



HAL
open science

Proportional marginal effects for global sensitivity analysis

Margot Herin, Marouane Il Idrissi, Vincent Chabridon, Bertrand Iooss

► **To cite this version:**

Margot Herin, Marouane Il Idrissi, Vincent Chabridon, Bertrand Iooss. Proportional marginal effects for global sensitivity analysis. 2022. hal-03825935v1

HAL Id: hal-03825935

<https://hal.science/hal-03825935v1>

Preprint submitted on 23 Oct 2022 (v1), last revised 16 Mar 2024 (v4)

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.

Proportional marginal effects for global sensitivity analysis

Margot Herin^a, Marouane El Idrissi^{b,c,d}, Vincent Chabridon^{b,c}, Bertrand Iooss^{b,c,d,e}

^a*Sorbonne Université, Laboratoire d'Informatique de Paris 6, 4 place Jussieu, 75005 Paris, France.*

^b*EDF Lab Chatou, 6 Quai Watier, 78401 Chatou, France*

^c*SINCLAIR AI Lab., Saclay, France*

^d*Institut de Mathématiques de Toulouse, 31062 Toulouse, France*

^e*Corresponding Author - Email: bertrand.iooss@edf.fr*

Abstract

Performing (variance-based) global sensitivity analysis (GSA) with dependent inputs has recently benefited from cooperative game theory concepts. By using this theory, despite the potential correlation between the inputs, meaningful sensitivity indices can be defined via allocation shares of the model output's variance to each input. The “Shapley effects”, i.e., the Shapley values transposed to variance-based GSA problems, allowed for this suitable solution. However, these indices exhibit a particular behavior that can be undesirable: an exogenous input (i.e., which is not explicitly included in the structural equations of the model) can be associated with a strictly positive index when it is correlated to endogenous inputs. In the present work, the use of a different allocation, called the “proportional values” is investigated. A first contribution is to propose an extension of this allocation, suitable for variance-based GSA. Novel GSA indices are then proposed, called the “proportional marginal effects” (PME). The notion of exogeneity is formally defined in the context of variance-based GSA, and it is shown that the PME allow the distinction of exogenous variables, even when they are correlated to endogenous inputs. Moreover, their behavior is compared to the Shapley effects on analytical toy-cases and more realistic use-cases.

Keywords: Cooperative game theory, Dependence, Proportional values, Sobol' indices, Shapley effects.

1. Introduction

When using phenomenological numerical models in science and engineering, the uncertainty quantification (UQ) process allows to consider and better quantify the various sources of uncertainties, most often by the way of probabilistic modeling [13]. Global sensitivity analysis (GSA) is a key step of this process, aiming to understand the effects of each uncertain model input (or set of inputs) on the quantity of interest related to one (or more) output variable of interest obtained from the numerical model [35, 23]. From a practical viewpoint, GSA aims at investigating four major settings [6]: (i.) model exploration, i.e., investigating the input-output relationship; (ii.) factor fixing, i.e., identifying non-influential inputs; (iii.) factor prioritization, i.e., quantifying the most important inputs using quantitative importance measures; (iv.) robustness analysis, i.e., quantifying the sensitivity of the quantity of interest with respect to probabilistic model uncertainty of the input distributions. In the present paper, one will more focus on the first three settings, without discussing much the fourth one.

Among a large panel of GSA indices, the variance-based sensitivity measures, also called “Sobol' indices” [38], are derived from the functional analysis of variance (FANOVA) decomposition [7] between all the independent inputs. Thus, these indices enable to provide interpretable answers to some of the previously mentioned GSA settings. Let $Y = G(X)$ denotes the input-output relationship under study, with $G(\cdot) : \mathbb{R}^d \rightarrow \mathbb{R}$ a deterministic (often black-box) numerical model, Y a scalar output and $X = (X_1, \dots, X_d)$ a vector of d scalar inputs. Moreover, let $\mathcal{P}(D)$ the set of all subsets of

$D = \{1, \dots, d\}$. For every subset of input $X_A = (X_i)_{i \in A}$, $A \in \mathcal{P}(D)$, the Sobol’ indices are defined as follows:

$$S_A = \frac{\sum_{B \subset A} (-1)^{|A|-|B|} \mathbb{V}(\mathbb{E}[G(X)|X_B])}{\mathbb{V}(G(X))} \quad (1)$$

where $|\cdot|$ denotes the number of elements in a subset. If the inputs are assumed to be independent, thanks to the FANOVA decomposition, Sobol’ indices lead to a well-defined allocation of an output’s variance share (i.e., S_A) to every subset of inputs $A \in \mathcal{P}(D)$. In this case, the variance’s shares $(S_A)_{A \in \mathcal{P}(D)}$ sum up to one while being nonnegative. As the indices can be interpreted as proportions of the output variance, they allow to determine which inputs of a numerical model contribute the most to the variability of the output, or, on the contrary, to identify the ones that are not influential, and possibly which inputs interact with each other. Therefore, Sobol’ indices can be directly used to answer to the factor fixing and factor prioritization settings (ii. and iii.).

However, in many applications, some inputs may have a statistical dependence structure, either initially imposed in their probabilistic modeling [27] or induced by physical constraints upon the input or the output space [26, 29]. In these cases, estimating and interpreting Sobol’ indices is not trivial as shown by many different analyses and interpretations proposed in the past (see [22] or [6] for an overview of this topic). In order to circumvent this issue, [31] proposed a new approach based on the “Shapley value” [36], a solution concept developed in cooperative game theory and powerfully used in economic modeling. It consists in distributing both gains and costs to several players working in coalition in an egalitarian way, ensuring that each player gains as much (or more) as they would have from playing individually. Therefore, based on Shapley values and Sobol’ indices, [31] proposed the so-called “Shapley effects” as new GSA indices in the context of dependent inputs. The underlying idea is to compute, similarly to a game involving coalition of players, the value assigned to a coalition of inputs X_A as the explanatory power of a part of output variance. This value corresponds to the so-called “closed Sobol’ indices” defined as:

$$S_A^{\text{clos}} = \frac{\mathbb{V}(\mathbb{E}[G(X)|X_A])}{\mathbb{V}(G(X))}. \quad (2)$$

In the GSA context, the two main properties and advantages of the Shapley effects are the following: firstly, they cannot be negative; secondly, their sum is equal to one, even in the dependent inputs’ case since they allow to bypass the intricate issue of variance decomposition [32, 22]. Let us remark that these two properties correspond to the two main desirability criteria for importance measures of linear regression models as reviewed in [14]. Moreover, the egalitarian principle driving the allocation rule states that, in the independent inputs’ case, an interaction effect is equally apportioned to each input involved in the interaction. Finally, several works have studied the Shapley effects estimation issues. Such estimates can be obtained via several techniques such as Monte Carlo-based algorithms [39], k -nearest neighbors [3] or Möbius inverses [33].

In [22], the Shapley effects have been claimed to be used for the factor fixing setting since an effect close to zero means that the input has no significant contribution to the variance of the output (neither by its interactions nor by its possible dependencies with other inputs). However, another phenomenon, observed by [22] and known as the “Shapley’s joke” [18], proves that the factor fixing setting cannot be fully achieved with Shapley effects: an exogenous variable (i.e., which is not explicitly included in the structural equations of the model) can be granted a non-negligible share of the output variance, as soon as it is sufficiently correlated with endogenous inputs. This means that Shapley effects do not respect the so-called “exclusion property” defined for the importance measures of linear regression models [24, 14]. This exclusion property states that, if an input’s linear regression coefficient equals zero, then its importance measure should be zero too.

In the context of statistical learning, if G is a linear regression model, an analogy can be made between the Sobol’ indices and the squared value of the standardized regression coefficients (denoted by SRC²). Moreover, the Shapley effects correspond to the so-called “LMG measure” (named after the authors’ names, Lindeman-Merenda-Gold, see [28, 4]), which partitions the explained variance

percentage R^2 in the same way that is done by the Shapley-based allocation rule. A weighted analog of LMG, called *proportional marginal variance decomposition* (PMVD), has been proposed by [9] in order to respect the exclusion property. It is based on the proportional value allocation rule coming from cooperative game theory. Its usefulness in relation to LMG has been described in details in [14, 15] and illustrated more recently in [20, 21]. In addition to the exclusion property, a more discriminating power between the influential inputs than the one obtained with the Shapley effects is also shown. Therefore, the PMVD is a good tool (in the linear regression context) to address the factor fixing setting.

In this paper, inspired on the one hand, by the work achieved in the linear regression context leading to the PMVD, and on the other hand, by the Shapley effects, we build and propose a set of novel sensitivity indices respecting the exclusion property and not restricted to the linear model case. To do so, the *proportional marginal effects* (PME) are introduced by using a new variance decomposition, based on the proportional values concept [30, 9], which encompasses the ability to detect exogenous variables. For the sake of clarity, Table 1 provides a first preliminary analogy to emphasize which category of problem one tries to address in the present paper.

R^2 decomposition (linear regression)	$\mathbb{V}(Y)$ decomposition (GSA)
SRC ²	Sobol’ indices
LMG	Shapley effects
PMVD	PME (proposed indices)

Table 1: Analogy between linear regression importance measures (R^2 decomposition) and variance-based GSA.

The rest of this paper is organized as follows. Section 2 focuses on the interaction between GSA and cooperative game theory and the existing literature. The Shapley effects are recalled, as well as their main shortcoming: the inability to detect exogenous inputs. To that end, the notion of L^2 -*exogeneity* is formally defined. Then, Section 3 defines the proportional values and presents the main result of this paper, an extension allowing for well-defined novel GSA indices: the PME. It is additionally shown that these novel indices allow to detect exogenous inputs, while remaining inherently interpretable. Section 4 illustrates the behavior of the novel PME by using analytical formulas obtained for analytical forms of G . Section 5 briefly recalls several strategies for the estimation of PME and provides the results obtained on several more challenging numerical test-cases. Section 6 discusses several possible improvements as well as some perspectives about the proposed work. A few appendices provide extra materials such as information about reproducibility of numerical results (Appendix Appendix A) and proofs (Appendix Appendix B).

Throughout this paper, let $\mathbb{E}[\cdot]$ and $\mathbb{V}(\cdot)$ denote the expectation and variance respectively. A *coalition of players* is a subset of the *grand coalition* denoted $D = \{1, \dots, d\}$. Moreover, $\forall A \subseteq D$, the restricted set of indices $A \setminus \{i\}$, for any $i \in A$, is denoted by A_{-i} . Additionally, for any $A \subseteq D$, $X_{D \setminus A}$ is denoted by $X_{\bar{A}}$. The distribution of the random inputs X is generically denoted by P_X and the marginal distribution of any subset of inputs X_A for any $A \subseteq D$ is generically denoted by P_{X_A} . The spaces $L^2(P_{X_A})$, for any $A \subseteq D$, denote the spaces of measurable functions with finite second-order moments. When a function is referred to as being nonnegative (resp. positive), it entails that it takes values in \mathbb{R}^+ (resp. \mathbb{R}_*^+). Whenever reference is made to a model G , it is always implicitly assumed that $G \in L^2(P_X)$. In this paper, almost sure statements are followed by the acronym “a.s.”.

2. Cooperative game theory for variance-based global sensitivity analysis

This section aims at reviewing the usefulness of cooperative game theory in the process of designing variance-based GSA indices. A particular class of allocations is presented: the random order model allocations, which contains the Shapley values. The Sobol’ cooperative games are introduced, as a formalization of the analogy between players and inputs of deterministic models. The Shapley effects

are presented as the application of Shapley values to a Sobol' cooperative game. The notion of dual of a cooperative game is also presented, and an analogy is drawn between backward-forward procedures and the random order model allocations. Finally, a specific Shapley effects' drawback (for factor fixing setting) is presented as a motivation for the proposed work: their inability to detect exogenous inputs.

2.1. Analogy between allocation and variance-based GSA indices

A cooperative game is a tuple (D, v) where $D = \{1, \dots, d\}$ is a set of d players and $v : \mathcal{P}(D) \rightarrow \mathbb{R}$ is the *value function*, i.e., an application that maps a value to every possible coalition of players. Usually, v is assumed to be *monotonically increasing*, meaning that, for any two sets T and A such that $T \subseteq A \in \mathcal{P}(D)$, one has $v(T) \leq v(A)$. In other words, the value of a coalition A cannot be lower than the value of a sub-coalition $T \subseteq A$. In the following, cooperative games with monotonically increasing value functions are referred to as “monotonic cooperative games”. Moreover, if the value function v takes values in \mathbb{R}_*^+ (resp. in \mathbb{R}^+), the corresponding cooperative game is referred to as “positive (resp. nonnegative) cooperative game”.

The analogy between the players D of a cooperative game (D, v) and the inputs $(X_i)_{i \in D}$ involved in a numerical model has been first used in [31]. The author proposed to use, as a value function, the closed Sobol' indices recalled in Eq. (2), allowing to define the Sobol' cooperative games.

Definition 1 (Sobol' cooperative game). *Let $X = (X_1, \dots, X_d)^\top$ be random inputs, let $G \in L^2(P_X)$ be a model and denote $Y = G(X)$ the random output. A Sobol' cooperative game is the cooperative game with value function S^{clos} defined as follows:*

$$S^{\text{clos}} : \mathcal{P}(D) \rightarrow \mathbb{R}^+$$

$$A \mapsto S_A^{\text{clos}} = \frac{\mathbb{V}(\mathbb{E}[Y | X_A])}{\mathbb{V}(Y)}.$$

The Sobol' cooperative game thus refers to the nonnegative, monotonic cooperative game (D, S^{clos}) .

By analogy with the cooperative game theory paradigm, the choice of S^{clos} as a value function entails measuring the value of every subset of players $A \subseteq D$ as the variance of the best approximation of Y on $L^2(P_{X_A})$, i.e., $\mathbb{V}(\mathbb{E}[Y | X_A])$.

One of the key aspects of cooperative games is the notion of *allocation*. In general, allocations can be understood as a decomposition of the quantity $v(D)$ in d elements, each one being allocated to a specific player. When it comes to Sobol' cooperative games, it translates to assigning a share of the output's variance $\mathbb{V}(Y)$ to each input in the model, with limited assumptions on the probabilistic structure between the inputs (in particular, no independence is assumed between the inputs). Formally, an allocation can be understood as a mapping ϕ that associates, to a cooperative game (D, v) , a real-valued vector $(\phi_1, \dots, \phi_d)^\top \in \mathbb{R}^d$.

The *Shapley values*, are a particular example of allocations. For any cooperative game (D, v) , it is uniquely characterized as the allocation $\phi((D, v))$ verifying a set of four distinct axioms:

1. **Efficiency:** $\sum_{i=1}^d \phi_i = v(D)$;
2. **Symmetry:** $\forall i, j \in D$ with $i \neq j$, if $v(A \cup \{i\}) = v(A \cup \{j\})$ for all $A \in \mathcal{P}(D)$, then $\phi_i = \phi_j$;
3. **Null player:** $\forall i \in D$, if $v(A \cup \{i\}) = v(A)$ for all $A \in \mathcal{P}(D)$, then $\phi_i = 0$;
4. **Additivity:** If two cooperative games (D, v) and (D, v') have Shapley values ϕ and ϕ' respectively, then the cooperative game $(D, v + v')$ has Shapley values $\phi_j + \phi'_j$ for $j \in D$.

For any cooperative game (D, v) , its Shapley values can be expressed analytically, for any $i \in D$, as:

$$\text{Shap}_i((D, v)) = \frac{1}{d} \sum_{A \subseteq D-i} \binom{d-1}{|A|}^{-1} [v(A \cup \{i\}) - v(A)]. \quad (3)$$

This original formulation attributed to [36] can be interpreted as a weighted average, over every possible coalition A , of the contribution of a player i to that coalition A . This contribution is quantified by the

quantity $v(A \cup \{i\}) - v(A)$, often called “*marginal contribution*” of the player i to the coalition A in the literature. The weighting scheme can be understood as the proportion of permutations (or orderings) of D such that i appears after the players in A . While this interpretation can be hard to understand, defining the Shapley values in terms of players permutations allows for a better understanding of its underlying sharing mechanism, as it is done in the following.

A particular class of allocations, known as *random order models* [40, 10], allows to define allocations based on orderings of players, instead of reasoning in terms of coalitions as in Eq. (3). Let \mathcal{S}_D be symmetric group on D (the set of all permutations of D). Let $\pi = (\pi_1, \dots, \pi_d) \in \mathcal{S}_D$ be a particular permutation, and for any $i \in D$, denote $\pi(i) = \pi_i^{-1}$ its inverse (i.e., the position of i in π , such that $\pi_{\pi(i)} = i$). Then, one can define the following set of players, for any $i \in \{0, \dots, d\}$:

$$C_i(\pi) = \{\pi_j : j \leq i\}. \quad (4)$$

$C_i(\pi)$ is the set of the i -th first players in the ordering π , with the convention that, for any permutation, $C_0(\pi) = \{\emptyset\}$. As an illustration, let $D = \{1, 2, 3\}$, and let $\pi = (2, 1, 3) \in \mathcal{S}_D$. Then,

$$\pi(1) = 2, \quad \pi(2) = 1, \quad \text{and} \quad \pi(3) = 3.$$

Moreover,

$$C_{\pi(1)}(\pi) = C_2(\pi) = \{1, 2\}, \quad C_{\pi(2)}(\pi) = C_1(\pi) = \{2\}, \quad C_{\pi(3)}(\pi) = C_3(\pi) = \{1, 2, 3\}$$

As their names suggest, random order models endow \mathcal{S}_D with a probabilistic structure. For any game (D, v) , the set of random order models allocations (or probabilistic allocations) contains every allocation $\phi((D, v))$ that can be written, for any $i \in D$, as:

$$\begin{aligned} \phi_i &= \sum_{\pi \in \mathcal{S}_D} p(\pi) [v(C_{\pi(i)}(\pi)) - v(C_{\pi(i)-1}(\pi))] \\ &= \mathbb{E}_{\pi \sim p} [v(C_{\pi(i)}(\pi)) - v(C_{\pi(i)-1}(\pi))] \end{aligned}$$

where p is a probability mass function over the orderings of D . For a player i , its random order allocation can be interpreted as the expectation over the permutations π of D with respect to p , of the marginal contributions of i to the coalitions formed by $C_{\pi(i)-1}(\pi)$. The random order model allocations are always *efficient* and, when dealing with monotonic games, *positive* (i.e., $\phi_i \geq 0$ for any $i \in D$) [40]. The Shapley values, in particular, can be expressed as a random order model allocation, under the particular choice of p as a discrete uniform distribution over \mathcal{S}_D , which echoes Eq. (3):

$$\text{Shap}_i((D, v)) = \frac{1}{d!} \sum_{\pi \in \mathcal{S}_D} [v(C_{\pi(i)}(\pi)) - v(C_{\pi(i)-1}(\pi))]. \quad (5)$$

Random order models allow to apprehend allocations *dynamically* (see Section 2.2), meaning that coalitions are formed regarding orderings, as opposed to the pure coalition point of view displayed in Eq. (3). In this setting, Shapley values can then be understood as a maximum entropy a priori (i.e., uniform over \mathcal{S}_D) about this dynamic. In the light of this equivalent expression, L. S. Shapley himself interpreted the Shapley values as “[...] *an a priori assessment of the situation, based on either ignorance or disregard of the social organization of the players*” [37].

When it comes to GSA, the Shapley values of the Sobol’ cooperative game (D, S^{clos}) associated to a numerical model $Y = G(X_1, \dots, X_d)$ allow to define the so-called *Shapley effects* [31]. For any $i \in D$, they can be written as:

$$\text{Sh}_i := \text{Shap}_i((D, S^{\text{clos}})) \quad (6a)$$

$$= \frac{1}{d} \sum_{A \subseteq D - i} \binom{d-1}{|A|}^{-1} (S_{A \cup \{i\}}^{\text{clos}} - S_A^{\text{clos}}) \quad (6b)$$

$$= \frac{1}{d!} \sum_{\pi \in \mathcal{S}_D} [S_{C_{\pi(i)}(\pi)}^{\text{clos}} - S_{C_{\pi(i)-1}(\pi)}^{\text{clos}}]. \quad (6c)$$

These indices, which have been extensively studied in [39, 32, 22], are a great tool to quantify variable importance in the context of dependent inputs [5]. They allow for a meaningful decomposition of $\mathbb{V}(Y)$ into positive shares attributed to each input, even in situations where the inputs are correlated.

2.2. Dual of a cooperative game

The notion of the *dual* of a cooperative game is also of interest in the present paper. On the one hand, under the game theory paradigm presented previously, the aim of the value function v is to quantify the “value produced” by a coalition of players (e.g., the monetary value). On the other hand, the dual of a cooperative game focuses on the “worth”, or “bargaining power” of a coalition, i.e., the shortfall in value due to a coalition [9, 10]. The dual of a cooperative game (D, v) is usually denoted by (D, w) where w is defined, for any $A \in \mathcal{P}(D)$ as:

$$w(A) = v(D) - v(D \setminus A). \quad (7)$$

The quantities $w(A)$ are often referred to as the *marginal contribution of a coalition A to the grand coalition D* in the literature, and is often interpreted as a measure of how crucial a coalition is in producing $v(D)$. For the sake of conciseness, in the following, one refers to $w(A)$ as the *marginal contribution of the coalition A*. The dual (D, w) of (D, v) is also a cooperative game, and thus one can seek to construct relevant allocations for this game.

Following up this idea of dual game, one can draw a parallel between random order model allocations and the well-known “forward” and “backward” variable selection procedures. Figure 1 illustrates this similarity. Formally, one can notice that, for a player i and any permutation $\pi \in \mathcal{S}_D$, one has:

$$w(C_{\pi(i)}(\pi)) - w(C_{\pi(i)-1}(\pi)) = v(D \setminus C_{\pi(i)-1}(\pi)) - v(D \setminus C_{\pi(i)}(\pi)). \quad (8)$$

A random order model allocation of the dual of a cooperative game can be understood as the expected (with respect to a probability mass function p over \mathcal{S}_D) marginal contribution of a player i to the players that *follows* in the orderings’ dynamic, whereas for the initial cooperative game, it is the expected marginal contribution of i to the players that *precedes* in the orderings’ dynamic.

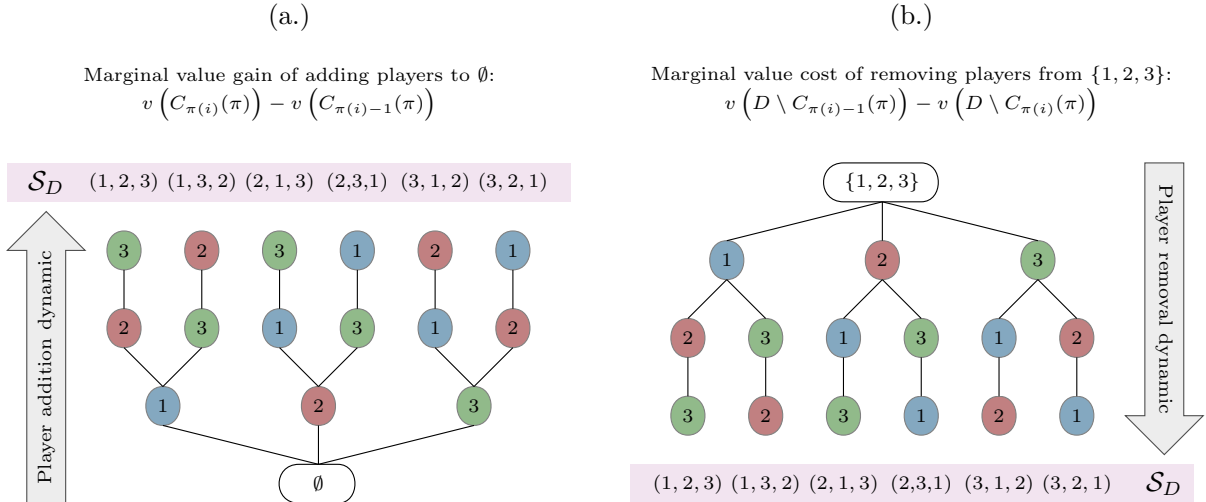


Figure 1: Analogy between random order model allocations and the forward-backward procedures for $D = \{1, 2, 3\}$: (a.) represents the allocation of a cooperative game as a forward procedure; (b.) illustrates the allocation of its dual as a backward procedure. The allocation of player 1 (resp. player 2 and 3) is the expected marginal gain (for a cooperative game (D, v)) or cost (for its dual (D, w)) computed for the blue (resp. red and green) ordering positions, weighted according to a probabilistic distribution over \mathcal{S}_D .

The Shapley values of a cooperative game are equal to the ones of its dual (see, [12] Lemma 2.7). The dual of a Sobol’ cooperative game (D, S^{clos}) is the nonnegative, monotonic cooperative

game (D, S^T) , where S^T denotes the total Sobol' indices, given for a model $Y = G(X)$, and for any $A \in \mathcal{P}(D)$, by

$$S_A^T = \frac{\mathbb{E}[\mathbb{V}(Y | X_{\bar{A}})]}{\mathbb{V}(Y)}. \quad (9)$$

The equivalence between the Shapley values of a Sobol' cooperative game and its dual has been highlighted by [39]. Taking S^{clos} or S^T as a value function leads to the same Shapley effects, allowing for alternate estimation schemes. However, it is important to note that this is a particular property of the Shapley values, and it is not inherent to every random order model allocation.

2.3. Detecting exogenous inputs

As noted in [18], the main drawback (for factor fixing setting) of the Shapley effect is their behavior when dealing with exogenous (or spurious) inputs. Formally, exogenous inputs, in the context of variance-based GSA, can be defined as follows.

Definition 2 (L^2 -exogeneity). *Let $X = (X_1, \dots, X_d)$ be random inputs of a model $G : \mathbb{R}^d \mapsto \mathbb{R}$ such that $Y = G(X)$, with Y the random output. Additionally, it is assumed that any input cannot be expressed as a deterministic transformation of some (or all) of the others. Let $E \subset D$. The subset of random inputs X_E are said to be (L^2 -)exogenous to G if, for any $B \subseteq E$, $\exists f \in L^2(P_{X_{\bar{B}}})$ such that:*

$$Y = f(X_{\bar{B}}) \quad a.s.$$

In other words, inputs gathered in a subset X_E are considered to be exogenous to a model G if it is possible to characterize the random output $G(X)$ using a function depending on $X_{\bar{E}}$, for every $B \subseteq E$. This definition is rather intuitive: a set of inputs X_E is considered as exogenous if one can find a function characterizing the random output which do neither involve the inputs in E , nor their possible interaction. It is important to notice that this definition is not too restrictive, since the exogenous inputs X_E and $X_{\bar{E}}$ can still be correlated. In situation where the random inputs are correlated, the Shapley effects can allocate shares of variance to exogenous inputs. This phenomenon, called the *Shapley's joke*, has been illustrated in [22, 18] through the following example.

Example 1 (Shapley's Joke). *Let $X = (X_1, X_2)^\top \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$, $-1 < \rho < 1$, and let the model be:*

$$Y = G(X) = X_1.$$

The Shapley effects of the random inputs are given by

$$Sh_1 = 1 - \frac{\rho^2}{2}, \quad Sh_2 = \frac{\rho^2}{2}.$$

Even if X_2 is an exogenous input, its Shapley effect is not zero as long as $\rho \neq 0$. While this behavior can be considered as valuable in a factor prioritization setting (effects due to correlation can be relevant), it can also be a drawback for spurious variable detection [6]. To overcome this drawback, other allocations can be considered. In particular, the *proportional values* (PV) [30] allow to detect exogenous variable while preserving the factor prioritization power and interpretability of cooperative game theory allocations.

3. From proportional values to proportional marginal effects

The proportional values of positive cooperative games are a particular random order model allocation. They are introduced and extended to nonnegative games in order to be computed for Sobol' cooperative games. This extended allocation applied to the dual of Sobol' cooperative games are introduced as the *proportional marginal effects* (PME). It is then shown that the PME allow to detect exogenous inputs by granting them zero allocation.

3.1. Proportional values as an alternative allocation strategy to Shapley values

The PV of a cooperative game (D, v) is a particular case of random order model allocation. PV arise from a particular case of chosen probability mass function [9], and can also be characterized recursively [11, 30].

Definition 3 (Proportional values). *Let (D, v) be a positive, monotonic cooperative game, where $v : \mathcal{P}(D) \rightarrow \mathbb{R}_*^+$. The proportional values of (D, v) , denoted $\text{PV}(D, v) \in \mathbb{R}^d$, are defined, for every $i \in D$, as a random order model allocation:*

$$\text{PV}_i = \sum_{\pi \in \mathcal{S}_D} p(\pi) [v(C_{\pi(i)}(\pi)) - v(C_{\pi(i)-1}(\pi))] \quad (10)$$

for the following particular choice of probability mass function over \mathcal{S}_D :

$$p(\pi) = \frac{L(\pi)}{\sum_{\sigma \in \mathcal{S}_D} L(\sigma)}, \quad \text{where } L(\pi) = \left(\prod_{j \in D} v(C_j(\pi)) \right)^{-1}. \quad (11)$$

Equivalently, PV can be characterized recursively, for every $i \in D$, as:

$$\text{PV}_i = \frac{R(D, v)}{R(D_{-i}, v)} \quad (12)$$

where, for all $A \in \mathcal{P}(D)$, $R(A, v) = v(A) \left(\sum_{j \in A} R(A_{-j}, v) \right)^{-1}$, and $R(\emptyset, v) = 1$. This recursive definition leads to the following identification [10]:

$$\text{PV}_i = \frac{\sum_{\pi \in \mathcal{S}_{D_{-i}}} \prod_{j=1}^{d-1} v(C_j(\pi))^{-1}}{\sum_{\sigma \in \mathcal{S}_D} \prod_{j=1}^d v(C_j(\sigma))^{-1}}. \quad (13)$$

Remark 1. *The recursive function $R(D, v)$ defined in Definition 3 is better known as a ratio potential in the cooperative game theory literature [11], which is central to define certain allocations. This formulation is especially useful for estimation purposes and when it comes to proving results related to the PV. However, for the sake of conciseness, this notion is considered to be out of the scope of the present work, and is not further discussed.*

They can also be characterized axiomatically (see [10]), as the unique allocation $\phi(D, v)$ respecting the following two axioms:

- **Efficiency:** $\sum_{i=1}^d \phi_i = v(D)$;
- **Equal proportional gains:** for all $A \in \mathcal{P}(D)$, and for all $i, j \in A$, $i \neq j$:

$$\frac{\phi_i((A, v))}{\phi_j((A, v))} = \frac{\phi_i((A_{-j}, v))}{\phi_j((A_{-i}, v))}.$$

Since the PV of positive monotonic games are efficient and nonnegative, they allow for the same meaningful interpretation as the Shapley values, i.e., as shares of $v(D)$. The equal proportional gains axiom allows to better interpret the redistribution dynamic of this particular allocation scheme. For any two different players i and j , the ratio of their allocations in any subgame (A, v) (for every $A \in \mathcal{P}(D)$ such that $i, j \in A$) must be invariant to removing each player's contribution to the other's allocation. In other words, the magnitude of the ratios must be preserved, independently of the possible interaction between i and j , within any coalition they can belong to. This implicitly entails that the

allocation tends to favor the players proportionally to their (marginal) contributions to every possible coalitions in the redistribution process.

As a frame of comparison, the Shapley values can also be characterized as the unique efficient allocation respecting the following axiom (see [10]):

- **Balanced contributions:** for all $A \in \mathcal{P}(D)$, and for all $i, j \in A, i \neq j$:

$$\phi_i(A, v) - \phi_i(A_{-j}, v) = \phi_j(A, v) - \phi_j(A_{-i}, v).$$

This axiom entails that for any two different players i and j , the difference in each allocation by removing the other player to any subgame (A, v) such that $i, j \in A$ must remain equal, for any $A \in \mathcal{P}(D)$. In other words, the difference in allocation of the two players induced by the removal of the other player must be equal, implicitly entailing a balanced redistribution process, where individual and coalitional contributions are favored equally.

Remark 2. *In a nutshell, one can remark that the redistribution processes in both allocations (Shapley values vs. PV) are fundamentally different: the PV redistribution process is proportional meanwhile the Shapley values are egalitarian.*

The different behaviors between PV and Shapley values can be illustrated by considering a two-player game, i.e., $(D = \{1, 2\})$. The allocation are given, for any $i \in D$, by

$$\text{PV}_i((D, v)) = v(\{i\}) + \frac{v(\{i\})}{v(\{1\}) + v(\{2\})} (v(D) - v(\{1\}) - v(\{2\})) \quad (14a)$$

$$\text{Shap}_i((D, v)) = v(\{i\}) + \frac{1}{2} (v(D) - v(\{1\}) - v(\{2\})). \quad (14b)$$

For both PV and Shapley values, each player receives its individual contribution, plus a weighted share of the value surplus generated due to their cooperation. In the literature, this surplus is referred to as the *Harsanyi dividend* of the coalition $\{1, 2\}$ [16]. The main difference between both allocations is the fact that the Shapley values redistribute exactly half of this dividend to each player (i.e., egalitarian way), while the PV redistributes them proportionally (i.e., proportional way) to each player's individual contribution.

It is important to notice that this allocation is only well-defined for positively defined value functions v . However, as stated in Definition 1, the value function of Sobol' cooperative game is inherently nonnegative. The following section presents a continuous extension of the PV to nonnegative games, enabling their use for variance-based GSA purposes.

3.2. Extension of proportional values to nonnegative games

The main contribution of the present work is to propose an adaptation of the PV to variance-based GSA purposes. However, PV are only well defined on cooperative games (D, v) with positive value function v , and Sobol' cooperative games are inherently endowed with nonnegative value functions. However, by leveraging the work proposed by [8], it is possible to define a continuous extension of the PV, allowing their definition for games containing coalitions which have null value. The following result builds upon this extension, and allows to extend the PV to monotonic cooperative games with nonnegative value functions.

Theorem 1 (PV extension to nonnegative games). *Let (D, v) be a nonnegative, monotonic cooperative game with value function $v : \mathcal{P}(D) \rightarrow \mathbb{R}^+$. Let k_{\max} denote the cardinal of the largest null coalition such that:*

$$k_{\max} := \max_{A \in \mathcal{P}(D)} \{|A| : v(A) = 0\}.$$

For any $i \in D$, let \mathcal{K}_{-i} denote the set of null coalitions of D_{-i} , having a cardinal equal to k_{\max} :

$$\mathcal{K}_{-i} := \{A \in \mathcal{P}(D_{-i}) : v(A) = 0 \text{ and } |A| = k_{\max}\}$$

and let $\mathcal{K} = \mathcal{K}_{-\emptyset}$ be the set of null coalitions of D having a cardinal equal to k_{\max} . Moreover, for any $A \in \mathcal{P}(D)$, and for any $B \subseteq \bar{A}$, let $v_A : \mathcal{P}(\bar{A}) \rightarrow \mathbb{R}^+$ be the function defined as:

$$v_A(B) = v(A \cup B).$$

The allocation $\text{PV}^0((D, v)) = (\text{PV}_1^0, \dots, \text{PV}_d^0)^\top \in \mathbb{R}$, defined for any $i \in D$ as:

$$\text{PV}_i^0 = \begin{cases} 0 & \text{if } \forall A \in \mathcal{K}, i \in A, \\ \frac{\sum_{A \in \mathcal{K}_{-i}} R(D_{-i} \setminus A, v_A)^{-1}}{\sum_{A \in \mathcal{K}} R(D \setminus A, v_A)^{-1}} & \text{otherwise,} \end{cases} \quad (15)$$

is a continuous extension of the PV on the set of nonnegative monotonic cooperative games, i.e., for any positive cooperative game (D, v) , one has that:

$$\text{PV}^0((D, v)) = \text{PV}((D, v)).$$

A detailed proof of this result can be found in Appendix [Appendix B](#). Additionally to extending the PV to nonnegative games, [Theorem 1](#) allows to clearly identify players who receives a zero allocation. More precisely, a player i receives a zero allocation if it is part of every largest coalitions with null value.

Remark 3. In the rest of this paper, any mention to the PV refers to their extended version to nonnegative monotonic games (i.e., PV^0), as the usual chosen value functions for variance-based GSA (e.g., typically, the closed Sobol' indices S^{clos}) can indeed be equal to zero.

3.3. Proportional marginal effects and exogeneity detection

From [Theorem 1](#), one can see that the PV of a cooperative game and the PV of its dual are not equal, unlike the Shapley values. In the case of the PV of Sobol' cooperative games, it is important to notice that focusing on the dual is more relevant to better detect exogenous variables (i.e., variables that are irrelevant to the studied model). Indeed, taking the dual of a Sobol' cooperative game, i.e., considering S^T as a value function instead of S^{clos} , allows to better detect exogenous inputs, thanks to the following classical result, echoing the work in [\[17\]](#).

Lemma 1. Let $X = (X_1, \dots, X_d)^\top$ be random inputs and $G \in L^2(P_X)$ denote a model. One has, $\forall A \subset D$,

$$S_A^T = \frac{\mathbb{E}[\mathbb{V}(G(X) \mid X_{\bar{A}})]}{\mathbb{V}(G(X))} = 0 \quad \iff \quad G(X) = \mathbb{E}[G(X) \mid X_{\bar{A}}] \text{ a.s.}$$

A proof of this result can be found in Appendix [Appendix B](#). In other words, for any subset of inputs $A \subseteq D$, whenever $S_A^T = 0$ indicates that $G(X)$ can be expressed as a function only depending on $X_{\bar{A}}$, which is strongly related to the definition of exogenous inputs (i.e., [Definition 2](#)). On the other hand, choosing S^{clos} as a value function leads to the interpretation that $\mathbb{E}[G(X) \mid X_A]$ is constant a.s., which is not necessarily equivalent to exogeneity. Hence, in the goal of exogenous input detection, choosing S^T as a value function is more relevant, especially in the case of the proportional values of Sobol' cooperative games.

This leads to the proposed cooperative game theory-inspired GSA indices called *proportional marginal effects* (PME). They are defined as follows:

Definition 4 (Proportional marginal effects). Let $X = (X_1, \dots, X_d)^\top$ be random inputs, and let $Y = G(X)$ be the random output of a model. The proportional marginal effects are the proportional values of the dual of the monotonic Sobol' cooperative game related to the model G . They are defined as:

$$\text{PME} = \text{PV}((D, S^T)) \in \mathbb{R}^d.$$

The PME are efficient and nonnegative, i.e., they allocate shares of the output’s variance to every inputs, and hence remain meaningful in practice even when inputs are dependent. As explained in Section 3.1, they differ from the Shapley effects on the underlying redistribution principle. More importantly, they allow to detect exogenous inputs to a model, thanks to the following result.

Proposition 1. *Let $X = (X_1, \dots, X_d)^\top$ be random inputs, let $G \in L^2(P_X)$ be a model, and let $E \subseteq D$. If X_E is the largest L^2 -exogenous subset inputs to G among X , then*

$$\text{PME}_i = 0, \forall i \in E, \text{ and } \text{PME}_j > 0, \forall j \in \bar{E}.$$

A proof of this result can be found in Appendix Appendix B. In addition to offer a tool for factor prioritization, the PME allow to detect exogenous inputs by granting them a null allocation. As explained in Section 2.3, they allow to circumvent one of the main practical drawback of the Shapley effects by enabling to detect exogenous inputs, while maintaining their strengths, i.e., a meaningful decomposition of the variance of the output of a numerical model when inputs are dependent. In the following, in order to better understand these novel sensitivity indices, their behavior is studied analytically on toy-cases, and compared to the Shapley effects, in the following section.

4. Illustration on analytical cases

Three analytical models are studied. The first one aims at illustrating the exclusion property of the PME ensured by Proposition 1, which means that PME avoid the Shapley’s joke as illustrated in Example 1. The second test-case is dedicated to highlighting the proportional principle of the PME (see Remark 2), by studying their behavior w.r.t. the magnitude of the linear coefficient of one input in a correlated setting. Finally, the third model introduces a trade-off between individual and interaction effects between two inputs, and further highlights the difference in repartition between the Shapley effects and the proposed PME.

4.1. A linear model with an exogenous input

This first analytical case aims at highlighting the exclusion property of the PME, as ensured by Proposition 1 and motivated by Example 1. This first model reads:

$$Y = G(X) = X_1 + X_2, \quad X = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & \rho \\ 0 & 1 & 0 \\ \rho & 0 & 1 \end{pmatrix} \right) \quad (16)$$

where $-1 < \rho < 1$. One can notice that X_3 is, on purpose, exogenous as it does not intervene explicitly in the computation of Y . It is, however, linearly correlated to X_1 . In such a case, traditional GSA tools such as first- and second-order Sobol’ indices fail to be interpreted as shares of variance [6].

In order to circumvent this problem, one can resort to compute the Shapley effects and PME of the inputs. For this test-case, reference analytical values for both indices are given in Table 4.1.

$\text{Sh}_1 = 1/2 - \rho^2/4$	$\text{PME}_1 = 1/2$
$\text{Sh}_2 = 1/2$	$\text{PME}_2 = 1/2$
$\text{Sh}_3 = \rho^2/4$	$\text{PME}_3 = 0$

Table 2: Reference analytical values for Shapley effects and PME (toy-case 4.1).

From the results provided in Table 4.1, one can first notice that X_3 can receive a non-zero Shapley effect, dependent on the value of ρ . In highly correlated settings, X_3 can be considered to be almost as important as X_1 . This behavior can be meaningful in practice, since X_3 is correlated with X_1 , which is not an exogenous input. However, this interpretation is based on the knowledge of the underlying

model, which is supposed to be black-box. As is, relying only on the Shapley effects, the practitioner would not be able to determine the exogenous nature of the inputs. If the aim of the sensitivity study is focused on better understanding the relationship between the model $G(\cdot)$ and its inputs, somewhat independently from their probabilistic structure, the Shapley effects are hence not suitable alone.

One can notice that the PME does indeed detect X_3 as being an exogenous input, by granting it a zero allocation. Moreover, in this setting, the PME is not influenced by the linear correlation ρ between X_1 and X_3 . In combination with the Shapley effects, additional insights on G can be extracted from the initial study: while X_3 can have an effect on G through its correlation with other inputs, it is exogenous to G . Additionally, by allocating half the output's variance to both X_1 and X_2 , the PME also indicates an equal influence. Hence, by combining the interpretation of both indices, one can interpret these results as follows: X_3 is an exogenous variable (PME), but it bears an effect on G through its dependence with other inputs (Shapley effects) and, moreover, X_1 and X_2 seem to bear an equal influence on the output's variance, whenever X_3 is detected as exogenous (PME).

It is important to note that neither G nor the dependence structure between the inputs need to be known for this interpretation. Hence, both indices are complementary and allow for a more precise interpretation of the studied model and its interaction with the inputs and their probabilistic structure.

4.2. Unbalanced linear model

Echoing Remark 2, beyond the detection of exogenous inputs, the Shapley effects and the PME fundamentally differ on their redistribution process. While the Shapley effects allocate importance in an egalitarian fashion, the PME follows a proportional principle. This toy-case aims at highlighting this difference, by introducing a coefficient in a linear model with three correlated Gaussian inputs. This use-case is referred to as *unbalanced* since the three linear coefficient are different. This toy-case writes:

$$Y = G(X) = X_1 + \beta X_2 + X_3, \quad X = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \rho \\ 0 & \rho & 1 \end{pmatrix} \right), \quad (17a)$$

$$\mathbb{V}(Y) = 2 + \beta^2 + 2\rho\beta. \quad (17b)$$

The analytical shares of output variance, according to the Shapley effects and the PME are given in Table 4.2.

$\mathbb{V}(Y) \times \text{Sh}_1 = 1$	$\mathbb{V}(Y) \times \text{PME}_1 = 1$
$\mathbb{V}(Y) \times \text{Sh}_2 = \beta^2 + \beta\rho + \frac{1}{2}\rho^2(1 - \beta^2)$	$\mathbb{V}(Y) \times \text{PME}_2 = \frac{\beta^2(1+\beta^2+2\rho\beta)}{(1+\beta^2)}$
$\mathbb{V}(Y) \times \text{Sh}_3 = 1 + \beta\rho - \frac{1}{2}\rho^2(1 - \beta^2)$	$\mathbb{V}(Y) \times \text{PME}_3 = \frac{(1+\beta^2+2\rho\beta)}{(1+\beta^2)}$

Table 3: Reference analytical values for Shapley effects and PME (toy-case 4.2).

One can notice that, by considering the *balanced* case (i.e., $\beta = 1$), the Shapley effects and PME are equal. However, as soon as the model is unbalanced, one can notice that both allocations behave in a completely different fashion as soon as ρ approaches 1. Using an asymptotic-analysis-based reasoning, one can obtain the following set of resulting approximation:

$$\text{Sh}_1 = \text{PME}_1 \xrightarrow{\rho \rightarrow 1} \frac{1}{2 + \beta^2 + 2\beta} \xrightarrow{\beta \rightarrow \infty} 0, \quad (18a)$$

$$\text{Sh}_2, \text{Sh}_3 \xrightarrow{\rho \rightarrow 1} \frac{\frac{1}{2}\beta^2 + \beta + \frac{1}{2}}{2 + \beta^2 + 2\beta} \xrightarrow{\beta \rightarrow \infty} \frac{1}{2}, \quad (18b)$$

$$\text{PME}_2 \xrightarrow{\rho \rightarrow 1} \frac{\beta^2(1 + \beta^2 + 2\beta)}{(2 + \beta^2 + 2\beta)(1 + \beta^2)} \xrightarrow{\beta \rightarrow \infty} 1, \quad (18c)$$

$$\text{PME}_3 \xrightarrow{\rho \rightarrow 1} \frac{(1 + \beta^2 + 2\beta)}{(2 + \beta^2 + 2\beta)(1 + \beta^2)} \xrightarrow{\beta \rightarrow \infty} 0. \quad (18d)$$

In other words, in extreme cases of positive linear correlation between X_2 and X_3 , the Shapley effects allocates half the importance to each input despite a fairly high β value in favor of X_2 . The PME, on the other hand, tend to favor X_2 by granting it the whole variance, despite the high correlation with X_3 . This behavior highlights the “egalitarian vs. proportional” behavior of both types of effects: the Shapley effects tend to consider X_2 and X_3 as equally important due to their high correlation, while the PME favor X_3 in regards of its high linear coefficient.

While these results inform on the asymptotic behavior of both indices, their difference can also be highlighted for punctual values of ρ and β . Figure 2 illustrates the behavior of both indices w.r.t. ρ , for two different values of β (namely, 2 and 10). Whenever $\beta = 2$, one can notice that PME_2 increases w.r.t. ρ , while Sh_2 decreases after $\rho \simeq -0.24$, and both indices are concave w.r.t. ρ . On the other hand, Sh_3 is convex w.r.t. ρ and becomes increasing at $\rho \simeq -0.54$, while PME_3 remains concave increasing. At extreme values of ρ (i.e., close to -1 or to 1), one can notice that Sh_2 and Sh_3 are considered equally important. Furthermore, one can notice that $\text{PME}_2 > \text{PME}_3$, whatever the magnitude of their correlation. Increasing β to 10 exacerbates this behavior of the Shapley effects. However, the PME behave differently: X_1 and X_3 are given a negligible part of variance, while X_2 is granted a seemingly constant share, w.r.t. ρ , hovering around 98%.

In conclusion, in this unbalanced case, the proportional redistribution property of the PME allows for a clearer importance hierarchy, even in situation of extreme correlation. On the other hand, the Shapley effect tends to the even importance out between the correlated inputs, leading to a potentially indecisive importance hierarchy.

4.3. Unbalanced linear model with interactions

This test-case aims at studying and comparing the behavior in a trade-off between individual and interaction effects. A parameter $\alpha \in [0, 1]$ is introduced as part of a linear model comprising an interaction term. This particular unbalanced linear model is given as:

$$Y = G(X) = X_1 + (1 - \alpha)X_2 + \alpha X_1 X_2, \quad X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right), \quad (19a)$$

$$\mathbb{V}(Y) = 2 + (1 - \alpha)^2 + 2(1 - \alpha)\rho + \rho^2. \quad (19b)$$

In other words, α controls the trade-off between the individual effect of X_2 and its interaction term with X_1 . When $\alpha = 0$, the model becomes balanced between X_1 and X_2 , and when $\alpha = 1$, X_2 only interacts with X_1 . In both cases, X_1 remains present in the model. Analytical formulas for the Shapley effects and the PME are given in Table 4.3.

$2\mathbb{V}(Y) \times \text{Sh}_1 = 3 + \rho^2(1 - \alpha)^2 + 2\rho(1 - \alpha)$	$\text{PME}_1 = \frac{2}{3 + (1 - \alpha)^2}$
$2\mathbb{V}(Y) \times \text{Sh}_2 = 1 + 2\rho^2 + (2 - \rho^2)(1 - \alpha)^2 + 2\rho(1 - \alpha)$	$\text{PME}_2 = \frac{(1 - \alpha)^2 + 1}{3 + (1 - \alpha)^2}$

Table 4: Reference analytical values for Shapley effects and PME (toy-case 4.3).

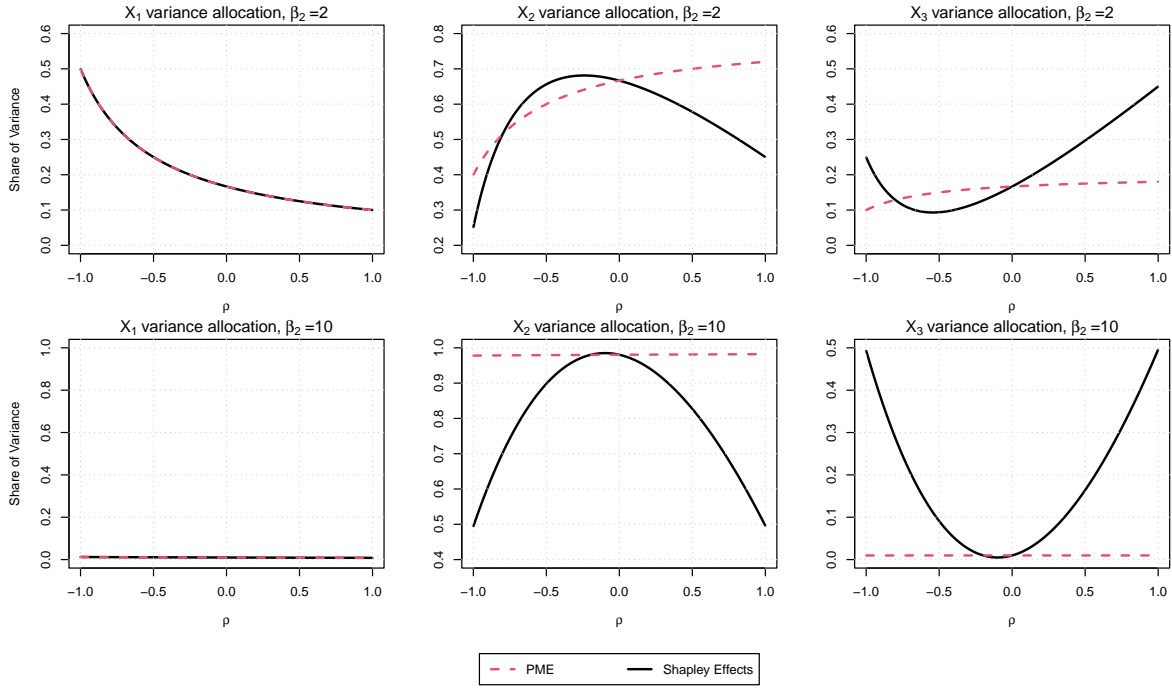


Figure 2: PME and Shapley effects w.r.t. ρ (toy-case 4.2). Top row depicts the allocations for $\beta = 2$ while the bottom row is for $\beta = 10$.

When $\alpha = 0$ the model is balanced, and the Shapley effects and the PME are equal, allocating half of the variance to each input. To better illustrate the redistribution principles w.r.t. both correlation and interaction, (α, ρ) -plane plots are provided, for each effect in Figure 3. One can notice that Sh_1 and PME_1 are increasing w.r.t. α , but Sh_2 and PME_2 show a decreasing behavior. However, the Shapley effects effectively depend on the correlation coefficient ρ in addition to α , while the PME only depend on α .

Focusing on the behavior of both effects w.r.t. the interaction, one can first focus on the α -axis of the plots in Figure 3. Whenever α is close to 0, one can notice that both indices tend to allocate an equal share of the output variance to both inputs. As α increases, the PME grants an increasing share of the output variance to X_1 , from half to two thirds of the variance, and a decreasing share to X_2 , from half to one third. As explained, the PME does not depend on the correlation between both inputs in this case.

Focusing on the Shapley effects, low values of α seem to share equally the variance. However, the higher α becomes, the more the sharing mechanism depends on ρ . When ρ is between -0.5 and 0.5 , and α is close to 1, Sh_1 increases, with a maximum allocation of 0.75 taken at $(\rho = 0, \alpha = 1)$, while Sh_2 decreases, with a minimum allocation of 0.25 at the same point. The redistribution proposed by the Shapley values is sensible in this case, since that when $\alpha = 1$, X_1 is endowed with its individual effect, and an interaction effect with X_2 , while X_2 only interacts with X_1 in this model. However, this interpretation is subject to a moderate level of correlation. As soon as the correlation rises, for values of ρ close to 1 and -1 , the Shapley effects still grant half the output variance to both inputs, whatever the value of α is.

In this toy-case, the interpretation of the PME is more robust to high level of correlation if the goal of the study is to gather insights on the intricacies of the model G . On the other hand, the Shapley effects tend to even the importance between the correlated inputs whenever their correlation levels are fairly high, resulting in a potentially indecisive importance ranking.

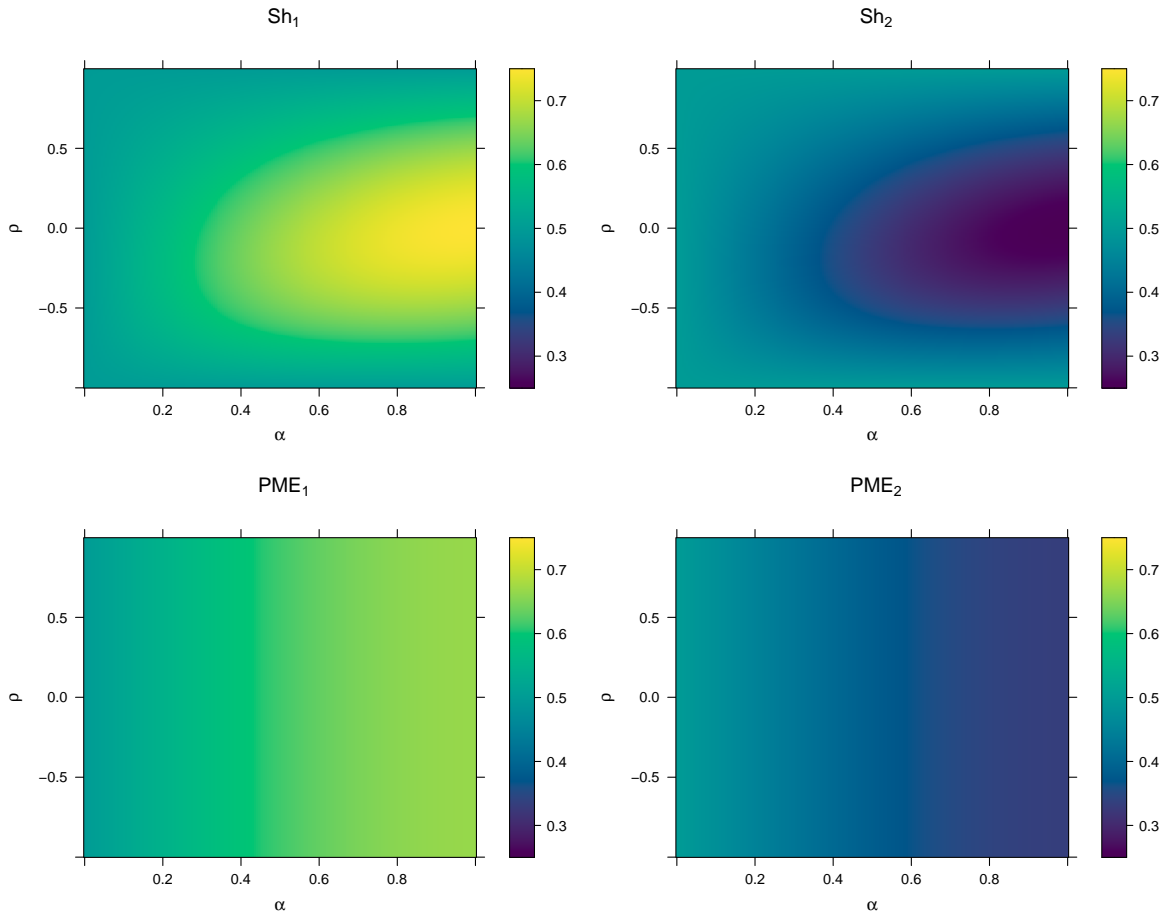


Figure 3: PME and Shapley effects in the (α, ρ) -plane for test-case 4.3.

4.4. First conclusions

From these few tests of linear models with a small number of correlated Gaussian inputs, and with the inclusion of simple interaction, the following conclusions can be drawn:

- whenever the inputs are correlated, the Shapley effect does not detect exogenous inputs, while the PME does;
- in the highly correlated cases, the Shapley effect can lead to indecisive importance ranking, while the PME allows for a more pronounced hierarchy in accordance with the model G ;
- overall, the PME is less sensitive to correlations, while the Shapley effect can vary greatly.

However, these conclusion are subject to the very specific presented toy-cases. The following section presents more ambitious models, closer to real-world applications, where both effects are estimated instead of computed analytically.

5. Estimation and numerical results

In this section, estimation schemes of the PME are presented: they rely on the same ingredients as the Shapley effects. Then, two numerical cases are studied using these schemes: a modified Ishigami function and a robot arm model.

5.1. Estimation strategies

The plug-in estimation of the PME relies on the exact same elements than the estimation of the Shapley effects. For the sake of completeness, the classical estimation framework is briefly stated. Following the two-steps methodology presented in [2], initially developed for Shapley effects' estimation, one can estimate the PME in two distinct steps:

- Step 1: Estimate the *conditional elements*, i.e., $S_A^T, \forall A \in \mathcal{P}(\mathcal{D})$;
- Step 2: Perform an *aggregation procedure* via a direct plug-in of the estimated conditional elements in Eq. (15).

Only the aggregation procedure differs between the estimation of the PME and the Shapley effects. This entails that the estimation cost in terms of model evaluations is exactly the same for the PME than for the Shapley effects. Furthermore, both indices can be evaluated “at-once”, by using the same conditional elements estimates. Two estimation schemes concerning the conditional elements are presented, followed by a break-down of the algorithmic logic behind the specific aggregation procedure of the PME.

A first estimation scheme for the total Sobol' indices relies on a double Monte Carlo procedure. It can be found in [39]. This estimator requires the ability to randomly sample from every possible conditional distributions of the conditional random variables $X_A|X_{\bar{A}}$, from every marginal distributions, i.e., to simulate i.i.d. observations of X_A , for all $A \in \mathcal{P}(\mathcal{D})$, as well as from the joint distribution of the random variable X . For any $A \subseteq D$, the required i.i.d. samples to estimate S_A^T are the following:

- an i.i.d. sample of size N_v of X , denoted by $(X^{(1)}, \dots, X^{(N_v)})$;
- another i.i.d. sample of size N_o drawn from $X_{\bar{A}}$ and denoted by $(X_{\bar{A}}^{(1)}, \dots, X_{\bar{A}}^{(N_o)})$;
- for each element $X_{\bar{A}}^{(i)}, i = 1, \dots, N_o$, a corresponding sample of size N_i drawn from $X_A|X_{\bar{A}} = X_{\bar{A}}^{(i)}$ denoted by $(X_{A|\bar{A},i}^{(1)}, \dots, X_{A|\bar{A},i}^{(N_i)})$. Additionally, for $i = 1, \dots, N_o, j = 1, \dots, N_i$, one conditional observation of the input $k \in D$ is given by

$$\tilde{X}_{(\bar{A},i),k}^{(j)} = \begin{cases} X_{\bar{A}}^{(i)} & \text{if } k \in \bar{A}, \\ X_{A|\bar{A},i}^{(j)} & \text{otherwise.} \end{cases}$$

Subsequently denote, for $j = 1, \dots, N_i$, a full conditional observation w.r.t. $X_{\bar{A}}^{(i)}$:

$$\tilde{X}_{\bar{A},i}^{(j)} = \left(\tilde{X}_{(\bar{A},i),k}^{(j)} \right)_{k=1,\dots,d} \in \mathbb{R}^d.$$

In other words, the $(N_i \times d)$ bloc-matrix composed of every full conditional observation, i.e.,

$$\tilde{X}_{\bar{A},i} = \begin{pmatrix} \tilde{X}_{\bar{A},i}^{(1)} \\ \vdots \\ \tilde{X}_{\bar{A},i}^{(N_i)} \end{pmatrix}$$

is a concatenation of columns of conditionally simulated samples (i.e., simulations from $X_A|X_{\bar{A}} = X_{\bar{A}}^{(i)}$) and columns composed of the values by which the distribution is conditioned on (i.e., $X_{\bar{A}}^{(i)}$).

Let $A \subseteq D$. One can first estimate $\mathbb{V}(G(X))$ using the unbiased classical variance estimator:

$$\widehat{\mathbb{V}(G(X))} = \frac{1}{N_v - 1} \sum_{i=1}^{N_v} \left(G(X^{(i)}) - \overline{G(X)} \right)^2$$

with $\overline{G(X)} = \frac{1}{N_v} \sum_{i=1}^n G(X^{(i)})$. Then, for every observation $i = 1, \dots, N_o$ of the sample of $X_{\overline{A}}$, the conditional variance $\mathbb{V}(G(X) \mid X_{\overline{A}} = X_{\overline{A}}^{(i)})$ is estimated by

$$\widehat{V}_{\overline{A}}^{(i)} := \frac{1}{N_i - 1} \sum_{j=1}^{N_i} \left(G(\tilde{X}_{\overline{A},i}^{(j)}) - \frac{1}{N_i} \sum_{k=1}^{N_i} G(\tilde{X}_{\overline{A},i}^{(k)}) \right)^2.$$

Finally, a consistent estimate of $S_{\overline{A}}^T$ is given by

$$\widehat{S}_{\overline{A}}^T = \left(\frac{1}{N_o} \sum_{i=1}^{N_o} \widehat{V}_{\overline{A}}^{(i)} \right) / \widehat{\mathbb{V}(G(X))}.$$

The ability to sample from these conditional distributions can be difficult in practice, especially in cases where only an i.i.d. sample of the input/output is available. A given-data estimation scheme for the conditional elements has then been proposed in the literature. It relies on the approximation of the conditional samples using a nearest-neighbor scheme, and thus allows to estimate every $S_{\overline{A}}^T$ with only one observed i.i.d. sample of X and its corresponding model output. For the sake of conciseness, this methodology is not detailed in the present paper, but is used in the use-case presented in Section 5.3. One can refer to [3, 19] for additional theoretical and computational details on this estimation method.

Given an estimation $\widehat{S}_{\overline{A}}^T$ for every conditional element $A \subseteq D$, the aggregation procedure leading to the PME can be computed using its recursive definition (see, Eq. (15)). It relies on the computation of the ratio potential, i.e., the function R in Eq. (12), as described in Algorithm 1.

Algorithm 1 Computation of the ratio potential $R(A, v)$.

Require: $A \subseteq D, v$.

- 1: $a \leftarrow |A|$
 - 2: $P^- \leftarrow (v(\{1\}), \dots, v(\{a\})) \in \mathbb{R}_+^d$
 - 3: **for** $k = 2, \dots, a - 1$ **do**
 - 4: $m \leftarrow \binom{a}{k}$
 - 5: $P \in \mathbb{R}^m$
 - 6: **for** $B \subset A$ s.t. $|B| = k$ **do** (possibly in parallel)
 - 7: $\mathcal{J} \leftarrow \{C \subset B \mid \forall j \in B, C = B \setminus \{j\}\}$
 - 8: $P_B = v(B) (\sum_{C \in \mathcal{J}} (P_C^-)^{-1})^{-1}$
 - 9: **end for**
 - 10: $P^- \leftarrow P$
 - 11: **end for**
 - 12: **return** $R(A, v) \leftarrow v(A) (\sum_{j=1}^m (P_j)^{-1})^{-1}$
-

With the ability to compute the ratio potential for any subset A and any function v , one can proceed to compute the PME, as detailed in Algorithm 2.

5.2. Ishigami model with a correlated exogenous input

In order to further study the behavior of the PME, the Ishigami model, well-known in GSA (see, e.g., [6]), is first considered. The Ishigami model is given by

$$G(X) = \sin(X_1) + 7 \sin^2(X_2) + 0.1 X_3^4 \sin(X_1).$$

In our study, the following probabilistic structure of the inputs is considered:

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_4 \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} (\pi/3)^2 & 0 & 0 & \rho \\ 0 & (\pi/3)^2 & 0 & 0 \\ 0 & 0 & (\pi/3)^2 & 0 \\ \rho & 0 & 0 & (\pi/3)^2 \end{pmatrix} \right).$$

Algorithm 2 PME plug-in estimation.

Require: $D, \widehat{S}_A^T, \forall A \in \mathcal{P}_D$.

- 1: $d \leftarrow |D|$
 - 2: $\text{PME} \in \mathbb{R}^d$
 - 3: $k_{\max} = \max_{A \subseteq D} \{|A| : \widehat{S}_A^T = 0\}$
 - 4: $\mathcal{K} = \{A \subseteq D : |A| = k_{\max}, \widehat{S}_A^T = 0\}$
 - 5: $E = \{i \in D : \forall A \in \mathcal{K}, i \in A\}$
 - 6: $\widetilde{S}_B^T(\cdot) \leftarrow \widehat{S}_{B \cup \cdot}^T$.
 - 7: **for** $i \in E$ **do**
 - 8: $\text{PME}_i = 0$
 - 9: **end for**
 - 10: $R_{\mathcal{K}} \leftarrow -\sum_{A \in \mathcal{K}} \left(R \left(D \setminus A, \widetilde{S}_A^T(\cdot) \right) \right)^{-1}$
 - 11: **for** $j \in D \setminus E$ **do** (possibly in parallel)
 - 12: $\mathcal{K}_j \leftarrow \{A \subseteq D_{-j} \mid |A| = k_{\max}, v(A) = 0\}$
 - 13: $\text{PME}_j \leftarrow \sum_{A \in \mathcal{K}_j} \left(R \left(D_{-j} \setminus A, \widetilde{S}_A^T(\cdot) \right) \right)^{-1} / R_{\mathcal{K}}$
 - 14: **end for**
 - 15: **return** PME
-

One can notice that X_4 is, by design, an exogenous input, but it is linearly correlated to X_1 by means of the parameter $\rho \in (-1, 1)$. The Shapley effects and the PME are estimated using a Monte Carlo procedure, as presented in Section 5.1, with chosen sample sizes $N_v = 10^5$, $N_o = 2 \times 10^3$ and $N_i = 300$, for various values of ρ (from -0.99 to 0.99 with a step of 0.01). Each Monte Carlo estimation has been independently repeated 200 times in order to obtain confidence intervals. The results are provided in Fig. 4.

First, one can notice a strong influence of X_2 , whose PME and Shapley effects are equal and constant along ρ . This result is expected, since X_2 has no interaction or correlation with other inputs in the Ishigami model, and hence its importance should not be subject to variation w.r.t. the correlation intensity. Second, focusing on X_1 and X_4 , one can notice the same behavior of the Shapley effects as depicted previously. Despite the fact that X_4 is exogenous, in situation of extreme correlation, Sh_4 can be as high as Sh_1 , which echoes the results in [22], but effectively grants a zero allocation to X_4 whenever both inputs are independent (i.e., $\rho = 0$). However, their PME differ, in the sense that PME_1 is constant w.r.t. ρ , while PME_4 is equal to zero whatever the correlation value. Hence X_4 is effectively detected as being exogenous. Third, one can notice that Sh_3 does vary w.r.t. ρ , which can be understood by the fact that X_3 interacts with X_1 in the model, which is itself correlated to X_4 . However, since the PME detects X_4 as being exogenous, PME_3 remains constant w.r.t the correlation structure. Finally, focusing on X_3 and X_1 whenever $\rho = 0$ (i.e., the inputs are independent), one can notice that $\text{Sh}_1 > \text{PME}_1$, and $\text{Sh}_3 < \text{PME}_3$. This can be understood as the expression of the proportional versus the egalitarian redistribution schemes. While the Shapley effects effectively grants half the interaction surplus to both inputs, the PME tend to favor X_3 . This can be understood by the fact that X_1 does not have an overwhelmingly higher overall effect on G than X_3 .

Overall, the PME are less sensitive to the correlation of exogenous inputs than the Shapley effects. In conclusion, this toy-case highlights further the fact that both effects are complementary when it comes to a more precise interpretation of the model when its inputs are correlated. It reinforces the previously found behavioral tendencies in a less straightforward model.

5.3. Robot arm model

In this use-case, one studies a model of the position (on the two-dimensional plane) of a robot arm with four segments [1]. The arm shoulder is fixed at the origin and the robot's segments have

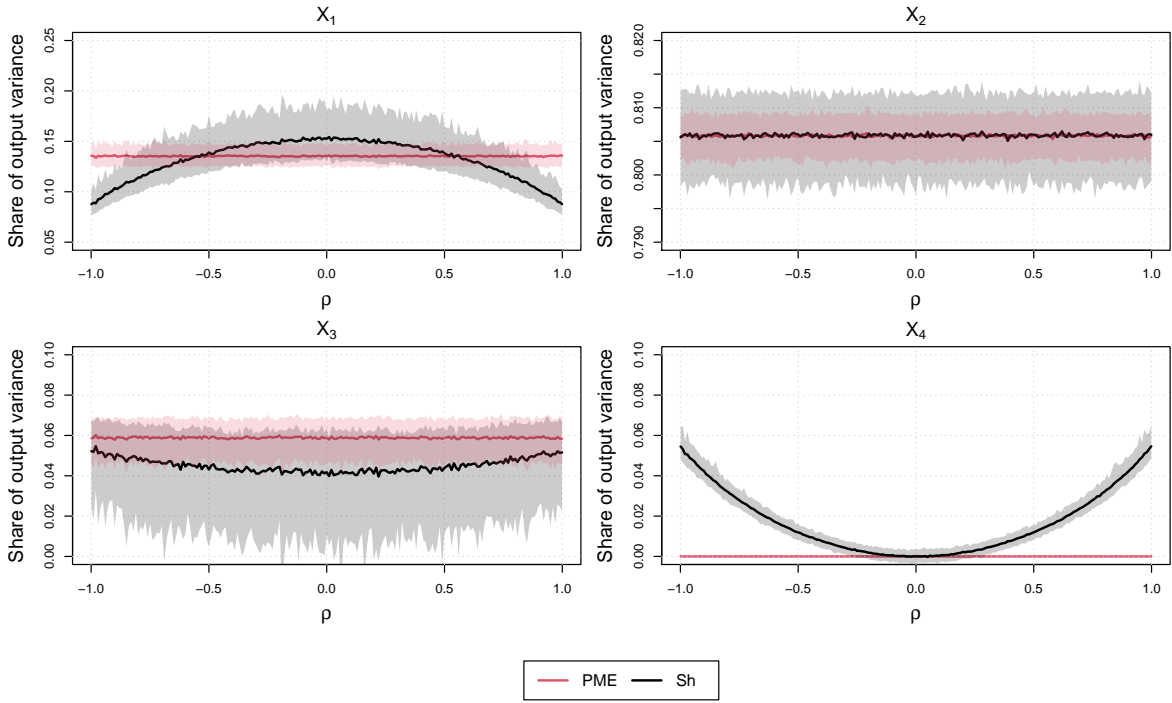


Figure 4: PME and Shapley effects for the Ishigami model with a spurious variable, with respect to the correlation coefficient ρ between X_1 and X_4 . The grey and red areas around the solid plots give the 95%-confidence intervals of the estimates.

lengths L_i ($i = 1, \dots, 4$). Each segment is positioned at an angle A_i ($i = 1, \dots, 4$) with respect to the horizontal axis. While, in the original model, the inputs are assumed to be independent, statistical dependence is introduced here between the angles and between the segment lengths. The probabilistic structure on the inputs can be described as follows:

- The angles A_i ($i = 1, \dots, 4$) follow a uniform distribution over $[0, 2\pi]$. They are pairwise correlated by the way of a Gaussian copula with correlation parameter equal to 0.95;
- The lengths are sequentially built: L_1 follows a uniform distribution over $[0, 1]$, while L_i ($i = 2, \dots, 4$) follows a uniform distribution over $[0, L_{i-1}]$. These inequality constraints create strong correlation between the lengths.

The model's output is the distance of the end of the robot arm to the origin, and writes:

$$Y = \left\{ \left[\sum_{i=1}^4 L_i \cos \left(\sum_{j=1}^i A_j \right) \right]^2 + \left[\sum_{i=1}^4 L_i \sin \left(\sum_{j=1}^i A_j \right) \right]^2 \right\}^{1/2}.$$

A unique i.i.d. sample of size 2000 of these 8 inputs has been simulated, on which the output of the model has been computed. Figure 5 illustrates this data sample by the way of the pairwise scatter-plots, the marginal distributions of each input by means of histograms, and the dependence structure with estimated correlation coefficients. One can also notice first-order tendencies of the different inputs on the output Y (last row).

Since, in this scenario, only an i.i.d. sample is available, the Shapley effects and PME have been computed using the nearest-neighbor procedure eluded in Section 5.1 (with an arbitrarily chosen number of nearest neighbors equal to 6). Figure 6 displays the Shapley effects and PME estimates with

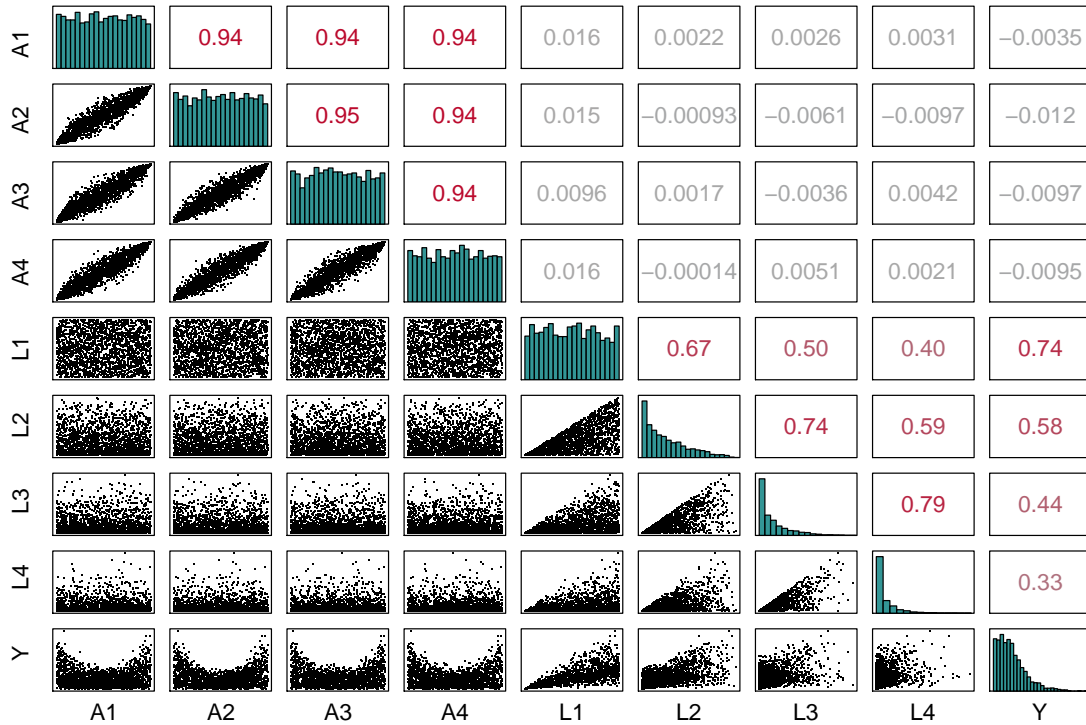


Figure 5: Pairwise plots of the 2000-size i.i.d. sample for the robot arm model. It gives the histograms of the variables (diagonal), as well as the scatterplots (lower part) and linear correlation coefficients (upper part) between all the variables pairs.

a 90%-confidence intervals computed on 100 replications of estimated effects by random selection of 80% of the dataset's observations. According to both effects, the most influential input is L_1 , with

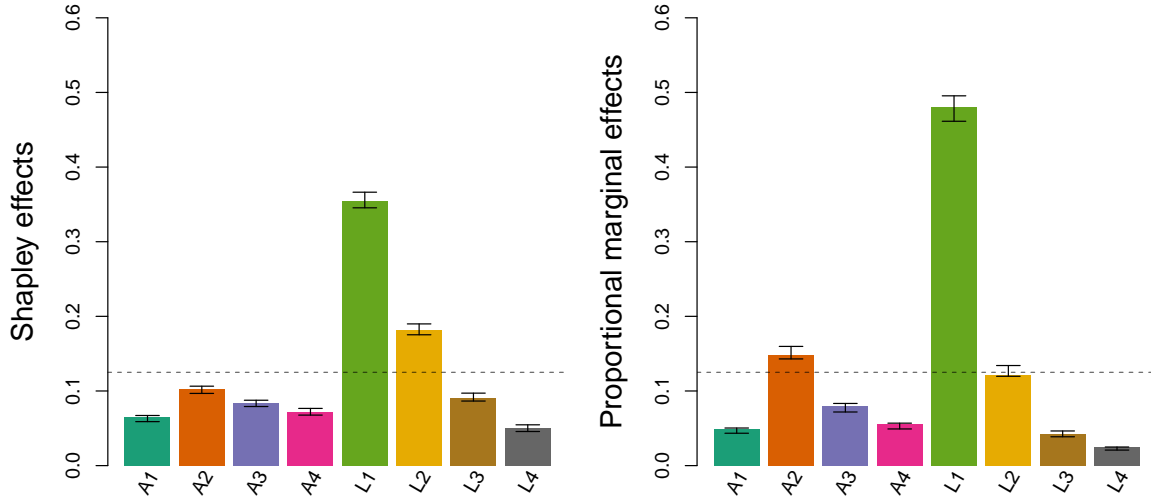


Figure 6: Shapley effects and PME estimators by nearest-neighbor procedure for the robot arm model. The vertical error bars represent the 90%-confidence intervals of the estimates. The horizontal grey dashed bar represents the average importance of $1/8$.

Shapley effect around 35% and PME around 48% of the output's variance. While both effects seem

to agree on the most influential input, they offer a fairly different influence hierarchy, as depicted in Table 5. This different influence hierarchy can be explained by the fairly high correlation between the inputs. For instance, L_2 has Shapley effects around 18% while having a linear correlation coefficient with L_1 equal to 0.67, whereas it has a PME of around 12.1%. Additionally, focusing on the angle inputs which are very linearly correlated, one can notice that their Shapley effects are relatively equal, varying between 6% and 10%. On the other hand, their PME grants A_2 nearly 15%, with reduced influence of the other angles. Only using the Shapley effects did not consider A_2 as an above-average influential variable, while the PME consider it as important. This highlights the more decisive ability of the PME for influence ranking in situations of highly correlated inputs, where the Shapley effects typically grants a similar output variance share to each correlated input.

Influence Rank	Shapley effects		PME	
	Input	Value	Input	Value
1	L_1	35.4%	L_1	48%
2	L_2	18.2%	A_2	14.9%
3	A_2	10.2%	L_2	12.1%
4	L_3	9%	A_3	8%
5	A_3	8.4%	A_4	5.5%
6	A_4	7.2%	A_1	4.9%
7	A_1	6.4%	L_3	4.2%
8	L_4	5.1%	L_4	2.4%

Table 5: Influence hierarchy between the inputs w.r.t. the Shapley effects and the PME, on the robot arm use-case.

Moreover, one can notice that no input is considered as exogenous, which is reassuring since every input is effectively involved in the computation of the output. However, for both effects, L_4 is the least influential, with Shapley effect of around 5% and PME around 2%.

To conclude, this use-case illustrates the more decisive discriminating power of the PME compared to the Shapley effects, in cases of highly correlated inputs. Overall, the PME favor the already most influential inputs at the expense of the inputs they are correlated with. This behavior is particularly interested in a screening setting, along with the ability of the PME to detect exogenous inputs, while maintaining a meaningful interpretation as shares of variance.

6. Discussion and perspectives

The main contribution of this paper is the application to GSA of the proportional values, an alternate allocation from cooperative game theory. An extension of the original allocation is proposed, in order to be used for Sobol' cooperative games, leading to novel GSA indices: the proportional marginal effects. They fundamentally differ from the Shapley effects in two ways. First, it is proved that they detect exogenous inputs, by granting them zero allocation despite their eventual correlation. Second, they allow for more decisive influence hierarchies than the Shapley effects, especially when inputs are highly correlated. They remain intrinsically interpretable, as shares of variance of the model's output.

It has been shown, by means of analytical toy-cases and use-cases, that the proposed PME allow for a more complete depiction of the intricacies of black-box models. However, this testing remains quite limited, and does not allow to derive general truth statements on their behavior. More in-depth studies (such as [22] for the Shapley effects) on challenging use-cases are mandatory in order to comfort their usefulness in practice.

Two estimation schemes are proposed, based on Monte Carlo sampling scheme or given-data using a nearest-neighbor procedure, initially developed for the Shapley effects. Moreover, their computation rely on the same conditional elements' estimations than the Shapley effects: their computation does

not require additional sampling or model evaluations. However, the computational burden associated to their estimation remains a drawback. They require the calculation of an exponential number $(2^d - 1)$ of Sobol indices. The exact same problem has been highlighted for the Shapley effects estimation.

This present work revolves around the idea of finding alternate allocation rules than the Shapley values which, while mainstream, may not be universally suitable. Introducing the framework of random order models, which allows to define allocations by simply choosing a probability mass function over the set of permutations of D , allows to further explore the richness of cooperative game theory. The subsequent diversity of properties inherited from the different allocations could be an interesting way to get complementary viewpoints on variance explanation. As seen in this paper, the Shapley effects and the PME are designed to give different messages, the interest of which depends on the UQ task one is dealing with. While the PME is a reasonable option for factor fixing and factor prioritization, the Shapley effects provide a tool for model exploration that allows for a good overview of all the inputs that might have an impact on the output, even though it is only due to correlation with other inputs. Other allocations, such as weighted Shapley values [25], may be defined with different specific UQ tasks in mind, allowing for domain-specific tools for more accurate and relevant sensitivity studies.

As discussed in [34] and highlighted in [21], machine learning (ML) interpretability and sensitivity analysis bear many resemblances. The PME are in part inspired from the *proportional marginal value decomposition*, introduced in [9] and further studied in [14, 15], which are importance measures related to linear regression models. Hence, their formal introduction to the field of sensitivity analysis is an expression of the relevant similarities between both fields and the exciting research opportunities that can arise by taking inspiration from either research domains.

Finally, while the cooperative game theory is an evident source for novel ideas, its intricacies remain poorly understood by both the UQ and ML interpretability communities. Cooperative games in general, and the construction of allocations in particular, are inherently player-centric, while UQ and interpretability studies have historically been model-centric. Despite the evident usefulness of cooperative game theory for defining relevant and interpretable tools, further works must be put into justifying their meaningfulness to broaden their use for practical studies.

Appendix A. Software and reproducibility of results

All the numerical tests have been performed using the R programming language. Every results and figures presented in this paper can be reproduced by means the openly accessible codes in a GitLab repository¹, as well as details on the packages used.

Appendix B. Proofs

Proof of Theorem 1. Let (D, v) be a nonnegative, monotonic cooperative game, let $A \subseteq D$ be a coalition, and denote $a = |A|$ the cardinal of A . Denote \mathcal{S}_A the set of permutations of players in A . Let $\pi \in \mathcal{S}_A$, and for the sake of clarity, denote $|\pi| = a$, the number of elements in the permutation. By monotonicity, one has, $\forall j \in \{1, \dots, |\pi|\}$,

$$v(C_{j+1}(\pi)) \geq v(C_j(\pi)) \geq 0.$$

Moreover, assume that $v(C_0(\pi)) = v(\emptyset) = 0$. Thus, there exists a level $k_\pi(v) \in \{0, 1, \dots, |\pi|\}$ such that $v(C_{k_\pi(v)}(\pi)) = 0$ and $v(C_{k_\pi(v)+1}(\pi)) > 0$. Formally:

$$k_\pi(v) = \begin{cases} \max\{j \in \{1, \dots, |\pi|\} : v(C_j(\pi)) = 0\} & \text{if } v(C_1(\pi)) = 0, \\ 0 & \text{otherwise.} \end{cases}$$

¹<https://gitlab.com/milidris/PME>

For the sake of conciseness and readability, the argument v is omitted and the notation $k_\pi := k_\pi(v)$ is adopted. Let $(\epsilon_p)_{p \in \mathbb{N}}$ be a sequence such that:

$$\forall p \in \mathbb{N}, \epsilon_p > 0, \quad \text{and} \quad \lim_{p \rightarrow \infty} \epsilon_p = 0.$$

Let $((D, v_p))_{p \in \mathbb{N}}$ be a sequence of positive, monotonic cooperative games defined, for any $p \in \mathbb{N}$ and for any $A \subseteq D$, as:

$$v_p(A) = \begin{cases} \epsilon_p & \text{if } v(A) = 0, \\ v(A) & \text{otherwise.} \end{cases}$$

Alternatively, one can notice that, $\forall \pi \in \mathcal{S}_A, \forall j \in \{1, \dots, |\pi|\}$,

$$v_p(C_j(\pi)) = \begin{cases} \epsilon_p & \text{if } j \leq k_\pi, \\ v(C_j(\pi)) & \text{otherwise.} \end{cases} \quad (\text{B.1})$$

Let $p \in \mathbb{N}$, and from the recursive definition of the PV (see, Definition 3) of the positive games (D, v_p) , one has, for any $i \in D$:

$$\text{PV}_i = \frac{\sum_{\pi \in \mathcal{S}_{D-i}} \prod_{m=1}^{d-1} v_p(C_m(\pi))^{-1}}{\sum_{\sigma \in \mathcal{S}_D} \prod_{m=1}^d v_p(C_m(\sigma))^{-1}}.$$

For the sake of conciseness and clarity, for any $\pi \in \mathcal{S}_A, \forall A \subseteq D$, let us introduce the following notation:

$$\Upsilon_k^l(\pi, v) = \begin{cases} \prod_{j=k}^l v(C_j(\pi))^{-1} & \text{if } k \leq l, \\ 1 & \text{otherwise.} \end{cases}$$

One then has that, for any $i \in D$:

$$\text{PV}_i = \frac{\sum_{\pi \in \mathcal{S}_{D-i}} \Upsilon_1^{d-1}(\pi, v_p)}{\sum_{\sigma \in \mathcal{S}_D} \Upsilon_1^d(\sigma, v_p)} = \frac{\sum_{\pi \in \mathcal{S}_{D-i}} \Upsilon_1^{k_\pi}(\pi, v_p) \Upsilon_{k_\pi+1}^{d-1}(\pi, v_p)}{\sum_{\sigma \in \mathcal{S}_D} \Upsilon_1^{k_\sigma}(\sigma, v_p) \Upsilon_{k_\sigma+1}^d(\sigma, v_p)}.$$

However, one can notice, from Eq. (B.1), that, for any $\pi \in \mathcal{S}_A, A \subseteq D$:

$$\Upsilon_1^{k_\pi}(\pi, v_p) = \prod_{j=1}^{k_\pi} v_p(C_j(\pi))^{-1} = \epsilon_p^{-k_\pi},$$

leading to:

$$\text{PV}_i = \frac{\sum_{\pi \in \mathcal{S}_{D-i}} \epsilon_p^{-k_\pi} \Upsilon_{k_\pi+1}^{d-1}(\pi, v_p)}{\sum_{\sigma \in \mathcal{S}_D} \epsilon_p^{-k_\sigma} \Upsilon_{k_\sigma+1}^d(\sigma, v_p)}.$$

Denote, for any $j \in D, k_{\max}^{-j}$ the size of the largest null coalition in D_{-j} , i.e.,

$$k_{\max}^{-j} = \max_{A \in \mathcal{P}(D_{-j})} \{|A| : v(A) = 0\},$$

with the convention that $|\emptyset| = 0$, and let $k_{\max} = k_{\max}^{-\emptyset}$, the size of the largest null coalition in D , and notice that necessarily,

$$\forall j \in D, \forall \pi \in \mathcal{S}_{D-j}, k_{\max}^{-j} \leq k_\pi. \quad (\text{B.2})$$

Moreover, denote, for any $j \in D \cup \{\emptyset\}$:

$$\mathcal{R}_{\max}^{-j} = \{\pi \in \mathcal{S}_{D-j} : k_\pi = k_{\max}^{-j}\}, \quad \text{and} \quad \mathcal{R}^{-j} = \{\pi \in \mathcal{S}_{D-j} : k_\pi < k_{\max}^{-j}\}$$

and notice that $\mathcal{R}_{\max}^{-j} \cup \mathcal{R}^{-j} = \mathcal{S}_{D-j}$ and moreover that,

$$\forall j \in D \cup \emptyset, \forall \pi \in \mathcal{R}^{-j}, \quad k_\pi < k_{\max}. \quad (\text{B.3})$$

Again, denote $\mathcal{R}_{\max} = \mathcal{R}_{\max}^{-\emptyset}$ and $\mathcal{R} = \mathcal{R}^\emptyset$.

Hence, for any $j \in D$, one has that:

$$\begin{aligned} \sum_{\pi \in \mathcal{S}_{D-j}} \epsilon_p^{-k_\pi} \Upsilon_{k_\pi+1}^{d-1}(\pi, v) &= \sum_{\pi \in \mathcal{R}_{\max}^{-j}} \epsilon_p^{-k_\pi} \Upsilon_{k_\pi+1}^{d-1}(\pi, v) + \sum_{\pi \in \mathcal{R}^{-j}} \epsilon_p^{-k_\pi} \Upsilon_{k_\pi+1}^{d-1}(\pi, v) \\ &= \sum_{\pi \in \mathcal{R}_{\max}^{-j}} \epsilon_p^{-k_{\max}^{-j}} \Upsilon_{k_\pi+1}^{d-1}(\pi, v) + \sum_{\pi \in \mathcal{R}^{-j}} \epsilon_p^{-k_\pi} \Upsilon_{k_\pi+1}^{d-1}(\pi, v) \end{aligned}$$

and in the particular case of $j = \emptyset$, one has that:

$$\begin{aligned} \sum_{\pi \in \mathcal{S}_D} \epsilon_p^{-k_\pi} \Upsilon_{k_\pi+1}^d(\pi, v) &= \sum_{\pi \in \mathcal{R}_{\max}} \epsilon_p^{-k_{\max}} \Upsilon_{k_\pi+1}^d(\pi, v) + \sum_{\pi \in \mathcal{R}} \epsilon_p^{-k_\pi} \Upsilon_{k_\pi+1}^d(\pi, v) \\ &= \epsilon_p^{-k_{\max}} \left(\sum_{\pi \in \mathcal{R}_{\max}} \Upsilon_{k_\pi+1}^d(\pi, v) + \sum_{\pi \in \mathcal{R}} \epsilon_p^{k_{\max}-k_\pi} \Upsilon_{k_\pi+1}^d(\pi, v) \right). \end{aligned}$$

It entails that:

$$\text{PV}_i = \frac{\sum_{\pi \in \mathcal{R}_{\max}^{-i}} \epsilon_p^{k_{\max}-k_\pi} \Upsilon_{k_\pi+1}^{d-1}(\pi, v) + \sum_{\pi \in \mathcal{R}^{-i}} \epsilon_p^{k_{\max}-k_\pi} \Upsilon_{k_\pi+1}^{d-1}(\pi, v)}{\sum_{\sigma \in \mathcal{R}_{\max}} \Upsilon_{k_\sigma+1}^d(\sigma, v) + \sum_{\sigma \in \mathcal{R}} \epsilon_p^{k_{\max}-k_\sigma} \Upsilon_{k_\sigma+1}^d(\sigma, v)}.$$

From Eq. (B.3), one can notice, for any $j \in D \cup \{\emptyset\}$:

$$\lim_{p \rightarrow \infty} \sum_{\pi \in \mathcal{R}^{-j}} \epsilon_p^{k_{\max}-k_\pi} \Upsilon_{k_\pi+1}^{d-|j|}(\pi, v) = 0$$

and additionally, from Eq. (B.2), notice that for any $j \in D$:

$$\lim_{p \rightarrow \infty} \sum_{\pi \in \mathcal{R}_{\max}^{-j}} \epsilon_p^{k_{\max}-k_\pi} \Upsilon_{k_\pi+1}^{d-1}(\pi, v) = \begin{cases} \sum_{\pi \in \mathcal{R}_{\max}^{-j}} \Upsilon_{k_\pi+1}^{d-1}(\pi, v) & \text{if } k_{\max} = k_{\max}^{-j} \\ 0 & \text{otherwise.} \end{cases}$$

Denote:

$$\text{PV}^0((D, v)) = \lim_{p \rightarrow \infty} \text{PV}(D, v_p),$$

and notice that, for any $i \in D$:

$$\text{PV}_i^0 = \begin{cases} \frac{\sum_{\pi \in \mathcal{R}_{\max}^{-i}} \Upsilon_{k_\pi+1}^{d-1}(\pi, v)}{\sum_{\sigma \in \mathcal{R}_{\max}} \Upsilon_{k_\sigma+1}^{d-1}(\sigma, v)} & \text{if } k_{\max} = k_{\max}^{-i}, \\ 0 & \text{otherwise.} \end{cases}$$

For any $j \in D$, the condition $k_{\max} = k_{\max}^{-j}$ is equivalent to the existence of a coalition $A \subseteq D-j$ such that $|A| = k_{\max}$ and $v(A) = 0$. On the other hand, the complement of this condition is that j must be in every coalition $A \subseteq D$ such that $|A| = k_{\max}$ and $v(A) = 0$, leading to the condition for which $\text{PV}_j^0 = 0$.

For any $j \in D \cup \{\emptyset\}$, and assuming that $k_{\max}^{-j} = k_{\max}$, one can notice that \mathcal{R}_{\max}^{-j} only contains the permutations $\pi \in \mathcal{S}_{D-j}$ such that $v(C_{k_{\max}}(\pi)) = 0$, and by monotonicity, this implies that for any $\pi \in \mathcal{R}_{\max}^{-j}$:

$$v(C_1(\pi)) = v(C_2(\pi)) = \dots = v(C_{k_{\max}-1}(\pi)) = v(C_{k_{\max}}(\pi)) = 0,$$

and that for $k_{\max} < k \leq |\pi|$,

$$v(C_k(\pi)) > 0.$$

For any $j \in D \cup \{\emptyset\}$, denote $\mathcal{K}_{-j} = \{A \subseteq D_{-j} : v(A) = 0 \text{ and } |A| = k_{\max}\}$, and notice that \mathcal{R}_{\max}^{-j} is necessarily composed of permutations having permutations of elements in \mathcal{K}_{-j} as their first k_{\max} elements. In other words, for every $\pi \in \mathcal{R}_{\max}^{-j}$,

$$C_{k_{\max}}(\pi) \in \mathcal{K}_{-j}.$$

Thus, for any $j \in D \cup \emptyset$, one has that:

$$\begin{aligned} \sum_{\pi \in \mathcal{R}_{\max}^{-j}} \Upsilon_{k_{\pi}+1}^{d-1}(\pi, v) &= \sum_{A \in \mathcal{K}_{-j}} k_{\max}! \sum_{\pi \in \mathcal{S}_{D_{-j} \setminus A}} \Upsilon_1^{|\pi|}(\pi, v_A) \\ &= k_{\max}! \sum_{A \in \mathcal{K}_{-j}} \sum_{\pi \in \mathcal{S}_{D_{-j} \setminus A}} \prod_{k=1}^{|\pi|} v(A \cup C_k(\pi))^{-1} \\ &= k_{\max}! \sum_{A \in \mathcal{K}_{-j}} R(D_{-j} \setminus A, v_A)^{-1} \end{aligned}$$

where for any $B \subseteq D \setminus A$, $v_A(B) = v(A \cup B)$, and using results from [10] on the ratio potential. This leads to the following rewriting of PV^0 , for any $i \in D$:

$$\text{PV}_i^0 = \begin{cases} 0 & \text{if } \forall A \in \mathcal{K}, i \in A, \\ \frac{\sum_{A \in \mathcal{K}_{-i}} R(D_{-i} \setminus A, v_A)^{-1}}{\sum_{A \in \mathcal{K}} R(D \setminus A, v_A)^{-1}} & \text{otherwise.} \end{cases}$$

Finally, notice that for any positive game (D, v) , i.e., , where v is positively valued, then necessarily, for any permutation and sub-permutations π of players $k_{\pi} = k_{\max} = 0$. Moreover, for any $j \in D \cup \{\emptyset\}$, $\mathcal{K}_{-j} = \emptyset$, and for any $i \in D$,

$$\text{PV}_i^0 = \frac{R(D, v)}{R(D_{-i}, v)} = \text{PV}_i,$$

and hence the allocation $\text{PV}^0((D, v))$ is a continuous extension of $\text{PV}((D, v))$ to cooperative games with nonnegative value function. \square

Proof of Lemma 1. Let $A \subseteq D$. First, focus on the implication

$$S_A^T = 0 \implies G(X) = \mathbb{E}[G(X) \mid X_{\bar{A}}] \text{ a.s.}$$

If $S_A^T = 0$, then necessarily,

$$\mathbb{V}(G(X) \mid X_{\bar{A}}) := \mathbb{E} \left[(G(X) - \mathbb{E}[G(X) \mid X_{\bar{A}}])^2 \mid X_{\bar{A}} \right] = 0 \text{ a.s.}$$

which can only be attained, by non-negativity of the squared distance, if

$$G(X) = \mathbb{E}[G(X) \mid X_{\bar{A}}] \text{ a.s.}$$

which proves the implication.

Now assume that $G(X) = \mathbb{E}[G(X) \mid X_{\bar{A}}]$ a.s.. Then necessarily,

$$\mathbb{V}(G(X) \mid X_{\bar{A}}) = 0 \text{ a.s.}$$

and thus $S_A^T = 0$, which proves the converse implication, and the equivalence stated in Lemma 1. \square

Proof of Proposition 1. Assume that X_E is the largest subset of X of L^2 -exogenous inputs to G . This entails, by Definition 2, $\forall B \subseteq E, \exists f \in L^2(P_{X_{\bar{B}}})$ such that:

$$G(X) = f(X_{\bar{B}}) \text{ a.s.}$$

Recall that for any $B \subseteq E$, the conditional expectation of $G(X)$ w.r.t. $X_{\bar{B}}$ is the unique projection defined as:

$$\mathbb{E}[G(X) | X_{\bar{B}}] = \underset{h \in L^2(P_{X_{\bar{B}}})}{\operatorname{argmin}} \mathbb{E} \left[(G(X) - h(X_{\bar{B}}))^2 \right],$$

One can notice that, since $f \in L^2(P_{X_{\bar{B}}})$ and $Y = f(X_{\bar{B}})$ a.s., then it necessarily minimizes the projection of $G(X)$ onto $L^2(P_{\bar{B}})$, leading to

$$G(X) = f(X_{\bar{B}}) = \mathbb{E}[G(X) | X_{\bar{B}}] \text{ a.s.}$$

and by Lemma 1, it entails that, for every $B \subseteq E$, $S_B^T = 0$. Furthermore, X_E being the largest subset of X being L^2 -exogenous, it entails that for any $A \subseteq D$ such that $|A| \geq |E|$ and $A \neq E$, $\nexists f \in L^2(P_{X_{\bar{A}}})$ s.t. $G(X) = f(X_{\bar{A}})$. Then, necessarily for any coalition $A \neq E$ of size larger or equal than E ,

$$\mathbb{E}[G(X) | X_{\bar{A}}] \neq G(X) \text{ a.s.} \iff S_A^T > 0.$$

Hence, E is the largest set of inputs with $S_E^T = 0$. Then, necessarily, $\mathcal{K} = \{E\}$ where \mathcal{K} is defined as in Theorem 1. Then $\forall i \in E, \forall A \in \mathcal{K}, i \in A$. Hence, it entails that $\text{PME}_i = 0$ for any $i \in E$, and, furthermore, $\forall i \in \bar{E}$, i is not contained in $E \in \mathcal{K}$. Finally, by Theorem 1, $\text{PME}_i > 0$ for any $i \in E$. \square

We are grateful to Nicolas Bousquet (EDF R&D), Fabrice Gamboa (IMT) and Christophe Labreuche (Thales) for their helpful comments.

References

- [1] J. An and A.B. Owen. Quasi-regression. *Journal of Complexity*, 17(588–607), 2001.
- [2] B. Broto. *Sensitivity analysis with dependent random variables : Estimation of the Shapley effects for unknown input distribution and linear Gaussian models*. PhD thesis, Université Paris-Saclay, September 2020.
- [3] B. Broto, F. Bachoc, and M. Depecker. Variance reduction for estimation of Shapley effects and adaptation to unknown input distribution. *SIAM/ASA Journal on Uncertainty Quantification*, 8:693–716, 2020.
- [4] D.V. Budescu. Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. *Psychological Bulletin*, 114:542–551, 1993.
- [5] S. Da Veiga. Kernel-based ANOVA decomposition and Shapley effects—Application to global sensitivity analysis. *Preprint*, 2021, arXiv:2101.05487.
- [6] S. Da Veiga, F. Gamboa, B. Iooss, and C. Prieur. *Basics and Trends in Sensitivity Analysis. Theory and Practice in R*. SIAM, 2021.
- [7] B. Efron and C. Stein. The jackknife estimate of variance. *The Annals of Statistics*, 9:586–596, 1981.
- [8] B. E. Feldman. A dual model of cooperative value. *Available at SSRN 317284*, 2002.
- [9] B. E. Feldman. Relative Importance and Value. *SSRN Electronic Journal*, 2005.
- [10] B. E. Feldman. A Theory of Attribution. *SSRN Electronic Journal*, 2007.

- [11] B. E. Feldman et al. The proportional value of a cooperative game. *Manuscript. Chicago: Scudder Kemper Investments*, 1999.
- [12] Y. Funaki. Dual axiomatizations of solutions of cooperative games, January 1996. preprint.
- [13] R. Ghanem, D. Higdon, and H. Owhadi, editors. *Springer Handbook on Uncertainty Quantification*. Springer, 2017.
- [14] U. Grömping. Estimators of relative importance in linear regression based on variance decomposition. *The American Statistician*, 61(2), 2007.
- [15] U. Grömping. Variable importance in regression models. *WIREs Comput Stat*, 7(137-152), 2015.
- [16] J. C. Harsanyi. A Simplified Bargaining Model for the n-Person Cooperative Game. *International Economic Review*, 4(2):194–220, 1963.
- [17] J. Hart and P. A. Gremaud. An approximation theoretic perspective of Sobol’ indices with dependent variables. *International Journal for Uncertainty Quantification*, 8(6), 2018.
- [18] M. Herin, M. Il Idrissi, V. Chabridon, and B. Iooss. Proportional marginal effects for sensitivity analysis with correlated inputs. In *Proceedings of the 10th International Conference on Sensitivity Analysis of Model Output (SAMO 2022)*, Tallahassee, Florida, USA, March 2022.
- [19] M. Il Idrissi, V. Chabridon, and B. Iooss. Developments and applications of Shapley effects to reliability-oriented sensitivity analysis with correlated inputs. *Environmental Modelling & Software*, 143:105115, 2021.
- [20] M. Il Idrissi, B. Iooss, and V. Chabridon. Mesures d’importance relative par décomposition de la performance de modèles de régression. In *Actes des 52èmes Journées de Statistique de la Société Française de Statistique (SFdS), Juin 2021, Nice, France*, pages 497–502, 2021. https://jds2021.sciencesconf.org/data/pages/book_jds2021_fr_compressed.pdf.
- [21] B. Iooss, V. Chabridon, and V. Thouvenot. Variance-based importance measures for machine learning model interpretability. In *Actes du 23ème Congrès de Maîtrise des Risques et de Sécurité de Fonctionnement ($\lambda\mu 23$)*, Saclay, France, october 2022. <https://hal.archives-ouvertes.fr/hal-03741384>.
- [22] B. Iooss and C. Prieur. Shapley effects for sensitivity analysis with dependent inputs: comparisons with Sobol’ indices, numerical estimation and applications. *International Journal for Uncertainty Quantification*, 9:493–514,, 2019.
- [23] B. Iooss and A. Saltelli. Introduction: Sensitivity analysis. In R. Ghanem, D. Higdon, and H. Owhadi, editors, *Springer Handbook on Uncertainty Quantification*, pages 1103–1122. Springer, 2017.
- [24] J.W. Johnson and J.M. LeBreton. History and use of relative importance indices in organizational research. *Organizational Research Methods*, 7:238–257, 2004.
- [25] E. Kalai and D. Samet. On weighted Shapley values. *International Journal of Game Theory*, 16(3):205–222, September 1987.
- [26] S. Kucherenko, O.V. Klymenko, and N. Shah. Sobol’ indices for problems defined in non-rectangular domains. *Reliability Engineering & System Safety*, 167:218–231, 2017.
- [27] D. Kurowicka and R. Cooke. *Uncertainty analysis with high dimensional dependence modelling*. Wiley, 2006.

- [28] R. H. Lindeman, P. F. Merenda, and R. Z. Gold. *Introduction to bivariate and multivariate analysis*. Scott Foresman and Company, Glenview, IL, 1980.
- [29] A. Marrel and V. Chabridon. Statistical developments for target and conditional sensitivity analysis: Application on safety studies for nuclear reactor. *Reliability Engineering & System Safety*, 214:107711, 2021.
- [30] K. M. Ortman. The proportional value for positive cooperative games. *Mathematical Methods of Operations Research (ZOR)*, 51(2):235–248, 2000.
- [31] A.B. Owen. Sobol’ indices and Shapley value. *SIAM/ASA Journal on Uncertainty Quantification*, 2:245–251, 2014.
- [32] A.B. Owen and C. Prieur. On Shapley value for measuring importance of dependent inputs. *SIAM/ASA Journal on Uncertainty Quantification*, 5:986–1002, 2017.
- [33] E. Plischke, G. Rabitti, and E. Borgonovo. Computing Shapley effects for sensitivity analysis. *SIAM/ASA Journal on Uncertainty Quantification*, 9:1411–1437, 2021.
- [34] S. Razavi, A. Jakeman, A. Saltelli, C. Prieur, B. Iooss, E. Borgonovo, E. Plischke, S. Lo Piano, T. Iwanaga, W. Becker, S. Tarantola, J.H.A. Guillaume, J. Jakeman, H. Gupta, N. Melillo, G. Rabitti, V. Chabridon, Q. Duan, X. Sun, S. Smith, R. Sheikholeslami, N. Hosseini, M. Asadzadeh, A. Puy, S. Kucherenko, and H.R. Maier. The Future of Sensitivity Analysis: An essential discipline for systems modeling and policy support. *Environmental Modelling and Software*, 137:104954, 2021.
- [35] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Salsana, and S. Tarantola. *Global sensitivity analysis - The primer*. Wiley, 2008.
- [36] L. S. Shapley. Notes on the n-Person Game – II: The Value of an n-Person Game. Research Memorandum ATI 210720, RAND Corporation, Santa Monica, California, August 1951.
- [37] L. S. Shapley. A value for n-person games. In H. Kuhn and A. W. Tucker, editors, *Contributions to the Theory of Games, Volume II*, Annals of Mathematics Studies, chapter 17, pages 307–317. Princeton University Press, Princeton, NJ, 1953.
- [38] I. M. Sobol. On sensitivity estimation for nonlinear mathematical models. *Mathematical Modelling and Computational Experiments*, 2(1):112–118, 1990.
- [39] E. Song, B.L. Nelson, and J. Staum. Shapley effects for global sensitivity analysis: Theory and computation. *SIAM/ASA Journal on Uncertainty Quantification*, 4:1060–1083, 2016.
- [40] R. J. Weber. Probabilistic values for games. In A. E. Roth, editor, *The Shapley value: essays in honor of Lloyd S. Shapley*, chapter 7, pages 101–120. Cambridge University Press, New York, NY, 1988.