# ON (ABSOLUTE) IRREDUCIBILITY CRITERIA FOR SOME POLYNOMIAL STRUCTURES VIA ARITHMETICAL PROPERTIES OF PERSISTENCE DIAGRAMS OF NEWTON POLYTOPES

DANNY A. J. GÓMEZ-RAMÍREZ, EDISSON GALLEGO-GONZALEZ, JAIBERTH PORRAS-BARRERA, DIEGO A. LOPEZ-CARDONA, AND SONIA A. AGUDELO

ABSTRACT. We establish basic results for an initially surprising connection between seminal notions of persistent homology, in the context of topological data analysis, and (absolutely) irreducibility criteria for polynomials in several variables over a field. Specifically, we prove (absolutely) irreducibility criteria for several sorts of polynomial in several variables in terms of arithmetical properties of the zero-dimensional persistence indexes of the corresponding persistence's skeletons of the polynomials based on a work of G. Gao. Additionally, we present a partially self-contained introduction to the classic and new notions of persistent homology and irreducibility criteria needed in the paper together with enlightening examples.

## INTRODUCTION

At the beginning of the third decade of the twenty-first century we, as humanity, produce very large collections of data sets with a titanic complexity and at a unprecedented rate. The excessive exposition and daily dependency on digital devices for the fulfillment of some of our most basic activities facilitate enormously that a lot of data if collected, stored and subsequently analysed, for example, in the form of cookies, etc. These clouds of millions of terabytes of information not only about us, but also regarding the industrial, economic, political, logistic, sociological, geological and cosmological dimensions of our universe contain and encode highly valuable secrets and relevant key features about the structure of the behaviour of us and of our cosmos.

Thus, it turns out that these seminal secrets of nature encoded within the data sets are gradually more challenging to be revealed due to the growth of the intrinsic complexity and size of them. Therefore, more general, finer, and deeper mathematical models and theories are required to analyse soundly

---

2020 *Mathematics Subject Classification.* 13B25, 13P05, 55N31.
*Key words and phrases.* Persistence's diagrams, Newton polytopes, absolute irreducibility, polynomial rings, topological data analysis.

such colossal clouds of information. In particular, structurally new modeling methodologies are required which should go far beyond the classic linear (polynomial), analytic and numerical methods [10], [12], [8], [9].

In this context, a relatively new mathematical modeling collection of techniques have emerged over time possessing a topological nature in contrast with the analytic and numerical one of the former approaches. These gathering of techniques are known as Topological Data Analysis (TDA) [1]. In a broader sense, TDA encompasses a group of methodologies of data analysis such as persistence homology, mode estimation, nonlinear dimension reduction, and clustering, among others [18]. Nonetheless, in a more concrete perspective, it refers directly to *persistence homology* [3], which is essentially an algebraic-homological method for measuring topological attributes of shapes such as connectedness and geometrical forms emerging from the peculiar local and global distribution of distances among (large) data sets, being configured and understood, for instance, as points in a $n-$dimensional space [14].

As a whole, TDA has a wide spectrum of applications in several academic and applied disciplines such as oncology [2], chemical engineering [17], biomedicine [16], aviation [11], spatial networks [4], among many others [1].

One of the main purposes of this work is to highlight and to describe precisely a new formal surprising connection of some of the main concepts and features of persistence homology with a sort of mathematical issues situated at the intersection of commutative algebra and number theory. More precisely, we establish new criteria for irreducibility (primality) of certain families of polynomials in several variables over an arbitrary coefficient's field in terms of some persistence properties of the corresponding cloud of integral points within the $n-$dimensional real space emerging from the multi-exponent of their non-zero monomials.

## 1. Preliminary Notions and Methodology

In this section, we introduce the central objects and results of study to establish our main results. Roughly speaking, they are divided into two classes: the first one including classic and new mathematical structures belonging to persistent homology and the second one regarding notions and theorems within the realm of (combinatorial) commutative algebra and number theory.

### 1.1. **Persistent Homology.**

In this chapter we introduce the concept of *persistence homology*, which plays a fundamental rol in the applications of TDA. To do this, we must first define what a *simplicial complex* is.

Let $X = \{x_i\}_{i\in\mathbb{N}_{\leqslant m}}$ be a finite cloud of points within the $n-$dimensional real space $\mathbb{R}^n$. A simplicial complex $\mathbb{K} \subset \mathcal{P}(X)$ over $X$ is a non-empty subset of parts of $X$ that satisfies:

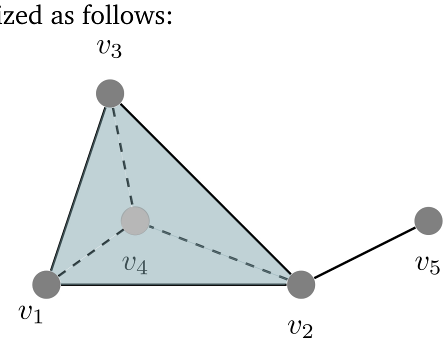
- $x_i \in \mathbb{K}$ for $i = 0, 1, 2, \ldots, m$.
- If $\sigma \in \mathbb{K}$ and $l \subset \sigma$, then $l \in \mathbb{K}$.

The elements of $\mathbb{K}$ are called *simplices* or *faces* of $\mathbb{K}$. If $\sigma = \{x_{i_0}, x_{i_1}, \ldots, x_{i_t}\}$ belongs to $\mathbb{K}$, we say that $\sigma$ is a $t-$dimensional face of $\mathbb{K}$ or that $\sigma$ is a $t-$simplex of $\mathbb{K}$. The $0-$simplices are called *vertices* of $\mathbb{K}$, and the $1-$simplices are called *edges*. We say that $\mathbb{K}$ is an $m-$dimensional simplicial complex if it contains at least one $m-$simplex.

**Example 1.1.** Let $X = \{v_1 = (1,3), v_2 = (2,4), v_3 = (7,6), v_4 = (2,-4), v_5 = (4,7)\}$. In the following definition of a simplicial complex $\mathbb{K}$ of $X$, the vertices appear in the first row, the edges in the second row, the $2-$simplices in the third row, and the $3-$simplex in the fourth row.

$$\mathbb{K} = \left\{ \begin{array}{l} \{v_1\}, \{v_2\}, \{v_3\}, \{v_4\}, \{v_5\}, \\ l_1 = \{v_1, v_2\}, l_2 = \{v_1, v_3\}, l_3 = \{v_1, v_4\}, l_4 = \{v_2, v_4\}, l_5 = \{v_2, v_5\}, \\ \sigma_1 = \{v_1, v_2, v_3\}, \sigma_2 = \{v_1, v_2, v_4\}, \sigma_3 = \{v_1, v_4, v_3\}, \sigma_4 = \{v_2, v_4, v_3\}, \\ \eta = \{v_1, v_2, v_3, v_4\} \end{array} \right\}$$

Since the simplicial complex $\mathbb{K}$ has dimension 3, we can associate it with a topological space in $\mathbb{R}^3$, which we will refer to as the *geometric realization* of $\mathbb{K}$. It can be visualized as follows:




Here, each vertex $v_i$ is identified with a point (in this example, in $\mathbb{R}^3$), usually labeled with the same name $v_i$ as the vertex. Each edge $l_i = \{v_r, v_s\}$ is identified with the line segment that goes from $v_r$ to $v_s$. For example, the edge $l_1$ is identified with the line segment that goes from $v_1$ to $v_2$. Each $2-$simplex $\sigma$ is identified with the filled triangle whose vertices are the vertices in $\sigma$. In general, any $n-$simplex $\sigma = \{v_0, v_1, \ldots, v_n\}$ can be identified with the set of points $\overline{X} = \sum_{i=0}^{n} \lambda_i v_i$ where each $\lambda_i \geqslant 0$ and $\sum_{i=0}^{n} \lambda_i = 1$.

**Definition 1.2.** Given a finite set of points $X = \{x_0, x_1, \ldots, x_n\}$ in a metric space $(M, d)$ and a positive real number $r$, we define the *Vietoris-Rips* complex of $X$ (or simply *Rips* of $X$) with parameter $r$ as the simplicial complex denoted by $\mathbb{K}_r(X)$ that satisfies:

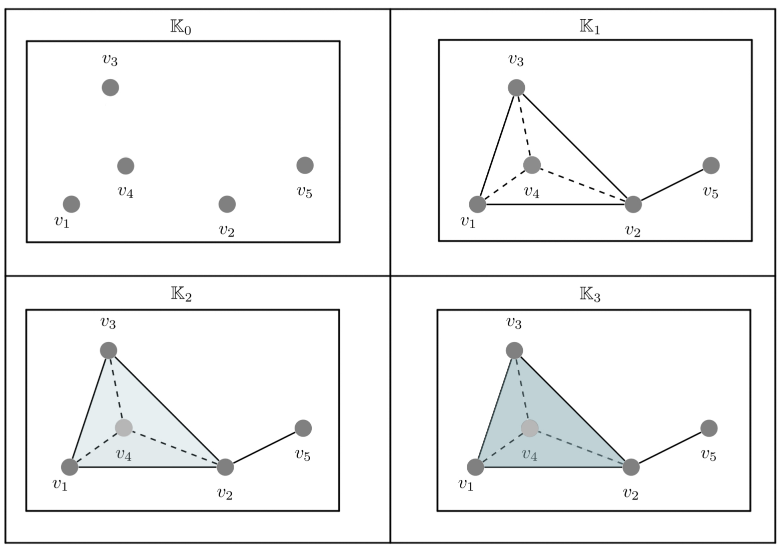
- The vertices of $\mathbb{K}_r(X)$ are the points of $X$.
- $\sigma = \{y_0, y_1, \ldots, y_t\} \in \mathbb{K}_r(X)$ if and only if $d(y_i, y_j) < r$ for all $i, j = 0, 1, 2, \ldots, t$.

In the subsequent examples, $X$ will be a finite set of points in $\mathbb{R}^d$. However, it may occur that $dim(\mathbb{K}_r(X)) > d$ for some values of $r$. Therefore, the geometric realization of $\mathbb{K}_r(X)$ may not be a topological subspace of $\mathbb{R}^d$ (in example (1.1), the point cloud $X$ is a subset of $\mathbb{R}^2$, but $\mathbb{K} \nsubseteq \mathbb{R}^2$).

Note that if $r < r'$, then $\mathbb{K}_r(X) \subset \mathbb{K}_{r'}(X)$. This inclusion induces a homomorphism between the $p-$th simplicial homologies $H_p(\mathbb{K}_r(X)) \to H_p(\mathbb{K}_{r'}(X))$, for all $p$; see, for instance [3].

**Definition 1.3.** A collection of simplicial subcomplexes $\{\mathbb{K}_r\}_{r\in J}$ of $\mathbb{K}$ (for some $J \subseteq \mathbb{R}$) such that $\mathbb{K}_r \subset \mathbb{K}_{r'}$, when $r < r'$, is called a *filtration* of $\mathbb{K}$ if $\bigcup \mathbb{K}_r = \mathbb{K}$.

**Example 1.4.** Below, we present a filtration of the simplicial complex $\mathbb{K}$ given in example (1.1). For simplicity, we only present the graphical representation of each subcomplex. $\mathbb{K}_0$ consists of all the $0$-simplices (vertices) of $\mathbb{K}$; $\mathbb{K}_1$ includes all the vertices and edges of $\mathbb{K}$; $\mathbb{K}_2$ is composed of all simplices of dimension less than or equal to $2$ of $\mathbb{K}$. Finally, $\mathbb{K}_3 = \mathbb{K}$ also contains the interior of the tetrahedron with vertices $v_1, v_2, v_3$ and $v_4$. Thus, $\mathbb{K}_0 \subset \mathbb{K}_1 \subset \mathbb{K}_2 \subset \mathbb{K}_3$.




Given a filtration of simplicial complexes

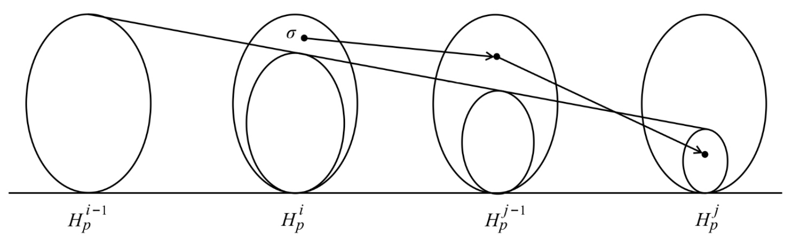
$$\mathbb{K}_0 \subset \mathbb{K}_1 \subset \cdots \subset \mathbb{K}_{n-1} \subset \mathbb{K}_n$$

and an integer $p > 0$, we can induce a sequence of linear maps in the homologies as follows:

$$H_p(\mathbb{K}_0) \xrightarrow{f_p^0} H_p(\mathbb{K}_1) \xrightarrow{f_p^1} \cdots \xrightarrow{f_p^{n-2}} H_p(\mathbb{K}_{n-1}) \xrightarrow{f_p^{n-1}} H_p(\mathbb{K}_n).$$

For $i < j$, let us denote $f_p^{ij} = f_p^j \circ \cdots \circ f_p^i$. Then, we define $H_p^{ij} = \dfrac{Z_p(\mathbb{K}_i)}{B_p(\mathbb{K}_j) \cap Z_p(\mathbb{K}_i)}$, where $Z_p(\mathbb{K}_i)$ represents the $p-$cycles of $\mathbb{K}_i$ and $B_p(\mathbb{K}_j)$ the $p-$boundaries of $\mathbb{K}_j$.

Let $\sigma \in H_p(\mathbb{K}_i)$ be an equivalence class of a $p-$simplex. We say that $\sigma$ is born in $\mathbb{K}_i$ if $\sigma \notin H_p^{i-1,i}$. Now, suppose that the simplex $\sigma$ is born in $\mathbb{K}_i$. We say that $\sigma$ dies in $\mathbb{K}_j$, if this simplex merges with an older class when going from $\mathbb{K}_{j-1}$ to $\mathbb{K}_j$, that is, if $f_p^{i,j-1}(\sigma) \notin H_p^{i-1,j-1}$ but $f_p^{i,j}(\sigma) \in H_p^{i-1,j}$. The following diagram schematically illustrates the notions of birth and death of an equivalence class:




Let $\mathbb{X} = \{x_l\}$ be a finite set of points in $\mathbb{R}^n$ and $\{\mathbb{K}_{r_i}(\mathbb{X}) = \mathbb{K}_{r_i}\}$ a filtration with Vietoris-Rips complexes of the convex hull of $\mathbb{X}$, ordered such that $\mathbb{K}_{r_i} \subset \mathbb{K}_{r_j}$ if $r_i < r_j$. Consider $\sigma \in H_p(\mathbb{K}_{r_i})$, a $p-$simplex that is born in $\mathbb{K}_{r_i}$ and dies in $\mathbb{K}_{r_j}$; we then associate to $\sigma$ the ordered pair $(r_i, r_j)$, which indicates its birth and death.

*Barcode.* The *barcode* of $\mathbb{X}$ in dimension $p$ is the set of horizontal lines that begin at $r_i$ and end at $r_j$, with one line for each $p-$simplex that is born at some $r_i$ and dies at some $r_j$.

*Persistence Diagram.* The *persistence diagram* is another way of representing the information gathered from the births and deaths of the $p-$simplices that appear and disappear as the filtration $\mathbb{K}_{r_i}$ progresses. The persistence diagram in dimension $p$ associated with $\mathbb{X}$ is the set of ordered pairs

$$D_p = \{(r_i, r_j) |\ \exists \sigma \text{ such that } \sigma \text{ is born at } r_i \text{ and dies at } r_j\}.$$

Additionally, for each ordered pair $(a, b) \in D_p$, we can define the function

$$f_{(a,b)}(x) = \begin{cases} x - a & \text{if } a < x \leqslant \frac{a+b}{2}, \\ b - x & \text{if } \frac{a+b}{2} < x < b, \\ 0 & \text{otherwise.} \end{cases}$$

This family of functions $f_{(a,b)}(x)$ allows us to introduce the following concept.

*Persistence Landscape.* For each $k \in \mathbb{N}$, we define the function $f_k(x)$ as the $k-$th largest value of the set $\{f_{(a,b)}(x)\}$, where $(a, b)$ varies over all elements of $D_p$. The function $f : \mathbb{N} \times \mathbb{R} \to \mathbb{R}$ given by $f(k, x) = f_k(x)$ is called the *persistence landscape* in dimension $p$ associated with $\mathbb{X}$.

**Example 1.5.** Consider the following point cloud $\mathbb{X}$ in $\mathbb{R}^3$:

$$\mathbb{X} = \{(0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1), (0, 1, 1)\}.$$

For this point set $\mathbb{X}$, we have the barcode, persistence diagram, and persistence landscape shown below. As is customary, the barcodes are presented in a single graph (see Table 1), assigning a different color to each dimension (in this example, black bars for dimension zero and red bars for dimension one). Similarly, the persistence diagram is presented with the same color scheme.

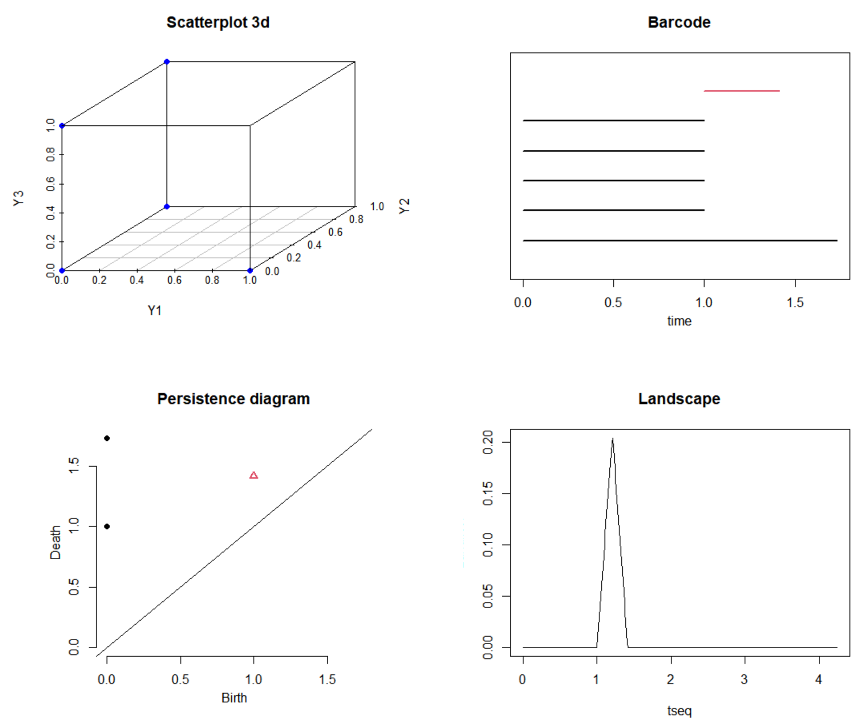



TABLE 1

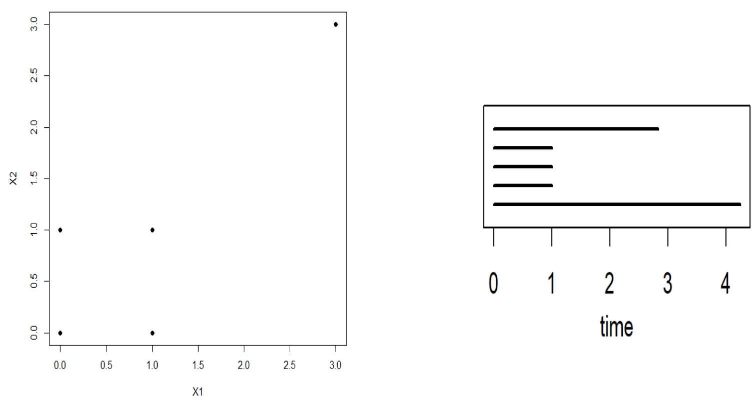
**Definition 1.6.** Let $\mathbb{X} = \{a_i\}_{i \in \mathbb{N}_{\leqslant m}}$ be a finite collection of $m$ points within the $n-$dimensional real space. Let $P_0(\mathbb{X})$ be the zero persistence barcode of $\mathbb{X}$. Then, the *first (Vietoris-Rips) persistence index of* $\mathbb{X}$ *of dimension zero* $pind_1(\mathbb{X})$

is the biggest horizontal finite length of $P_0(\mathbb{X})$, with *(Vietoris-Rips) multiplicity* $mind_1(\mathbb{X})$ equal to the number of times that there are lines in $P_0(\mathbb{X})$ of exactly this length.[1] If $mind_1(\mathbb{X})) = m-1$, then the next indices and multiplicities of persistence are defined as zero. Otherwise, *the second persistence index of* $\mathbb{X}$ *of dimension zero,* $pind_2(\mathbb{X})$, is the second biggest horizontal finite length appearing in $P_0(\mathbb{X})$, with its similar corresponding multiplicity $mind_2(\mathbb{X})$. If $mind_1(\mathbb{X}) + mind_2(\mathbb{X}) = m-1$, then the next indices and multiplicities of persistence are defined as zero, and so on.

Note that the persistence's indices and multiplicities are defined until the $(m-1)$th step. Moreover, from the former definition one can immediately deduce the following formula:

$$\sum_{i=1} mind_i(\mathbb{X}) = m-1.$$

**Example 1.7.** Let's consider the polynomial $q(x_1, x_2) = 1 + x_1 + x_2 + x_1x_2 + x_1^3x_2^3$. We can pick the exponents of its terms to obtain the collection $\mathbb{X} = \{(0,0),(1,0),(0,1),(1,1),(3,3)\}$ of points in $\mathbb{R}^2$. In the following figure it is drawn the cloud of points $\mathbb{X}$ and its persistence barcode in dimension $0$:




(A) Geometric representation of the finite cloud of points emerging from the exponents of the polynomial $q(x_1, x_2)$.

(B) Representation of the persistence's indexes of the corresponding finite cloud of points via its barcode.

FIGURE 1

[1] In the sequel, we will omit, for simplicity, the Vietoris-Rips name for the indexes as well as for the multiplicities.

Here, we can see that its first persistent index $pind_1(\mathbb{X})$ is approximately 2.8, $mind_1(\mathbb{X}) = 1; pind_2(\mathbb{X}) = 1$, and $mind_2(\mathbb{X}) = 3$, finally

$$\sum_i mind_i(\mathbb{X}) = 5 - 1 = 4.$$

1.2. **Absolute Irreducibility of Polynomials, Newton Polytopes and Persistence's Skeletons.** One question of tantamount importance in polynomial and commutative algebra naturally derived from classic arithmetic is primality criteria for polynomials. In other words, to find sufficient and necessary conditions for a polynomial (or similar entities such as formal power series or rational functions) to be primes. Now, it is worth to note that from elementary abstract algebra we know that in the context of polynomial rings in several variables over a unique factorization domain being prime is equivalent to being irreducible, since UFD-s remains so by adding finitely many variables [5].

This specific field of research possesses on its own a huge amount of results and techniques which at first sight seem to be kind of overwhelming and surprising (see, for example, [15]).

Here, our center of attention will be focus on irreducibility criteria based on properties of the so called *Newton polytope of a polynomial*. One of the cornerstones for our results will be the work of Shuhong Gao as presented in [6].

For the sake of completeness and for a more easier understanding from the reader's perspective, now we will present explicitly some definitions and results needed in the next section.

Let $R := k[x_1, \cdots, x_n]$ denote the ring of polynomials in several variables over a arbitrary field $k$. For each polynomial $f \in R$ we define the *Newton polytope of* $f$, denoted by $P_f$, as the convex hull in $\mathbb{R}^n$ of the finite collection of integral $n-$dimensional points formed by the exponents of the non-zero monomials appearing in $f$. More generally, if $N = \{c_1, \cdots, c_s\} \subseteq \mathbb{R}^n$ a finite collection of points, then the *convex hull of* $N$ is defined as

$$conv(N) = conv(c_1, \cdots, c_s) := \left\{ \sum_{i=1}^{s} z_i c_i : z_i \geqslant 0, \sum_{i=1}^{s} z_i = 1 \right\},$$

and is called a *polytope*. Additionally, a point $a_j$ is called a *vertex of* $conv(N)$ if it not lie strictly inside the line segment of any other two points of the polytope. Clearly, any non-empty polytope has a non-empty collection of vertices. A polytope is *integer* if all the coordinates of each of its vertices are integer numbers. The greatest common divisor of the points $c_1, \cdots, c_s$, denoted as $gcd(c_1, \cdots, c_s)$, is simply the greatest common divisor of all the entries of each $c_j$, together, for $j = 1, \cdots, s$.

For example, the integer Newton polytope of $f = x_1^2 + x_2^3 \in k[x_1, x_2]$ is

$$P_f = \{(a,b) = w_1(2,0) + w_2(0,3) \in \mathbb{R}^2 : w_1, w_2 \geqslant 0, w_1 + w_2 = 1\},$$

and the vertices of $P_f$ are $(2,0)$ and $(0,3)$.
If $A, B \subseteq \mathbb{R}^s$, then the *Minkowsky sum of* $A$ *and* $B$ is defined as usual

$$A + B = \{a + b \in \mathbb{R}^n : a \in A \text{ and } b \in B\}.$$

An integral polytope $N$ is called *integrally indecomposable* (or simply *indecomposable* in our presentation) if it cannot be written as $A + B$, where $A$ and $B$ are integral polytopes with at least two points each one.

On the other hand, a polynomial $f \in R$ is called *absolutely irreducible* if it is irreducible over any algebraic extension of the coefficient field $k$, in other words, if $f$ is irreducible in $R_L = L[x_1, \cdots, x_n]$, where $k \subseteq L$ is any algebraic field extension of the original coefficients' field.

The following proposition is the classic key element connecting absolute irreducibility of polynomials and indecomposability of their corresponding Newton polypotes and was firstly discovered by A. M. Ostrowski in [13]. We state explicitly the main result here in similar terms of the presentation given in [6].

**Proposition 1.8.** *(Irreducibility Criterion) Let $f \in R\backslash\{0\}$ be a polynomial non-divisible by any of the variables $x_j$. If the Newton polytope $P_f$ is (integrally) indecomposable, then $f$ is absolutely irreducible.*

The main ingredient of the proof of this criterion is the fact that if $f, g, h \in R$ are polynomials such that $h = fg$, then $P_h = P_f + P_g$ [6, Lemma 2.1].

For the sake of completeness in our presentation, we state explicitly the following results, which play a key role for the proofs given in the next section, they correspond to Theorem 4.2 and corollaries 4.3, 4.5 a 4.7 of [6], respectively.

**Theorem 1.9.** *Suppose that the polytope $conv(v_1, v_2, \cdots, v_m) \subseteq \mathbb{R}^n$ lies in a hyperplane $H$. Let $v \in \mathbb{R}^n$ not belonging to $H$. Then the composed polytope $conv(v_1, \cdots, v_m, v)$ is integrally indecomposable if and only if*

$$gcd(v - v_1, \cdots, v - v_m) = 1.$$

**Proposition 1.10.** *Let $a$ and $b$ be two (distinct) integral points in $\mathbb{R}^n$, then the line segment $ab := conv(a,b)$ is integrally indecomposable if and only if $gcd(a - b) = 1$.*

**Proposition 1.11.** *Let $a, b, c \in \mathbb{R}^n$ be three (distinct) integral non-collinear points. Then, the triangle $conv(a,b,c)$ is integrally indecomposable if and only if $gcd(a - b, a - c) = 1$.*

**Proposition 1.12.** *Let $a, b, c, d \in \mathbb{R}^n$ be four (distinct) integral noncoplanar points. The tetrahedron $conv(a, b, c, d)$ is integrally indecomposable if and only if $gcd(a-b, a-c, a-d) = 1$.*

The next notion is new in our context and explore the external and essential points of the Newton polytope of a polynomial regarding the kind of irreducibility criteria developed in [6].

**Definition 1.13.** Let $f \in R\backslash\{0\}$. The *persistence's skeleton of $f$*, denoted as $Psk(f)$, is the collection of vertices of the Newton polytope of $f$.

In other words, the persistence's skeleton of a non-zero polynomial $f$ corresponds exactly to the multi-indices of it that structures the whole boundary and interior of $P_f$, omitting any multi-index corresponding to a point in the interior of $P_f$.

## 2. Main Results

In this section, we will state and prove our main results describing (necessary and) mostly sufficient absolutely irreducibility criteria for several collections of polynomials via arithmetical properties of their corresponding persistence's skeletons.

In the structure of our formal presentation we obtain a lot of inspiration by the general heuristics and cognitive nature of the multidisciplinary research program of Cognitive-Computational Metamathematics (CCMM) or Artificial Mathematical Intelligence (AMI) presented in [7]. Explicitly, we will derive the results gradually in a very intuitive, constructive and natural fashion, focusing on both the rigor and completeness of the proofs, and in a sequential edification of the more general facts. So, we could allow ourselves that similar lines of argumentation can occur at different levels of generality in order to stress that one can capture quite interesting and enlightening ideas even in the most concrete cases, and, on the other hand, one can loss a little bit of solid intuition among more general settings.

The first proposition establishes a complete characterization of irreducibility for suitable classes of binomials in two variables over an algebraically closed (coefficient) field. In our whole presentation $n$ denotes a natural number $\geqslant 1$.

**Proposition 2.1.** *Let $E$ be an algebraically closed field and $R_E = E[x, y]$. Let $f \in R_E$ be a binomial not divisible by any variable. Then, $f$ is irreducible in $R_E$ if and only if there exists one of the nonzero integral coordinates of some of the vertices of $P_f$, let us say $c$, such that $gcd(c, (pind_1(Psk(f)))^2) = 1$.*

*Proof.* Set $d = (pind_1(Psk(f)))^2$, and $Psk(f) = \{(a_1, b_1), (a_2, b_2)\} \subseteq \mathbb{Z}^2$, in other words, $f = h_1x^{a_1}y^{b_1} + h_2x^{a_2}y^{b_2}$, with $h_1, h_2 \neq 0$. From the fact that $f$ is not divisible by any of the variables we deduce that $a_1 = 0$ or $a_2 = 0$, and, $b_1 = 0$ or $b_2 = 0$. Now, from the definition of persistent barcode we see that $d = (a_1 - a_2)^2 + (b_1 - b_2)^2$, which, by the former fact, has the form of $d = a_j^2 + b_i^2$, for some $i, j \in \{1, 2\}$.

($\Leftarrow$) Suppose that $gcd(c, d) = 1$, where $c$ is some (nonzero) of the coordinates of a vertex of $P_f$, i.e., $c \in \{a_1, b_1, a_2, b_2\}$. Then, let us show that $gcd(a_1 - a_2, b_1 - b_2) = 1$. Effectively, assume for the sake of contradiction that there exists a prime number $p$ dividing $a_1 - a_2$ and $b_1 - b_2$. Then, by the former general facts we know that up to sign $c$ is equal either to $a_1 - a_2$ or to $b_1 - b_2$, since $c \neq 0$. So, $p$ divides $c$. Clearly, $p|d$. In conclusion, $p|gcd(d, c)$, which is absurd, since they are coprime numbers by hypothesis. Thus, $gcd(a_1 - a_2, b_1 - b_2) = 1$, then, by Proposition 1.10 and Proposition 1.8 $f$ is (absolutely) irreducible in $R_E$.

($\Rightarrow$) Suppose that $f$ is irreducible and assume by contradiction that there exists a nonzero coordinate $c$ of one of the vertex such that $d$ and $c$ are both divisible by a prime number $q$. Again, due to $c \neq 0$ and the general hypothesis, it holds that either $c = \pm(a_1 - a_2) = \pm a_j$ or $c = \pm(b_1 - b_2) = \pm b_j$, for some suitable $i, j \in \{1, 2\}$. From this fact, we conclude that $q$ divides any nonzero exponent of each of the variables appearing in each of the two monomials of $f$. So, each of the monomials of $f$ can be written as the $q$-th power of another monomial. Due to the fact that $E$ is an algebraically closed field, there exists $t_1, t_2 \in E$ such that $t_1^q = h_1$ and $t_2^q = -h_2$. Then,

$$f = (t_1 x^{a_1/q} y^{b_1/q})^q - (t_2 x^{a_2/q} y^{b_2/q})^q =$$

$$(t_1 x^{a_1/q} y^{b_1/q} - t_2 x^{a_2/q} y^{b_2/q})((t_1 x^{a_1/q} y^{b_1/q})^{q-1} + \cdots)$$

is a reducible polynomial in $R_E$, which contradicts the hypothesis. This completes the proof. $\square$

The next proposition generalizes only one direction the former criterion. The main challenge for polynomials in three (or more) variables is that the arithmetic condition given before in terms of the greatest common divisor of the first persistence index of dimension zero and one of the coordinates of some of the vertices does not directly implies that the coordinates of the difference of the entries of the vertices are coprime. This is due to the fact that there could be a collection of (difference of) integral coordinates such that the sum of their squares (which is equal to the first persistence index of dimension zero) is not coprime with each of them, but all of them are coprime.

**Proposition 2.2.** *Let $k$ be an arbitrary field and $R = k[x_1, \cdots, x_n]$, with $n \geqslant 3$. Let $f \in R$ be a binomial (polynomial) not divisible by any variable. If there exists one of the integral coordinates of some of the two vertices of $P_f$, let us say $c$, such that*

$$gcd(c, (pind_1(Psk(f)))^2) = 1,$$

*then $f$ is absolutely irreducible over $R$.*

*Proof.* Set $d = (pind_1(Psk(f)))^2$, and

$$Psk(f) = \{s = (s_1, \cdots, s_n), r = (r_1, \cdots, r_n)\} \subseteq \mathbb{Z}^n.$$

From the fact that $f$ is not divisible by any of the variables we deduce that for each $j \in \{1, \cdots, n\}$, either $s_i = 0$ or $r_i = 0$. Now, from the definition of persistent barcode we see that $d = (s_1 - r_1)^2 + \cdots + (s_n - r_n)^2$, which, by the former fact, has the form of $d = w_1^2 + \cdots + w_n^2$, where for each $j$, either $w_j = \pm s_j$ or $w_j = \pm r_j$.

Let us consider two cases. The first one is $c = 0$. In this case $d = 1$, which implies necessarily that there exists exactly one index $i \in \{1, \cdots, n\}$ such that $s_i = r_i \pm 1$. This entails to $gcd(s - r) = gcd(s_1 - r_1, \cdots, s_n - r_n) = 1$. Thus, by Propositions 1.10 and 1.8, $f$ is absolutely irreducible.

For the second case $c \neq 0$, we deduce from the former facts that $c = \pm(s_j - r_j)$, for some index $j$. So, if $gcd(s - r) \neq 1$, then a common prime divisor $p$ would necessarily divide $c$ and $d$, which is a contradiction since $gcd(c, d) = 1$. Thus, $gcd(s - r) = 1$, which implies by Propositions 1.10 and 1.8 that $f$ is absolutely irreducible. □

**Remark 2.3.** In the former proposition the arithmetical hypothesis is strong enough to guarantee that any polynomial satisfying it must be a binomial. Effectively, let $f \in R$ be a polynomial not divisible by any variable such that $P_f$ consists of two vertices and fulfilling the former hypothesis. Then, it is an elementary fact to verify that from the hypothesis one deduces that the greatest common divisor the differences of the entries of the two vertices of $P_f$ is one. So, if $f$ is formed by more than two different monomials, then at least one of them would correspond to a point strictly in the middle of the two vertices of $P_f$. Otherwise, $P_f$ would not be a line, but at least a triangle having strictly more than two vertices. Thus, this implies that there exists a third point $u \in \mathbb{R}^n$ with integral entries strictly in between the two vertices. Now, this fact implies that the greatest common divisor of the entries of the difference of the vertices is greater than one. In fact, after translation one can assume that one of the vertices is the origin. Additionally, one can prove by an elementary argument that any point (different from the origin) of a line passing through the origin in $\mathbb{R}^n$ and with integer entries (if there exists such a point) should be an integer multiple of one of the points of the finite set consisting of elements in $\mathbb{Z}^n$ within the line with smallest (non-zero) Euclidean norm, which, in fact, corresponds with the set of integral points of the line whose entries are coprime. So, from the former general property and from the fact that $u$ should have Euclidean norm strictly smaller that the non-zero vertex of $P_f$, we conclude that the entries of the this non-zero vertex (or, more generally, of the difference of the originally two vertices) are not coprime, which is absurd. So, our original polynomial $f$ should be a binomial. This justifies the addition of the word *polynomial* in parentheses in hypothesis of the last proposition.

Our next result involves a criterion for the case of trinomials involving a simple arithmetic condition for two of the persistence indexes of the corresponding persistence skeleton.

**Proposition 2.4.** *Let $k$ be an arbitrary field and $R = k[x_1, \cdots, x_n]$. Let $f \in R$ be a polynomial not divisible by any variable. Suppose that the persistence's skeleton of $f$ is a triangle such that the first two persistence indexes exists, they are both different from zero and*

$$gcd\left((pind_1(Psk(f)))^2, (pind_2(Psk(f)))^2\right) = 1.$$

*Then $f$ is absolutely irreducible over $R$.*

*Proof.* Set $d = (pind_1(Psk(f)))$ and $e = (pind_2(Psk(f)))$. First, note that by basic properties of persistence homology and by the definition of persistence barcode $d$ and $e$ are precisely the lengths of two of the sides of the triangle $Psk(f)$. So, since the vertices of $Psk(f)$ are integral points, the numbers $d^2$ and $e^2$ are positive integers. Now, since these distances $d$ and $e$ fulfill the condition that $gcd(d^2, e^2)$, by hypothesis, they should involve exactly one vertex of $Psk(f)$, let us say $v_1$. Denote by $v_2$ and $v_3$ the other two vertices of $Psk(f)$. We claim that $gcd(v_1 - v_2, v_1 - v_3) = 1$. Effectively, if not choose a prime number $p$ such that $p$ divides all the entries of $v_1 - v_2$ and $v_1 - v_3$. Then, by definition of $d^2$ and $e^2$, $p$ would divide both of them, because $\|v_1 - v_2\|^2 = d^2$ and $\|v_1 - v_3\|^2 = e^2$. This contradicts the fact that $gcd(d^2, e^2) = 1$. In conclusion, $gcd(v_1 - v_2, v_1, v_3) = 1$. Finally, since by the former fact and by Proposition 1.11 $P_f = conv(Psk(f)) = conv(v_1, v_2 - v_3)$ is integrally indecomposable, this implies by Proposition 1.8 and by hypothesis that $f$ is absolutely irreducible over $R$. □

In the next proposition we give simple, buy slightly more technical conditions for guaranteeing the absolutely irreducibility in the case that the corresponding persistence's skeletons consists of four vertices. As the reader may appreciate, as the number of vertices slightly increase the geometrical complexity of the possible configurations of persistence indices increase considerably fast.

**Proposition 2.5.** *Let $k$ be an arbitrary field and $R = k[x_1, \cdots, x_n]$. Let $f \in R$ be a polynomial not divisible by any variable. Suppose that the persistence's skeleton of $f$ is a tetrahedron with noncoplanar vertices $Psk(f) = \{v_1, v_2, v_3, v_4\}$. Assume that there exist indices $h, i, j \in \{1, 2, 3\}$ and a vertex of $Psk(f)$, let us* *say $v_r$, such that $pind_h(Psk(f)), pind_i(Psk(f)$ and $pind_j(Psk(f))$ represent nonzero numbers and correspond to distances among the vertex $v_r$ and the other three vertices of $Psk(f)$. Additionally, suppose that the following condition holds*

$$gcd\left((pind_h(Psk(f)))^2, (pind_i(Psk(f)))^2, (pind_j(Psk(f)))^2\right) = 1.$$

*Then $f$ is absolutely irreducible over $R$.*

*Proof.* Set $d = pind_h(Psk(f)), e = pind_i(Psk(f))$ and $g = pind_j(Psk(f))$. After re-enumeration we can assume that that $v_r = v_1$. Moreover, note that in the hypothesis we are not necessarily counting multiplicity among the persistence indexes. In other words, some of the indexes $h, i$ and $j$ could be equal. The crucial fact here is that the correspond to distances among the fixed vertex $v_1$ and all the other vertices of $Psk(f)$. We claim that $gcd(v_1 - v_2, v_1 - v_3.v_1 - v_4) = 1$. If not, then let $p$ be a prime dividing all the entries of $v_1 - v_2, v_1 - v_3$ and $v_1 - v_4$. So, $p$ would divide the squares of their corresponding Euclidean norms, i.e., $p|d^2 = \|v_1 - v_2\|^2, p|e^2 = \|v_1 - v_3\|^2$ and $p|g^2 = \|v_1 - v_4\|^2$. This contradicts directly the hypothesis. Thus, by Proposition 1.12 $Psk(f) = conv(v_1, v_2, v_3, v_4)$ is integrally indecomposable, and then the hypothesis and Proposition 1.8 implies that $f$ is absolutely irreducible. □

In our next statement, we establish suitable geometric conditions for persistence skeletons with arbitrarily many vertices that implies again absolutely irreducibility.

**Proposition 2.6.** *Let $k$ be an arbitrary field and $R = k[x_1, \cdots, x_n]$. Let $f \in R$ be a polynomial not divisible by any variable. Suppose that the persistence's skeleton of $f$ is a pyramid with vertices $Psk(f) = \{v, v_1, \cdots, v_m\}$, where the vertices $v_1, \cdots, v_m$ lie in a hyperplane $H$ and $v \notin H$. Suppose that there exist indices $i_1, \cdots, i_p \in \{1, \cdots, m\}$, with $i_p \leqslant m$, such that*

$$pind_{i_1}(Psk(f)), \cdots, pind_{i_p}(Psk(f))$$

*correspond exactly to distances between $v$ and some vertices $v_{j_1}, \cdots, v_{j_p}$, with $j_1, \cdots, j_p \in \{1, \cdots, m\}$.*

*Thus, if*

$$gcd\left((pind_{i_1}(Psk(f)))^2, \cdots, (pind_{i_p}(Psk(f)))^2\right) = 1,$$

*then $f$ is absolutely irreducible over $R$.*

*Proof.* Following similar conventions as in the former proofs, we claim that our hypothesis implies that $gcd(v - v_1, \ldots, v - v_m) = 1$. Otherwise, let $p$ be a prime number that divides all the coordinates of $v - v_1, \cdots, v - v_m$. Then, in particular, $p$ would divide all the square of the numbers

$$pind_{i_1}(Psk(f)))^2 = \|v - v_{j_1}\|^2, \cdots, (pind_{i_p}(Psk(f)))^2 = \|v - v_{j_p}\|^2,$$

which contradicts the main condition regarding the coprimality of these numbers. Then, by Propositions 1.8 and Theorem 1.9 $f$ is absolutely irreducible. □

Complementing the methodological remarks made at the beginning of this section, it is worth noting that although one could derive formally some of the former propositions from Proposition 2.6, we believe that it is more convenient for the sake of a deep understanding of the reader about the core intuitions behind these results, to present them more organically and in coherence with the way in which they were discovered; emphasizing expressly the highlights at each particular stage of generally, e.g. with 1, 2, 3, 4, and more vertices. We hope that this kind of presentation will offer not only the formal demonstrations, but also valuable information about the 'out of the blue' dimension of the mathematical creation, which is, in general, hidden within the formalism and purely logic presentation of most of the research articles published in (pure) mathematics.

## 3. General Conclusions and Further Potential Future Work

We believe that all the results presented in this article represent just the top of the iceberg consisting of very foundational and enlightening formal connections between notions coming from and inspired by persistence homology and topological data analysis, on the one hand; and from combinatorial commutative algebra and number theory, on the other hand.

A natural question to further study would be the connection of arithmetical-geometric properties of persistence indexes of higher dimension for guaranteeing and characterizing (absolute) irreducibility of arbitrary polynomial in several variables over a field or over suitable unique factorization domains, for example.

## Acknowledgements

Danny A. J. Gomez-Ramirez would like to thank Michelle Gomez Villa, Irene Villa, Sebastian Rendon and Diego Hincapie for all the love, kindness and support.

## References

1. Nils A Baas, Gunnar E Carlsson, Gereon Quick, Markus Szymik, and Marius Thaule, *Topological data analysis*, Springer, 2020.
2. Anuraag Bukkuri, Noemi Andor, and Isabel K Darcy, *Applications of topological data analysis in oncology*, Frontiers in artificial intelligence **4** (2021), 659037.
3. Herbert Edelsbrunner, John Harer, et al., *Persistent homology-a survey*, Contemporary mathematics **453** (2008), no. 26, 257–282.

4. Michelle Feng and Mason A Porter, *Spatial applications of topological data analysis: Cities, snowflakes, random structures, and spiders spinning under the influence*, Physical Review Research **2** (2020), no. 3, 033426.
5. John B Fraleigh, *A first course in abstract algebra*, Pearson Education India, 2003.
6. S. Gao, *Absolute irreducibility of polynomials via Newton polytopes*, J. Algebra **237** (2001), no. 2, 501–520. MR 1816701
7. D. A. J. Gómez Ramírez, *Artificial mathematical intelligence—cognitive, (meta)mathematical, physical and philosophical foundations*, Springer, Cham, 2020. MR 4225288
8. Richard M Heiberger, Erich Neuwirth, Richard M Heiberger, and Erich Neuwirth, *Polynomial regression*, R Through Excel: A Spreadsheet Interface for Statistics, Data Analysis, and Graphics (2009), 269–284.
9. Vishal Jagota, Aman Preet Singh Sethi, and Khushmeet Kumar, *Finite element method: an overview*, Walailak Journal of Science and Technology (WJST) **10** (2013), no. 1, 1–8.
10. Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, and Jonathan Taylor, *Linear regression*, An introduction to statistical learning: With applications in python, Springer, 2023, pp. 69–134.
11. Max Z Li, Megan S Ryerson, and Hamsa Balakrishnan, *Topological data analysis for aviation applications*, Transportation Research Part E: Logistics and Transportation Review **128** (2019), 149–174.
12. Eva Ostertagová, *Modelling using polynomial regression*, Procedia Engineering **48** (2012), 500–506.
13. Alexander M Ostrowski, *On multiplication and factorization of polynomials, i. lexicographic orderings and extreme aggregates of terms*, aequationes mathematicae **13** (1975), no. 3, 201–228.
14. Nina Otter, Mason A Porter, Ulrike Tillmann, Peter Grindrod, and Heather A Harrington, *A roadmap for the computation of persistent homology*, EPJ Data Science **6** (2017), 1–38.
15. Andrzej Schinzel, *Polynomials with special regard to reducibility*, vol. 77, Cambridge University Press, 2000.
16. Yara Skaf and Reinhard Laubenbacher, *Topological data analysis in biomedicine: A review*, Journal of Biomedical Informatics **130** (2022), 104082.
17. Alexander D Smith, Paweł Dłotko, and Victor M Zavala, *Topological data analysis: concepts, computation, and applications in chemical engineering*, Computers & Chemical Engineering **146** (2021), 107202.
18. Larry Wasserman, *Topological data analysis*, Annual Review of Statistics and Its Application **5** (2018), 501–532.



VISIÓN REAL COGNITIVA (COGNIVISIÓN) S.A.S. ITAGUÍ, COLOMBIA..
*Email address*: `daj.gomezramirez@gmail.com`

UNIVERSITY INSTITUTION TECNOLOGICO DE ANTIOQUIA, MEDELLÍN, COLOMBIA.
*Email address*: `egalleg@gmail.com`

UNIVERSITY INSTITUTION TECNOLOGICO DE ANTIOQUIA, MEDELLÍN, COLOMBIA.
*Email address*: `jaiberth.porras@udea.edu.co`

UNIVERSITY INSTITUTION TECNOLOGICO DE ANTIOQUIA, MEDELLÍN, COLOMBIA.
*Email address*: `diego.lopez@tdea.edu.co`

UNIVERSIDAD DE ANTIOQUIA, MEDELLÍN, COLOMBIA.
*Email address*: `sonia.agudeloa@udea.edu.co`