  varying uniformly over  $(0, 2\pi]$ ,  $\varphi_j \in (0, 2\pi]$  is a randomly chosen shift parameter and  $\omega_j$  is an integer frequency
2. Perform the transformation  $x_j = F^{-1}(u_j), \forall j = 1, \dots, n$  to get  $x_j \sim p_{x_j}(x_j)$
3. Evaluate  $f(\mathbf{x}(s))$  and compute the Fourier coefficients,

$$c_\omega = \frac{1}{\pi} \int_0^{2\pi} f(\mathbf{x}(s)) e^{-i\omega s} ds, \forall \omega \in \mathbb{N}^* \quad (18)$$

4. Compute the first-order sensitivity indices,

$$S_{x_i}^{FAST} = \frac{\sum_{l=1}^{+\infty} |c_{l\omega_i}|^2}{\sum_{\omega=1}^{+\infty} |c_\omega|^2}, \forall i = 1, \dots, n. \quad (19)$$

168 To compute the total sensitivity index of  $x_i$ , that is  $ST_{x_i}$ , Saltelli et al. [11] propose to assign a high frequency to  
 169  $x_i$  (typically  $\omega_i = 2M \times \max(\omega_{\sim i})$ , where  $\omega_{\sim i} = \omega/\omega_i$ ) and small values to the other frequencies  $\omega_{\sim i}$ . These  
 170 latter do not need to be free of interferences although recommended. Consequently, in the Fourier spectrum of the  
 171 model response the amount of variance attributed to  $x_i$  (including its marginal effects and its interactions with the  
 172 other inputs) is localized in the high frequency range ( $\omega > M \times \max(\omega_{\sim i})$ ). The total sensitivity index is estimated  
 173 as follows:

$$\hat{S}T_{x_i} = \frac{\sum_{\omega=\frac{\omega_i}{2}+1}^{N/2} |\hat{c}_\omega|^2}{\sum_{\omega=1}^{N/2} |\hat{c}_\omega|^2} \quad (20)$$

174 with  $\hat{c}_\omega = \frac{2}{N} \sum_{k=1}^N f(\mathbf{x}(s_k)) e^{-i\omega s_k}$  an estimator of Eq. (18).

175 The drawback of the proposed approach (called EFAST) is that the computational effort to estimate  $(S_{x_i}, ST_{x_i})$   
 176 is high since the Nyquist criterion imposes that  $N > 2M\omega_i$ . But,  $S_{x_i}$  can be estimated simultaneously with  $ST_{x_i}$   
 177 (i.e. with no extra cost). Hence,  $n \times N$  model runs are necessary to compute all individual first-order and total  
 178 sensitivity indices. The author in [38] proposes a slight different version of EFAST that do not alleviate much the  
 179 computational burden, but allows for the calculation of the total sensitivity indices for groups of inputs.

## 180 5.2 EFAST and the inverse Rosenblatt transformation

181 Using EFAST and IRT it is possible to derive an algorithm to compute the variance-based sensitivity indices in the  
 182 case of dependent input variables (from now on, this procedure is named EFAST-IRT). Indeed, in the both EFAST  
 183 and IRT algorithms the vector  $\mathbf{x}$  is generated from uniformly and independently distributed variables. Consequently,

---

**Algorithm 2:** EFAST with the inverse Rosenblatt transformation (EFAST-IRT)
 

---

1. Set  $M$  (usually 4 or 6 but sometimes even 10 if we know a priori that the model has strong non linearities). Select a set of  $n - 1$  integer frequencies  $\omega_{\sim i}$  and infer  $\omega_i = 2M \max(\omega_{\sim i})$
2. Generate  $u_j = \frac{1}{2} + \frac{1}{\pi} \arcsin(\sin(\omega_j s + \varphi_j)), \forall j = 1, \dots, n$
3. Generate the vector  $\mathbf{x}$  from IRT (Eq. (15)) with  $\mathbf{x}_1 = x_i$  and  $\bar{\mathbf{x}}_2 = \mathbf{x}_{\sim i}$
4. Run the model and save the model responses of interest  $f(\mathbf{x})$
5. Compute the Fourier coefficients and deduce the variance-based sensitivity indices as follows,

$$\hat{S}_{x_i} = \frac{\sum_{l=1}^M |\hat{c}_{l\omega_i}|^2}{\sum_{\omega=1}^{N/2} |\hat{c}_\omega|^2} \quad (22a)$$

$$\hat{ST}_{x_i} = \frac{\sum_{\omega=\frac{\omega_i}{2}+1}^{N/2} |\hat{c}_\omega|^2}{\sum_{\omega=1}^{N/2} |\hat{c}_\omega|^2}. \quad (22b)$$


---

184 one can take advantage of the (deterministic) periodical sampling of EFAST to generate the dependent variables in  
 185 conjunction with IRT. The new algorithm to compute  $(S_{x_i}, ST_{x_i})$  is given by algorithm 2.

186 It can be noticed that one sample of size  $N > 2M\omega_i$  is necessary to compute both indices for a given  $x_i$  and  
 187  $n \times N$  model runs are necessary to compute all first order and total sensitivity indices. To evaluate the independent  
 188 sensitivity indices  $(S_{x_i}^{ind}, ST_{x_i}^{ind})$  one must proceed as just shown by operating the inverse Rosenblatt transformation  
 189 from Eq. (14) with  $\bar{\mathbf{x}}_1 = x_i$  and  $\mathbf{x}_2 = \mathbf{x}_{\sim i}$ . The sensitivity indices are estimated as previously, namely,

$$\hat{S}_{x_i}^{ind} = \frac{\sum_{l=1}^M |\hat{c}_{l\omega_i}|^2}{\sum_{\omega=1}^{N/2} |\hat{c}_\omega|^2} \quad (21a)$$

$$\hat{ST}_{x_i}^{ind} = \frac{\sum_{\omega=\frac{\omega_i}{2}+1}^{N/2} |\hat{c}_\omega|^2}{\sum_{\omega=1}^{N/2} |\hat{c}_\omega|^2}. \quad (21b)$$

190 Using samples of size  $N$ , an overall of  $2N \times n$  model runs are necessary to compute the four sensitivity indices  
 191  $(S_{x_i}, ST_{x_i}, S_{x_i}^{ind}, ST_{x_i}^{ind}), \forall i = 1, \dots, n$ . As compared to the non-parametric methods described in [1], which require  
 192  $4N \times n$  samples to compute the same set of sensitivity indices, the computational effort is halved.

### 193 5.3 EFAST and the inverse Nataf transformation

194 The algorithm to perform GSA via the inverse Nataf transformation (from now on named EFAST-INT) is more subtle  
 195 than with IRT. Indeed, the calculation of  $(S_{x_i}, ST_{x_i})$  or  $(S_{x_i}^{ind}, ST_{x_i}^{ind})$  depends on the position of  $z_i$  in the vector  $\mathbf{z}$

**Algorithm 3:** EFAST with the inverse Nataf transformation (EFAST-INT)

1. Choose a value for  $M$  and set  $i = 1$ . Select a set of  $n - 1$  integer frequencies  $\omega_{\sim i}$  and set  $\omega_i = 2M \max(\omega_{\sim i})$
2. Generate  $u_j = \frac{1}{2} + \frac{1}{\pi} \arcsin(\sin(\omega_j s + \varphi_j)), \forall j = 1, \dots, n$  with  $s$  regularly sampled over  $(0, 2\pi]$  using  $N$  points
3. Deduce the independent standard normal variables,  $z_j = \Phi^{-1}(u_j), \forall j = 1, \dots, n$ . Consider the following vector ordering  $\mathbf{z} = (z_i, \mathbf{z}_{\sim i})$  to estimate  $S_{x_i}$  and  $ST_{x_i}$  (resp.  $\mathbf{z} = (\mathbf{z}_{\sim i}, z_i)$  to estimate  $S_{x_i}^{ind}$  and  $ST_{x_i}^{ind}$ )
4. Set  $\mathbf{R}_z = \mathbf{R}_x$
5. Find  $\mathbf{L}$ , get  $\mathbf{z}^c$  from Eq. (17) and generate  $\mathbf{x}$  from Eq. (16)
6. Calculate the sample correlation matrix  $\hat{\mathbf{R}}_x$  from the sample  $\mathbf{x}$ . If  $\hat{\mathbf{R}}_x$  is not satisfactory, in the sense that it is not close enough to the desired correlation structure  $\mathbf{R}_x$ , modify  $\mathbf{R}_z$  and resume from 5, otherwise, continue
7. Run the model on  $\mathbf{x}$  and save the model responses of interest  $f(\mathbf{x})$ . Finally, compute the variance-based sensitivity indices  $(\hat{S}_{x_i}, \hat{ST}_{x_i})$  (resp.  $(\hat{S}_{x_i}^{ind}, \hat{ST}_{x_i}^{ind})$ ) as in Eq. (22) (resp. Eq. (21))
8. if  $i = n$  then stop. Otherwise, set  $i = i + 1$ ,  $\mathbf{R}_x = \mathbf{P}\mathbf{R}_z\mathbf{P}^T$  and resume from 3.

196 of standard normal variables used in the Cholesky transformation (Eq. (17)). As explained in [1], if we consider the  
 197 set  $(z_i, \mathbf{z}_{\sim i})$  and apply the Cholesky transformation,  $(S_{x_i}, ST_{x_i})$  can be computed. If the Cholesky transformation  
 198 is applied to the set  $(\mathbf{z}_{\sim i}, z_i)$ , then  $(S_{x_i}^{ind}, ST_{x_i}^{ind})$  can be obtained. In both cases, EFAST-INT assigns the highest  
 199 frequency  $\omega_i$  to  $z_i$ .

200 The procedure must also include an algorithm to find the optimal  $\mathbf{R}_z$  that produces a sample with the desired  
 201 correlation matrix  $\mathbf{R}_x$ . Given  $\mathbf{R}_x$  the correlation matrix and the marginal densities of each input variables, EFAST-  
 202 INT proceeds as in Algorithm 3 in which the following permutation matrix is employed,

$$\mathbf{P} = \begin{bmatrix} 0 & \mathbf{e}_{n-1}^T \\ \mathbf{e}_{n-1} & \mathbf{I}_{n-1} \end{bmatrix} \quad (23)$$

203 with  $\mathbf{e}_{n-1}^T = (0, \dots, 1)$  and  $\mathbf{I}_{n-1}$  the  $(n - 1) \times (n - 1)$  identity matrix.

204 The entire set of sensitivity indices  $(\hat{S}_{x_i}, \hat{ST}_{x_i}, \hat{S}_{x_i}^{ind}, \hat{ST}_{x_i}^{ind}), \forall i = 1 \dots, n$  are obtained with  $2n$  samples  
 205 of size  $N$  by considering circular permutations of the set  $\mathbf{z} = (z_1, \dots, z_n)$ . More specifically, considering the  
 206 set  $(z_1, z_2, \dots, z_n)$ , with one sample  $(\hat{S}_{x_1}, \hat{ST}_{x_1})$  can be evaluated by assigning the highest frequency to  $u_1$ , and  
 207 with another sample  $(\hat{S}_{x_n}^{ind}, \hat{ST}_{x_n}^{ind})$  are obtained by assigning the highest frequency to  $u_n$ . By considering the set  
 208  $(z_2, \dots, z_n, z_1)$  and the correlation matrix associated to  $(x_2, \dots, x_n, x_1)$ ,  $(\hat{S}_{x_2}, \hat{ST}_{x_2}, \hat{S}_{x_1}^{ind}, \hat{ST}_{x_1}^{ind})$  can be estimated

209 with other two samples, and so on. At the  $i$ -th iteration, step 8) transforms the correlation matrix  $\mathbf{R}_x$  so that the latter  
 210 corresponds to the correlation matrix of the circularly permuted vector  $(x_i, x_{i+1}, \dots, x_n, x_1, \dots, x_{i-1})$ .

## 211 6. NUMERICAL TEST CASES

### 212 6.1 Example with EFAST-IRT

213 To illustrate the EFAST-IRT approach, let us consider the following non-linear function:  $f(\mathbf{x}) = x_1x_2 + x_3x_4$  where  
 214  $(x_1, x_2) \in [0, 1]^2$  is uniformly distributed within the triangle  $x_1 + x_2 \leq 1$  and  $(x_3, x_4) \in [0, 1]^2$  is uniformly distributed  
 215 within the triangle  $x_3 + x_4 \geq 1$ . This function was studied in [1] with an non-parametric approach based on Quasi-  
 216 Monte Carlo sampling. The inverse Rosenblatt transformation that yields  $(x_1, x_2)$  in the domain  $x_1 + x_2 \leq 1$  from  
 217 the independent variables  $(u_1, u_2) \in [0, 1]^2$  is (see details in [1]),

$$\begin{cases} x_1 &= 1 - \sqrt{1 - u_1} \\ x_2 &= u_2 \sqrt{1 - u_1} \end{cases}. \quad (24)$$

218 Because of the symmetry, the IRT of  $(x_2, x_1)$  is obtained by simply inverting  $x_1$  and  $x_2$  in Eq. (24). In the same way,  
 219 the IRT of  $(x_3, x_4)$  writes,

$$\begin{cases} x_3 &= \sqrt{u_3} \\ x_4 &= (u_4 - 1)\sqrt{u_3} + 1 \end{cases}. \quad (25)$$

220 The symmetry of the problem implies that the sensitivity indices of  $x_1$  and  $x_2$  are equal, as well as those of  $x_3$   
 221 and  $x_4$ . We have computed 100 replicate estimates of these indices with the EFAST-IRT approach. This has been  
 222 achieved by randomly drawing the four shift parameters  $\{\varphi_i, i = 1, \dots, 4\}$  at each replication. We imposed a sample  
 223 size  $N = 4\,095$ , which corresponded to a total number of function calls per replicate equal to  $4 \times 4,095 = 16\,380$ ,  
 224 and we selected  $M = 8$ . Consequently,  $\omega_i = (N - 1)/M = 255$ . Finally, we chose the following set of low  
 225 frequencies  $\boldsymbol{\omega}_{\sim i} = (5, 9, 14)$ , although other choices were equally suitable.

226 Fig. 1 depicts a few draws obtained with EFAST-IRT. We note that the input values are sampled accordingly with  
 227 the desired constraints. The sample fills the input space quite well. Tab. 1 shows good agreement between the mean  
 228 estimates of the sensitivity indices and the analytical values. The results indicate that  $(x_3, x_4)$  are the most relevant  
 229 inputs for the response variance.

230 By referring to the work of [1], it can be inferred that one possible ANOVA decomposition (in the Sobol' sense

231 [5]) of  $f(\mathbf{x})$  is:

$$f(\mathbf{x}) = f_0 + f_1(x_1) + f_2(\bar{x}_2) + f_{12}(x_1, \bar{x}_2) + f_3(x_3) + f_4(\bar{x}_4) + f_{34}(x_3, \bar{x}_4) \quad (26)$$

232 where the functions in Eq. (26) have the same properties than those of the Sobol's ANOVA decomposition [5]. In  
 233 particular, they are orthogonal. This property of orthogonality allows to cast the variance of  $f(\mathbf{x})$ , denoted  $V$ , as  
 234 follows:

$$V = V_1 + V_2^{ind} + V_{12} + V_3 + V_4^{ind} + V_{34} \quad (27)$$

235 with  $V_i = E[f_i^2(x_i)]$ ,  $V_i^{ind} = E[f_i^2(\bar{x}_i)]$  and  $V_{ij} = E[f_{ij}^2(x_i, \bar{x}_j)]$ . By denoting  $VT_i = V_i + V_{ij}$  yields the following  
 236 variance decomposition,

$$V_y = VT_1 + V_2^{ind} + VT_3 + V_4^{ind}. \quad (28)$$

The normalization of the latter equation by  $V$  yields the following relationship between the variance-based sensitivity indices,

$$ST_{x_1} + S_{x_2}^{ind} + ST_{x_3} + S_{x_4}^{ind} = 1.$$

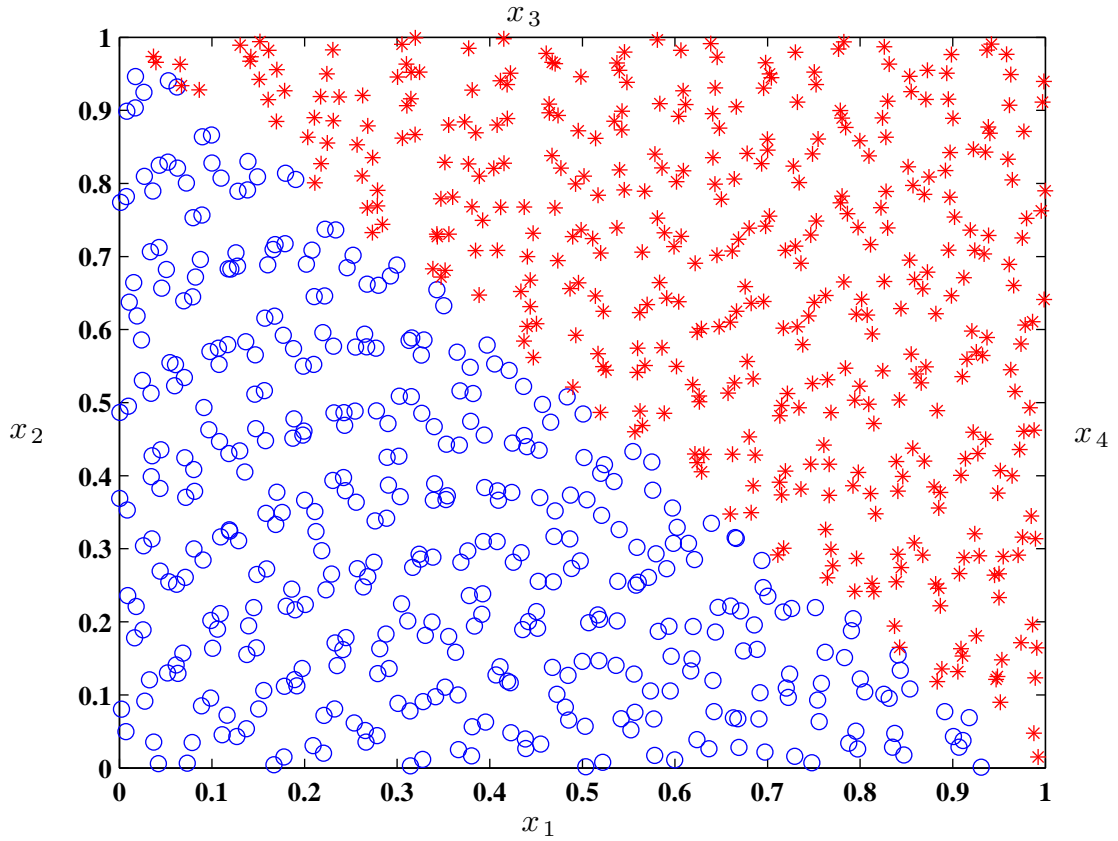
237 The analytical variance-based sensitivity indices reported in Tab. 1 satisfy this relationship. Moreover, we can infer  
 238 that  $ST_{x_1} + S_{x_2}^{ind} = 0.10$ , which indicates that the pair  $(x_1, x_2)$  explains 10% of the response variance since the pairs  
 239  $(x_1, x_2)$  and  $(x_3, x_4)$  are independent.

**TABLE 1:** One hundred replicate estimates of the sensitivity indices computed with EFAST-IRT.

		$S_{x_1}$	$ST_{x_1}$	$S_{x_2}^{ind}$	$ST_{x_2}^{ind}$	$S_{x_3}$	$ST_{x_3}$	$S_{x_4}^{ind}$	$ST_{x_4}^{ind}$
	Analytical	0.033	0.044	0.056	0.067	0.233	0.400	0.500	0.666
EFAST-IRT	2.5 <sup>th</sup> perc.	0.028	0.043	0.051	0.066	0.209	0.361	0.475	0.639
	mean	0.032	0.048	0.055	0.071	0.226	0.395	0.497	0.669
	97.5 <sup>th</sup> perc.	0.037	0.056	0.061	0.077	0.248	0.448	0.523	0.705

## 240 6.2 The Ishigami function with EFAST-INT

241 The Ishigami function is one of the benchmarks model for assessing the efficiency of GSA methods [39]. It has been  
 242 intensively used by statisticians to test their sensitivity analysis approaches in the case of independent inputs (e.g.  
 243 [40,41], among others). Recently, the authors in [21] introduced a new method for computing  $S_{x_i}$  and  $ST_{x_i}^{ind}$  in the  
 244 case of models with correlated inputs and tested it on the Ishigami function. We repeat here the same example with



**FIG. 1:** Some draws obtained with FAST after the inverse Rosenblatt transformation. We have  $x_1 + x_2 \leq 1$  (blue circles) and  $x_3 + x_4 \geq 1$  (red stars).

245 EFAST-INT but compute  $(S_{x_i}, ST_{x_i}^{ind}, S_{x_i}^{ind}, ST_{x_i}), \forall i = 1, 2, 3$ . The Ishigami function writes,

$$f(\mathbf{x}) = \sin x_1 + 7 \sin^2 x_2 + 0.1 x_3^4 \sin x_1 \quad (29)$$

246 with input variables being uniformly distributed:  $-\pi \leq x_i \leq \pi, i = 1, 2, 3$ . Note that this function is non-linear in all  
 247 inputs and  $x_2$  does not interact with the other two variables. Although this function has a low dimension (three inputs  
 248 only), it is a challenging function for EFAST because the last term, which represents the interaction between  $x_1$  and  
 249  $x_3$ , is strongly non-linear and its Fourier coefficients have considerable amplitudes at high frequencies. This strong  
 250 non-linearity led us to choose a considerably high value of  $M$ . The difficulty of the test was emphasized because a  
 251 correlation  $r_{13} \in ] - 1, 1[$  was imposed between  $x_1$  and  $x_3$ .

252 Likewise the previous exercise, 100 replication estimates of the sensitivity indices were performed. The design

253 of the EFAST-INT was the following:  $M = 15$ ,  $N = 8\,191$ ,  $\omega_{\sim i} = (5, 8)$  and  $\omega_i = 273$ . We chose higher  $N$  here  
 254 than in the previous example because of the strong non-linearities mentioned above. Following the work of [21], the  
 255 exercise was conducted by varying the correlation coefficient  $r_{13}$  within  $]-1, 1[$ .

256 The results are depicted in Fig. 2. To our best knowledge, no analytical sensitivity indices are available for this  
 257 test function. We can infer that, because  $x_2$  is not correlated to the other two inputs, its full and independent sensitivity  
 258 indices are equal. Besides, because  $x_2$  does not interact with the other two inputs, its total and first-order indices are  
 259 also equal. Moreover, we note that the full index  $S_{x_i}$  is always greater or equal to the independent index  $S_{x_i}^{ind}$  (resp.  
 260  $ST_{x_i} \geq ST_{x_i}^{ind}$ ). However, we may also find  $S_{x_i}^{ind} > S_{x_i}$  (or  $ST_{x_i}^{ind} > ST_{x_i}$ ) as also shown in the previous exercise  
 261 (see Tab. 1).

262 As far as  $x_1$  and  $x_3$  are concerned we note that, when the correlation coefficient is zero, as expected the full and  
 263 independent sensitivity indices are equal. When  $r_{13}$  is close to  $\pm 1$ , the independent sensitivity indices (both first and  
 264 total order) of  $x_1$  and  $x_3$  are null. This makes sense because if  $r_{13} = \pm 1$ , then all the information in  $f(x)$  is captured  
 265 by only one of the pairs  $(x_1, x_2)$  or  $(x_2, x_3)$ . Indeed, in this case,  $x_1$  and  $x_3$  contain the same information and it is  
 266 not possible to distinguish their individual contributions in the model response. This finding is peculiarly important  
 267 for the modeller as it indicates that the output uncertainty is explained by one of these two pairs only, thus, allowing  
 268 some kind of dimensionality reduction. Indeed, the modeller now knows that to obtain narrower output uncertainty,  
 269 he/she should pay some further effort to reduce the uncertainty either in the pair  $(x_1, x_2)$  or in  $(x_2, x_3)$ .

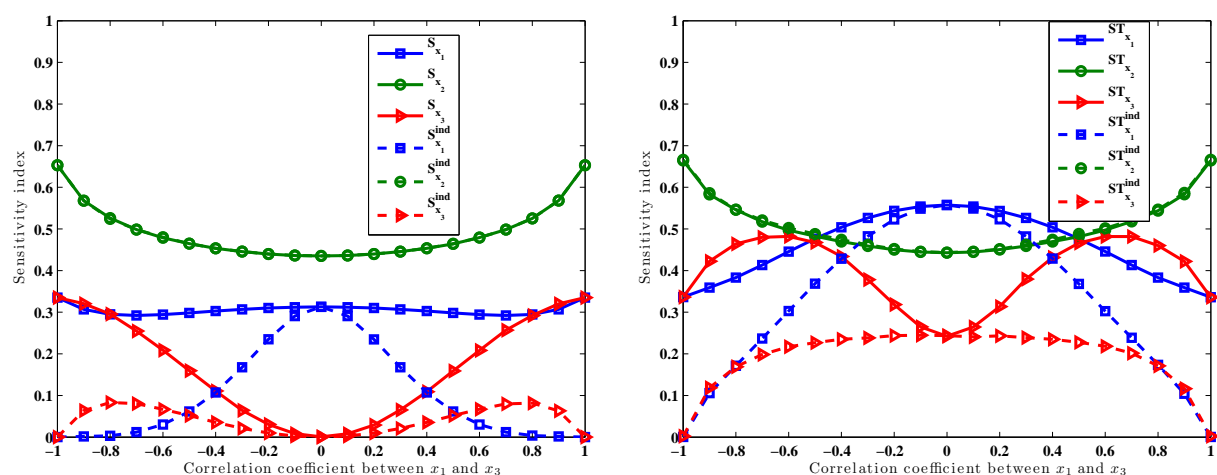


FIG. 2: The variance-based sensitivity indices estimated with EFAST-INT versus the correlation coefficient between  $x_1$  and  $x_3$ .



## 270 7. CONCLUSION

271 Performing global sensitivity analysis of model output with dependent inputs is a challenging issue. Variance-based  
272 sensitivity indices have been defined in [1,19,21] and different approaches have been proposed to estimate them. Four  
273 types of sensitivity indices can be of interest: (i) the full first-order (resp. full total) sensitivity index that measures  
274 the amount of model response variance explained by an input factor which takes into account its dependence with the  
275 other inputs; (ii) the full total-order sensitivity index that measures the amount of model response variance explained  
276 by an input factor which takes into account the both its dependence and interactions with the other inputs; (iii)  
277 the independent first-order sensitivity index of an input that measures its relative contribution alone to the response  
278 variance by ignoring its dependence with the other inputs; and finally (iv) the independent total-order sensitivity index  
279 of an input that measures its relative contribution to the response variance by ignoring its dependence with the other  
280 input variables but by accounting for its interactions with the latter.

281 The Fourier amplitude sensitivity test is one of the first methods for variance-based global sensitivity analysis  
282 [10]. Since that, the method has been extended by several authors [11,29,30,37,38]. Specifically, in [30] FAST has  
283 been adapted to account for correlations among inputs by using the sampling technique of Iman and Conover [24].  
284 In this work, we extend FAST to compute the four sensitivity indices defined above. The main idea of our approach  
285 is to impose either a dependence structure amongst the inputs with the inverse Rosenblatt transformation [22], with  
286 Algorithm 2 denoted EFAST-IRT, or a correlation structure with the inverse Nataf transformation [23], denoted  
287 EFAST-INT (see Algorithm 3). The numerical tests shown in the paper confirm the suitability of both EFAST-IRT  
288 and EFAST-INT.

289 The sampling strategy proposed by [21] allows for estimating the overall full first-order sensitivity indices and  
290 the independent total sensitivity indices with  $(2n+2)$  (quasi) Monte Carlo samples. The sampling strategies proposed  
291 in [1] allows for assessing the four sensitivity indices of all the input variables with  $4n$  (quasi) Monte Carlo samples.  
292 In the present work, we show that  $2n$  samples are sufficient to compute the four sensitivity indices of all the inputs  
293 with either EFAST-IRT or EFAST-INT.

## 294 ACKNOWLEDGMENTS

295 The corresponding author would like to thank the French National Research Agency for its financial support (Research  
296 project RESAIN n° ANR-12-BS06-0010-02).

297 **REFERENCES**

- 298 1. Mara, T.A., Tarantola, S., and Annoni, P., Non-parametric methods for global sensitivity analysis of model output with  
299 dependent inputs, *Environmental Modelling and Software*, 72:173–183, 2015.
- 300 2. Blatman, G. and Sudret, B., Adaptive sparse polynomial chaos expansion based on least angle regression, *Journal of Compu-*  
301 *tational Physics*, 230(6):2345–2367, 2011.
- 302 3. Anstett-Collin, F., Goffart, J., Mara, T., and Denis-Vidal, L., Sensitivity analysis of complex models: Coping with dynamic  
303 and static inputs, *Reliability Engineering and System Safety*, 134:268–275, 2015.
- 304 4. Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S., *Global Sensitivity*  
305 *Analysis: The Primer*, Probability and Statistics, John Wiley and Sons, Chichester, 2008.
- 306 5. Sobol', I.M., Sensitivity estimates for nonlinear mathematical models, *Math. Mod. and Comput. Exp.*, 1:407–414, 1993.
- 307 6. Chun, M.H., Han, S.J., and Tak, N.I., An uncertainty importance measure using distance metric for the change in a cumulative  
308 distribution function, *Reliability Engineering and System Safety*, 70:313–321, 2000.
- 309 7. Borgonovo, E., Measuring uncertainty importance: Investigation and comparison of alternative approaches, *Risk Analysis*,  
310 26(5):1349–1361, 2006.
- 311 8. Sobol', I.M., Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates, *Mathematics and*  
312 *Computers in Simulation*, 55:271–280, 2001.
- 313 9. Saltelli, A., Making best use of model evaluations to compute sensitivity indices, *Computational Physics Communications*,  
314 145:280–297, 2002.
- 315 10. Cukier, R.I., Fortuin, C.M., Shuler, K.E., Petschek, A.G., and Schaibly, J.H., Study of the sensitivity of coupled reaction  
316 systems to uncertainties in rate coefficients. I. theory, *J. Chemical Physics*, 59:3873–3878, 1973.
- 317 11. Saltelli, A., Tarantola, S., and Chan, K., A quantitative model independent method for global sensitivity analysis of model  
318 output, *Technometrics*, 41:39–56, 1999.
- 319 12. Oakley, J.E. and O'Hagan, A., Probabilistic sensitivity analysis of complex models: a Bayesian approach, *J. Royal Statist.*  
320 *Soc. B*, 66:751–769, 2004.
- 321 13. Ratto, M., Pagano, A., and Young, P., State dependent parameter metamodelling and sensitivity analysis, *Computer Physics*  
322 *Communications*, 117(11):863–876, 2007.
- 323 14. Pearson, K. On the general theory of skew correlation and non-linear regression. In *Mathematical contributions to the theory*  
324 *of evolution*, Vol. XIV. Drapers's Company Research Memoirs, 1905.
- 325 15. McKay, M.D., Variance-based methods for assessing uncertainty importance, Tech. Rep., Technical Report NUREG-  
326 1150,UR-1996-2695, Los Alamos National Laboratory, 1996.

- 327 16. Homma, T. and Saltelli, A., Importance measures in global sensitivity analysis of nonlinear models, *Reliability Engineering*  
328 *and System Safety*, 52:1–17, 1996.
- 329 17. Saltelli, A. and Tarantola, S., On the relative importance of input factors in mathematical models: Safety assessment for  
330 nuclear waste disposal, *Journal of the American Statistical Association*, 97:702–709, 2002.
- 331 18. Xu, C. and Gertner, G.Z., Uncertainty and sensitivity analysis for models with correlated parameters, *Reliability Engineering*  
332 *and System Safety*, 93:1563–1573, 2008b.
- 333 19. Mara, T.A. and Tarantola, S., Variance-based sensitivity indices for models with dependent inputs, *Reliability Engineering*  
334 *and System Safety*, 107:115–121, 2012.
- 335 20. Zhou, C., Lu, Z., Zhang, L., and Hu, J., Moment independent sensitivity analysis with correlations, *Applied Mathematical*  
336 *Modelling*, 38:4885–4896, 2014.
- 337 21. Kucherenko, S., Tarantola, S., and Annoni, P., Estimation of global sensitivity indices for models with dependent variables,  
338 *Computer Physics Communications*, 183:937–946, 2012.
- 339 22. Rosenblatt, M., Remarks on the multivariate transformation, *Annals of Mathematics and Statistics*, 43:470–472, 1952.
- 340 23. Nataf, A., Détermination des distributions dont les marges sont données, *Comptes Rendus de l'Académie des Sciences*,  
341 225:42–43, 1962.
- 342 24. Iman, R.I. and Conover, W.J., A distribution-free approach to inducing rank correlation among input variables, *Communica-*  
343 *tions in Statistics: Simulation and Computation*, 11:311–334, 1982.
- 344 25. Li, G., Rabitz, H., Yelvington, P.E., Oluwole, O.O., Bacon, F., Kolb, C.E., and Schoendorf, J., Global sensitivity analysis for  
345 systems with independent and/or correlated inputs, *J. Physical Chemistry*, 114:6022–6032, 2010.
- 346 26. Caniou, Y., Analyse de sensibilité globale pour les modèles imbriqués et multiéchelles, PhD thesis, University Blaise Pascal,  
347 Clermont-Ferrand, 2012.
- 348 27. Chastaing, G., Gamboa, F., and Prieur, C., Generalized Hoeffding-Sobol decomposition for dependent variables – application  
349 to sensitivity analysis, *Electronic Journal of Statistics*, 6:2420–2448, 2012.
- 350 28. Da Veiga, S., Wahl, F., and Gamboa, F., Local polynomial estimation for sensitivity analysis of models with correlated inputs,  
351 *Technometrics*, 51(4):452–463, 2009.
- 352 29. Xu, C., Decoupling correlated and uncorrelated uncertainty contributions for nonlinear models, *Applied Mathematical Mod-*  
353 *elling*, 37:9950–9969, 2013.
- 354 30. Xu, C. and Gertner, G.Z., A general first-order global sensitivity analysis method, *Reliability Engineering and System Safety*,  
355 93:1060–1071, 2008a.
- 356 31. Fisher, R.A., The distribution of the partial correlation coefficient, *Metron*, pp. 329–332, 1924.

- 357 32. Sobol', I.M., Turchaninov, V.I., Levitan, Y.L., and Shukman, B.V. *Quasi-random sequence generator (routine LPTAU51)*.  
358 Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, 1992.
- 359 33. Li, H., Lu, Z., and Yuan, X., Nataf transformation based point estimate method, *Chinese Science Bulletin*, 53(17):2586–2592,  
360 2008.
- 361 34. Ditlevsen, O. and Madsen, H.O., *Structural Reliability Methods*, John Wiley and Sons, New York, 1996.
- 362 35. Cukier, R.I., Schaibly, J.H., and Shuler, K.E., Study of the sensitivity of coupled reaction systems to uncertainties in rate  
363 coefficients. iii. analysis of the approximations, *J. Chemical Physics*, 63:1140–1149, 1975.
- 364 36. Satterthwaite, F.E., Random balance experimentation, *Technometrics*, 1:111–137, 1959.
- 365 37. Tarantola, S., Gatelli, D., and Mara, T.A., Random balance designs for the estimation of first-order global sensitivity indices,  
366 *Reliability Engineering and System Safety*, 91:717–727, 2006.
- 367 38. Mara, T.A., Extension of the rbd-fast method to the computation of global sensitivity indices, *Reliability Engineering and*  
368 *System Safety*, 94:1274–1281, 2009.
- 369 39. Saltelli, A., Chan, K., and Scott, E.M., *Sensitivity analysis*, John Wiley and Sons, Chichester, 2000.
- 370 40. Sudret, B., Global sensitivity analysis using polynomial chaos expansions, *Reliability Engineering and System Safety*, 93:964–  
371 979, 2008.
- 372 41. Marrel, A., Iooss, B., Laurent, B., and Roustant, O., Calculations of Sobol indices for the Gaussian process metamodel,  
373 *Reliability Engineering and System Safety*, 94:742–751, 2009.