An inference method for global sensitivity analysis

Gildas Mazo, Laurent Tournier

An inference method for global sensitivity analysis

Gildas Mazo* Laurent Tournier*

September 7, 2023

Abstract

Although there is a plethora of methods to estimate sensitivity indices associated with individual inputs, there is much less work on interaction effects of every order, especially when it comes to make inferences about the true underlying values of the indices. To fill this gap, a method that allows one to make such inferences simultaneously from a single Monte Carlo sample is given. One advantage of this method is its simplicity: one leverages the fact that Shapley and Sobol indices are only linear transformations of total indices, so that standard asymptotic theory suffices to get confidence intervals, p-values, and the possibility of carrying out statistical tests. To do the numerical computations, Möbius inversion formulas are used and linked to the fast Möbius transform algorithm. An illustration on a model in cellular biology involving 12 inputs and 4096 input combinations and interaction effects is given.

Keywords: Möbius inversion, Möbius transform, Sobol index, global sensitivity analysis, inference

1 Introduction

Sensitivity analysis seeks to quantify, for a given mathematical or computer model f, the extent to which a change in the inputs contributes to a change in the output. Contrary to local methods, which quantify rates of change at particular points in the input space, global sensitivity analysis assesses the inputs' contributions over the whole input space [34, 35], often with statistical methods. One then adopts a probabilistic framework in which the inputs $X = (X_1, \ldots, X_d)$ are seen as random variables. From now on, by sensitivity analysis we shall mean global sensitivity analysis.

Different aspects of sensitivity analysis are captured by different kinds of sensitivity indices. The Sobol index $\sigma^*(A)$, $A \subset \{1, \ldots, d\} =: D$, quantifies the interaction effect of order $|A|$ between the components of X_A, the subvector of X corresponding to the subset A. (See, e.g. [35].) In the special case $A = \{j\}$, $j = 1, \ldots, d$, the quantity $\sigma^*(\{j\}) = \operatorname{Var} \mathrm{E}(f(X)|X_j)$, called the first-order Sobol index of X_j, is interpreted as the expected variance reduction should X_j be fixed [15, 20]. The total index $\tau(A)$ is defined as the sum over all interaction

*Université Paris-Saclay, INRAE, MaIAGE, 78350, Jouy-en-Josas, France

effects $\sigma^*(B)$, $B \cap A \neq \emptyset$, or, equivalently, as the output variability induced by a random change in X_A, averaged over the remaining inputs $X_{D\setminus A}$. Shapley indices stem from the application of an allocation method coming from game theory [37, 45]. One assigns a "value" to each of the players, based on the performance of all the possible coalitions among them. By analogy, the Shapley index $\theta(\{j\})$ of the individual input X_j summarizes its global importance, based on all the possible input combinations.

Estimation of the above sensitivity indices has attracted much attention in the literature. In particular, a plethora of methods have been proposed to estimate first-order Sobol indices. Many methods rest on the fact that $\operatorname{Var} \mathrm{E}(f(X) - \mathrm{E}[f(X)]|X_j)$ can be re-expressed as a simple expectation and use the method-of-moments to estimate it from a Monte Carlo or quasi-Monte Carlo sample of inputs [38, 19, 18, 8, 14, 32, 33, 27]. Asymptotic normality of some estimators is studied in [8, 18]. To avoid the waste of data, more elaborate ways of combining outputs have been proposed [42, 10, 33, 32], the goal being always to have the most efficient estimates with the least model runs [29]. See also [36] for another way of exploring the input space. Methods based on U-statistics are found in [9]. There are also methods that offer a different look, statistically speaking. For instance, nonparametric regression techniques, or techniques based on kernels, are found in [5, 4, 39]. The use of meta-models to compute sensitivity indices has been proposed as well [41]. It is a fact that the estimation of first-order Sobol indices is an active area of research, and many other methods have been, and continue to be, proposed. See, e.g. the books [31, 5] for more details and possibly other estimation methods. A list of methods is also available in [17]. A numerical comparison of some of the above methods is given in [24].

Although a plethora of methods exist to estimate first-order indices, the literature is more thin about the estimation of higher order interaction effects. Attention has increased recently, though, as they are useful to compute Shapley indices [3, 16, 30, 40, 11]. The full consideration of the possibility of making inferences about the true underlying values for each of the interaction effects based on sound statistical theory (e.g. confidence intervals, p-values, hypothesis testing) is still lacking, however.

A method to make inferences about the true underlying values of the sensitivity indices of every order simultaneously from a single Monte Carlo sample is given. The main advantage of this method is its simplicity: one leverages the fact that Shapley and Sobol indices are only linear transformations of total indices, so that the standard central limit theorem and delta method suffice to get asymptotic confidence intervals, p-values, and the possibility of carrying out statistical tests. In practice, one pitfall that may impede the calculation of inferences using asymptotic theory is the requirement of analytic expressions and/or efficient algorithms to compute the corresponding asymptotic variance-covariance matrices [23]. Since the number of interaction effects grows exponentially with the number of inputs, efficient algorithms must be found to do the numerical computations. For this purpose, it is explained how Möbius inversion formulas and the fast Möbius transform algorithm can be used to reduce the computational expense. An application on a model in cellular biology involving 12 inputs and 4096 input combinations and interaction effects is presented.

The rest of the paper is organized as follows. Section 2 is a brief review of sensitivity analysis. Section 3 presents the inference method. Section 4 presents

the sensitivity analysis of a cellular biology model. The results are discussed in a Conclusion.

2 Sensitivity analysis

Let $f(X)$ be the output of some function f with random inputs $X := (X_1, \ldots, X_d)$. Denote $D = \{1, \ldots, d\}$. Let $\mathbf{2}^D$ be the power set of D, that is, the set of the subsets of D. For $A \in \mathbf{2}^D \setminus \emptyset$, the symbol X_A denotes the subvector of $X = X_D = (X_1, \ldots, X_d)$ that corresponds to A. For instance if $d = 4$ and $A = \{3, 1, 4\}$ then $X_A = (X_1, X_3, X_4)$.

2.1 Variance decomposition and Sobol indices

If the components of X are independent and $\mathrm{E}\, f(X)^2 < \infty$ then it holds that $f(X) = \sum_{A \subset D} f_A(X_A)$, where the functions f_A satisfy

$$\mathrm{E}(f_A(X_A)|X_B) = 0 \text{ if } A \not\subset B. \tag{1}$$

By convention, $f(X_\emptyset; \emptyset)$ denotes a constant. Equation (1) characterizes the decomposition of $f(X)$: if there are functions f'_A that satisfy (1) and $f(X) = \sum_{A \subset D} f'_A(X_A)$ then each $f'_A(X_A)$ must be equal to $f_A(X_A)$ almost surely. The above decomposition is called the Sobol decomposition, or Sobol-Hoeffding decomposition. For more details see [13, 38, 44].

The Sobol-Hoeffding decomposition entails a decomposition of the variance of $f(X)$. Indeed, equation (1) implies that the random variables $f_A(X_A)$ are uncorrelated, and hence $\mathrm{Var}\, f(X) = \sum_{A \subset D} \mathrm{Var}\, f_A(X_A)$. The quantities $\sigma^*(A) := \mathrm{Var}\, f_A(X_A)$ are called the Sobol indices [31, 38]. Note that $\sigma^*(\emptyset) = 0$. The Sobol indices associated with the singletons are called first order Sobol indices, or main effects, and the remaining indices are called interaction indices, or interaction effects. The Sobol index vector, that is, the vector with components $\sigma^*(A)$, is denoted by σ^*.

First order indices have an easy interpretation. Indeed, it is easy to see that

$$\sum_{B \subset A} f_B(X_B) = \mathrm{E}(f(X)|X_A) \tag{2}$$

and hence, taking $A = \{j\}$, one gets $\sigma^*(j) = \sigma^*(\{j\}) = \mathrm{Var}\, \mathrm{E}(f(X)|X_j) = \mathrm{Var}\, f(X) - \mathrm{E}\, \mathrm{Var}(f(X)|X_j)$, which is the expected reduction in variance induced by fixing X_j to a random value. Taking the variance in both sides of (2), one gets similarly that

$$\sum_{B \subset A} \sigma^*(B) = \mathrm{Var}\, \mathrm{E}(f(X)|X_A) =: \tau^*(A). \tag{3}$$

The index $\tau^*(A)$ is sometimes called the closed Sobol index of A [32, 19]. We shall call it the dual total index of A, or simply the dual of A. The reason stems from the duality relatioship between $\tau^*(A)$ and the total index $\tau(A)$ defined in the next section. (See (5.) The dual of A is the expected reduction in variance induced by fixing the components of X_A. The vector of dual indices will be denoted by τ^*.

2.2 Total indices

The total index $\tau(A)$ of A is defined as

$$\tau(A) = \sum_{B \cap A \neq \emptyset} \sigma^*(B). \tag{4}$$

The quantity $\tau(A)$ measures the sum of all interactions of X_A [14]. Note that $\tau(D) = \operatorname{Var} f(X)$ and $\tau(\emptyset) = 0$. By (4), it holds

$$\tau(D) - \tau(A) = \sum_{B \subset D} \sigma^*(B) - \sum_{B \cap A \neq \emptyset} \sigma^*(B) = \tau^*(D \setminus A), \tag{5}$$

where the last equality holds because of (3). It follows [31]

$$\tau(A) = \operatorname{E} \operatorname{Var}(f(X)|X_{D \setminus A}). \tag{6}$$

Therefore, the total index of A is also interpreted as the output variance induced by a random change in X_A, averaged over all the remaining inputs. It is easy to see that $\tau(A) = 0$ if and only if the function f does not depend on X_A. Thus, total indices are used in a "factor fixing" setting [34, 38]. The vector of total indices will be denoted by τ.

2.3 Shapley indices

Shapley indices assess the global importance of each individual input [28]. They result from an allocation method of game theory [37, 45]. If D denotes the set of players, then to each coalition $B \subset D$ there corresponds a real number, denoted by $\operatorname{val}(B)$, that represents the "value" of B. It is assumed that $\operatorname{val}(\emptyset) = 0$. The Shapley value of player j, $j = 1, \ldots, d$, is defined as

$$\theta(\{j\}) = \frac{1}{d} \sum_{B \subset D \setminus \{j\}} \frac{1}{\binom{d-1}{|B|}} (\operatorname{val}(B \cup \{j\}) - \operatorname{val}(B)).$$

The value of player j is based on the change in value of the coalitions, should the player be added in them. In fact, it was shown by Shapley [37] that this way of assigning values is the only possible way that satisfy certain axioms; see [37, 45] for more details. By an analogy between game theory and global sensitivity analysis (the individual inputs X_j play the role of the players), $\operatorname{val}(B)$ was set to $\operatorname{Var} \operatorname{E}(f(X)|X_B)$ in [28] and $\operatorname{E} \operatorname{Var}(f(X)|X_{D \setminus B})$ in [40]. It was shown in [40] that both choices lead to the same Shapley indices. Shapley indices have the advantage that their sum equals the model output variance. Moreover, it holds that

$$\theta(\{j\}) = \sum_{B \cap \{j\} \neq \emptyset} \frac{\sigma^*(B)}{|B|}, \tag{7}$$

for every $j = 1, \ldots, d$. (See, e.g. Lemma 1 of [28].) The vector of Shapley indices, that is, the vector with components $\theta(\{j\})$, is denoted by θ.

3 An inference method

Consider the following steps:

1. Estimate the total indices and the duals using Janssen's formula [20].
2. Compute the variance decomposition and Sobol indices using Möbius inversion formulas and the fast Möbius transform.
3. Use the fact that the vector of Sobol indices is a linear transformation of the vector of total indices to make inferences.
4. Make inferences about Shapley values using this same idea.

The steps above are described in more detail below.

3.1 Estimation of the total indices and the duals

Let $X' = (X'_1, \ldots, X'_d)$ be an independent copy of X. If $A \in \mathbf{2}^D$, denote by $X^{\setminus A}$ the vector made out of X and X' such that the components of X indexed by the members of A are replaced by those of X'. For instance, if $d = 4$ and $A = \{3, 1, 4\}$ then $X^{\setminus A} = (X'_1, X_2, X'_3, X'_4)$, $X^{\setminus \emptyset} = X$ and $X^{\setminus D} = X'$.

Janssen's formula reads

$$\tau(A) = \mathrm{E}\,\frac{1}{2}\left(f(X) - f(X^{\setminus A})\right)^2. \tag{8}$$

Janssen's formula has the advantage that total indices are expressed as simple expectations, and hence simple and yet efficient moment estimators are available. It suffices to draw a sample from (X, X'), say of size n, and substitute the empirical average for the expectation in (8). One then obtains an estimator $\widehat{\tau}$ of the vector τ. The central limit theorem immediately yields that

$$\sqrt{n}(\widehat{\tau} - \tau) \xrightarrow{\mathrm{d}} \mathrm{N}(0, T),$$

where T is the variance-covariance matrix with elements

$$\mathrm{Cov}((f(X) - f(X^{\setminus A}))^2/2, (f(X) - f(X^{\setminus B}))^2/2),$$

$A, B \subset D$. The matrix T can be estimated by taking empirical covariances. To avoid possible complications due to the fact that the entries of T corresponding to the empty set will be null, we may remove from $\widehat{\tau}$ and τ the component corresponding to the empty set.

Asymptotic p-values and confidence intervals are easily built. For instance, letting $T(A, B)$ stand for the element of T at the row and column corresponding to the sets A and B, respectively, an asymptotic confidence interval for $\tau(A)$ of level $1 - \alpha$, $\alpha \in (0, 1)$ (that is, a random interval containing the true value $\tau(A)$ with probability approaching $1 - \alpha$ as n increases) is given by $[\widehat{\tau}(A) \pm q_{1-\alpha/2}\sqrt{T(A, A)/n}]$, where here $q_{1-\alpha/2}$ is the quantile of order $1 - \alpha/2$ of the standard normal distribution. An asymptotic p-value for testing the null hypothesis "$\tau(A) = 0$" against "$\tau(A) > 0$" is given by the probability that a standard normal random variable exceeds the value $\sqrt{n}\widehat{\tau}(A)/\sqrt{T(A, A)}$ actually observed in the data.

Estimation of the dual index vector τ^* can be done from (5). There is a matrix L such that $\widehat{\tau}^* = L\widehat{\tau}$. A standard delta-method yields that

$$\sqrt{n}(\widehat{\tau}^* - \tau^*) \xrightarrow{\mathrm{d}} \mathrm{N}(0, T^*), \tag{9}$$

where

$$T^* = LTL^\top. \tag{10}$$

Confidence intervals and p-values can be built are before.

3.2 Computation of the variance decomposition and Sobol indices

3.2.1 Möbius inversion formulas

A decomposition of the variance $\mathrm{Var}\, f(X)$ is easily obtained from Möbius inversion formulas [26]. Remember that $\tau^*(A)$ can be estimated as in Section 3.1. Now define

$$\sigma^*(A) = \sum_{B \subset A} (-1)^{|A \setminus B|} \tau^*(B), \tag{11}$$

for every $A \in \mathbf{2}^D$. Möbius inversion formulas state that, for arbitrary real maps τ^* and σ^* defined on $\mathbf{2}^D$ with $\tau^*(\emptyset) = \sigma^*(\emptyset) = 0$, equation (11) holds if and only if

$$\tau^*(A) = \sum_{B \subset A} \sigma^*(B) \tag{12}$$

for every $A \in \mathbf{2}^D$. Since $\tau^*(D) = \mathrm{Var}\, f(X)$, taking $A = D$ in (12) *de facto* provides a decomposition of the variance $\mathrm{Var}\, f(X)$. It is not difficult to show that, in fact, the quantities $\sigma^*(A)$ coincide with the Sobol indices of Section 2. The maps τ^* and σ^* are sometimes known as Möbius transforms of one another [1, 12, 21].

Remark 1. *Notice that, since Möbius inversion formulas hold for arbitrary maps, and since $\tau(A)$ has been defined in Section 3.1 along with $\tau^*(A)$, we can define σ from τ the same way σ^* was defined from τ^*. Since $\tau(D)$ is also equal to* $\mathrm{Var}\, f(X)$*, this defines another variance decomposition.*

From (11), it is clear that

$$\sigma^* = M^* \tau^*, \tag{13}$$

where M^* is the matrix with entries

$$M^*(A, B) = \begin{cases} (-1)^{A \setminus B} & \text{if } B \subset A, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathrm{nnz}(M^*)$ denote the number of nonzero elements of M^*. Recall the "Big Theta" notation [22]: two functions f and g of d satisfy $g(d) = \Theta(f(d))$ if $g(d)/f(d)$ goes to a positive constant as d goes to infinity. Since, asymptotically as $d \to \infty$, $\mathrm{nnz}(M^*) = \Theta(3^d)$ (see, e.g. [25]), standard sparse multiplication yields that the computation cost of (13), measured by the number of arithmetic operations, is $\Theta(3^d)$.

3.2.2 The fast Möbius transform

The computation cost of (13) can be as little as $\Theta(d2^d)$. To achieve such a low cost, arrange the elements of the powerset in the following order: associate with each $A \in \mathbf{2}^D$ a Boolean vector of size d and order the Boolean vectors according to the lexicographical order. An example is given in Table 1.

(0,0,0)	$\emptyset$
(0,0,1)	$\{3\}$
(0,1,0)	$\{2\}$
(0,1,1)	$\{2,3\}$
(1,0,0)	$\{1\}$
(1,0,1)	$\{1,3\}$
(1,1,0)	$\{1,2\}$
(1,1,1)	$\{1,2,3\}$

Table 1: Arrangement of the subsets of $D = \{1,\ldots,d\}$ according to the lexicographical order in the Boolean space $\{0,1\}^d$ for $d = 3$.

Assuming that the components of τ^* and σ^* are arranged as described above, it can be shown that the matrix M^* in (13) is a d-fold Kronecker autoproduct of the 2×2 matrix

$$M_1^* = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}. \tag{14}$$

In symbols, we write

$$M^* = \otimes^d M_1^* = M_1^* \otimes \cdots \otimes M_1^*, \tag{15}$$

where $\otimes$ denotes the Kronecker product. A proof of (15) is given in Appendix A. See also [12].

Efficient recursive algorithms to compute arbitrary Kronecker auto-products can then be used [1, 6, 7, 12, 21, 46]. Let τ^* be an arbitrary vector of size 2^d and M_1^* be an arbitrary nonnull matrix of size 2×2. Let $\sigma^{*(0)}, \sigma^{*(1)}, \ldots, \sigma^{*(d)}$ be defined through the recurrence relations

$$\left\{ \begin{array}{lcl} \sigma^{*(0)} & = & \tau^* \\ \sigma^{*(j)} & = & (I_{2^{d-j}} \otimes M_1^* \otimes I_{2^{j-1}})\sigma^{*(j-1)} \quad (j = 1,\ldots,d), \end{array} \right. \tag{16}$$

where and $I_{2^{d-j}}$ and $I_{2^{j-1}}$ denote the identity matrices of size 2^{d-j} and 2^{j-1}, respectively. (In general, if k is an integer then I_k will denote the identity matrix of size k.) The recurrence relations (16) define an algorithm in which $\sigma^{*(j)}$ is first initialized to the vector τ^* and then updated at each iteration $j = 1,\ldots,d$. (Notice that each $\sigma^{*(j)}$ must be a vector of length 2^d.) It can be shown that the last update of (16) is equal to

$$\sigma^{*(d)} = (\otimes^d M_1^*)\tau^* = M^*\tau^* = \sigma^* \tag{17}$$

and that the total number of arithmetic operations needed to reach $\sigma^{*(d)}$ is $\Theta(d2^d)$. For completeness, a proof of these facts is given in Appendix A. Since $\mathrm{nnz}(\otimes^d M_1^*) = \mathrm{nnz}(M_1*)^d$, Algorithm (16) is less costly than standard sparse

matrix multiplication as soon as $\mathrm{nnz}(M_1^*) > 2$. In terms of computer memory usage, algorithm (16) is also preferable because it does not need to store the whole matrix $\otimes^d M_1^*$.

Taking for τ^* the actual dual total index vector and M_1^* the actual matrix (14) yields that the Sobol index vector σ^* can be computed from τ^* with computation cost $\Theta(d2^d)$.

Remark 2. *Algorithm (16) with M_1 the absolute value of (14) is refered to as the* fast Möbius transform *in [1, 21]. Indeed, If σ^* is a real map defined on $\mathbf{2}^D$ then the map defined by $\tau^* : \mathbf{2}^D \to \mathbf{R}$, $\tau^*(A) = \sum_{B \subset A} \sigma^*(B)$ is the Möbius transform of σ^*, see [1, 21]. The fast Möbius transform is then a "fast" algorithm to compute the Möbius transform from its inverse (as the fast Fourier transform computes the Fourier transform rapidly, hence the name). Computing the inverse from the Möbius transform (that is, algorithm* (16) *with M_1 as in* (14)*) is, in essence, the same algorithm. An early version was given in [46].*

Remark 3. *As noted in [12, 30], the absolute value of $M^\top$ is equal to the Sierpinski matrix, a fractal matrix.*

3.3 Making the inferences

Sobol indices

From (13) and (9), it follows that the estimator $\widehat{\sigma}^* = M^* \widehat{\tau}^*$ satisfies

$$\sqrt{n}(\widehat{\sigma}^* - \sigma^*) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \Sigma^*), \tag{18}$$

where

$$\Sigma^* = M^* T^* M^{*\top}. \tag{19}$$

The matrix Σ^* can be computed with algorithm (16): for each $j = 1, \ldots, 2^d$, apply algorithm (16) to multiply M^* by the jth column of T^*. Once all columns of M^*T^* have been calculated, compute $M^*(M^*T^*)^\top$ similarly. The total computation cost of Σ^* is thus $\Theta(2^d) \times \Theta(d2^d) = \Theta(d4^d)$. Notice that the cost of standard sparse matrix multiplication is $\Theta(6^d)$. Also observe that an estimate $\widehat{\Sigma}^*$ of Σ^* is readily obtained by substituting $\widehat{T}^*$ for T^* in (19), where $\widehat{T}^*$ is an estimate of T^* obtained by substituting $\widehat{T}$ for T in (10).

The asymptotic normality in (18) allows us to make inferences about the values of the Sobol indices, interactions included. An example is given below. Suppose we want to report, with "confidence" $(1-\alpha)100\%$, the contribution of k Sobol indices $\sigma^*(A_1), \ldots, \sigma^*(A_k)$ to the total variance $\mathrm{Var}\, f(X)$, where here $A_1, \ldots, A_{2^d-1}$ is some enumeration of the elements of $\mathbf{2}^D \setminus \emptyset$. Let us consider the test

> "H_0: the contribution is less than or equal to $(1-\eta)100\%$"

against the alternative "H_1: the contribution is more than $(1-\eta)100\%$"; more formally, the null hypothesis is given by

$$\mathrm{H}_0 : \sum_{i=1}^{k} \sigma^*(A_i) \leq (1-\eta)\, \mathrm{Var}\, f(X).$$

Let us rearrange the Sobol indices so that the ith component of σ^* becomes $\sigma^*(A_i)$. Put $c_i = \eta/(1-\eta)$ if $i = 1, \ldots, k$ and -1 otherwise. Since $\mathrm{Var}\, f(X) = \sum_{i=1}^{2^d-1} \sigma^*(A_i)$, the null hypothesis is equivalent to "$\mathrm{H}_0 : \sum_{i=1}^{2^d-1} c_i \sigma^*(A_i) \leq 0$", which is further rewritten as "$\mathrm{H}_0 : c^\top \sigma \leq 0$", where here c is the vector with components c_i. We reject the hypothesis H_0 if the test statistic $T := \sqrt{n} c^\top \widehat{\sigma}^* / \sqrt{c^\top \widehat{\Sigma}^* c}$ is greater than $q_{1-\alpha}$, where $q_{1-\alpha}$ is the quantile of the standard normal distribution of order $(1-\alpha)$. The type-I error is controlled, since $\lim_{n\to\infty} P_{\mathrm{H}_0}(T > q_{1-\alpha}) \leq \lim_{n\to\infty} P_{\mathrm{H}_0}(N > q_{1-\alpha}) \leq \alpha$, where here N is a standard normal random variable. Replace $q_{1-\alpha}$ by the actual value of the test statistic to get a bound for the p-value $\pi = P_{\mathrm{H}_0}(N > T^{\mathrm{obs}})$, where T^{obs} is the value of the test statistic observed on the actual set of simulations. The test depends on η. The higher the contribution level $1-\eta$, the less we will tend to reject H_0. Indeed, conditionally on the data, $T = T(\eta)$ is an increasing function of η on $[0,1]$ and $\pi = \pi(\eta)$ is a decreasing function of η on $[0,1]$.

Shapley indices

In view of (7), there is a matrix N^* of size $d \times 2^d$ such that $\theta = N^* \sigma^*$. From (7), it is easily seen that if $\{k\}$ and B are the ith and jth elements in some arbitrary list of the subsets of D, then the entry at the ith row and jth column of N^* is given by

$$N^*(\{k\}, B) = \begin{cases} 1/|B| & \text{if } \{k\} \cap B \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, it holds that the estimator $\widehat{\theta} := N^* \widehat{\sigma}^*$ satisfies

$$\sqrt{n}(\widehat{\theta} - \theta) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \Theta),$$

where $\Theta = N^* \Sigma^* N^{*\top}$. An estimate $\widehat{\Theta}$ of Θ is obtained by substituting $\widehat{\Sigma}^*$ for Σ. Confidence intervals and p-values can be obtained as usual.

4 Sensitivity analysis of a Boolean model of cell fate decision

We consider a mathematical model, originally proposed in [2], of the decision-making process of a eukaryote cell between different "death pathways". When subject to specific membrane receptors such as TNFα (Tumor Necrosis Factor), cells can trigger inner antagonistic signaling pathways ultimately leading to the programmed death of the cell or, on the contrary, to its survival. Two types of death are considered here: apoptosis, which is the classical programmed cell death, and necroptosis, which is a programmed death sharing apoptotic pathways yet presenting some necrotic features (thereafter, this phenotype will be abusively designated by "necrosis"). The mathematical model of [2] consists in an asynchronous Boolean model modeling the main regulatory network responsible for the decision between the three phenotypes: apoptosis, necrosis and survival. In a Boolean network, signaling molecules are represented by Boolean variables denoting their absence or presence, and regulatory processes are modeled through logical functions. In an asynchronous network, variables update asynchronously implying that a state may transition to several possible

successors. When associating probabilities to such transitions, the transition graph of the network straightforwardly translates to a Markov chain (see *e.g.* [43] for more details on the construction and analysis of asynchronous Boolean networks).

Using the model of [2], it is thus possible to compute the probability to enter apoptosis, necrosis and survival starting from a given "physiological" initial condition. (In the computational experiments, a reduced version of [2] was used.) However, the direct interpretation of these probabilities is difficult since they are based on the assumption that, for any given state, all asynchronous successors have an equal probability. A suitable way to circumvent this assumption would be to adjust these probabilities taking into account the relative velocities of all the regulatory mechanisms involved. Nevertheless, such experimental information is generally unavailable as too many complex biochemical processes are intertwined. Therefore, a global sensitivity analysis is well suited to detect which regulatory mechanisms or combinations of regulatory mechanisms have the most impact on the model's dynamics, and are therefore priority targets for further investigations.

A first sensitivity analysis of the model was performed in [47], although a "one-at-a-time" approach was used. The idea was to affect a positive weight to each variable: a weight greater than 1 means that the processes involving the variable are rapid, so the transitions updating the variable are favored, while a weight lower than 1 means that the processes are slow and the transitions are hindered. Here, we refine this approach by considering two weights for each variable instead of one: one weight for the variable's activation (*i.e.* from 0 to 1) and one weight for its deactivation (*i.e.* from 1 to 0). This comes from the basic observation that in cell biology, the production and the degradation of a given compound may have very different time scales.

In the following computational experiments, we choose six variables central to the decision process: C8 (caspase 8), RIP1, NFkB, cIAP, MOMP and MPT. The reader is referred to [2] for a comprehensive description of these biological variables and the regulatory processes they are involved in. This leads us to $d = 2 \times 6 = 12$ inputs. We then generate a dataset of $n2^d = 1000 \times 4096$ model simulations. Each simulation consists in the computation of three outputs: the probabilities of the three expected phenotypes (apoptosis, necrosis and survival). Total, Sobol and Shapley indices are then estimated for each output, as depicted in Figure 1.

The first observation that can be made is that necrosis seems qualitatively different from the other two, with generally lower index values, indicating a globally low variability of this output. We confirmed this result by drawing histograms of the three outputs in Figure 2. By inspecting highest Sobol index estimates (Fig. 1, middle row) we find that the three most important variables are C8, RIP1 and NFkB, as was originally observed in [47]. Nevertheless, here we get a more comprehensive view. The activations of proteins C8 and RIP1 (denoted by C8+ and RIP1+) seem determinant in the choice between apoptosis and survival, as these inputs contribute to more than 85% (p-value < 0.03) and 60% (p-value < 0.002) of the total variance for the "apoptosis" and "survival" outputs, respectively. Although the combination $\{\text{C8+}, \text{RIP1+}\}$ stands out both in apoptosis and in survival (in the latter, it is the fourth highest Sobol index), statistical inference has revealed that its effect is nonsignificant. The activation of transcription factor NFkB appears to play a critical role in

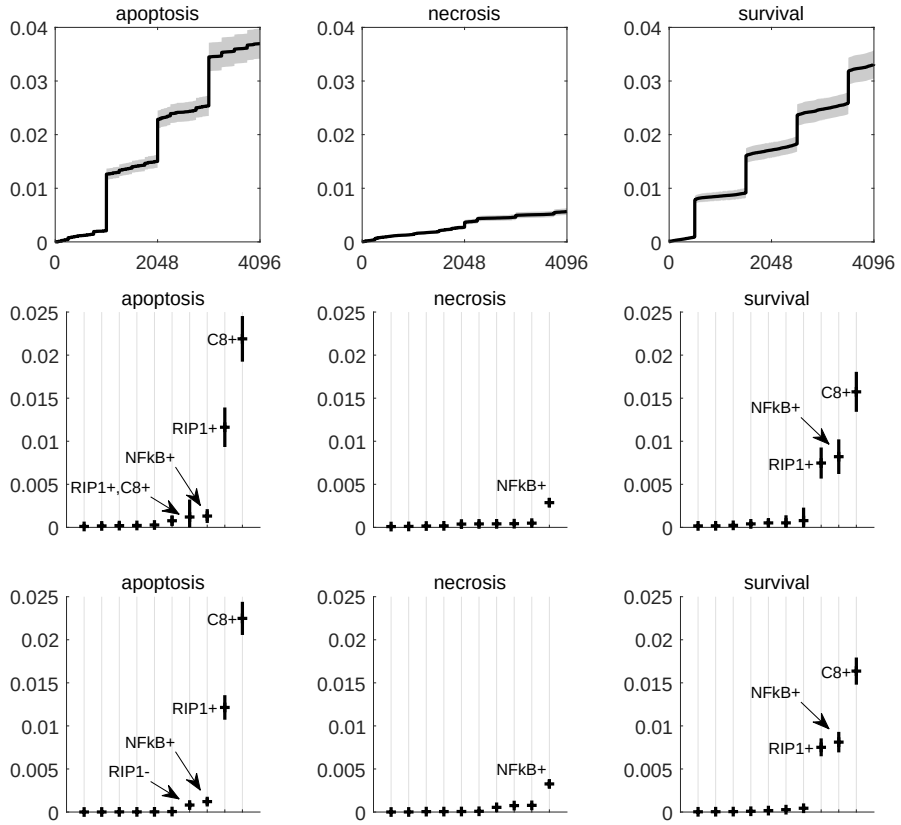

Figure 1: Global sensitivity indices of the probability of each phenotype, computed from the model in [2]. The first row shows total index estimates for all $2^d = 4096$ combinations of inputs, arranged in ascending order. The second row shows the 10 highest Sobol index estimates, with the corresponding input combination indicated only for the most important ones. The third row shows the 10 highest Shapley index estimates (again the corresponding input is only indicated for the most important). Each estimate is drawn with its 95% confidence interval.

the choice between necrosis and survival. Indeed, the input NFkB+ is the sole significant detected effect in the sensitivity analysis for the output "necrosis". and belongs to the three significant detected effects for the output "survival". The share of the total variance explained by these three inputs is more than 85% (p-value < 0.01), while it was only 60% without NFkB+. Intriguingly, however, it was calculated that while NFkB+ can be said to contribute to more than 45% of the total variance for the output "necrosis" with high confidence (p-value < 0.025), a higher level of contribution is no more significant. (For instance, testing for a contribution less than or equal to 50% leads to a p-value greater than 0.32.) Therefore, half of the total variance for the "necrosis" is left unexplained. This suggests that the choice between necrosis and non-necrosis, although strongly influenced by the activation of NFkB, cannot be reduced to that sole input. The choices between apoptosis and non-apoptosis, and between survival and non-survival, seem to be less complex, as they almost exclusively depend on a few key inputs. From the bottom row of Figure 1, similar insights can be drawn from Shapley index estimates.

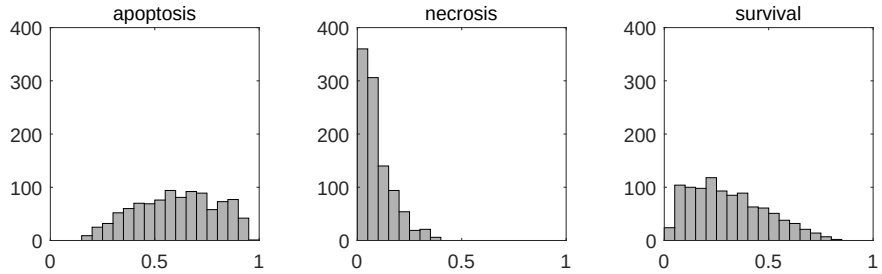

Figure 2: Histograms of 1000 independent model simulations.

Finally, it is interesting to note that total index estimates (Fig. 1, top row) appear to be relatively well separated into small numbers of clusters, especially for apoptosis and survival. Furthermore, there seems to be a striking relationship between those plateaus and highest Sobol indices. For instance, the highest plateau of apoptosis (top left, $\tau > 0.03$) exactly corresponds to all 1024 combinations of inputs that contain both C8+ and RIP1+ at the same time. A more thorough examination of the relationship between total and Sobol indices is currently investigated.

5 Discussion

An inference method for global sensitivity analysis was designed and practically tested on a model with 12 inputs. Although many estimation methods already exist in the literature, they largely focus on first-order Sobol indices only, or, when interactions are considered, the asymptotic normality of the estimators is rarely considered to quantify the uncertainties of the estimates. Thus, the method designed in this paper bridges a gap in the literature: this is a method to perform the inferences for everything from a single Monte Carlo sample, thanks to the established asymptotic normality of the estimators. In particular, it was designed a procedure to test the contribution of an arbitrary number of

Sobol indices of arbitrary orders. Computational aspects were also addressed with the fast Möbius transform algorithm. The method is general: although total indices were used in step 1 of the method, they can be replaced by any sensitivity measure for which there is an asymptotically normal estimator. The major limitation of the method is that it cannot deal with a large number of inputs. Beyond, say 20 inputs, the method may be computationally unfeasible. Moreover, for a large number of inputs, the statistical performance of the estimators is probably poor, unless a very large sample is generated. Finally, it should be noted that, for a large number of inputs, model runtime will impede the method to get a chance to be applied anyway.

A Proofs of the statements of Section 3.2.2

A.1 Proof that $\sigma^{*(d)} = (\otimes^d M_1^*)\sigma^{*(0)}$

We shall show that, in general,

$$\sigma^{*(j)} = (I_{2^{d-j}} \otimes (\otimes^j M_1^*))\sigma^{*(0)}, \tag{20}$$

for all $j = 1, \ldots, d$. The proof is by mathematical induction. Equation (20) is true for $j = 1$ because of (16). Suppose that it is true for a given j and let us show it is true for $j + 1$. From (16) and (20), we have

$$\begin{aligned}
\sigma^{*(j+1)} &= (I_{2^{d-j-1}} \otimes M_1^* \otimes I_{2^j})\, \sigma^{*(j)} \\
&= ((I_{2^{d-j-1}} \otimes M_1^*) \otimes I_{2^j}) \left(I_{2^{d-j}} \otimes (\otimes^j M_1^*)\right) \sigma^{*(0)} \\
&= ((I_{2^{d-j-1}} \otimes M_1^*) I_{2^{d-j}}) \otimes \left(I_{2^j} (\otimes^j M_1^*)\right) \sigma^{*(0)} \\
&= (I_{2^{d-j-1}} \otimes M_1^*) \otimes \left(\otimes^j M_1^*\right) \sigma^{*(0)} \\
&= (I_{2^{d-j-1}} \otimes (\otimes^{j+1} M_1^*))\sigma^{*(0)}.
\end{aligned}$$

To show that the computation cost is $\Theta(d2^d)$, notice that each matrix $I_{2^{d-j}} \otimes M_1^* \otimes I_{2^{j-1}}$ is a block-diagonal matrix of the form

$$\begin{bmatrix} M_1^* \otimes I_{2^{j-1}} & & \\ & \ddots & \\ & & M_1^* \otimes I_{2^{j-1}} \end{bmatrix}.$$

Let m_{ij} denote the element at the ith row and jth column of M_1^* ($i, j \in \{1, 2\}$). There are 2^{d-j} block-rows of the form

$$M_1^* \otimes I_{2^{j-1}} = \begin{bmatrix} m_{11} I_{2^{j-1}} & m_{12} I_{2^{j-1}} \\ m_{21} I_{2^{j-1}} & m_{22} I_{2^{j-1}} \end{bmatrix},$$

and, if multiplied by a column vector to the right, each of them leads to $\Theta(\text{nnz}(M_1^* \otimes I_{2^{j-1}})) = \Theta(\text{nnz}(M_1^*)2^{j-1}) = \Theta(2^{j-1})$ arithmetic operations, yielding $2^{d-j}\Theta(2^{j-1}) = \Theta(2^{d-1}) = \Theta(2^d)$ operations per update. Since there are d updates, the proof is complete.

A.2 Proof that $M^* = \otimes^d M_1^*$

Let ω be the inverse of the one-to-one map that with each $A \in \mathbf{2}^D$ associates $\sum_{i=1}^d b_i\, 2^{d-i} \in \{0, \ldots, 2^d - 1\}$, where $b_i = 1$ if $i \in A$ and $b_i = 0$ otherwise. (Note

that ω depends on d. For instance, for $d = 2$, $\omega(3) = \{1, 2\}$ but $\omega(3) = \{2, 3\}$ for $d = 3$.) The arrangement of the components of τ^* described at the beginning of Section 3.2.2 implies that

$$\tau_i^* = \tau^*(\omega(i)),$$

for all $i = 0, \ldots, 2^d - 1$, where in the left-hand side τ_i^* denotes the ith component (starting at zero) of the dual total index vector and in the right-hand side $\tau^*(\omega(i))$ denotes the dual total index of the set $\omega(i)$. The same holds for σ^*.

We show that M^* in (13) coincides with $\otimes^d M_1^*$ where M_1^* is as in (14). In this proof, d, and hence ω, since it depends on d, are fixed. A few results are collected in the following lemma.

Lemma 1. *Let $1 \le n \le d$ and $i, j \in \{0, \ldots, 2^{n+1} - 1\}$. If $0 \le i \le 2^n - 1$ and $2^n \le j \le 2^{n+1} - 1$ then the following statements are true:*

(i) $\omega(i) \subset \omega(j)$ *if and only if* $\omega(i) \subset \omega(j - 2^n)$.

(ii) $|\omega(j - 2^n) \setminus \omega(i)|$ *is even if and only if* $|\omega(j) \setminus \omega(i)|$ *is odd.*

(iii) $\omega(i) \not\subset \omega(j)$

We need to show that

$$(\otimes^d M_1^*)_{ij} = \begin{cases} -1 & \text{if } \omega(j) \subset \omega(i) \text{ and } |\omega(i) \setminus \omega(j)| \text{ is odd,} \\ 1 & \text{if } \omega(j) \subset \omega(i) \text{ and } |\omega(i) \setminus \omega(j)| \text{ is even,} \\ 0 & \text{otherwise.} \end{cases} \tag{21}$$

Let us show a slightly more general result. Let M_n^*, $n = 1, \ldots, d$, be a finite sequence of matrices of increasing size defined by

$$M_1^* = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \qquad M_n^* = \underbrace{M_1^* \otimes \cdots \otimes M_1^*}_{n \text{ times}} \quad (n = 1, \ldots, d).$$

Let us show that M_n^* satisfies (21) for every $i, j \in \{0, \ldots 2^n - 1\}$ and all $1 \le n \le d$. This will show in particular that $M_d^* = M^*$. The proof is by mathematical induction. If $n = 1$ then it is clear that (21) holds because $\omega(0) = \emptyset$ and $\omega(1) = \{d\}$. Let $n \ge 1$ and suppose that M_n^* satisfies (21) for every $i, j \in \{0, \ldots 2^n - 1\}$. Let us show that M_{n+1}^* satisfies (21) for every $i, j \in \{0, \ldots 2^{n+1} - 1\}$. By definition,

$$M_{n+1}^* = M_1^* \otimes M_n^* = \begin{pmatrix} M_n^* & 0 \\ -M_n^* & M_n^* \end{pmatrix}. \tag{22}$$

Case $i, j \in \{0, \ldots, 2^n - 1\}$: We have that $(M_{n+1}^*)_{i,j}$ is equal to $(M_n^*)_{i,j}$, which satisfies (21) by assumption.

Case $i \notin \{0, \ldots, 2^n - 1\}$ and $j \in \{0, \ldots, 2^n - 1\}$: We have $(M_{n+1}^*)_{i,j} = (-M_n^*)_{i-2^n,j}$. By assumption,

$$(-M_n^*)_{i-2^n,j} = \begin{cases} 1 & \text{if } \omega(j) \subset \omega(i - 2^n) \text{ and } |\omega(i - 2^n) \setminus \omega(j)| \text{ is odd,} \\ -1 & \text{if } \omega(j) \subset \omega(i - 2^n) \text{ and } |\omega(i - 2^n) \setminus \omega(j)| \text{ is even,} \\ 0 & \text{otherwise.} \end{cases}$$

Notice that, compared to (21), "1" and "-1" have been interchanged. By Lemma 1, it holds that $\omega(j) \subset \omega(i)$ is equivalent to $\omega(j) \subset \omega(i - 2^n)$. Thus, it remains to show that $|\omega(i-2^n) \setminus \omega(j)|$ is even if and only if $|\omega(i) \setminus \omega(j)|$ is odd. But again this is true from Lemma 1.

Case $i \notin \{0, \ldots, 2^n-1\}$ and $j \notin \{0, \ldots, 2^n-1\}$: Here $(M^*_{n+1})_{i,j} = (M^*_n)_{i-2^n, j-2^n}$. That $(M^*_n)_{i-2^n, j-2^n}$ satisfies (21) follows from the same considerations as in the previous case.

Case $i \in \{0, \ldots, 2^n - 1\}$ and $j \notin \{0, \ldots, 2^n - 1\}$: We know from (22) that $(M^*_{n+1})_{i,j} = 0$. To show that (21) is satisfied, it suffices to see that $\omega(j) \not\subset \omega(i)$. But this is true from Lemma 1, since $j > i$.

References

[1] A. Björklund, T. Husfeldt, P. Kaski, and M. Koivisto. Fourier meets Möbius: fast subset convolution. In *Proc. of the 39th ACM symposium on Theory of computing*, pages 67–74, 2007.

[2] L. Calzone, L. Tournier, S. Fourquet, D. Thieffry, B. Zhivotovsky, E. Barillot, and A. Zinovyev. Mathematical modelling of cell-fate decision in response to death receptor engagement. *PLOS Computational Biology*, 6(3):e1000702, 2010.

[3] J. Castro, D. Gómez, and J. Tejada. Polynomial calculation of the Shapley value based on sampling. *Computers & Operations Research*, 36(5):1726–1730, May 2009.

[4] S. Da Veiga and F. Gamboa. Efficient estimation of sensitivity indices. *Journal of Nonparametric Statistics*, 25(3):573–595, 2013. Publisher: Taylor & Francis.

[5] S. Da Veiga, F. Gamboa, A. Lagnoux, T. Klein, and C. Prieur. New estimation of Sobol'indices using kernels. *arXiv preprint arXiv:2303.17832*, 2023.

[6] T. Dayar and M. C. Orhan. On Vector-Kronecker Product Multiplication with Rectangular Factors. *SIAM Journal on Scientific Computing*, 37(5):S526–S543, Jan. 2015. Publisher: Society for Industrial and Applied Mathematics.

[7] P. Fernandes, B. Plateau, and W. J. Stewart. Efficient descriptor-vector multiplications in stochastic automata networks. *Journal of the ACM*, 45(3):381–414, 1998.

[8] F. Gamboa, A. Janon, T. Klein, A. Lagnoux, and C. Prieur. Statistical inference for Sobol pick-freeze Monte Carlo method. *Statistics*, 50(4):881–902, 2016.

[9] F. Gamboa, T. Klein, A. Lagnoux, and L. Moreno. Sensitivity analysis in general metric spaces. *Reliability Engineering & System Safety*, 212:107611, 2021.

[10] L. Gilquin, E. Arnaud, C. Prieur, and A. Janon. Making the best use of permutations to compute sensitivity indices with replicated orthogonal arrays. *Reliability Engineering & System Safety*, 187:28–39, 2019. Sensitivity Analysis of Model Output.

[11] T. Goda. A simple algorithm for global sensitivity analysis with Shapley effects. *Reliability Engineering & System Safety*, 213:107702, Sept. 2021.

[12] M. Grabisch, J.-L. Marichal, and M. Roubens. Equivalent Representations of Set Functions. *Mathematics of Operations Research*, 25(2):157–178, May 2000. Publisher: INFORMS.

[13] W. Hoeffding. A class of statistics with asymptotically normal distribution. *The Annals of Mathematical Statistics*, 19(3):293–325, 1948.

[14] T. Homma and A. Saltelli. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, 52(1):1–17, 1996.

[15] R. L. Iman and S. C. Hora. A robust measure of uncertainty importance for use in fault tree system analysis. *Risk Analysis*, 10(3):401–406, 1990.

[16] B. Iooss and C. Prieur. Shapley effects for sensitivity analysis with correlated inputs: Comparisons with Sobol' indices, numerical estimation and applications. *International Journal for Uncertainty Quantification*, 9(5), 2019. Publisher: Begel House Inc.

[17] B. Iooss, S. D. Veiga, A. Janon, G. Pujol, with contribu-tions from Baptiste Broto, K. Boumhaout, T. Delage, R. E. Amri, J. Fruth, L. rent Gilquin, J. Guillaume, L. L. Gratiet, P. Lemaitre, A. Marrel, A. Meynaoui, B. L. Nelson, F. ippo Monari, R. Oomen, O. Rakovec, B. Ramos, O. Roustant, E. hye Song, J. Staum, R. Sueur, T. Touati, and F. Weber. *sensitivity*, 2020. R package version 1.23.0.

[18] A. Janon, T. Klein, A. Lagnoux, M. Nodet, and C. Prieur. Asymptotic normality and efficiency of two Sobol index estimators. *ESAIM: Probability and Statistics*, 18:342–364, 2014.

[19] M. J. Jansen. Analysis of variance designs for model output. *Computer Physics Communications*, 117(1):35–43, 1999.

[20] P. Janssen. Assessing sensitivities and uncertainties in models: a critical evaluation. In J. Grasman and G. van Straten, editors, *Predictability and Nonlinear Modelling in Natural Sciences and Economics*. Springer-Science+Business Media, B.V., 1994.

[21] R. Kennes. Computational aspects of the Mobius transformation of graphs. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(2):201–223, Mar. 1992. Conference Name: IEEE Transactions on Systems, Man, and Cybernetics.

[22] D. E. Knuth. *The art of computer programming*. Addison Wesley Longman, 1997. Volume 1. Fundamental Algorithms.

[23] E. Koehler, E. Brown, and S. J.-P. A. Haneuse. On the assessment of Monte Carlo error in simulation-based statistical analyses. *The American Statistician*, 63(2):155–162, 2009.

[24] S. Kucherenko and S. Song. Different numerical estimators for main effect global sensitivity indices. *Reliability Engineering and System Safety*, 165:222–238, 2017.

[25] R. La Haye. Binary relations on the power set of an n-element set. *Journal of Integer Sequences*, 12(2):3, 2009.

[26] R. Liu and A. B. Owen. Estimating mean dimensionality of analysis of variance decompositions. *Journal of the American Statistical Association*, 101(474):712–721, 2006.

[27] A. B. Owen. Better estimation of small Sobol'sensitivity indices. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 23(2):1–17, 2013.

[28] A. B. Owen. Sobol' Indices and Shapley Value. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1):245–251, Jan. 2014.

[29] S. L. Piano, F. Ferretti, A. Puy, D. Albrecht, and A. Saltelli. Variance-based sensitivity analysis: The quest for better estimators and designs between explorativity and economy. *Reliability Engineering and System Safety*, 206, 2021.

[30] E. Plischke, G. Rabitti, and E. Borgonovo. Computing Shapley effects for sensitivity analysis. *SIAM/ASA Journal on Uncertainty Quantification*, 9(4):1411–1437, 2021.

[31] C. Prieur and S. Tarantola. Variance-based sensitivity analysis: theory and estimation algorithms. In *Handbook of Uncertainty Quantification*, pages 1–23. Springer, 2015.

[32] A. Saltelli. Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications*, 145(2):280–297, 2002.

[33] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola. Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. *Computer Physics Communications*, 181:259–270, 2010.

[34] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. *Global Sensitivity Analysis. The Primer*. Wiley, 2008.

[35] A. Saltelli, S. Tarantola, and F. Campolongo. Sensitivity anaysis as an ingredient of modeling. *Statistical Science*, 15(4):377–395, 2000.

[36] A. Saltelli, S. Tarantola, and K. P.-S. Chan. A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, 41(1):39–56, 1999.

[37] L. S. Shapley. The value of an n-person game. Technical report, RAND CORP SANTA MONICA CA, 1951. Technical Report.

[38] I. M. Sobol. Sensitivity estimates for nonlinear mathematical models. *Mathematical modelling and computational experiments*, 1(4):407–414, 1993.

[39] M. Solís. Non-parametric estimation of the first-order Sobol indices with bootstrap bandwidth. *Communications in Statistics-Simulation and Computation*, pages 1–16, 2019.

[40] E. Song, B. L. Nelson, and J. Staum. Shapley Effects for Global Sensitivity Analysis: Theory and Computation. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1):1060–1083, Jan. 2016. Publisher: Society for Industrial and Applied Mathematics.

[41] B. Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety*, 93:964–979, 2008.

[42] J.-Y. Tissot and C. Prieur. A randomized orthogonal array-based procedure for the estimation of first- and second-order Sobol' indices. *Journal of Statistical Computation and Simulation*, 85(7):1358–1381, 2015.

[43] L. Tournier and M. Chaves. Uncovering operational interactions in genetic networks using asynchronous Boolean dynamics. *Journal of Theoretical Biology*, 260(2):196–209, 2009.

[44] A. W. van der Vaart. *Asymptotic Statistics*. Cambridge University Press, 1998.

[45] E. Winter. Chapter 53 The Shapley value. In *Handbook of Game Theory with Economic Applications*, volume 3, pages 2025–2054. Elsevier, Jan. 2002.

[46] F. Yates. *The design and analysis of factorial experiments*. Imperial Bureau of Soil Science, Harpenden, 1937.

[47] A. Zinovyev, S. Fourquet, L. Tournier, L. Calzone, and E. Barillot. Cell death and life in cancer: mathematical modeling of cell fate decisions. In *Advances in Systems Biology*, pages 261–274. Springer, 2012.