Visualization of H_{k-1} Homology Groups of k-dimensional Pareto-optimal Front Using PaletteViz

Khaled Talukdar and Kalyanmoy Deb
Computational Optimization and Innovation (COIN) Laboratory
Michigan State University
428 S. Shaw Lane, 2120 EB
East Lansing, MI 48864, USA {talukde1, kdeb}@msu.edu

COIN Report Number 2021006

Abstract

PaletteViz is a novel k-dimensional ($3 \leq k \leq 10$) data visualization technique which produces a visual representation of the Pareto-optimal front in terms of its level set (i.e. depth contours). In this paper, we show that PaletteViz can also be utilized for a vis-a-vis representation of different homology groups. In this paper, we only consider the $(k-1)^{th}$ homology group $\mathbf{H}_{k-1}$ in the data point cloud. Homology groups can be extremely important in multi-criteria decision making since they describe the non-convexity and the disconnected pockets of infeasible regions in the Pareto-optimal front. None of the existing visualization methods does not have the capability to visually represent such information in a high-dimensional objective space. In this paper, we show that some simple modifications on the PaletteViz procedure can give us a novel way of visual analytics and visual multi-criteria decision making.

1 Introduction

With the current advancement in many-objective optimization (MOP) algorithms to solve problems with three or more conflicting objective functions, a set of posteriori issues need to be addressed on an urgent basis. One of the most crucial issues is to visually analyze the obtained non-dominated points in the objective space produced by an MOP solver, so that it facilitates a more efficient and practical decision making process.

A good visualization technique should enable a graphical exploration of the objective space and should help the decision maker (DM) to obtain better insights into the problem and compare different alternative Pareto-optimal solutions (i.e. in terms of the objective function values).

Such visual analyses should provide more control over the decision maker (DM) to correctly analyze different options – not just in terms of their dominance relationships or trade-offs but also in terms of *vis-a-vis* representation of the space where the data points reside. An ideal case would be to provide a DM with such a tool that could represent a high-dimensional space in an intuitive and natural way. Understanding the *shape* induced by a set of data points in a k-dimensional space is the first step to achieve such a goal. Fortunately, the field of Topological Data Analysis (TDA) can offer us a number of alternative directions, that might be extremely helpful to completely change the way the field of Multi-criteria Decision Making (MCDM) has been approaching the problems of visual data analytics for a long time. Although there exist a number of different visualization techniques in the literature for generic data analytics, however, none of them are suitable for topological analysis and visualization of topological information. Evolutionary Multi-objective Optimization (EMO) and Multi-criteria Decision Analysis (MCDA) researchers have simply borrowed them so far and used in their particular application domains [9].

While most of these methods are based on mapping high-dimen-sional data points onto a lower-dimensional (i.e. two or three-dimensional) space, they are not particularly good at conveying any of the geometric, topological, functional, and

decision-making properties that a DM would be interested in. Although in [4], authors propose a new way to visualize high-dimensional Pareto-optimal front using a Star-coordinate [6] based approach, however the proposed method doesn't address the topological nature of the data points. Another important aspect of visual analytics of Pareto-optimal data is to understand the infeasible regions and non-convexity of the space. Such features creates *pockets* or *holes* in the high-dimensional Pareto-optimal front. A DM might be interested in visually analyze the surrounding points near those infeasible or void regions, since the surrounding points might give a better insight about the underlying MOP.

In this paper, we propose a way to devise a visualization procedure that does exactly what we have described. We show a method to identify the points surrounding the holes or voids in the high-dimensional Pareto-optimal front. Formally, they are described with simplicial homology. We will apply a number of TDA methodologies to analyze and detect those points and use them in a recently proposed visualization technique called *PaletteViz* [8]. The principal motivation behind this work is show that simplicial homology can be important in MCDM, also to show PaletteViz can be a good tool to visualize those topological features in an intuitive manner.

The paper is organized as follows. In Section 2, we go an extensive length of building the basics that are essential to understand the procedures described in the next section. Then, in Section 5, we describe our proposed technique in detail. In Section 6, we present some interesting results found from applying our method on a bench mark multi-objective optimization problem in three, four and five dimensions. In the end, we conclude our paper with a discussion on its limitations, possible improvement and future directions.

2 Formal Background

In this section we describe some basic definitions that will be necessary to understand the algorithms and procedures discussed in this paper. Due to the space constraints, we will exclude the relevant algebra which can be found in more details in [7].

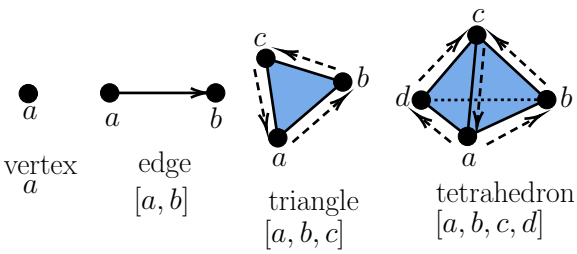


Figure 1: Oriented k-simplices in R^3, $0 \leq k \leq 3$. The orientations are shown on the edges. In the case of 3-simplices, the orientations will be drawn on the surfaces.

2.1 Simplicial Complexes

A *simplicial complex* is a set K, where $\mathcal{S}$ is a collection of subsets of K called *simplices* (simplex, if singular) such that $\forall v \in K, \{v\} \in \mathcal{S}$, and if $\tau \subseteq \sigma \in \mathcal{S}$, then $\tau \in \mathcal{S}$. The sets $\{v\}$ are the vertices of K. When it is apparent from the context what $\mathcal{S}$ means, we call set K simply as a *complex*. We say $\sigma \in \mathcal{S}$ is a k-simplex of dimension k if $|\sigma| = k + 1$. If $\tau \subseteq \sigma$, τ is a *face* of σ, and σ is a *coface* of τ.

An *orientation* of a k-simplex σ, $\sigma = \{v_0, \ldots, v_k\}$, is an equivalence class of orderings of the vertices of σ, where $(v_0, \ldots, v_k) \sim (v_{\tau(0)}, \ldots, v_{\tau(k)})$ are equivalent if the sign of τ is 1. We denote an *oriented simplex* by $[\sigma]$. A simplex may be realized geometrically as the convex hull of $k+1$ affinely independent points in $R^d, d \geq k$. This realization gives us the familiar low-dimensional k-simplices: vertices, edges, triangles, and tetrahedra, for $0 \leq k \leq 3$, shown in Figure 1. In a realized complex, the simplices must meet at common faces. A *subcomplex* of K is a subset $L \subseteq K$ that is also a simplicial complex. A *filtration* of a complex K is a nested subsequence of complexes $\emptyset = K^0 \subseteq K^1 \subseteq \ldots \subseteq K^m = K$. For generality, we set $K^i = K^m$, $\forall i \geq m$. We call K a *filtered complex*. We show an example of a filtered complex in Figure 2.

2.2 Vietoris-Rips Complex and α Complex

The geometrical realization of simplicial complexes can be defined based on how the simplices are included in K – which results into different no-

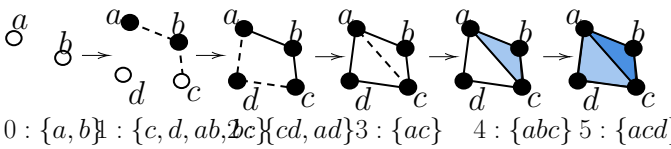


Figure 2: A filtered complex with newly added simplices highlighted.

tions of complexes. In this paper we only focus on *Vietoris-Rips Complex* and α *Complex*. The former is needed to compute the boundary points of $\mathbf{H}_{k-1}$ homology groups and the later is the basis for PaletteViz visualization method [8].

2.2.1 Vietoris-Rips Complex

Given a set $\mathcal{X}$ of points in Euclidean space E^k, the *Vietoris-Rips* complex $\mathcal{R}_\epsilon(\mathcal{X})$ is the abstract simplicial complex whose k-simplices are determined by subsets of $k+1$ points in $\mathcal{X}$ with diameter at most ϵ.

2.2.2 α Complex

The notion of α *complex* comes from the so called *union of balls* in a *Delaunay triangulation* [5].

(Union of balls) Let $\mathcal{X}$ be a finite set of points in R^k and ρ a non-negative real number. For each $p \in \mathcal{X}$, we let $\mathcal{B}_p(\rho) = p + \rho B^k$ be the closed ball with center p and radius ρ. The union of these balls is the set of points at distance at most ρ from at least one of the points in $\mathcal{X}$,

$$U(r) = \|x - p\| \leq r, s.t. x \in R^k \mid \exists p \in \mathcal{X} \qquad (1)$$

To decompose the union, we intersect each ball with the corresponding *Voronoi cell* [5] V_p, i.e. $\mathcal{R}_p(r) = \mathcal{B}_p(r) \cap V_p$. Since balls and Voronoi cells are convex, the $\mathcal{R}_p(r)$ are also convex. Any two of them are disjoint or overlap along a common piece of their boundaries, and together the $\mathcal{R}_p(r)$ cover the entire union. The α complex $\mathcal{K}_\alpha(r)$ is isomorphic to *the nerve* [7] of the attained cover,

$$\mathcal{K}_\alpha(r) = \left\{ \sigma \subseteq \mathcal{X} \mid \bigcap_{p \in \sigma} \mathcal{R}_p(r) \neq \emptyset \right\} \qquad (2)$$

Since $\mathcal{R}_p(r) \subseteq V_p$, the α complex is a subcomplex of the *Delaunay complex* (i.e. simplicial complexes arose from the triangulation). It follows that for a set $\mathcal{X}$ in general position [7], the geometric realization is defined by the convex hulls. In order to count the homology, we need understand how each k-dimensional simplicial complex are formed by a set of lower dimensional simplicial complexes. In the next section, we formalize this structure.

2.3 Chain Complex

The k^{th} *chain group* $\mathbf{C}_k$ of K is the free Abelian group on its set of oriented k-simplices, where $[\sigma] = -[\tau]$ if $\sigma = \tau$ and σ and τ are oriented opposite. An element $c \in \mathbf{C}_k$ is a k-chain, $c = \sum_i n_i [\sigma_i]$, $\sigma_i \in K$ with coefficients $n_i \in Z$. The boundary operator $\partial_k : \mathbf{C}_k \to \mathbf{C}_{k-1}$ is a homomorphism defined linearly on a chain c by its action on any simplex $\sigma = [v_0, v_1, \ldots, v_k] \in c$,

$$\partial_k \sigma = \sum_i (-1)^i [v_0, v_1, \ldots, \hat{v}_i, \ldots, v_k] \qquad (3)$$

where $\hat{v}_i$ indicates that v_i is deleted from the sequence. The boundary operator connects the chain groups into a chain complex $\mathbf{C}_*$: $\cdots \to \mathrm{C}_{k+1} \xrightarrow{\partial_{k+1}} \mathrm{C}_k \xrightarrow{\partial_k} \mathrm{C}_{k-1} \to \cdots$

We also define subgroups of $\mathbf{C}_k$ using the *boundary operator*, which is defined with the cycle group $\mathbf{Z}_k = \ker \partial_k$ and the boundary group $\mathbf{B}_k = im\, \partial_{k+1}$. We show examples of cycles in 3. An important property of the boundary operators is that the boundary of a boundary is always empty, $\partial_k \partial_{k+1} = \emptyset$. This fact, along with the definitions, implies that the defined subgroups are nested, $\mathbf{B}_k \subseteq \mathbf{Z}_k \subseteq \mathbf{C}_k$, as in Figure 4. For generality, we define *null boundary operators* in dimensions where $\mathbf{C}_k$ is empty.

2.4 Homology

The k^{th} *homology group* is $\mathbf{H}_k = \mathbf{Z}_k / \mathbf{B}_k$. Its elements are classes of homologous cycles. To describe its structure, the Abelian groups are viewed

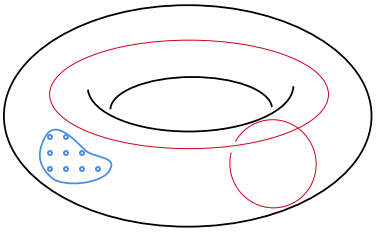


Figure 3: The dotted 1-boundary rests on the surface of a torus (in blue). The two solid 1-cycles (red lines) form a basis for the first homology class of the torus. Cycles are non-bounding, i.e. neither is a boundary of a piece of surface.

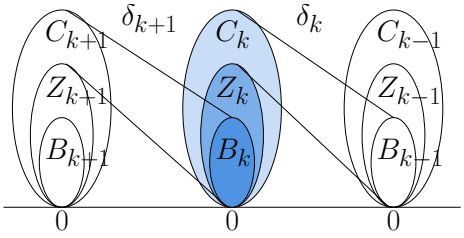


Figure 4: A chain complex with its components: chain, cycle, and boundary groups, and their images under the boundary operators.

as modules over the integers. This view allows alternate ground rings of coefficients, including fields. If the ring is a Principal Ideal Domain (PID) D, $\mathbf{H}_k$ is a D-module, then the rank of the module $\beta \in Z$ is the Betti number of the module, and $d_i \in D$ are its torsion coefficients. When the ground ring is Z, the homology group describes the structure of finitely generated Abelian groups. Over a field, such as R, Q, or Zp for p a prime, the torsion submodule disappears. The module is a vector space that is fully described by a single integer, its rank β, which depends on the chosen field.

2.5 Reduction

The standard method for computing homology is the reduction algorithm. We describe this method for integer coefficients as it is the more familiar ring. The method extends to modules over arbitrary PIDs, however.

As $\mathbf{C}_k$ is free, the oriented k-simplices form the standard basis for it. We represent the boundary operator $\partial_k : \mathbf{C}_k \to \mathbf{C}_{k-1}$ relative to the standard bases of the chain groups as an integer matrix M_k with entries in $\{-1, 0, 1\}$. The matrix M_k is called the standard matrix representation of ∂_k and we denote i^{th} row and j^{th} column with $M_k[i][j]$. It has m_k columns and m_{k-1} rows (the number of k- and $(k-1)$-simplices, respectively). The nullspace of M_k corresponds to $\mathbf{Z}_k$ and its range-space to $\mathbf{B}_{k-1}$, as manifested in Figure 4. The reduction algorithm derives alternate bases for the chain groups, relative to which the matrix for ∂_k is diagonal. The algorithm utilizes the following elementary row operations on M_k:

1. exhange row i and row j.
2. multiply row i by -1.
3. replace row i by $\forall k : M_k[i][k] + qM_k[j][k]$, where $q \in Z$ and $j \neq i$.

The algorithm also uses elementary column operations that are similarly defined. Each column (row) operation corresponds to a change in the basis for $\mathbf{C}_k(\mathbf{C}_{k-1})$. For example, if e_i and e_j are the i^{th} and j^{th} basis elements for $\mathbf{C}_k$, respectively, a column operation of type (3) amounts to replacing e_i with $e_i + qe_j$. A similar row operation on basis elements $\hat{e}_i$ and $\hat{e}_j$ for $\mathbf{C}_{k-1}$, however, replaces $\hat{e}_j$ by $\hat{e}_j \ - q\hat{e}_i$. We shall make use of this fact in Section 4. The algorithm systematically modifies the bases of $\mathbf{C}_k$ and $\mathbf{C}_{k-1}$ using elementary operations to reduce M_k to its *(Smith) normal form*:

$$\tilde{M}_k = \left[\begin{array}{ccc|c} b_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & b_{l_k} & \\ \hline & 0 & & 0 \end{array}\right], \qquad (4)$$

where $l_k = rank\, M_k = rank\, \tilde{M}_k$, $b^i \geq 1$, and $b_i | b_{i+1}$ for all $1 \leq i < l_k$. The algorithm can also compute corresponding bases $\{e_j\}$ and $\{\hat{e}_i\}$ for $\mathbf{C}_k$ and $\mathbf{C}_{k-1}$, respectively, although this is unnecessary if a decomposition is all that is needed. Computing the normal form in all dimensions, we get a full characterization of $\mathbf{H}_k$:

1. the torsion coefficients of $\mathbf{H}_{k-1}$ ($d_i \in D$) are precisely the diagonal entries b_i greater than one.
2. $\{e_i \,|\, l_k + 1 \leq i \leq m_k\}$ is a basis for $\mathbf{Z}_k$. Therefore, $rank\, \mathbf{Z}_k = m_k - l_k$.

3. $\{b_i \hat{e}_i \,|\, 1 \le i \le l_k\}$ is a basis for $\mathbf{B}_{k-1}$. Equivalently, $rank\, \mathbf{B}_k = rank\, M_{k+1} = l_{k+1}$.

Therefore, applying (2) and (3):

$$\beta_k = rank\, \mathbf{Z}_k - rank\, \mathbf{B}_k = m_k - l_k - l_{k+1} \quad (5)$$

2.5.1 Example

For the complex in Figure 1, the standard matrix representation of ∂_1 is

$$M_1 = \left[\begin{array}{c|ccccc} & ab & bc & cd & ad & ac \\ \hline a & -1 & 0 & 0 & -1 & -1 \\ b & 1 & -1 & 0 & 0 & 0 \\ c & 0 & 1 & -1 & 0 & 1 \\ d & 0 & 0 & 1 & 1 & 0 \end{array}\right], \quad (6)$$

where we show the bases within the matrix. Reducing the matrix, we get the normal form

$$\tilde{M}_1 = \left[\begin{array}{c|ccccc} & cd & bc & ab & z_1 & z_2 \\ \hline d-c & 1 & 0 & 0 & 0 & 0 \\ c-b & 0 & 1 & 0 & 0 & 0 \\ b-a & 0 & 0 & 1 & 0 & 0 \\ a & 0 & 0 & 0 & 0 & 0 \end{array}\right], \quad (7)$$

where $z_1 = ad - bc - cd - ab$ and $z_2 = ac - bc - ab$ form a basis for $\mathbf{Z}_1$ and $\{d-c, c-b, b-a\}$ is a basis for $\mathbf{B}_0$.

We can use a similar procedure to compute homology over PIDs. A homogeneous basis is a basis of homogeneous elements. We begin by representing ∂_k relative to the standard basis of $\mathbf{C}_k$ and a basis for $\mathbf{Z}_{k-1}$. Reducing to normal form, we read off the description provided by direct sum using the new basis $\{\hat{e}_j\}$ for $\mathbf{Z}_{k-1}$:

1. zero row i contributes a free term with shift $\alpha_i = deg\, \hat{e}_i$,

2. a row with diagonal term b_i contributes a torsional term with homogeneous $d_j = b_j$ and shift $\gamma_j = deg\, \hat{e}_j$.

The reduction algorithm requires $O(m^3)$ elementary operations, where m is the number of simplices in K. The operations must be performed in exact integer arithmetic.

2.6 Persistence

The algorithms to find the boundary points of a $\mathbf{H}_{k-1}$ homology group are solely based on the computing of *persistence*. Given a filtered complex, the i^{th} complex K^i has associated boundary operators ∂_k^i, matrices M_k^i, and groups $\mathbf{C}_k^i$, $\mathbf{Z}_k^i$, $\mathbf{B}_k^i$ and $\mathbf{H}_k^i$ for all $i, k \ge 0$[1]. The *p-persistent* k^{th} *homology group of* K^i is:

$$\mathrm{H}_k^{i,p} = \mathbf{Z}_k^i / \left(\mathrm{B}_k^{i+p} \cap \mathbf{Z}_k^i\right) \quad (8)$$

Here both groups in the denominator are subgroups of $\mathbf{C}_k^{l+p}$, so their intersection is also a group, a subgroup of the numerator. The p-persistent k^{th}*Betti number* of K^i is $\beta_k^{i,p}$ the rank of the free subgroup of $\mathbf{H}_k^{i,p}$. We can also define*persistent homology groups* using the injection $\eta_k^{i,p} : \mathbf{H}_k^i \to \mathbf{H}_k^{i+p}$, that maps a homology class into the one that contains it, i.e. $im\, \eta_k^{i,p} \simeq \mathbf{H}_k^{i,p}$ [3]. This definition can be extended over arbitrary PIDs as well.

3 The Persistence Module

In order to understand the structure of a persistent homology, we need to analyze it through a *persistence module*. Intuitively, the computation of persistence requires compatible bases for $\mathbf{H}_k^i$ and $\mathbf{H}_k^{i+p}$. It is not clear when a precise description is available for the compatible bases. In this section, we show how to combine the homology of all the complexes in the filtration into a single algebraic structure. We then construct a correspondence that reveals a simple description over fields. Most importantly, we show that the persistent homology of a filtered complex is the standard homology of a particular graded module over a polynomial ring. In order to explain this, we need couple of more definitions.

(persistence complex) A *persistence complex* $\mathcal{C}$ is a family of chain complexes $\{\mathbf{C}_*^i\}_{i \ge 0}$ over a commutative ring R with unity, equipped with chain map's $f^i : \mathrm{C}_*^i \to \mathrm{C}_*^{i+1}$, so that we have the following diagram: $\mathrm{C}_*^0 \xrightarrow{f^0} \mathrm{C}_*^1 \xrightarrow{f^1} \mathrm{C}_*^2 \xrightarrow{f^2} \cdots$

The *filtered complex* K with inclusion maps for the simplices becomes a persistence complex. We

[1]Note that superscripts indicate the filtration index and are not related to cohomology.

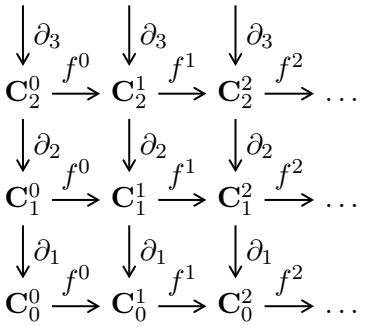


Figure 5: a portion of a persistence complex, with the chain complexes expanded.

show a portion of a persistence complex, with the chain complexes expanded in Figure 5. The filtration index increases horizontally to the right under the chain maps f_i, and the dimension decreases vertically to the bottom under the boundary operators ∂_k.

(persistence module) A persistence module $\mathcal{M}$ is a family of R-modules M^i, together with homomorphisms $\phi^i : M^i \to M^{i+1}$.

Intuitively, we are building a single structure that contains all the complexes in the filtration. We begin by computing a direct sum of the complexes, arriving at a much larger space that is graded according to the filtration ordering. We then remember the time each simplex enters using a polynomial coefficient. For instance, simplex a enters the filtration in Figure 1 at time 0. To shift this simplex along the grading[2], we must multiply the simplex using t. Therefore, while a exists at time 0, $t \cdot a$ exists at time 1, $t^2 \cdot a$ at time 2, and so on. The key idea is that the filtration ordering is encoded in the coefficient polynomial ring.

4 Level Sets & Persistence Diagrams

So far we have developed a fairly comprehensive idea on the computation of simplicial homology of a point cloud in dimension k. However, the goal of this paper is to find the boundary points of a $\mathbf{H}_{k-1}$ homology groups. In order to find those points, we need to understand how the simplicial homology is represented in terms of *level sets* and *persistence diagrams*.

(Level Sets) Given a piecewise-linear function $f : |K| \to R$ and a value $\lambda \in R$, we define the upper[3] λ-level set of f as

$$L_\lambda = \{x : f(x) \geq \lambda\} = f^{-1}([\lambda, \infty)) \qquad (9)$$

For any two level thresholds $\lambda_1 > \lambda_2$, the corresponding level sets satisfy $L_{\lambda_1} \subset L_{\lambda_2}$. Thus, the collection of level sets $\mathcal{L} = \bigcup_\lambda \{L_\lambda\}$ forms a filtration with the level as the index set.

4.1 Persistence Diagrams

For each level set L_λ, the corresponding topological features are captured by the generators of its homology groups. In a simpler term, the 0^{th} order homology groups (0^{th} order topological feature) capture the connected components (i.e. chain), the 1^{st} order homology groups capture regions forming a cycle, and 2^{nd} order homology groups capture regions forming a void structure (i.e. boundary of a 2-simplex), etc.

If we decrease the level λ, new generators for the homology groups may be created (e.g., the formation of new components), existing generators may merge together (e.g., two connected components joining together), and existing generators may be eliminated (e.g., a loop getting filled in). The level at which a generator is created is called its *birth time* and the level at which a generator is destroyed (or merges with another generator that has an earlier birth time) is called its *death time*. Therefore, every generator in $\mathcal{L}$ encodes three features – homology order, birth time, and death time. The persistence diagram is a collection of all these tuples found from of filtration formed by a given f. Thus, if a function's filtration has $|T|$ generators, the corresponding persistence diagram is

$$\mathsf{Dgm} = \left\{ \left(r_j, t_{b_j}, t_{d_j} \right) : j = 1, \cdots, |T| \right\} \qquad (10)$$

where r_j, t_{b_j} and t_{d_j} are the homology order, birth time, and death time of the j^{th} generator, respectively. The norm $|T|$ specifies the number of off-diagonal elements in the persistence diagram Dgm.

[2]The ring R is graded.

[3]Some literatures consider the lower level set as $f^{-1}((-\infty, \lambda))$.

5 Finding The $\mathbf{H}_{k-1}$ Homology Boundary

At this point, we have successfully set up the entire scene for the basics of TDA concepts. In this section, we are going to describe the procedure that has been used to compute the homology and detection of boundary points.

5.1 Basic Idea

The first step of the algorithm is to construct the persistence diagram from $\mathcal{X}$. Which can be done by solving the Smith-normal form found from the Vietoris-Rips complex of the data points. Once this is done, we will also have the r_j, t_{b_j} and t_{d_j} values of the persistence diagram Dgm for all homologies up to $\mathbf{H}_{k-1}$. To find the chain complex of $\mathbf{H}_{k-1}$, we locate the highest off-diagonal value in the Dgm for $\mathbf{H}_{k-1}$. Which basically will give us the death time for the last $k-1$-dimensional hole that has been eliminated. The second highest off-diagonal value in Dgm is the birth time for the same complex. Therefore, we collect all the representative cocyles that are generated at the mid-point of highest and the second highest off-diagonal values in Dgm. We then use that value as threshold to find all the representative cocyles. The cocycle is a collection of 1-simplices that, if removed, would break the cycle in $\mathcal{X}$. Now in order to find all the points that bounds all the cocylces, we go through each of them and apply *union find* algorithm to construct a chain of vertices that corresponds to the boundary.

5.2 Algorithmic Visualization Steps

One way is to compute the persistent homology directly, and the other is involuted homology computation. Involuted homology computes cohomology first and then uses its result to recompute cycles. Cohomology is the dual of homology, and that a coboundary operator is the adjoint operator of the boundary operator, similar concept is true for cochain. In our case, we collected the cocycles from the tortion components from the matrix $\tilde{M}_k$. The reconstruction of the shortest cycle from the set of cocycle is done using the *union find* algorithm. Once the boundary vertices are found, we store their indices to be used in the PaletteViz for the final visualization.

6 Visualization Results

In order to present the $\mathbf{H}_{k-1}$ homology boundary vertices, we utilize a novel visualization method called *PaletteViz* [8]. The PaletteViz algorithm employs a monotonically increasing function $f(\cdot)$ acting on the k-dimensional data point cloud $\mathcal{X} \subset R^k$. The function is computed in terms of so called *depth contour*. A depth contour is computed with respect to the *inward-outward* relationship of the points in $\mathcal{X}$ in terms their distance from its *core*[4]. The farthest points from the core are the *boundary of* $\mathcal{X}$. PaletteViz works by finding the boundary points using a non-convex hull algorithm (i.e. α-shape) and successively separating them out from $\mathcal{X}$. At every separation step, the points are mapped on a 2-dimensional plane using RadViz or Star-Coordinate plot. Thus, the PaletteViz algorithm can capture the global topology of $\mathcal{X}$ in terms of its level sets $\mathcal{L}$. In the case of PaletteViz, the $\mathcal{L}$ is computed by finding the α hull. The set of points $\mathcal{X}_0$ found from the first pass of α-hull have depth-contour of 0, then they are removed from $\mathcal{X}$ (i.e. $\mathcal{X}' = \mathcal{X} - \mathcal{X}_0$). Next, the set of points $\mathcal{X}_1$ found from the second pass of α-hull have depth-contour of 1, then they are removed from $\mathcal{X}'$ and so on. For more details of this algorithm, readers are referred to [8].

In our case, once we identify the boundary vertices, we color them with a more conspicuous shade (i.e. red) and keep the rest of the data points in a uniform color (i.e. grey). Instead of using the color coding described in the PaletteViz, we adopt this approach since we are not interested in visualization of their distances from the mean or their feasibility in terms of distance from a constraint function. For this experiment, we apply our method on the Pareto-optimal front found by solving a scalable DTLZ2 problem [2] (in 3, 4 and 5 dimensions). In our case, we apply a constraint function to create a hole in the middle of the Pareto-optimal surface (i.e. hyper-surface):

The results for 3-dimensional DTLZ2 problem is presented in Figure 6. Although, PaletteViz does

[4]The authors use centroid/mean of $\mathcal{X}$ as *core*.

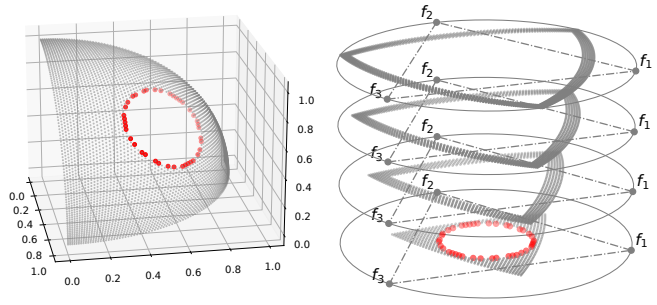


Figure 6: (a) A 3-dimensional DTLZ2 problem with a constraint function causing infeasible region in the Pareto-optimal surface. (b) PaletteViz visualization of the Pareto-front. The points surrounding the infeasible region appear on the bottom layer. In this case, infeasible region is a homology group of $\mathbf{H}_2$.

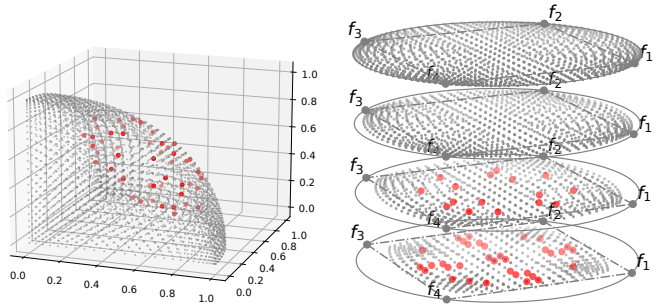


Figure 7: (a) A 4-dimensional DTLZ2 problem with a constraint function causing infeasible region in the Pareto-optimal surface. The scatter plot is shown on the first three objectives. (b) PaletteViz visualization of the Pareto-front. The points surrounding the infeasible region appear on multiple layers. In this case, infeasible region is a homology group of $\mathbf{H}_3$.

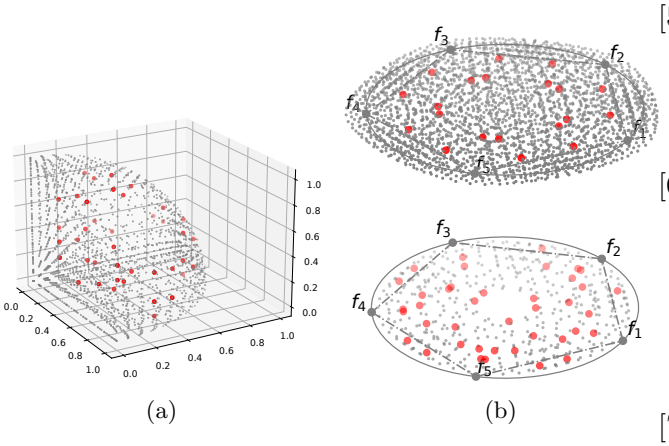
not give any special insight of $\mathcal{X}$ in 3-dimension, we include this example since it will be helpful to understand other results. In Figure 6b, we can see how the 2-dimensional hole (points in red) appears in the bottom layer of PalettViz plot. The red points in the plot represents the Pareto-optimal solutions that are very close to the constraint function.

Next, we apply our method on a 4-dimensional DTLZ2 problem, which is presented in Figure 7. Figure 7a is the Pareto-optimal front in first three objective. Here we can see that points surrounding a $\mathbf{H}_3$ are highlighted in red. The corresponding points are depicted on the PaletteViz plot in Figure 7b. Since $\mathcal{X}$ is in 4 dimensional space, the hole created by the infeasible region is bounded by a chain of 3-dimensional simplices. As a result, the boundary points are found on multiple layers of PaletteViz.

In the end, we apply our method on a 5-dimensional DTLZ2 problem, which is shown in Figure 8. Similarly, the points surrounding a $\mathbf{H}_4$ are highlighted in red. The corresponding points are depicted on the PaletteViz plot in Figure 8b. Since in this case, $\mathcal{X}$ is in 5 dimensional space, the hole created by the infeasible region is bounded by a chain of 4-dimensional simplices. As a result, the boundary points are found on all layers of PaletteViz. Due to the high-dimensionality of the underlying space, it is very difficult to find enough points in DTLZ2 surface in *general positions*. As a result, the triangulation algorithm does not find enough simplices that are not coplanar. Consequently, the α-hull algorithm produces only two level sets. However, still we can see the boundary points that surround the infeasible region.

7 Conclusions and Future Work

The procedure described in this paper to find the (co)homology boundary is based on Vietoris-Rips complex. However, finding this complex is computationally very expensive, especially for a higher dimensional space. Moreover, in this particular application, we actually might not need to compute this complex since a Cech or α complex should be good enough to bound the $\mathbf{H}_{k-1}$ homology groups. This is one direction of improvement that might be worth pursuing. Another way to improve the running time for Vietoris-Rips computation is to apply an *approximate sparse filtration* [1], to approximate the highest off-diagonal value in Dgm. This tool can be very useful to the MCDM community due the reasons explained earlier in the paper. Also our proposed method shows that topological information can be represented vis-a-vis to a user utilizing



(a) (b)

Figure 8: (a) A 5-dimensional DTLZ2 problem with a constraint function causing infeasible region in the Pareto-optimal surface. The scatter plot is shown on the first three objectives. (b) PaletteViz visualization of the Pareto-front. The points surrounding the infeasible region appear on both layers. In this case, infeasible region is homology group of $\mathbf{H}_4$.

other methods like PaletteViz – which is an interesting and a useful application to both MCDM and TDA practitioners.

References

[1] Nicholas J. Cavanna, Mahmoodreza Jahanseir, and Donald R. Sheehy. A geometric perspective on sparse filtrations. In *Proceedings of the Canadian Conference on Computational Geometry*, 2015.

[2] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler. Scalable multi-objective optimization test problems. In *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02*, pages 825–830, 2002.

[3] Edelsbrunner, Letscher, and Zomorodian. Topological persistence and simplification. *Discrete & Computational Geometry*, 28(4):511–533, Nov 2002.

[4] Z. He and G. G. Yen. Visualization and Performance Metric in Many-Objective Optimization. *IEEE Transactions on Evolutionary Computation*, 20(3):386–402, June 2016.

[5] R. A. Jarvis. On the identification of the convex hull of a finite set of points in the plane. *Information Processing Letters*, 2(1):18–21, March 1973.

[6] Eser Kandogan. Visualizing Multi-dimensional Clusters, Trends, and Outliers Using Star Coordinates. In *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 107–116, New York, NY, 2001. ACM.

[7] J.R. Munkres. *Elements Of Algebraic Topology*. Avalon Publishing, 1996.

[8] A. K. A. Talukder and K. Deb. Paletteviz: A visualization method for functional understanding of high-dimensional Pareto-optimal data-sets to aid multi-criteria decision making. *IEEE Computational Intelligence Magazine*, 15(2):36–48, 2020.

[9] T. Tušar and B. Filipič. Visualization of Pareto Front Approximations in Evolutionary Multiobjective Optimization: A Critical Review and the Prosection Method. *IEEE Transactions on Evolutionary Computation*, 19(2):225–245, April 2015.