An inference method for global sensitivity analysis

Gildas Mazo, Laurent Tournier

An inference method for global sensitivity analysis

Gildas Mazo
Université Paris-Saclay, INRAE, MaIAGE, 78350, Jouy-en-Josas, France
and
Laurent Tournier
Université Paris-Saclay, INRAE, MaIAGE, 78350, Jouy-en-Josas, France

May 29, 2024

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 Monte Carlo sample is given. One advantage of this method is its simplicity: it leverages the fact that Shapley effects and Sobol indices are only linear transformations of total indices, so that standard asymptotic theory suffices to get confidence intervals and to carry out statistical tests. To perform the numerical computations efficiently, Möbius inversion formulas are used, and linked to the fast Möbius transform algorithm. The method is illustrated on two dynamical systems, both with an application in life sciences: a Boolean network modeling a cellular decision-making process involving 12 inputs, and a system of ordinary differential equations modeling some population dynamics involving 10 inputs.

Keywords: Möbius inversion, Möbius transform, Sobol index, high order interaction effect

1 Introduction

The sensitivity analysis of mathematical models is important to gain insight into one's model, eventually suggesting potential avenues for model reduction, refinement, or guidance in the design of future experiments.

Sensitivity analysis seeks to quantify, for a given mathematical or computational 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 around particular points in the input space, global sensitivity analysis assesses the inputs' contributions over the whole input space (Saltelli et al., 2008, 2000), 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. We also assume throughout that the inputs $X_1, \ldots, X_d$ are mutually independent.

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. Saltelli et al. (2000)). In the special case where $A = \{j\}$, $j = 1, \ldots, d$, 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 of the output, should the input X_j be fixed (Iman and Hora, 1990; Janssen, 1994). The closed Sobol index $\tau^*(A)$, given by $\operatorname{Var} \mathrm{E}(f(X)|X_A)$, is a generalization of the first-order Sobol index to any subset $A \subset D$. We shall call it the dual index for a reason explained later. The total index $\tau(A)$ is defined as the sum over all interaction 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 effects stem from the application of an allocation method coming from cooperative game theory (Shapley, 1951; Winter, 2002). One assigns a "value" to each of the players, based

on the performance of all the possible coalitions among them. By analogy, one assigns a value—called the Shapley effect—to each of the inputs, based on all the possible input combinations. The Shapley effect $\theta(\{j\})$ of the individual input X_j summarizes its global importance. The sum of the Shapley effects sums up to the output variance. (This is not true for the individual total indices or the individual Sobol indices.)

Estimation of sensitivity indices has attracted much attention in the literature. In particular, numerous 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 (Sobol, 1993; Jansen, 1999; Janon et al., 2014; Gamboa et al., 2016; Homma and Saltelli, 1996; Saltelli, 2002; Saltelli et al., 2010; Owen, 2013). Asymptotic normality of some estimators is studied in Gamboa et al. (2016); Janon et al. (2014). To avoid the waste of data, more elaborate ways of combining outputs have been proposed (Tissot and Prieur, 2015; Gilquin et al., 2019; Saltelli et al., 2010; Saltelli, 2002), the goal being always to have the most efficient estimates with the least model runs (Piano et al., 2021). See also Saltelli et al. (1999) for another way of exploring the input space. Methods based on U-statistics are found in Gamboa et al. (2021). There are also methods that offer a different look, statistically speaking. For instance, nonparametric regression techniques, or techniques based on kernels, are found in Da Veiga et al. (2023); Da Veiga and Gamboa (2013); Solís (2019). The use of meta-models to compute sensitivity indices has been proposed as well (Sudret, 2008; Bénard et al., 2022). 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. For instance, the reader is referred to the books by Prieur and Tarantola (2015) or by Da Veiga et al. (2023) for more details and possibly other estimation methods. A list of methods is also available in Iooss et al. (2020). A numerical comparison of some of the above methods is given in Kucherenko and Song

(2017). Although a plethora of methods exist to estimate first-order indices, the literature is thinner about the estimation of higher-order interaction effects. Attention has increased recently, though, as they are useful to compute Shapley effects (Castro et al., 2009; Iooss and Prieur, 2019; Plischke et al., 2021; Song et al., 2016; Goda, 2021; Broto et al., 2020).

Rather than narrowing the estimation to particular sensitivity indices or effects, we can look at the problem of inference globally: the goal is then to extract, combine and assess the information provided by the four kinds of sensitivity indices and effects of all orders, thus paving the way for richer analyses. One advantage of such a global approach is that it takes into account the effects of interactions between inputs. These effects, indeed, carry important information when considering mathematical models of complex natural phenomena (Ball, 2023; Boyle et al., 2017). For instance, they explain a significant share of the total output variability in the two dynamical models (Calzone et al., 2010; De Paepe et al., 2016) analysed in Section 5 below. The first is a Boolean network modeling cell fate decision (Calzone et al., 2010) and the second is a differential system modeling microbial population dynamics (De Paepe et al., 2016). While these models were originally analysed with local and one-at-a-time appraoches, we shall see that using a global approach and estimating higher order sensitivity indices gives the modeler valuable insight about the system's dynamical behavior, emphasizing possibly complex interactions between the system's parameters. These interactions may be relevant to the modeled system, or they can be mathematical artifacts of the model. In either case, estimating higher order indices will provide useful information when analyzing a complex mathematical model.

To address the problem of inference globally, an approach that exploits the linear mappings between the four kinds of sensitivity indices is proposed. The asymptotic normality of the estimator of the vector of total indices, the vector of dual indices, the vector of Sobol indices and the vector of Shapley effects is established. The formulas of the asymptotic variance-covariance matrices of the estimators are given in closed form. Algorithms to

compute the estimators and their asymptotic variance-covariance matrices are presented, analysed, and discussed. To assess the statistical significance of the extracted information, several statistical tests are built.

The method is applied to two published dynamical systems, both with application in life sciences. The two systems are different in nature: the first one is a stochastic Boolean network modeling a cellular decision-making process introduced by Calzone et al. (2010); the second one is a deterministic differential system modeling population dynamics introduced by De Paepe et al. (2016). Both systems were analyzed with local approaches; here a global sensitivity analysis is proposed, including the estimation of total and Sobol indices as well as Shapley effects, and the analysis of those indices based on statistical testing. We show how such inferences can be used to gain insight on the systems' dynamics.

The rest of the paper is organized as follows. Section 2 is a brief review of the four major kinds of sensitivity indices. Section 3 presents the inference method, from the estimation of the sensitivity indices to asymptotic normality and hypothesis testing. The computational aspects are addressed in Section 4. Section 5 presents the sensitivity analysis of the two dynamical systems. Section 6 summarizes the results and discusses limitations and possible future developments.

2 Global sensitivity indices

The definition of the four major kinds of sensitivity indices—namely, Sobol, total, dual indices and Shapley effects—are recalled in this section. Let $f(X)$ be the output of some real function f with real random inputs $X := (X_1, \ldots, X_d)$. Recall that inputs are assumed to be mutually independent. 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 $X_A = (X_1, X_3, X_4)$.

If the components of X are mutually 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

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

Throughout, the subset symbol $\subset$ is to be understood in the non-strict sense, meaning that $A \subset A$ always holds. By convention, $f(X_\emptyset; \emptyset)$ is 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 Hoeffding (1948); Sobol (1993); van der Vaart (1998). According to the standard theory of global sensitivity analysis, this decomposition is the starting point to define the sensitivity indices.

Sobol indices. 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 (Prieur and Tarantola, 2015; Sobol, 1993). 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. 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 $\sigma^*(\{j\}) = \mathrm{Var}\, \mathrm{E}(f(X)|X_j) = \mathrm{Var}\, f(X) - \mathrm{E}\, \mathrm{Var}(f(X)|X_j)$, which is the expected reduction in variance induced by fixing X_j to a random value.

Example 1. *Let* $f(X) = f(X_1, X_2, X_3) = \sin X_1 + a \sin^2 X_2 + bX_3^4 \sin X_1$, *where* X_1, X_2, X_3 *are independent random variables uniformly distributed on the interval* $(-\pi, \pi)$. *The Sobol indices are given by* $\sigma^*(\{1\}) = 1/2 + b\pi^4/5 + b^2\pi^8/50$, $\sigma^*(\{2\}) = a^2/8$, $\sigma^*(\{1,3\}) = (1/18 - 1/50)b^2\pi^8$, $\sigma^*(\{3\}) = \sigma^*(\{1,2\}) = \sigma^*(\{2,3\}) = \sigma^*(\{1,2,3\}) = \sigma^*(\emptyset) = 0$, *see e.g. Sobol and Levitan (1999).*

We shall denote by σ^* the vector with components $\sigma^*(A)$ and call it the Sobol index vector.

Dual indices (closed Sobol indices). Taking the variance in both sides of (2), one gets that

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

The index $\tau^*(A)$ is sometimes called the closed Sobol index of A (Saltelli, 2002; Jansen, 1999). We shall call it the dual total index of A, or simply the dual index of A. The reason stems from the duality relationship between $\tau^*(A)$ and the total index $\tau(A)$ defined in the next section, see equation (5) below. (See also Herin et al. (2022).) The dual index of A is the expected reduction in variance induced by fixing the components of X_A.

Example 2 (Continuation of Example 1)**.** *We have*

$$\begin{pmatrix} \tau^*(\emptyset) \\ \tau^*(\{1\}) \\ \tau^*(\{2\}) \\ \tau^*(\{3\}) \\ \tau^*(\{1,2\}) \\ \tau^*(\{1,3\}) \\ \tau^*(\{2,3\}) \\ \tau^*(\{1,2,3\}) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \sigma^*(\emptyset) \\ \sigma^*(\{1\}) \\ \sigma^*(\{2\}) \\ \sigma^*(\{3\}) \\ \sigma^*(\{1,2\}) \\ \sigma^*(\{1,3\}) \\ \sigma^*(\{2,3\}) \\ \sigma^*(\{1,2,3\}) \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ \frac{1}{2} + \frac{b\pi^4}{5} + \frac{b^2\pi^8}{50} \\ \frac{a^2}{8} \\ 0 \\ \frac{1}{2} + \frac{b\pi^4}{5} + \frac{b^2\pi^8}{50} + \frac{a^2}{8} \\ \frac{1}{2} + \frac{b\pi^4}{5} + \frac{b^2\pi^8}{18} \\ \frac{a^2}{8} \\ \frac{1}{2} + \frac{b\pi^4}{5} + \frac{b^2\pi^8}{18} + \frac{a^2}{8} \end{pmatrix}.$$

The vector of dual indices will be denoted by τ^*.

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 (Homma and Saltelli, 1996). 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 (Prieur and Tarantola, 2015)

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

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 (Saltelli et al., 2008; Sobol, 1993).

Example 3 (Continuation of Example 1)**.** *We have*

$$\begin{pmatrix} \tau(\emptyset) \\ \tau(\{1\}) \\ \tau(\{2\}) \\ \tau(\{3\}) \\ \tau(\{1,2\}) \\ \tau(\{1,3\}) \\ \tau(\{2,3\}) \\ \tau(\{1,2,3\}) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \sigma^*(\emptyset) \\ \sigma^*(\{1\}) \\ \sigma^*(\{2\}) \\ \sigma^*(\{3\}) \\ \sigma^*(\{1,2\}) \\ \sigma^*(\{1,3\}) \\ \sigma^*(\{2,3\}) \\ \sigma^*(\{1,2,3\}) \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ \frac{1}{2} + \frac{b\pi^4}{5} + \frac{b^2\pi^8}{50} \\ \frac{a^2}{8} \\ 0 \\ \frac{1}{2} + \frac{b\pi^4}{5} + \frac{b^2\pi^8}{18} + \frac{a^2}{8} \\ \frac{1}{2} + \frac{b\pi^4}{5} + \frac{b^2\pi^8}{50} \\ \frac{a^2}{8} + \frac{16b^2\pi^8}{225} \\ \frac{1}{2} + \frac{b\pi^4}{5} + \frac{b^2\pi^8}{18} + \frac{a^2}{8} \end{pmatrix}.$$

The vector of total indices will be denoted by τ.

Shapley effects. Shapley effects assess the global importance of each individual input (Owen, 2014). They result from an allocation method of cooperative game theory (Shapley, 1951; Winter, 2002). If D is seen as a set of players, then to each coalition $B \subset D$ there corresponds a real number, denoted by $\text{val}(B)$, which represents the "value" of B. It is assumed that $\text{val}(\emptyset) = 0$. The Shapley value of player j, $j = 1, \ldots, d$, is defined as

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

The value of player j is based on the change in value of the coalitions, should that player be added in them. In fact, it was shown by Shapley (Shapley, 1951) that this way of assigning values is the only possible way that satisfy certain axioms, among which $\text{val}(D) = \sum_{j=1}^{d} \theta_{\text{val}}(\{j\})$. See Shapley (1951); Winter (2002) for more details. The Shapley value can be re-expressed (Harsanyi, 1963) as

$$\theta_{\text{val}}(\{j\}) = \sum_{B \cap \{j\} \neq \emptyset} \frac{\sum_{A \subset B} (-1)^{|B \setminus A|} \text{val}(A)}{|B|}, \tag{7}$$

a formula that will be useful later on. If val is replaced by its dual $\text{val}^*(B) = \text{val}(D) - \text{val}(D \setminus B)$ then $\theta_{\text{val}}(\{j\}) = \theta_{\text{val}^*}(\{j\})$ for all $j = 1, \ldots, d$. We say that the Shapley value is self-dual, see Funaki (1998); Oishi et al. (2016).

By an analogy between cooperative game theory and global sensitivity analysis (the individual inputs X_j play the role of the players), we can choose

$$\text{val}(B) = \text{Var}\, \text{E}(f(X)|X_B) = \tau^*(B)$$

or

$$\text{val}(B) = \text{E}\, \text{Var}(f(X)|X_{D \setminus B}) = \tau(B),$$

leading *a priori* to two sets of indices $\theta_{\tau^*}(\{j\})$ and $\theta_{\tau}(\{j\})$. The first choice was made in Owen (2014) and the second in Song et al. (2016). But since τ and τ^* are duals

of each other, it holds that $\theta_\tau(\{j\}) = \theta_{\tau^*}(\{j\})$ for every $j = 1, \ldots, d$. This property was noticed in Song et al. (2016) and is a direct consequence of the self-duality of the Shapley value in cooperative game theory. Thus, it makes sense to define

$$\theta(\{j\}) := \theta_\tau(\{j\}) = \theta_{\tau^*}(\{j\}), \tag{8}$$

which are called Shapley effects in sensitivity analysis.

Shapley effects can be expressed in terms of Sobol indices as

$$\theta(\{j\}) = \sum_{B \cap \{j\} \neq \emptyset} \frac{\sigma^*(B)}{|B|}; \tag{9}$$

see Owen (2014). This formula is a consequence of (7) and the Möbius inversion formula discussed in Section 3 below.

Example 4 (Continuation of Example 1)**.** *We have*

$$\begin{pmatrix} \theta(\{1\}) \\ \theta(\{2\}) \\ \theta(\{3\}) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 & 1/2 & 1/2 & 0 & 1/3 \\ 0 & 0 & 1 & 0 & 1/2 & 0 & 1/2 & 1/3 \\ 0 & 0 & 0 & 1 & 0 & 1/2 & 1/2 & 1/3 \end{pmatrix} \begin{pmatrix} \sigma^*(\emptyset) \\ \sigma^*(\{1\}) \\ \sigma^*(\{2\}) \\ \sigma^*(\{3\}) \\ \sigma^*(\{1,2\}) \\ \sigma^*(\{1,3\}) \\ \sigma^*(\{2,3\}) \\ \sigma^*(\{1,2,3\}) \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{2} + \frac{b\pi^4}{5} + \frac{17b^2\pi^8}{450} \\ \frac{a^2}{8} \\ \frac{4b^2\pi^8}{225} \end{pmatrix}.$$

Shapley effects in sensitivity analysis are interesting because they allow one to assess the global importance of the individual inputs by taking interactions into account in

a "fair" way. Indeed, as seen in (9), the interaction effects captured by the Sobol indices are divided by the size of the interaction. This leads to a fair allocation of the output variance: we have that $\mathrm{Var}\, f(X) = \sum_{j=1}^{d} \theta(\{j\})$. Let us mention another advantage of Shapley effects: they can be used when the inputs are not mutually independent (Owen and Prieur, 2017).

We shall denote by θ the vector of Shapley effects.

3 A global inference method

According to the standard theory of global sensitivity analysis, Sobol indices are defined first and, then, total indices are defined as cumulative effects of Sobol indices. It is more convenient, however, to go the other way around: in the following total indices are first estimated as global effects of combinations of inputs and Sobol index estimates ensue through a linear transformation of the vector of total index estimates. Estimation and inference is then made straightforward even for higher-order Sobol indices. In summary, the method is based on the following steps:

1. Express the total indices using Jansen's formula.

2. Express the vector of dual indices as a linear transformation of the vector of total indices.

3. Express the vector of Sobol indices as a linear transformation of the vector of dual indices.

4. If needed, express the vector of Shapley effects as a linear transformation of the vector of Sobol indices.

Remark 1. *The mention "if needed" above is to emphasize that the proposed inference method should* not *have as its primary goal the estimation of Shapley indices. For if that*

were the objective then there are methods that may be more appropriate as they bypass the estimation of Sobol indices altogether (Goda, 2021). Here, the estimation of Shapley indices should be seen as a by-product: once the Sobol indices are calculated, the estimation of Shapley effects is almost free.

By inference it is meant the ability to take into account the Monte Carlo error in the insights, statements and conclusions drawn from the data. This shall take the form of confidence intervals and p-values associated with statistical tests of hypotheses.

The steps above are described in more detail below. Section 3.1 presents the estimators and Section 3.2 presents results and methods to make inferences.

3.1 Estimation of the sets of indices

The estimation of the four kinds of sensitivity index vectors by following the four steps above is addressed next.

Total indices. 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 $X^{\setminus D} = X'$. Since we assumed mutually independent inputs, the total index associated with A given in (6) can be rewritten as

$$\tau(A) = \mathrm{E}\,\frac{1}{2}\left(f(X) - f(X^{\setminus A})\right)^2. \tag{10}$$

The above formula, known as Jansen's formula (Jansen, 1999), has the advantage that total indices are expressed as simple expectations, and hence simple 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 (10). More precisely, for each $i = 1, \ldots, n$, draw two independent copies $X^{(i)} = (X_1^{(i)}, \ldots, X_d^{(i)})$ and $X'^{(i)} =$

$(X_1'^{(i)}, \ldots, X_d'^{(i)})$ and then perform the steps below:

$$\begin{cases} \text{for each } A \in \mathbf{2}^D \\ \quad \text{make } X^{(i)\setminus A} \text{ out of } X^{(i)} \text{ and } X'^{(i)} \\ \quad \text{compute } f(X^{(i)\setminus A}). \end{cases} \tag{11}$$

One then obtains an estimator

$$\widehat{\tau}(A) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{2}\left(f(X^{(i)}) - f(X^{(i)\setminus A})\right)^2 \tag{12}$$

for each $A \subset D$. Denote by $\widehat{\tau}$ the estimator of the vector τ obtained as above.

Remark 2. *While the estimator $\widehat{\tau}$ is conceptually simple, more efficient estimators can be built, for instance with a smaller variance or necessitating less function evaluations. More details can be found in the Supplementary material (Section C).*

Dual indices. Estimation of the dual index vector τ^* can be done from (5). Note that there is a matrix L such that $\tau^* = L\tau$. More precisely, if $L(A, B)$ denotes the element of L at the row and column corresponding to the nonempty sets A and B, respectively, then

$$L(A,B) = \begin{cases} -1 & \text{if } B = D \setminus A \\ 1 & \text{if } B = D \\ 0 & \text{otherwise.} \end{cases}$$

(By convention we may choose $L(\emptyset, \emptyset) = 1$ and $L(\emptyset, B) = 0$ for all $B \neq \emptyset$. For instance, if $d = 2$ then

$$\begin{pmatrix} 0 \\ \tau^*(\{2\}) \\ \tau^*(\{1\}) \\ \tau^*(\{1,2\}) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ \tau(\{2\}) \\ \tau(\{1\}) \\ \tau(\{1,2\}) \end{pmatrix}.$$

Note that the expression of matrix L depends on the chosen order of the components of τ. We shall denote by $\widehat{\tau}^*$ the estimator of the vector τ^*.

Sobol indices. A decomposition of the output variance $\operatorname{Var} f(X)$ is easily obtained from the vector of dual indices τ^*, see Liu and Owen (2006). Let τ^* and σ^* momentarily denote *arbitrary* real maps defined on $\mathbf{2}^D$, with $\tau^*(\emptyset) = \sigma^*(\emptyset) = 0$. Then it holds that

$$\tau^*(A) = \sum_{B \subset A} \sigma^*(B) \quad (\text{for every } A \in \mathbf{2}^D) \tag{13}$$

and

$$\sigma^*(A) = \sum_{B \subset A} (-1)^{|A \setminus B|} \tau^*(B) \quad (\text{for every } A \in \mathbf{2}^D) \tag{14}$$

are equivalent. The pair of equations (13) and (14) is known as Möbius inversion formulas. The equivalence is a direct consequence of a result from Rota (1964) on finite partially ordered sets. Now, let us come back to sensitivity analysis and let the total index vector be defined through (10). Let τ^* be the dual index vector of τ. Both vectors can be estimated directly as seen above. Then let σ^* be defined through (14). Since $\tau^*(D) = \operatorname{Var} f(X)$, taking $A = D$ in (13) *de facto* provides a decomposition of the output variance $\operatorname{Var} f(X)$. It is not difficult to show that, in fact, the quantities $\sigma^*(A)$ coincide with the Sobol indices as defined in Section 2. The maps τ^* and σ^* are sometimes known as Möbius transforms of one another (Björklund et al., 2007; Grabisch et al., 2000; Kennes, 1992).

Remark 3. *Notice that, since Möbius inversion formulas hold for arbitrary maps, we can define a vector σ from the vector τ the same way σ^* was defined from τ^*. Since $\tau(D)$ is also equal to $\operatorname{Var} f(X)$, this defines another variance decomposition. By the same reasoning, every real map τ defined on $\mathbf{2}^D$ with $\tau(\emptyset) = 0$ and $\tau(D) = \operatorname{Var} f(X)$ leads to a variance decomposition. According to Il Idrissi et al. (2023), the choice (14) is justified because of the Sobol-Hoeffding decomposition for mutually independent inputs.*

From (14), it is then clear that

$$\sigma^* = M^* \tau^*, \tag{15}$$

where M^* is the matrix with entries

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

(As noted in Grabisch et al. (2000); Plischke et al. (2021), the absolute value of $M^{*\top}$ is equal to the Sierpinski matrix, a fractal matrix.) We shall denote by $\widehat{\sigma^*}$ the estimator of the vector σ^* obtained by substituting $\widehat{\tau}^*$ for τ^* in (15).

Shapley effects. In view of (9), there is a matrix N^* of size $d \times 2^d$ such that

$$\theta = N^* \sigma^*. \tag{16}$$

From (9), 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

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

We shall denote by $\widehat{\theta}$ the estimator of the vector θ obtained by substituting $\widehat{\sigma}^*$ for σ^* in (16).

3.2 Making the inferences

The asymptotic normality of the estimators is established and utilized to calculate asymptotic confidence intervals, p-values, and test various hypotheses. A reminder about asymptotic confidence intervals and hypothesis testing is given first. For numerical examples and illustrations, see Section 5.

Preliminaries

Let ψ denote any of $\tau, \tau^*, \sigma^*, \theta$ and let $\widehat{\psi}$ denote its estimate. Consider first the situation where it is known that $\widehat{\psi}$ follows a multivariate normal distribution with mean ψ and variance-covariance matrix Ψ/n. Then, for any given $A \subset D$, we have that $\widehat{\psi}(A)$ is normally distributed with mean $\psi(A)$ and variance $\Psi(A, A)/n$ and hence we could calculate a confidence interval of level, say $1 - \alpha \in (0, 1)$ under the form $[\widehat{\psi}(A) \pm q_{1-\alpha/2}\sqrt{\Psi(A, A)/n}]$, where $q_{1-\alpha/2}$ is the quantile of order $1 - \alpha/2$ of a standard normal distribution.

In all but the most simple cases, however, it is unknown whether $\widehat{\psi}$ actually follows a multivariate normal distribution. But if we can establish that $\widehat{\psi}$ approximately follows a multivariate normal distribution, that is, if we can establish that

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

for some asymptotic variance-covariance matrix Ψ, then we could do as above and get an approximate confidence interval.

The same goes for hypothesis testing, which is another way of making inferences. Remember that to decide between a null hypothesis H_0 and an alternative H_1, one computes a statistic S_n from the data and rejects H_0 if S_n belongs to some "rejection set". For simplicity, assume that rejection set is of the form (r, ∞), so that rejection occurs if $S_n > r$. The so-called type I and type II errors of the test are given by $\Pr_{\mathrm{H}_0}(S_n > r)$ (probability of rejecting H_0 while it is true) and $\Pr_{\mathrm{H}_1}(S_n \leq r)$ (probability of accepting H_0 while H_1 is true), respectively. The p-value of the test is given by $\Pr_{\mathrm{H}_0}(S_n > s_n^{(\mathrm{obs})})$ (probability under H_0 of observing a statistic beyond the one actually observed in the data), where $s_n^{(\mathrm{obs})}$ stands for the statistic actually observed in the dataset at hand. Given some level of significance $1 - \alpha \in (0, 1)$, one seeks the least critical value r such that $\Pr_{\mathrm{H}_0}(S_n > r) \leq \alpha$. If it is known that S_n obeys a standard normal distribution then $r = q_{1-\alpha}$ and the test can be carried out. Otherwise, as for confidence intervals, one wants to establish that S_n

is approximately normal and get approximate critical values.

To establish that $\widehat{\psi}$ and S_n are approximately normal, one uses limit theorems of asymptotic statistics. The obtained approximate confidence intervals and p-values are then sometimes said to be asymptotic. The approximation improves as the sample size n increases. Such limit theorems are established in the next section.

Asymptotic normality and confidence intervals

It is established below that the sensitivity index estimators converge to multivariate normal distributions as the sample size grows. Based on this result, various tests are then built.

Proposition 1. *It holds that*

(i) $\sqrt{n}(\widehat{\tau} - \tau) \xrightarrow{\mathrm{d}} \mathrm{N}(0, T)$, *where* T *is the variance-covariance matrix with elements* $T(A, B) = \mathrm{Cov}((f(X) - f(X^{\setminus A}))^2/2, (f(X) - f(X^{\setminus B}))^2/2)$, $A, B \subset D$;

(ii) $\sqrt{n}(\widehat{\tau}^* - \tau^*) \xrightarrow{\mathrm{d}} \mathrm{N}(0, T^*)$, *where* $T^* = LTL^\top$;

(iii) $\sqrt{n}(\widehat{\sigma}^* - \sigma^*) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \Sigma^*)$, *where* $\Sigma^* = M^* T^* M^{*\top}$.

(iv) $\sqrt{n}(\widehat{\theta} - \theta) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \Theta)$, *where* $\Theta = N^* \Sigma^* N^{*\top}$.

The proof of Proposition 1 follows from a direct application of the central limit theorem and the delta-method of classical mathematical statistics; see e.g. van der Vaart (1998). Note that the matrix T can be estimated by taking empirical covariances. Let us denote by $\widehat{T}$ the estimate of T. From $\widehat{T}$, estimates $\widehat{T}^*$, $\widehat{\Sigma}^*$ and $\widehat{\Theta}$ of T^*, Σ^* and Θ, respectively, can be constructed by applying the formulas in Proposition 1. 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. Proposition 1 allows us to build asymptotic confidence intervals as explained above.

Hypothesis testing

Proposition 1 allows us to build various tests. If $\beta \in (0,1)$, let q_β denote the quantile of order β of a standard normal distribution. Let Φ denote the cumulative distribution function of a standard normal distribution.

Nullity of a given sensitivity index. Recall that ψ denotes any of τ, τ^*, σ^*, θ. A formal test is given by

$$\mathrm{H}_0 : \psi(A) = 0 \text{ versus } \mathrm{H}_1 : \psi(A) \neq 0. \tag{17}$$

To carry out this test, we can take the statistic $S_n = \sqrt{n}\widehat{\psi}(A)/\sqrt{\widehat{\Psi}(A,A)}$ and the critical value $r = q_{1-\alpha/2}$. The p-value is given by $1 - \Phi(S_n)$.

Presence of interactions between a given input and the other ones. It is well known that $\tau^*(\{j\}) \leq \tau(\{j\})$ and that $\tau(\{j\}) - \tau^*(\{j\})$ quantifies the strength of interactions involving input $\{j\}$. A formal test is given by

$$\mathrm{H}_0 : \tau^*(\{j\}) = \tau(\{j\}) \text{ versus } \mathrm{H}_1 : \tau^*(\{j\}) < \tau(\{j\}). \tag{18}$$

To carry out this test, we can take the statistic $S_n = \sqrt{n}a_j^\top \widehat{\tau}/\sqrt{a_j^\top \widehat{T} a_j}$, where a_j is the vector of length 2^d comprising only zeros, except at the positions corresponding to $\{j\}$, $D \setminus \{j\}$ and D where the components are equal to 1, 1 and -1, respectively. The critical value is given by $r = q_{1-\alpha}$. The p-value is given by $1 - \Phi(S_n)$.

Contribution of a given set of Sobol indices to the total variance. Suppose we want to test whether a given set of Sobol indices contributes to less than a certain share of the total variance. The above test can be formulated as

$$\begin{aligned} &\mathrm{H}_0 : \sum_{i=1}^{k} \sigma^*(A_i) \leq (1-\eta) \operatorname{Var} f(X), \text{ versus} \\ &\mathrm{H}_1 : \sum_{i=1}^{k} \sigma^*(A_i) > (1-\eta) \operatorname{Var} f(X), \end{aligned} \tag{19}$$

where $\eta \in (0,1)$ is some desired threshold and $\sigma^*(A_1), \ldots, \sigma^*(A_k)$ are chosen Sobol indices. Here $A_1, \ldots A_k$ can be any subsets of D (except the empty set). To carry out this test, proceed as follows. Let j_i denote the position of $\sigma^*(A_i)$ in the vector σ^*. Define the vector $c = (c_1, \ldots, c_{2^d-1})$, where $c_{j_i} = \eta/(1-\eta)$ for $i = 1, \ldots, k$, and -1 otherwise. Take the test statistic and the critical value

$$S_n = \frac{\sqrt{n} c^\top \widehat{\sigma}^*}{\sqrt{c^\top \widehat{\Sigma}^* c}}, \quad r = q_{1-\alpha}, \tag{20}$$

respectively. Note that 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, $S_n = S_n(\eta)$ is an increasing function of η on $[0,1]$.

Proposition 2. *With the given S_n and r, it holds that:*

(i) The tests (17) and (18) satisfy $\lim_{n\to\infty} \Pr_{\mathrm{H}_0}(S_n > r) = \alpha$.

(ii) The test (19) satisfies $\lim_{n\to\infty} \Pr_{\mathrm{H}_0}(S_n > r) \leq \alpha$.

The first statement of Proposition 2 is a direct application of standard theory. The second statement is proved in the Supplementary Material.

Examples of applications of the above tests are given in Section 5.

4 Computational aspects

Estimators and their asymptotic variance-covariance matrices have been defined in Section 3. However, it may not be immediately apparent from the given formulas how to actually perform the computations, and at what expense. In this section, we assume that the sample of model outputs has already been generated. Thus, in the following *computation expense* is understood as the expense of the inference method *per se*, and not the expense of the sensitivity analysis as a whole. A discussion about the computation expense of the sensitivity analysis as a whole is given in Section 6.

In what follows, computation expense will be synonymous with number of arithmetic operations. Also, we shall use the "Big Theta" notation (Knuth, 1997) : two functions f and g of d satisfy $g(d) = \Theta(f(d))$ if $|g(d)/f(d)|$ remains bounded from above and away from zero, as d goes to infinity. For instance, the computation expense of $\widehat{\tau}$ is $\Theta(n2^d)$, meaning that it is proportional to $n2^d$ as $d \to \infty$. Finally, if M is a matrix, let $\mathrm{nnz}(M)$ denote the number of nonzero elements of M.

The computations and their implementation is greatly facilitated once the elements of the power set have been arranged in some specific 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, most significant bit first. 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 $d = 3$.

With the above arrangement, "building" the matrix M^* in (15) becomes straightforward, as shows Proposition 3.

Proposition 3. *If the components of τ^* and σ^* are arranged as described above then the*

matrix M^* in (15) is a d-fold Kronecker autoproduct of the 2×2 matrix

$$M_1^* = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}. \tag{21}$$

In other words,

$$M^* = \otimes^d M_1^* = M_1^* \otimes \cdots \otimes M_1^*, \tag{22}$$

where $\otimes$ denotes the Kronecker product.

A proof of (22) is given in the Supplementary Material. See also Grabisch et al. (2000). To compute $\sigma^* = M^*\tau^*$ from τ^*, one can simply perform the product between the sparse matrix M^* and the vector τ^*. Indeed, the matrix M^* is sparse since $\mathrm{nnz}(M^*) = \mathrm{nnz}(\otimes^d M_1^*) = \mathrm{nnz}(M_1^*)^d = 3^d$; see also, e.g. La Haye (2009). The expense of computing σ^* from τ^* is then $\Theta(\mathrm{nnz}(M^*)) = \Theta(3^d)$. Similarly, the matrix $\Sigma^* = M^*T^*M^{*\top}$ defined in Proposition 1 can be computed by multiplying T^* to the left by M^* and then multiplying the result to the right by $M^{*\top}$. By the same token, the expense for computing Σ^* from T^* is $\Theta(6^d)$.

It turns out that there are—at least in principle—less expensive algorithms to compute σ^* from τ^* and Σ^* from T^*. These algorithms are presented below.

4.1 The fast Möbius transform

The so-called fast Möbius transform can be used to compute σ^* from τ^*. Suppose momentarily that τ^* is an arbitrary vector of size 2^d and M_1^* is an arbitrary matrix of size 2×2. Let $\sigma^{*(0)}, \sigma^{*(1)}, \ldots, \sigma^{*(d)}$ be defined through the recurrence relations

$$\begin{cases} \sigma^{*(0)} &= \tau^* \\ \sigma^{*(j)} &= (I_{2^{d-j}} \otimes M_1^* \otimes I_{2^{j-1}})\sigma^{*(j-1)}, \quad j = 1 \ldots d, \end{cases} \tag{23}$$

where $I_{2^{d-j}}$ 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 (23) define an algorithm in which the vector $\sigma^{*(j)}$ is first initialized to the vector τ^* and then updated at each iteration $j = 1, \ldots, d$.

Proposition 4. *If τ^* is a vector of size 2^d and M_1^* is a nonnull matrix of size 2×2 then the last update of (23) is equal to*

$$\sigma^{*(d)} = (\otimes^d M_1^*)\tau^*. \tag{24}$$

Moreover, the total number of arithmetic operations needed to reach $\sigma^{*(d)}$ is $\Theta(d2^d)$.

A proof of Proposition 4 is given in the Supplementary Material. See also Björklund et al. (2007); Dayar and Orhan (2015); Fernandes et al. (1998); Grabisch et al. (2000); Kennes (1992); Yates (1937).

It follows from Proposition 4 that the computation of $(\otimes^d M_1^*)\tau^*$ is less expensive with (23) than with standard sparse matrix multiplication as soon as $\mathrm{nnz}(M_1^*) > 2$, because $\mathrm{nnz}(\otimes^d M_1^*) = \mathrm{nnz}(M_1^*)^d$. In terms of computer memory usage, algorithm (23) is preferable because it does not need to store the whole matrix $\otimes^d M_1^*$. The efficiency of (23) in terms of execution time is depicted in Figure 1 (left panel).

Remember that in Proposition 4 the vector τ^* and the matrix M_1^* were arbitrary. Now take τ^* to be the actual dual index vector with components arranged as in Table 1 and M_1^* the matrix (21). Then, in view of (15) and Proposition 4, the vector $\sigma^{*(d)}$ is the Sobol index vector with components arranged as in Table 1. Note that instead of using matricial operations, the recurrence relations (23) can be implemented directly by "bit-shift" manipulations as follows. Let us introduce some notation first. If σ^* denotes a vector of size 2^d then let $\sigma^*(a_{i,1} \cdots a_{i,d})$ denote its component corresponding to the subset of D represented by the Boolean vector $(a_{i,1} \cdots a_{i,d})$. The components of σ^* are assumed to be arranged in the lexicographical order of the Boolean vectors $(a_{i,1} \cdots a_{i,d})$, as in Table 1. To compute $\sigma^{*(d)}$ in (23), proceed as follows:

for each $j = 1, \ldots, d$

for each $i = 1, \ldots, 2^d$

if $a_{i,d-j+1} = 1$

$$\sigma^{*(j)}(a_{i,1} \cdots a_{i,d}) \leftarrow \sigma^{*(j-1)}(a_{i,1} \cdots a_{i,d-j}, 1, a_{i,d-j+2} \cdots a_{i,d}) - \sigma^{*(j-1)}(a_{i,1} \cdots a_{i,d-j}, 0, a_{i,d-j+2} \cdots a_{i,d})$$

else

$$\sigma^{*(j)}(a_{i,1} \cdots a_{i,d}) \leftarrow \sigma^{*(j-1)}(a_{i,1} \cdots a_{i,d}).$$

A flow diagram of algorithm (23) with $d = 3$ and M_1^* as in (21) is depicted in supplementary Figure 1. A Matlab implementation is available as a supplementary file (see Supplementary material). This "bit-shift" implementation has the additional advantage of saving memory usage, as there is no need to store the matrices $(I_{2^{d-j}} \otimes M_1^* \otimes I_{2^{j-1}})$ in (23) anymore.

Remark 4. *Algorithm (23) with M_1^* the absolute value of (21) is refered to as the* fast Möbius transform *in Björklund et al. (2007); Kennes (1992). 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 Björklund et al. (2007); Kennes (1992). The fast Möbius transform is then a "fast" algorithm to compute the Möbius transform from its inverse (in a similar way than the fast Fourier transform computes the Fourier transform rapidly, hence the name). Computing the inverse from the Möbius transform (that is, algorithm* (23) *with M_1^* as in* (21)*) is, in essence, the same algorithm. An early version was given in Yates (1937).*

4.2 The extended fast Möbius transform

Along with the sensitivity index estimates, we need to compute their asymptotic variance-covariance matrices, too. Algorithm (23) can be extended to compute Σ^* from T^* the same way σ^* was computed from τ^*. As before suppose momentarily that T^* is an arbitrary

matrix of size $2^d \times 2^d$ and M_1^* is an arbitrary matrix of size 2×2. Let $\Sigma^{*(0)}, \Sigma^{*(1)}, \ldots, \Sigma^{*(d)}$ be defined through the recurrence relations

$$\begin{cases} \Sigma^{*(0)} &= T^* \\ \Sigma^{*(j)} &= (I_{2^{d-j}} \otimes M_1^* \otimes I_{2^{j-1}})\Sigma^{*(j-1)}(I_{2^{d-j}} \otimes M_1^{*\top} \otimes I_{2^{j-1}}) \quad (j = 1, \ldots, d), \end{cases} \tag{25}$$

The difference between (23) and (25) is that we now consider a matrix which is updated with both left and right multiplications.

A proposition similar to Proposition 4 holds.

Proposition 5. *If T^* is an arbitrary matrix of size $2^d \times 2^d$ and M_1^* is an arbitrary nonnull matrix of size 2×2 then the last update of (25) is equal to*

$$\Sigma^{*(d)} = (\otimes^d M_1^*)T^*(\otimes^d M_1^{*\top}). \tag{26}$$

Moreover, the total number of arithmetic operations needed to reach $\Sigma^{*(d)}$ is $\Theta(d4^d)$.

The proof of Proposition 5 is an easy adaptation of the proof of Proposition 4 and is not given. The expense $\Theta(d4^d)$ is to be compared with $\Theta(6^d)$ that was obtained with sparse matrix multiplication. The effect on computation times are depicted in Figure 1 (right panel).

As in Section 4.1, it now suffices to take T^* to be the variance-covariance matrix defined in Proposition 1 with its rows and columns arranged as in Table 1 and M_1^* the matrix (21). We then get that $\Sigma^{*(d)} = (\otimes^d M_1^*)T^*(\otimes^d M_1^{*\top}) = (\otimes^d M_1^*)T^*(\otimes^d M_1^*)^\top = M^*T^*M^{*\top}$ coincides with the matrix Σ^* of Proposition 1. As for (23), a direct implementation is possible but there is a "bit-shift" version which is more efficient. After initialization, proceed in two steps:

S1. For each $j = 1, \ldots, d$

for each $i = 1, \ldots, 2^d$

if $a_{i,d-j+1} = 1$

$$\text{row}^{(j)}(a_{i,1}\cdots a_{i,d}) \leftarrow \text{row}^{(j-1)}(a_{i,1}\cdots a_{i,d-j}1a_{i,d-j+2}\cdots a_{i,d}) - \text{row}^{(j-1)}(a_{i,1}\cdots a_{i,d-j}0a_{i,d-j+2}\cdots a_{i,d})$$

else

$$\text{row}^{(j)}(a_{i,1}\cdots a_{i,d}) \leftarrow \text{row}^{(j-1)}(a_{i,1}\cdots a_{i,d});$$

here $\text{row}^{(j)}(a_{i,1}\cdots a_{i,d})$ stands for the row of $\Sigma^{*(j)}$ that corresponds to the index $(a_{i,1}\cdots a_{i,d})$.

S2. Repeat S1 above but with "col" in place of "row", where $\text{col}^{(j)}(a_{i,1}\cdots a_{i,d})$ stands for the column of $\Sigma^{*(j)}$ that corresponds to the index $(a_{i,1}\cdots a_{i,d})$.

The above algorithm coincides with (25) when M_1^* is the matrix (21). As above, this implementation saves memory usage as there is no need to store the matrices in (25) anymore. A Matlab implementation is available as a supplementary file (see Supplementary material).

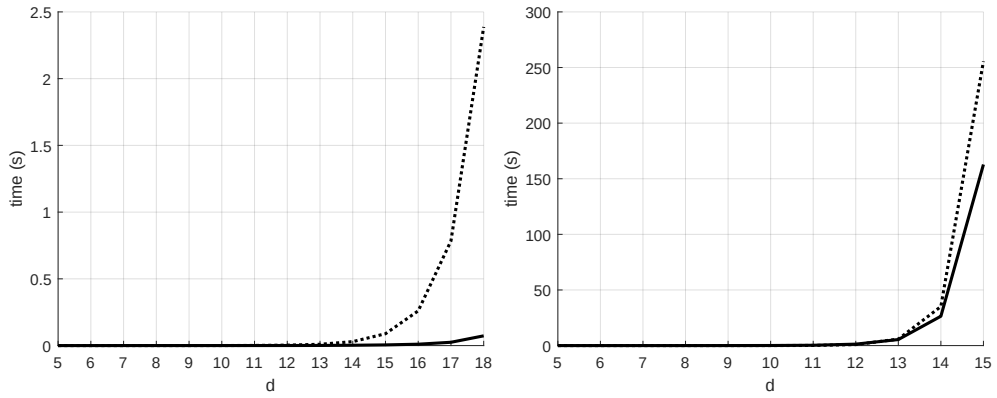

Figure 1: Execution times of the fast Möbius transform (straight line) and straightforward sparse matrix multiplications (dotted line) for the computation of Sobol indices σ^* from τ^*. Left: only the indices are computed. Right: computations include $2^d \times 2^d$ covariance matrices Σ^* (computations were carried out in MATLAB).

5 Application to the sensitivity analysis of two dynamical systems

In this section we apply the inference method to analyze the sensitivity of two different dynamical systems, both with an application in life sciences. The first one is a discrete system, more precisely a stochastic Boolean model of a cellular regulatory network, and the second one is a system of deterministic ordinary differential equations (ODE) modeling some population dynamics. In both cases, after identifying d key parameters, a $n \times 2^d$ dataset is simulated according to part 3.1. The different sets of sensitivity indices are then estimated and are used, both qualitatively and quantitatively, to gain insight on the systems' dynamical behavior. All the following computations were performed with MATLAB (R2023b).

5.1 A Boolean network modeling cell fate decision

First we consider a Boolean network, originally proposed by Calzone et al. (2010), modeling 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: apoptosis, which is the classical programmed cell death, and necroptosis, which is a programmed death sharing apoptotic regulators yet presenting necrotic features (thereafter, this phenotype will be abusively denoted by "necrosis"). The mathematical model of Calzone et al. (2010) consists in an asynchronous Boolean model modeling the main regulatory network responsible for the decision between three phenotypes: apoptosis, necrosis and survival. In such networks, 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 an absorbing Markov chain. More details on the construction of asynchronous Boolean networks and their translation into Markov chains can be found for instance in Tournier and Chaves (2009). With such a model, it becomes possible for instance to compute the probability to enter apoptosis, necrosis or survival starting from a given initial condition.

A critical point in the interpretation of the model's outputs lies in the choice of the transition probabilities and by default, all successors of a state are assumed equiprobable. In reality those probabilities are linked to the relative speeds of the regulatory processes involved. Defining them from biological experiments for such a complex network, with multiple and heterogeneous intertwined pathways, is out of reach. This makes a global sensitivity analysis particularly appealing, to help detect which regulatory mechanisms, or combinations of regulatory mechanisms, have the most impact on the dynamics, and are therefore priority targets for further investigation.

A first sensitivity analysis of the model was performed by Zinovyev et al. (2012). The idea was to affect a positive weight to each Boolean variable: a weight greater than 1 means that processes involving the variable are rapid and the transitions updating the variable are favored, while a weight lower than 1 means that the processes are slow and the transitions are hindered. Mathematically, this boils down to parametrize the transition matrix of the Markov chain. Since the network is asynchronous, each nonnull entry corresponds to a switch of a single variable, and now contains the (normalized) weight associated with that variable. A local sensitivity analysis around the reference vector $(1, \ldots, 1)$ was performed by using a "one-at-a-time" approach, each weight varying independently around its reference value 1.

Here, we refine this approach by considering two types of weights instead of one: one

type controls the variable's activation (from 0 to 1) and one type its deactivation (from 1 to 0). This comes from the basic observation that in cell biology, production and degradation of a given compound may have very different time scales. Also, we now perform a global sensitivity analysis, which is preferable because it does not give any particular importance to the reference vector $(1, \ldots, 1)$. Six variables central to the decision process are chosen: C8 (caspase 8), RIP1, NFkB, cIAP, MOMP and MPT. This leads to $d = 2 \times 6 = 12$ inputs, corresponding to the 12 weights that parametrize the transition matrix of the Markov chain. The reader is referred to Calzone et al. (2010) for a comprehensive description of these biological variables and the regulatory processes they are involved in.

A dataset of $n2^d = 1000 \times 4096$ model simulations was generated according to (11). The inputs were drawn independently from log-uniform distributions. See Table 1 in the Supplementary material for details. Each simulation consists in the computation of three types of outputs: the probabilities to reach the three expected phenotypes (apoptosis, necrosis and survival) starting from a given initial condition. The detailed procedure is described in the Supplementary material. For each type of output, total and Sobol indices as well as Shapley effects are then estimated as described in Section 3. The estimates and their asymptotic variance-covariance matrices are computed with the "bit-shift" versions of the fast Möbius transform and the extended fast Möbius transform algorithms, respectively, given in Section 4. The choice of the size n of the Monte Carlo sample was arbitrary, but the obtained confidence intervals are narrow enough (see Figure 2) to draw reliable inferences. (Remember that to halve the standard error, four times more simulations are needed.)

The estimates are depicted in Figure 2. The first observation that can be made is that the "necrosis" phenotype 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 3. By inspecting highest Sobol index estimates in "apoptosis" and "survival" phenotypes (Fig. 2, middle row), we find that

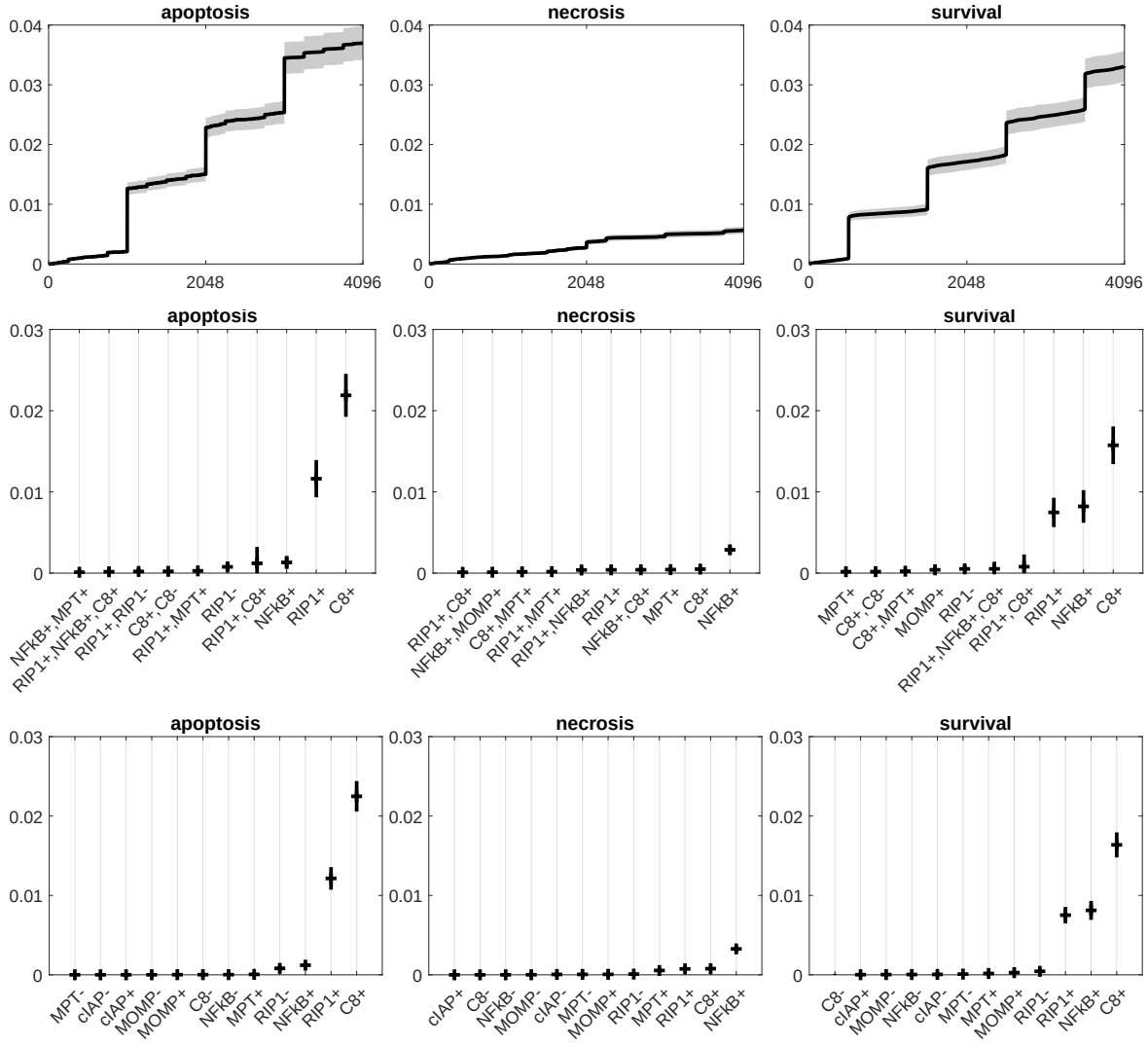

Figure 2: Global sensitivity analysis of the cell fate Boolean model by Calzone et al. (2010). The first row shows total index estimates for all $2^d = 4096$ combinations of inputs, arranged in ascending order. The second row shows the 10 highest Sobol index estimates in ascending order; the third row shows Shapley effect estimates, also in ascending order. Each estimate is drawn within its 95% confidence interval.

the three most important inputs are C8+, RIP1+ and NFkB+, as was originally observed by Zinovyev et al. (2012). To go beyond this qualitative observation, we apply Test (19) above and find these three inputs contribute to more than 85% of the total variance in each case, with respective p-values 2.5×10^{-3} and 9.8×10^{-3}. Although the combination $A = \{\text{C8+}, \text{RIP1+}\}$ stands out as the fourth highest Sobol index in both cases, statistical inference invalidates this result. Indeed, testing the nullity of $\sigma(\{\text{C8+}, \text{RIP1+}\})^*$ (test (17) above) leads to high p-values: $p > 0.12$ for apoptosis and $p > 0.15$ for survival, indicating the uncertainty on the estimation is too high to be conclusive. The interpretation of the "necrosis" phenotype is slightly more delicate. If on a qualitative level input NFkB+ stands out as the sole significant one, it can only be shown to contribute to 45% of total variance with high confidence (p-value < 0.05). To explain more than 85% of the variance with p-value < 0.05, one has to consider the seven highest main or interaction effects $\sigma^*(\{\text{NFkB+}\})$, $\sigma^*(\{\text{C8+}\})$, $\sigma^*(\{\text{MPT+}\})$, $\sigma^*(\{\text{NFkB+}, \text{C8+}\})$, $\sigma^*(\{\text{RIP1+}\})$, $\sigma^*(\{\text{NFkB+}, \text{RIP1+}\})$ and $\sigma^*(\{\text{RIP1+}, \text{MPT+}\})$. This indicates that, although the variability of the phenotype is globally lower (Fig. 3), it is much more spread out than for the other two phenotypes. The choices between apoptosis and non-apoptosis, and between survival and non-survival are less complex, as they almost exclusively depend on a few key inputs, namely C8+ and RIP1+. Looking at the bottom row of Figure 2, very similar interpretations can be drawn from Shapley effect estimates, although the seemingly complex interactions between inputs for necrosis are globally missed.

Finally, it is interesting to note that total index estimates (Fig. 2, top row) are 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

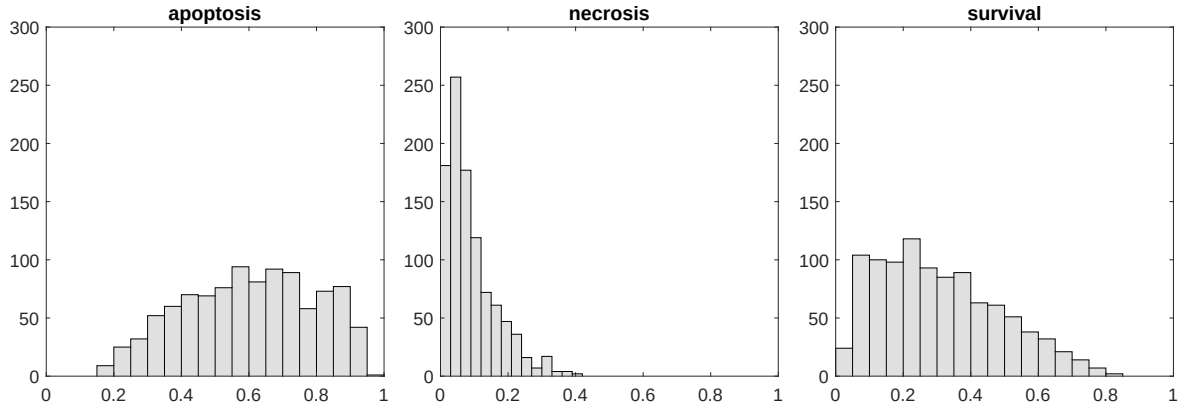

Figure 3: Histograms of 1000 independent model simulations.

indices is currently investigated.

5.2 A differential system modeling population dynamics

We consider now a system of five nonlinear ordinary differential equations (ODE) modeling the dynamics of infection of a bacterial population of *E. coli* by the bacteriophage λ in the mouse gut, originally proposed by De Paepe et al. (2016). Lambda is a temperate phage, meaning that its infection of a bacterium may lead to two opposite outcomes: either it enters the "lysogenic" phase, where the virus disappears and its DNA resides within the bacterium in a dormant phase, the bacterial host is then called a lysogen; or it enters the "lytic" phase, where it uses the bacterium's machinery to replicate itself, eventually leading to the destruction of the host and the release of new, free viruses. Lysogens continue to grow and survive normally, until a future change in external conditions may induce the lytic phase, leading to the bacterium's demise and the subsequent release of free viruses.

In the model, two bacterial populations exist initially: L denotes lysogens, containing the virus in a dormant phase, and S denotes "susceptible", uninfected bacteria. Newly lysogenized susceptible are denoted by S^l, in order to distinguish them from originally resident lysogens L. Once infected, susceptible bacteria enter lytic phase (variable Q) with

probability $(1-g)$. As for lysogens L and S^l, they enter lytic phase at constant rate x. After some time, lytic bacteria Q die and release free viruses, denoted by variable V. The system is given by the five ODE:

$$\begin{cases} dL/dt &= r(1-N/k)L - xL - dL, \\ dS/dt &= r(1-N/k)S - aVS - dS, \\ dS^l/dt &= r(1-N/k)S^l - xS^l + gaVS - dS^l, \\ dQ/dt &= x(L+S^l) + (1-g)aVS - lQ - dQ, \\ dV/dt &= ylQ - aV(N+Q) - dV, \end{cases} \tag{27}$$

where N denotes total bacterial population: $N = L + S + S^l$. In total, the system contains eight parameters, denoted by small letters and listed in supplementary Table 2. Viable bacterial populations L, S and S^l all grow according to the same logistic term with parameters r and d, while lytic bacteria Q are supposed not to grow anymore. Besides induction rate x and lysogenisation probability g, key infection parameters are the burst size y, which is the average number of released free viruses after lysis, and the adsorption constant a, governing the attachment of a free virus to a bacterial cell. A more detailed description of the parameters can be found in the original paper by De Paepe et al. (2016) and its supplementary material, along with an extensive presentation of the equations and the modeling choices behind them.

A mathematical analysis of the system's equilibria was carried out in De Paepe et al. (2016), indicating four different steady states, characterized thereafter by the eventual extinction or persistence of bacterial populations L and S:

- A "wash-out" equilibrium, where all microbial populations vanish.
- An equilibrium where $L = L^* > 0$ is maintained while $S = 0$.
- An equilibrium where $S = S^* > 0$ is maintained while $L = 0$.
- Finally, an equilibrium where both $L = L^* > 0$ and $S = S^* > 0$ coexist.

In this mathematical analysis parameters were treated symbolically, leading to explicit formulas for the equilibria as well as explicit relations between parameters determining equilibria's reachability. For conciseness, those explicit formulas are not reproduced here but can be found in De Paepe et al. (2016).

While this type of symbolic computations is relatively classical to analyze such systems, it is worth noting that it can rapidly become cumbersome, even impossible if the dimension is too high or if the system contains highly nonlinear terms. Here we propose to use a global sensitivity analysis as an alternative to such calculations. Although the output will necessarily be less informative than explicit formulas, in the following we show that global sensitivity indices capture interesting properties and provide valuable information about the effects of the parameters on the dynamics. Such analyses can thus become particularly helpful in the analysis of complex and large ODE systems, pointing for instance to specific groups of parameters that can be further investigated with dedicated tools.

We consider the system's eight parameters, to which we add the initial values of populations L and S, denoted by L_0 and S_0, leading to $d = 10$ inputs. A dataset of $n2^d = 4987 \times 1024$ model simulations was then generated, according to (11). The inputs were drawn independently from uniform or log-uniform distributions. See Table 2 in the Supplementary material for details. Each simulation is carried out by the matlab routine ode45, until the steady state is reached. The detailed procedure can be found in the Supplementary material. As before total, Sobol and Shapley indices were estimated from this dataset as in Section 3 and the computations were done with the Möbius transform and extended Möbius transform algorithms presented in Section 4. The size n was again arbitrary and initially set to 5000 but some simulations had to be discarded because of numerical anomalies (see Supplementary for more information). In view of the obtained confidence intervals, it was assessed that no more simulations was needed.

Figure 4 presents the ten highest Sobol index estimates. A rapid inspection of those

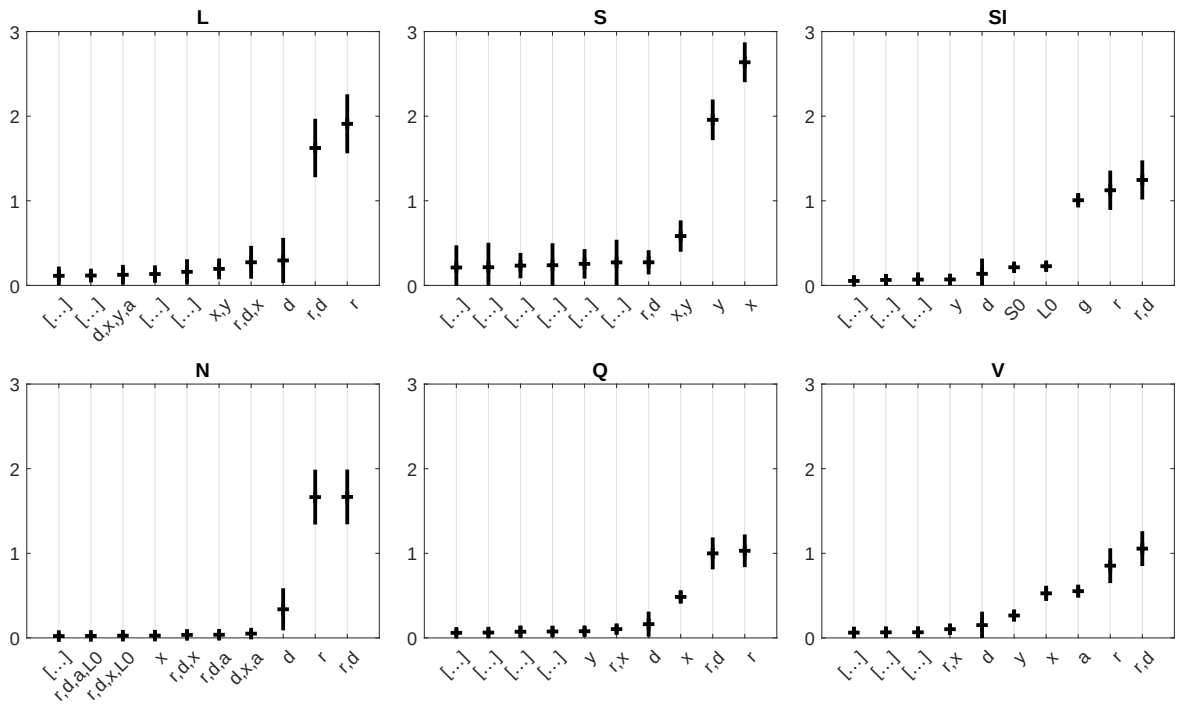

Figure 4: Global sensitivity analysis of the differential model by De Paepe et al. (2016). Graphs show the 10 highest Sobol index estimates for each of the five coordinates of the steady state of system (27) and the total bacterial population $N = L + S + S^l$. For readability, interactions are omitted when they contain more than 6 inputs.

indices shows that L is mainly impacted by parameters r and d, while S is mainly impacted by x and y. More precisely, the abundance of L is mostly dependent on the interaction between r and d, emphasized by the significantly high value of $\sigma^*(\{r, d\})$. This was observed in De Paepe et al. (2016), as the first and second equilibria of the system are mainly separated according to the relative position of r and d. Parameters r and d mainly govern bacterial growth, while x and y mainly govern viral infection, therefore the fact that susceptible are more sensitive to infection than lysogens is relatively unsurprising. However, the sensitivity analysis reveals a higher importance of x and y with respect to other viral parameters such as the adsorption constant a, for instance. This was rather unexpected.

Additional observations can be made about Figure 4. First, the viral population V depends on more parameters that bacterial populations: parameters r and d suffice to contribute to more than 80% of the variance of N (Test (19) with the two highest Sobol indices $\sigma^*(\{r, d\})$ and $\sigma^*(\{r\})$ returned a p-value $< 1.5 \times 10^{-3}$), whereas five parameters r, d, a, x and y are necessary to contribute to more than 80% of the variance of V (Test (19) with the five highest Sobol indices $\sigma^*(\{r, d\})$, $\sigma^*(\{r\})$, $\sigma^*(\{a\})$, $\sigma^*(\{x\})$ and $\sigma^*(\{y\})$ returned a p-value $< 1.5 \times 10^{-2}$). Interestingly, this phenomenon can be observed when looking at the explicit formulas in the original paper. Second, parameters S_0 and L_0, which are part of the initial condition, only appear with significant nonzero values for variable S^l, newly lysogenized bacteria. Usually in a non-degenerate ODE system, this type of dependence should not appear: equilibria are generally independent of the initial condition. Nevertheless in the original paper, a singular property of the system was noted: rewriting (27) in the system of variables (N, S, S^l, Q, V) implies that S^l does not influence the rest of the system and can be omitted, thus dropping the dimension to 4 instead of 5. This implies that no explicit formula can be retrieved for the equilibrium of S^l, but instead for the sum $L + S^l$. In the sensitivity analysis, the significant effect of the initial condition on the steady state of S^l is a hint that such phenomenon may occur. Those two observations: multiple

dependencies of variable V and dependence of S^l with respect to initial conditions, illustrate the interest of using a global sensitivity analysis as a first step to study this type of dynamical systems. It becomes particularly interesting for systems where explicit formulas are unavailable.

Contrary to the previous Boolean model where interactions of inputs were relatively insignificant, here they are more relevant. We noted earlier the significance and importance of the interaction effects $\{r, d\}$ or $\{x, y\}$. As a matter of fact, this system seems to exhibit a generally high influence of interactions between parameters. To quantify this, we implemented Test (18) described above. The results are presented in Figure 5. If for a singleton $\{i\}$, one observes a significant difference $\tau(\{i\}) - \tau^*(\{i\})$ (grey area of the bar), then one concludes that the influence of the input $\{i\}$ is strongly dependent on interactions with other inputs. This is notably the case for parameters r and d, which we already know from Figure 4. To a lesser extent, interaction effects are also important for other parameters as well, see Figure 5. A notable counterexample is the effect of parameter g for variable S^l. This parameter denotes the probability of lysogenisation for an newly infected bacterium. The fact that it has a high effect on the population of new lysogens is therefore expected; however, statistical testing indicates this effect is relatively isolated from other parameters, which is a new result.

6 Discussion

An inference method for global sensitivity analysis was designed and practically tested on two dynamical systems. 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: it

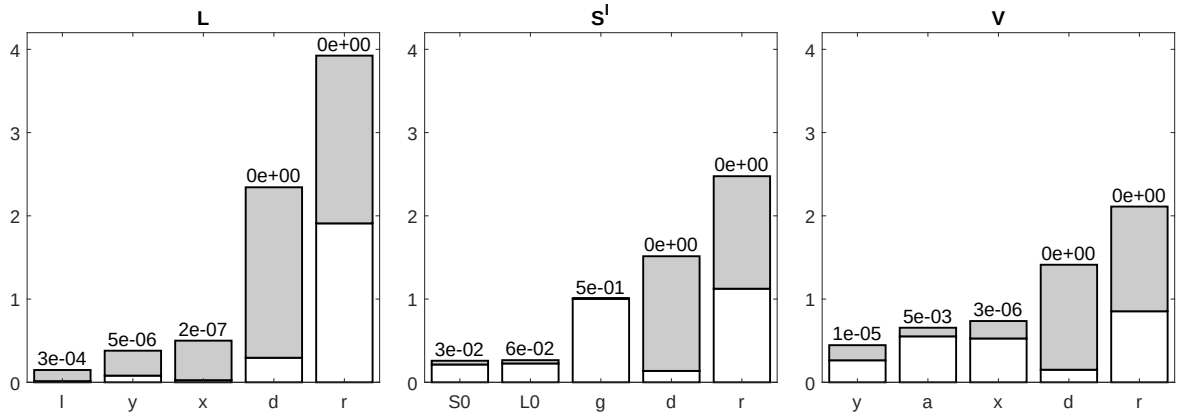

Figure 5: Testing the importance of interactions between inputs in System (27). Bars compare $\tau^*(\{i\})$ (white) and $\tau(\{i\})$ (white+grey), for the five highest $\tau(\{i\})$. The values above the bars indicate the p-values of test (18).

performs the inferences for everything thanks to the established asymptotic normality of the estimators. In particular, a procedure to test the contribution of an arbitrary number of Sobol indices of arbitrary orders was designed. To compute the estimates and their asymptotic variance-covariance matrices, the fast Möbius transform and the extended fast Möbius transform algorithms were proposed and implemented. 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.

Questions remain, however. Potential limitations and extensions are discussed below.

Computational aspects. It was seen in Section 4 that the fast Möbius transform and the extended fast Möbius transform algorithms can compute σ^* from τ^* with expense $\Theta(d2^d)$ instead of $\Theta(3^d)$ and Σ^* from T^* with expense $\Theta(d4^d)$ instead of $\Theta(6^d)$, respectively. These gains have to be put into perspective, however. Indeed, all in all, most of the computation time is likely to come from the generation of the dataset of the model outputs. Besides, even if we ignore the generation of this dataset, the computational savings observed in Figure 1 remain modest (mere seconds). This is in

part because of the high efficiency of sparse matrix manipulations in Matlab (a low level implementation of the fast Möbius transform, say in C or C++, certainly would improve the observed computation times) and in part because the difference between the numbers $\Theta(d2^d)$ and $\Theta(3^d)$ (or $\Theta(d4^d)$ and $\Theta(6^d)$) is most felt as $d \to \infty$, but going beyond about $d \approx 15$ becomes execessively expensive because of the manipulation of enormous $2^d \times 2^d$ matrices (in practice, we stopped at $d = 16$, involving > 30 Gigabyte matrices).

Dimension of the input space. Calculating the index estimates of all input combinations is feasible if the number of inputs is not too large. If the model runs fast, then perhaps up to a dozen inputs can be considered. Otherwise, owing to the sparsity-of-effects principle—aka Pareto principle (Becker, 2020)—it is reasonable to expect that a sufficiently large number of total indices will be close to zero. Thus, we may still apply the inference method proposed, but to a well-chosen selection of inputs. For instance, we might proceed as follows. First, we perform a screening experiment by calculating estimates of the total indices and their duals associated with individual inputs only. Second, we spot the inputs for which total index estimates are below some threshold and fix the corresponding inputs to some arbitrary values. Finally, we apply the inference method proposed to the remaining inputs only.

Global sensitivity analysis with dependent inputs. If the inputs are not mutually independent then the Sobol-Hoeffding decomposition presented in Section 2 no longer holds and hence Sobol indices are undefined. Interpretation of the right-hand side of (6) remains valid whether or not the inputs are mutually independent: in either case, it is the output variance induced by a random change in X_A, averaged over all the remaining inputs. Therefore, we can *define* $\tau(A)$ through (6) and, in turn, define and calculate Sobol indices through the identity $\tau(D) - \tau(A) = \tau^*(D \setminus A)$ and

the Möbius inversion formula (14), respectively. The inference method of Section 3 remains valid as long as we have an asymptotically normal estimator of the vector of total indices. The estimator given in Section 3.1 cannot be used anymore but one may use, adapt and/or extend other estimation methods (Da Veiga et al., 2023; Gamboa et al., 2022).

Improving the efficiency of estimators. Computational aspects and efficiency of estimators are intimately connected in sensitivity analysis: building more efficient estimators leads to computational savings, since we need less data for a given level of precision. To this aim, it might be interesting to consider adaptive estimation as in Broto et al. (2020). Quasi-Monte Carlo samples may also improve estimates, although controlling the statistical error is more difficult.

Acknowledgments

We thank two anonymous reviewers and an associate editor for very detailed and constructive comments that led to an improved version of this manuscript.

SUPPLEMENTARY MATERIAL

Supplementary file: Details about the generation of the datasets of Section 5, proofs of Proposition 2, 3 and 4, construction of two additional estimators of the total index, flow diagram of the fast Möbius transform. (.pdf file)

Code: MATLAB code to compute the estimates of the total, dual and Sobol indices, and the Shapley effects, as well as all the variance-covariance matrices, with the fast and extended fast Möbius transform algorithms. (.m files)

Dataset: A dataset of outputs from the cellular biology model. (.mat file)

References

Ball, P. (2023). *How life works: a user's guide to the new biology.* Chicago: The University of Chicago Press.

Becker, W. (2020). Metafunctions for benchmarking in sensitivity analysis. *Reliability Engineering & System Safety 204*, 107189.

Bénard, C., G. Biau, S. Da Veiga, and E. Scornet (2022). Shaff: Fast and consistent shapley effect estimates via random forests. In *International Conference on Artificial Intelligence and Statistics*, pp. 5563–5582. PMLR.

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

Boyle, E. A., Y. I. Li, and J. K. Pritchard (2017, June). An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell 169*(7), 1177–1186.

Broto, B., F. Bachoc, and M. Depecker (2020, January). Variance Reduction for Estimation of Shapley Effects and Adaptation to Unknown Input Distribution. *SIAM/ASA Journal on Uncertainty Quantification 8*(2), 693–716. Publisher: Society for Industrial and Applied Mathematics.

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

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

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

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

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

De Paepe, M., L. Tournier, E. Moncaut, O. Son, P. Langella, and M.-A. Petit (2016). Carriage of λ latent virus is costly for its bacterial host due to frequent reactivation in monoxenic mouse intestine. *PLoS Genetics 12*(2), e1005861.

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

Funaki, Y. (1998). Dual axiomatizations of solutions of cooperative games.

Gamboa, F., P. Gremaud, T. Klein, and A. Lagnoux (2022, November). Global sensitivity analysis: A novel generation of mighty estimators based on rank statistics. *Bernoulli 28*(4), 2345–2374. Publisher: Bernoulli Society for Mathematical Statistics and Probability.

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

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

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

Goda, T. (2021, September). A simple algorithm for global sensitivity analysis with Shapley effects. *Reliability Engineering & System Safety 213*, 107702.

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

Harsanyi, J. C. (1963). A Simplified Bargaining Model for the n-Person Cooperative Game. *International Economic Review 4*(2), 194–220. Publisher: [Economics Department of the University of Pennsylvania, Wiley, Institute of Social and Economic Research, Osaka University].

Herin, M., M. I. Idrissi, V. Chabridon, and B. Iooss (2022). Proportional marginal effects for global sensitivity analysis.

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

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

Il Idrissi, M., N. Bousquet, F. Gamboa, B. Iooss, and J.-M. Loubes (2023, November). On the coalitional decomposition of parameters of interest. *Comptes Rendus. Mathématique 361*, 1653–1662.

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

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

Iooss, B., 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. Mey-naoui, 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 (2020). *sensitivity*. R package version 1.23.0.

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

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

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

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

Knuth, D. E. (1997). *The art of computer programming*. Addison Wesley Longman. Volume 1. Fundamental Algorithms.

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

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

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

Oishi, T., M. Nakayama, T. Hokari, and Y. Funaki (2016, March). Duality and anti-duality in TU games applied to solutions, axioms, and axiomatizations. *Journal of Mathematical Economics 63*, 44–53.

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

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

Owen, A. B. and C. Prieur (2017, January). On Shapley Value for Measuring Importance of Dependent Inputs. *SIAM/ASA Journal on Uncertainty Quantification 5*(1), 986–1002. Publisher: Society for Industrial and Applied Mathematics.

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

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

Prieur, C. and S. Tarantola (2015). Variance-based sensitivity analysis: theory and estimation algorithms. In *Handbook of Uncertainty Quantification*, pp. 1–23. Springer.

Rota, G.-C. (1964). On the Foundations of Combinatorial Theory I. Theory of Mobius Functions. *Z. Wahrseheinlichkeitstheorie 2*, 340–368.

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

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

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

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

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

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

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

Sobol, I. M. and Y. L. Levitan (1999). On the use of variance reducing multipliers in monte carlo computations of a global sensitivity index. *Computer Physics Communications 117*(1), 52–61.

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

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

Sudret, B. (2008). Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety 93*, 964–979.

Tissot, J.-Y. and C. Prieur (2015). 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.

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

van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press.

Winter, E. (2002, January). Chapter 53 The Shapley value. In *Handbook of Game Theory with Economic Applications*, Volume 3, pp. 2025–2054. Elsevier.

Yates, F. (1937). *The design and analysis of factorial experiments*. Harpenden: Imperial Bureau of Soil Science.

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