



HAL
open science

Sensitivity indices for independent groups of variables

Baptiste Broto, François Bachoc, Marine Depecker, Jean-Marc Martinez

► **To cite this version:**

Baptiste Broto, François Bachoc, Marine Depecker, Jean-Marc Martinez. Sensitivity indices for independent groups of variables. 2018. hal-01680687v1

HAL Id: hal-01680687

<https://hal.science/hal-01680687v1>

Preprint submitted on 11 Jan 2018 (v1), last revised 11 Apr 2019 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sensitivity indices for independent groups of variables

Baptiste Broto¹, François Bachoc², Marine Depecker¹, and Jean-Marc Martinez³

¹CEA, LIST, Université Paris-Saclay, F-91120, Palaiseau, France

²Institut de Mathématiques de Toulouse, Université Paul Sabatier, F-31062 Toulouse, France

³CEA, DEN-STMF, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France

January 11, 2018

Abstract

In this paper, we study sensitivity indices in an additive model and for independent groups of variables. We show in this case that most of the Sobol indices are equal to zero and that Shapley effects can be estimated more efficiently. We then apply this study to Gaussian linear models and we provide an efficient algorithm to compute the theoretical sensitivity indices. In numerical experiments, we show that this algorithm compares favourably to other existing methods.

1 Introduction

In global sensitivity analysis, we consider a variable of interest Y which is a function of several variables X_1, \dots, X_p . We aim to associate a value to each input variable, that quantifies its impact on the output variable. We call these values "sensitivity indices". The first sensitivity indices for a general framework were introduced by Sobol in 1993 [Sob93] and are called "Sobol indices". They are based on the output variance. Since that date, many researchers have defined other sensitivity indices. For instance, we can find in [BHP16] a general framework for defining sensitivity indices based on variances, on densities, or on distributions. These sensitivity indices are very useful in many applications, for example in physics or in the industry. The most used sensitivity indices remain the Sobol indices.

Recently, Owen used the notion of Shapley value in order to define new variance-based sensitivity indices in [Owe14] and we call them "Shapley effects".

Shapley value originates from game theory in [Sha53]. This quantity can be useful in very different fields (see for example [MvLG⁺08] or [HI03]). Hence, there is an active strand of literature focusing on the computation of Shapley values ([CBSV16], [FWJ08]...) but they are adapted to a more general framework than sensitivity analysis. Only a few articles focus on this computation in the field of global sensitivity analysis (see [Owe14], [SNS16], [OP17], [IP17]).

In this paper, we focus on two popular variance-based sensitivity indices (Sobol indices and Shapley effects) when the input variables form independent groups and the model is additive. We first show an analytical formula on the variances of the conditional expectations. Then, we use this formula to prove that the sensitivity indices have much simpler expressions in this framework. In particular, we show that many Sobol indices are equal to zero and that the computation of the Shapley effects is much easier in this case than in the general framework. Finally, we provide an algorithm for evaluating these sensitivity indices in the linear Gaussian case, which is particularly efficient when the covariance matrix of the input Gaussian distribution is block diagonal.

The paper is organized as follows: in Section 2, we recall the definition of Sobol indices and Shapley effects and their properties. In Section 3, we give our main theoretical results about these sensitivity indices when there are independent groups of input variables. In Section 4, we focus on the linear Gaussian framework and give an algorithm to compute the variance-based sensitivity indices in this case. Then, we apply our theoretical results about the variance-based sensitivity indices to give an efficient computation when the covariance matrix is block diagonal. In Section 5, we highlight our results comparing these algorithms with the existing algorithm of [SNS16] designed to compute the Shapley effects. The proofs are postponed to the appendix.

2 Background and notations

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $E = \times_{i=1}^p E_i$ be endowed with the product σ -algebra $\mathcal{E} = \otimes_{i=1}^p \mathcal{E}_i$ and X be a random variable from Ω to E in L^2 . The $(X_i)_{i \in [1:p]}$ will be the input variables of the model.

Let $f : (E, \mathcal{E}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be in L^2 and $Y = f(X)$. We will call Y the output variable. We assume that $\text{Var}(Y) \neq 0$.

Let $[1 : p]$ be the set of all integers between 1 and p . If $u \subset [1 : p]$ and $x \in \mathbb{R}^p$, we will write $x_u := (x_i)_{i \in u}$.

For all $u \subset [1 : p]$, we write $V_u := \text{Var}(\mathbb{E}(Y|X_u))$. We can now define the Sobol indices (see [SCSo00]) for a group of variable X_u as:

$$S_u := \frac{1}{\text{Var}(Y)} \sum_{v \subset u} (-1)^{|u|-|v|} V_v, \quad (1)$$

where $|u|$ is the cardinal of u . This quantity assesses the impact of the subset of inputs X_u on the model output variability. These sensitivity indices were the first ones induced in [Sob93] for non-linear models. The Sobol indices are well defined and the sum of the Sobol indices over $u \subset [1 : p]$ is still equal

to one, even if the inputs are dependent. Moreover, these sensitivity indices are positive when the inputs are independent. However, this is not true in the general case, so that the Sobol indices are less amenable to interpretation with dependent input variables. Mara and a.l. suggested in [MTA15] an alternative definition for Sobol indices in the dependent case, allowing them to remain positive. Nevertheless, Iooss and a.l. exhibited in [IP17] an analysis supporting the use of Shapley effects as sensitivity indices when the inputs are dependent.

We can define the Shapley effects as in [Owe14] for a variable X_i as:

$$\eta_i := \frac{1}{p\text{Var}(Y)} \sum_{u \subset -i} \binom{p-1}{|u|}^{-1} (V_{u \cup \{i\}} - V_u) \quad (2)$$

where $-i$ is the set $[1 : p] \setminus \{i\}$.

The Shapley effects have interesting properties for global sensitivity analysis. Indeed, there is only one Shapley effect for each variable (contrary to the Sobol indices). Moreover, the sum of all the Shapley effects is equal to 1 (see [Owe14]) and all these values lie in $[0, 1]$ even with dependent inputs, which is very convenient for the interpretation of these sensitivity indices.

To conclude this part, we remark that if we replace Y by $Y - \mathbb{E}(Y)$, we do not change the values of $(V_u)_{u \subset [1:p]}$, and so on the Sobol indices and the Shapley effects. We then assume without loss of generality that the expectation of the output variable Y is equal to zero.

3 Variance-based sensitivity indice properties for independent block-additive models

In this section, we give theoretical results about variance-based sensitivity indices for additive models with independent groups of variables.

Let $\mathcal{C} = \{C_1, \dots, C_k\}$ be a partition of $[1 : p]$ such that the groups of random variables $(X_{C_j})_{j \in [1:k]}$ are independent. In this section, we assume that:

$$Y = \sum_{j=1}^k g_j(X_{C_j}). \quad (3)$$

We will study the Sobol indices and the Shapley effects in this case. First, we show a proposition about the $(V_u)_{u \subset [1:p]}$.

Proposition 1. *For all $u \subset [1 : p]$, we have:*

$$V_u = \sum_{j=1}^k V_{u \cap C_j}. \quad (4)$$

This first proposition states that all the $(V_u)_{u \subset [1:p]}$ are functions of only the V_u , for subsets u of C_1, \dots, C_k .

We then provide a consequence of Proposition 1 on Sobol indices:

Proposition 2. *For all $u \subset [1 : p]$ such that u is not a subset of a C_j , we have:*

$$S_u = 0. \quad (5)$$

Remark 1. *We have the same result for the "generalized sensitivity indices" suggested in [Cha13] but the proof is not a consequence of Proposition 1 and is straightforward.*

This proposition gives the majority of Sobol indices immediately. It remains only $\sum_{j=1}^k 2^{C_j} - 1$ unknown non-zero Sobol indices instead of $2^p - 1$.

We also provide a consequence of Proposition 1 on Shapley effects:

Proposition 3. *Let $i \in [1 : p]$. Let $j(i) \in \mathbb{N}$ be so that $i \in C_{j(i)}$. Then, we have:*

$$\eta_i = \frac{1}{|C_{j(i)}| \text{Var}(Y)} \sum_{u \subset C_{j(i)} \setminus \{i\}} \left(\frac{|C_{j(i)}| - |u|}{|u|} \right)^{-1} (V_{u \cup \{i\}} - V_u). \quad (6)$$

In an equivalent manner, let us write $\tilde{\eta}_i$ for the Shapley effect for the following output:

$$Y_{C_{j(i)}} := g_j(X_{C_j})$$

(we just focus on the variables which are in the same group as X_i). Then, the Shapley effect η_i for the whole output ($Y = f(X)$) is equal to :

$$\eta_i = \frac{\text{Var}(Y_{C_{j(i)}})}{\text{Var}(Y)} \tilde{\eta}_i. \quad (7)$$

For example, if we apply this proposition in the case where X_i is the only variable in its group, then we have $\eta_i = S_i$.

To compute the Shapley effect η_i , the last proposition reduces the sum from all the subsets of $[1 : p] \setminus \{i\}$ to all the subset of $C_{j(i)} \setminus \{i\}$. Then, the computational gain is the same as in Proposition 1.

4 Explicit computation of sensitivity indices for Gaussian linear models

4.1 Linear Gaussian framework

In this section, we assume that $X \sim \mathcal{N}(\mu, \Gamma)$, that Γ is invertible and that $f : x \mapsto \beta_0 + \beta^T x$, for a fixed $\beta_0 \in \mathbb{R}$ and a fixed vector β . This framework is widely used to model physical phenomena (see for example [KHF⁺06], [HT11], [Ros04]). Indeed, uncertainties are often modelled as Gaussian variables and

an unknown function is commonly estimated by its linear approximation. We can assume without loss of generality that $\mu = 0$ and $\beta_0 = 0$. For now, we will not assume that there are different groups of independent variables. In this framework, the sensitivity indices can be calculated explicitly. First, we write the sensitivity indices with expectations of conditional variances:

$$S_u := (-1)^{|u|} \sum_{v \subset u} (-1)^{|v|+1} \frac{\mathbb{E}(\text{Var}(Y|X_u))}{\text{Var}(Y)}, \quad \forall u \neq \emptyset, \quad (8)$$

and

$$\eta_i := \frac{1}{p \text{Var}(Y)} \sum_{u \subset -i} \binom{p-1}{|u|}^{-1} (\mathbb{E}(\text{Var}(Y|X_u)) - \mathbb{E}(\text{Var}(Y|X_{u \cup \{i\}}))). \quad (9)$$

We will now exploit the Gaussian framework as in [OP17], using that for all subset $u \subset [1 : p]$,

$$\text{Var}(Y|X_u) = \text{Var}(\beta_{-u}^T X_{-u} | X_u) = \beta_{-u}^T (\Gamma_{-u, -u} - \Gamma_{-u, u} \Gamma_{u, u}^{-1} \Gamma_{u, -u}) \beta_{-u} \quad (10)$$

where $\beta_u := (\beta_i)_{i \in u}$ and $\Gamma_{u, v} := (\Gamma_{i, j})_{i \in u, j \in v}$. These conditional variances are constant so they are equal to their expectation. Then, we can use these formulae to compute explicitly these sensitivity indices in the Gaussian linear case.

Remark 2. *If the matrix Γ is not invertible, there exist subsets u such that $\Gamma_{u, u}$ is not invertible. However, Equation (10) remains if we replace $\Gamma_{u, u}^{-1}$ by the generalized inverse (for symmetric matrices) of $\Gamma_{u, u}$.*

Remark 3. *One can show a similar result when X follows an asymmetric Laplace distribution $AL_p(m, \Gamma)$. However, the conditional variances are not constant in this case and their expectations must be estimated, for instance by Monte-Carlo.*

One issue remains though, namely computing numerically the sum in (9). Indeed, we have to sum over all the subsets of $[1 : p]$ which do not contain i . We also have to group the subsets u and $u \cup \{i\}$. For this purpose, we suggest to use the following bijective map:

$$h : \begin{array}{ccc} \mathcal{P}([1 : p]) & \longrightarrow & [0 : 2^p - 1] \\ u & \longmapsto & \sum_{i \in u} 2^{i-1}. \end{array}$$

We remark that:

$$u \subset -i \iff \left\lfloor \frac{h(u)}{2^{i-1}} \right\rfloor \equiv 0 \pmod{2}.$$

Finally, we can see that if $u \subset -i$, then $h(u \cup \{i\}) = h(u) + h(\{i\})$. Based on this map and Equations (8) and (9), we suggest an algorithm that we call "LG-Indices" (for Linear Gaussian). This algorithm computes the variance-based sensitivity indices in the linear Gaussian framework.

LG-Indices: Inputs: β, Γ .

1. Let $\text{Var}(Y) = \beta^T \Gamma \beta$ and let $\text{Var}(Y|X) = 0$.
2. (Compute the conditional variances) For $j = 0, \dots, 2^p - 1$, do the following:
 - (a) Compute $u = h^{-1}(j)$.
 - (b) Compute $V_j := \text{Var}(Y|X_u)$ using (10).
3. Initialise $S = (0, \dots, 0) \in \mathbb{R}^{2^p}$
4. (Compute the Sobol indices) For $j = 0, \dots, 2^p - 1$, do the following:
 - (a) Let $v = h^{-1}(j)$.
 - (b) For all u such that $v \subset u$, let:

$$S_{h(u)} = S_{h(u)} + (-1)^{|v|+1} V_j \quad (11)$$

5. Let $S_0 = 0$
6. For $j = 1, \dots, 2^p - 1$, let $u = h^{-1}(j)$ and

$$S_j = \frac{(-1)^{|u|}}{\text{Var}(Y)} S_j. \quad (12)$$

7. (Compute the Shapley effects) For $i = 1, \dots, p$, do the following:
 - (a) Initialize $\eta = (0, \dots, 0) \in \mathbb{R}^p$.
 - (b) For $k = 0, \dots, 2^p - 1$, do the following:
 - i. If $\lfloor \frac{k}{2^{i-1}} \rfloor \equiv 0 \pmod{2}$, then update :

$$\eta_i = \eta_i + \left(\frac{p-1}{|h^{-1}(k)|} \right)^{-1} (V_k - V_{k+2^{i-1}}). \quad (13)$$

- (c) Let $\eta_i = \eta_i / (p \text{Var}(Y))$.

Outputs (S, η) .

Remark 4. *We can use the previous algorithm for any f and any law of X if we can estimate the expectation of conditional variances (or the variance of conditional expectations equivalently).*

4.2 Linear Gaussian framework with independent groups of variables

Despite the analytical formula (10), the computational cost remains an issue when the number of input variables p is too large. Based on an implementation in the R software, LG-Indices provides almost instantaneous results for $p \leq 15$, but becomes impracticable for $p \geq 30$. Indeed, we have to store 2^p values,

namely the $(\text{Var}(Y|X_u))_{u \subset [1:p]}$, and this can be a significant issue. Fortunately, when p is large, it can frequently be the case that there are independent groups of random variables. That is, after a permutation of the variables, the covariance matrix Γ is a block diagonal matrix. In this case, Propositions 2 and 3 show that this high dimensional computational problem boils down to a collection of lower dimensional problems.

In this framework, we have seen in Proposition 1 that we only have to calculate the $\sum_{j=1}^k 2^{|C_j|}$ values $\{\text{Var}(Y|X_u), u \subset C_j, j \in [1:k]\}$ instead of all the 2^p values $\{\text{Var}(Y|X_u), u \subset [1:p]\}$. We detail this idea in the algorithm "LG-GroupsIndices".

LG-GroupsIndices: Inputs : β, Γ .

1. Let C_1, \dots, C_k be the independent groups of variables, for example using the function "graph_from_adjacency_matrix" of the R package "igraph" (see [CN06]).
2. Let η be a vector of size p .
3. For $j = 1, \dots, k$, do the following:

(a) Let $(\tilde{S}, \tilde{\eta})$ be the output of the algorithm LG-Indices with the inputs β_{C_j} and Γ_{C_j, C_j} .

(b) Let

$$S^j = \frac{\beta_{C_j}^T \Gamma_{C_j, C_j} \beta_{C_j}}{\beta_T^T \Gamma \beta_T} \tilde{S}$$

(c) Let

$$\eta_{C_j} = \frac{\beta_{C_j}^T \Gamma_{C_j, C_j} \beta_{C_j}}{\beta_T^T \Gamma \beta_T} \tilde{\eta}.$$

Ouputs : (S^1, \dots, S^k, η) .

We have used LG-GroupsIndices for computing Shapley effects on an industrial study in the field of nuclear safety. In this model, the twelve inputs were modelled by a Gaussian vector with two independent groups of six variables. Proposition 3 enables us to compute the Shapley effects computing only $2^6 + 2^6 = 128$ conditional variances instead of $2^{12} = 4096$. These results have been presented by Pietro Mosca in the sixteen International Symposium on Reactor Dosimetry.

The computational time does not really depend of the coefficients of β and Γ , that is why we prefer to consider simulated toy examples to compare the different algorithms in the next section.

5 Numerical experiments: Shapley effects computations

To position our work with respect to the state of art, we compare in this section the algorithms "LG-Indices" and "LG-GroupsIndices" with an existing algorithm designed to compute the Shapley effects for global sensitivity analysis.

5.1 Random permutations Algorithm: Shapley effects estimation

To the best of our knowledge, the only existing algorithm designed to compute the Shapley effects for global sensitivity analysis is suggested by Song in [SNS16]. This algorithm is introduced for a general function f and a general distribution of X and does not focus specifically on the linear Gaussian model. In this general context, the expectation of the conditional variances are estimated by a double Monte-Carlo procedure.

This Shapley effect estimation suggested in [SNS16] relies on the following formulation of the η_i :

$$\eta_i = \frac{1}{p! \text{Var}(Y)} \sum_{\sigma \in \mathcal{S}_p} (\text{E}(\text{Var}(Y|X_{P_i(\sigma)})) - \text{E}(\text{Var}(Y|X_{P_i(\sigma) \cup \{i\}}))), \quad (14)$$

where \mathcal{S}_p is the set of all permutations of $[1 : p]$ and $P_i(\sigma) := \{\sigma(j), j \in [1 : i]\}$. Then, in order to circumvent the $p!$ estimations of expectations of conditional variances, [SNS16] suggests to generate m ($m < p!$ for large values of p) permutations $\sigma_1, \dots, \sigma_m$ uniformly in \mathcal{S}_p and to let:

$$\hat{\eta}_i = \frac{1}{m \text{Var}(Y)} \sum_{j=1}^m (\text{E}(\text{Var}(Y|X_{P_i(\sigma_j)})) - \text{E}(\text{Var}(Y|X_{P_i(\sigma_j) \cup \{i\}}))). \quad (15)$$

It is suggested in [IP17] to choose m as large as possible and to choose little values for the double Monte-Carlo procedure. This algorithm is already implemented in the R package "sensitivity" and is called "shapleyPermRand". However, as we focus on the linear Gaussian framework, for a fair comparison, we adapt the algorithm suggested in [SNS16] to this particular framework replacing the estimations of $(\text{E}(\text{Var}(Y|X_u)))_{u \subset [1:p]}$ by their theoretical values given by (10). We will write "random permutations Algorithm" for this algorithm.

A variant of this algorithm, called "shapleyPermEx", is implemented in the R package "sensitivity". Although, this algorithm is not clearly suggested in [SNS16]. "shapleyPermEx" differs from "shapleyPermRand" by computing the sum over all the permutations when the former algorithm was only estimating the sum thanks to a Monte-Carlo method. We used the linear Gaussian framework to replace the expectation of conditional variances by their theoretical values, such that the algorithm gives the exact values of the Shapley effects. We will write "exact permutations Algorithm" for this algorithm. This method

still remains very long due to the computation of $(p - 1)!$ conditional variances. For example, it spends more than ten minutes computing the Shapley effects for only $p = 10$.

5.2 Shapley effects in the linear Gaussian framework

From now on, as we focus on the Shapley effects, we do not carry out the steps related to the computation of Sobol indices in the algorithms "LG-indices" and "LG-GroupsIndices".

Let us consider a simulated toy example. We generate β by a $\mathcal{N}(0, I_p)$ random variable and Γ by writing $\Gamma = AA^T$, where the coefficients of A are generated independently with a standard normal distribution.

First, we compare LG-Indices with exact permutations Algorithm. Both provide the exact Shapley values but with different computational times. Table 1 provides the computation times in seconds for different values of p , the number of input variables. We remark that LG-indices is faster than exact permutations

Table 1: Computational time (in seconds) for exact permutations Algorithm and LG-Indices for different values p .

	$p = 6$	$p = 7$	$p = 8$	$p = 9$	$p = 10$
exact permutations	0.11	0.78	7.31	77.6	925
LG-Indices	0.004	0.008	0.018	0.039	0.086

Algorithm.

We can also compare LG-Indices with random permutations Algorithm. For the latter, we choose m , the number of permutations generated in Song, so that the computational time is the same as LG-Indices. Yet, while our algorithm gives the exact Shapley effect, the random permutations Algorithm provides an estimation of it. Hence, the performance of Song is evaluated by computing the coefficients of variation in %, for different values of p . We give in Table 2 the average of the p coefficients of variations. We recall that the coefficient of variation corresponds to the ratio of the standard deviation over the mean value. We see that the algorithm suggested in [SNS16] has quite large coefficients of

Table 2: Mean of the coefficients of variation of Shapley effects estimated by random permutations Algorithm for the same computational time as LG-GroupsIndices.

	$p = 3$	$p = 4$	$p = 5$	$p = 6$	$p = 7$
m chosen	10	12	18	30	50
mean of coefficients of variation	31%	26%	22%	18%	13%

variation when we choose m so that the computational time is the same as

our algorithm. However, this variation decreases with the number of inputs p . We can explain that saying that the computational time of LG-Indices is exponential with p . So, we can see that the precision of random permutations Algorithm increases with m .

5.3 Shapley effects in the linear Gaussian framework with independent groups of variables

For now, to the best of our knowledge, LG-GroupsIndices is the only algorithm which can compute the exact Shapley effects for large values of p (the number of inputs). Indeed, random permutations Algorithm can handle large values of p but always computes estimations of Shapley effects. On an other hand, LG-Indices computes exact Shapley effects but becomes too long for $p \geq 20$ (the computational time is exponential in p).

First, we compare the computational time of LG-Indices and LG-GroupsIndices for low values of p on a toy simulated example as in 5.2. We generate k independent groups of n variables. We give these results in Table 3.

Table 3: Computational time (in seconds) for LG-Indices and LG-GroupsIndices for different values of k and n .

	$k = 3$ $n = 3$	$k = 4$ $n = 3$	$k = 4$ $n = 4$	$k = 5$ $n = 4$
LG-Indices	0.04	0.47	8.45	168.03
LG-GroupsIndices	0.002	0.003	0.004	0.007

Now, we compare LG-GroupsIndices with random permutations Algorithm as in 5.2: we choose m so that the computational time is the same and we give the average of the p coefficients of variation of random permutations Algorithm in Table 4.

Table 4: Mean of the coefficients of variation of Shapley effects estimated by random permutations Algorithm for the same computational time as LG-GroupsIndices.

	$k = 3$ $n = 3$	$k = 4$ $n = 4$	$k = 5$ $n = 5$	$k = 6$ $n = 6$	$k = 10$ $n = 5$	$k = 5$ $n = 10$
m chosen	7	8	9	10	5	98
mean of coefficients of variation	34%	38%	34%	36%	40%	12%

Here, the mean of the coefficients of variation remains quite large (around 35%) when we chose $k = n$. However, when we choose k larger (resp. lower) than n , this variation increases (resp. decreases).

6 Conclusion

In this paper, we give new theoretical results about the variance-based sensitivity indices for independent groups of inputs and additive models. These results drastically reduce the computational cost of these indices. Then, we apply these results on the linear Gaussian framework and we suggest two algorithms: the first one for the general case and the second one for a block diagonal covariance matrix. These algorithms compute efficiently the theoretical values of the variance-based sensitivity indices. Numerical experiments on Shapley effects computations highlight this efficiency and the benefit compared to existing methods.

Acknowledgements

We are very grateful to Bertrand Iooss for his ideas and his advises. We also would like to thank Pietro Mosca for his contribution to the implementation of the industrial application, and Arnaud Grivet Sebert for the constructive discussions that helped to improve the paper.

Appendix

Proof of Proposition 1:

We will use a classical result on the conditional expectation that says that if \mathcal{H} is an independent σ -algebra of $\sigma(\sigma(Z), \mathcal{G})$, then:

$$\mathbb{E}(Z|\sigma(\mathcal{G}, \mathcal{H})) = \mathbb{E}(Z|\mathcal{G}). \quad (16)$$

First, we compute the conditional expectation:

$$\begin{aligned}
\mathbb{E}(Y|X_u) &= \mathbb{E}\left(\sum_{j=1}^k g_j(X_{C_j}) \middle| X_u\right) \\
&= \sum_{j=1}^k \mathbb{E}(g_j(X_{C_j})|X_u) \\
&= \sum_{j=1}^k \mathbb{E}(g_j(X_{C_j})|\sigma((X_{u \cap C_i}, i \neq j), \sigma(X_{u \cap C_j}))) \\
&= \sum_{j=1}^k \mathbb{E}(g_j(X_{C_j})|X_{u \cap C_j}), \text{ thanks to (16)} \\
&= \sum_{j=1}^k \mathbb{E}\left(\sum_{i=1}^k (g_i(X_{C_i})) \middle| X_{u \cap C_j}\right) \\
&= \sum_{j=1}^k \mathbb{E}(Y|X_{u \cap C_j})
\end{aligned}$$

Now, taking the variance, we have:

$$\begin{aligned}
V_u &= \text{Var}(\mathbb{E}(Y|X_u)) \\
&= \sum_{j=1}^k \text{Var}(\mathbb{E}(Y|X_{u \cap C_j})) + \sum_{j_1 \neq j_2} \text{Cov}(\mathbb{E}(Y|X_{u \cap C_{j_1}}), \mathbb{E}(Y|X_{u \cap C_{j_2}})) \\
&= \sum_{j=1}^k \text{Var}(\mathbb{E}(Y|X_{u \cap C_j})) \\
&= \sum_{j=1}^k V_{u \cap C_j}.
\end{aligned}$$

□

Proof of Proposition 2:

$$\begin{aligned}
S_u &= \frac{1}{\text{Var}(Y)} \sum_{v \subset u} (-1)^{|u|-|v|} V_v \\
&= \frac{1}{\text{Var}(Y)} \sum_{v \subset u} (-1)^{|u|-|v|} \sum_{j=1}^k V_{v \cap C_j} \\
&= \sum_{j=1}^k \sum_{w \subset u \cap C_j} \sum_{\substack{v, \\ w \subset v \subset u}} (-1)^{|u|-|v|} V_w \\
&= \sum_{j=1}^k \sum_{w \subset u \cap C_j} V_w \sum_{\substack{v, \\ w \subset v \subset u}} (-1)^{|u|-|v|} \\
&= \sum_{j=1}^k \sum_{w \subset u \cap C_j} V_w \cdot 0 \\
&= 0.
\end{aligned}$$

In effect, if u is not a subset of a C_j , then, for all j , for all $w \subset u \cap C_j$, we have $w \subsetneq u$. So, using the Newton formula, we have:

$$\begin{aligned}
\sum_{v, w \subset v \subset u} (-1)^{|v|-|w|} &= \sum_{n=|w|}^{|u|} \sum_{\substack{v, \\ |v|=n, \\ w \subset v \subset u}} (-1)^{|v|-|w|} \\
&= \sum_{n=|w|}^{|u|} \binom{|u|-|w|}{n-|w|} (-1)^{n-|w|} \\
&= \sum_{m=0}^{|u|-|w|} \binom{|u|-|w|}{m} (-1)^m \\
&= 0.
\end{aligned}$$

□

Proof of Proposition 3:

$$\begin{aligned}
\eta_i &= \frac{1}{p\text{Var}(Y)} \sum_{u \subset C_{-i}} \binom{p-1}{|u|}^{-1} (V_{u \cup \{i\}} - V_u) \\
&= \frac{1}{p\text{Var}(Y)} \sum_{u \subset C_{-i}} \binom{p-1}{|u|}^{-1} \sum_{j=1}^k (V_{(u \cup i) \cap C_j} - V_{u \cap C_j}) \\
&= \frac{1}{p\text{Var}(Y)} \sum_{u \subset C_{-i}} \binom{p-1}{|u|}^{-1} \left(\left[\sum_{j \neq j(i)} V_{u \cap C_j} - V_{u \cap C_j} \right] + (V_{(u \cup \{i\}) \cap C_{j(i)}} - V_{u \cap C_{j(i)}}) \right) \\
&= \frac{1}{p\text{Var}(Y)} \sum_{u \subset C_{-i}} \binom{p-1}{|u|}^{-1} (V_{(u \cup \{i\}) \cap C_{j(i)}} - V_{u \cap C_{j(i)}}) \\
&= \frac{1}{p\text{Var}(Y)} \sum_{u \subset C_{j(i)-i}} \sum_{v \subset C_{j(i)}} \binom{p-1}{|u \cup v|}^{-1} (V_{(u \cup v \cup \{i\}) \cap C_{j(i)}} - V_{u \cup v \cap C_{j(i)}}) \\
&= \frac{1}{p\text{Var}(Y)} \sum_{u \subset C_{j(i)-i}} \sum_{v \subset C_{j(i)}} \binom{p-1}{|u| + |v|}^{-1} (V_{u \cup \{i\}} - V_u) \\
&= \frac{1}{p\text{Var}(Y)} \sum_{u \subset C_{j(i)-i}} \sum_{j=0}^{p-|C_{j(i)}|} \binom{p-|C_{j(i)}|}{j} \binom{p-1}{|u| + j}^{-1} (V_{u \cup \{i\}} - V_u).
\end{aligned}$$

It remains to prove the following equation:

$$\frac{1}{p} \sum_{j=0}^{p-|C_{j(i)}|} \binom{p-|C_{j(i)}|}{j} \binom{p-1}{|u| + j}^{-1} = \frac{1}{|C_{j(i)}|} \binom{C_{j(i)}-1}{|u|}^{-1}. \quad (17)$$

In the interest of simplifying notation, until the end of the proof, we will write u (resp. c) instead of $|u|$ (resp. $|C_{j(i)}|$). We can verify that the equation (17) is equivalent to the following equations:

$$\begin{aligned}
&\frac{1}{p} \sum_{j=0}^{p-c} \frac{(p-c)!}{j!(p-c-j)!} \frac{(u+j)!(p-1-u-j)!}{(p-1)!} = \frac{1}{c} \frac{u!(c-1-u)!}{(c-1)!} \\
\iff &\sum_{j=0}^{p-c} \binom{u+j}{u} \binom{p-1-u-j}{c-u-1} = \binom{p}{c}. \quad (18)
\end{aligned}$$

We will show (18). Now, we can remark that we have:

$$\frac{x^c}{(1-x)^{c+1}} = x \frac{x^u}{(1-x)^{u+1}} \frac{x^{c-u-1}}{(1-x)^{c-u}}. \quad (19)$$

Giving their power series, we have:

$$x \left(\sum_{k \geq 0} \binom{k}{u} x^k \right) \left(\sum_{k' \geq 0} \binom{k'}{c-u-1} x^{k'} \right) = \sum_{k'' \geq 0} \binom{k''}{c} x^{k''}. \quad (20)$$

We have the equality of the coefficient of x^p . Then:

$$\begin{aligned} \binom{p}{c} &= \sum_{k+k'=p-1} \binom{k}{u} \binom{k'}{c-u-1} = \sum_{k=u}^{p-1} \binom{k}{u} \binom{p-1-k}{c-u-1} = \sum_{j=0}^{p-1-u} \binom{u+j}{u} \binom{p-1-u-j}{c-u-1} \\ &= \sum_{j=0}^{p-c} \binom{u+j}{u} \binom{p-1-u-j}{c-u-1}. \end{aligned}$$

For the last equality, we remark that if $j > p-c$, then $p-1-u-j < c-u-1$ so the last terms of the sum are equal to zero. We have proven (18), that concludes the proof. \square

References

- [BHP16] Emanuele Borgonovo, Gordon B. Hazen, and Elmar Plischke. A Common Rationale for Global Sensitivity Measures and Their Estimation. *Risk Analysis: An Official Publication of the Society for Risk Analysis*, 36(10):1871–1895, October 2016.
- [CBSV16] Riccardo Colini-Baldeschi, Marco Scarsini, and Stefano Vaccari. Variance allocation and Shapley value. *Methodology and Computing in Applied Probability*, pages 1–15, 2016.
- [Cha13] Galle Chastaing. *Indices de Sobol gnraliss pour variables dpendantes*. phdthesis, Universit de Grenoble, September 2013.
- [CN06] Gabor Csardi and Tamas Nepusz. The igraph software package for complex network research. *InterJournal, Complex Systems*:1695, 2006.
- [FWJ08] Shaheen S. Fatima, Michael Wooldridge, and Nicholas R. Jennings. A linear approximation method for the Shapley value. *Artificial Intelligence*, 172(14):1673–1699, September 2008.
- [HI03] Franz Hubert and Svetlana Ikonnikova. Strategic investment and bargaining power in supply chains: A Shapley value analysis of the Eurasian gas market. March 2003.
- [HT11] Hugo Hammer and Hkon Tjelmeland. Approximate forwardbackward algorithm for a switching linear Gaussian model. *Computational Statistics & Data Analysis*, 55(1):154–167, January 2011.
- [IP17] Bertrand Iooss and Clementine Prieur. Shapley effects for sensitivity analysis with dependent inputs: comparisons with Sobol’ indices, numerical estimation and applications. *arXiv:1707.01334 [math, stat]*, July 2017. arXiv: 1707.01334.

- [KHF⁺06] T. Kawano, K. M. Hanson, S. Frankle, P. Talou, M. B. Chadwick, and R. C. Little. Evaluation and Propagation of the ²³⁹Pu Fission Cross-Section Uncertainties Using a Monte Carlo Technique. *Nuclear Science and Engineering*, 153(1):1–7, May 2006.
- [MTA15] Thierry A. Mara, Stefano Tarantola, and Paola Annoni. Non-parametric methods for global sensitivity analysis of model output with dependent inputs. *Environmental Modelling & Software*, 72:173–183, October 2015.
- [MvLG⁺08] Stefano Moretti, Danitsja van Leeuwen, Hans Gmuender, Stefano Bonassi, Joost van Delft, Jos Kleinjans, Fioravante Patrone, and Domenico Franco Merlo. Combining Shapley value and statistics to the analysis of gene expression data in children exposed to air pollution. *BMC Bioinformatics*, 9:361, September 2008.
- [OP17] Art B. Owen and Clmentine Prieur. On Shapley value for measuring importance of dependent inputs. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1):986–1002, 2017.
- [Owe14] A. Owen. Sobol’ Indices and Shapley Value. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1):245–251, January 2014.
- [Ros04] Antti-Veikko Ilmari Rosti. *Linear Gaussian models for speech recognition*. PhD thesis, University of Cambridge, 2004.
- [SCSo00] Andrea Saltelli, Karen Chan, E. Marian Scott, and others. *Sensitivity analysis*, volume 1. Wiley New York, 2000.
- [Sha53] L. S. Shapley. A value for n-person games. Contribution to the Theory of Games. *Annals of Mathematics Studies*, 2:28, 1953.
- [SNS16] E. Song, B. Nelson, and J. Staum. Shapley Effects for Global Sensitivity Analysis: Theory and Computation. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1):1060–1083, January 2016.
- [Sob93] Ilya M. Sobol. Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiments*, 1(4):407–414, 1993.