



SINCLAIR



COALITIONAL DECOMPOSITIONS OF QUANTITIES OF INTEREST

GENERALIZED MÖBIUS INVERSION AND THE INPUT/MODEL-CENTRIC PARADIGMS

¹EDF R&D - Lab Chatou - PRISME Department

²Institut de Mathématiques de Toulouse

³SINCLAIR AI Lab

2023 Annual GdR MASCOT-NUM Meeting

Le Domaine Port aux Rocs, Le Croisic, France

April 3-6, 2023

One of the PhD. goal: Theoretical characterization of what GSA (and XAI) methods **quantify** when the **inputs are not independent**.

One of the PhD. goal: Theoretical characterization of what GSA (and XAI) methods **quantify** when the **inputs are not independent**.

Traditionally in global sensitivity analysis (GSA), **QoI decompositions** have been shown using a **model-centric approach**.

One of the PhD. goal: Theoretical characterization of what GSA (and XAI) methods **quantify** when the **inputs are not independent**.

Traditionally in global sensitivity analysis (GSA), **QoI decompositions** have been shown using a **model-centric approach**.

Recently, new indices inspired by **cooperative game theory** allowed to circumvent the problem of **dependence between inputs** (Owen 2014; Herin et al. 2022).

One of the PhD. goal: Theoretical characterization of what GSA (and XAI) methods **quantify** when the **inputs are not independent**.

Traditionally in global sensitivity analysis (GSA), **QoI decompositions** have been shown using a **model-centric approach**.

Recently, new indices inspired by **cooperative game theory** allowed to circumvent the problem of **dependence between inputs** (Owen 2014; Herin et al. 2022).

It also allows apprehending QoI decompositions in a **new light**: the **input-centric approach**.

One of the PhD. goal: Theoretical characterization of what GSA (and XAI) methods **quantify** when the **inputs are not independent**.

Traditionally in global sensitivity analysis (GSA), **QoI decompositions** have been shown using a **model-centric approach**.

Recently, new indices inspired by **cooperative game theory** allowed to circumvent the problem of **dependence between inputs** (Owen 2014; Herin et al. 2022).

It also allows apprehending QoI decompositions in a **new light**: the **input-centric approach**.

Both of these approaches are linked by a combinatorial mechanism: the **Möbius inversion formula**.

One of the PhD. goal: Theoretical characterization of what GSA (and XAI) methods **quantify** when the **inputs are not independent**.

Traditionally in global sensitivity analysis (GSA), **QoI decompositions** have been shown using a **model-centric approach**.

Recently, new indices inspired by **cooperative game theory** allowed to circumvent the problem of **dependence between inputs** (Owen 2014; Herin et al. 2022).

It also allows apprehending QoI decompositions in a **new light**: the **input-centric approach**.

Both of these approaches are linked by a combinatorial mechanism: the **Möbius inversion formula**.

Paving the way towards solutions for general QoI decomposition with dependent inputs.

What is a coalitional decomposition ?

Definition (Coalitional decomposition of a quantity of interest).

Let $X = (X_1, \dots, X_d)^\top$ be random inputs, and let $G(X)$ be a random output.

Let $QoI(G(X))$ be a quantity of interest (QoI) on the output.

Let $D = \{1, \dots, d\}$, and let $\mathcal{P}(D)$ denote the set of subsets of D (power-set).

If $QoI(G(X))$ can be written as:

$$QoI(G(X)) = \sum_{A \in \mathcal{P}(D)} \psi(A)$$

then the right-hand side of the equality is called a coalitional decomposition of $QoI(G(X))$.

Two ways to define coalitional decompositions: an **input-centric** approach and a **model-centric** approach.

Let's take an example with the variance decomposition.

Model-centric : Sobol' indices

Let X_1, \dots, X_d be **mutually independent** inputs. Let $G(X)$ be real-valued random variable such that $V(G(X)) < 1$.

Model-centric : Sobol' indices

Let X_1, \dots, X_d be **mutually independent** inputs. Let $G(X)$ be real-valued random variable such that $V(G(X)) < 1$.

From Hoeffding (1948), one has that:

$$L^2(P_X) = \bigoplus_{A \in \mathcal{P}(D)} \overline{V}_A, \quad \text{and} \quad G(X) = \sum_{A \in \mathcal{P}(D)} G_A(X_A)$$

where the summands are **pairwise orthogonal**.

Model-centric : Sobol' indices

Let X_1, \dots, X_d be **mutually independent** inputs. Let $G(X)$ be real-valued random variable such that $V(G(X)) < 1$.

From Hoeffding (1948), one has that:

$$L^2(P_X) = \bigoplus_{A \in \mathcal{P}(D)} \overline{V}_A, \quad \text{and} \quad G(X) = \sum_{A \in \mathcal{P}(D)} G_A(X_A)$$

where the summands are **pairwise orthogonal**.

Moreover, for any $A \in \mathcal{P}(D)$ (Da Veiga et al. 2021):

$$V(E[G(X) | X_A]) = \sum_{B \in \mathcal{P}(A)} V(G_B(X_B))$$

which implies that (Sobol' 1990), $\forall A \in \mathcal{P}(D)$:

$$V(G_A(X_A)) = \sum_{B \in \mathcal{P}(A)} (1)^{|A|-|B|} V(E[G(X) | X_B]) = V(G(X)) S_A,$$

and in particular ($A = D$),

$$V(G(X)) = \sum_{A \in \mathcal{P}(D)} V(G_A(X_A))$$

Input-centric : Cooperative games for variance-based GSA

Now suppose that X_1, \dots, X_d are **not mutually independent**, and $V(G(X)) < 1$.

By analogy between Cooperative Game Theory and GSA, Owen (2014) proposed to view **dependent inputs** as **players**, whose value is chosen to be:

$$v(A) = V(E[G(X) | X_A]).$$

Input-centric : Cooperative games for variance-based GSA

Now suppose that X_1, \dots, X_d are **not mutually independent**, and $V(G(X)) < 1$.

By analogy between Cooperative Game Theory and GSA, Owen (2014) proposed to view **dependent inputs** as **players**, whose value is chosen to be:

$$v(A) = V(E[G(X) | X_A]).$$

The Harsanyi (1963) dividends of this game are, $\forall A \subseteq P(D)$:

$$\psi(A) = \sum_{B \in P(A)} (-1)^{|A|-|B|} V(E[G(X) | X_B]),$$

and it implies that (Bilbao 2000), $\forall A \subseteq P(D)$:

$$V(E[G(X) | X_A]) = \sum_{B \in P(A)} \psi(B),$$

Input-centric : Cooperative games for variance-based GSA

Now suppose that X_1, \dots, X_d are **not mutually independent**, and $V(G(X)) < 1$.

By analogy between Cooperative Game Theory and GSA, Owen (2014) proposed to view **dependent inputs** as **players**, whose value is chosen to be:

$$v(A) = V(E[G(X) | X_A]).$$

The Harsanyi (1963) dividends of this game are, $\delta_A \in P(D)$:

$$\psi(A) = \sum_{B \in P(A)} (-1)^{|A|-|B|} V(E[G(X) | X_B]),$$

and it implies that (Bilbao 2000), $\delta_A \in P(D)$:

$$V(E[G(X) | X_A]) = \sum_{B \in P(A)} \psi(B),$$

and in particular ($A = D$),

$$V(G(X)) = \sum_{A \in P(D)} \psi(A)$$

Input-centric and Model-centric approaches for variance decomposition

(Traditional) Model-centric approach

X_1, \dots, X_d mutually independent.

Decompose $G(X)$ into orthogonal $G_A(X_A)$.

One then has $\exists A \subseteq P(D)$:

$$V(E[G(X) | X_A]) = \sum_{A \in P(D)} V(G_A(X_A)).$$

which implies that:

$$V(G_A(X_A)) = \sum_{B \in P(A)} (-1)^{|A|-|B|} V(E[G(X) | X_B])$$

Input-centric approach

X_1, \dots, X_d not necessarily mutually independent.

Choose to value X_A by $V(E[G(X) | X_A])$.

Set, $\exists A \subseteq P(D)$:

$$\psi(A) = \sum_{B \in P(A)} (-1)^{|A|-|B|} V(E[G(X) | X_B]),$$

which implies that:

$$V(G(X)) = \sum_{A \in P(D)} \psi(A)$$

Input-centric and Model-centric approaches for variance decomposition

(Traditional) Model-centric approach

X_1, \dots, X_d mutually independent.

Decompose $G(X)$ into orthogonal $G_A(X_A)$.

One then has $\delta_A \perp P(D)$:

$$V(E[G(X) | X_A]) = \sum_{A \in \mathcal{P}(D)} V(G_A(X_A)).$$

which implies that:

$$V(G_A(X_A)) = \sum_{B \in \mathcal{P}(A)} (-1)^{|A|-|B|} V(E[G(X) | X_B])$$

If the inputs are **actually mutually independent**:

- $\psi(A) = V(G_A(X_A))$ and both approaches are equivalent.
- The **input-centric** approach **did not require the** $G_A(X_A)$ **to be pairwise orthogonal.**

Input-centric approach

X_1, \dots, X_d not necessarily mutually independent.

Chose to value X_A by $V(E[G(X) | X_A])$.

Set, $\delta_A \perp P(D)$:

$$\psi(A) = \sum_{B \in \mathcal{P}(A)} (-1)^{|A|-|B|} V(E[G(X) | X_B]),$$

which implies that:

$$V(G(X)) = \sum_{A \in \mathcal{P}(D)} \psi(A)$$

How are both approaches linked ?

How are both approaches linked ?

It is due to a **generalization of the Möbius inversion formula** to **locally finite partially ordered sets** (Rota 1964).

One of the **cornerstones** of the field of **combinatorics** (Kung, Rota, and Hung Yan 2012).

The (very general) result of Rota admits a particular form when **dealing with power-sets**.

It can be understood as a generalization of the **inclusion-exclusion principle**.

Möbius inversion on power-sets

Corollary (Möbius inversion on power-sets (Rota 1964; Kung, Rota, and Hung Yan 2012)).

Let $D = \{d_1, \dots, d_n\}$, and any two set functions:

$$f : \mathcal{P}(D) \rightarrow A, \quad g : \mathcal{P}(D) \rightarrow A,$$

where A is an **abelian group**. Then the following equivalence holds:

$$f(A) = \sum_{B \in \mathcal{P}(A)} g(B), \quad \forall A \in \mathcal{P}(D) \quad (\Leftrightarrow) \quad g(A) = \sum_{B \in \mathcal{P}(A)} (-1)^{|A|-|B|} f(B), \quad \forall A \in \mathcal{P}(D).$$

Möbius inversion on power-sets

Corollary (Möbius inversion on power-sets (Rota 1964; Kung, Rota, and Hung Yan 2012)).

Let $D = \{d_1, \dots, d_n\}$, and any two set functions:

$$f : \mathcal{P}(D) \rightarrow A, \quad g : \mathcal{P}(D) \rightarrow A,$$

where A is an **abelian group**. Then the following equivalence holds:

$$f(A) = \sum_{B \in \mathcal{P}(A)} g(B), \quad \forall A \in \mathcal{P}(D) \quad (\Leftrightarrow) \quad g(A) = \sum_{B \in \mathcal{P}(A)} (-1)^{|A|-|B|} f(B), \quad \forall A \in \mathcal{P}(D).$$

Three remarks:

- **Left to right:** traditional **model-centric approach**.
- **Right to left:** **Input-centric approach**.
- The set functions f and g can be valued in an **abelian group**, and not necessarily \mathbb{R} :
we can generalize input-centric decompositions to a broad range of QoIs.

Examples of abelian groups: \mathbb{R} , \mathbb{R}^d , spaces of matrices, polynomials, vector spaces...

Möbius inversion mechanism

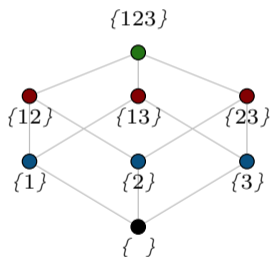
The Möbius inversion is a **mechanical process**.

Variance decomposition: Let's compute $\sum_{A \in \mathcal{P}(D)} \sum_{B \in \mathcal{P}(A)} (-1)^{|A|-|B|} V(E[G(X) | X_B])$ with $d = 3$.

Möbius inversion mechanism

The Möbius inversion is a **mechanical process**.

Variance decomposition: Let's compute $\sum_{A \in \mathcal{P}(D)} \sum_{B \in \mathcal{P}(A)} (-1)^{|A|-|B|} V(E[G(X) | X_B])$ with $d = 3$.



Powerset of $\{1, 2, 3\}$

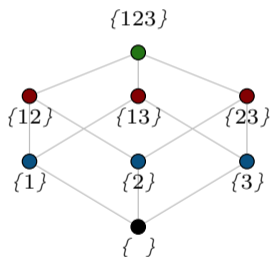
Let $V_A = V(E[G(X) | X_A])$.

A							
123	$V(G(X))$	$-V_{12}$	$-V_{23}$	$-V_{13}$	$+V_1$	$+V_2$	$+V_3$
12		$+V_{12}$			$-V_1$	$-V_2$	
23			$+V_{23}$			$-V_2$	$-V_3$
13				$+V_{13}$	$-V_1$		$-V_3$
1					$+V_1$		
2						$+V_2$	
3							$+V_3$
Sum	$V(G(X))$	$+0$	$+0$	$+0$	$+0$	$+0$	$+0$

Möbius inversion mechanism

The Möbius inversion is a **mechanical process**.

Variance decomposition: Let's compute $\sum_{A \in \mathcal{P}(D)} \sum_{B \in \mathcal{P}(A)} (-1)^{|A|-|B|} V(E[G(X) | X_B])$ with $d = 3$.



Powerset of $\{1, 2, 3\}$

Let $V_A = V(E[G(X) | X_A])$.

A							
123	$V(G(X))$	$-V_{12}$	$-V_{23}$	$-V_{13}$	$+V_1$	$+V_2$	$+V_3$
12		$+V_{12}$			$-V_1$	$-V_2$	
23			$+V_{23}$			$-V_2$	$-V_3$
13				$+V_{13}$	$-V_1$		$-V_3$
1					$+V_1$		
2						$+V_2$	
3							$+V_3$
Sum	$V(G(X))$	$+0$	$+0$	$+0$	$+0$	$+0$	$+0$

This is just a fancy way to write $QoI(G(X)) = QoI(G(X)) + 0$.

Meaningfulness

Many coalitional decompositions built using Möbius inversion formula **are not meaningful**.

For instance:

$$v(A) = \begin{cases} V(G(X)) & \text{if } A = D, \\ c_A & \text{for any } c_A \geq A \text{ otherwise.} \end{cases}$$

leads to a Möbius variance decomposition.

Meaningfulness

Many coalitional decompositions built using Möbius inversion formula **are not meaningful**.

For instance:

$$v(A) = \begin{cases} V(G(X)) & \text{if } A = D, \\ c_A & \text{for any } c_A \geq A \text{ otherwise.} \end{cases} \quad \text{leads to a Möbius variance decomposition.}$$

Thus, we need to define **desirability properties** on **input-centric** coalitional decompositions.

Meaningfulness

Many coalitional decompositions built using Möbius inversion formula **are not meaningful**.

For instance:

$$v(A) = \begin{cases} V(G(X)) & \text{if } A = D, \\ c_A & \text{for any } c_A \geq 0 \text{ otherwise.} \end{cases} \quad \text{leads to a Möbius variance decomposition.}$$

Thus, we need to define **desirability properties** on **input-centric** coalitional decompositions.

Definition (*Gradual coalitional decomposition* (I. et al. 2023)).

Let $X = (X_1, \dots, X_d)^\top$ be random inputs, and let $QoI(G(X))$ be an A -valued QoI on G .

For any $A \in \mathcal{P}(D)$, let $f_A(X_A)$ be a $\sigma(X_A)$ -measurable **representation** of $G(X)$. If the coalitional decomposition can be written as:

$$QoI(G(X)) = \sum_{A \in \mathcal{P}(D)} \sum_{B \in \mathcal{P}(A)} (-1)^{|A|-|B|} QoI(f_B(X_B))$$

it is said to be **gradual**.

Recipe

To build an **input-centric** gradual coalitional QoI decomposition:

- Choose **candidates** $f_A(X_A)$ to **represent** $G(X)$ as a function of X_A .
(In the previous examples, $f_A(X_A) = E [G(X) \mid X_A]$)

Recipe

To build an **input-centric** gradual coalitional QoI decomposition:

- Choose **candidates** $f_A(X_A)$ to **represent** $G(X)$ as a function of X_A .
(In the previous examples, $f_A(X_A) = E[G(X) \mid X_A]$)
- Make sure that $f_D(X) = G(X)$.

Recipe

To build an **input-centric** gradual coalitional QoI decomposition:

- Choose **candidates** $f_A(X_A)$ to **represent** $G(X)$ as a function of X_A .
(In the previous examples, $f_A(X_A) = E[G(X) | X_A]$)
- Make sure that $f_D(X) = G(X)$.
- Compute the same A-valued QoI on each of the $f_A(X_A)$, **provided they exist**.
(e.g., a (super)-quantile, a failure probability, a covariance matrix)

Recipe

To build an **input-centric** gradual coalitional QoI decomposition:

- Choose **candidates** $f_A(X_A)$ to **represent** $G(X)$ as a function of X_A .
(In the previous examples, $f_A(X_A) = E[G(X) | X_A]$)
- Make sure that $f_D(X) = G(X)$.
- Compute the same A -valued QoI on each of the $f_A(X_A)$, **provided they exist**.
(e.g., a (super)-quantile, a failure probability, a covariance matrix)
- Compute, $\forall A \subseteq P(D)$, the quantities:

$$\psi(A) = \sum_{B \in \mathcal{P}(A)} (-1)^{|A|-|B|} \text{QoI}(f_B(X_B)).$$

Recipe

To build an **input-centric** gradual coalitional QoI decomposition:

- Choose **candidates** $f_A(X_A)$ to **represent** $G(X)$ as a function of X_A .
(In the previous examples, $f_A(X_A) = E[G(X) | X_A]$)
- Make sure that $f_D(X) = G(X)$.
- Compute the same A -valued QoI on each of the $f_A(X_A)$, **provided they exist**.
(e.g., a (super)-quantile, a failure probability, a covariance matrix)
- Compute, $\forall A \subseteq P(D)$, the quantities:

$$\psi(A) = \sum_{B \in \mathcal{P}(A)} (-1)^{|A|-|B|} \text{QoI}(f_B(X_B)).$$

- Then, by the Möbius inversion:

$$\text{QoI}(G(X)) = \sum_{A \in \mathcal{P}(D)} \psi(A).$$

Recipe

To build an **input-centric** gradual coalitional QoI decomposition:

- Choose **candidates** $f_A(X_A)$ to **represent** $G(X)$ as a function of X_A .
(In the previous examples, $f_A(X_A) = E[G(X) | X_A]$)
- Make sure that $f_D(X) = G(X)$.
- Compute the same A -valued QoI on each of the $f_A(X_A)$, **provided they exist**.
(e.g., a (super)-quantile, a failure probability, a covariance matrix)
- Compute, $\forall A \subseteq P(D)$, the quantities:

$$\psi(A) = \sum_{B \in \mathcal{P}(A)} (-1)^{|A|-|B|} \text{QoI}(f_B(X_B)).$$

- Then, by the Möbius inversion:

$$\text{QoI}(G(X)) = \sum_{A \in \mathcal{P}(D)} \psi(A).$$

But, how can these decompositions be interpreted ?

Interpretation

The interpretation of each:

$$\psi(A) = \sum_{B \in \mathcal{P}(A)} (-1)^{|A|-|B|} \text{QoI}(f_A(X_A))$$

is **subject to the** $f_A(X_A)$, which are **subject to the model** and the **distribution of the inputs**.

Interpretation

The interpretation of each:

$$\psi(A) = \sum_{B \in \mathcal{P}(A)} (1 - 1)^{|A|-|B|} \text{QoI}(f_A(X_A))$$

is **subject to the** $f_A(X_A)$, which are **subject to the model** and the **distribution of the inputs**.

For the **variance decomposition**:

- If the inputs are **mutually independent**, and we choose $f_A(X_A) = E[G(X) | X_A]$, we saw that **both approaches are equivalent**:

$$\forall A \subseteq P(D), \quad \psi(A) = V(G_A(X_A)) = V(G(X)) \quad S_A \quad (\text{the Sobol' indices})$$

and hence the $\psi(A)$ can be interpreted as **pure interaction effects**.

Interpretation

The interpretation of each:

$$\psi(A) = \sum_{B \in \mathcal{P}(A)} (1)^{|A|-|B|} \text{QoI}(f_A(X_A))$$

is **subject to the** $f_A(X_A)$, which are **subject to the model** and the **distribution of the inputs**.

For the **variance decomposition**:

- If the inputs are **mutually independent**, and we choose $f_A(X_A) = E[G(X) | X_A]$, we saw that **both approaches are equivalent**:

$$\forall A \subseteq P(D), \quad \psi(A) = V(G_A(X_A)) = V(G(X)) \cdot S_A \quad (\text{the Sobol' indices})$$

and hence the $\psi(A)$ can be interpreted as **pure interaction effects**.

- If the inputs are **not mutually independent**, $V(E[G(X) | X_A])$ and $\psi(A)$ can vary **according to the dependence structure**, and hence **cannot quantify pure interaction**.

Illustration: Linear model with interaction and gaussian inputs

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

Let, $\mathcal{A} = \{1, 2, 3\}$:

$$\psi(\mathcal{A}) = \frac{1}{V(G(X))} \sum_{B \in \mathcal{P}(\mathcal{A})} \binom{3}{|B|} \rho^{|B|} V(E[G(X) | X_B])$$

Independent case ($\rho = 0$)

(The $\psi(\mathcal{A})$ are equal to the Sobol' indices)

$$\begin{aligned} S_1 &= 0.5 & S_2 &= 0, & S_3 &= 0, \\ S_{12} &= 0, & S_{13} &= 0, & S_{23} &= 0.5, \\ S_{123} &= 0 \end{aligned}$$

Correlated case ($\rho \neq 0$)

$$\begin{aligned} \psi(1) &= 0.5 & \psi(2) &= 0, & \psi(3) &= \rho^2/2, \\ \psi(12) &= \rho^2/2, & \psi(13) &= \rho^2/2, & \psi(23) &= 0.5, \\ \psi(123) &= \rho^2/2 \end{aligned}$$

Illustration: Linear model with interaction and gaussian inputs

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

Let, $\mathcal{A} = \{1, 2, 3\}$:

$$\psi(\mathcal{A}) = \frac{1}{V(G(X))} \sum_{B \in \mathcal{P}(\mathcal{A})} \binom{|A|}{|B|} \rho^{|A|-|B|} V(E[G(X) | X_B])$$

Independent case ($\rho = 0$)

(The $\psi(\mathcal{A})$ are equal to the Sobol' indices)

$$\begin{aligned} S_1 &= 0.5 & S_2 &= 0, & S_3 &= 0, \\ S_{12} &= 0, & S_{13} &= 0, & S_{23} &= 0.5, \\ S_{123} &= 0 \end{aligned}$$

Correlated case ($\rho \neq 0$)

$$\begin{aligned} \psi(1) &= 0.5 & \psi(2) &= 0, & \psi(3) &= \rho^2/2, \\ \psi(12) &= \rho^2/2, & \psi(13) &= \rho^2/2, & \psi(23) &= 0.5, \\ \psi(123) &= \rho^2/2 \end{aligned}$$

In both cases $\sum_{A \in \mathcal{P}(D)} \psi(A) = 1$, but in the **correlated case**, we cannot precisely characterize what $\psi(A)$ quantifies.

Shapley effects with dependent inputs

Hence, the precise interpretation of

$$\psi(A) = \frac{1}{V(G(X))} \sum_{B \in \mathcal{P}(A)} (-1)^{|A|-|B|} V(E[G(X) | X_A])$$

is still an open question : it clearly is a mixture of interaction and dependence effects.

But which mixture ?

Shapley effects with dependent inputs

Hence, the precise interpretation of

$$\psi(A) = \frac{1}{\text{V}(G(X))} \sum_{B \in \mathcal{P}(A)} (-1)^{|A|-|B|} \text{V}(E[G(X) | X_A])$$

is still an open question : it clearly is a mixture of interaction and dependence effects.

But which mixture ?

The **Shapley effects** for an input $i \in D$ can be written as (Harsanyi 1963):

$$\text{Sh}_i = \sum_{A \in \mathcal{P}(D), i \in A} \frac{\psi(A)}{|A|}$$

which is an **egalitarian aggregation of a (not so clear) mixture of interaction and dependence effects**.

Shapley effects with dependent inputs

Hence, the precise interpretation of

$$\psi(A) = \frac{1}{V(G(X))} \sum_{B \in \mathcal{P}(A)} (-1)^{|A|-|B|} V(E[G(X) | X_A])$$

is still an open question : it clearly is a mixture of interaction and dependence effects.

But which mixture ?

The **Shapley effects** for an input $i \in D$ can be written as (Harsanyi 1963):

$$\text{Sh}_i = \sum_{A \in \mathcal{P}(D), i \in A} \frac{\psi(A)}{|A|}$$

which is an **egalitarian aggregation** of a (not so clear) mixture of interaction and dependence effects.

Choosing $v(A) = V(E[G(X) | X_A])$, leads to an uncharacterized quantification.

Coalitional decompositions of QoIs:

- We saw two approaches: **Input-centric** and **Model-centric**.
- Defining **input-centric gradual** QoI decomposition **reduces to the choice of a representant** $f_A(X_A)$.
- The **input-centric** approach **bypasses the need** for **input independence** and, in the case of L^2 , an **orthogonal functional decomposition**.
- The interpretation of these decompositions **vary w.r.t. the dependence structure and the choice of representant**.

Conclusions

Coalitional decompositions of Qols:

- We saw two approaches: **Input-centric** and **Model-centric**.
- Defining **input-centric gradual** Qol decomposition **reduces to the choice of a representant** $f_A(X_A)$.
- The **input-centric** approach **bypasses the need for input independence** and, in the case of L^2 , an **orthogonal functional decomposition**.
- The interpretation of these decompositions **vary w.r.t. the dependence structure and the choice of representant**.

Cooperative games based GSA indices:

- Allocations are aggregations of **input-centric** coalitional Qol decompositions, driven by the choice of value function $v(A)$.
- The Shapley effects (for dependent inputs) are an **egalitarian redistribution** of the gradual Qol decomposition with $v(A) = V(E[G(X) | X_A])$.
- **At this time, we cannot characterize exactly what they quantify.**

But...

But...

For **(not necessarily mutually independent)** inputs $X = (X_1, \dots, X_d)^\top$, is it possible to **find representants** $f_A(X_A)$ such that each term

$$\psi(A) = \sum_{B \in \mathcal{P}(A)} (-1)^{|A|-|B|} V(f_B(X_B))$$

of the **input-centric gradual variance decomposition**

$$V(G(X)) = \sum_{A \in \mathcal{P}(D)} \psi(A)$$

quantifies pure interaction ?

But...

For **(not necessarily mutually independent)** inputs $X = (X_1, \dots, X_d)^\top$, is it possible to find **representants** $f_A(X_A)$ such that each term

$$\psi(A) = \sum_{B \in \mathcal{P}(A)} (-1)^{|A|-|B|} V(f_A(X_A))$$

of the **input-centric** gradual variance decomposition

$$V(G(X)) = \sum_{A \in \mathcal{P}(D)} \psi(A)$$

quantifies pure interaction ?

Our intuition:

- **Model-centric** approach to find the representants $f_A(X_A)$.
- **Input-centric** approach to define a gradual variance decomposition using these representants.

In the case of models in L^2 , the **model-centric** approach would amount to show that:

$$L^2(P_X) = \bigoplus_{A \in \mathcal{P}(D)} \overline{V_A},$$

hold whenever the inputs are **not necessarily mutually independent**, where the $\overline{V_A}$ are **not necessarily pairwise orthogonal**.

In the case of models in L^2 , the **model-centric** approach would amount to show that:

$$L^2(P_X) = \bigoplus_{A \in \mathcal{P}(D)} \overline{V}_A,$$

hold whenever the inputs are **not necessarily mutually independent**, where the \overline{V}_A are **not necessarily pairwise orthogonal**.

If so, the **projections** of $G(X)$ onto each \overline{V}_A could allow to define **promising representants**.

For fixed marginals and a fixed model, there would be **one set of representants for a particular dependence structure**.

In the case of models in L^2 , the **model-centric** approach would amount to show that:

$$L^2(P_X) = \bigoplus_{A \in \mathcal{P}(D)} \overline{V}_A,$$

hold whenever the inputs are **not necessarily mutually independent**, where the \overline{V}_A are **not necessarily pairwise orthogonal**.

If so, the **projections** of $G(X)$ onto each \overline{V}_A could allow to define **promising representants**.

For fixed marginals and a fixed model, there would be **one set of representants for a particular dependence structure**.

What would be the properties of gradual variance decompositions with this choice of representants ?

In the case of models in L^2 , the **model-centric** approach would amount to show that:

$$L^2(P_X) = \bigoplus_{A \in \mathcal{P}(D)} \overline{V}_A,$$

hold whenever the inputs are **not necessarily mutually independent**, where the \overline{V}_A are **not necessarily pairwise orthogonal**.

If so, the **projections** of $G(X)$ onto each \overline{V}_A could allow to define **promising representants**.

For fixed marginals and a fixed model, there would be **one set of representants for a particular dependence structure**.

What would be the properties of gradual variance decompositions with this choice of representants ?

We don't know yet... But we're working on it :)

For a more in-depth (and more general) study of the **relationship** between **Möbius inversion** and **coalitional decompositions** of Qols, check-out our **pre-print** (HAL/arXiv):

On the coalitional decomposition of parameters of interest

Marouane El Idrissi^{a,b,c,e}, Nicolas Bousquet^{a,b,d}, Fabrice Gamboa^c, Bertrand Iooss^{a,b,c}, Jean-Michel Loubes^c

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

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

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

^d*Sorbonne Université, LPSM, 4 place Jussieu, Paris, France*

References i

- Bilbao, J. M. 2000. *Cooperative Games on Combinatorial Structures* [in en]. Edited by W. Leinfellner and G. Eberlein. Vol. 26. Theory and Decision Library. Boston, MA: Springer US. isbn: 978-1-4613-6976-9 978-1-4615-4393-0.
<https://doi.org/10.1007/978-1-4615-4393-0>. <http://link.springer.com/10.1007/978-1-4615-4393-0>.
- Da Veiga, S., F. Gamboa, B. Iloos, and C. Prieur. 2021. *Basics and Trends in Sensitivity Analysis: Theory and Practice in R* [in en]. Philadelphia, PA: Society for Industrial / Applied Mathematics, January. isbn: 978-1-61197-668-7 978-1-61197-669-4.
<https://doi.org/10.1137/1.9781611976694>. <https://epubs.siam.org/doi/book/10.1137/1.9781611976694>.
- Gamboa, F., A. Janon, T. Klein, and A. Lagnoux. 2013. "Sensitivity indices for multivariate outputs." *Comptes Rendus Mathematique* 351 (7): 307–310. issn: 1631-073X. <https://doi.org/10.1016/j.crma.2013.04.016>.
- Harsanyi, J. C. 1963. "A Simplified Bargaining Model for the n-Person Cooperative Game." Publisher: [Economics Department of the University of Pennsylvania, Wiley, Institute of Social and Economic Research, Osaka University], *International Economic Review* 4 (2): 194–220. issn: 0020-6598. <https://doi.org/10.2307/2525487>. <https://www.jstor.org/stable/2525487>.
- Herin, M., M. I., V. Chabridon, and B. Iloos. 2022. *Proportional marginal effects for global sensitivity analysis* [in en], October.
<https://hal.science/hal-03825935>.
- Hoeffding, W. 1948. "A Class of Statistics with Asymptotically Normal Distribution." Publisher: Institute of Mathematical Statistics, *The Annals of Mathematical Statistics* 19, no. 3 (September): 293–325. issn: 0003-4851, 2168-8990.
<https://doi.org/10.1214/aoms/1177730196>.
<https://projecteuclid.org/journals/annals-of-mathematical-statistics/volume-19/issue-3/A-Class-of-Statistics-with-Asymptotically-Normal-Distribution/10.1214/aoms/1177730196.full>.

- I., M., N. Bousquet, F. Gamboa, B. Iooss, and J. Loubes. 2023. *On the coalitional decomposition of parameters of interest* [in en], January. Accessed April 2, 2023. <https://hal.science/hal-03927476>.
- Kung, J. P. S., G-C. Rota, and C. Hung Yan. 2012. *Combinatorics: the Rota way*. OCLC: 1226672593. New York: Cambridge University Press. ISBN: 978-0-511-80389-5.
- Owen, Art B. 2014. "Sobol' Indices and Shapley Value" [in en]. *SIAM/ASA Journal on Uncertainty Quantification* 2, no. 1 (January): 245–251. ISSN: 2166-2525. <https://doi.org/10.1137/130936233>. <http://epubs.siam.org/doi/10.1137/130936233>.
- Rota, G-C. 1964. "On the foundations of combinatorial theory I. Theory of Möbius Functions." *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 2 (4): 340–368. ISSN: 1432-2064. <https://doi.org/10.1007/BF00531932>.
- Sobol', I M. 1990. "On sensitivity estimation for nonlinear mathematical models" [in Russian]. *Mathematical Modelling and Computational Experiments* 2 (1): 112–118.

THANK YOU FOR YOUR ATTENTION!

ANY QUESTIONS?

Cooperative game theory

In a nutshell, cooperative game theory can be summarized as “**the art of cutting a cake**”.



Given a **set of players** $D = \{d_1, \dots, d_d\}$, who produces a **quantity** $v(D)$, how can one allocate shares of $v(D)$ among the d players?

The “**cake cutting process**” is often described through **axioms** (i.e., desired properties), and results in an **allocation**.

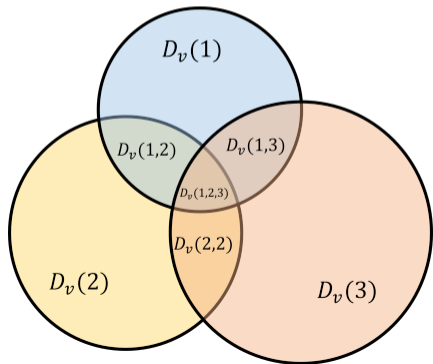
Formally, a cooperative game is denoted (D, v) where D is a **set of players**, and $v : P(D) \rightarrow \mathbb{R}$ is a **value function**, mapping every possible subset of players to a real value.

Interpreting the Shapley values: Harsanyi dividends

Another equivalent enlightening representation of the Shapley values can be done using **Harsanyi dividends** (Harsanyi 1963).

Let (D, v) be a cooperative game, and for any $A \subseteq D$, let the **Harsanyi dividend** of the coalition A be:

$$D_v(A) = \sum_{B \subseteq A} (-1)^{|A|-|B|} v(B).$$



The Harsanyi dividends can be interpreted as the **surplus (or shortfall)** that a coalition generates:

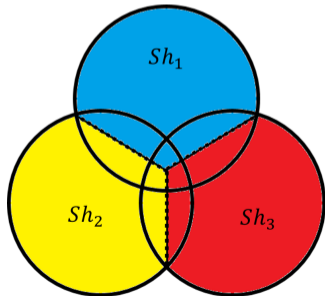
$$\begin{aligned} D_v(1) &= v(1), & D_v(2) &= v(2), \\ D_v(1,2) &= v(1,2) - v(1) - v(2). \end{aligned}$$

Interpreting the Shapley values: Harsanyi dividends

The Shapley values are then defined as:

$$Sh_i = \sum_{A \subseteq D: i \in A} \frac{D_v(A)}{|A|},$$

or, in other words, each dividend of a coalition is **equally** redistributed between the players that composes it.



Quick example: Eve and John are two developers, Eve produces 10.000 lines of code, John produces 8.000 lines of code.

However, John really likes to play babyfoot, but Eve is a hard-worker.

When working together, they only produce 10.000 lines of code. This means that the dividend of their coalition is 8.000.

Is it fair to attribute Eve 4.000 lines of code, even if she did all the work ?

Example - Covariance Matrix decomposition

Suppose that $G(X) = (G_1(X), \dots, G_k(X))^T$ is valued in \mathbb{R}^k , and that $G(X) \in L^2(P_X, \mathbb{R}^k)$ (Gamboa et al. 2013).

The QoI is the covariance matrix of the outputs $V(G(X)) \in \mathbb{R}^{k \times k}$.

Let $\Sigma(A) = V(E[G(X) | X_A]) \in \mathbb{R}^{k \times k}$ be defined element-wise as:

$$\Sigma_{i,j}(A) = \text{Cov}(E[G_i(X) | X_A], E[G_j(X) | X_A]).$$

Let, $\delta_A \in P(D)$:

$$\psi(A) = \sum_{B \in \mathcal{P}(A)} (-1)^{|A|-|B|} \Sigma(B) \in \mathbb{R}^{k \times k}.$$

Then, using the Möbius inversion on power-sets, one has the following coalitional decomposition of the output covariance matrix:

$$V(G(X)) = \sum_{A \in \mathcal{P}(D)} \psi(A).$$

Posets, incidence algebra and Möbius inverse

A *partially ordered set* (poset) is defined as a pair (S, \leq) where S is a non-empty set, and \leq is a partial order binary relation on elements of S . A poset (S, \leq) is said to be *locally finite* if, for any $x, z \in S$, the sets $f_{y \in S : x \leq y \leq z}$ (also called *segments* of S) are finite.

Denote $I_A(S)$ the incidence algebra of a locally finite poset (S, \leq) over a commutative ring with identity A , i.e., the set of functions $f : S \times S \rightarrow A$ such that $f(x, y) = 0$ if $x \not\leq y$. $(I_A(S), +, \cdot)$ forms an A -algebra with the usual pointwise addition $+$ and the usual convolution \cdot , i.e., for any $f, g \in I_A(S)$, and any $x, z \in S$ such that the segment $f_{y \in S : x \leq y \leq z}$ is non-empty,

$$(f \cdot g)(x, z) = \sum_{x \leq y \leq z} f(x, y)g(y, z).$$

The zeta function $\zeta \in I_A(S)$ is the convolutional identity of the incidence algebra, and is defined as, $\forall x, y \in S$:

$$\zeta(x, y) = \begin{cases} 1 & \text{if } x = y, \\ 0 & \text{otherwise.} \end{cases}$$

Posets, incidence algebra and Möbius inverse

The Möbius function, denoted $\mu \in I_A(S)$, in the case of locally finite posets S , is defined as the *inverse of the zeta function for the convolution operator* defined on the incidence algebra of S , and can be computed recursively, for any $x, y \in S$ with $x \leq y$, as

$$\mu(x, y) = \begin{cases} 1 & \text{if } x = y \\ -\sum_{x \leq z < y} \mu(x, z) & \text{otherwise.} \end{cases}$$

Theorem (*Möbius inversion formula on locally finite posets*). Let S be any non-empty set and (S, \leq) form a locally finite poset, where \leq is a binary relation. Let φ and ψ be functions from S to A . Then, the following equivalence hold:

$$\varphi(x) = \sum_{y: y \leq x} \psi(y), \quad \forall x \in S \quad (\Leftrightarrow) \quad \psi(x) = \sum_{y: y \leq x} \varphi(y) \mu(y, x), \quad \forall x \in S.$$

where μ is the Möbius function.

Posets, incidence algebra and Möbius inverse

Definition (*Quantity of interest*). An A -valued Qol on a model G with random inputs $X \in P_X$, is an application:

$$\phi : P(E) \rightarrow M(E) \rightarrow A$$

$$P \rightarrow H \mapsto \phi_P(H).$$

onto G and P_X , i.e., $\phi_{P_X}(G)$.

Lemma (*Möbius decomposition*). Let $G \in M$ a model with E -valued random inputs $X \in P_X \in P(E)$. Let $\phi_{P_X}(G)$ be a Qol on G . Let $\varphi : P(D) \rightarrow A$ be a set function such that:

$$\varphi_D = \phi_{P_X}(G).$$

and $\forall A \in P(D)$, φ_A is well-defined. Then, $\phi_{P_X}(G)$ admits the following coalitional decomposition:

$$\phi_{P_X}(G) = \sum_{A \in P(D)} \psi_A,$$

where, $\forall A \in D$, $\psi_A = \sum_{B \subseteq A} (-1)^{|A|-|B|} \varphi_B$.