Sensitivity analysis of discrete preference functions using Koszul simplicial complexes

Jose Divasón
jose.divason@unirioja.es
Universidad de La Rioja
Spain

Fatemeh Mohammadi
fatemeh.mohammadi@kuleuven.be
UiT The Arctic University of Norway
and KU Leuven
Belgium

Eduardo Sáenz-de-Cabezón
esaenz-d@unirioja.es
Universidad de La Rioja
Spain

Henry P. Wynn
h.wynn@lse.ac.uk
London School of Economics
and the Alan Turing Institute
UK

ABSTRACT

We use a monomial ideal I to model a discrete preference function on a set of n factors. We can measure the sensitivity of each point represented by a monomial m by calculating its formal partial derivatives with respect to each variable. These derivatives can be used to define the Koszul simplicial complex of the ideal I at m. We refer to points at which the homology of their Koszul complex is not null as sensitive corners. In the context of preference analysis, the ranks of the homology groups are not precise enough to distinguish between sensitive corners that have the same homology but correspond to different sensitivity behaviors. To address this issue, we propose using a filtration on the Koszul complexes of the sensitive corners based on the lcm-lattice of the ideal I. This filtration induces a persistent homology at each corner m. We then use unsupervised Machine Learning methods to classify the corners based on the distance between their persistence diagrams.

CCS CONCEPTS

• **Applied computing → Mathematics and statistics**; **Multi-criterion optimization and decision-making**.

KEYWORDS

sensitivity analysis, monomial ideals, Koszul simplicial complexes, persistent homology, unsupervised machine learning

ACM Reference Format:
Jose Divasón, Fatemeh Mohammadi, Eduardo Sáenz-de-Cabezón, and Henry P. Wynn. 2023. Sensitivity analysis of discrete preference functions using Koszul simplicial complexes. In *Proceedings of ISSAC 2023 (ISSAC '23)*. ACM, New York, NY, USA, 9 pages. https://doi.org/XXXXXXX.XXXXXXX

ISSAC '23, July 24–27, 2023, Tromsø, Norway

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Sensitivity analysis is a method for studying how uncertainty in the output of a model can be attributed to different sources of uncertainty in the model input [33]. This technique explores the relationship between the input and output variables of a model. There are two main categories of sensitivity analysis techniques: local and global. Local sensitivity analysis is performed by varying model parameters around specific values to investigate how small input perturbations affect the model's performance. In contrast, global techniques vary the uncertain factors within the whole space of variable model responses [30]. Sensitivity analysis is used for various purposes, including exploratory modeling, model evaluation, model simplification, and model refinement. Depending on the goals, sensitivity analysis is used in three main modes: factor prioritization, factor fixing, and factor mapping [30, 33, 36].

Our model is a multi-factor or multi-objective decision-making system that takes several factors into account to make a binary decision. Sensitivity analysis in this context focuses on identifying which combinations of inputs are most important in determining the decision and should be revised carefully. Applications of this type of system and sensitivity analysis range from Energy Management [22] to Computer-Aided Design [14] and Machine Learning [1]. There are various techniques for dealing with multifactor decision-making systems, including Bayesian networks [18], weighting scoring rules [37], and maximum likelihood estimates [29], among others. This is a wide research area in which several different configurations and definitions of systems, factors and preference functions have been described and applied.

In this paper, we introduce an algebraic-combinatorial approach based on monomial ideals as a way to encode monotone preference multi-factor systems. First, we use the homological properties of the ideal associated with the system to sample the model's variability space. We choose those points (multi-degrees in this formulation) at which the Betti numbers are nonzero. At each of those points, which we call sensitive corners, we then perform a local analysis to find out how small perturbations of the factors influence the decision-making system. This local sensitivity analysis is a variation of the derivative-based methods used in the literature (cf. [4]). A contribution of this methodology is that the sampling of the relevant points is based precisely on the local interaction between factors, therefore taking advantage of a global and local approach. The novelty of this proposed methodology lies in the combination of mathematical and computer science techniques, including simplicial complexes, monomial ideals, and machine learning, which have not been used together before for this specific problem. This approach has the potential to open new avenues of research in the field.

Our set-up is the following. We have a decision system that uses n factors to determine whether or not to accept a particular asset. Each factor can have discrete and possibly finite levels, which represent increasing levels of satisfaction with that factor. In a binary system, there are multiple *minimum acceptance points*, which are combinations of factor scores where any decrease in satisfaction in any of the factors results in rejection of the asset. The *acceptance region* of the system is defined as all points that are above at least one of the minimum acceptance points. A multi-level preference system has a certain number of preference levels, and each level l has its own minimum acceptance points and acceptance region. It's important to note that if level l has more preference levels than level l', then any acceptance point for level l is also an acceptance point for level l'. In other words, the acceptance region for level l is always included in the acceptance region for level l'.

Example 1.1. As a simple example, consider a binary system that decides whether to accept a hotel offer based on two factors: price and distance to train station. Assuming that factor *price* has six score levels ranging from 0, meaning the most expensive, to 5, meaning the most affordable, and that factor *distance-to-station* has three score levels, ranging from 0, meaning more than $3km$ to the train station, to 2, meaning less than $500m$ to the train station, the system has determined three minimal acceptance points: Point 1 is $(4, 0)$, i.e. a hotel is acceptable if its price score is at least 4 regardless of the distance to the station, Point 2 is $(2, 1)$ meaning that a hotel is acceptable if the price score is at least 2 provided the score of the distance-to-station factor is at least 1. Finally, Point 3 is $(1, 2)$.

To study these types of systems, we use an algebraic modeling approach via monomial ideals. The monomials in the monomial ideal generated at the multi-degrees that correspond to the minimal acceptance points represent the acceptance regions. There are specific multi-degrees at the boundary between the ideal and its complement that contain homological information about the ideal. This information can be interpreted as the simplicial homology of certain simplicial complexes located at these multi-degrees. In the context of the decision system, these particular multi-degrees correspond to *corners* or points where the relations among the scores of the factors are most important in determining the output choice. The word corner here emphasizes the fact that these particular points correspond to the least common multiples of some subset of the set of minimal acceptance points. The local sensitivity analysis of the system is performed by carefully examining these corners. This examination is done in algebraic terms using the persistent homology of the simplicial complexes at the sensitive corners. To define such persistent homology, we use a filtration based on the least common multiples of the generators of the ideal that models the system. This filtration reflects the structure of the interaction between the factors. Finally, with this local analysis, we can use unsupervised clustering algorithms to classify the relevant points of the system based on their homology.

Throughout the paper, well-known results and standard mathematical notions are used; hence, their proofs are omitted.

2 MAIN ALGEBRAIC CONCEPTS

In this section, we provide an overview of the main concepts and results concerning monomial ideals and their homological structure, which are relevant to the sensitivity analysis described in Section 3.

2.1 Monomial ideals and Betti numbers

Let I be a monomial ideal $I \subseteq S = \mathbf{k}[x_1, \ldots, x_n]$. Since I has a graded module structure over S, we can consider a free S-resolution of I, which we denote by $\mathbb{F}(I)$:

$$\mathbb{F}(I) : 0 \longrightarrow F_l \xrightarrow{\varphi_l} F_{l-1} \longrightarrow \cdots \longrightarrow F_1 \xrightarrow{\varphi_1} F_0 \xrightarrow{\varphi_0} I \longrightarrow 0,$$

where each F_i is a free S-module. Among these resolutions there is a *minimal* one, that is unique up to isomorphism. In this case, the rank of each of the i-th modules of $\mathbb{F}(I)$is an invariant of I known as the i-th Betti number of I, denoted by $\beta_i(I)$. In the case of monomial ideals, there are graded and multi-graded versions of the Betti numbers, $\beta_{i,j}(I)$ and $\beta_{i,\mu}(I)$, for $i, j \in \mathbb{N}$ and μ a multi-degree in $\mathbb{N}^n$. The morphisms φ_i in $\mathbb{F}(I)$ are given by matrices with monomial entries. The resolution is minimal if none of those matrices has nonzero scalar entries or equivalently, if $\mathrm{im}(\varphi_i)$ minimally generates the i-th syzygy module of I, cf. [9, 15]. Moreover, the (multi-)graded Betti numbers of I can be seen as the dimensions of the Tor modules of I with respect to $\mathbf{k}$:

$$\beta_{i,\mu}(I) = \dim_{\mathbf{k}} \mathrm{Tor}^S_{i,\mu}(\mathbf{k}, I).$$

2.2 Stanley-Reisner and Koszul simplicial complexes

A remarkable result in combinatorial commutative algebra is the equivalence between the Betti numbers of monomial ideals and the dimensions of the homology groups of simplicial complexes. Hochster's formula, based on the Stanley-Reisner equivalence, establishes this correspondence and can be used for the efficient computation of the homology of simplicial complexes [3]. Given an abstract simplicial complex Δ, its Stanley-Reisner ideal is defined as $I_\Delta = \langle \mathbf{x}^\sigma : \sigma \notin \Delta \rangle$, where the minimal generators of I_Δ correspond to the minimal non-faces of Δ. I_Δ is a square-free monomial ideal in the variables associated with the vertices of Δ. Conversely, for every square-free monomial ideal I, there exists a Stanley-Reisner complex Δ_I, whose faces are the square-free monomials not in I. We have that $I_{\Delta_I} = I$.

The Betti numbers of the ideal I_Δ are equal to the dimensions of the (co-)homology groups of all the subcomplexes $\Delta|_\sigma = \{\tau \in \Delta : \tau \subseteq \sigma\}$ via Hochster's formula:

Theorem 2.1 (Hochster [16])**.**

$$\beta_{i-1,\sigma}(I_\Delta) = \dim \widetilde{H}^{|\sigma|-i-1}(\Delta|_\sigma; \mathbf{k})$$

The correspondence between square-free monomial ideals and simplicial complexes is a fundamental concept in combinatorial commutative algebra, known as Stanley-Reisner theory. However, this theory only applies to square-free monomial ideals. To extend this theory to general monomial ideals, we can use the polarization operation, as described in [15]. This operation transforms any monomial ideal I into a square-free monomial ideal $P(I)$, which can be analyzed using the theory of simplicial complexes. Another approach for analyzing general monomial ideals using simplicial complexes is through the use of local simplicial complexes at each multi-degree $\mu \in S$. This is achieved by constructing Koszul simplicial complexes, as detailed in [23]. By studying the homology groups of these simplicial complexes, we can gain insights into the homological structure of the monomial ideal and its relationship to the original decision system.

Definition 2.2. Let I be a monomial ideal and $\mu \in \mathbb{N}^n$. The upper and lower Koszul simplicial complexes at x^μ with respect to I are respectively defined as

$$K^\mu(I) = \{\tau \subset \operatorname{supp}(\mu) \,:\, x^{\mu-\tau} \in I\}$$

$$K_\mu(I) = \{\tau \subset \operatorname{supp}(\mu) \,:\, x^{\mu'+\tau} \notin I\}$$

where μ' is defined by subtracting one from each nonzero coordinate of μ, i.e. $\mu' = \mu - \operatorname{supp}(\mu)$.

Remark 2.3. *Observe that if I is a square-free monomial ideal, then $K_{(1,\dots,1)}(I) = \Delta_I$. Hence these complexes generalize the Stanley-Reisner correspondence.*

Hochster's formula can be stated in terms of Koszul complexes at each multi-degree $\mu \in S$.

Theorem 2.4 (cf. [23], Theorem 1.34 and Theorem 5.11)**.**

$$\beta_{i,\mu}(I) = \dim_{\mathbf{k}} \tilde{H}_{i-1}(K^\mu(I))$$

$$\beta_{i-1,\mu}(I) = \dim_{\mathbf{k}} \tilde{H}_{|\mu|-i-1}(K_\mu(I))$$

For ease of notations, in this paper we use mainly upper Koszul complexes, but the analysis applies dually to lower complexes.

2.3 The lcm-lattice

Suppose we have a set of monomials $M = \{\mathbf{x}^{\mu_1}, \dots, \mathbf{x}^{\mu_r}\}$ in the polynomial ring $S = \mathbf{k}[x_1, \dots, x_n]$ over a field $\mathbf{k}$. Let m_i be the largest exponent of variable x_i among the monomials in M. We can compute the least common multiple of the elements in M as $\operatorname{lcm}(M) = \prod_i m_i$.

Definition 2.5. Let I be a monomial ideal in S and let $G(I) = \langle \mathbf{x}^{\mu_1}, \dots, \mathbf{x}^{\mu_r} \rangle$ be its minimal monomial generating set. Let $L_I = \{\mathbf{x}^\mu \,:\, \mu = \operatorname{lcm}(\sigma) \text{ for some } \sigma \subseteq G(I)\}$. The set L_I, ordered by divisibility, forms a finite atomic lattice, called the lcm-lattice of I.

The lcm-lattice of I contains most of the homological information about I. Although there are infinitely many possible multi-degrees to examine when applying Theorem 2.4, only a finite number of them are nonzero. In fact, nonzero Betti numbers occur only at multi-degrees that are the least common multiples of sets of minimal generators of I, as shown in [12]. The following theorem shows that the Betti numbers of I can be directly determined from the homology of the lcm-lattice, making explicit the strong relation between the homological structure of monomial ideals and their *lcm*-lattice.

Theorem 2.6 (Hochster's formula for the lcm-lattice)**.** *Let I be a monomial ideal and let L_I be its* lcm*-lattice. For any $\mathbf{x}^\mu$ in L_I the multigraded Betti numbers of I with multi-degree μ are given by the reduced homology of the order complex of the open interval $(\hat{0}, \mathbf{x}^\mu)$.*

The lcm-lattice has been extensively studied in the context of free resolutions of monomial ideals and their relations to atomic lattices. See [20, 28] and references therein for details on this theory.

2.4 The lcm-filtration of a monomial ideal

Definition 2.7. Let I be a monomial ideal. A chain of ideals $I_1 \subseteq I_2 \subseteq \cdots \subseteq I_k = I$ is called a *filtration* of I. Similarly, a chain of ideals $I = I_1 \supseteq I_2 \supseteq \cdots \supseteq I_k$ is called a *reverse filtration* of I.

Definition 2.8. Let $I \subseteq S = \mathbf{k}[x_1, \dots, x_n]$ be a monomial ideal and $G(I) = \{m_1, \dots, m_r\}$ be a minimal monomial generating system of I. Let I_k be the ideal generated by the least common multiples of all sets of k distinct monomial generators of I,

$$I_k = \langle \operatorname{lcm}(\sigma) : \sigma \subseteq \{1, \dots, r\}, |\sigma| = k \rangle$$

where $\operatorname{lcm}(\sigma) = \operatorname{lcm}(\{m_i\}_{i \in \sigma})$. We call I_k the k-fold lcm-ideal of I. The ideals I_k form a descending filtration

$$I = I_1 \supseteq I_2 \supseteq \cdots \supseteq I_r$$

which we call the lcm-*filtration* of I.

The lcm-lattice-based filtration of a monomial ideal I is a legitimate structural filtration of I as it is based on the lcm-lattice of I. The underlying concept of this filtration is to investigate the changes in the features of the ideal when considering sets of generators instead of individual generators. For example, in [31, 32], monomial ideals have been used to study failure events and reliability of coherent systems, where each monomial generator represents a basic working or failure event of the system. To investigate simultaneous events and the signature analysis of coherent systems, ideals generated by consecutively taking least common multiples of a monomial ideal have been utilized [24, 25].

Example 2.9. The corresponding ideal of the system in Example 1.1 is $I = \langle x_1^4, x_1^2x_2, x_1x_2^2 \rangle \subseteq \mathbf{k}[x_1, x_2]$. The elements of the lcm-lattice of I are $\{x_1^4, x_1^2x_2, x_1x_2^2, x_1^4x_2, x_1^4x_2^2, x_1^2x_2^2\}$. The lcm-filtration of I is given by:

$$I_1 = I = \langle x_1^4, x_1^2x_2, x_1x_2^2 \rangle,\ I_2 = \langle x_1^4x_2, x_1^2x_2^2 \rangle,\ I_3 = \langle x_1^4x_2^2 \rangle.$$

Figure 1 depicts the staircase diagram of the elements of the lcm-filtration of I and the elements of the lcm-lattice of I. Observe that at $\mu = x_1^4x_2^2$ the corresponding Koszul complexes are:

$$K^\mu(I_1) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\},\ K^\mu(I_2) = \{\emptyset, \{1\}, \{2\}\},\ K^\mu(I_3) = \{\emptyset\}$$

$$K_\mu(I_1) = \{\},\ K_\mu(I_2) = \{\emptyset\},\ K_\mu(I_3) = \{\emptyset, \{1\}, \{2\}\}.$$

Using Theorems 2.4 or 2.6, $\beta_{1,x_1^4x_2^2}(I_2) = 1$, $\beta_{0,x_1^4x_2^2}(I_3) = 1$. All other Betti numbers at $x_1^4x_2^2$ are zero for all ideals in the filtration.

The lcm-filtration is a natural structural (reverse) filtration of a monomial ideal. Let $G(I) = \{m_1, \dots, m_r\}$ be the set of minimal generators of the monomial ideal $I \subseteq R = \mathbf{k}[x_1, \dots, x_n]$. The lcm-filtration has the form

$$I = I_1 \supseteq I_2 \supseteq \cdots \supseteq I_r \supseteq \langle \emptyset \rangle,$$

where $\langle \emptyset \rangle$ is the ideal generated by the empty set of monomials in n variables. Using the Stanley-Reisner correspondence, we have that for each k, $\Delta_{I_k} \subseteq \Delta_{I_{k+1}}$, also, $\Delta_{\langle \emptyset \rangle}$ is the full simplex on n vertices, Δ_n. Therefore, the lcm-filtration induces a filtration on the full simplex Δ_n which starts at Δ_I. In the forthcoming paper [10], we further investigate the properties of the induced lcm-filtrations.

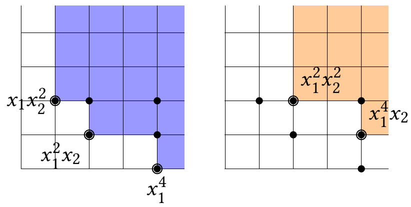


Figure 1: Staircase diagrams of the lcm-filtration $I = I_1$, I_2, I_3 with the elements of the lcm-lattice of $I = I_1$.

3 SENSITIVE CORNERS OF MULTI-FACTOR DECISION SYSTEMS

Consider a decision system $\mathcal{S}$ on n factors with minimal acceptance points $P_1, \dots, P_r$, where each $P_i = (p_{i,1}, \dots, p_{i,n}) \in \mathbb{N}^n$. Let $I_{\mathcal{S}} \subseteq S = \mathbf{k}[x_1, \dots, x_n]$ be the monomial ideal generated by $\{\mathbf{x}^{P_1}, \dots, \mathbf{x}^{P_r}\}$. Let M be the set of monomials in the ring S, and $M(I)$ be the set of monomials in I. We define $f_I : M \longrightarrow \{0,1\}$ as the indicator function for I on M. For any $\mathbf{a} = (a_1, \dots, a_n)$ in the acceptance region of $\mathcal{S}$, we have $\mathbf{x}^{\mathbf{a}} \in M(I_{\mathcal{S}})$, and $f_{I_{\mathcal{S}}}(\mathbf{a}) = 1$. We now consider the homology of sensitive corners. The upper Koszul simplicial complex $K^{\mathbf{a}}(I_{\mathcal{S}})$ is empty for monomials corresponding to points outside the acceptance region of $\mathcal{S}$. For corners strictly inside the acceptance region, their upper Koszul complex is a full simplex. Note that even if our formulation is based on upper Koszul complexes, a dual formulation based on lower complexes is analogous and might be more convenient in some problems.

Definition 3.1. Let $\mathbf{a}$ be a point in the acceptance region of the decision system $\mathcal{S}$. We say that $\mathbf{a}$ is on the *boundary of the acceptance region* of $\mathcal{S}$ if $K^{\mathbf{a}}(I_{\mathcal{S}}) \neq \emptyset$ and $K^{\mathbf{a}}(I_{\mathcal{S}}) \neq \Delta$, where Δ is the full simplex on the vertices given by the support of $\mathbf{a}$.

We can characterize the boundary of the acceptance region of $\mathcal{S}$ in terms of the indicator function of $I_{\mathcal{S}}$. Specifically, for each $\mathbf{a} \in \mathbb{N}^n$, we consider the function from $\{0,1\}^n$ to $\{0,1\}$ that maps each set σ to $f_I(\mathbf{x}^{\mathbf{a}}/\mathbf{x}^{\sigma})$, which is the indicator function evaluated at the point given by the formal partial derivative of $\mathbf{a}$ with respect to the variables in σ. The points in the boundary region are those for which this function is not constant. Observe that the only monomials that have non-null homology in their Koszul complex correspond to points in the boundary region. Therefore, we focus our attention on these points in order to gain a better understanding of the homology of the complex.

Definition 3.2. We say that a point $\mathbf{a}$ is *sensitive at degree i* or *i-sensitive* for the system $\mathcal{S}$ if $\beta_{i,\mathbf{a}}(I_{\mathcal{S}}) \neq 0$. If $\mathbf{a}$ is sensitive for at least one degree, we say that it is a *sensitive corner.*

Our methodology is based on the sensitive corners of a multi-factor decision system, which correspond to the sampling points that capture the variability of the system's output values as a function of the input factors. These sensitive corners are characterized by their non-null homology values, which indicate local changes in the behavior of the system around these points. It is important to note that the number of sensitive corners is finite, while the boundary region, where the homology values are not null, is not.

Figure 2: Two different simplicial complexes on 4 vertices for which $\beta_1(I_{\mathcal{S}}) = 1$

Example 3.3. Consider the system in Example 1.1. The 1-sensitive points of this system are $(4,1)$ and $(2,2)$. The minimal acceptance points are the only 0-sensitive points. The point $(3,1)$ is a boundary point of the acceptance region that is not a sensitive corner. The other points in the boundary region have the form $(x,0)$ for $x > 4$ or $(1,y)$ for $y > 2$.

Sensitive corners have a simplicial interpretation. The fact that $\beta_{i,\mathbf{a}}(I_{\mathcal{S}}) \neq 0$ for some point $\mathbf{a}$ is equivalent, by Theorem 2.4, to the non-null homology of the Koszul simplicial complex at $\mathbf{a}$ in dimension $i' = |\mathbf{a}| - i$. This implies the existence of an i'-dimensional hole, i.e., some i'-cycle of $K^{\mathbf{a}}(I_{\mathcal{S}})$ that is not a boundary of any $(i'+1)$-face. In the context of the system, this indicates an interaction among $(i'+1)$-factors that takes the point outside the acceptance zone, while all its i'-factor sub-iterations remain within the acceptance zone. These corners are critical because they involve combinations of inputs that play a particularly significant role in determining the decision outcome and must be analyzed with great care, which is precisely the objective of sensitivity analysis. Note that the Koszul complex at $\mathbf{a}$ consists of those $\sigma \in \{0,1\}^n$ such that $f_{I_{\mathcal{S}}}(x^{\mathbf{a}}/x^{\sigma}) = 1$. The relationship between Betti numbers and the complexity of hierarchical models was studied in [21].

Example 3.4. Consider the system in Example 1.1. Point $(4,1)$ is in the acceptance zone. If we decrease the score of any of the factors by 1, arriving at points $(3,1)$ or $(4,0)$, we are still in the acceptance zone. However, if we decrease both scores simultaneously, we arrive at point $(3,0)$, which is outside the acceptance zone. This behavior of the two-factor interaction is different from both one-factor interactions and is detected by the 0-th simplicial homology of $K^{(4,1)}(I_{\mathcal{S}})$ or, equivalently, $\beta_{1,(4,1)}(I_{\mathcal{S}})$.

3.1 Persistent homology of sensitive corners

The second tool in our methodology is a close examination of the local behavior of the decision function around our sampling points, specifically around the sensitive corners. To carry out this analysis, we require invariants that are more detailed than the Betti numbers. To see why, consider the following two simplicial complexes in four vertices represented by their sets of facets: $\Delta = \{\{1,2\},\{2,3\},\{3,4\},\{1,4\}\}$ and $\Delta' = \{\{1,2,3\},\{3,4\},\{1,4\}\}$ (see Figure 2). The vertices can be thought of as factors in a system $\mathcal{S}$, and an i-dimensional face represents i-fold interactions among the factors. Both complexes have a non-null 1-st homology group of rank 1, meaning that they both have a one-dimensional hole. However, they correspond to different patterns of 3-fold interactions with respect to 2-fold interactions.

One of the main advantages of persistent homology is that it can capture the topological features of a simplicial complex at different scales, which allows us to detect topological changes as we move through the filtration. This can be useful in identifying important features of a system that may be missed by traditional homology techniques. Additionally, persistent homology can be used to construct a summary of the topological structure of a simplicial complex in the form of a persistence diagram, which can be easier to interpret than a set of Betti numbers. In the context of sensitivity analysis, persistent homology can be used to identify which interactions among the factors of a system are responsible for the appearance of certain homological features. By analyzing the persistence intervals associated with these features, we can determine which interactions are important for maintaining certain properties of the system, such as being in the acceptance zone. This information can be used to guide the design of experiments aimed at further exploring the system's behavior. Overall, persistent homology provides a powerful tool for analyzing the topological structure of simplicial complexes and can be particularly useful for sensitivity analysis where we need to understand how the behavior of a system changes in response to different inputs.

To study the homological features that persist in a filtration of a simplicial complex Δ, we consider the sequence of homology groups through the filtration. Let $\mathcal{F}: \Delta_0 \subseteq \cdots \subseteq \Delta_k = \Delta$ be a filtration of Δ. Persistence homology is interested in the sequence of homology groups through the filtration. For every $i \leq j$, consider the inclusion map $\Delta_i \xrightarrow{\mathrm{i}^{i,j}} \Delta_j$. These maps induce homomorphisms of homology groups for each dimension p, $f_p^{i,j}: H_p(\Delta_i) \longrightarrow H_p(\Delta_j)$. Hence, for this filtration we have, for each dimension p, a sequence of homology groups connected by homomorphisms,

$$H_p(\Delta_0) \longrightarrow H_p(\Delta_1) \longrightarrow \cdots \longrightarrow H_p(\Delta_{k-1}) \longrightarrow H_p(\Delta_k).$$

These modules provide a measure of the homological features that persist in the filtration, and are used to study the underlying dataset or simplicial complex at different scales.

Definition 3.5. The p-th persistent homology groups of Δ with respect to $\mathcal{F}$ are defined as the images of the homomorphisms induced by the inclusions in $\mathcal{F}$, i.e., $H_p^{i,j}(\Delta, \mathcal{F}) = \mathrm{im}(f_p^{i,j})$, where $f_p^{i,j}: H_p(\Delta_i) \to H_p(\Delta_j)$ is the homomorphism induced by the inclusion map $\Delta_i \hookrightarrow \Delta_j$. The rank of $H_p^{i,j}(\Delta, \mathcal{F})$, denoted by $\beta_p^{i,j}(\Delta, \mathcal{F})$, is called the p-th persistent Betti number of Δ according to $\mathcal{F}$.

The persistent homology groups consist of the classes

$$Z_p(\Delta_i)/(B_p(\Delta_j) \cap Z_p(\Delta_i)),$$

i.e. the homology classes of Δ_i that are still alive at Δ_j. Here $Z_p(\Delta_i)$ is the p-cycle of Δ_i and $B_p(\Delta_j)$ is the p-boundaries of Δ_j. If κ is a class in $H_p(\Delta_i)$, we say that it is born at step i of the filtration, if $\kappa \notin H_p^{i-1,i}(\Delta, \mathcal{F})$, and that it dies at step j if it merges with an older class when going from Δ_{j-e} to Δ_j, i.e. $f_p^{i,j-1}(\kappa) \notin H_p^{i-1,j-1}(\Delta, \mathcal{F})$ and $f_p^{i-1,j}(\kappa) \in H_p^{i-1,j}(\Delta, \mathcal{F})$.

One of the main applications of persistent homology is in the field of Topological Data Analysis (TDA). The goal of TDA is to extract topological features from a given set of data points or space. In order to do so, the space is first represented as a simplicial complex. To apply persistent homology, two key elements are required. First, a notion of distance or a measure of dissimilarity between the points or elements of the set is needed. Second, a parameter is selected that drives the construction of a filtration on the simplicial complex. Both the parameter and the underlying distance are chosen based on some property of the initial set of points, such as Euclidean distance, time, scale of resolution, etc.

The Vietoris-Rips filtration, based on the Vietoris complex [35], is one of the most widely used filtrations in TDA. It can be used alone or in combination with the Čech complex [6, 7]. These filtrations, along with other constructions, allow persistent homology to extract topological information from complex data sets, making it a powerful tool in fields such as biology, neuroscience, computer science, and more.

In some cases, there may not be a natural notion of distance or time to build a meaningful filtration when studying a simplicial complex itself, or phenomena that can be modeled by simplicial complexes. In such situations, a structural or intrinsic filtration based solely on the structure of the simplicial complex must be considered. This is the case for the preference systems described in this paper. In the following sections, we utilize the lcm-filtration to examine the local behavior of the decision function near the sensitive corners of the system. An advantage of this filtration in this context is that at each step k of the filtration, we obtain a new region given by the combinations of any k minimal acceptance points of the system $\mathcal{S}$. When $k = 1$ this is the acceptance region of $\mathcal{S}$, when $k > 1$ we speak of the k-fold acceptance region. The *lcm*-filtration allows the study of the behaviour of the sensitive corners with respect to the k-fold acceptance regions for $k \geq 1$. This provides a finer analysis of the sensitivity of the corner.

4 CLUSTERING OF SENSITIVE CORNERS BY THEIR PERSISTENCE DIAGRAMS

This paper presents a methodology for conducting sensitivity analysis of multi-factor decision systems. To achieve this, we introduce two tools: the first tool involves sampling based on the homological properties of the ideal associated with the system, with the resulting points referred to as sensitive corners. The second tool involves a local analysis of these sensitive corners, which we perform through their algebraic and combinatorial properties, and in particular, using persistent homology.

As demonstrated in Example 3.4, we use the Betti numbers of the system ideal at multi-degrees that correspond to the sensitive corners. These corners represent a subset of the lcm-lattice of $I_{\mathcal{S}}$. For larger systems with a high number of factors, there may be several i-sensitive corners for each i. We can analyze these corners by utilizing the local information surrounding them, which we obtain through the persistent homology based on the lcm-filtration. Consequently, we propose the following approach for classifying sensitive corners:

(1) Compute the persistence homology of each i-sensitive corner based on the induced lcm-filtration on $I_{\mathcal{S}}$.
(2) Compute the matrix of distances among the persistence diagrams obtained.
(3) Use a machine learning algorithm to cluster the persistence diagrams into groups based on their distances.

Step 1 is based on the definitions in the previous section, and can be performed by several software packages like [26, 34].

Step 2 needs a concept of distance. Several distances have been defined for persistence diagrams. The most used ones are *bottleneck* and *Wasserstein* distances [8, 11]. Before explaining Step 3, we first recall the definitions of p-Wasserstein and the bottleneck distances.

Definition 4.1. Let D_1 and D_2 be two persistence diagrams, $M \subseteq D_1 \times D_2$ an optimal matching, and M^c the set of unmatched points. The p-Wasserstein distance between D_1 and D_2 is defined as

$$W_p(D_1, D_2) := \inf_M \left(\sum_{(x,y)\in M} \|x-y\|_\infty^p + \sum_{(x,y)\in M^c} |x-y|^p \right)^{\frac{1}{p}},$$

where $\|\cdot\|_\infty$ denotes the ∞-norm and M ranges over all matchings between D_1 and D_2. Here, $p > 0$.

Definition 4.2. The bottleneck distance is defined as the limit:

$$W_\infty(D_1, D_2) := \lim_{p\to\infty} W_p(D_1, D_2).$$

Step 3 can be performed using a variety of methods. In general, unsupervised machine learning (ML) involves dividing n observations into k clusters. In our approach, we emphasize that the statistical units for the clustering analysis are persistence diagrams, and the distance metric used is a distance between persistence diagrams. There are many clustering algorithms available, but as with other ML problems, there is no single technique that is universally best. The most popular unsupervised ML algorithm is the k-means algorithm, which minimizes within-cluster squared Euclidean distances and calculates cluster centers as the mean of the cluster points. However, k-means is not a suitable choice for our purposes because we need to work directly with the distance matrix since the statistical units are persistence diagrams. Therefore, we have opted to use another well-known unsupervised ML algorithm: the k-medoids algorithm. The k-medoids algorithm is similar to k-means but with two main differences. Firstly, the k-medoids algorithm chooses actual data points as centers, and secondly, it can be used with arbitrary distance measures (while k-means requires Euclidean distance). As a result, it can be adapted to cluster observations into k groups, using an initial distance matrix to find the new medoids at each iterative step [17]. Other algorithms, such as agglomerative clustering [27] and OPTICS clustering [2], are also capable of working with a precomputed distance matrix and are valid alternatives for Step 3.

5 A FULLY COMPUTED EXAMPLE

Let us consider a system $\mathcal{S}$ on eight factors $x_1, \ldots, x_8$, some of which are binary (factors 1, 2, 4, 6, 7), and some multi-level (factors 3 and 5 have three levels, factor 8 has five levels) . The output of the system is binary. The system is described by the following set of eleven minimal acceptance points:

$$\{x_1, x_2, x_3^2, x_3x_5, x_3x_8^2, x_4, x_5^2, x_5x_8, x_6, x_7, x_8^4\}. \tag{1}$$

The monomial ideal $I_\mathcal{S}$ generated by the monomials corresponding to these eleven points has the following Betti diagram (computed using `Macaulay2` [13]):

	0	1	2	3	4	5	6	7
total:	**11**	**48**	**113**	**160**	**141**	**76**	**23**	**3**
1 :	5	10	10	5	1	.	.	.
2 :	4	24	61	85	70	34	9	1
3 :	1	7	21	35	35	21	7	1
4 :	1	7	21	35	35	21	7	1

The ideal $I_\mathcal{S}$ of the system has an associated lcm-lattice with 895 corners. Among these corners, 575 exhibit homology at some dimension. Of the eleven multi-degrees with 0-dimensional homology, five have degree 1, four have degree 2, and so on. The 48 corners with 1-dimensional homology correspond to holes in two-factor interaction. Although these corners have different degrees, we can distinguish them using the procedure presented in this paper. Additionally, there are 113 corners with 2-dimensional homology, 160 with 3-dimensional homology, and so on.

We computed the persistent homology of the sensitive corners of the ideal using the lcm-filtration of $I_\mathcal{S}$ and the software package `Dionysus` [26]. To analyze the resulting persistence diagrams, we computed a distance matrix using the bottleneck distance and employed the k-*medoids* algorithm [17] for clustering. This allowed us to identify distinct clusters of persistent homology diagrams that correspond to different features of the multi-factor decision system.

The k-medoids algorithm (as well as k-means) raises the main question of how to determine the optimal number of clusters k. Unfortunately, there is no definitive answer to this question, and the optimal number of clusters is subjective. The goal is to determine an appropriate number of clusters that satisfies two fundamental properties: compactness (how closely related the points in a cluster are) and separation (how well-separated a cluster is from other clusters). To evaluate the quality of a clustering and select the appropriate number of clusters, internal clustering validation measures can be used [19]. Concretely, two of the most commonly used techniques for determining the optimal k in both k-means and k-medoids are:

- Elbow method: a heuristic that attempts to choose the smallest number of clusters that explain the greatest amount of variation in the data. The method plots the *inertia* (the sum of distances of samples to their closest cluster center) as a function of the number of clusters and one has to pick the elbow of the curve as the number of clusters to use. Such a shift point is hard to determine and usually subjective.
- Silhouette index: an index validating the clustering performance based on the pairwise difference of between-and within-cluster distances. Its value ranges from -1 to 1, where a high value indicates that the object is well matched to its own cluster and poorly matched to neighboring clusters. The computation of Silhouette values is based on distance and, indeed, they can be calculated with any distance metric.

There exist well-known alternative clustering criteria [19], such as the Calinski–Harabasz [5] indices. However, most of them cannot deal with a precomputed distance matrix, since they usually need the points of the space, or imply Euclidean distance in order to have geometrically sensible meaning. For such reasons, we have proceeded with the elbow method and the analysis of Silhouette scores to choose the appropriate number of clusters k for each dimension.

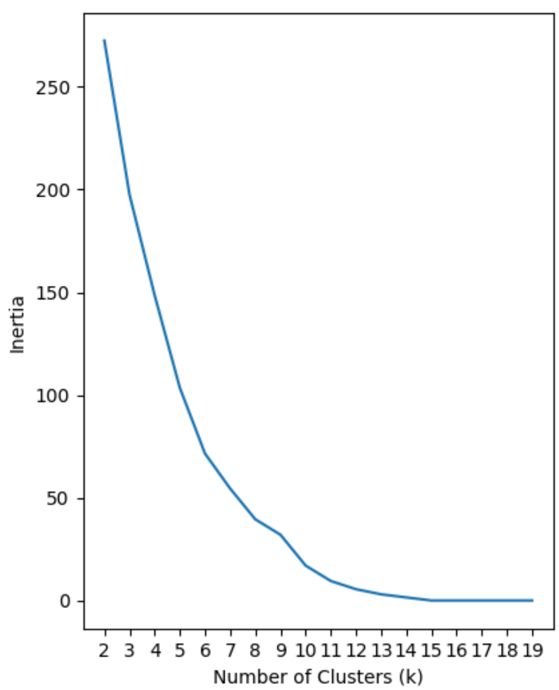


Figure 3: Elbow method for $\dim = 0$ **using** k**-medoids.**

We applied these techniques to the eight dimensions (from 0 to 7), clustering from 2 groups up to 20. For example, let us consider the 0-dimensional case. Figure 3 shows the elbow method plot for this dimension. We analyzed the results of both the elbow method and the Silhouette scores for each k, and they suggest that clustering with $k = 10$ and $k = 11$ produces good results in this dimension (Silhouette indices are 0.928 and 0.96, respectively). Handling infinity values that arise when computing the bottleneck distances is crucial since clustering algorithms cannot deal with infinity values. We set them to 0, but we have also tried with other values (for instance, with a number larger than the highest distance in each dimension) and obtained very similar results. In dimension 1, we found that $k = 9$ and $k = 13$ are good choices. However, from this dimension on, the number of optimal clusters is greatly reduced because of the high number of infinity distances. If k is too high, especially in high dimensions, the k-medoids algorithm may leave clusters empty.

We briefly explain two examples of the application of these techniques to the corners of the system $\mathcal{S}$ given by the minimal acceptance points (1).

For the first example, we consider the 48 corners that have 1-dimensional homology. Based on bottleneck distance, we have identified three groups of corners, each representing a different type of simplicial complex that starts as two separate points but progresses differently. Figure 4 shows representatives of these three groups, where rows in the picture represent the clusters, and the columns represent different stages of the lcm-filtration. As the figure illustrates, at the initial step of the filtration, all simplicial complexes in these corners are the same. However, as the filtration proceeds,

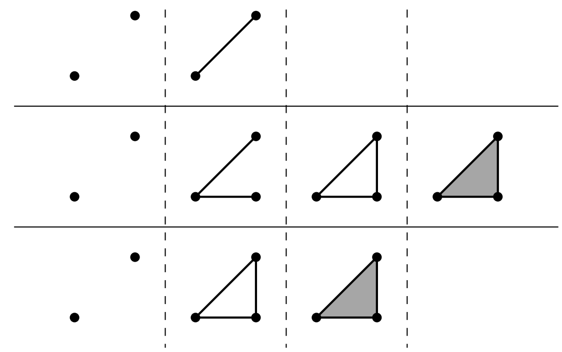


Figure 4: Filtrations of the three types of simplicial complexes for which $\beta_1(I_{\mathcal{S}}) \neq 0$

the simplicial complexes exhibit different patterns of interactions among the factors surrounding these sensitive corners. This classification constitutes an initial step in the finer analysis of these corners, which represent the few-fold interactions among factors.

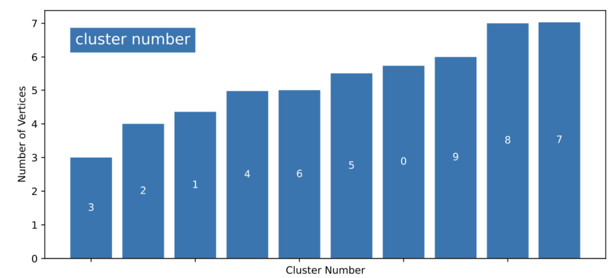
As a second example of the use of these techniques, we focus on the behavior of the 0-dimensional persistence homology modules along the lcm-filtration, which provides an overview of how the minimal acceptance points relate to each other at the sensitive corners. In particular, we would like to take a first glance at how the number of involved factors varies around these corners, by establishing groups based on the 0-dimensional homology of the Koszul simplicial complexes that correspond to the multi-degrees at the lcm-lattice of $I_{\mathcal{S}}$. Table 1 shows, for the case $\dim = 0$ and $k = 10$, the number of persistence diagrams that belong to each cluster. Most of them are grouped in cluster 1, whereas clusters 8 and 9 have significantly fewer instances. The table also shows that the results are very similar if one chooses $k = 11$, but the new cluster contains 15 persistence diagrams, most of which previously belonged to cluster 1 when $k = 10$. Figure 5 shows the mean values of basic size-related characteristics of the complexes at the corners in each of the clusters when $k = 10$. We can observe that the clusters capture the difference in the number of vertices, but more clearly the differences in the number of simplices. Hence, a first exploration using this classification indicates that the complexes in clusters 7 and 8 are bigger complexes, and those in clusters 2 and 3 are smaller. Although a first analysis of the 0-dimensional persistence diagrams provides some information, a closer examination of the persistence diagrams of these corners to see how the 0-dimensional homology evolves within the filtration will provide further knowledge about the interaction of factors around the choice problem modeled by this system.

6 CONCLUSIONS

On an acceptance region defined for a decision problem whose factor values lie on the non-negative integer lattice, and which is compatible with the usual partial ordering, there exist what we define as "sensitive points". These are special boundary points whose topologies are complex and therefore may require special attention in decision making because of intricate interactions and trade-offs between factors. The three main contributions are: (i)

Table 1: Number of persistence diagrams that belong to each cluster for $\dim = 0$, with $k = 10$ and $k = 11$ using k-medoids.

k=10		k=11	
Cluster	Size	Cluster	Size
0	115	0	113
1	207	1	196
2	150	2	150
3	90	3	90
4	159	4	157
5	64	5	64
6	30	6	30
7	35	7	35
8	15	8	15
9	30	9	30
		10	15



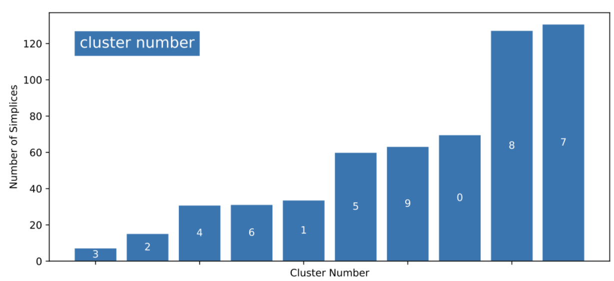


Figure 5: Basic size-related characteristics of the simplicial complexes in the 0-dim clusters: number of vertices and number of simplices.

to investigate the local topology at each sensitive point, (ii) to employ, at each such point, a new and solely algebraic version of Topological Data Analysis based on the *lcm*-filtration, and (iii) to set up distances between the points based on this local topological "data" and thereby allow the use of Machine Learning clustering methods to classify the sensitive points.

7 ACKNOWLEDGMENTS

J.D. and E.S.C. are supported by grant PID2020-116641GB-100 funded by MCIN/AEI/10.13039/501100011033. F.M. is partially supported by the KU Leuven grant iBOF/23/064, the FWO grants G0F5921N and G023721N, and the UiT Aurora project MASCOT.

REFERENCES

[1] R. Ali, S. Lee, and T.C. Chung. 2017. Accurate multi-criteria decision making methodology for recommending machine learning algorithm. *Expert Systems with Applications* 71 (2017), 257 – 278. https://doi.org/10.1016/j.eswa.2016.11.034
[2] Mihael Ankerst, Markus M Breunig, Hans-Peter Kriegel, and Jörg Sander. 1999. OPTICS: Ordering points to identify the clustering structure. *ACM Sigmod record* 28, 2 (1999), 49–60.
[3] A. Bigatti, J. Heras, and E. Sáenz de Cabezón. 2019. Monomial resolutions for efficient computation of simplicial homology. *Proceedings of ISSAC 2019* (2019), 50–57.
[4] E. Borgonovo. 2008. Sensitivity analysis of model output with input constraints: a generalized rationale for local methods. *Risk Analysis: An International Journal* 28, 3 (2008), 667–680.
[5] Tadeusz Caliński and Jerzy Harabasz. 1974. A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods* 3, 1 (1974), 1–27.
[6] E. Carlsson, G. Carlsson, and V. de Silva. 2006. An algebraic topological method for feature identification. *International Journal of Computational Geometry and Applications* 16(4) (2006), 291–314.
[7] F. Chazal, D. Cohen-Steiner, L. J. Guibas, F. Mémoli, and S. Y. Oudot. 2009. Gromov-Hausdorff stable signatures for shapes using persistence. *Comput. Graph. Forum* 28(5) (2009), 1393–1403.
[8] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. 2007. Stability of persistence diagrams. *Discrete and Computational Geometry* 37(1) (2007), 1103–120.
[9] D. Cox, J. Little, and D. O'Shea. 1998. *Using algebraic geometry*. Springer.
[10] J. Divasón, F. Mohammadi, E. Sáenz-de-Cabezón, and H.P. Wynn. 2023. Discrete persistent homology of simplicial complexes via filtrations of monomial ideals. *In preparation* (2023).
[11] H. Edelsbrunner and J. Harer. 2010. *Computational topology: an introduction*. American Mathematical Society.
[12] V. Gasharov, I. Peeva, and V. Welker. 1999. The lcm-lattice in monomial resolutions. *Mathematical Research Letters* 6 (1999), 521–532.
[13] Daniel R. Grayson and Michael E. Stillman. [n. d.]. Macaulay2, a software system for research in algebraic geometry. Available at http://www.math.uiuc.edu/Macaulay2/.
[14] J. Hao, S. Luo, and L. Pan. 2021. Computer-aided intelligent design using deep multi-objective cooperative optimization algorithm. *Future Generation Computer Systems* 124 (2021), 49–53. https://doi.org/10.1016/j.future.2021.05.014
[15] J. Herzog and T. Hibi. 2011. *Monomial Ideals*. Springer.
[16] M. Hochster. 1977. Cohen-Macaulay rings, combinatorics, and simplicial complexes. In *Ring theory, II (Proc. Second Conf., Univ. Oklahoma, Norman, Okla., 1975)*. Lecture Notes in Pure and Appl. Math, Vol. 26. Dekker, 171–223.
[17] C.H. Jun H.S. Park. 2009. A simple and fast algorithm for K-medoids clustering. *Expert systems with applications* 36, 2 (2009), 3336–3341.
[18] D.R. Insua and S. French. 1991. A framework for sensitivity analysis in discrete multi-objective decision-making. *European Journal of Operational Research* 54, 2 (1991), 176–190. https://doi.org/10.1016/0377-2217(91)90296-8
[19] Y. Liu, Z. Li, H. Xiong, X. Gao, and J. Wu. 2010. Understanding of internal clustering validation measures. In *2010 IEEE international conference on data mining*. IEEE, 911–916.
[20] S. Mapes. 2013. Finite atomic lattices and resolutions of monomial ideals. *Journal of Algebra* 379 (2013), 259–276.
[21] H. Maruri-Aguilar, E. Sáenz-de Cabezón, and H.P. Wynn. 2012. Betti numbers of polynomial hierarchical models for experimental designs. *Ann. Math. Artif. Intell.* 64 (2012), 411–426.
[22] R. McKenna, V. Bertsch, K. Mainzer, and W. Fichtner. 2018. Combining local preferences with multi-criteria decision analysis and linear optimization to develop feasible energy concepts in small communities. *European Journal of Operational Research* 268, 3 (2018), 1092–1110. https://doi.org/10.1016/j.ejor.2018.01.036 Community Operational Research: Innovations, internationalization and agenda-setting applications.
[23] E. Miller and B. Sturmfels. 2005. *Combinatorial commutative algebra*. Springer.
[24] F. Mohammadi, E. Sáenz-de-Cabezón, and H. P. Wynn. 2015. Types of signature analysis in reliability based on Hilbert series. *Journal of Symbolic Computation* 79(1) (2015), 140–155.
[25] F. Mohammadi, E. Sáenz-de-Cabezón, and H. P. Wynn. 2018. Efficient multi-cut enumeration of k-out-of-n:F and consecutive k-out-of-n:F systems. *Pattern Recognition Letters* 102 (2018), 82 – 88.
[26] D. Morozov. 2022. Dionysus. https://github.com/mrzv/dionysus. Online; accessed February 2022.
[27] Fionn Murtagh and Pedro Contreras. 2012. Algorithms for hierarchical clustering: an overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 2, 1 (2012), 86–97.
[28] J. Phan. 2006. *Order properties of monomial ideals and their free resolutions*. Ph. D. Dissertation. Columbia University.
[29] Vladislav V. Podinovski. 2020. Maximum likelihood solutions for multicriterial choice problems. *European Journal of Operational Research* 286, 1 (2020), 299–308. https://doi.org/10.1016/j.ejor.2020.03.028

[30] Patrick M. Reed, Antonia Hadjimichael, Keyvan Malek, Tina Karimi, Chris R. Vernon, Vivek Srikrishnan, Rohini S. Gupta, David F. Gold, Ben Lee, Klaus Keller, Travis B. Thurber, and Jennie S. Rice. 2022. *Addressing Uncertainty in Multisector Dynamics Research.* Zenodo. https://doi.org/10.5281/zenodo.6110623

[31] E. Sáenz-de-Cabezón and H. P. Wynn. 2009. Betti numbers and minimal free resolutions for multi-state system reliability bounds. *Journal of Symbolic Computation* 44 (2009), 1311–1325.

[32] E. Sáenz-de-Cabezón and H. P. Wynn. 2015. Hilbert functions for design in reliability. *IEEE Trans. Rel.* 64 (2015), 83–93.

[33] A. Saltelli, S. Tarantola, Campolongo F., and M. Ratto. 2004. *Sensitivity analysis in practice: a guide toassessing scientific models.* Wiley Online library, Vol. 1. Wiley.

[34] G. Tauzin, U. Lupo, L. Tunstall, J. Burella-Pérez, M. Caorsi, A. Medina-Mardones, A. Dassatti, and K. Hess. 2020. giotto-tda: A Topological Data Analysis Toolkit for Machine Learning and Data Exploration. arXiv:2004.02551 [cs.LG]

[35] L. Vietoris. 1927. Über den höheren Zusammenhang kompakter Räume und eine Klasse von zusammenhangstreuen Abbildungen. *Math. Ann.* 97 (1927), 454–472.

[36] T. Wagener and F. Pianosi. 2019. What has Global Sensitivity Analysis ever done for us? A systematic review to support scientific advances and to inform policy making in earth system modelling. *Earth Science Reviews* 194 (2019), 1–18.

[37] Y. Zhao, F. Wu, and Y. Chen. 2021. Sensitivity analyses and measurements for group decisions using weighted scoring rules. *International Transactions in Operational Research* 28, 2 (2021), 959 – 975. https://doi.org/10.1111/itor.12588

Received 31 January 2023; revised 25 April 2023; accepted 2 May 2023