

Persistent Homology: Intro

Tao Hou, University of Oregon

Outline for studying persistent homology

1. Intro to persistent homology
 - Build intuitions of persistent homology: what it does, what it produces
2. Formalizing persistent homology
 - Introduce its input (filtration) and study an algorithm for computation
3. Different ways for building filtrations
 - Vietoris-Rips filtration, sub-levelset filtration
 - Cubical complexes (for images)
4. Interpretation and stability of persistence diagram

Outline for studying persistent homology

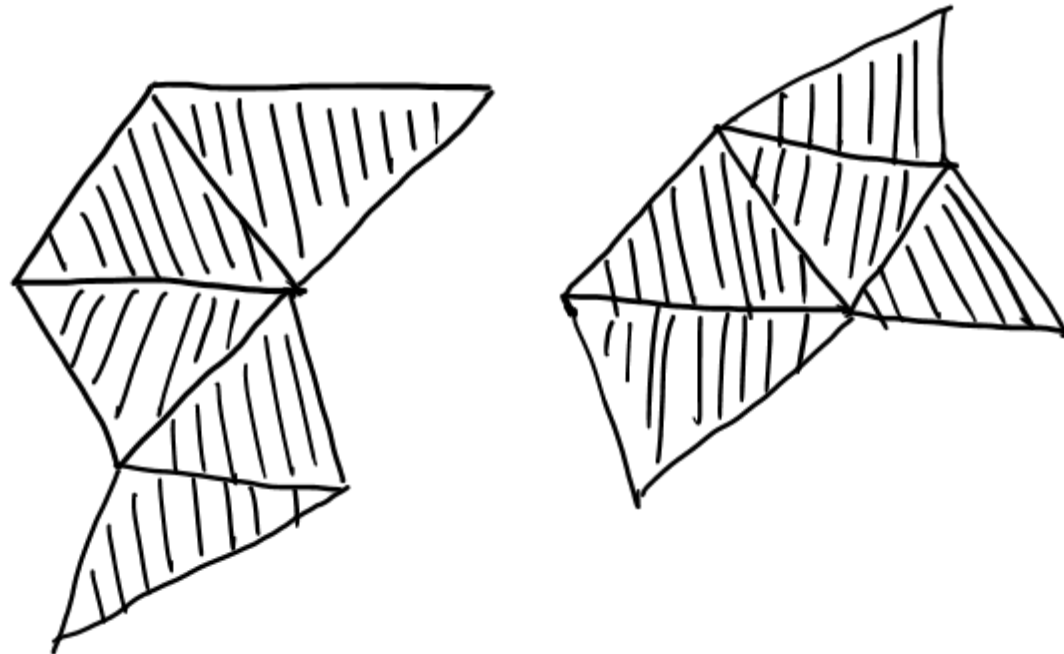
1. Intro to persistent homology
 - Build intuitions of persistent homology: what it does, what it produces
2. Formalizing persistent homology
 - Introduce its input (filtration) and study an algorithm for computation
3. Different ways for building filtrations
 - Vietoris-Rips filtration, sub-levelset filtration
 - Cubical complexes (for images)
4. Interpretation and stability of persistence diagram

Homology inference

- We know now that, given a topological space (e.g., a simplicial complex), we can use homology (e.g., *Betti number* or *homology basis*) to infer the shape of the data in different dimensions

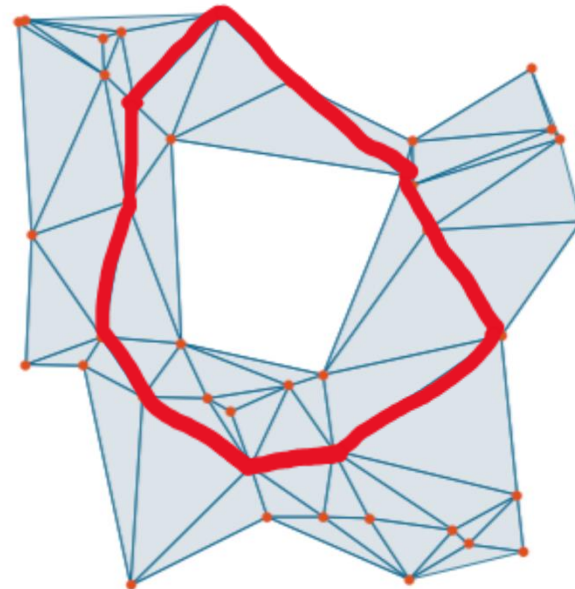
Homology inference

- We know now that, given a topological space (e.g., a simplicial complex), we can use homology (e.g., *Betti number* or *homology basis*) to infer the shape of the data in different dimensions
- **Ex:** There is a 0-dimensional hole of the following complex because of the gap between the two connected components



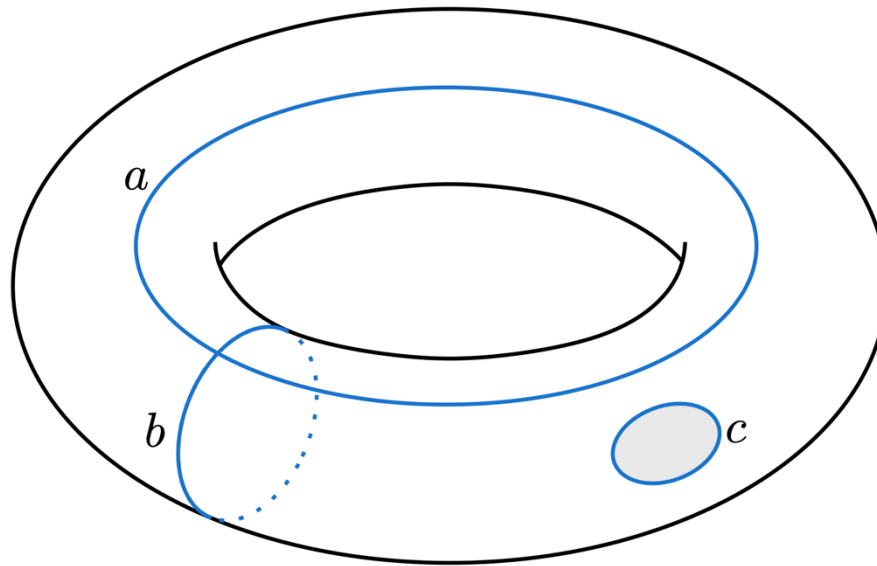
Homology inference

- We know now that, given a topological space (e.g., a simplicial complex), we can use homology (e.g., *Betti number* or *homology basis*) to infer the shape of the data in different dimensions
- **Ex:** The homology basis for the 1-cycles in the below simplicial complex contains the single red 1-cycle.
 - So that we can use the red cycle to represent the 1-dimensional “*homological features*” of the space



Homology inference

- We know now that, given a topological space (e.g., a simplicial complex), we can use homology (e.g., *Betti number* or *homology basis*) to infer the shape of the data in different dimensions
- **Ex:** The 1-dimensional homological features of a torus can be characterized by two cycles:
 - a (longitude) and b (meridian)



Homology inference

- Homology theory was invented by **Poincaré** about 100 years ago
- So far, the inference of the shape of topological spaces using homology theory seems perfect
- But is there any problem?



Homology inference

- Homology theory was invented by **Poincaré** about 100 years ago
- So far, the inference of the shape of topological spaces using homology theory seems perfect
- But is there any problem?
- We shall look at at least two problems with it



Homology inference: Problem 1

- Homology inference relies on given a simplicial complex as input

Homology inference: Problem 1

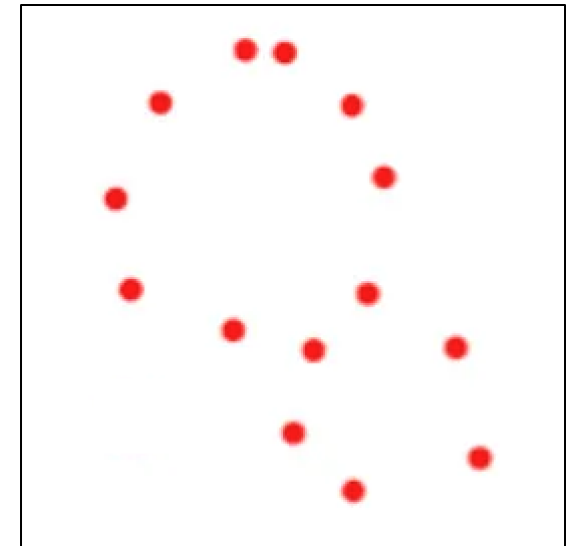
- Homology inference relies on given a simplicial complex as input
- Simplicial complex is “highly structured” data, while in practice we don’t have the luxury of always having data rich structure

Homology inference: Problem 1

- Homology inference relies on given a simplicial complex as input
- Simplicial complex is “highly structured” data, while in practice we don’t have the luxury of always having data rich structure
- Typically, data come in as “unstructured” (e.g., point clouds)

Homology inference: Problem 1

- Homology inference relies on given a simplicial complex as input
- Simplicial complex is “**highly structured**” data, while in practice we don’t have the luxury of always having data rich structure
- Typically, data come in as “**unstructured**” (e.g., point clouds)
- For the right point cloud (which is unstructured), everyone could see that it consists of two rings (1-cycles)
- But we have to **construct a simplicial complex from the point cloud** first to infer this information



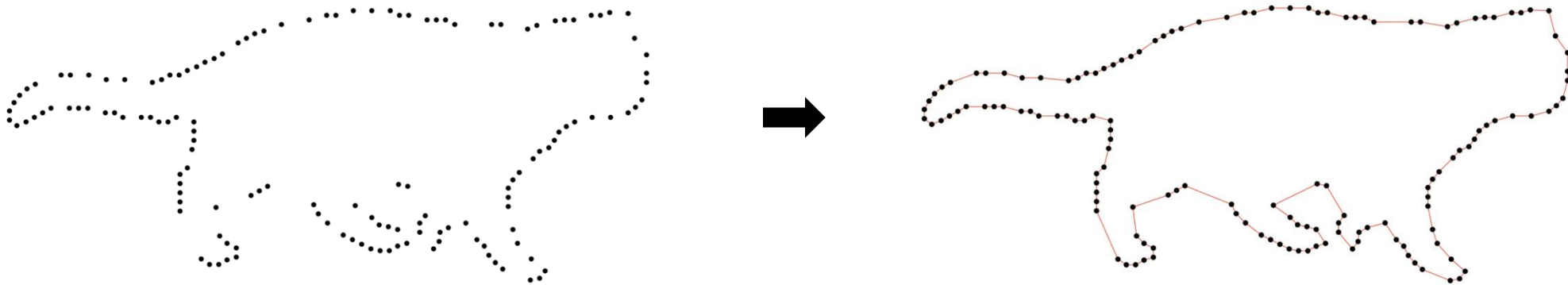
Homology inference: Problem 1

- There are mature methods on **reconstruction from point clouds**.
- In 2D:



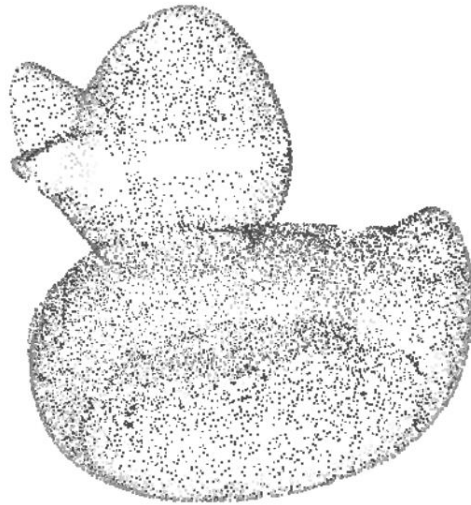
Homology inference: Problem 1

- There are mature methods on **reconstruction from point clouds**.
- In 2D:



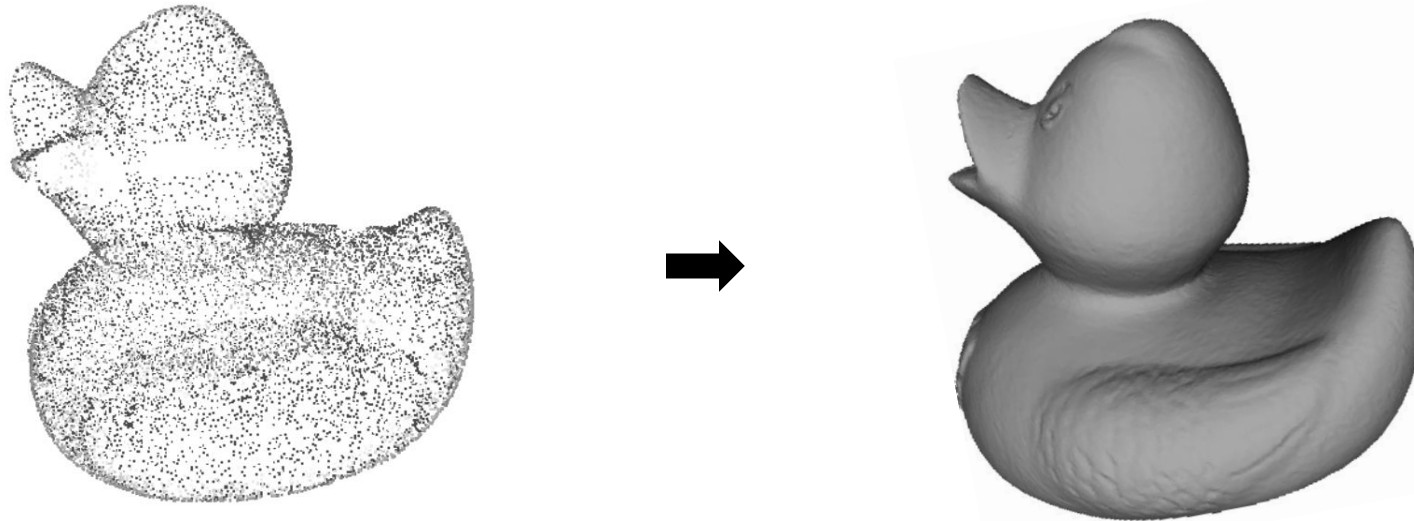
Homology inference: Problem 1

- There are mature methods on **reconstruction from point clouds**.
- In 3D:



Homology inference: Problem 1

- There are mature methods on **reconstruction from point clouds**.
- In 3D:



Homology inference: Problem 1

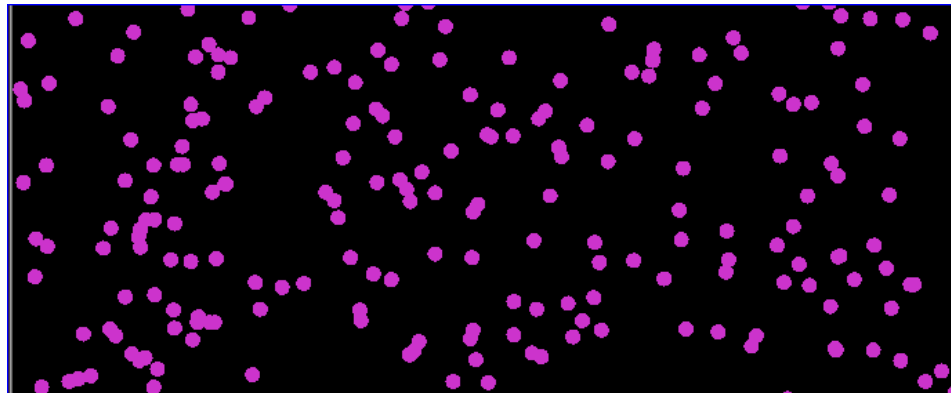
- There are mature methods on **reconstruction from point clouds**.
- But there still are problems:
 1. The reconstructions process can be **costly**

Homology inference: Problem 1

- There are mature methods on **reconstruction from point clouds**.
- But there still are problems:
 1. The reconstructions process can be **costly**
 2. There are probably **more information** in the original unstructured data **than is reconstructed**

Homology inference: Problem 1

- There are mature methods on **reconstruction from point clouds**.
- But there still are problems:
 1. The reconstructions process can be **costly**
 2. There are probably **more information** in the original unstructured data **than is reconstructed**
 3. Reconstruction from point clouds which are not nicely shaped is **very hard** if at all possible



Homology inference: Problem 2

- In the following space, there are **seven 1-dimensional holes** (i.e., *homology basis contains seven non-trivial 1-cycles*)



Homology inference: Problem 2

- In the following space, there are **seven 1-dimensional holes** (i.e., *homology basis contains seven non-trivial 1-cycles*)
- We all could see that the three circled ones are more “**significant features**”, while the remaining ones could be well due to some “**artifacts**” or “**noise**”



Homology inference: Problem 2

- In the following space, there are **seven 1-dimensional holes** (i.e., *homology basis contains seven non-trivial 1-cycles*)
- We all could see that the three circled ones are more “**significant features**”, while the remaining ones could be well due to some “**artifacts**” or “**noise**”
- i.e., with some “perturbation” on the data, the four small holes could simply disappear



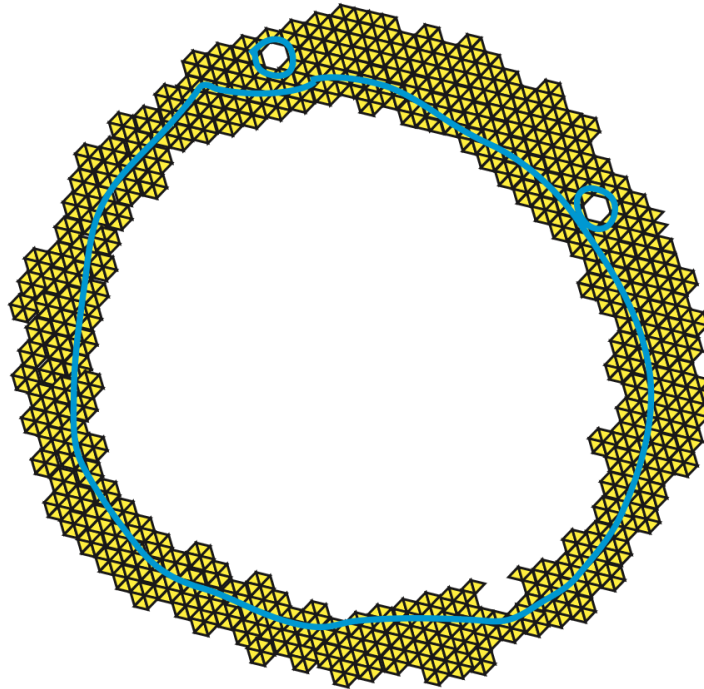
Homology inference: Problem 2

- In the following space, there are **seven 1-dimensional holes** (i.e., *homology basis contains seven non-trivial 1-cycles*)
- We all could see that the three circled ones are more “**significant features**”, while the remaining ones could be well due to some “**artifacts**” or “**noise**”
- i.e., with some “perturbation” on the data, the four small holes could simply disappear
- But using homology basis we could not differentiate the “**more significant holes**” from the “**less significant ones**”



Homology inference: Problem 2

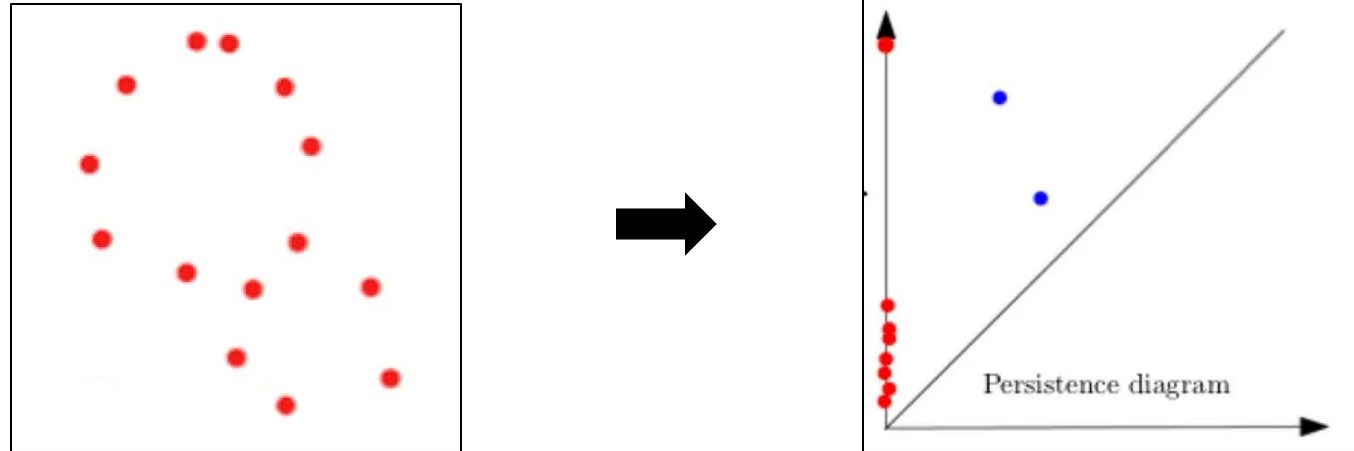
- Similarly, in the following space, there are **three 1-dimensional holes**, but there is clearly a “more significant” one and two “less significant” ones which also be some artifacts
- Again, using just homology basis we could not differentiate them



Solution: Persistent homology

Solution: Persistent homology

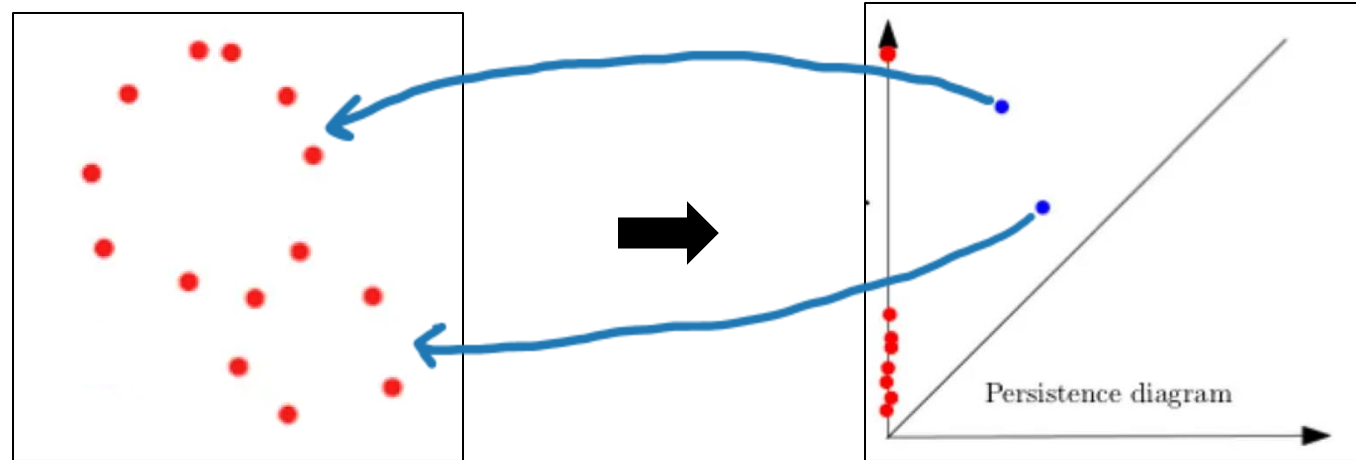
- Solving problem 1:



- For the point cloud, persistent homology produces a “topological signature” called **persistence diagram**
- In the diagram, the **blue dots** represents the two rings, thus correctly inferring the topological structure of the point cloud

Solution: Persistent homology

- Solving problem 1:



- For the point cloud, persistent homology produces a “topological signature” called **persistence diagram**
- In the diagram, the **blue dots** represents the two rings, thus correctly inferring the topological structure of the point cloud

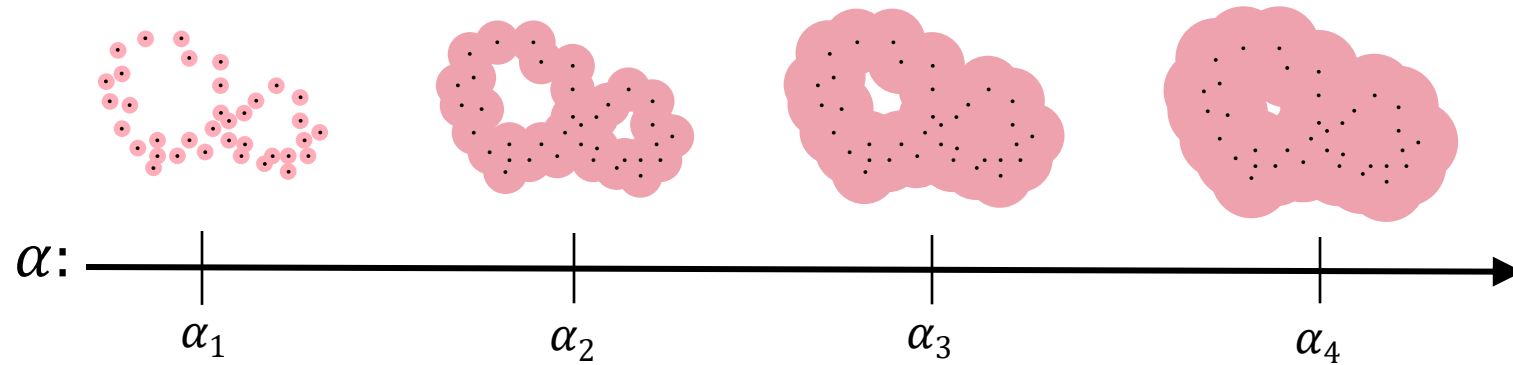
Solution: Persistent homology

- Solving problem 2:



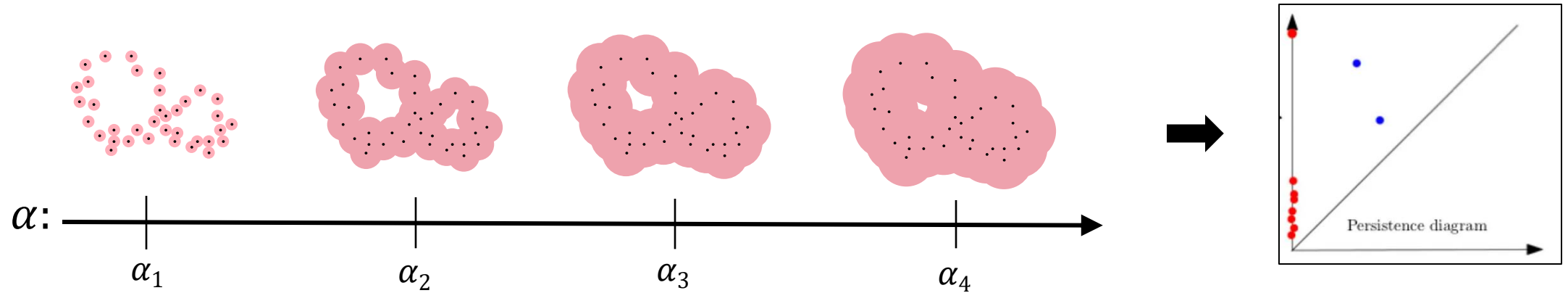
- For the above shape, its **persistence diagram** provides a **measure of the “size”** (i.e., “significance”) of the 1-dimensional holes so that we can differentiate the three more significant ones from the remaining

Persistent homology, more formally



- The input to persistent homology is a **growing topological space**

Persistent homology, more formally



- The input to persistent homology is a **growing topological space**
- Given this, it produces a persistence diagram, which is a **robust** (i.e., stable) “topological signature” that **captures the multi-scale topological features** (aka. **holes**) of the data **in arbitrary dimensions**

Persistent homology, more formally

Idea:

- The growing space can be more formally defined as follows:

Persistent homology, more formally

Idea:

- The growing space can be more formally defined as follows:
 - We let a value α ranges, say, from 0 to ∞

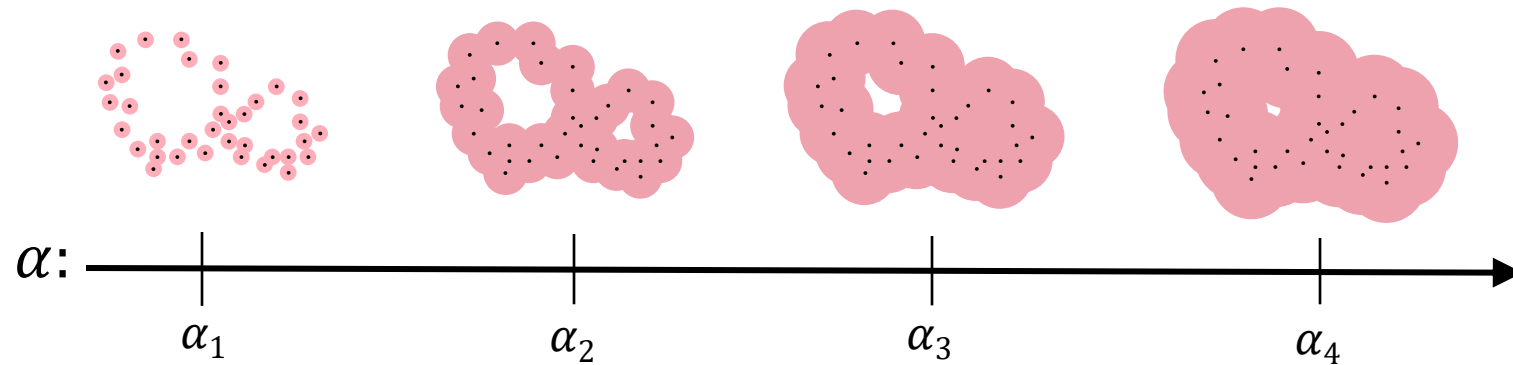
Persistent homology, more formally

α : 

Idea:

- The growing space can be more formally defined as follows:
 - We let a value α ranges, say, from 0 to ∞

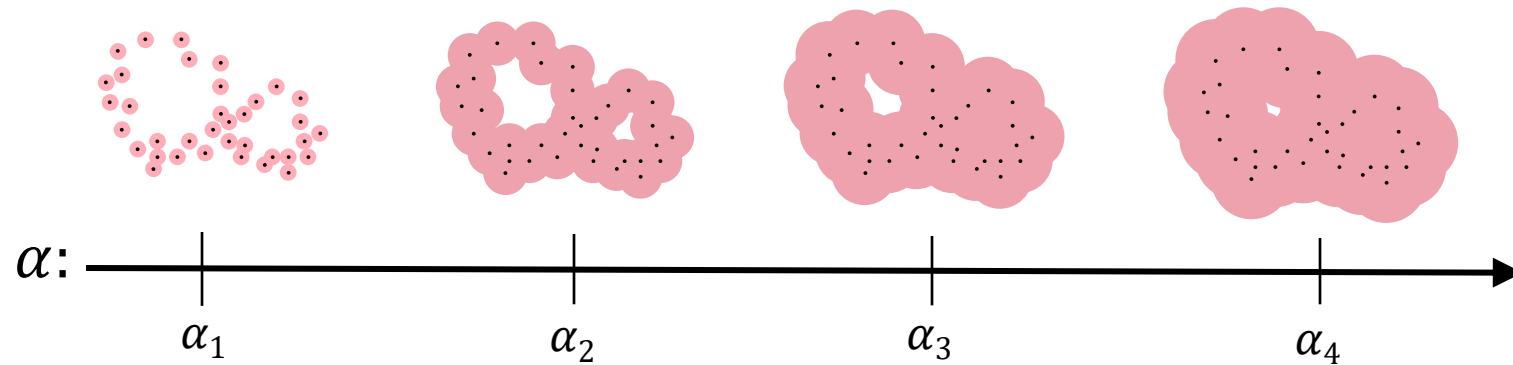
Persistent homology, more formally



Idea:

- The growing space can be more formally defined as follows:
 - We let a value α ranges, say, from 0 to ∞
 - Let each value α corresponds to a topological space so that
 - The topological space grows as α increases from 0 to ∞

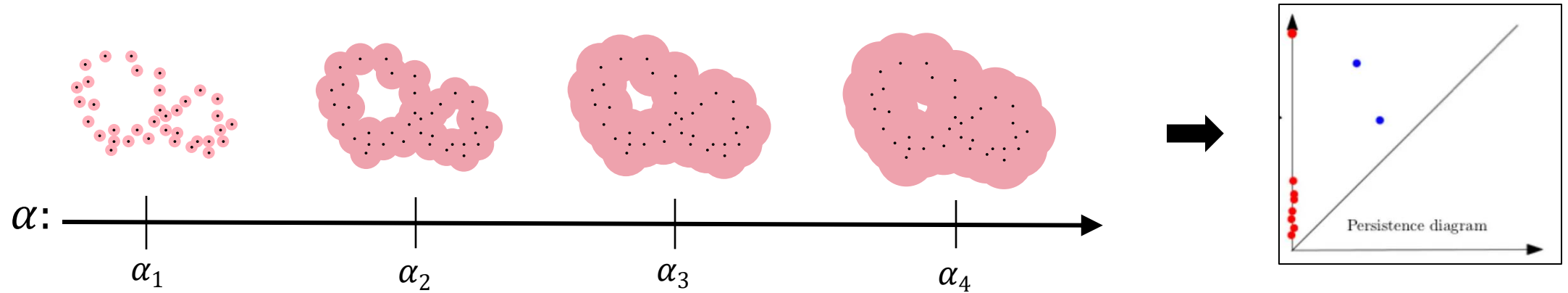
Persistent homology, more formally



Idea:

- The growing space can be more formally defined as follows:
 - We let a value α ranges, say, from 0 to ∞
 - Let each value α corresponds to a topological space so that
 - The topological space grows as α increases from 0 to ∞
- Then, as α increase, we track the changes of the homology features of the corresponding spaces

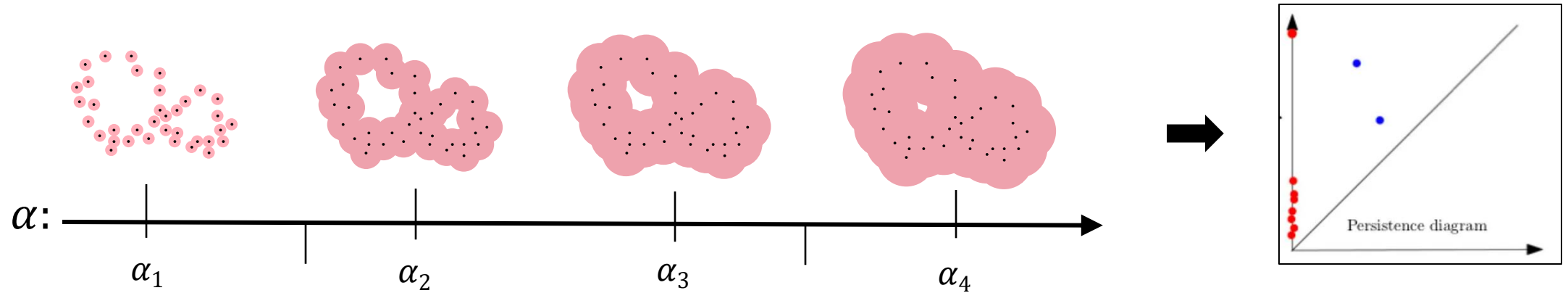
Persistent homology, more formally



Examples:

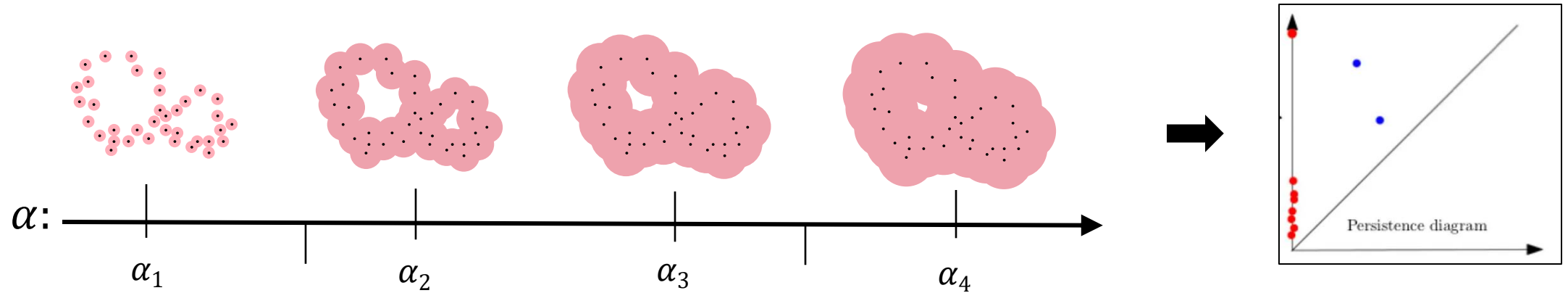
- https://gjkoplik.github.io/pers-hom-examples/0d_pers_2d_data_widget.html
- https://gjkoplik.github.io/pers-hom-examples/1d_pers_2d_data_widget.html

Persistent homology, more formally



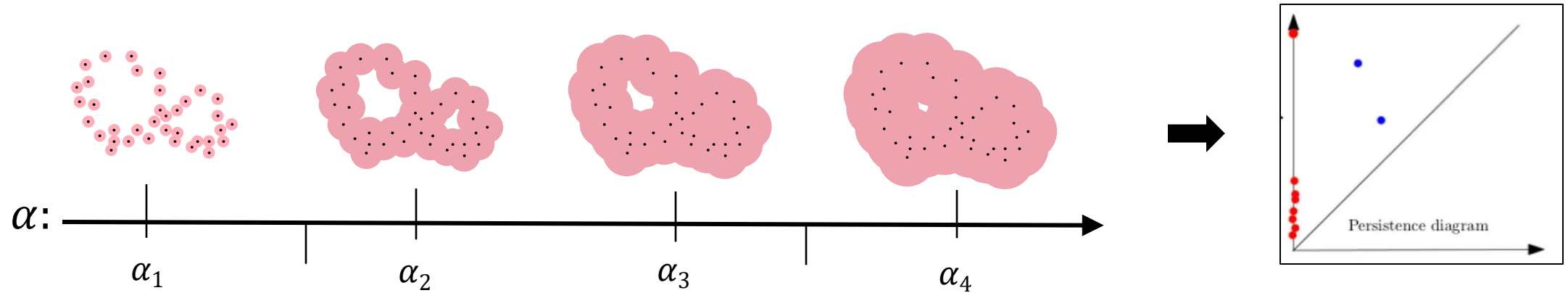
- **Definition:** A *persistence diagram (PD)* is a set of points on the 2D plane above the diagonal such that:

Persistent homology, more formally



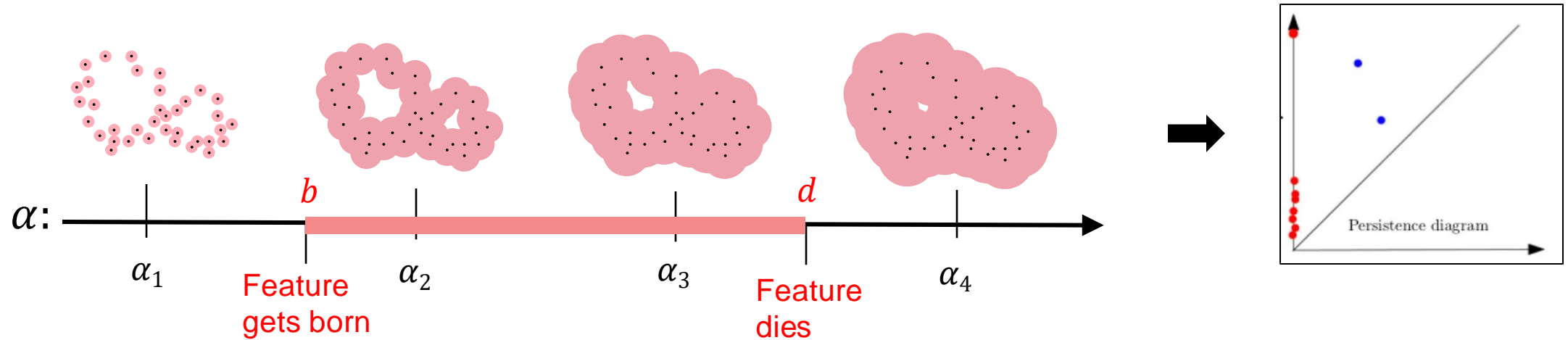
- **Definition:** A *persistence diagram (PD)* is a set of points on the 2D plane above the diagonal such that:
 - Each point in the PD represents the **birth** and **death** a homological feature (aka. cycle / hole) of the data in a certain dimension.

Persistent homology, more formally



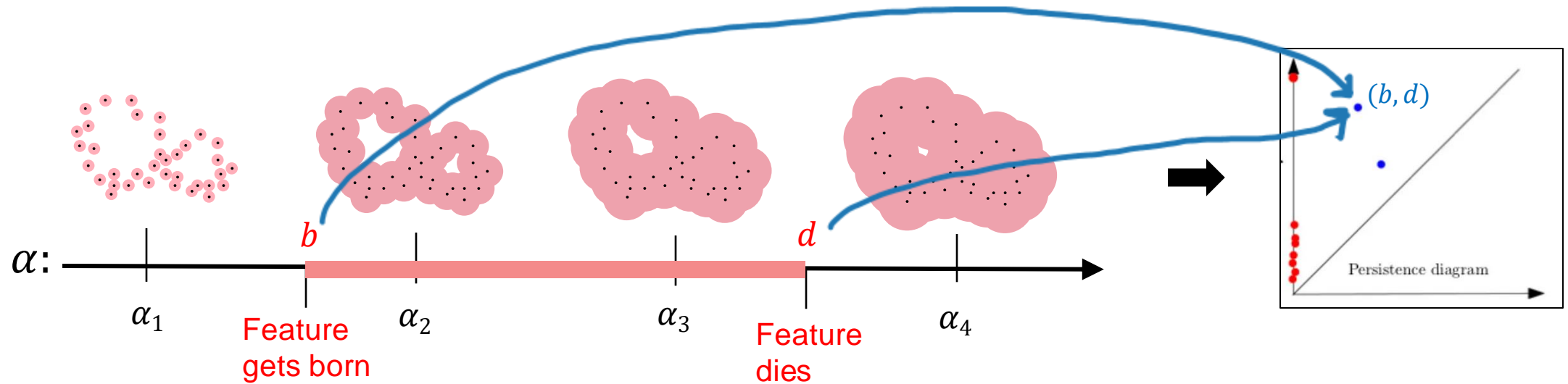
- **Definition:** A *persistence diagram (PD)* is a set of points on the 2D plane above the diagonal such that:
 - Each point in the PD represents the **birth** and **death** a homological feature (aka. cycle / hole) of the data in a certain dimension.
 - A point (b, d) :
 - b indicates **birth value** (the α value in which the feature is born)
 - d indicates **death value** (the α value in which the feature is dies)

Persistent homology, more formally



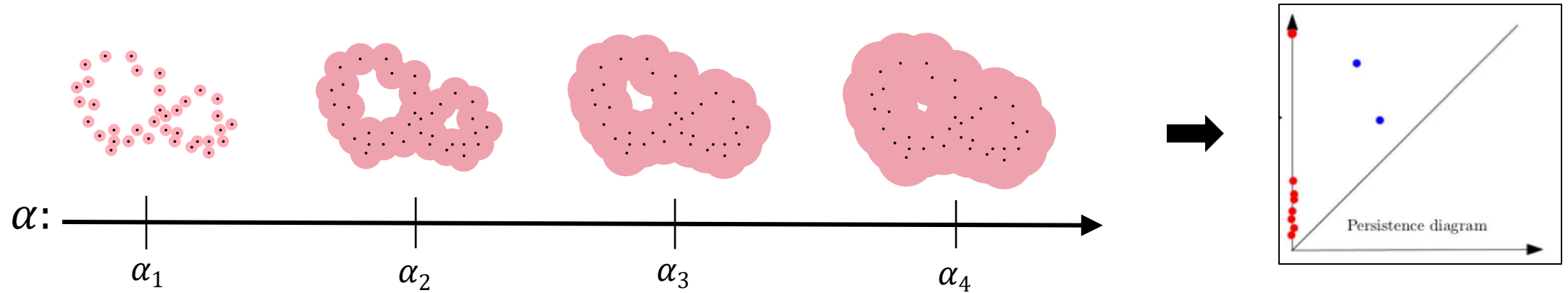
- **Definition:** A *persistence diagram (PD)* is a set of points on the 2D plane above the diagonal such that:
 - Each point in the PD represents the **birth** and **death** a homological feature (aka. cycle / hole) of the data in a certain dimension.
 - A point (b, d) :
 - b indicates **birth value** (the α value in which the feature is born)
 - d indicates **death value** (the α value in which the feature is dies)

Persistent homology, more formally



- **Definition:** A *persistence diagram (PD)* is a set of points on the 2D plane above the diagonal such that:
 - Each point in the PD represents the **birth** and **death** a homological feature (aka. cycle / hole) of the data in a certain dimension.
 - A point (b, d) :
 - b indicates **birth value** (the α value in which the feature is born)
 - d indicates **death value** (the α value in which the feature is dies)

Persistent homology, more formally



- Notice that homology features / holes are in different dimensions
- The PD where points corresponding to d -dimensional holes is also called the *d -dimensional / d -th PD* which is typically denoted as PD_d
- And of course, we could also have the PD *in all dimensions* (this is the PD *by default*)

Persistent homology: History

- Persistent homology is proposed roughly around 2000 (or earlier) by several works
- The following is *by no means a comprehensive list of works*:
 - Edelsbrunner, Letscher and Zomorodian, 2002. Topological persistence and simplification.
 - Zomorodian, A. and Carlsson, G., 2004, June. Computing persistent homology.
 - Carlsson, G., 2009. Topology and data.
 - Ghrist, R., 2008. Barcodes: the persistent topology of data.
 - Singh, G., Mémoli, F. and Carlsson, G.E., 2007. Topological methods for the analysis of high dimensional data sets and 3d object recognition.

Motivation: Homology inference from points clous

- We try to infer the homology for the following point cloud data

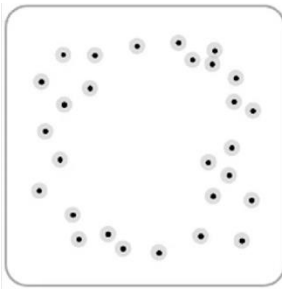


Image source: Bobrowski
O, Skraba P. A universal
null-distribution for
topological data analysis.

Motivation: Homology inference from points clous

- We try to infer the homology for the following point cloud data
- For this, we need to [build a meaningful topological space](#)

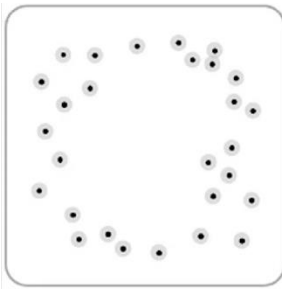
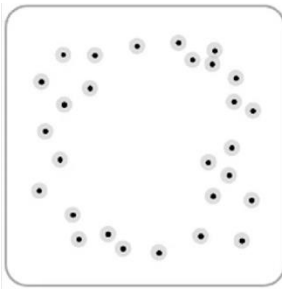


Image source: Bobrowski
O, Skraba P. A universal
null-distribution for
topological data analysis.

Motivation: Homology inference from points clous

- We try to infer the homology for the following point cloud data
- For this, we need to **build a meaningful topological space**
- Our strategy **is to connect the dots by increasing their size**, as before



Motivation: Homology inference from points cloud

- We try to infer the homology for the following point cloud data
- For this, we need to **build a meaningful topological space**
- Our strategy **is to connect the dots by increasing their size**, as before
- Notice that there are **different choices of the size**

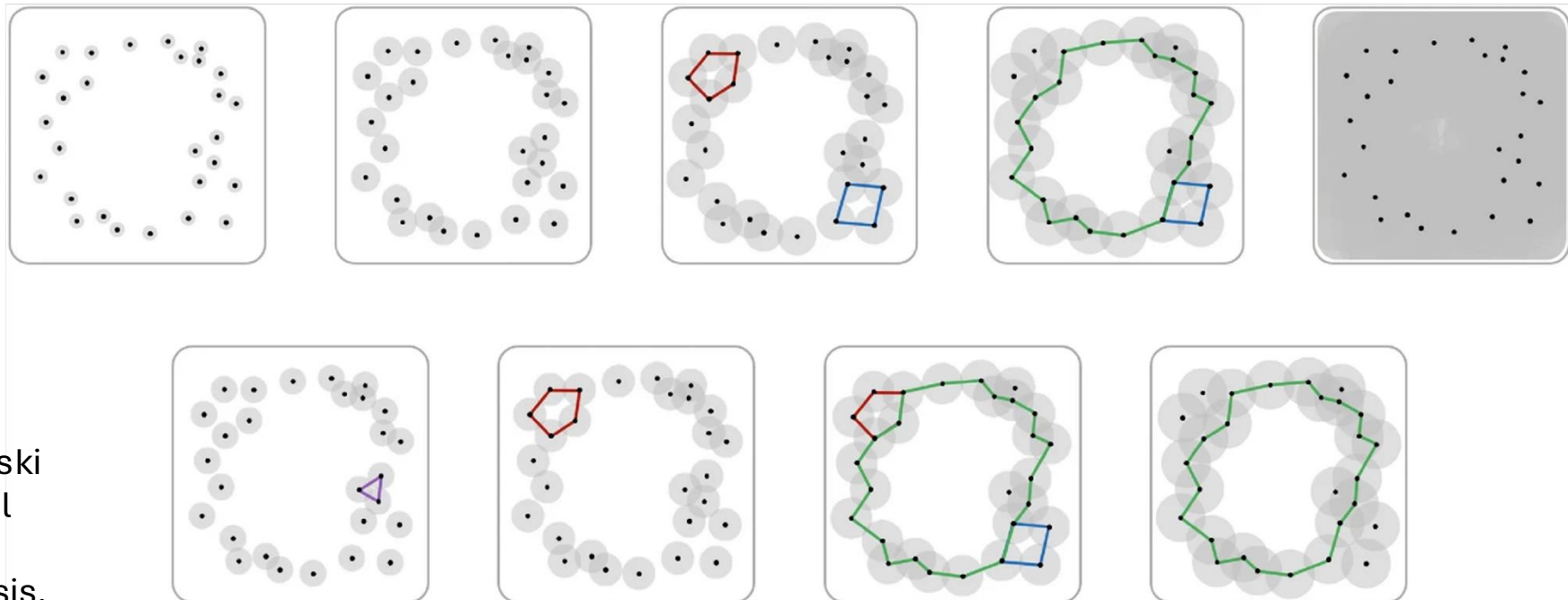


Image source: Bobrowski
O, Skraba P. A universal
null-distribution for
topological data analysis.

Motivation: Homology inference from points cloud

- Technically, a point does not have “size”, so what we are actually doing here is that we put a 2-dimensional ball around each point, where all such balls have the same radius.

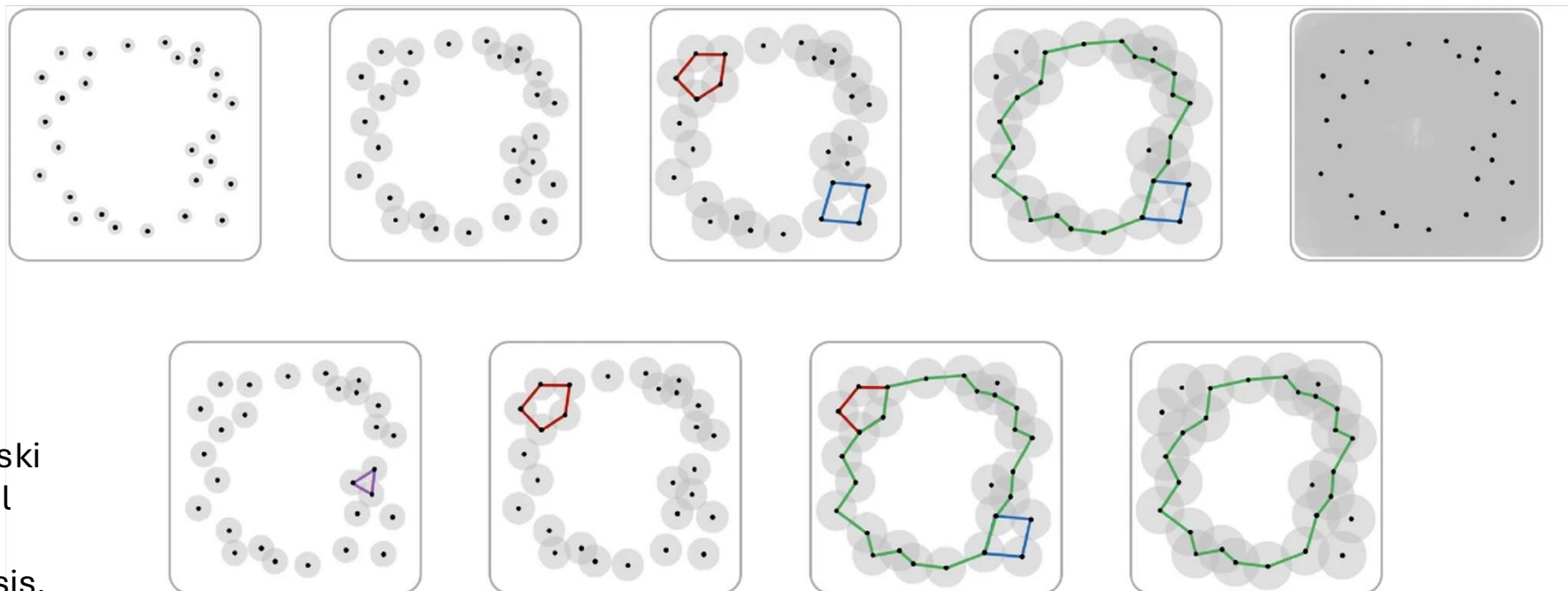


Image source: Bobrowski
O, Skraba P. A universal
null-distribution for
topological data analysis.

Motivation: Homology inference from points cloud

- Technically, a point does not have “size”, so what we are actually doing here is that we **put a 2-dimensional ball around each point, where all such balls have the same radius**.
- For each different radius, the homology can be **vastly different**, with different cycles in the homology basis corresponding to the different radii
 - We focus on the **1-cycles** (1-dimensional holes) in the example
 - For each radius, the **colored** cycles form the homology basis

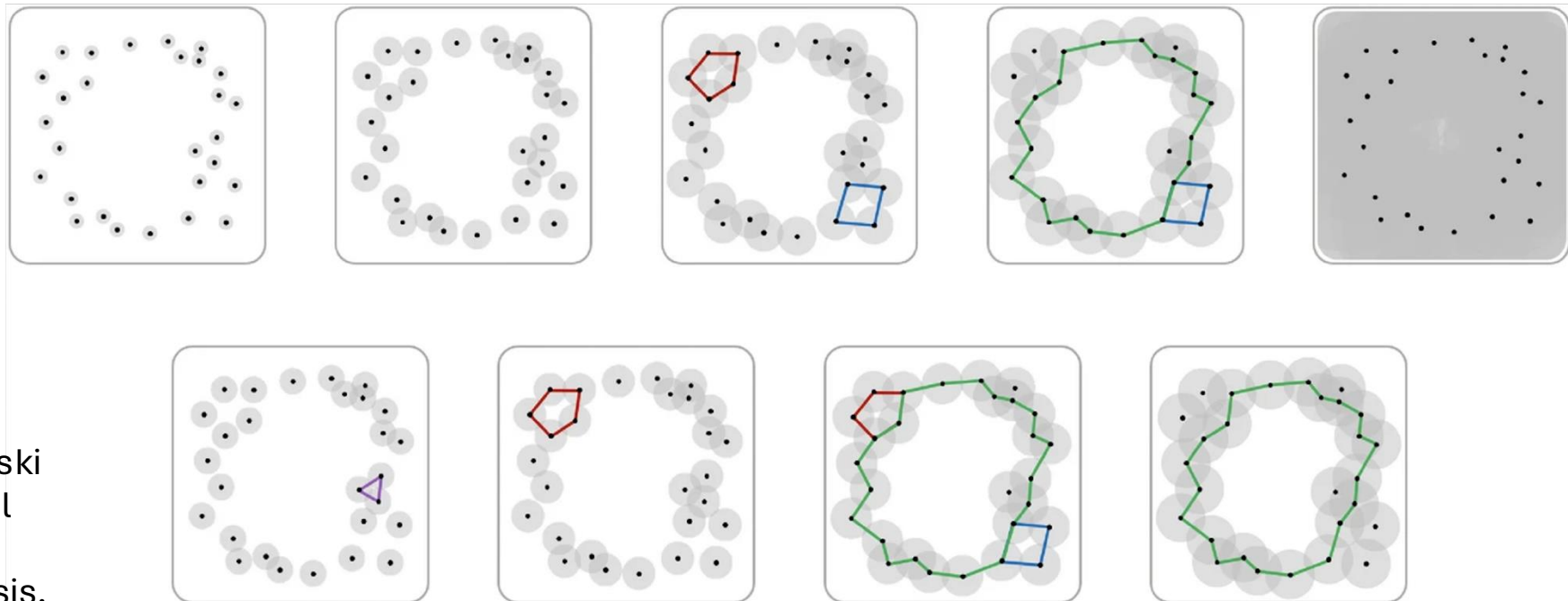


Image source: Bobrowski
O, Skraba P. A universal
null-distribution for
topological data analysis.

Motivation: Homology inference from points cloud

- **Question:** What is a **correct radius** to infer the shape of the point cloud?

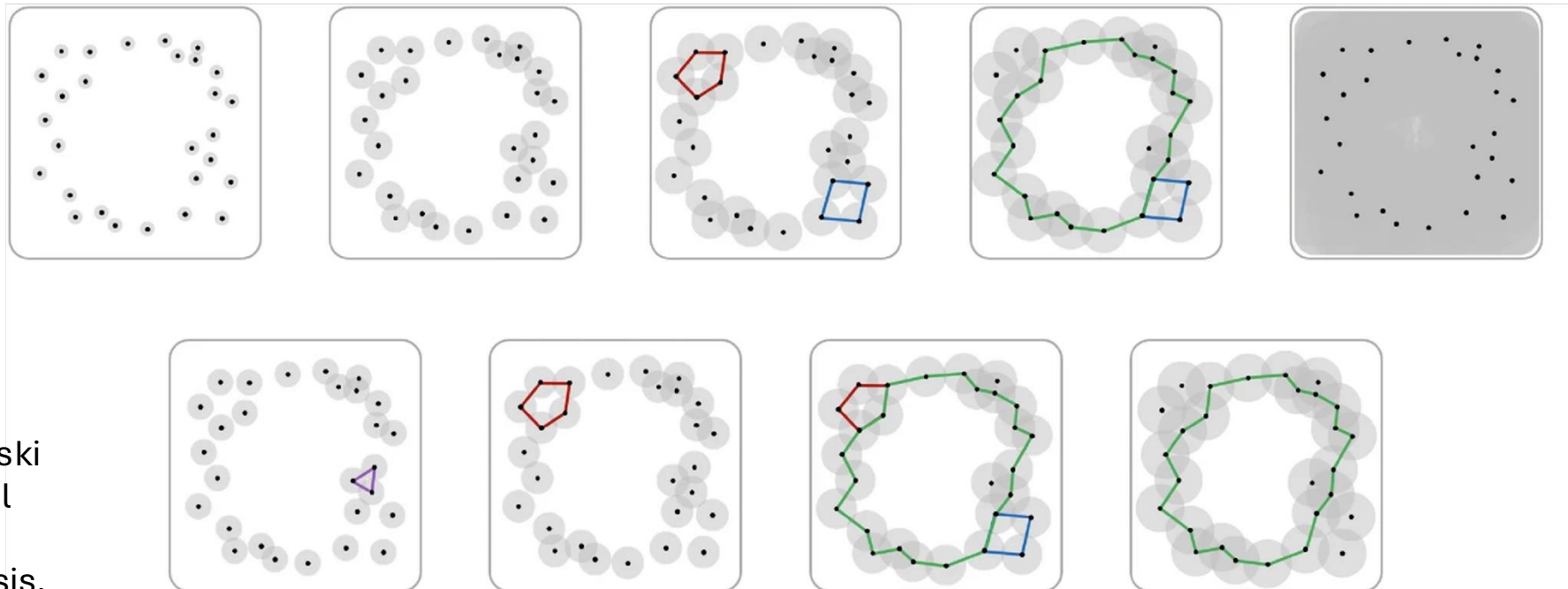


Image source: Bobrowski
O, Skraba P. A universal
null-distribution for
topological data analysis.

Motivation: Homology inference from points cloud

- **Question:** What is a **correct radius** to infer the shape of the point cloud?
- **Answer:** It's really hard to know, and there probably is no such “**correct**” radius

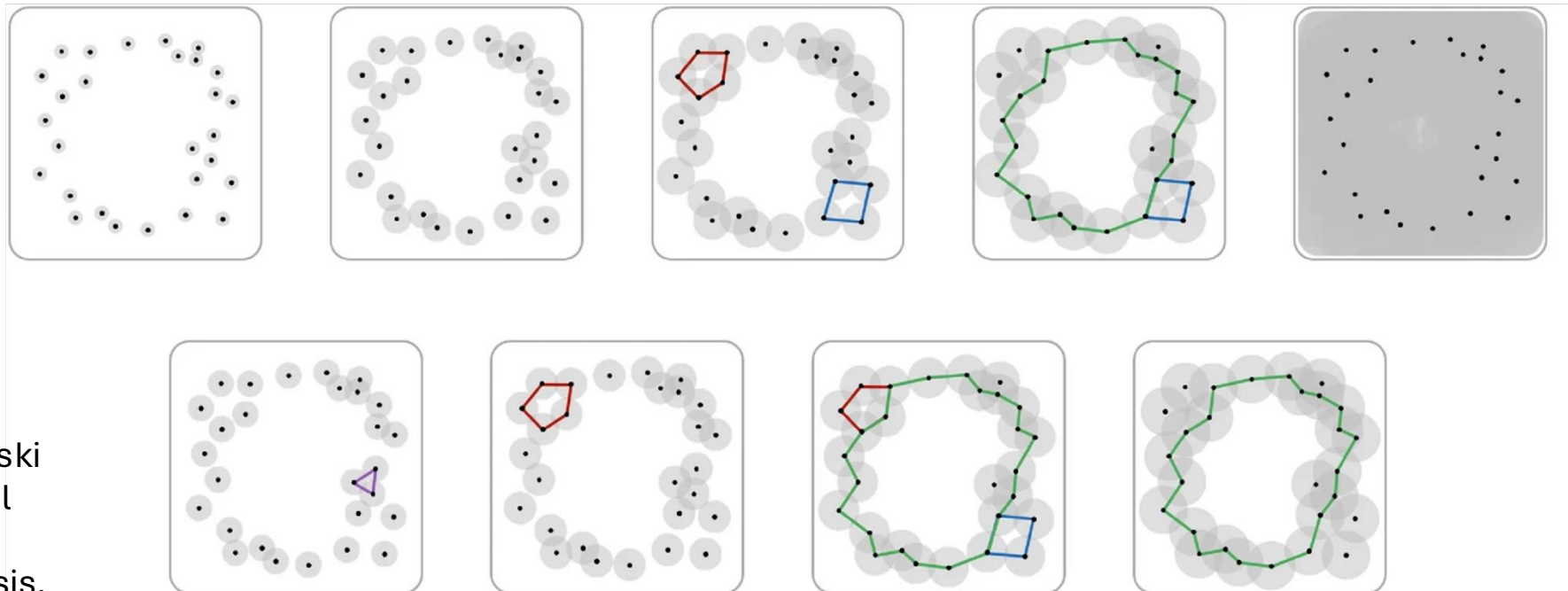


Image source: Bobrowski
O, Skraba P. A universal
null-distribution for
topological data analysis.

Motivation: Homology inference from points cloud

- **Solution:** Consider **all radius**, and **track the changes of the 1-cycles in the homology basis as we increase the radius**

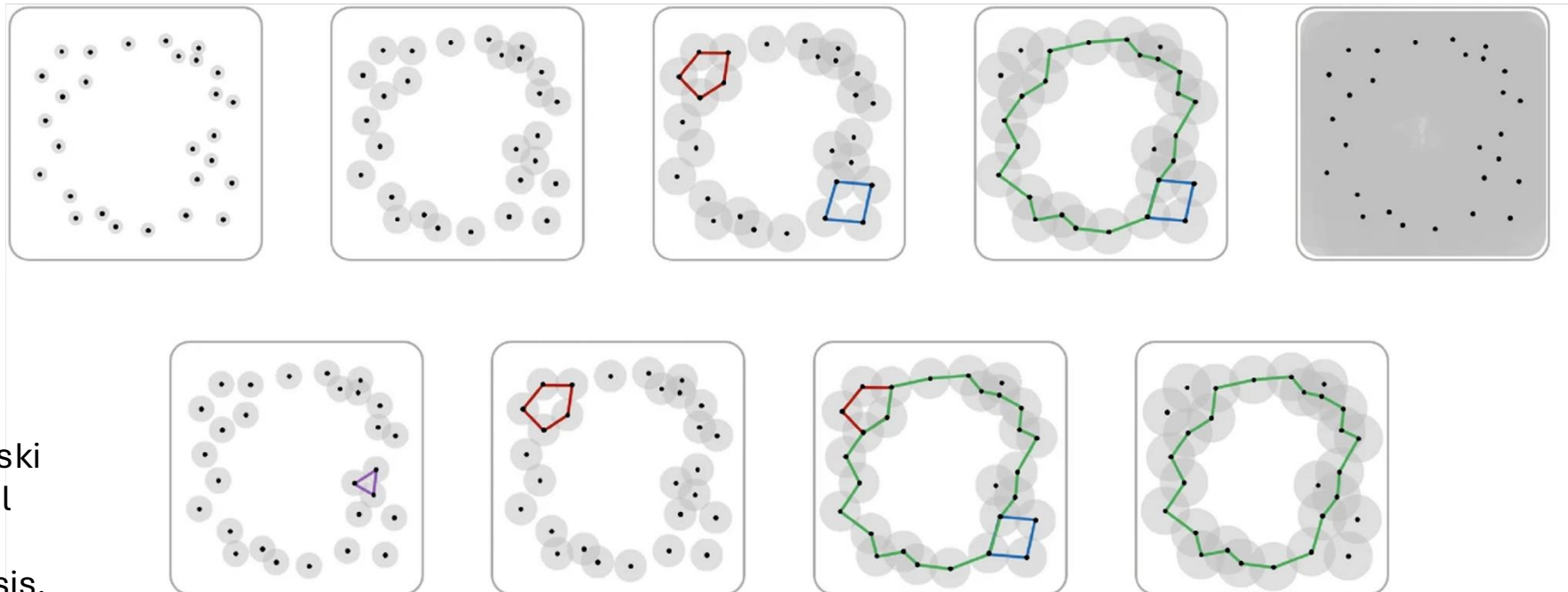


Image source: Bobrowski
O, Skraba P. A universal
null-distribution for
topological data analysis.

Motivation: Homology inference from points cloud

- **Solution:** Consider **all radius**, and **track the changes of the 1-cycles in the homology basis as we increase the radius**
- As the radius increases, different cycles in the basis could **appear** (getting **born**) or **becomes trivial** (**dies**).

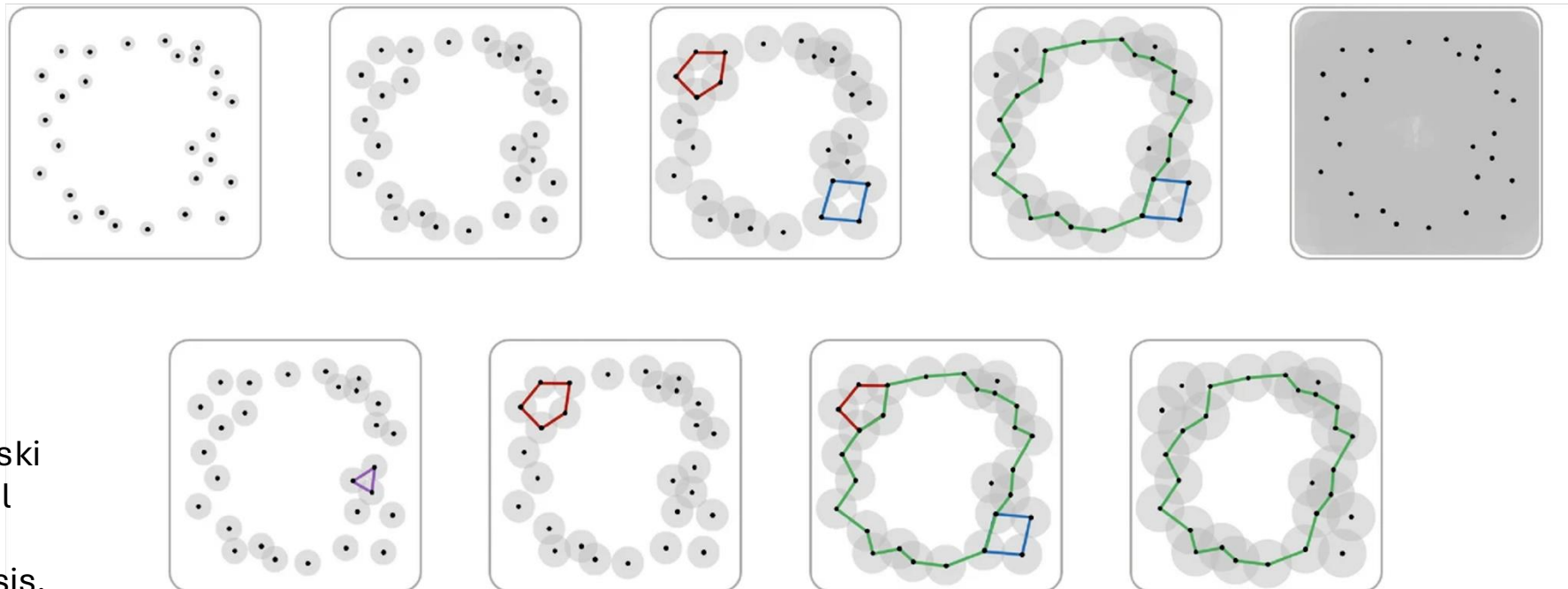


Image source: Bobrowski
O, Skraba P. A universal
null-distribution for
topological data analysis.

Motivation: Homology inference from points cloud

- **Solution:** Consider **all radius**, and **track the changes of the 1-cycles in the homology basis as we increase the radius**
- As the radius increases, different cycles in the basis could **appear** (getting **born**) or **becomes trivial** (**dies**).
- We **pair the births and deaths**, which are the points in the PD

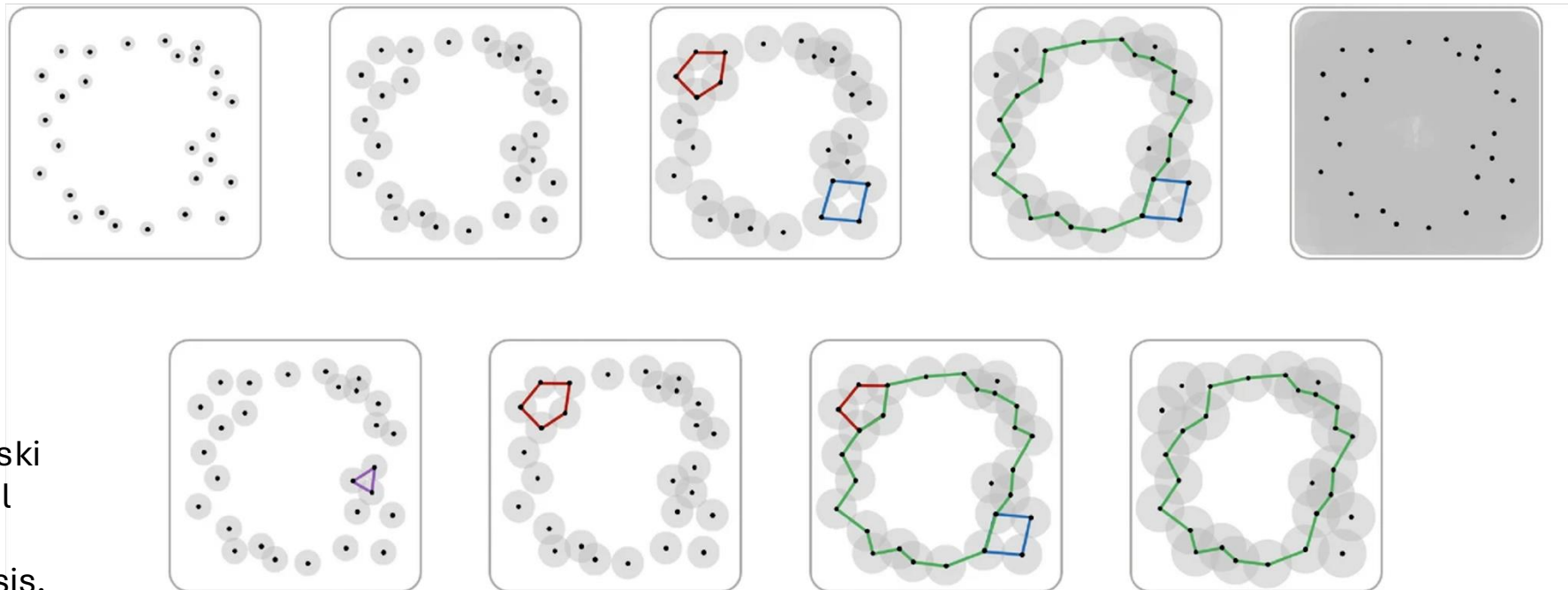
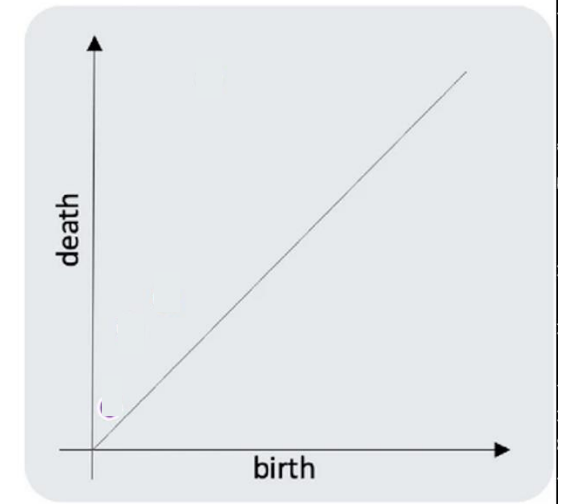
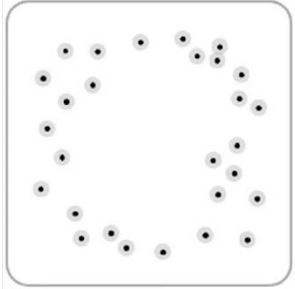


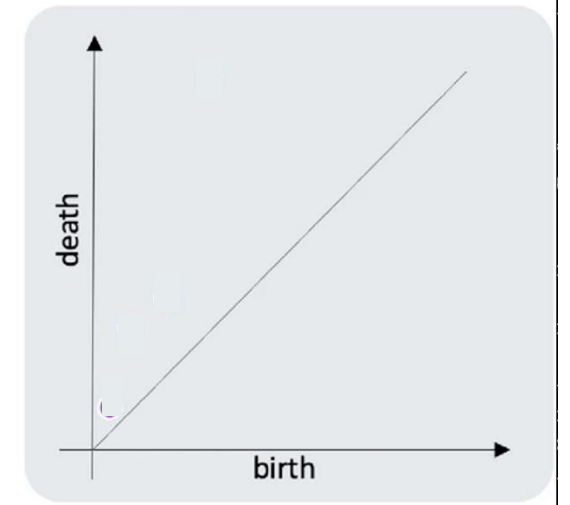
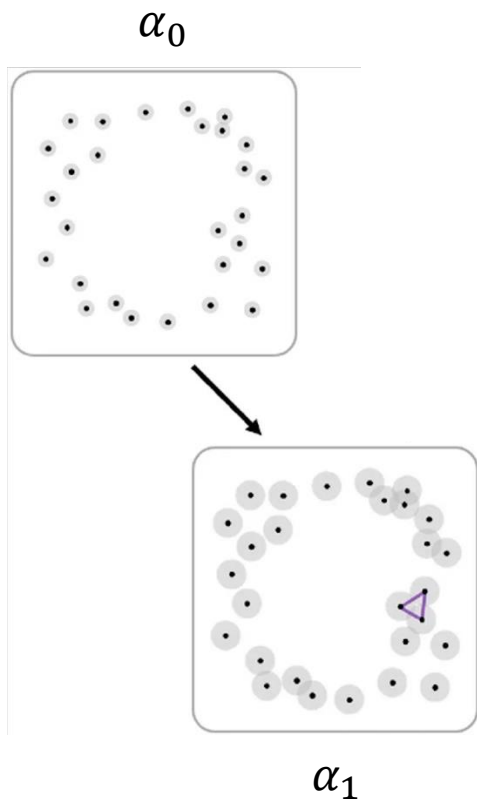
Image source: Bobrowski
O, Skraba P. A universal
null-distribution for
topological data analysis.

- α_0 : nothing happens.

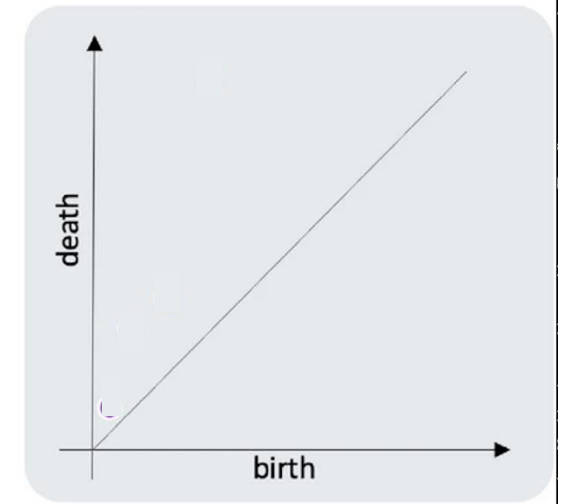
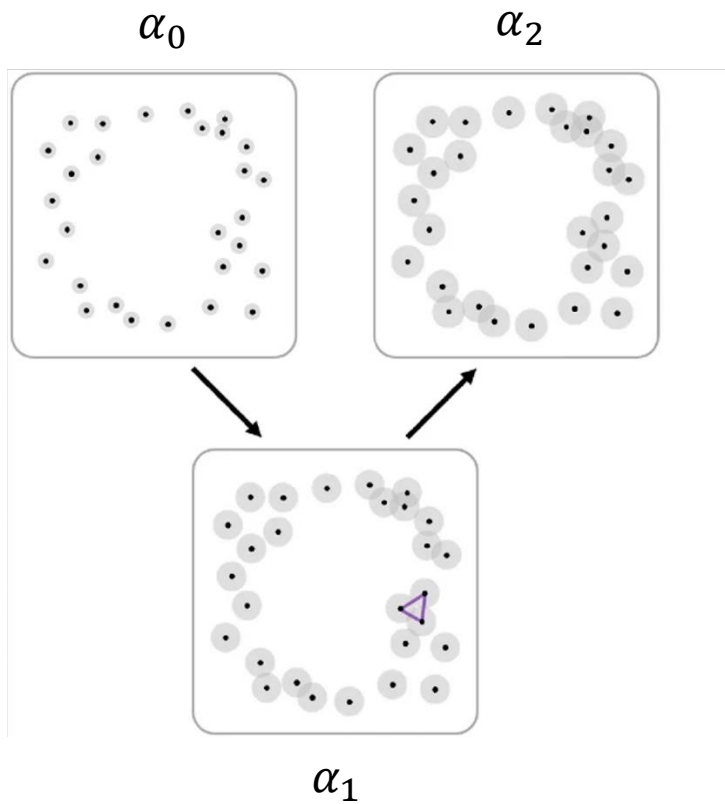
α_0



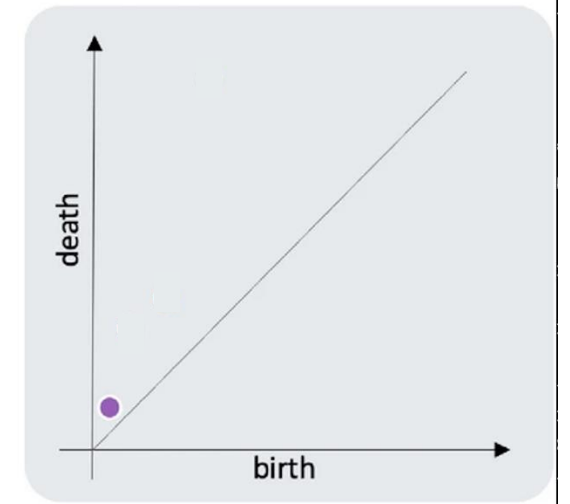
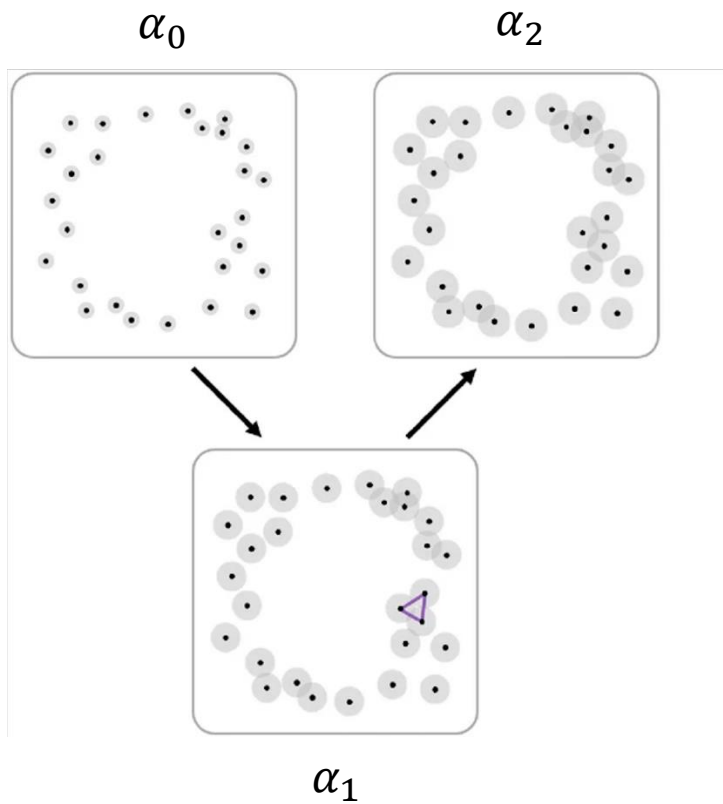
- α_0 : nothing happens.
- α_1 : purple cycle born



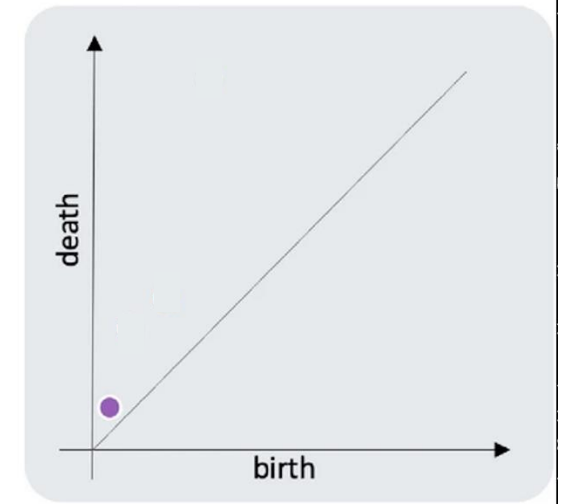
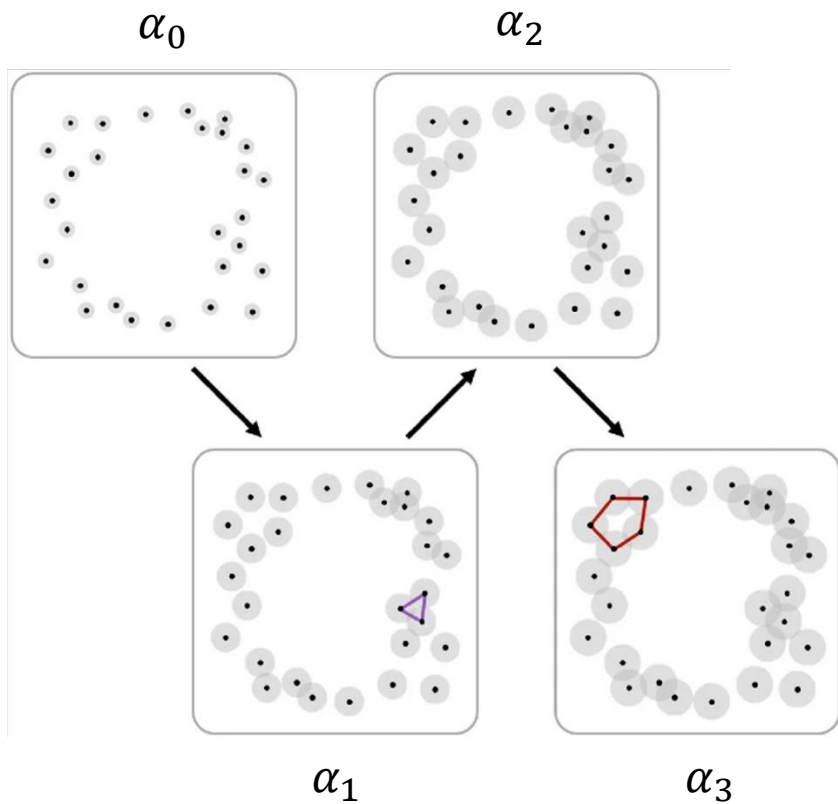
- α_0 : nothing happens.
- α_1 : purple cycle born
- α_2 : purple cycle dies



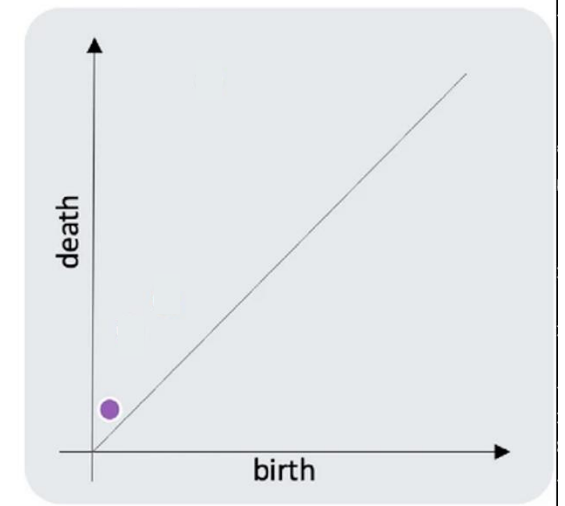
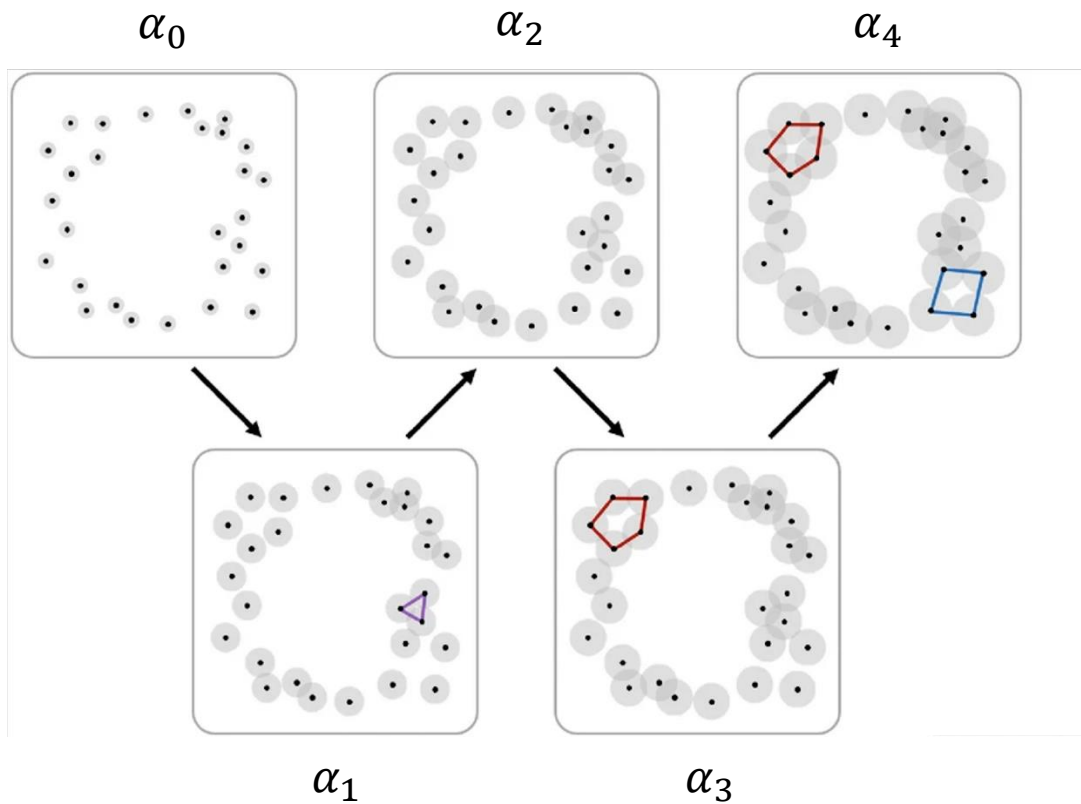
- α_0 : nothing happens.
 - α_1 : purple cycle born
 - α_2 : purple cycle dies
- $\Rightarrow (\alpha_1, \alpha_2)$



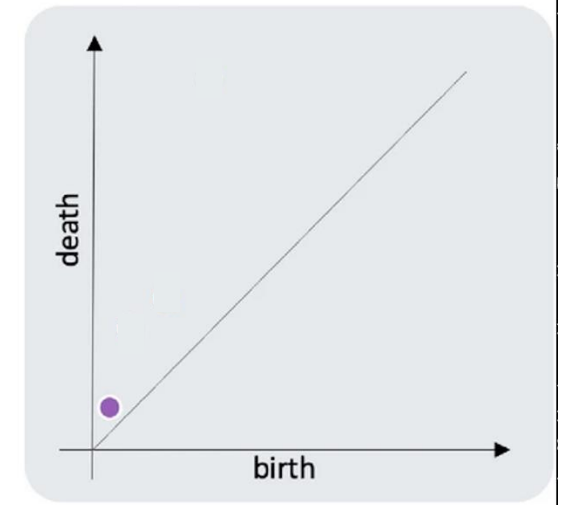
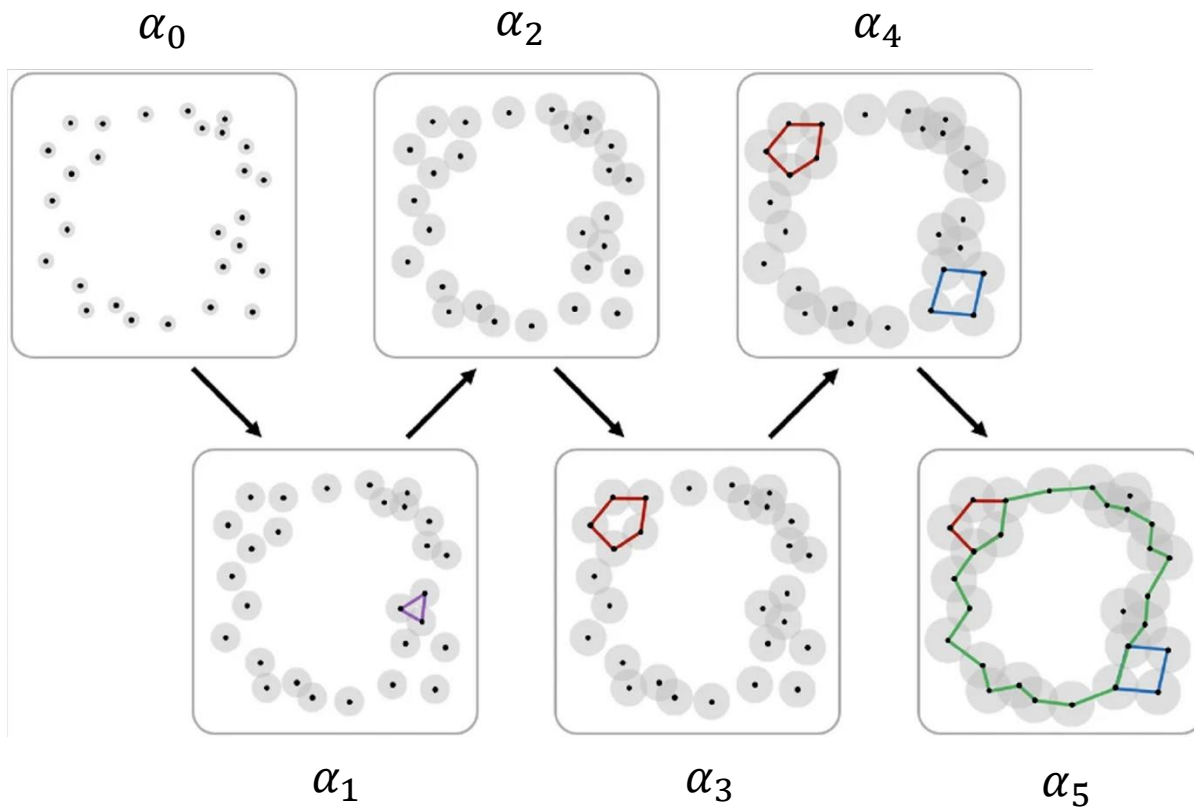
- α_0 : nothing happens.
 - α_1 : purple cycle born
 - α_2 : purple cycle dies
 - α_3 : red cycle born
- $\Rightarrow (\alpha_1, \alpha_2)$



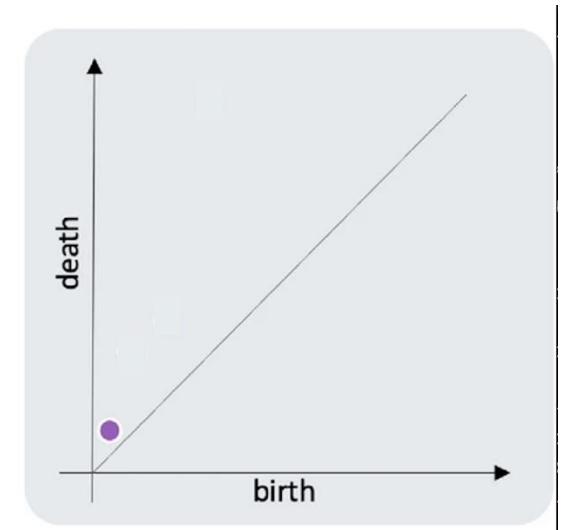
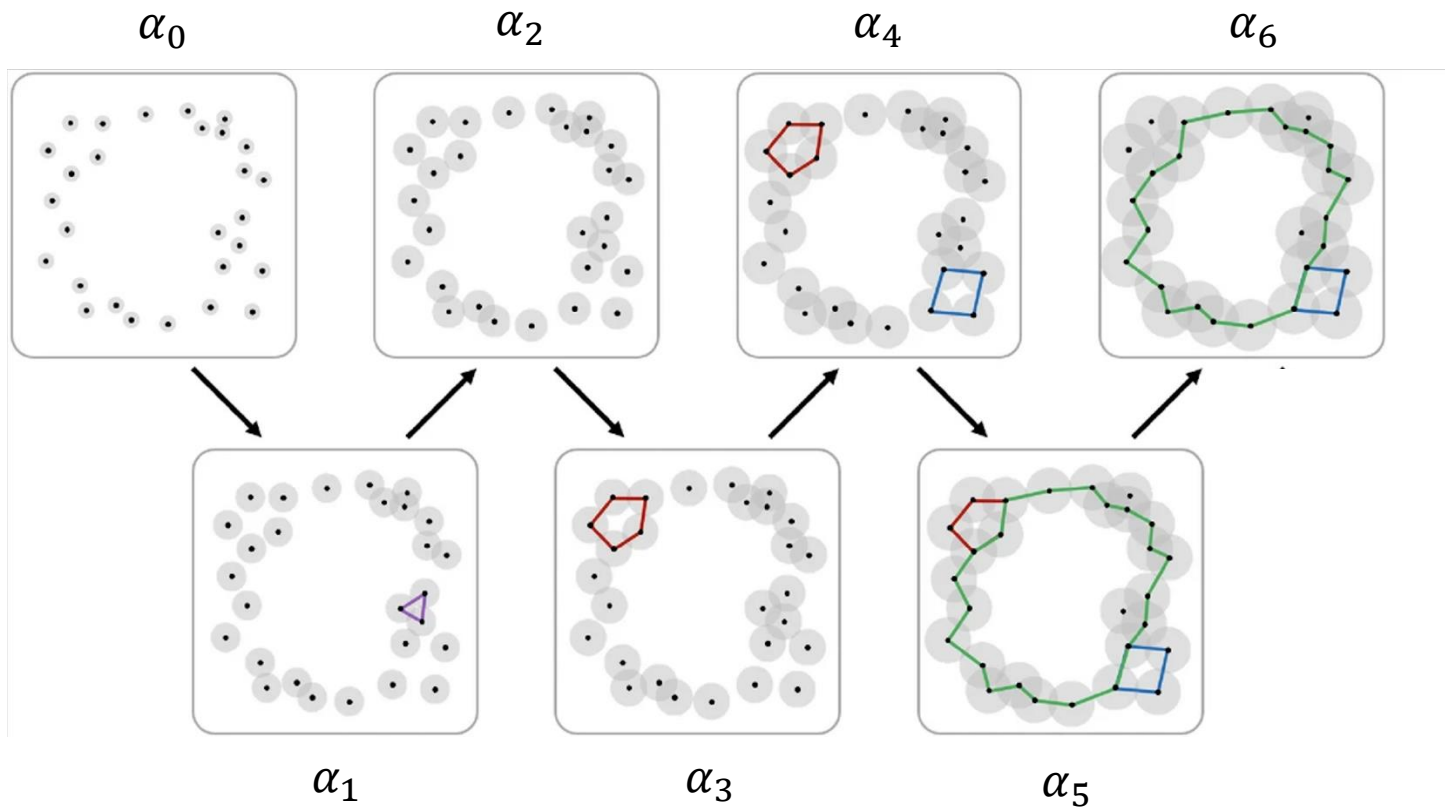
- α_0 : nothing happens.
 - α_1 : purple cycle born
 - α_2 : purple cycle dies
 - α_3 : red cycle born
 - α_4 : blue cycle born
- $\Rightarrow (\alpha_1, \alpha_2)$



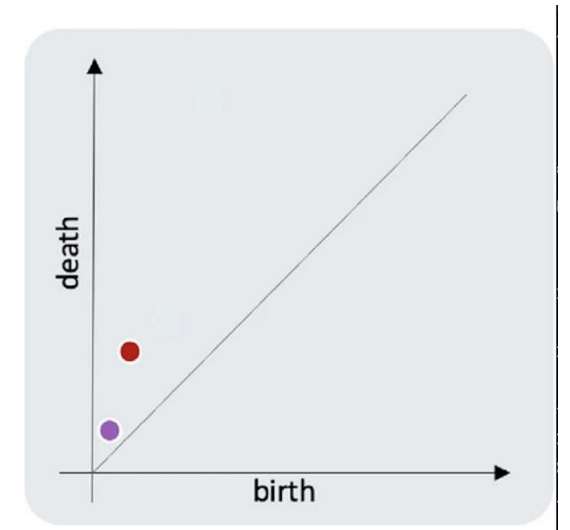
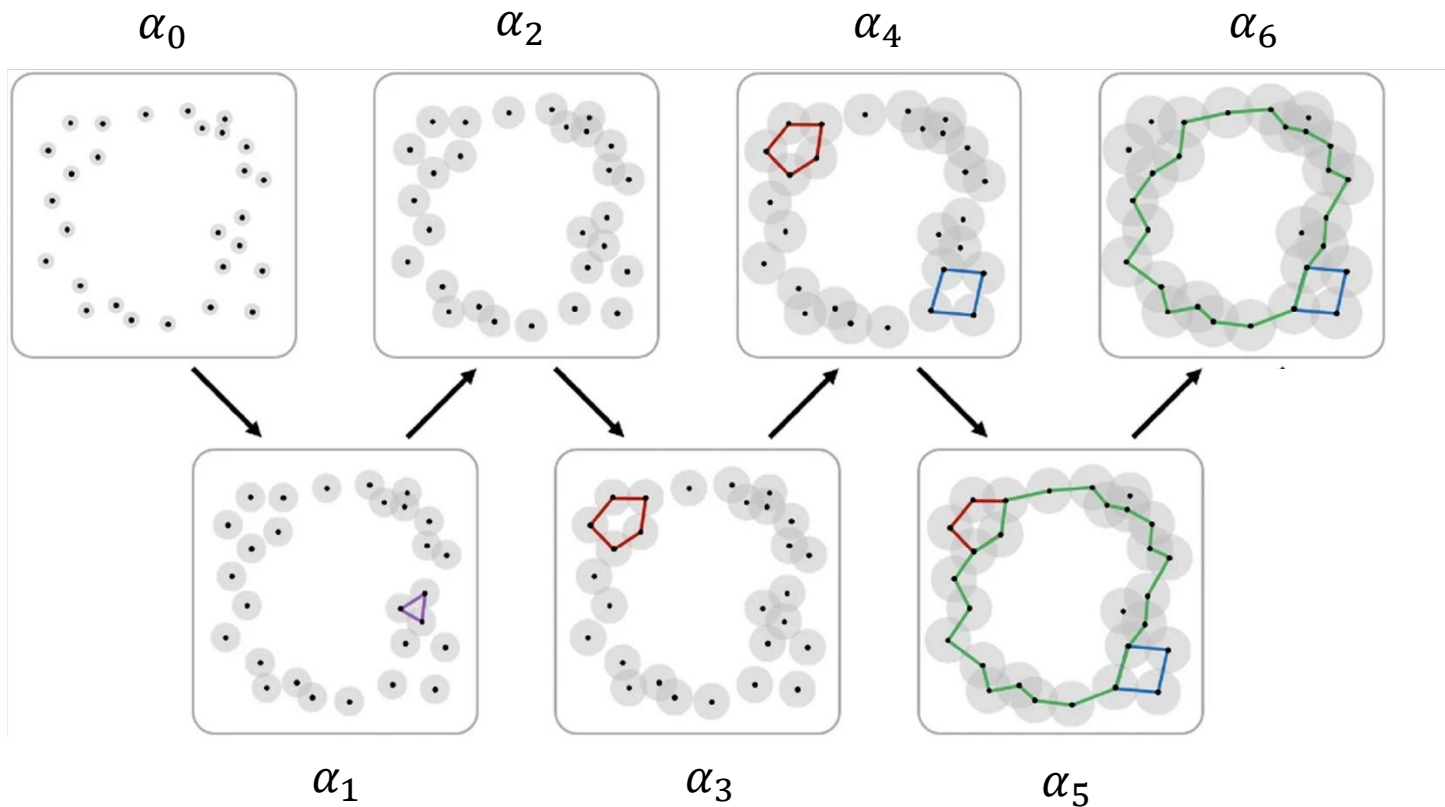
- α_0 : nothing happens.
 - α_1 : purple cycle born
 - α_2 : purple cycle dies
 - α_3 : red cycle born
 - α_4 : blue cycle born
 - α_5 : green cycle born
- $\Rightarrow (\alpha_1, \alpha_2)$



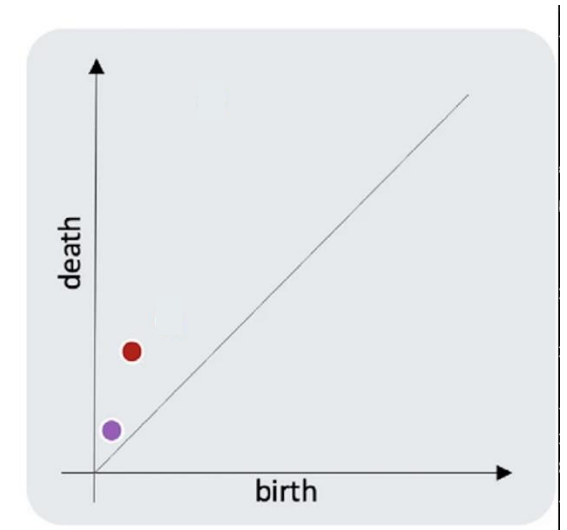
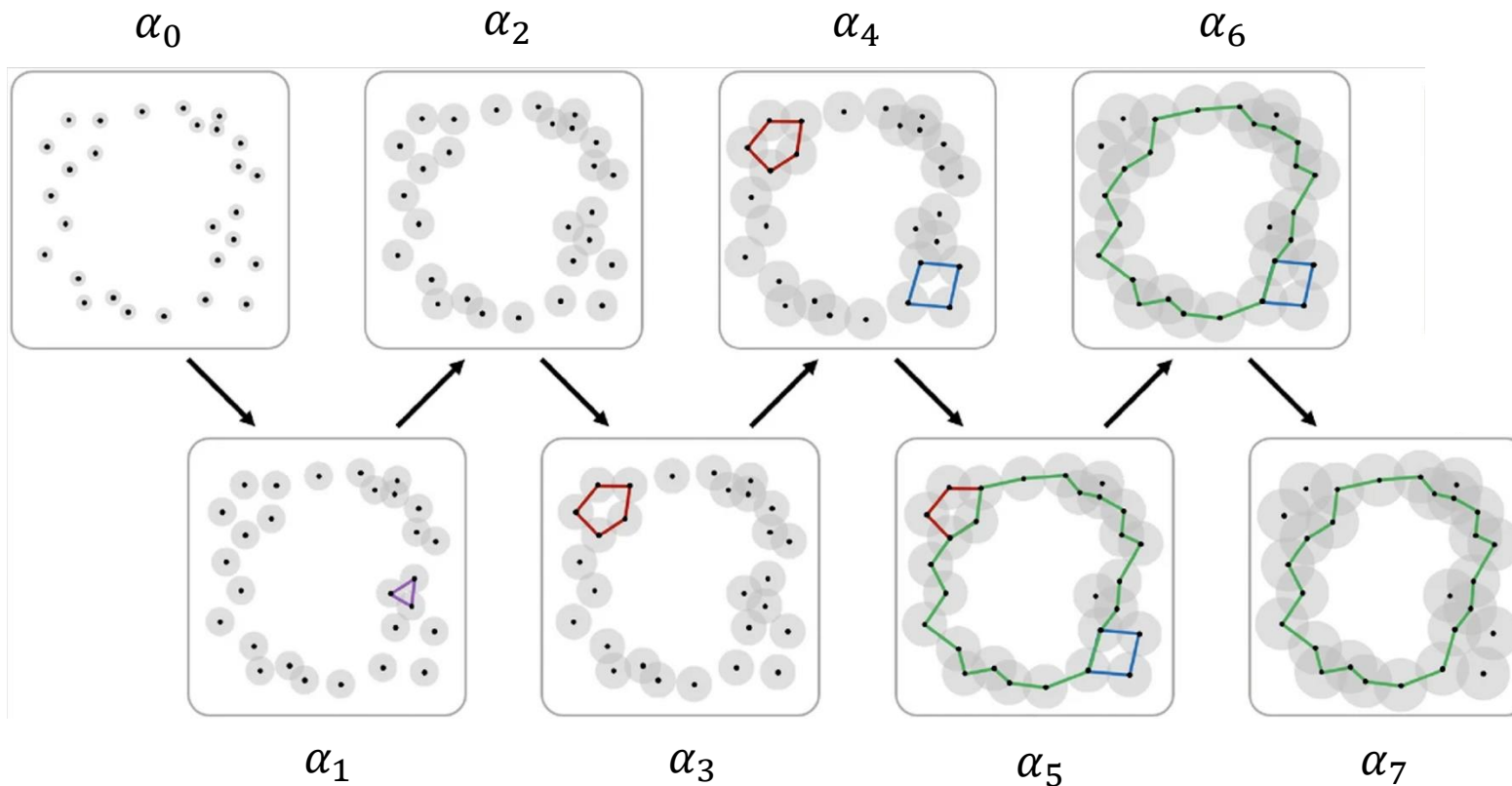
- α_0 : nothing happens.
 - α_1 : purple cycle born
 - α_2 : purple cycle dies
 - α_3 : red cycle born
 - α_4 : blue cycle born
 - α_5 : green cycle born
 - α_6 : red cycle dies
- $\Rightarrow (\alpha_1, \alpha_2)$



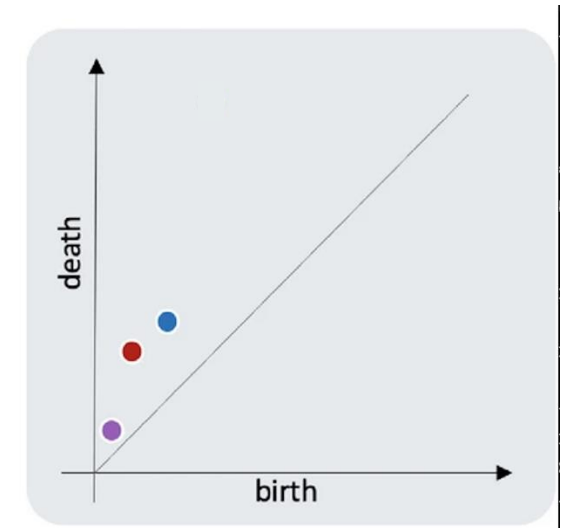
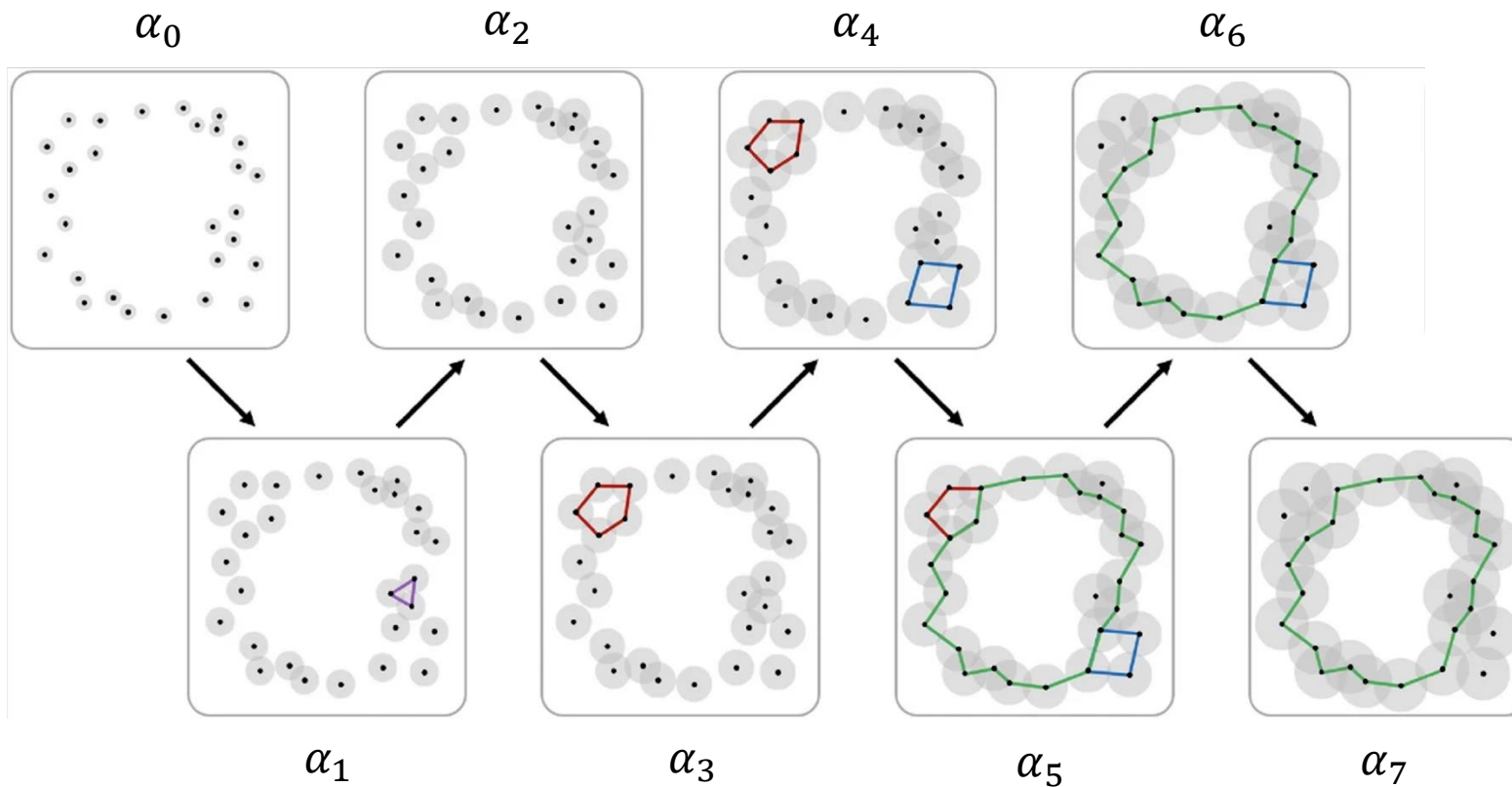
- α_0 : nothing happens.
- α_1 : purple cycle born
- α_2 : purple cycle dies $\Rightarrow (\alpha_1, \alpha_2)$
- α_3 : red cycle born
- α_4 : blue cycle born
- α_5 : green cycle born
- α_6 : red cycle dies $\Rightarrow (\alpha_3, \alpha_6)$



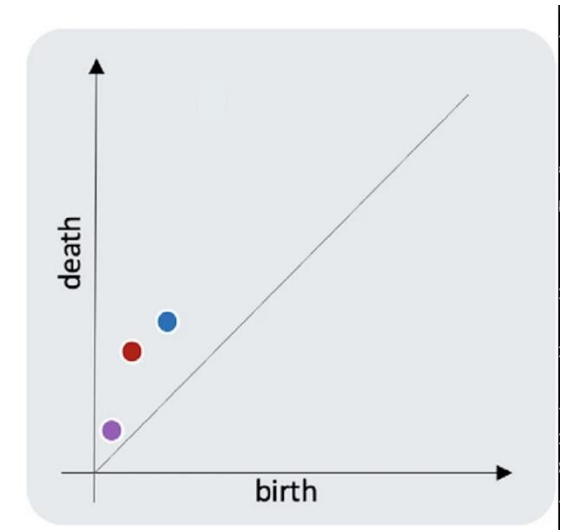
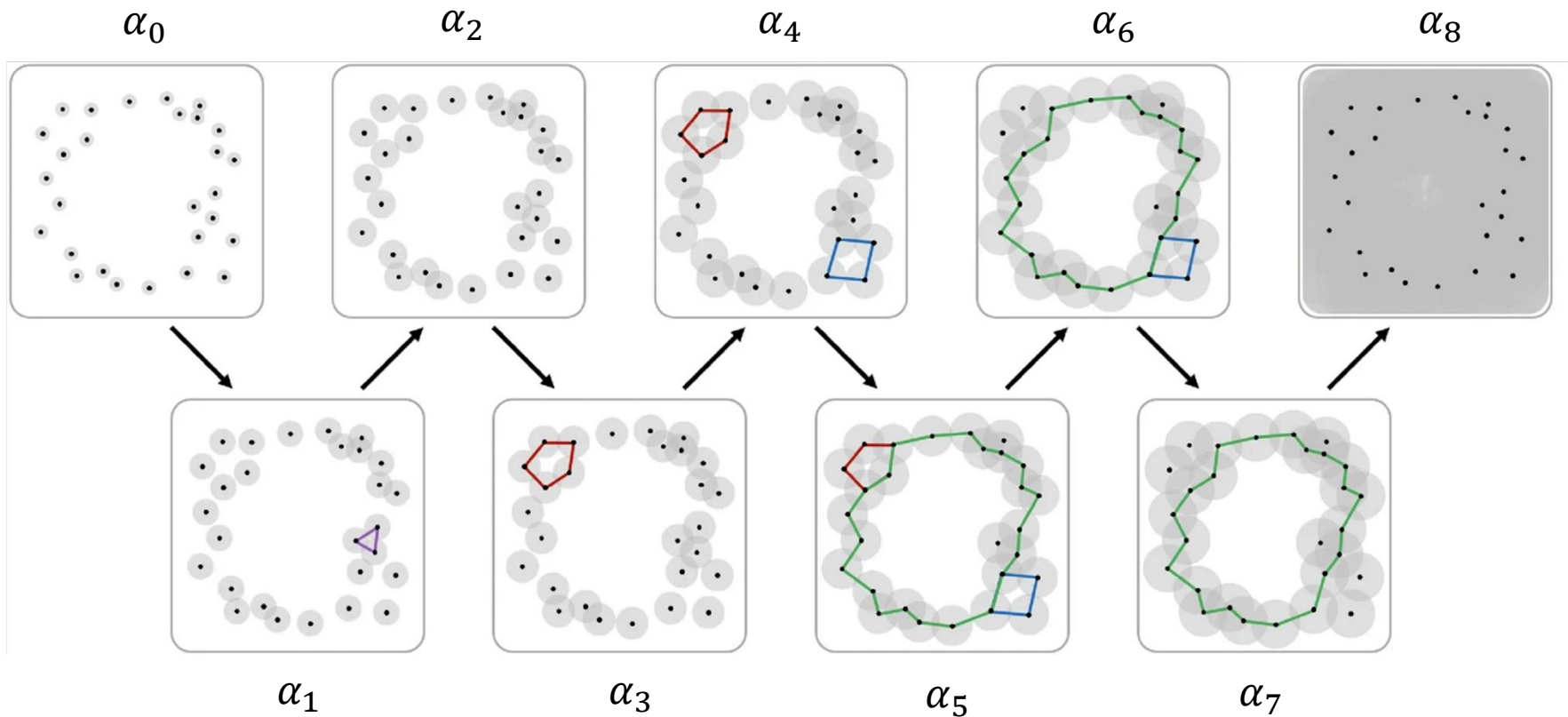
- α_0 : nothing happens.
- α_1 : purple cycle born
- α_2 : purple cycle dies $\Rightarrow (\alpha_1, \alpha_2)$
- α_3 : red cycle born
- α_4 : blue cycle born
- α_5 : green cycle born
- α_6 : red cycle dies $\Rightarrow (\alpha_3, \alpha_6)$
- α_7 : blue cycle dies



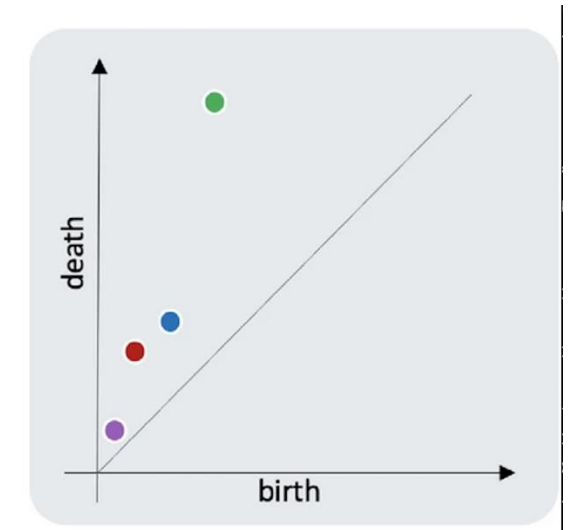
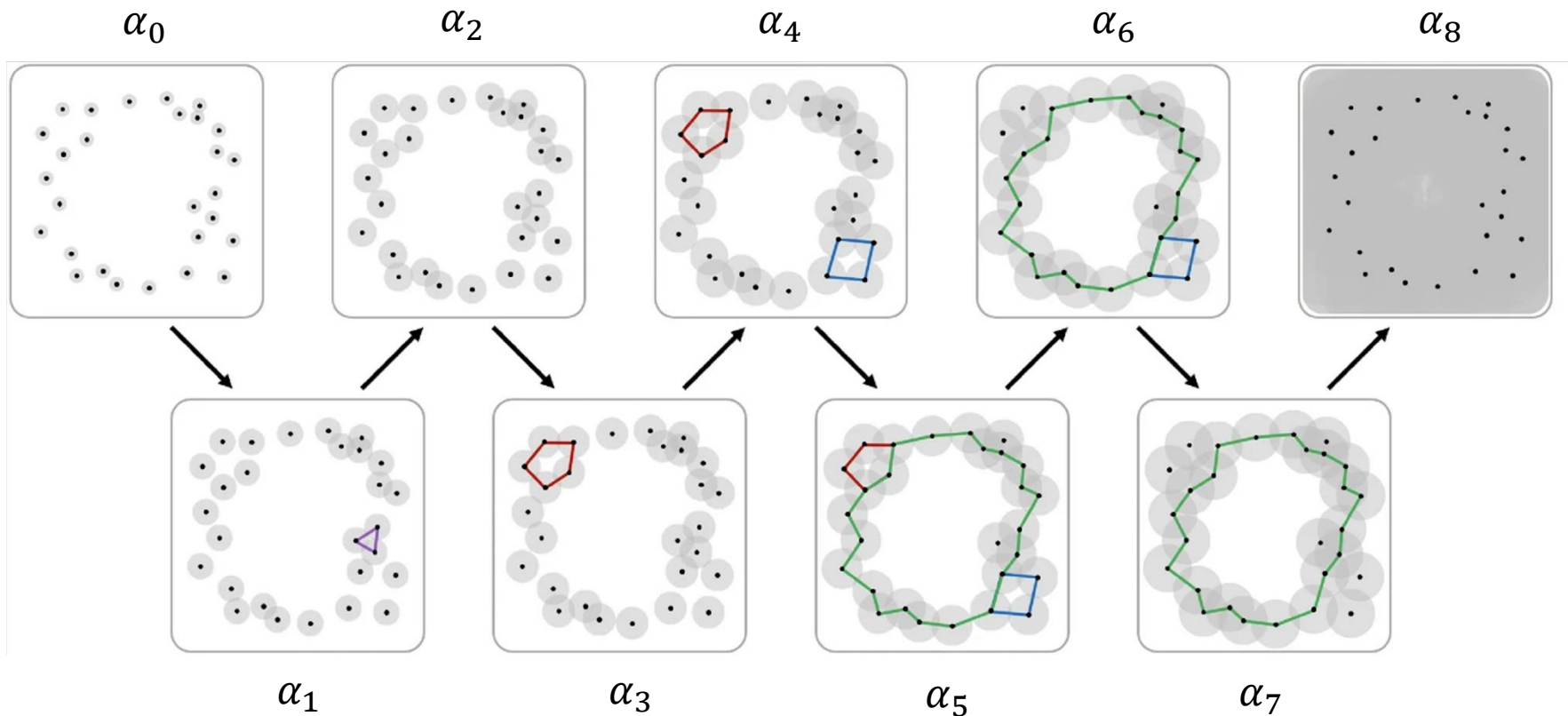
- α_0 : nothing happens.
- α_1 : purple cycle born
- α_2 : purple cycle dies $\Rightarrow (\alpha_1, \alpha_2)$
- α_3 : red cycle born
- α_4 : blue cycle born
- α_5 : green cycle born
- α_6 : red cycle dies $\Rightarrow (\alpha_3, \alpha_6)$
- α_7 : blue cycle dies $\Rightarrow (\alpha_4, \alpha_7)$



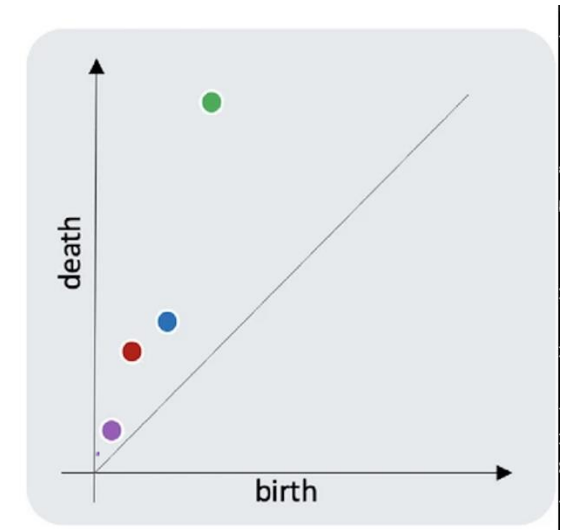
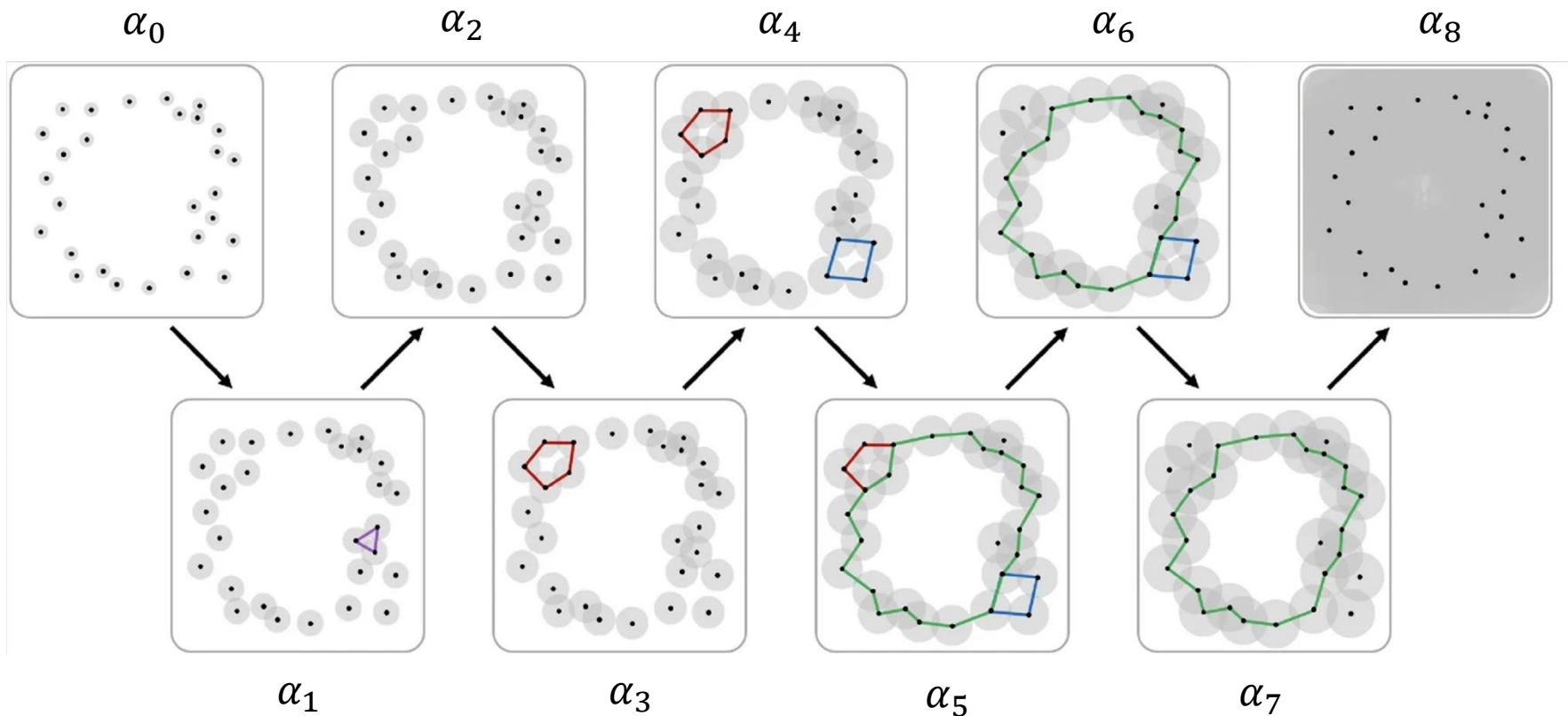
- α_0 : nothing happens.
- α_1 : purple cycle born
- α_2 : purple cycle dies $\Rightarrow (\alpha_1, \alpha_2)$
- α_3 : red cycle born
- α_4 : blue cycle born
- α_5 : green cycle born
- α_6 : red cycle dies $\Rightarrow (\alpha_3, \alpha_6)$
- α_7 : blue cycle dies $\Rightarrow (\alpha_4, \alpha_7)$
- α_8 : green cycle dies



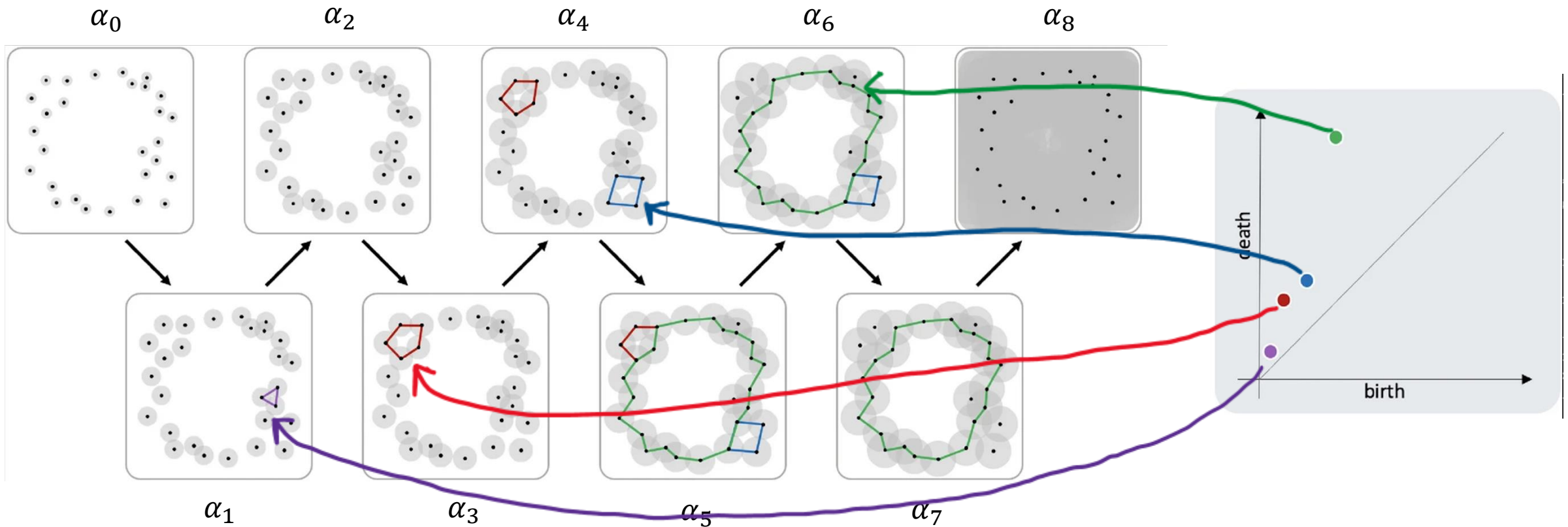
- α_0 : nothing happens.
- α_1 : purple cycle born
- α_2 : purple cycle dies $\Rightarrow (\alpha_1, \alpha_2)$
- α_3 : red cycle born
- α_4 : blue cycle born
- α_5 : green cycle born
- α_6 : red cycle dies $\Rightarrow (\alpha_3, \alpha_6)$
- α_7 : blue cycle dies $\Rightarrow (\alpha_4, \alpha_7)$
- α_8 : green cycle dies $\Rightarrow (\alpha_5, \alpha_8)$



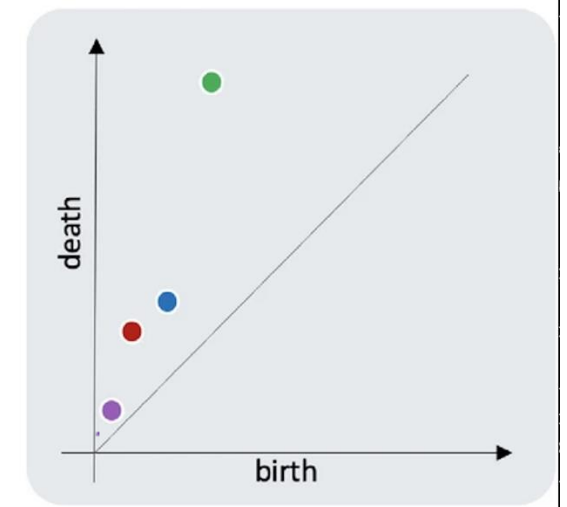
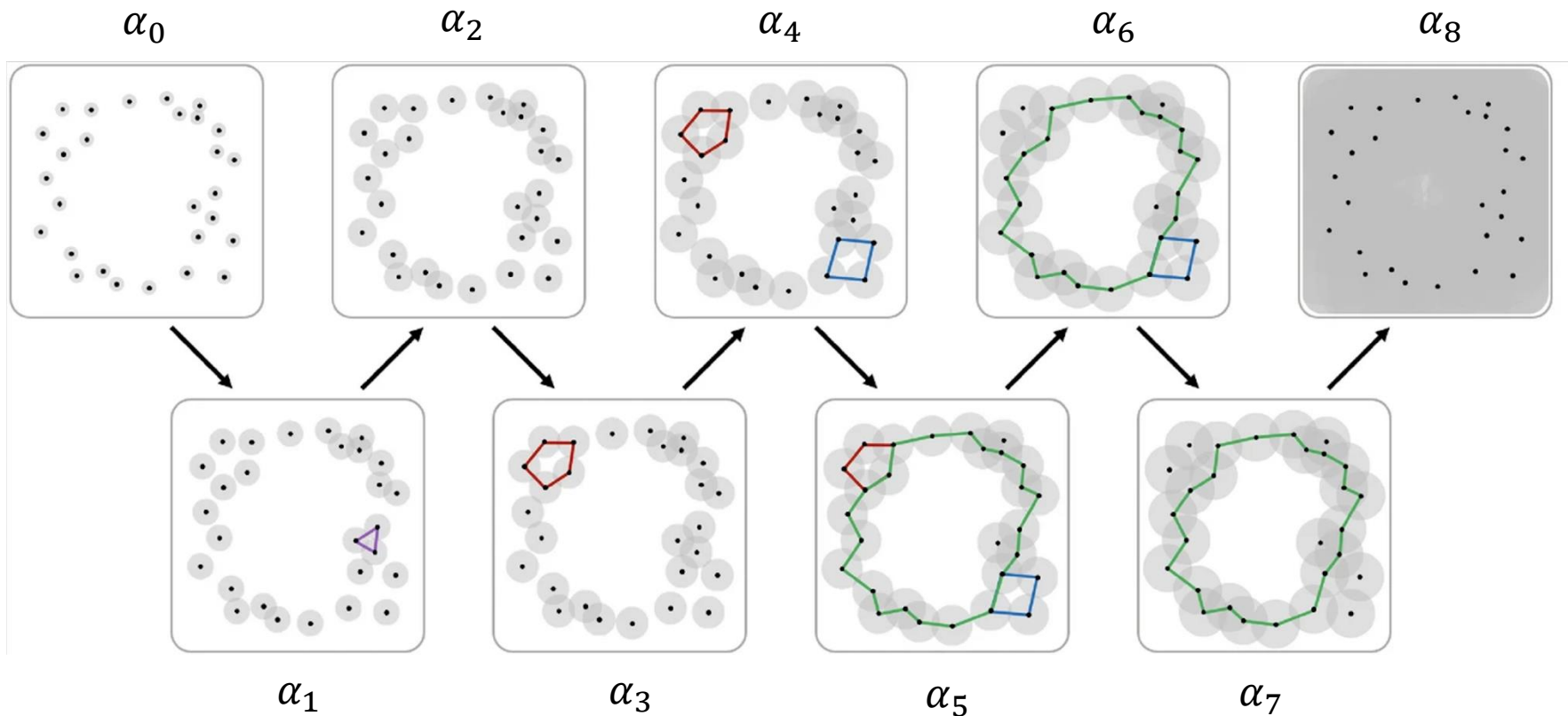
- So we have a 1-dimensional PD on the left with the four points **corresponding to the different cycles born and died** in the growing spaces with different α value, matching the colors



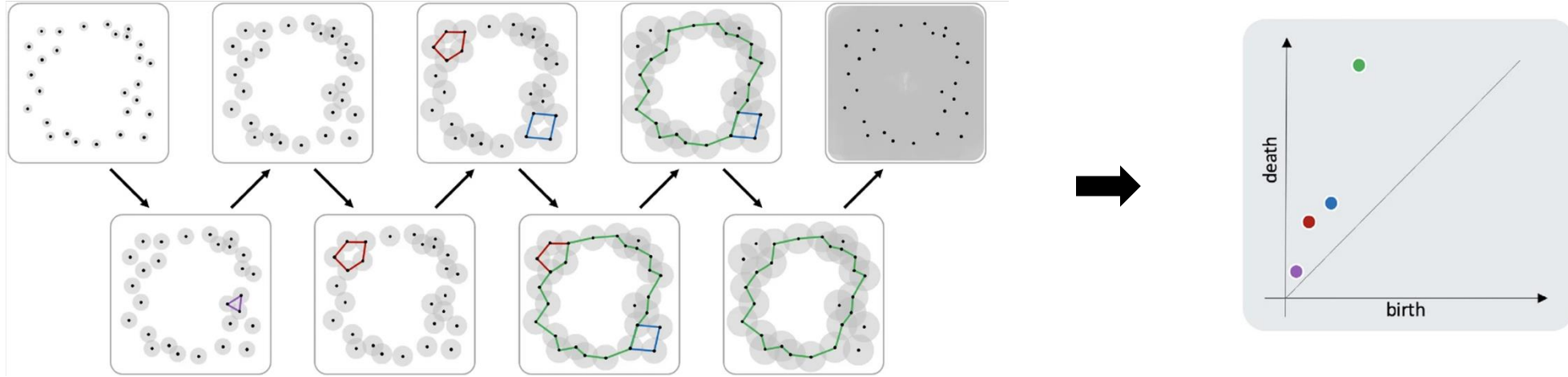
- So we have a 1-dimensional PD on the left with the four points **corresponding to the different cycles born and died** in the growing spaces with different α value, matching the colors



- Furthermore, we have that **distances of the points to diagonal indicate the difference of birth and death** (how long a cycle persist), which in turn **indicate the significance of the feature**



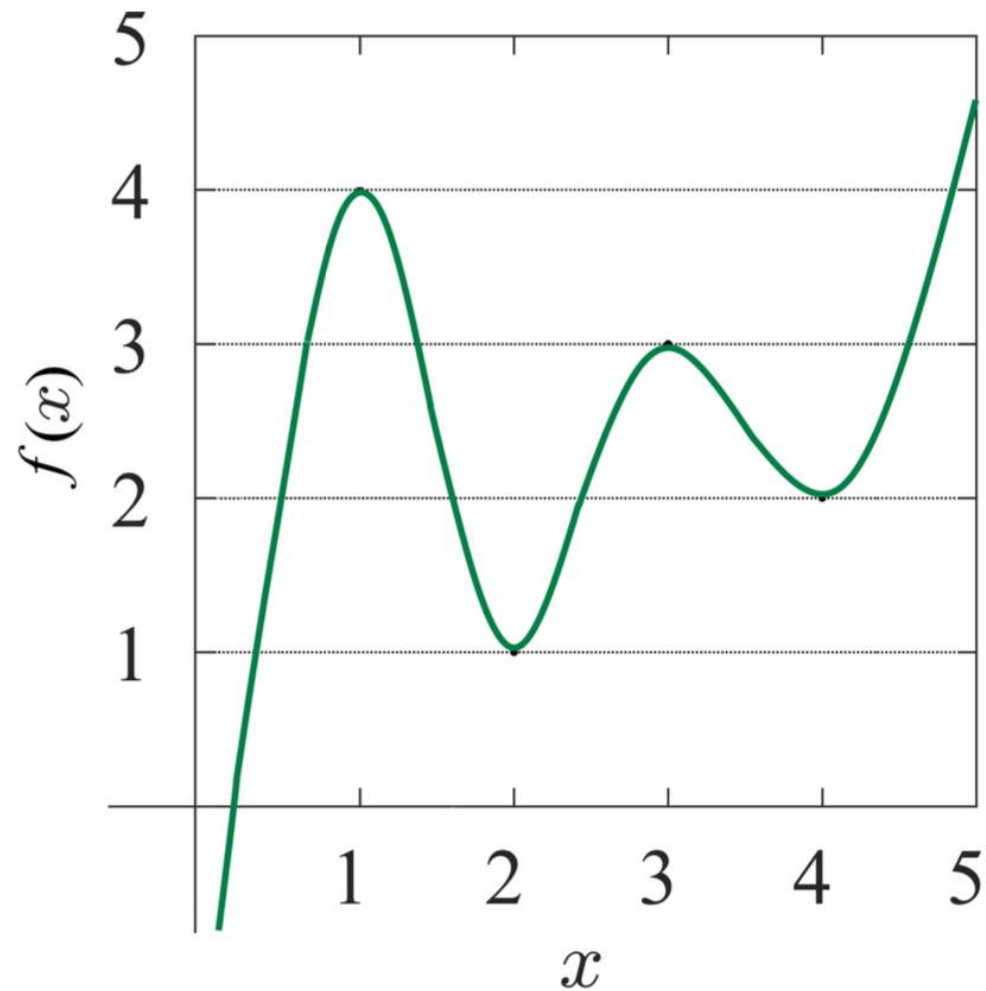
Persistent homology: Brief Summary



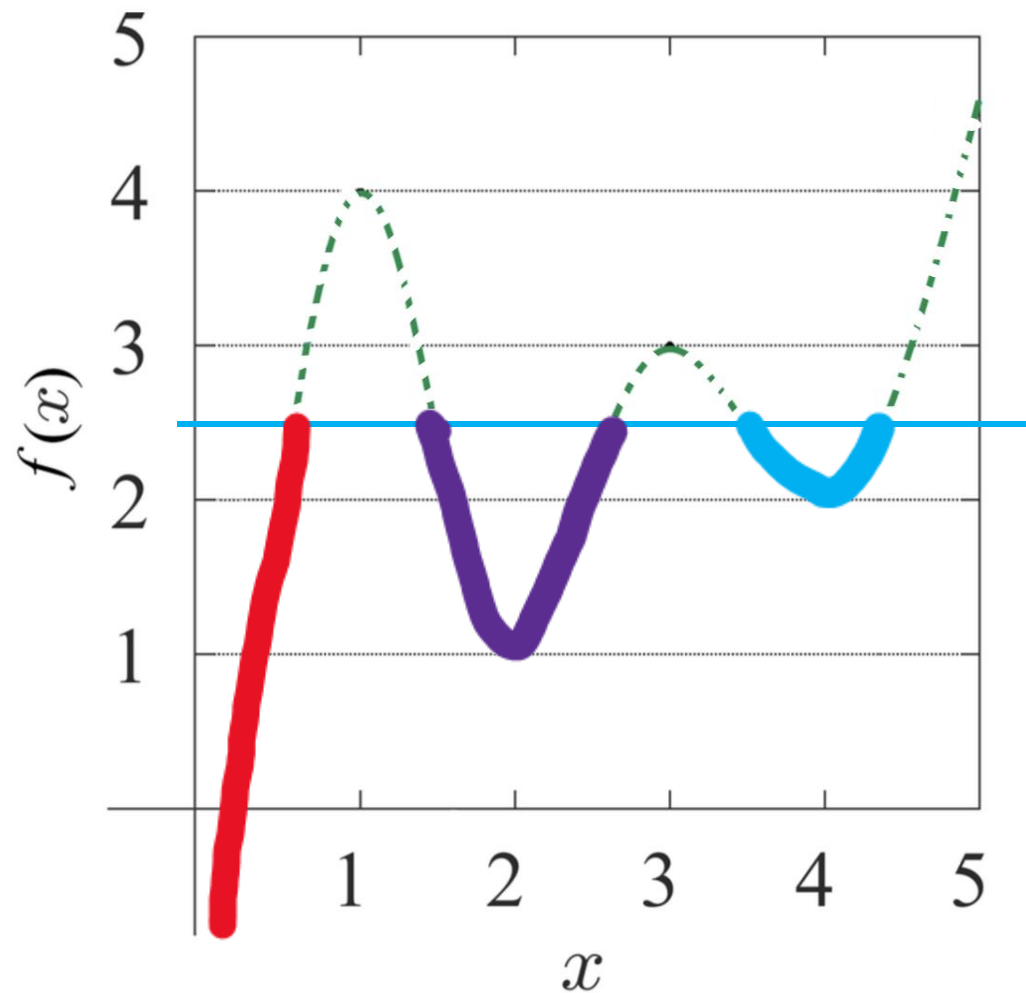
- Given a **growing topological space**, produce a **set of points on the 2D plane** (above the diagonal) called **persistence diagram (PD)** such that:
 - each point in the PD represents a homological feature (aka. cycle / hole) of the data in a certain dimension.

Online resources

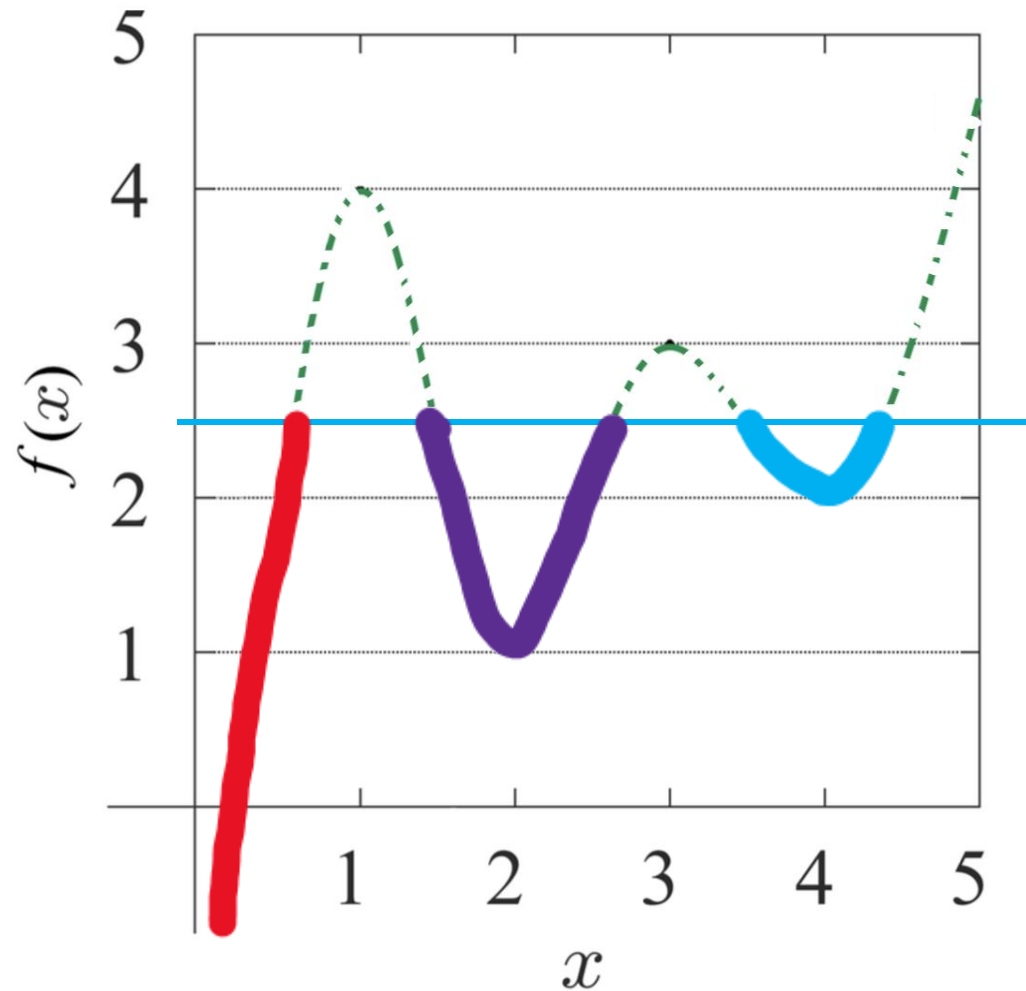
- A webpage for visualizing 1–dim PD: https://gjkoplik.github.io/pers-hom-examples/1d_pers_2d_data_widget.html



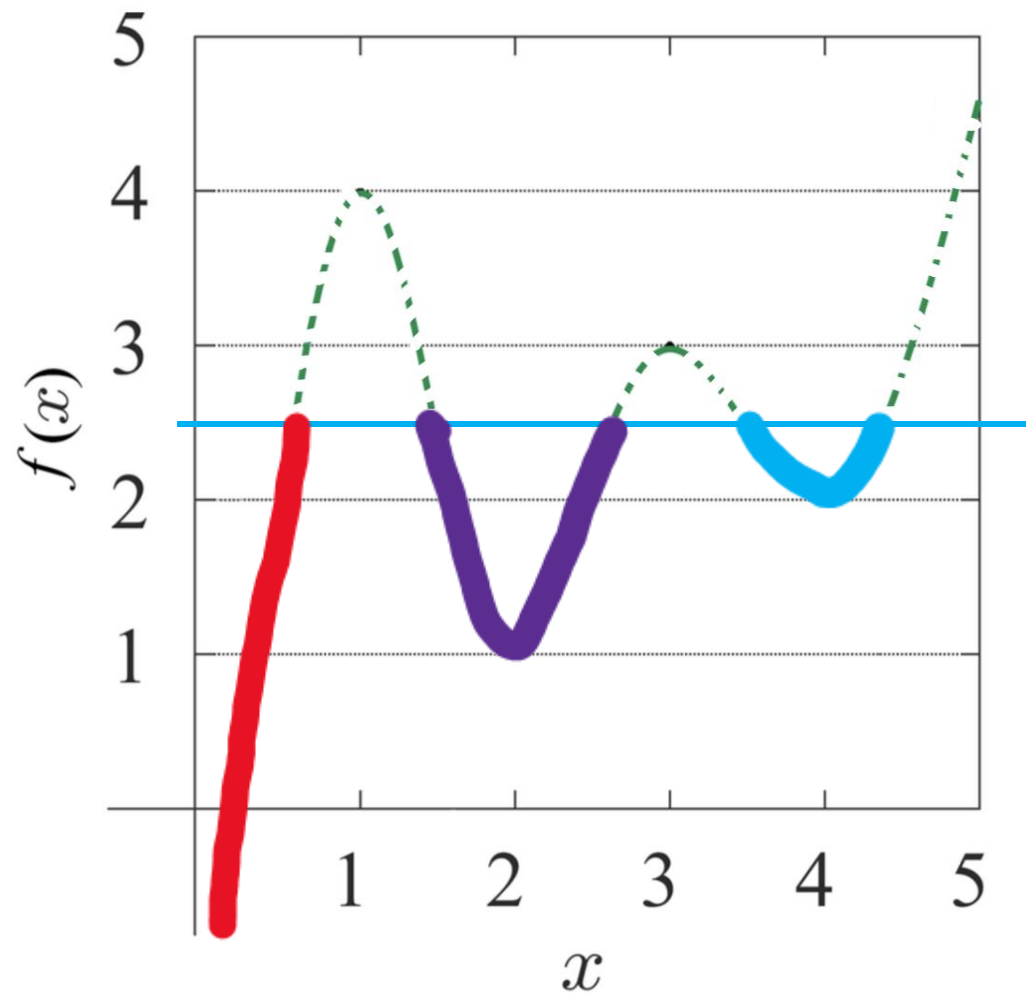
- For another example of persistent homology, we look at the left curve $y = f(x)$
- Again, we consider a growing space
- Each space in the growing sequence is part of the curve below a certain horizontal line



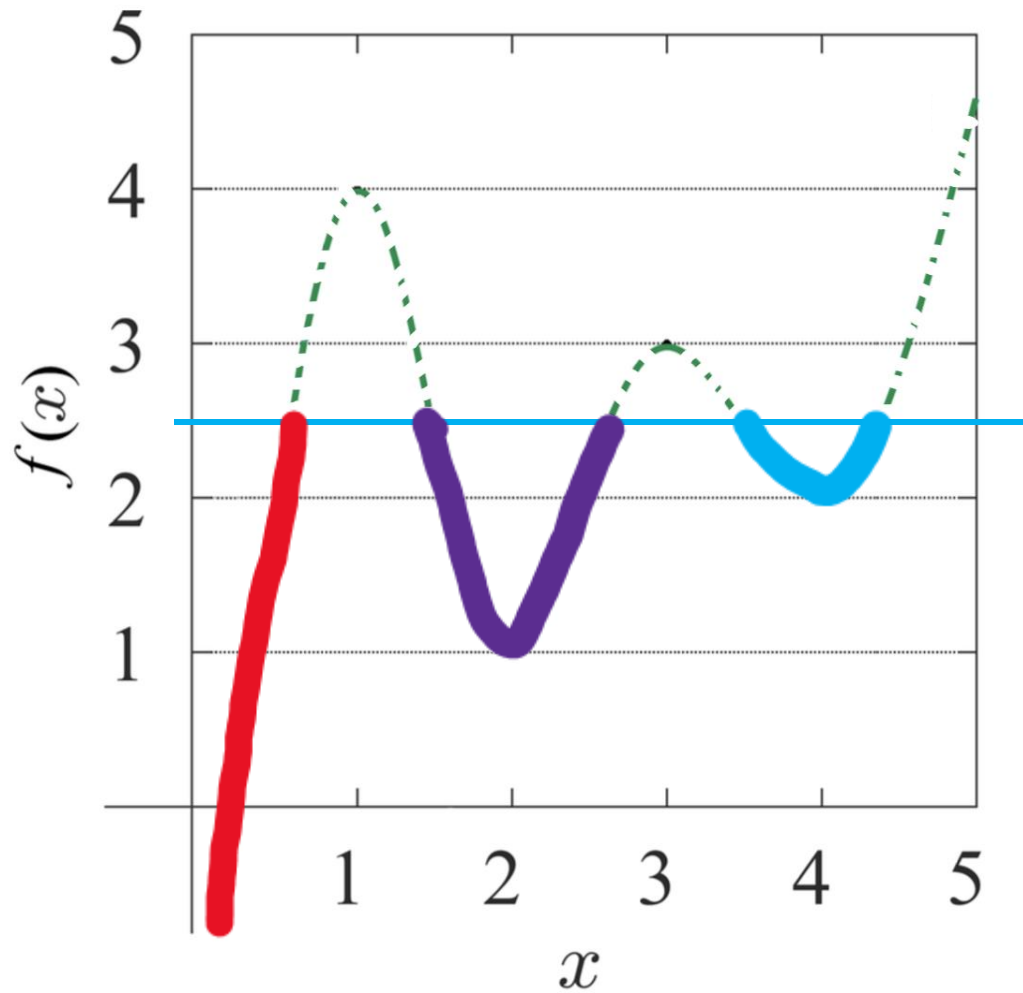
- For another example of persistent homology, we look at the left curve $y = f(x)$
- Again, we consider a growing space
- Each space in the growing sequence is part of the curve below a certain horizontal line
- Left is an example for horizontal line $y = 2.5$



- For another example of persistent homology, we look at the left curve $y = f(x)$
- Again, we consider a growing space
- Each space in the growing sequence is part of the curve below a certain horizontal line
- Left is an example for horizontal line $y = 2.5$
- As the space grows, we track the changes of *0-dimensional homology*

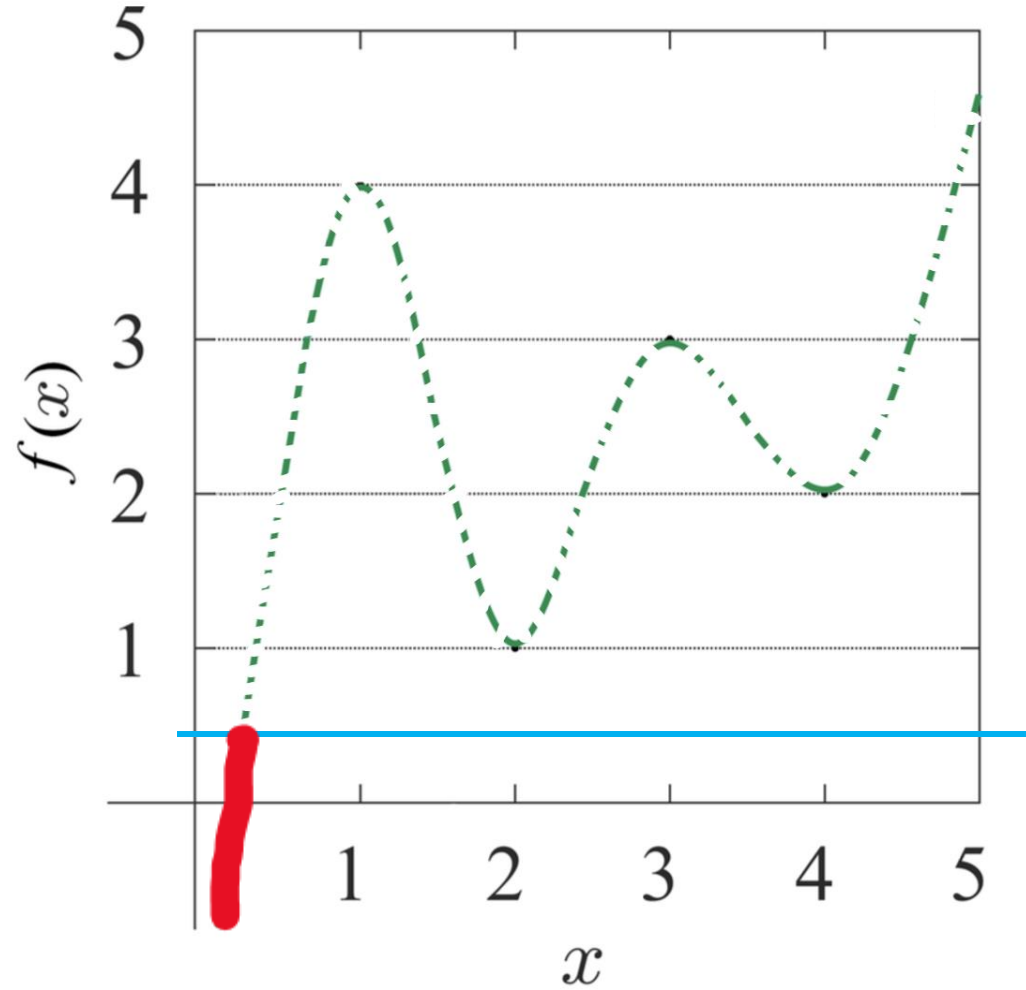


- For another example of persistent homology, we look at the left curve $y = f(x)$
- Again, we consider a growing space
- Each space in the growing sequence is part of the curve below a certain horizontal line
- Left is an example for horizontal line $y = 2.5$
- As the space grows, we track the changes of *0-dimensional homology*
- i.e., we track the **changes of the connected components and the gaps in between**



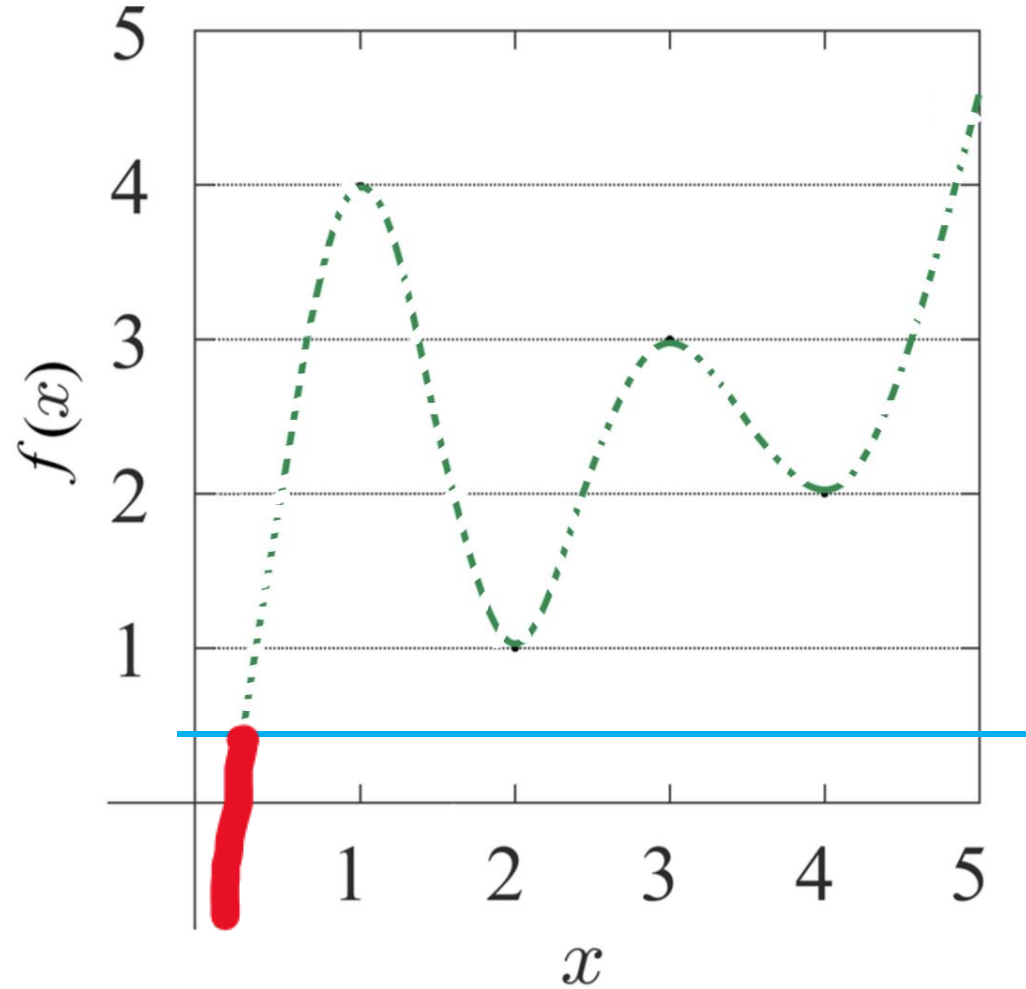
- For another example of persistent homology, we look at the left curve $y = f(x)$
- Again, we consider a growing space
- Each space in the growing sequence is part of the curve below a certain horizontal line
- Left is an example for horizontal line $y = 2.5$
- As the space grows, we track the changes of *0-dimensional homology*
- i.e., we track the **changes of the connected components and the gaps in between**
- On the left, there are *three connected components with two gaps in between*

$$y = 0.5$$



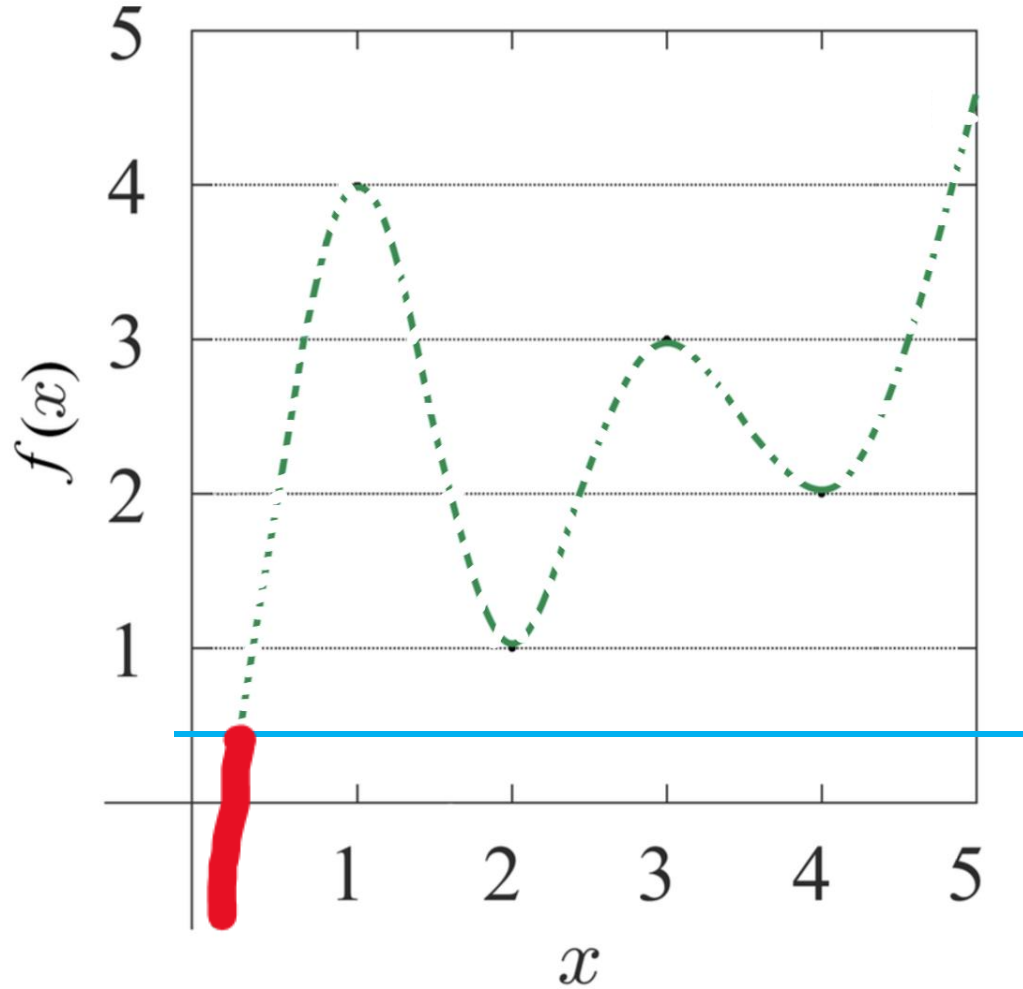
- We have that there is a single connected component (red) below the line $y = 0.5$

$$y = 0.5$$



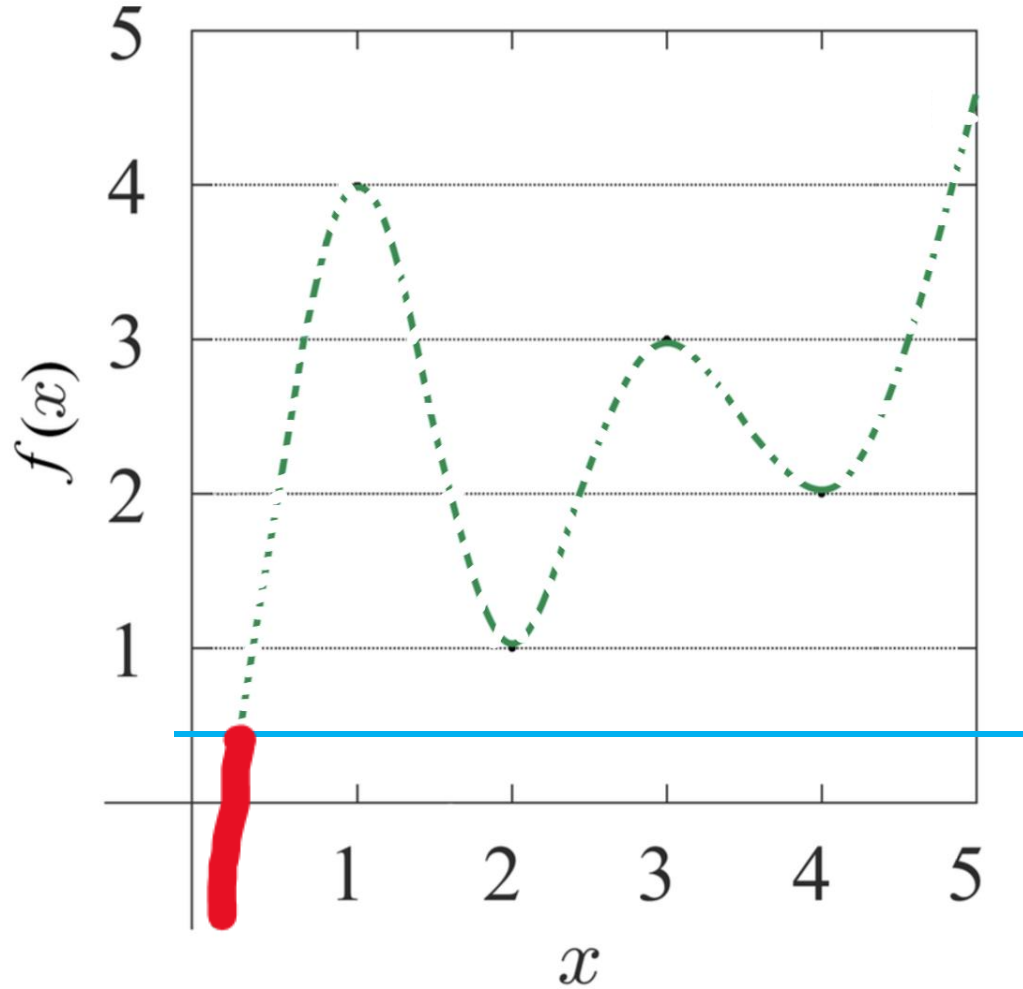
- We have that there is a single connected component (red) below the line $y = 0.5$
- In general, suppose that $f(x)$ approaches $-\infty$ as x approaches 0, we have that there is a single connected component below the line $y = \alpha$ for any $\alpha \leq 0.5$

$$y = 0.5$$



- We have that there is a single connected component (red) below the line $y = 0.5$
- In general, suppose that $f(x)$ approaches $-\infty$ as x approaches 0, we have that there is a single connected component below the line $y = \alpha$ for any $\alpha \leq 0.5$
- So we can assume the red connected component is born at the value $-\infty$

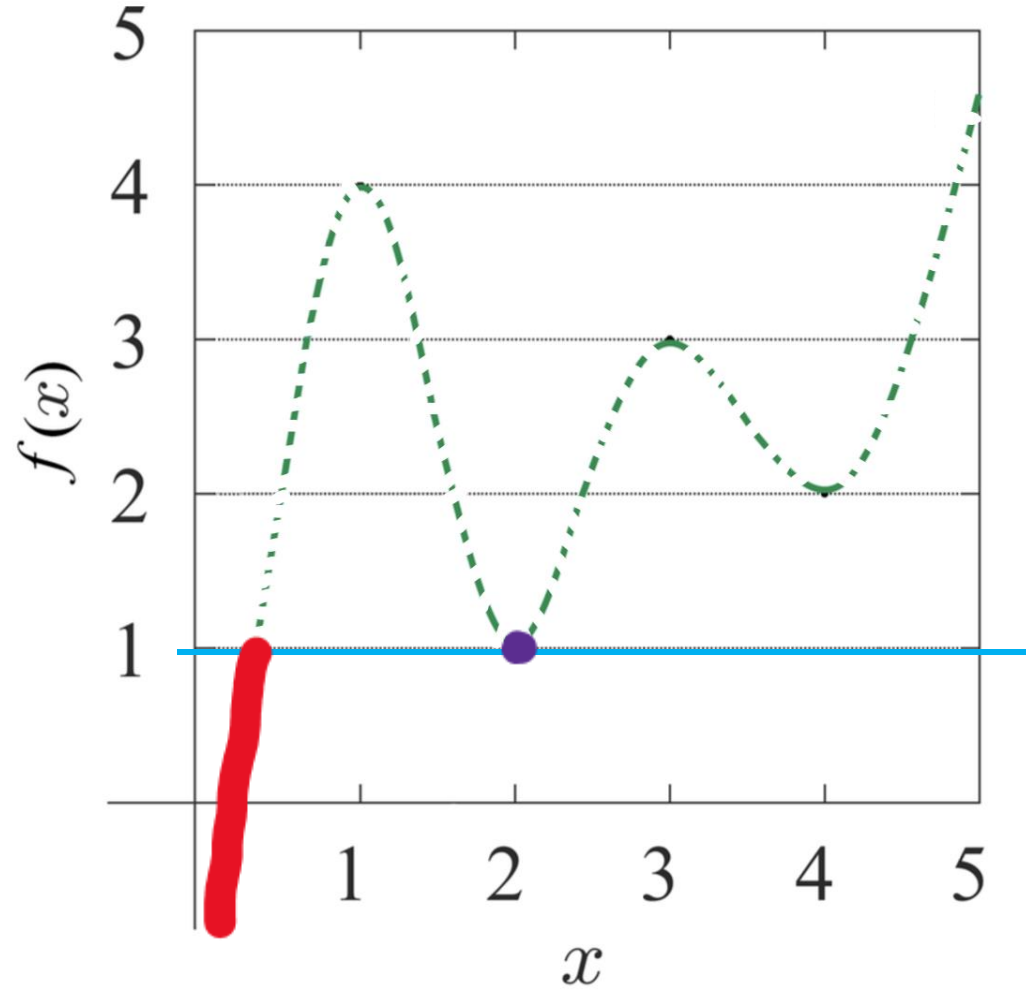
$$y = 0.5$$



- We have that there is a single connected component (red) below the line $y = 0.5$
- In general, suppose that $f(x)$ approaches $-\infty$ as x approaches 0, we have that there is a single connected component below the line $y = \alpha$ for any $\alpha \leq 0.5$
- So we can assume the red connected component is born at the value $-\infty$

- Red: born at $-\infty$

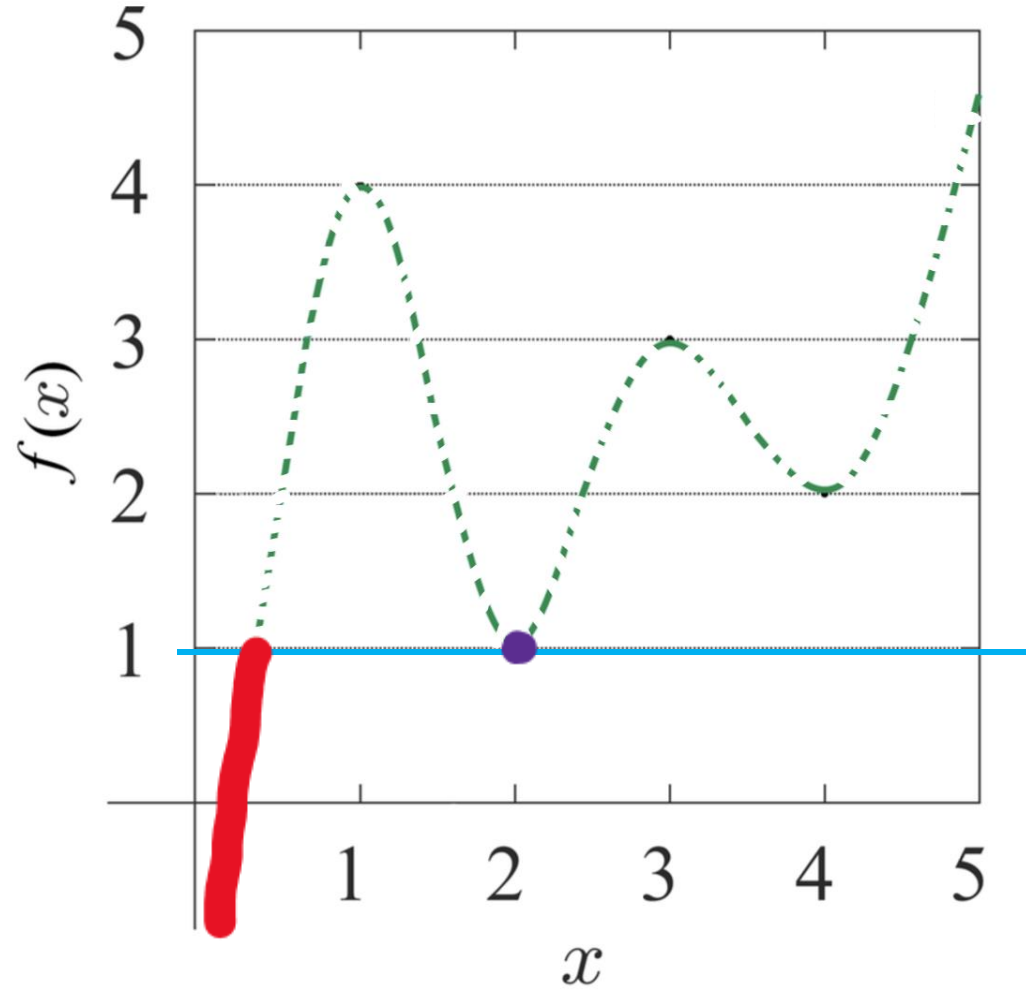
$$y = 1.0$$



- Red component continues
- A new purple component is born

- Red: born at $-\infty$

$$y = 1.0$$

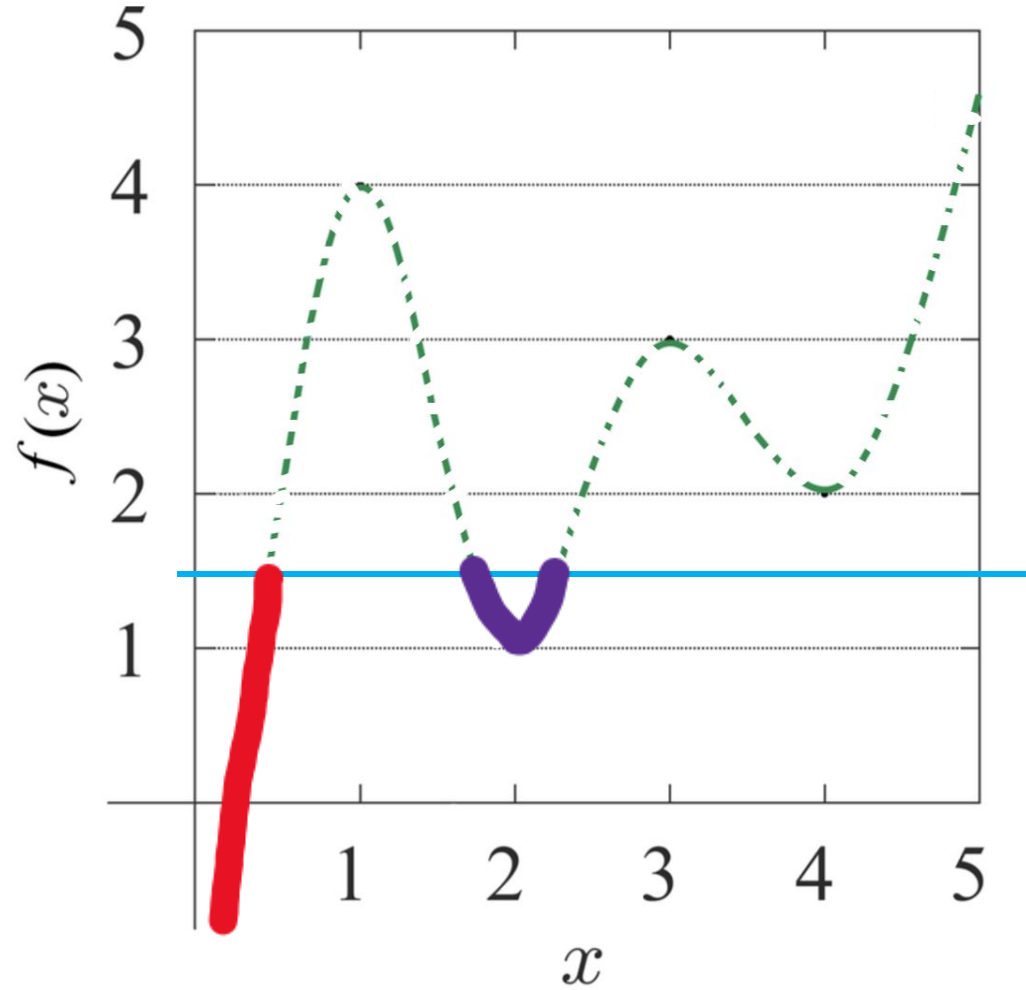


- Red component continues
- A new purple component is born

• Red: born at $-\infty$

• Purple: born at 1.0

$$y = 1.5$$

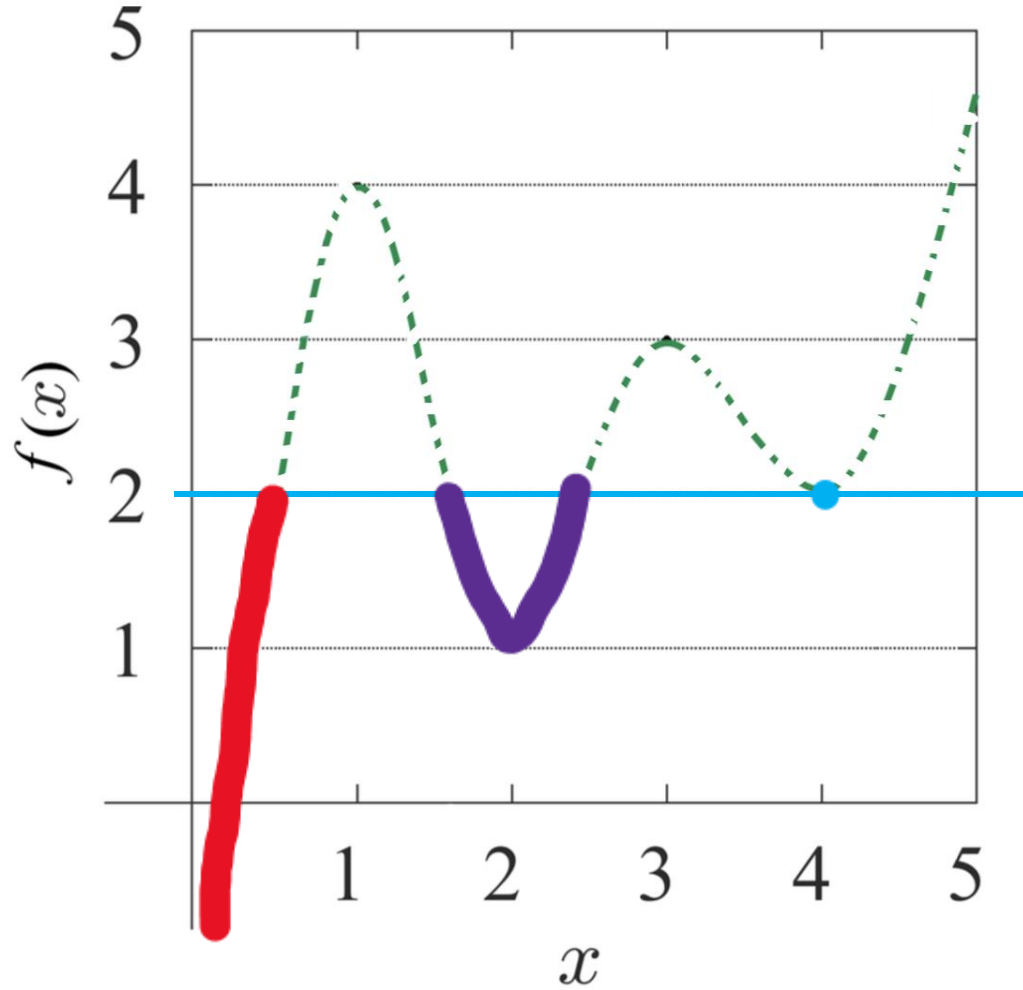


- Red and purple components continue

• Red: born at $-\infty$

• Purple: born at 1.0

$$y = 2.0$$

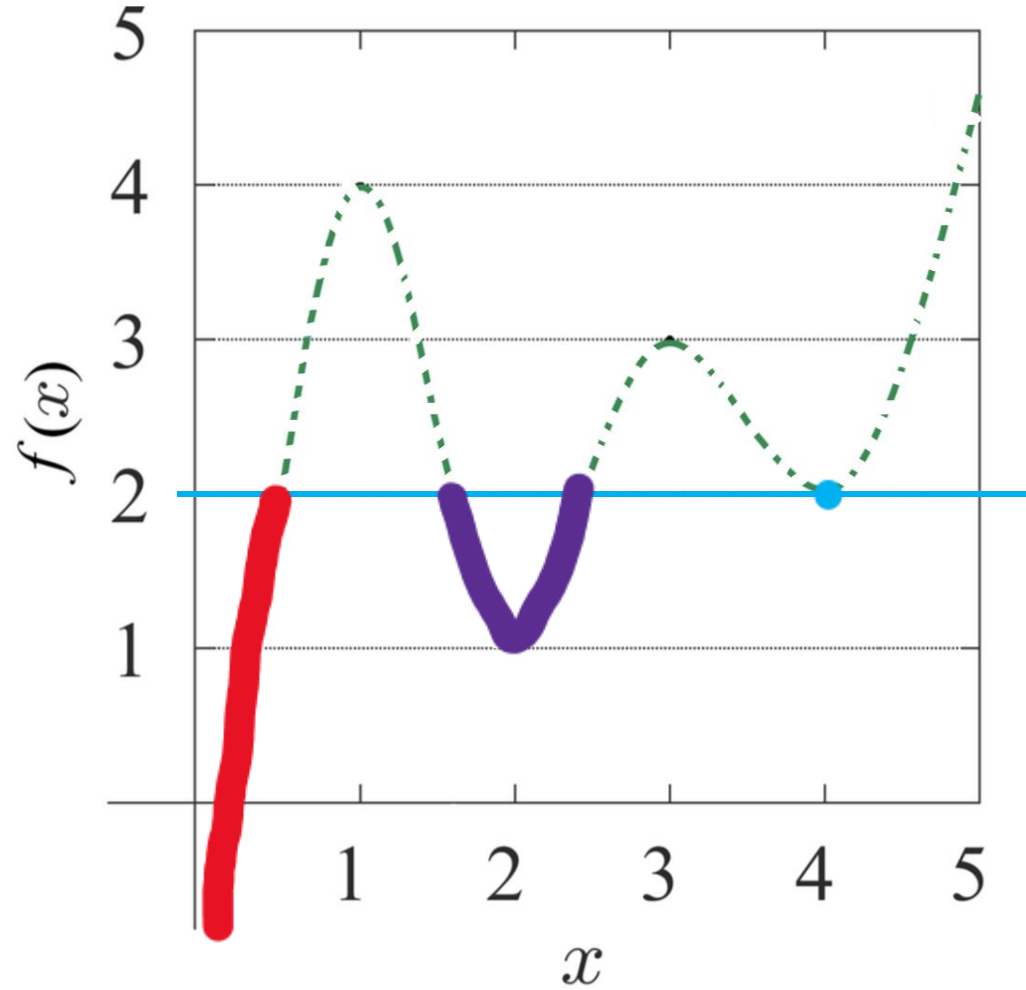


- Red and purple components continue
- A new blue component is born

• Red: born at $-\infty$

• Purple: born at 1.0

$$y = 2.0$$



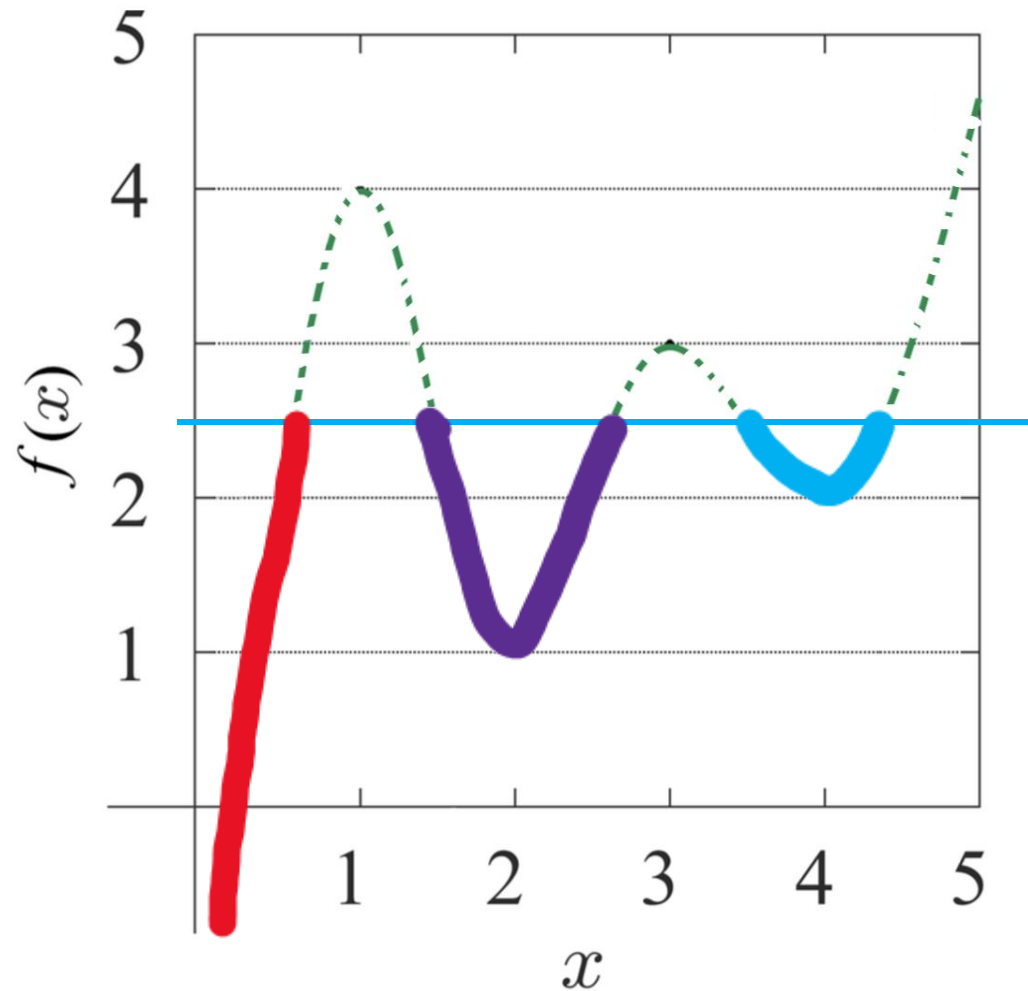
- Red and purple components continue
- A new blue component is born

• Red: born at $-\infty$

• Purple: born at 1.0

• Blue: born at 2.0

$$y = 2.5$$



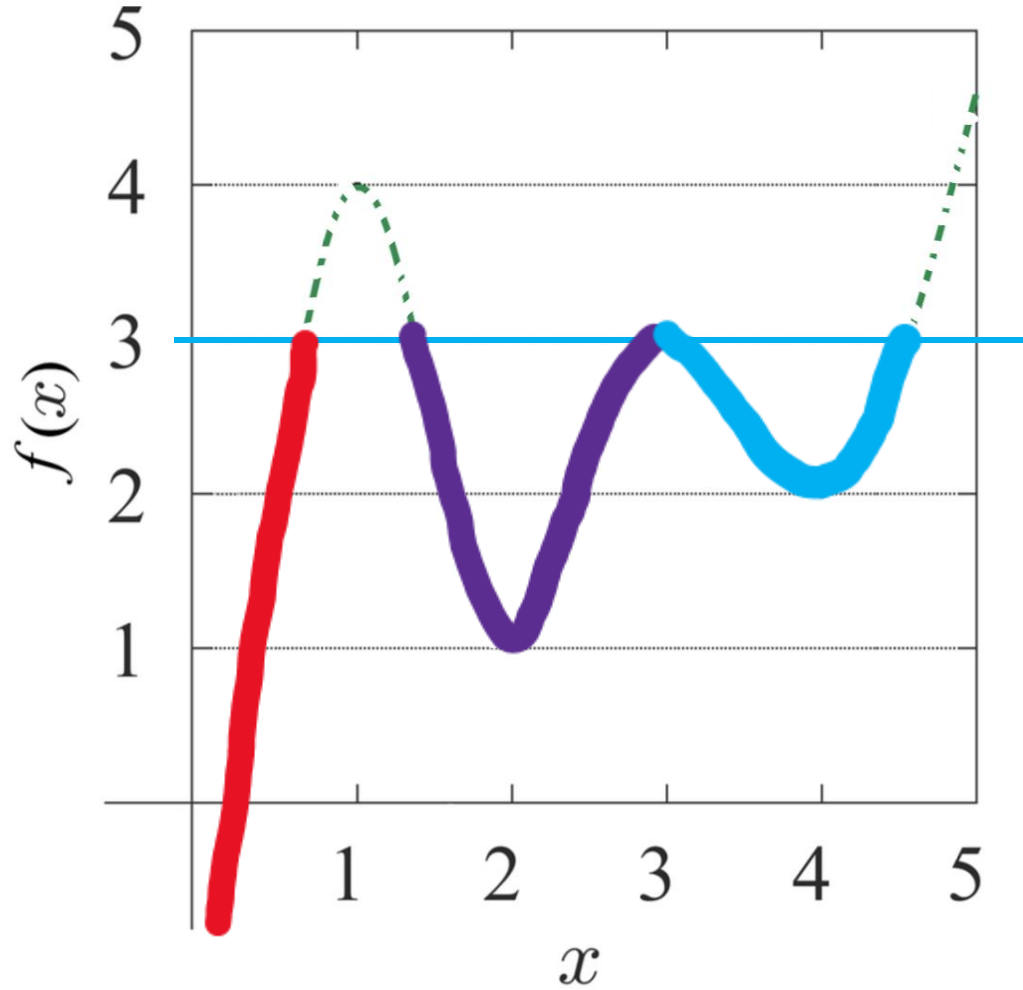
- Three components continue

• Red: born at $-\infty$

• Purple: born at 1.0

• Blue: born at 2.0

$$y = 3.0$$



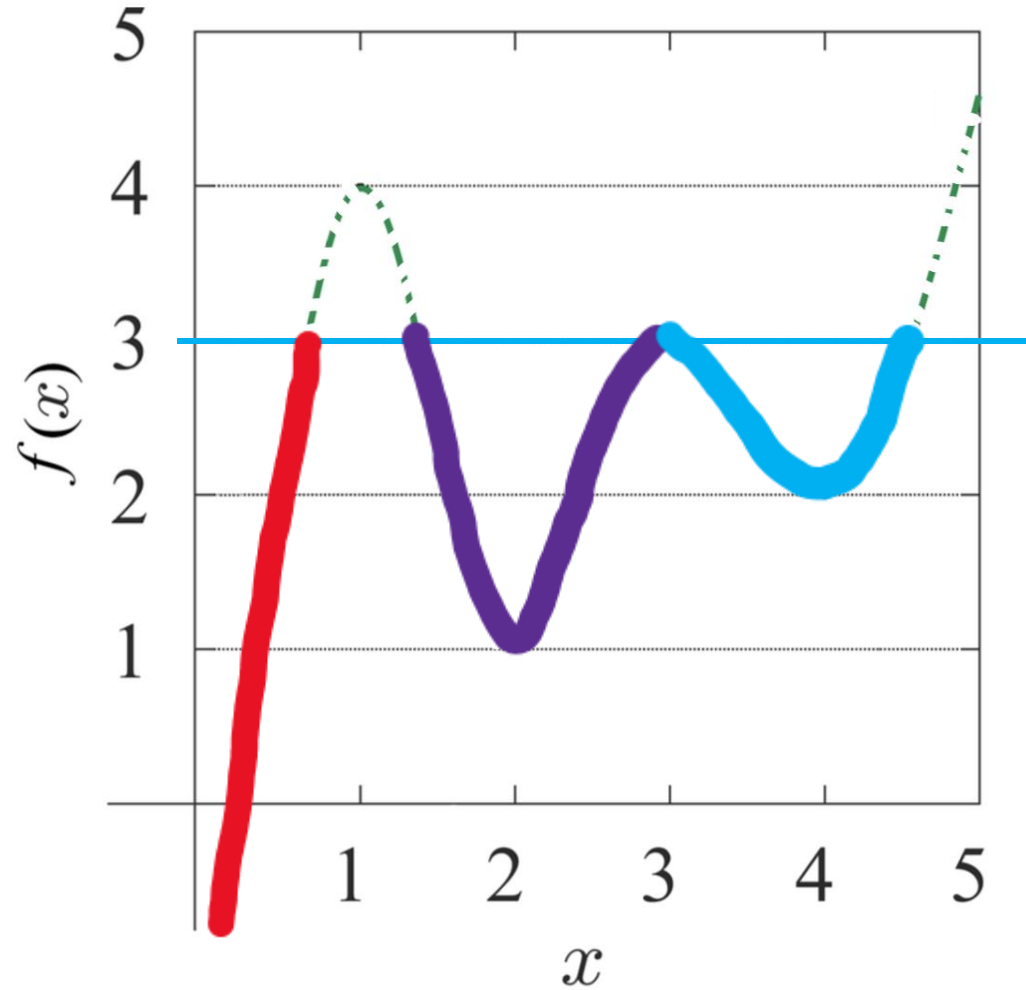
- The **purple** and **blue** components merge into one (gaps between them disappear)

• **Red**: born at $-\infty$

• **Purple**: born at 1.0

• **Blue**: born at 2.0

$$y = 3.0$$



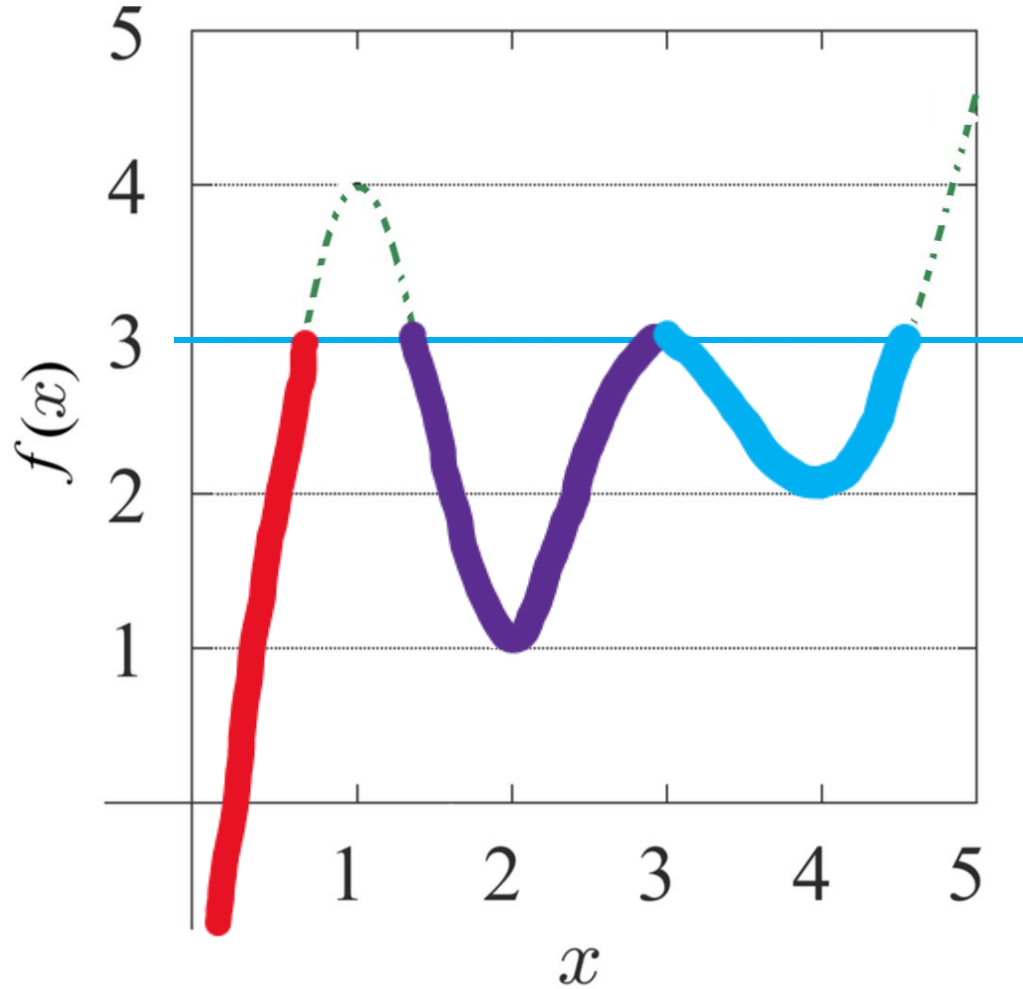
- The **purple** and **blue** components merge into one (gaps between them disappear)
- This means that a **0-dimensional homology hole disappears (dies)**

• **Red**: born at $-\infty$

• **Purple**: born at 1.0

• **Blue**: born at 2.0

$$y = 3.0$$



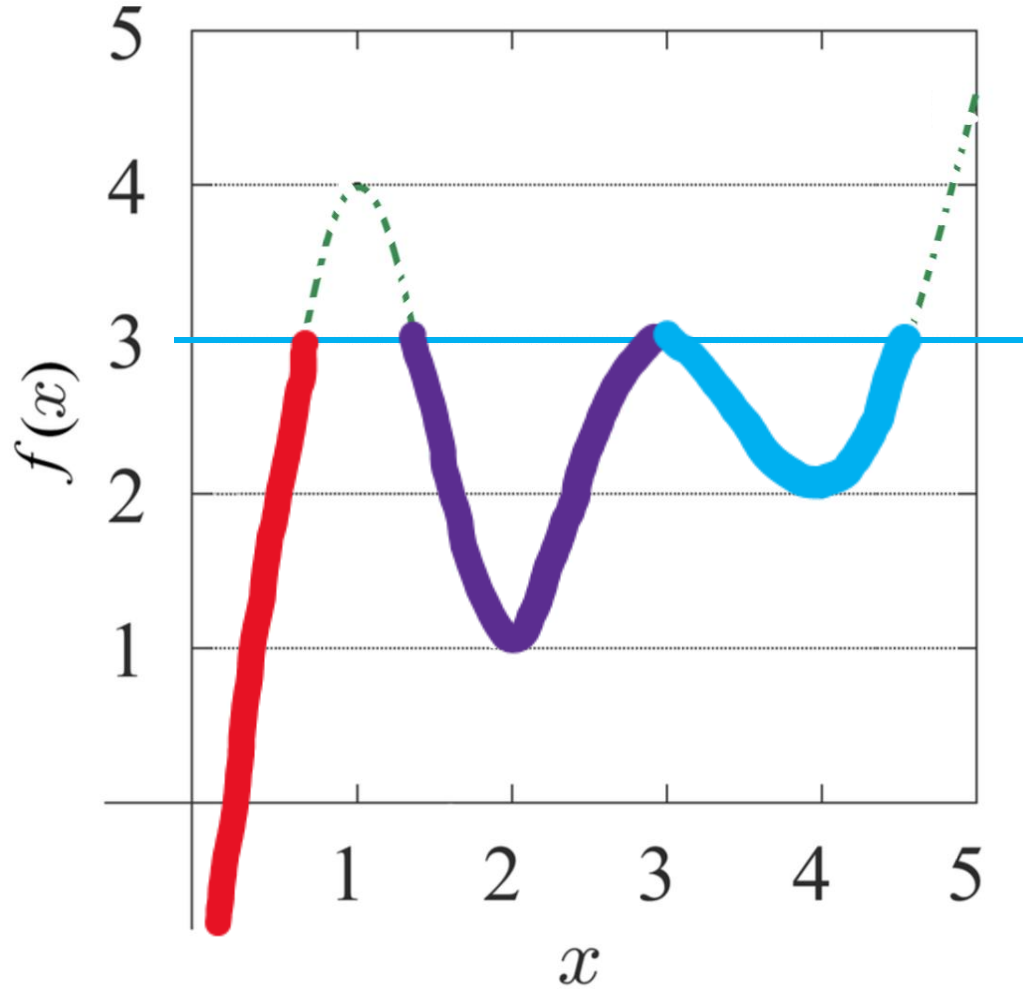
- The **purple** and **blue** components merge into one (gaps between them disappear)
- This means that a **0-dimensional homology hole disappears (dies)**
- The gap between **purple** and **blue** components appears because of birth of the **blue** component

• **Red**: born at $-\infty$

• **Purple**: born at 1.0

• **Blue**: born at 2.0

$$y = 3.0$$



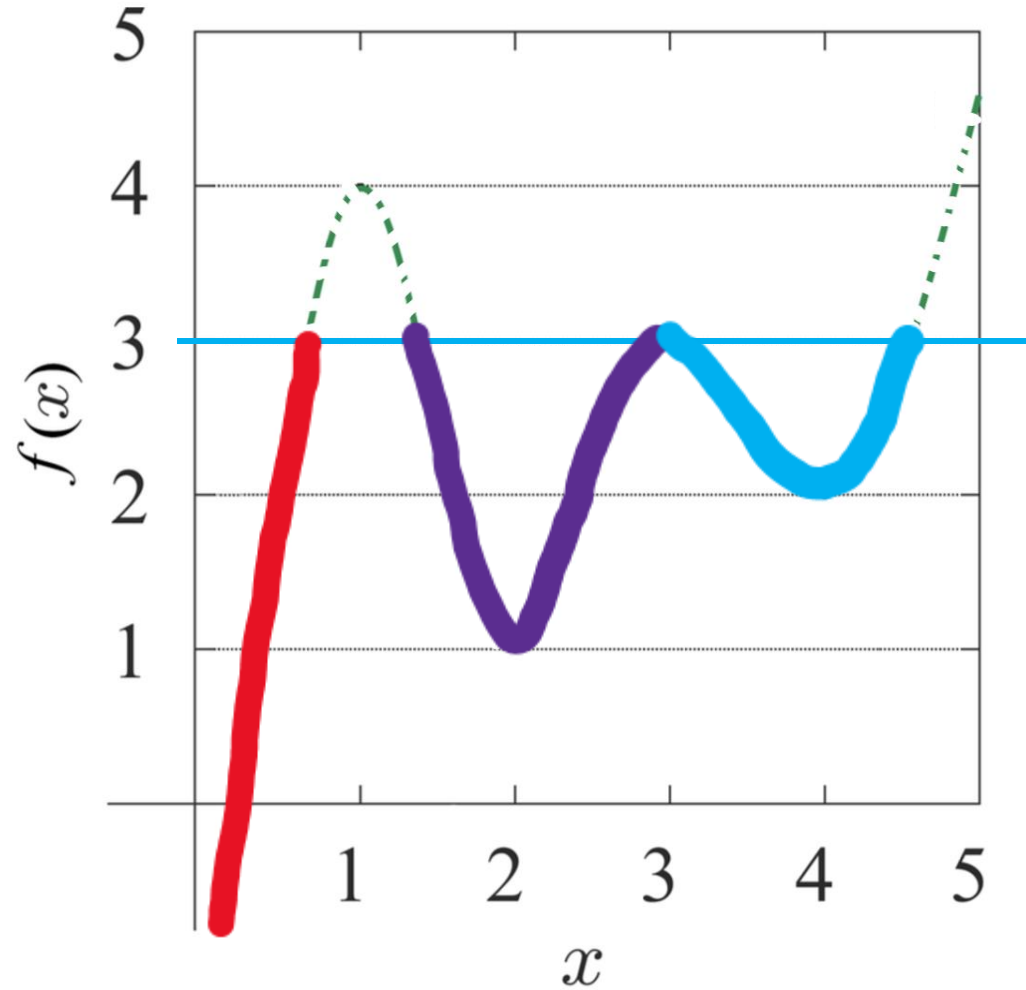
- The **purple** and **blue** components merge into one (gaps between them disappear)
- This means that a **0-dimensional homology hole disappears (dies)**
- The gap between **purple** and **blue** components appears because of birth of the **blue** component
- So we consider the gap to be born when the **blue** component is born, i.e., **at 2.0**

• **Red**: born at $-\infty$

• **Purple**: born at 1.0

• **Blue**: born at 2.0

$$y = 3.0$$



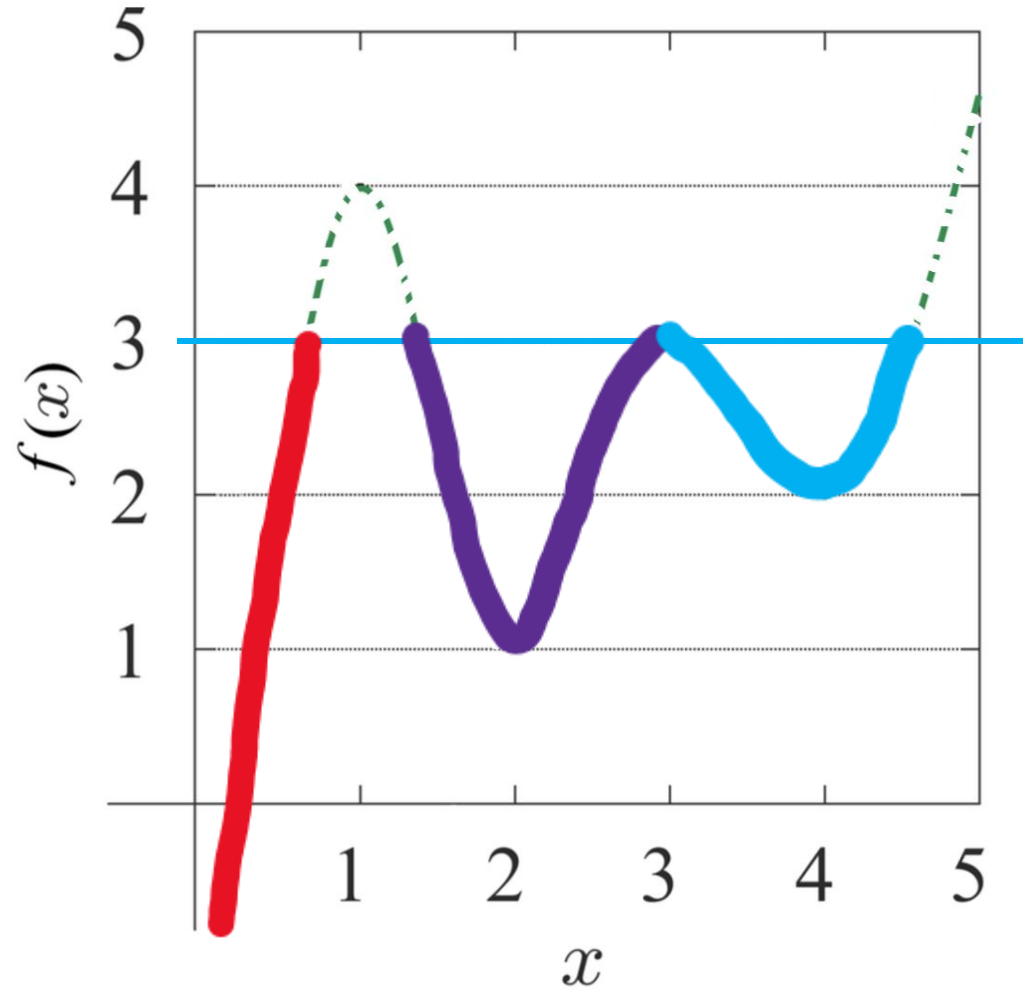
- The **purple** and **blue** components merge into one (gaps between them disappear)
- This means that a **0-dimensional homology hole disappears (dies)**
- The gap between **purple** and **blue** components appears because of birth of the **blue** component
- So we consider the gap to be born when the **blue** component is born, i.e., **at 2.0**
- So we have a 0-dimensional hole **born at 2.0** and **dies at 3.0**

• **Red**: born at $-\infty$

• **Purple**: born at 1.0

• **Blue**: born at 2.0

$$y = 3.0$$



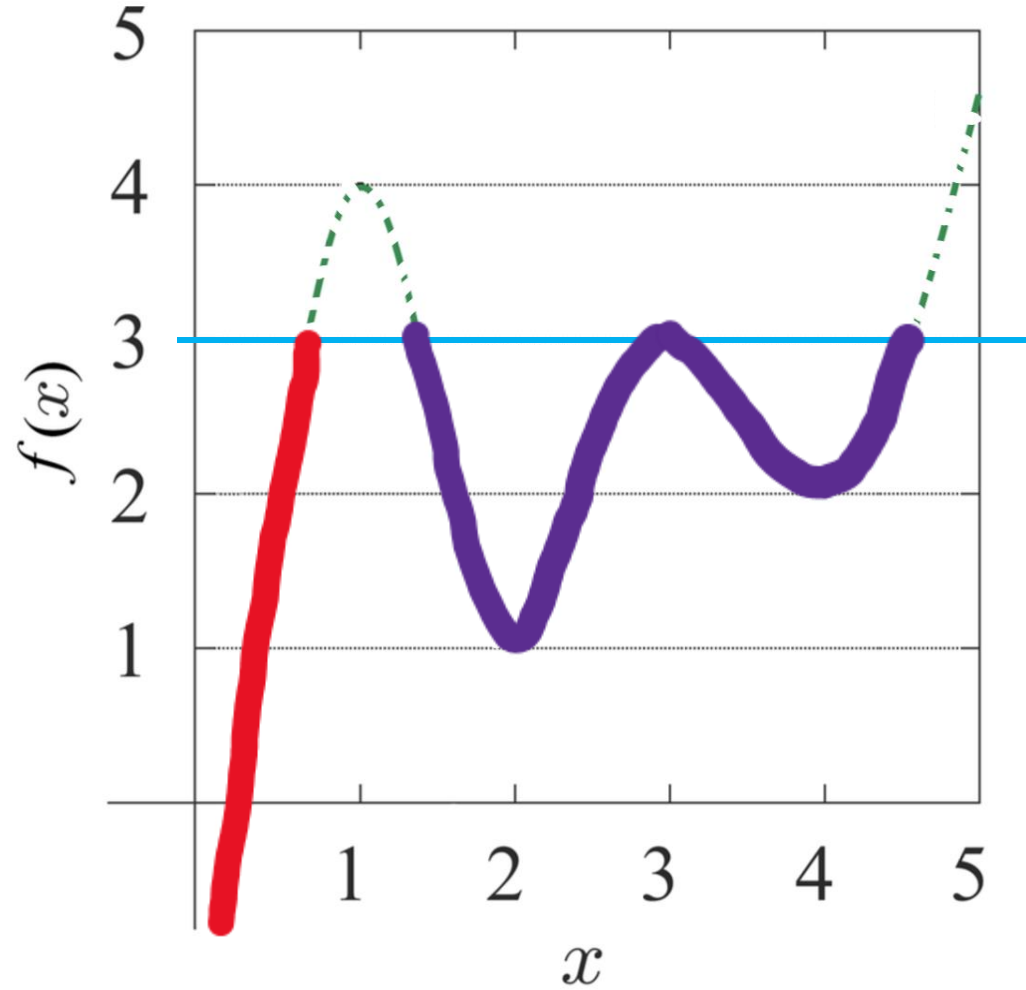
- The **purple** and **blue** components merge into one (gaps between them disappear)
- This means that a **0-dimensional homology hole disappears (dies)**
- The gap between **purple** and **blue** components appears because of birth of the **blue** component
- So we consider the gap to be born when the **blue** component is born, i.e., **at 2.0**
- So we have a 0-dimensional hole **born at 2.0** and **dies at 3.0**

• **Red**: born at $-\infty$

• **Purple**: born at 1.0

• PD: **(2.0, 3.0)**

$$y = 3.0$$

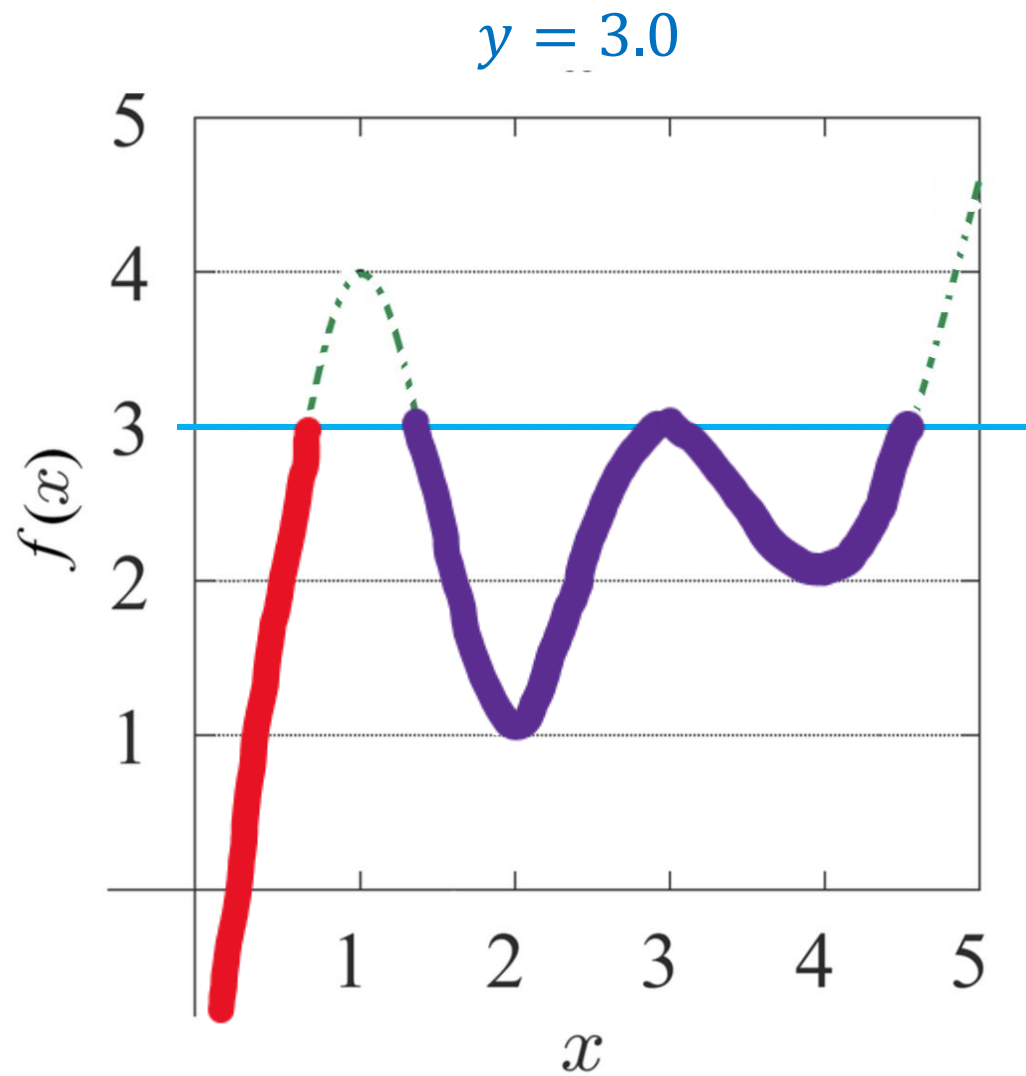


- For the merged component, we keep the one born earlier (**purple**), and kill the one born later (**blue**)

• **Red**: born at $-\infty$

• **Purple**: born at 1.0

• PD: **(2.0, 3.0)**



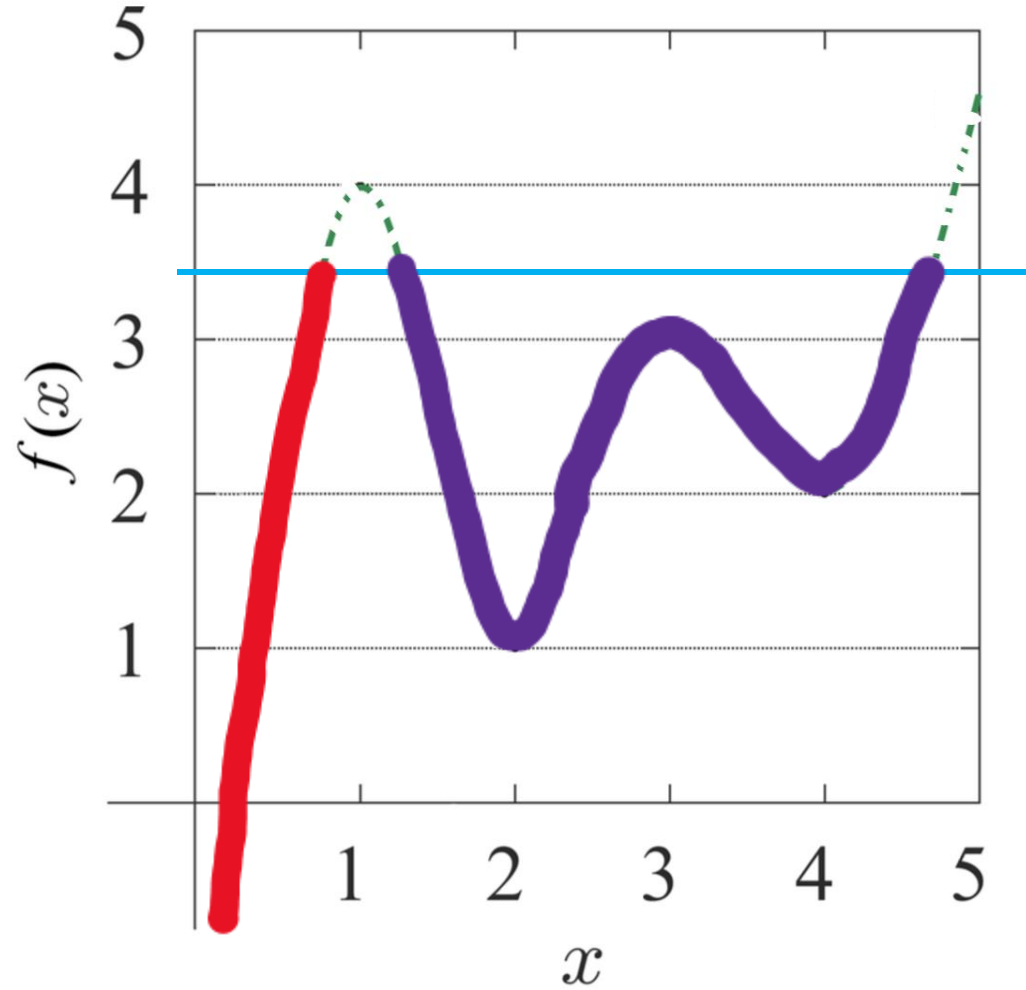
- For the merged component, we keep the one born earlier (**purple**), and kill the one born later (**blue**)
- So we have a larger **purple** component born at **1.0**

• **Red**: born at $-\infty$

• **Purple**: born at 1.0

• PD: **(2.0, 3.0)**

$$y = 3.5$$



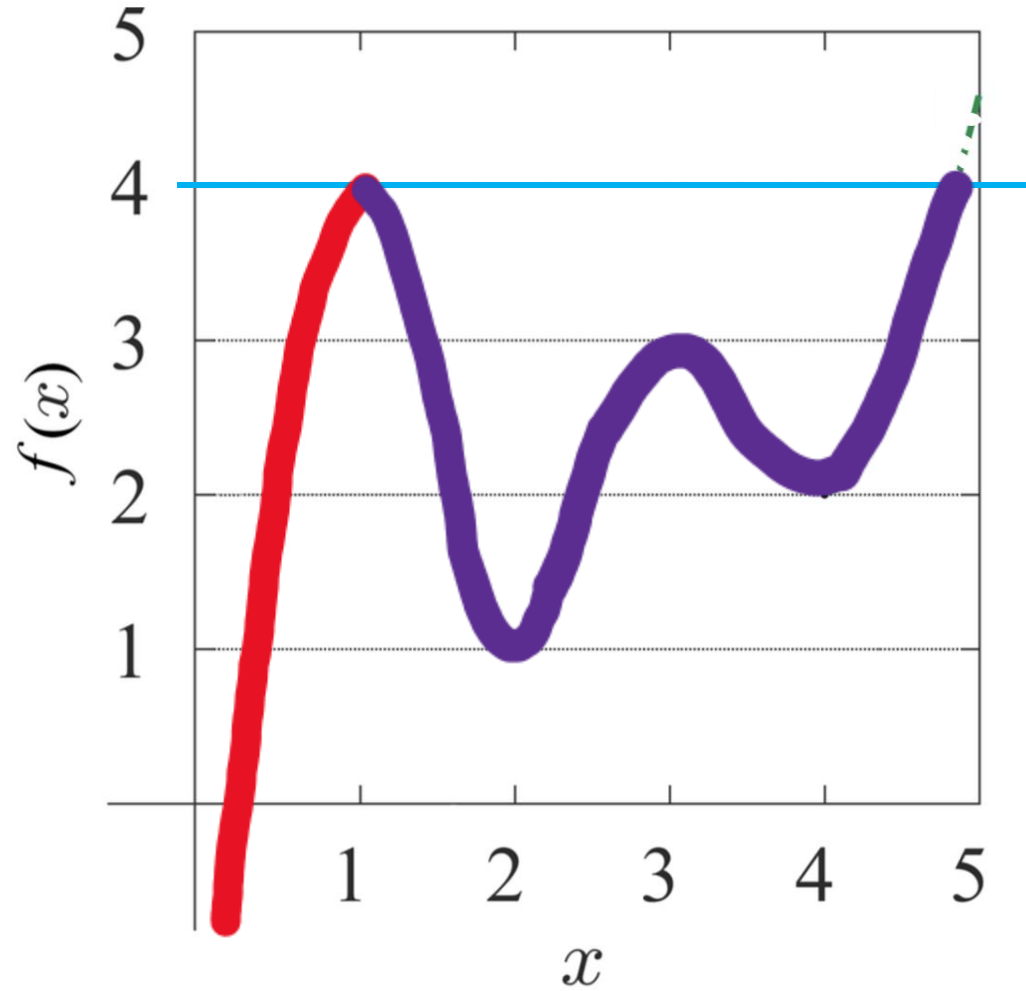
- Two components continue

• Red: born at $-\infty$

• Purple: born at 1.0

• PD: $(2.0, 3.0)$

$$y = 4.0$$



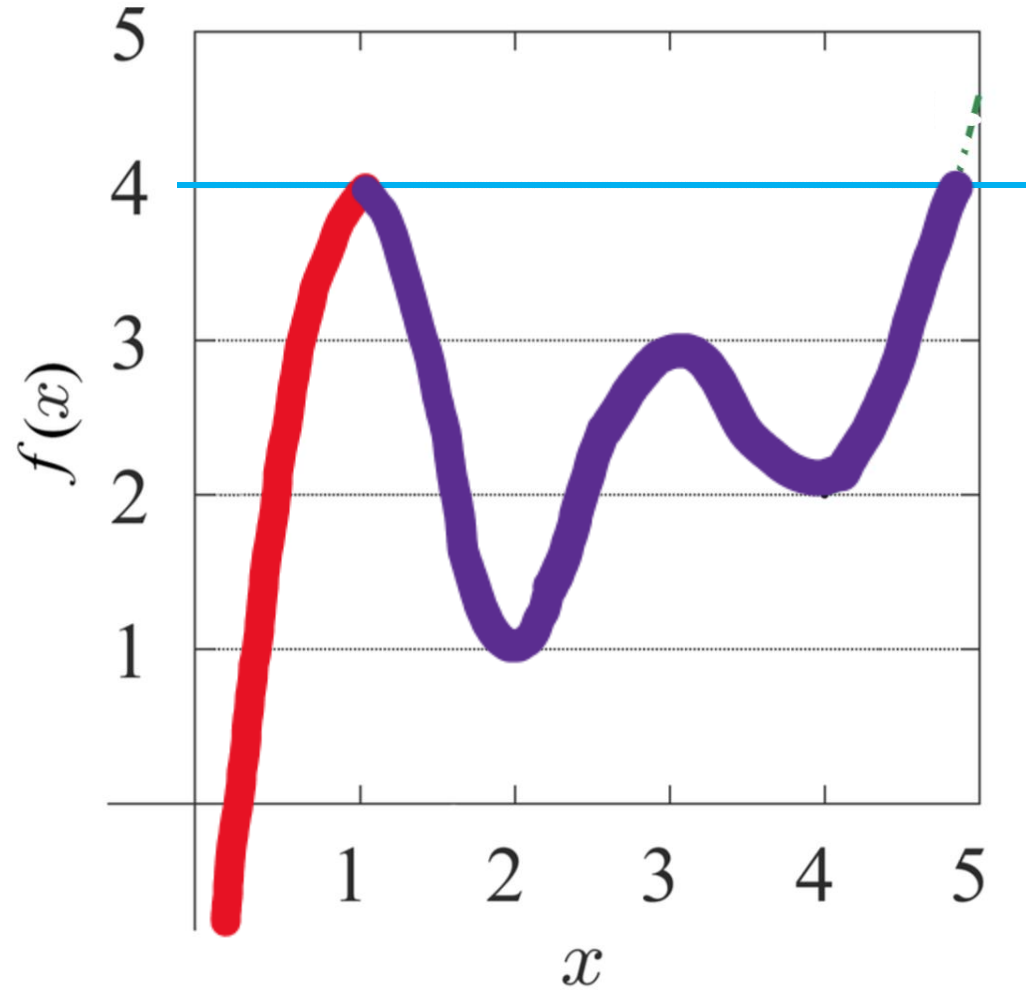
- The **red** and **purple** components merge into one (gaps between them disappear)

- **Red**: born at $-\infty$

- **Purple**: born at 1.0

- PD: (2.0, 3.0)

$$y = 4.0$$



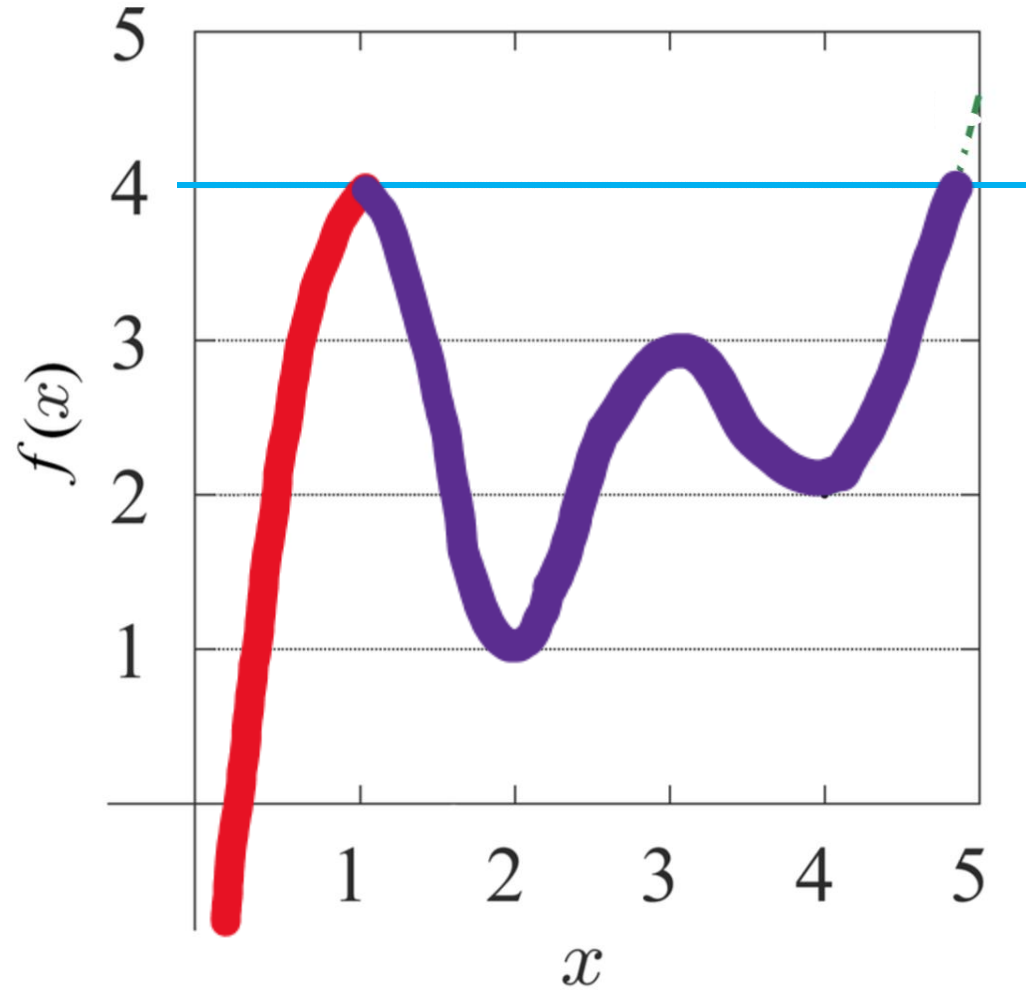
- The **red** and **purple** components merge into one (gaps between them disappear)
- A 0-dimensional homology hole disappears (dies)

• **Red**: born at $-\infty$

• **Purple**: born at 1.0

• PD: **(2.0, 3.0)**

$$y = 4.0$$

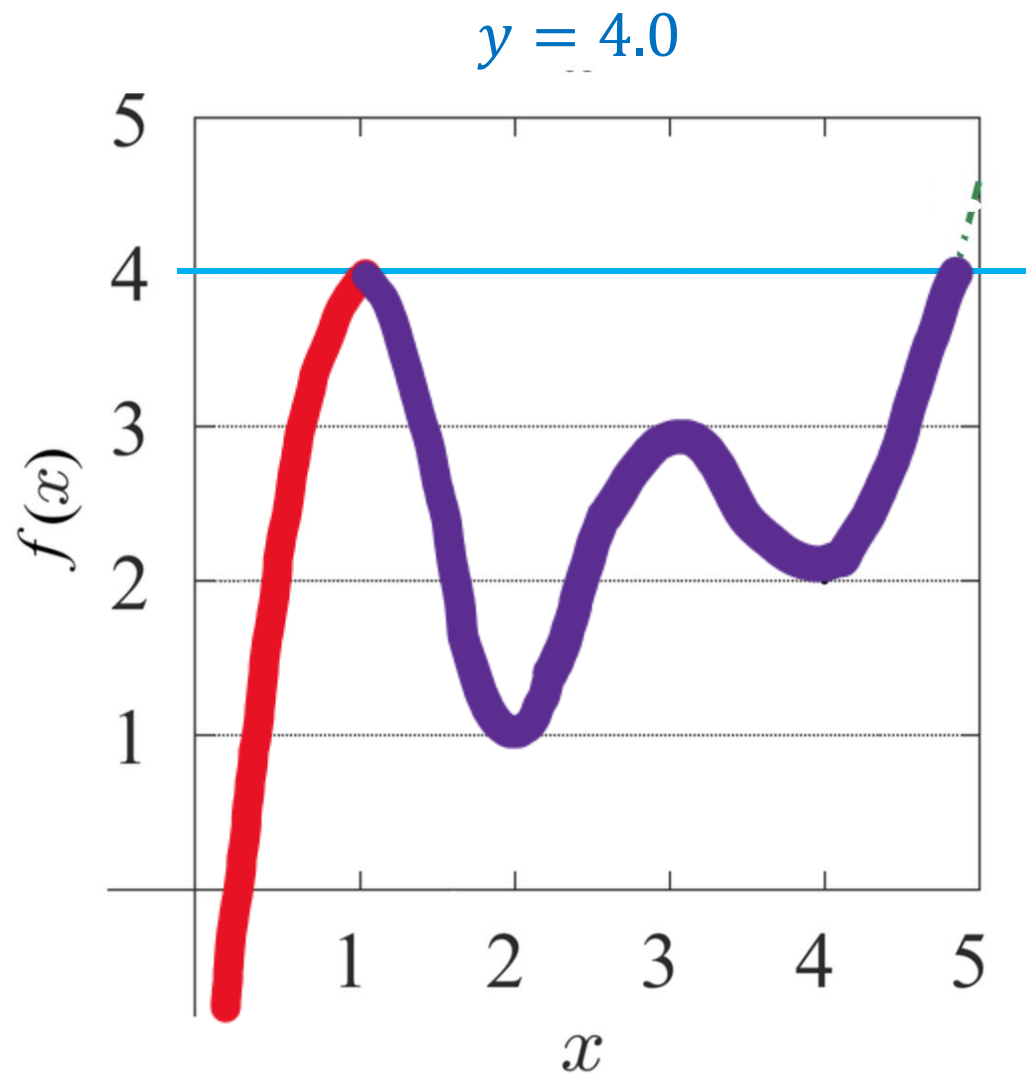


- The **red** and **purple** components merge into one (gaps between them disappear)
- A 0-dimensional homology hole disappears (dies)
- The gap between **red** and **purple** components appears because of birth of the **purple** component

• **Red**: born at $-\infty$

• **Purple**: born at 1.0

• PD: (2.0, 3.0)

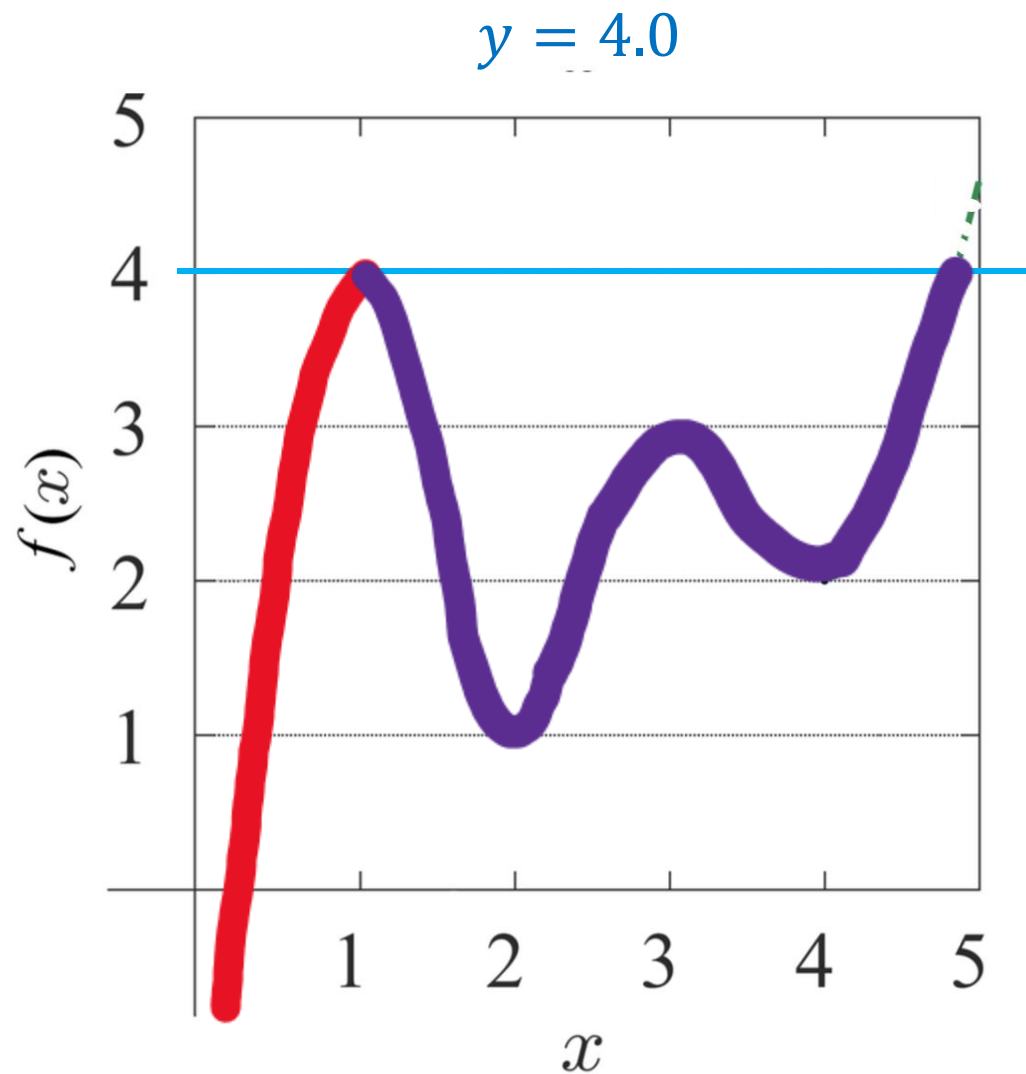


- The **red** and **purple** components merge into one (gaps between them disappear)
- A 0-dimensional homology hole disappears (dies)
- The gap between **red** and **purple** components appears because of birth of the **purple** component
- So the gap is born **at 1.0**

• **Red**: born at $-\infty$

• **Purple**: born at 1.0

• PD: **(2.0, 3.0)**

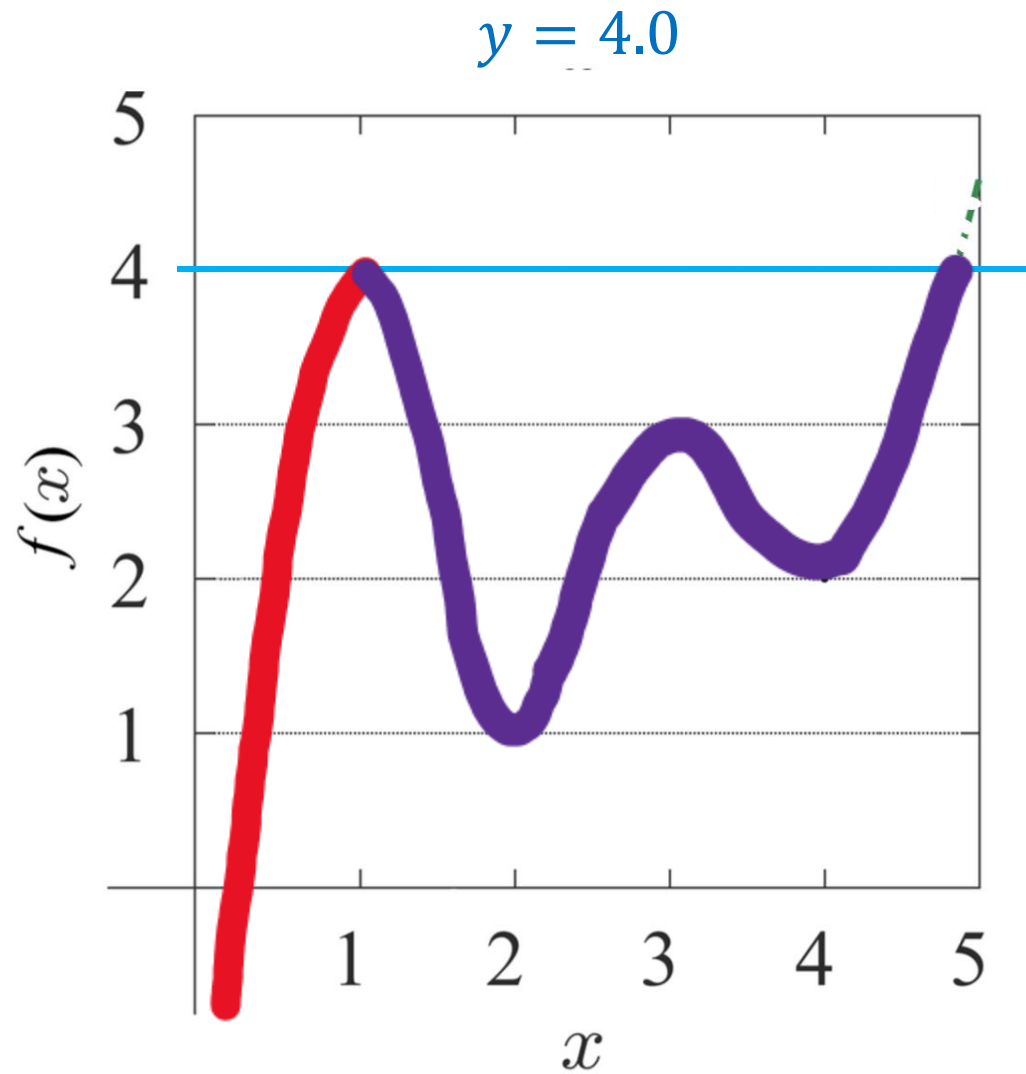


- The **red** and **purple** components merge into one (gaps between them disappear)
- A 0-dimensional homology hole disappears (dies)
- The gap between **red** and **purple** components appears because of birth of the **purple** component
- So the gap is born **at 1.0**
- So we have a 0-dimensional hole **born at 1.0** and **dies at 4.0**

• **Red**: born at $-\infty$

• **Purple**: born at 1.0

• PD: **(2.0, 3.0)**



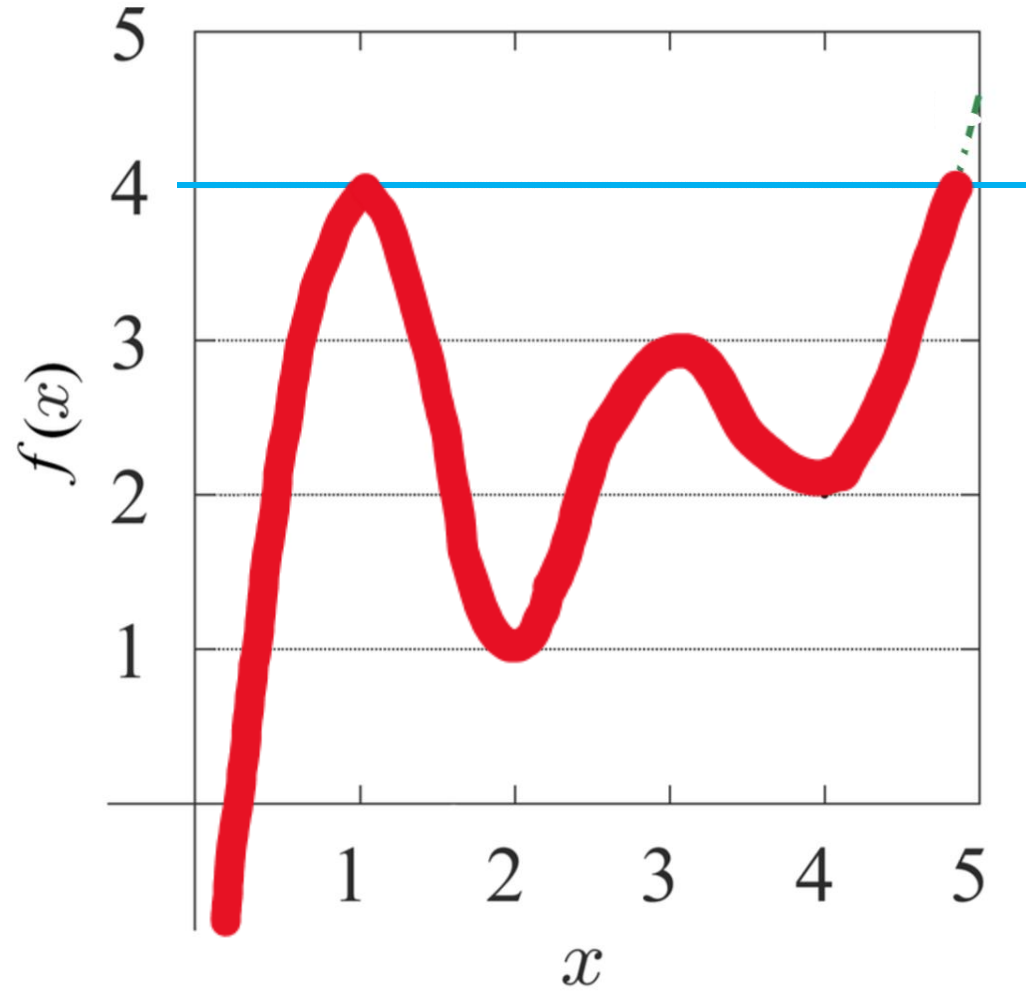
- The **red** and **purple** components merge into one (gaps between them disappear)
- A 0-dimensional homology hole disappears (dies)
- The gap between **red** and **purple** components appears because of birth of the **purple** component
- So the gap is born **at 1.0**
- So we have a 0-dimensional hole **born at 1.0** and **dies at 4.0**

• **Red**: born at $-\infty$

• PD: **(1.0, 4.0)**

• PD: **(2.0, 3.0)**

$$y = 4.0$$



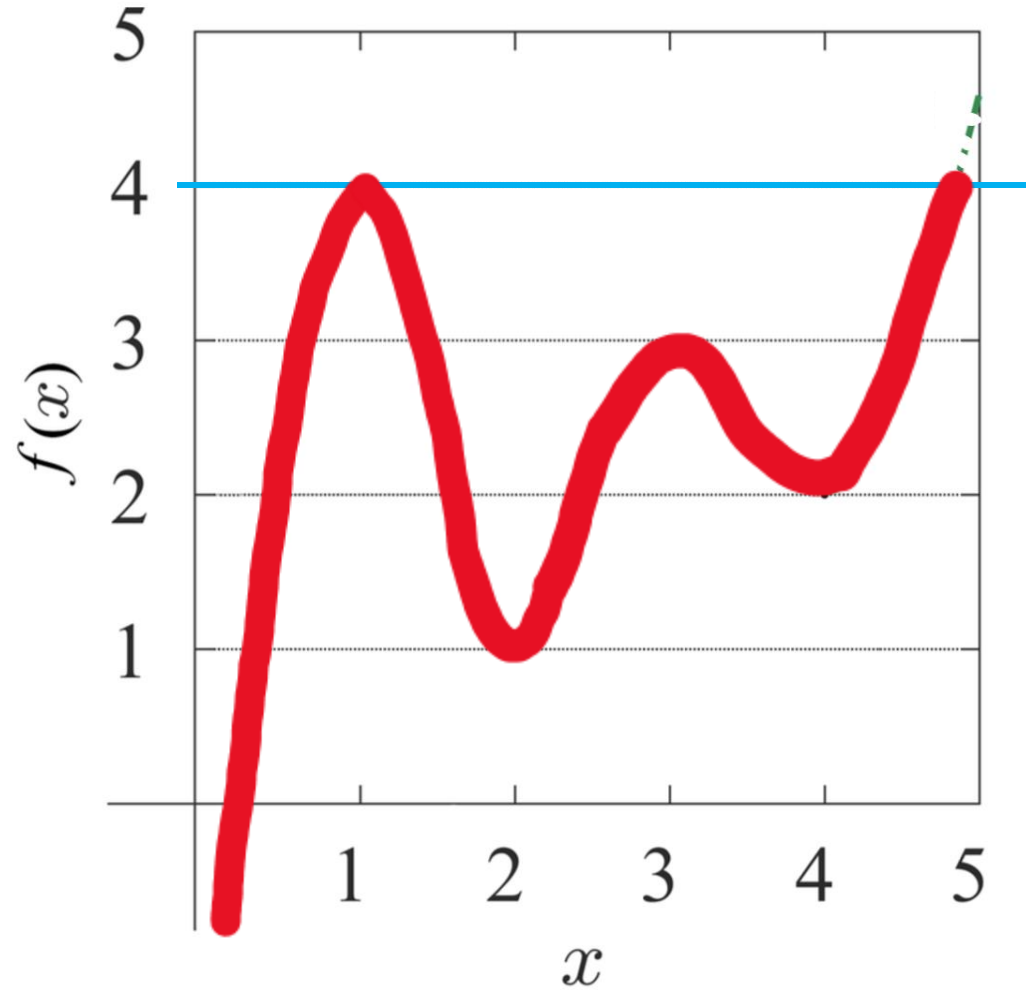
- For the merged component, we keep the one born earlier (**red**), and kill the one born later (**purple**)

• **Red**: born at $-\infty$

• PD: (1.0, 4.0)

• PD: (2.0, 3.0)

$$y = 4.0$$



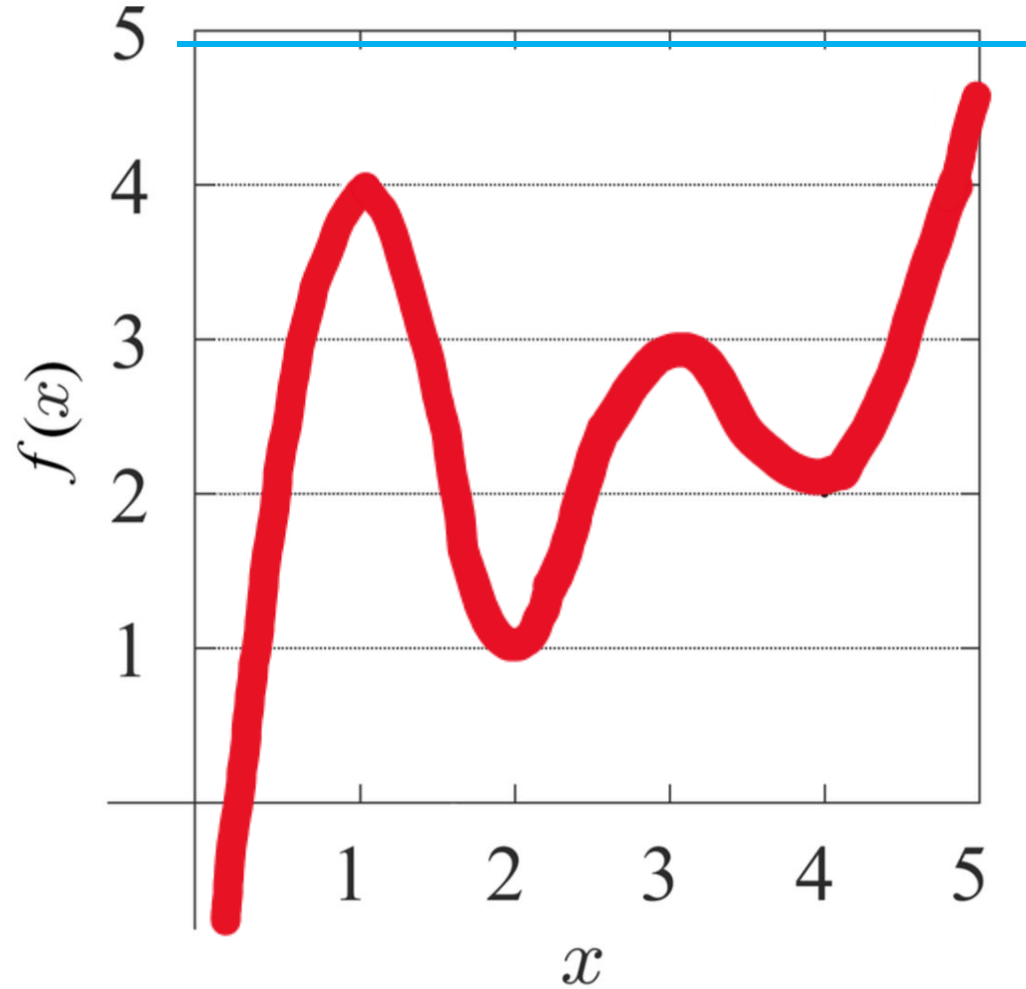
- For the merged component, we keep the one born earlier (**red**), and kill the one born later (**purple**)
- So we have a single **red** component born at $-\infty$

• **Red**: born at $-\infty$

• PD: $(1.0, 4.0)$

• PD: $(2.0, 3.0)$

α arbitrary large



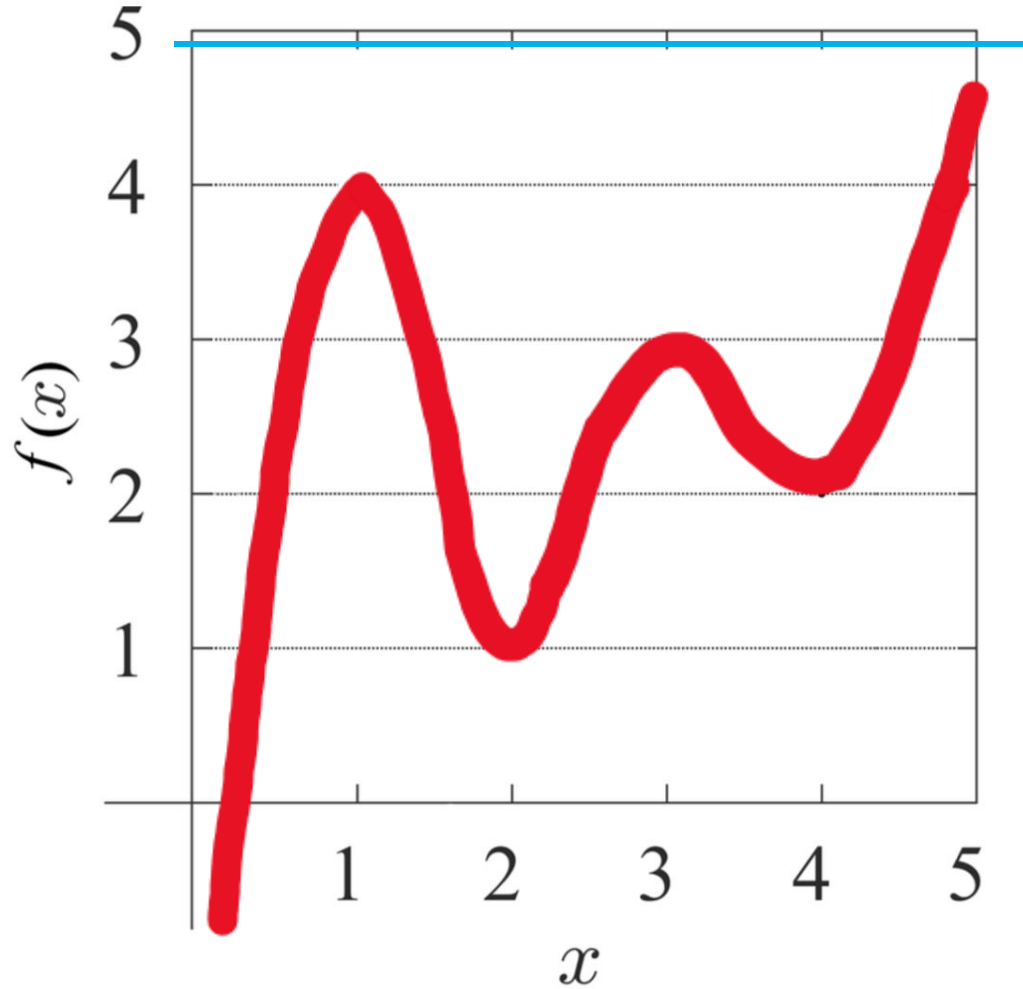
- As the value for the line keeps on increasing to $+\infty$, the single **red** component will keep on persisting
- So we have the **red** component born at $-\infty$ and dies at $+\infty$

• **Red**: born at $-\infty$

• PD: $(1.0, 4.0)$

• PD: $(2.0, 3.0)$

α arbitrary large

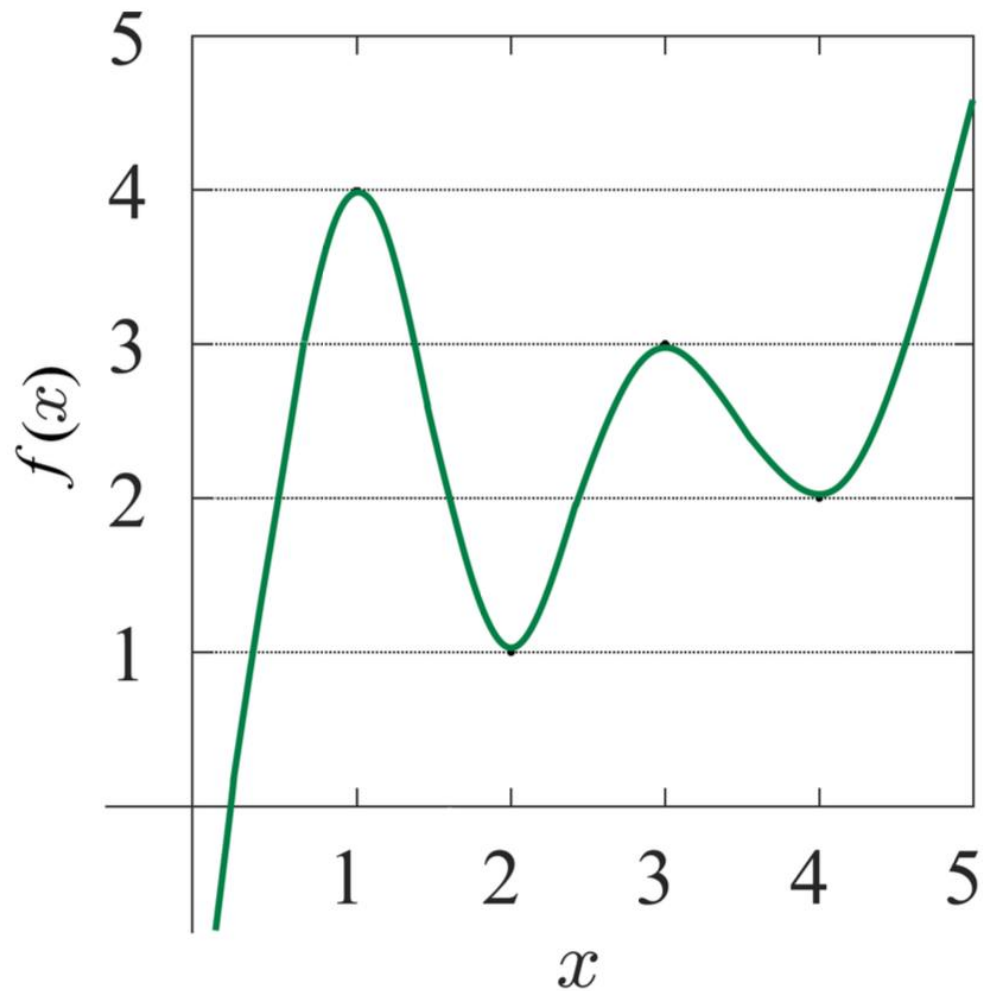


- As the value for the line keeps on increasing to $+\infty$, the single red component will keep on persisting
- So we have the red component born at $-\infty$ and dies at $+\infty$

• PD: $(-\infty, +\infty)$

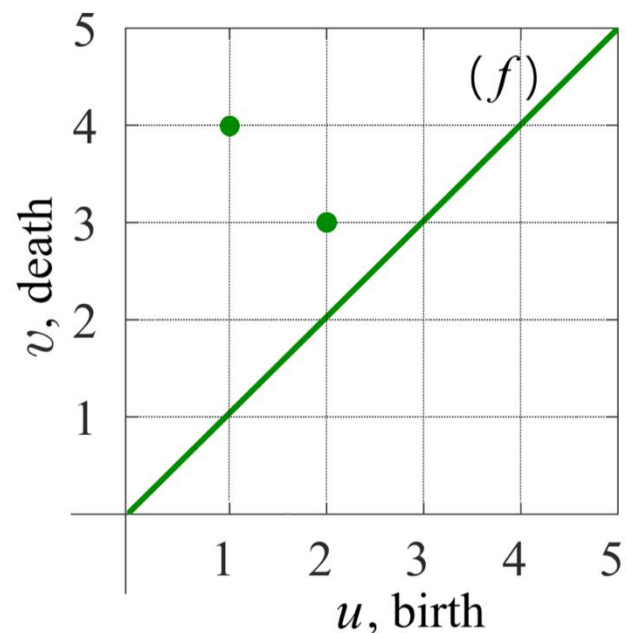
• PD: $(1.0, 4.0)$

• PD: $(2.0, 3.0)$



Summary:

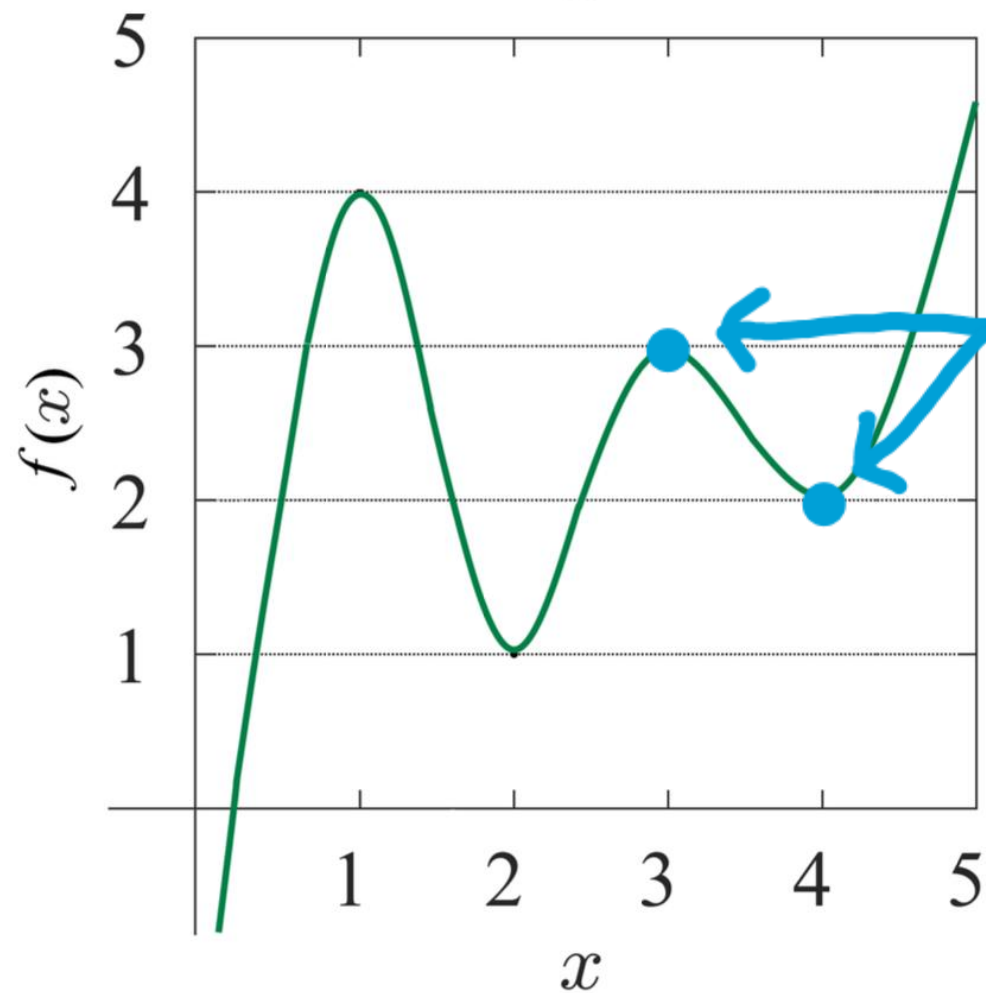
- We have three points in the 0-dimension PD
- Each point is tracking the birth and death of a connect component (or gap in between)



• PD: $(-\infty, +\infty)$

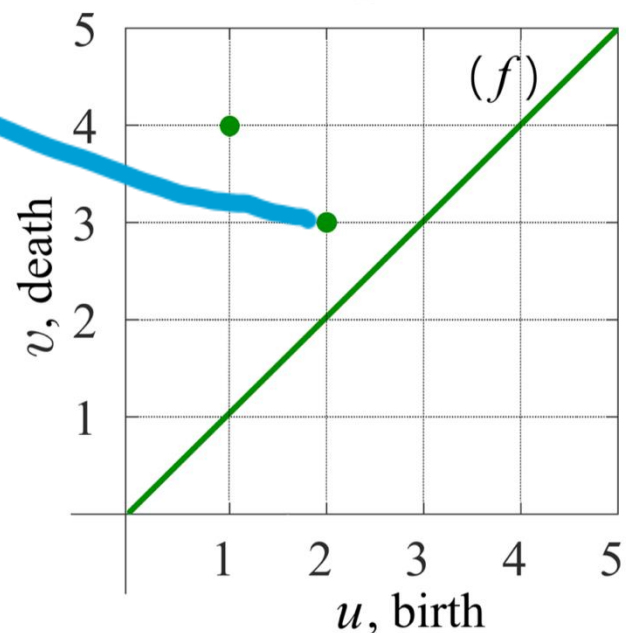
• PD: $(1.0, 4.0)$

• PD: $(2.0, 3.0)$



Summary:

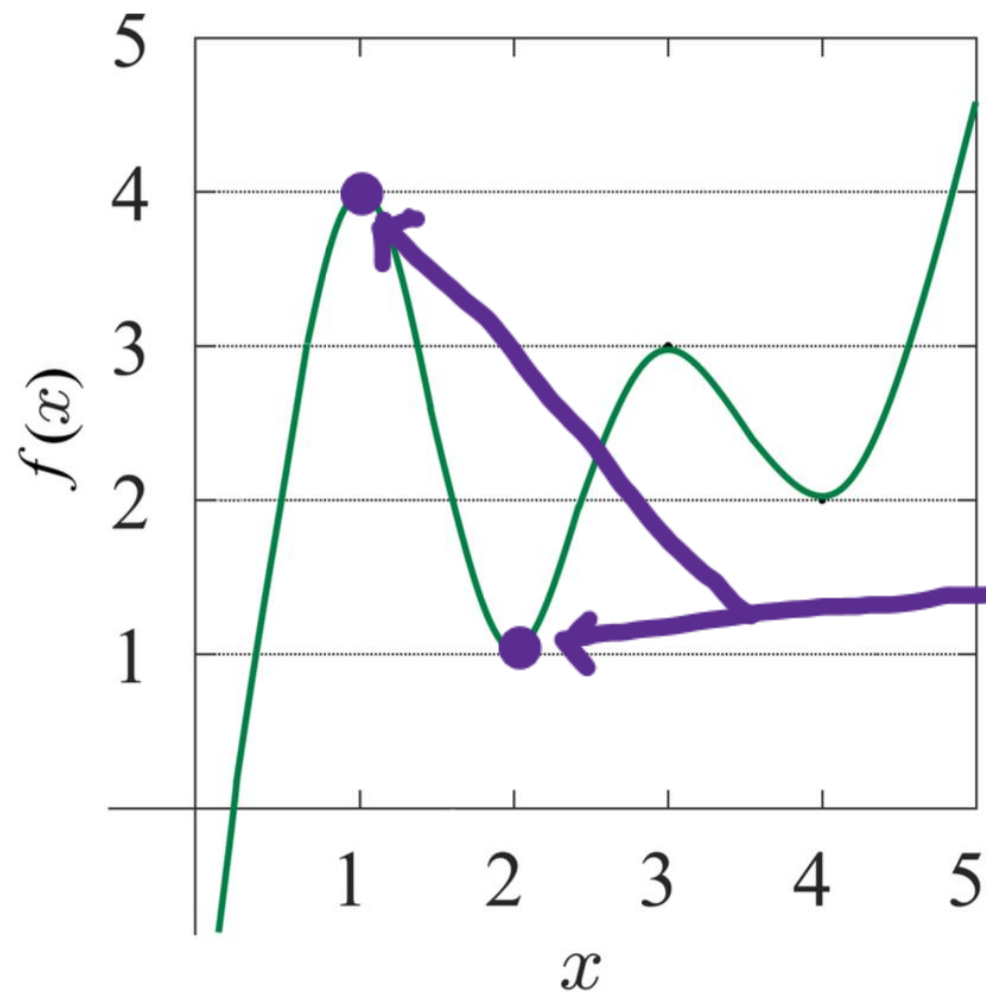
- We have three points in the 0-dimension PD
- Each point is tracking the birth and death of a connect component (or gap in between)



• PD: $(-\infty, +\infty)$

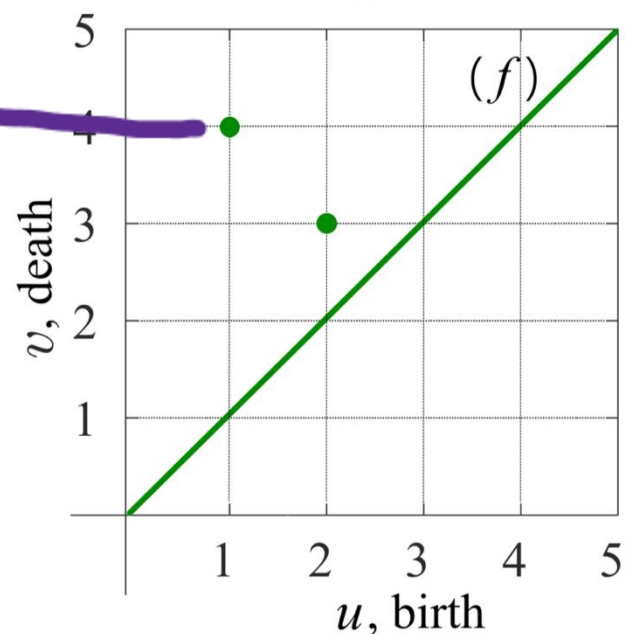
• PD: $(1.0, 4.0)$

• PD: $(2.0, 3.0)$



Summary:

- We have three points in the 0-dimension PD
- Each point is tracking the birth and death of a connect component (or gap in between)



• PD: $(-\infty, +\infty)$

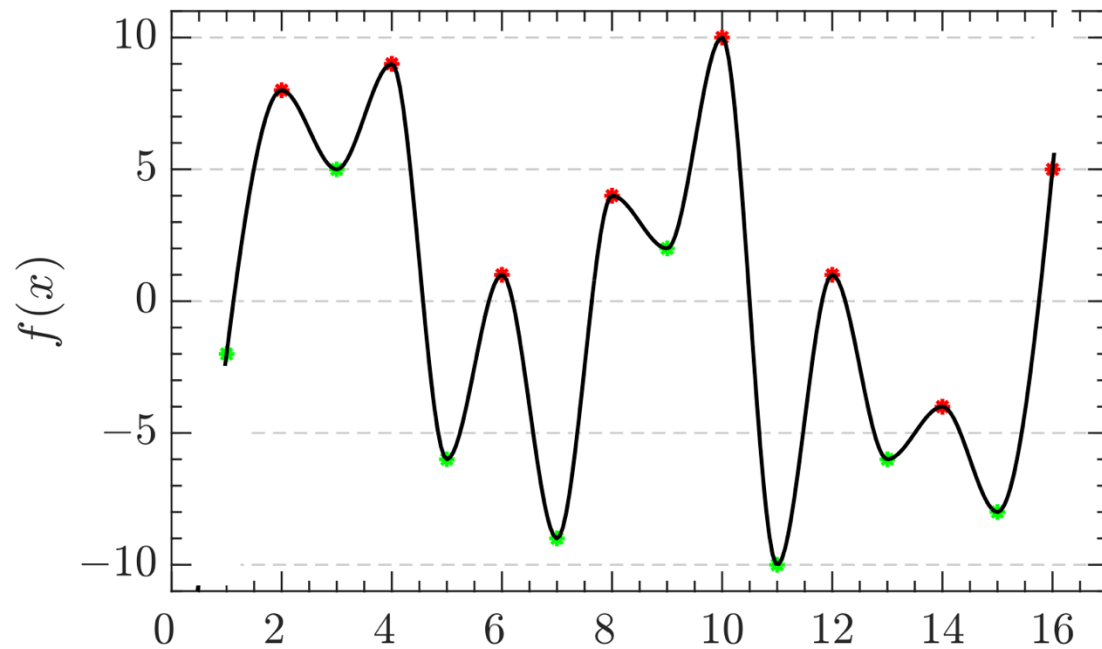
• PD: $(1.0, 4.0)$

• PD: $(2.0, 3.0)$

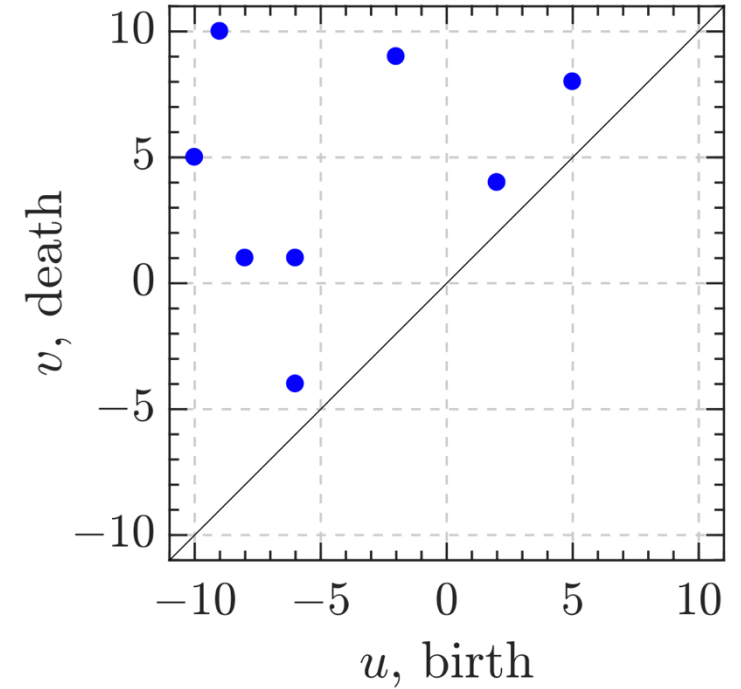
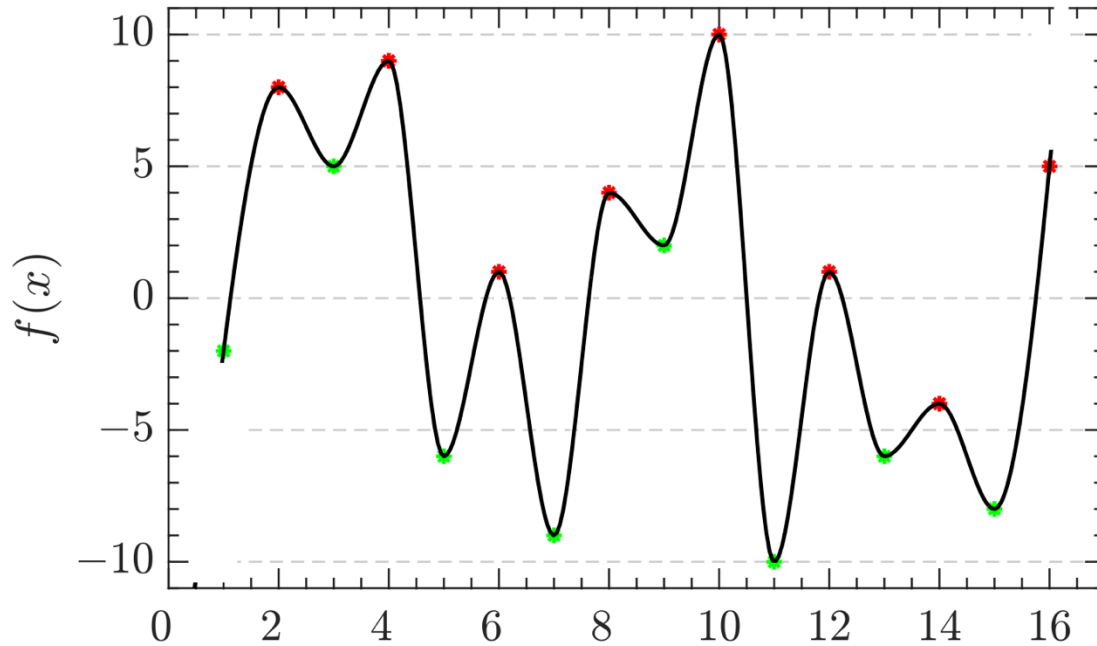
Online resources

- A webpage for visualizing 0-th PD: https://gjkoplik.github.io/pers-hom-examples/0d_pers_2d_data_widget.html

A similar but more involved example



A similar but more involved example



Persistent homology on 2D function

- Let's visualize another example on a 2D function

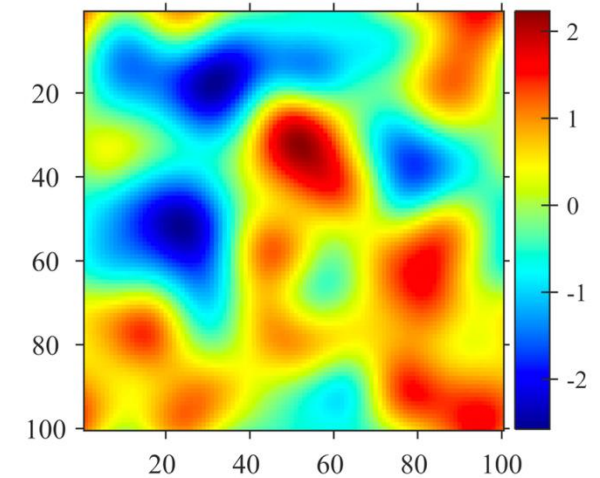
$$f: \mathbb{R}^2 \rightarrow \mathbb{R}$$

Persistent homology on 2D function

- Let's visualize another example on a 2D function

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}$$

- Right is an example where the value is indicated by color (**red** for high and **blue** for low)

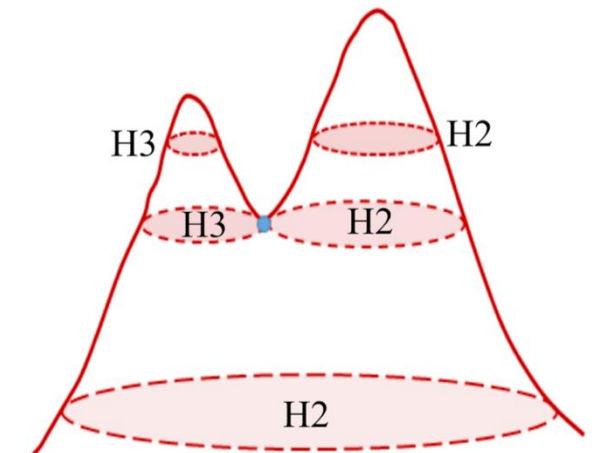
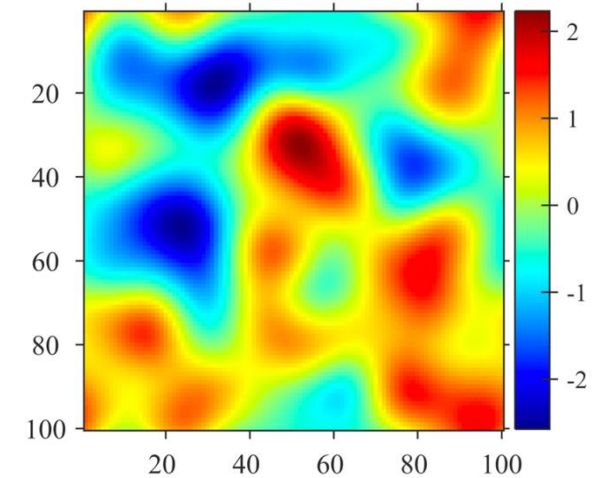


Persistent homology on 2D function

- Let's visualize another example on a 2D function

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}$$

- Right is an example where the value is indicated by color (red for high and blue for low)
- You can also treat the value on each point of \mathbb{R}^2 as a “height”, and plot the function like the bottom one

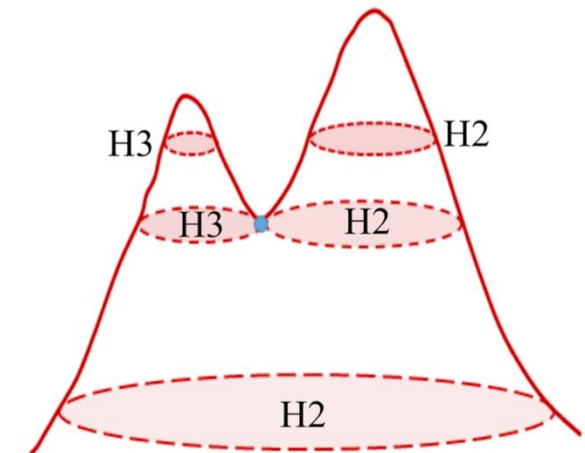
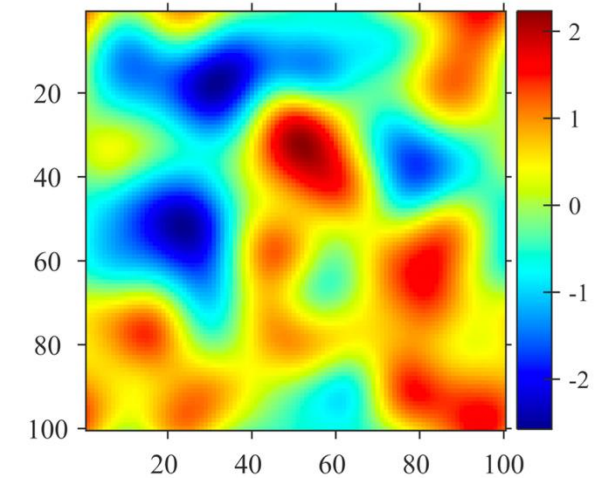


Persistent homology on 2D function

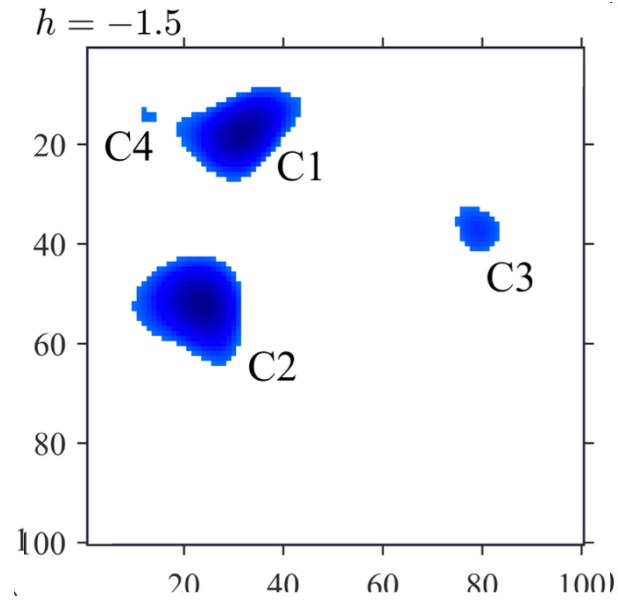
- Let's visualize another example on a 2D function

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}$$

- Right is an example where the value is indicated by color (red for high and blue for low)
- You can also treat the value on each point of \mathbb{R}^2 as a “height”, and plot the function like the bottom one
- Similar to the previous 1D function, as we increase the value α , we consider the part (subset) of the domain \mathbb{R}^2 whose values are below α

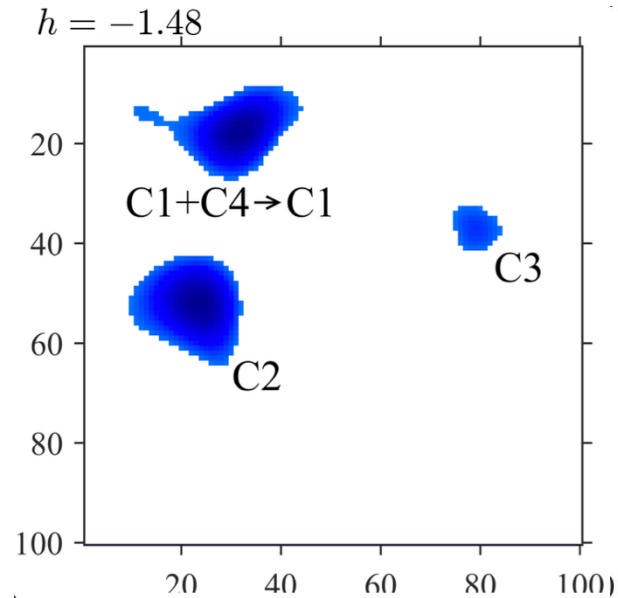


Persistent homology on 2D function



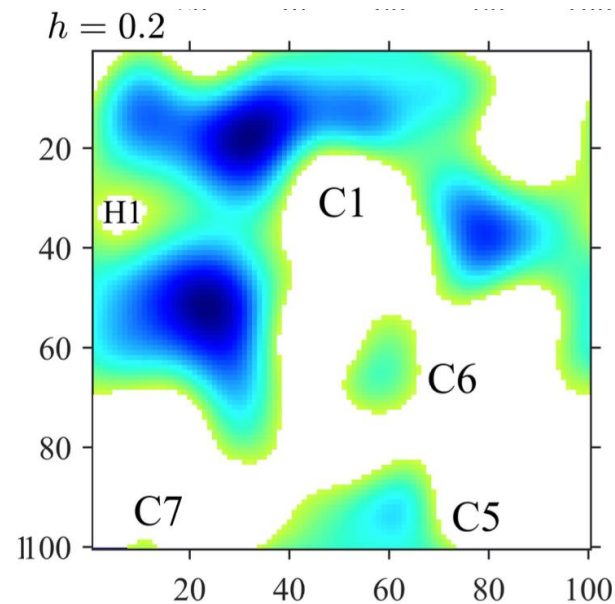
- Similar to the previous 1D function, as we increase the value α , we consider the part (subset) of the domain \mathbb{R}^2 whose values are below α

Persistent homology on 2D function



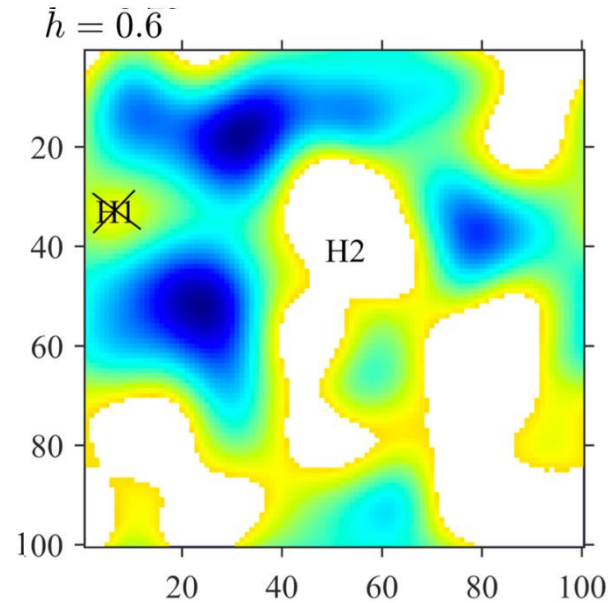
- Similar to the previous 1D function, as we increase the value α , we consider the part (subset) of the domain \mathbb{R}^2 whose values are below α

Persistent homology on 2D function



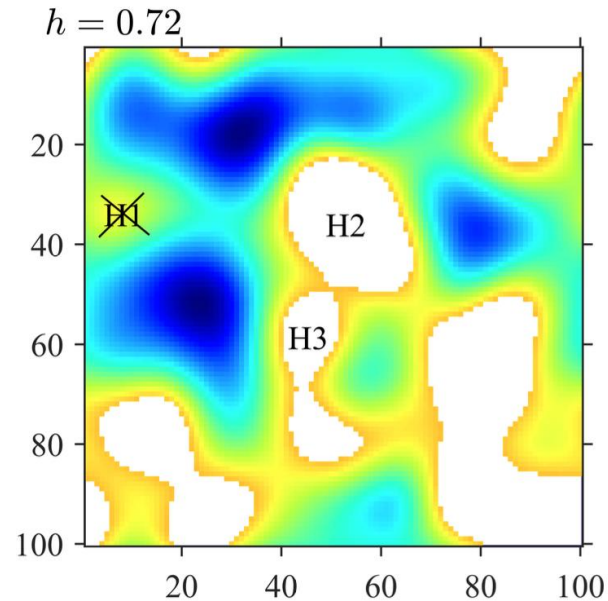
- Similar to the previous 1D function, as we increase the value α , we consider the part (subset) of the domain \mathbb{R}^2 whose values are below α

Persistent homology on 2D function



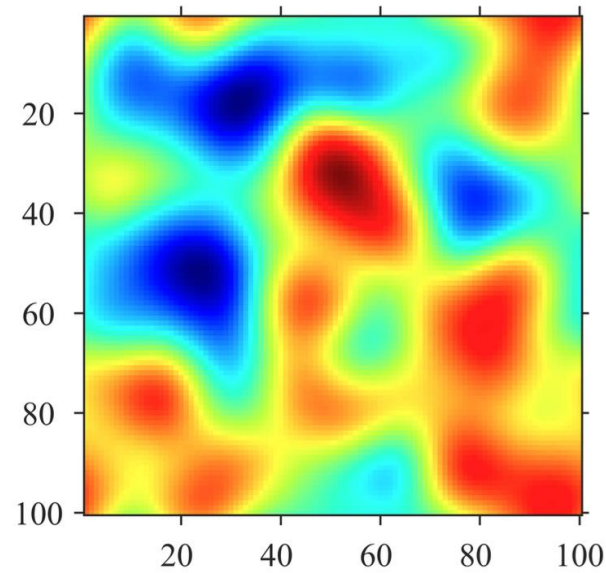
- Similar to the previous 1D function, as we increase the value α , we consider the part (subset) of the domain \mathbb{R}^2 whose values are below α

Persistent homology on 2D function



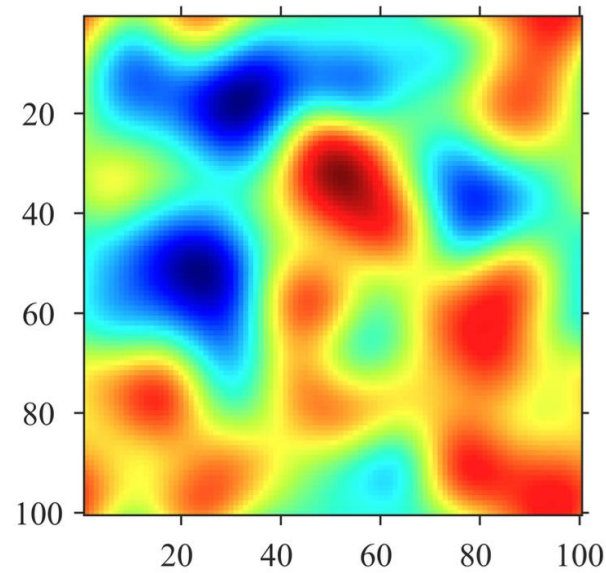
- Similar to the previous 1D function, as we increase the value α , we consider the part (subset) of the domain \mathbb{R}^2 whose values are below α

Persistent homology on 2D function



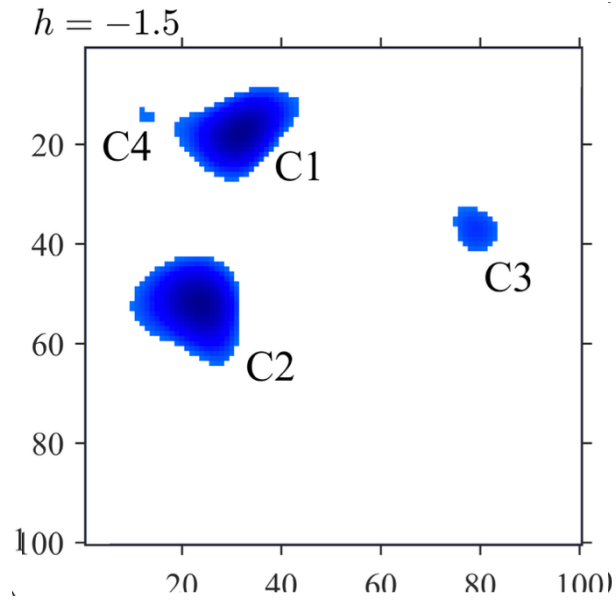
- Similar to the previous 1D function, as we increase the value α , we consider the part (subset) of the domain \mathbb{R}^2 whose values are below α

Persistent homology on 2D function



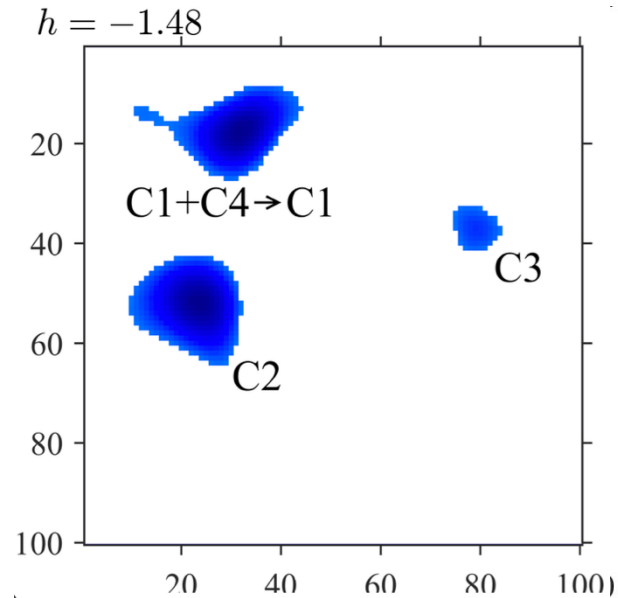
- Similar to the previous 1D function, as we increase the value α , we consider the part (subset) of the domain \mathbb{R}^2 whose values are below α
- Now let's track the birth and death of 0D/1D holes

Persistent homology on 2D function



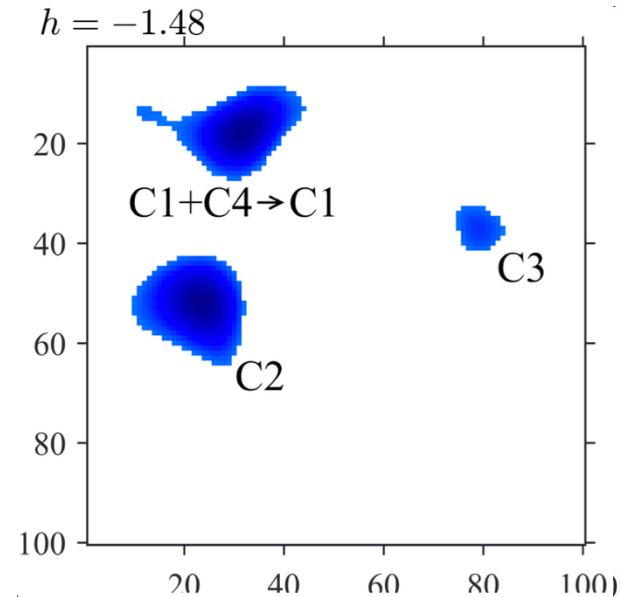
- Four connected components are born at different values
- (Will not display the birth of each component though)

Persistent homology on 2D function



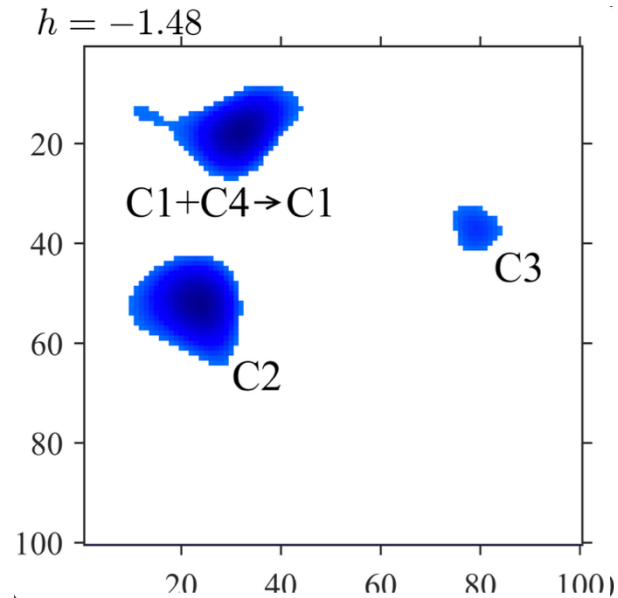
- C1 and C4 merged into the same connected component, thus the gap between them is filled

Persistent homology on 2D function



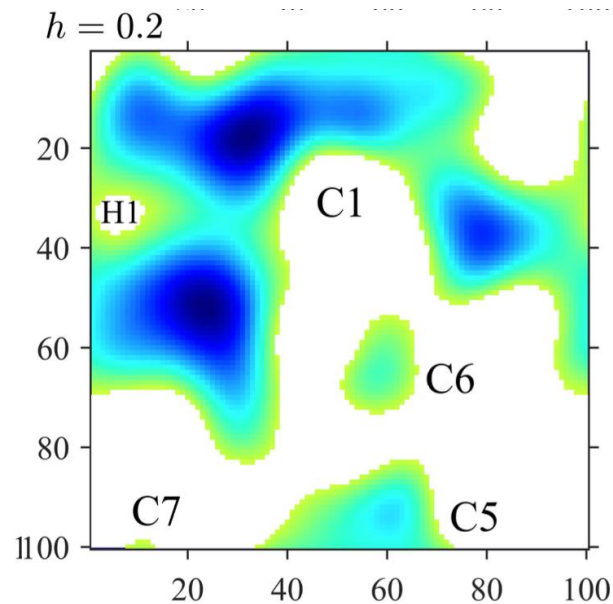
- C1 and C4 merged into the same connected component, thus the gap between them is filled
- Since C1 is born earlier, we keep C1 and kill C4 (the rule adopted by persistent homology)

Persistent homology on 2D function



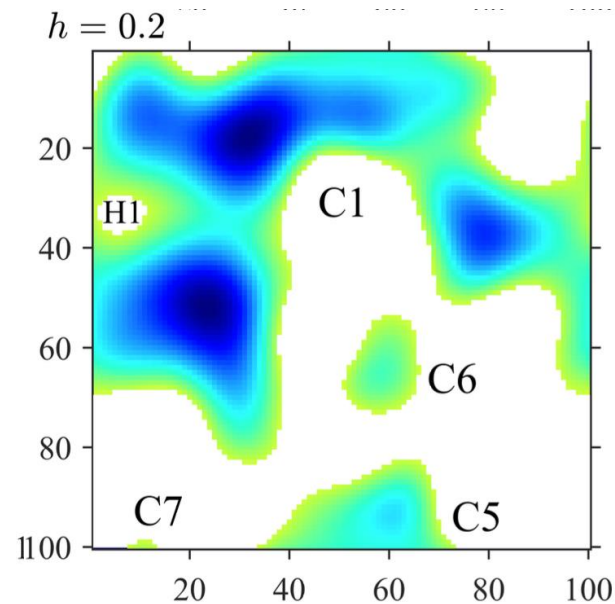
- C1 and C4 merged into the same connected component, thus the gap between them is filled
- Since C1 is born earlier, we keep C1 and kill C4 (the rule adopted by persistent homology)
- We then add a point (b, d) to the 0-d PD where b is the value in which C4 is born and d is current values where C4 dies (merges with other)

Persistent homology on 2D function



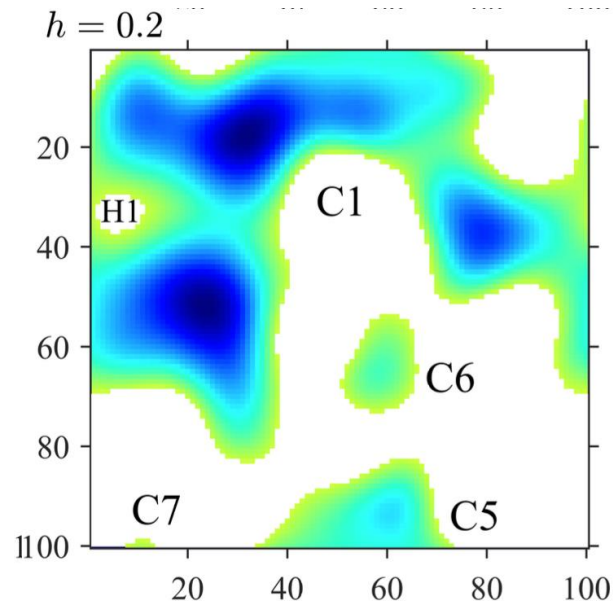
- As the value keep increasing, C1, C2 and C3 merged into the same connected component, producing two additional points in 0-d PD

Persistent homology on 2D function



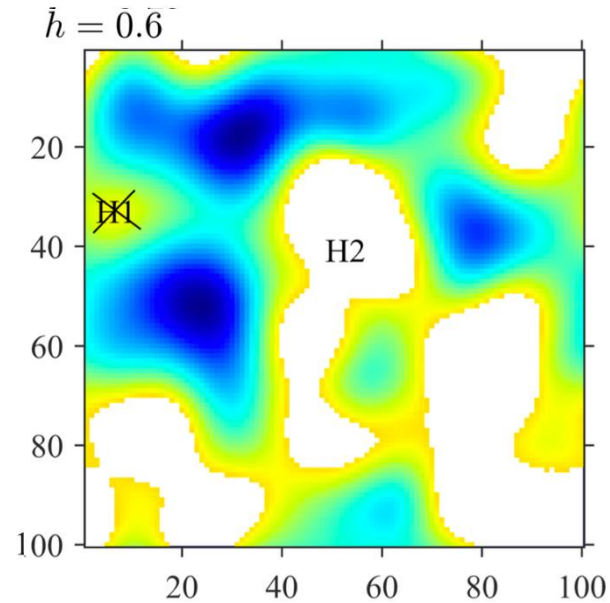
- As the value keep increasing, C1, C2 and C3 merged into the same connected component, producing two additional points in 0-d PD
- Three additional components C5, C6 and C7 are born

Persistent homology on 2D function



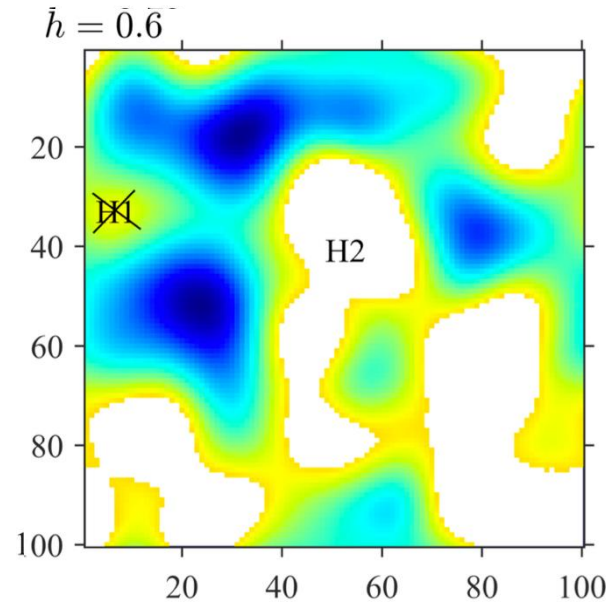
- As the value keep increasing, C1, C2 and C3 merged into the same connected component, producing two additional points in 0-d PD
- Three additional components C5, C6 and C7 are born
- Also, a 1-dimensional hole H1 is born

Persistent homology on 2D function



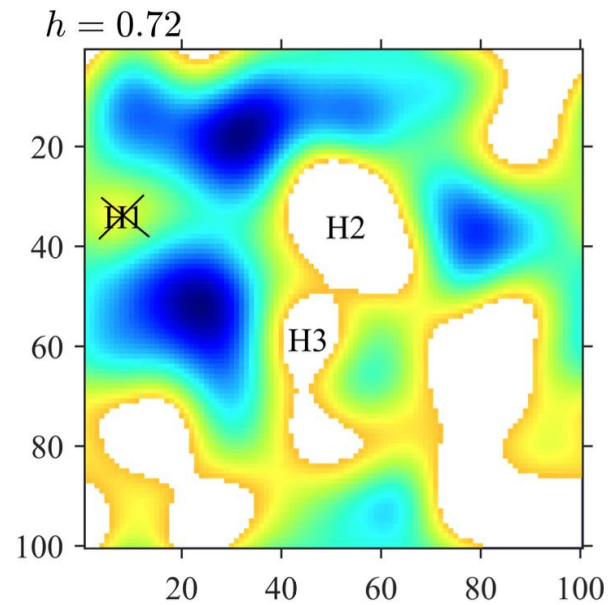
- H1 dies, producing a point in the 1-d PD

Persistent homology on 2D function



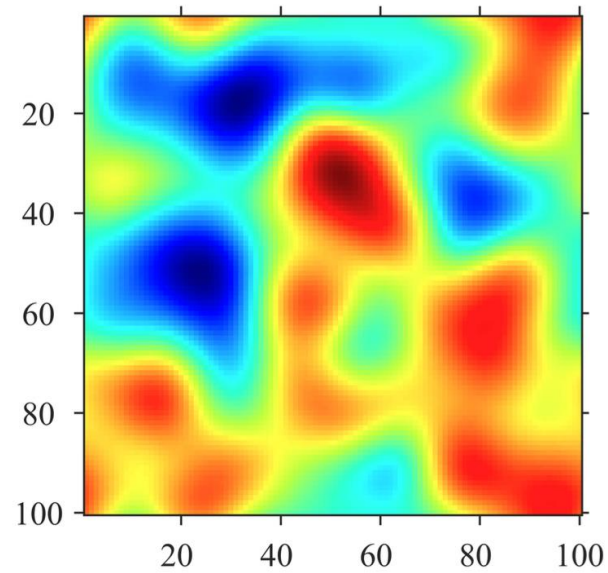
- H1 dies, producing a point in the 1-d PD
- A 1-dimensional hole H2 is born

Persistent homology on 2D function



- A 1-dimensional hole H3 is born

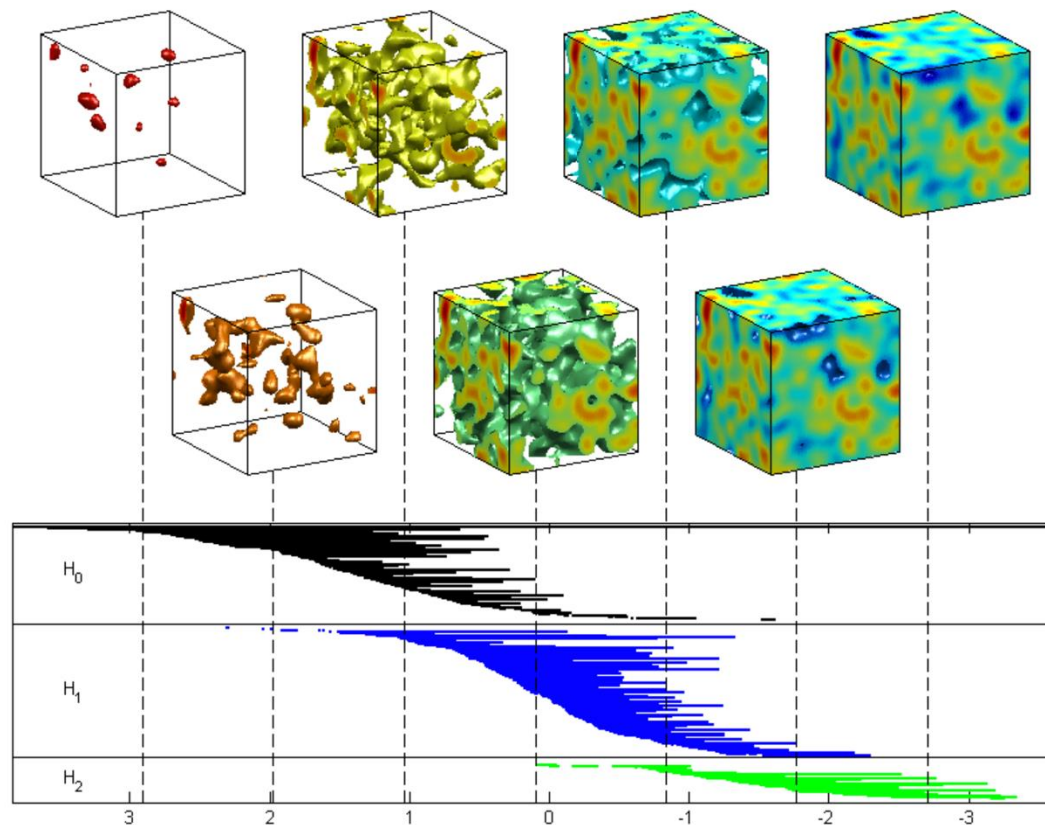
Persistent homology on 2D function



- H_2 and H_3 die, producing two additional points in the 1-d PD

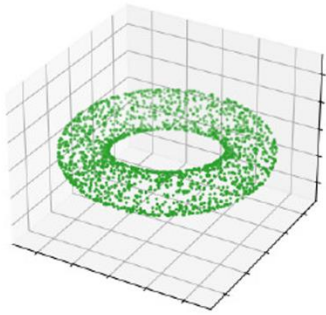
Persistent homology on 3D function

- We can also extend the prev. idea and define persistence on 3D function:
 $f: \mathbb{R}^3 \rightarrow \mathbb{R}$
- Similarly, as we increase the value α , we consider the part (subset) of the domain \mathbb{R}^3 (or a cube) whose values are below α

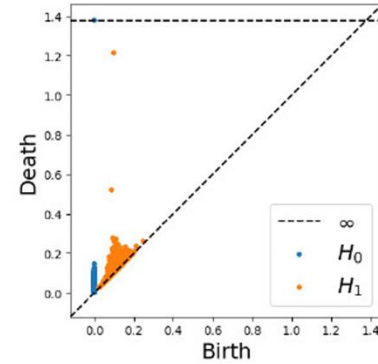


Adler, Robert J., Omer Bobrowski, Matthew S. Borman, Eliran Subag, and Shmuel Weinberger. "Persistent homology for random fields and complexes."

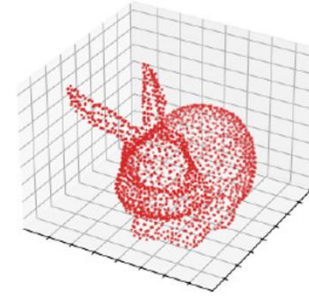
Persistence diagrams for differentiating point clouds



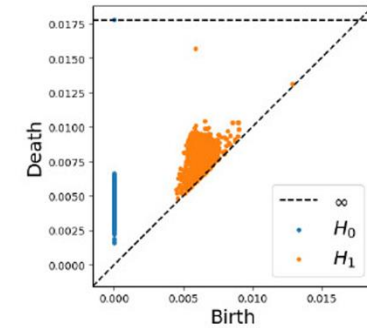
(c) 2000 points on a 3D torus.



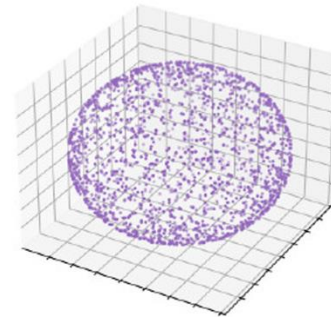
(d) Corresponding diagram.



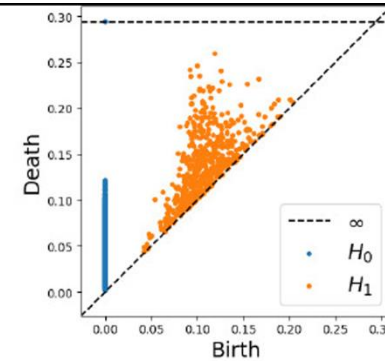
(e) Stanford bunny with 1889 points.



(f) Corresponding diagram.

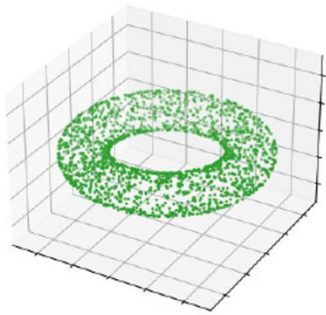


(a) 2000 points on a 3D sphere.

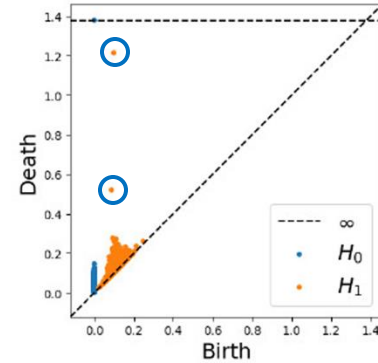


(b) Corresponding diagram.

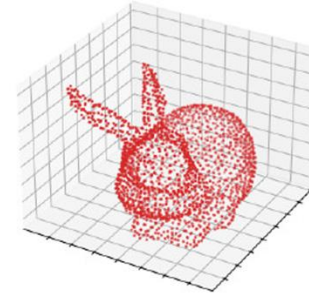
Persistence diagrams for differentiating point clouds



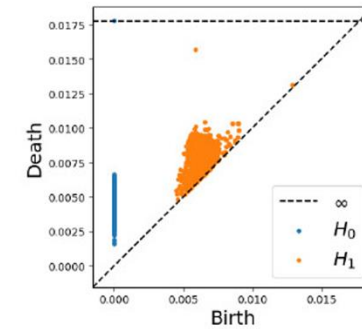
(c) 2000 points on a 3D torus.



(d) Corresponding diagram.

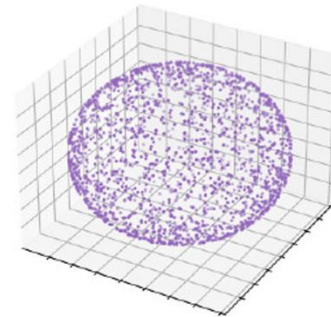


(e) Stanford bunny with 1889 points.

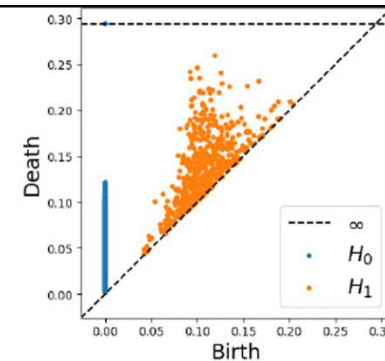


(f) Corresponding diagram.

Corresponding to
meridian and
longitude

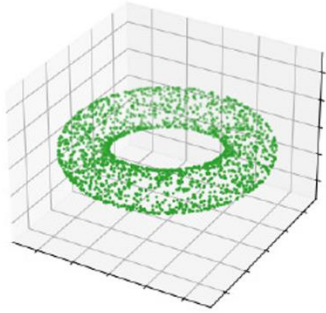


(a) 2000 points on a 3D sphere.

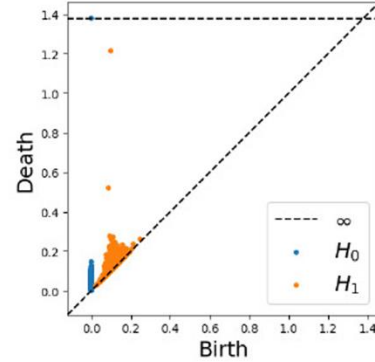


(b) Corresponding diagram.

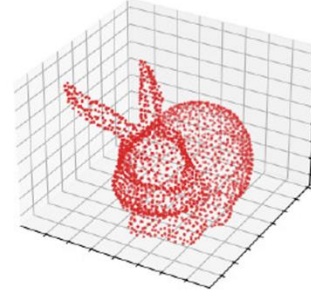
Persistence diagrams for differentiating point clouds



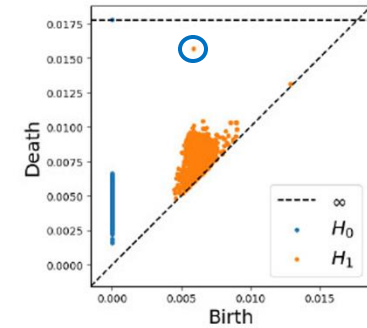
(c) 2000 points on a 3D torus.



(d) Corresponding diagram.

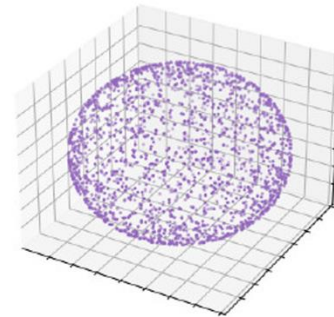


(e) Stanford bunny with 1889 points.

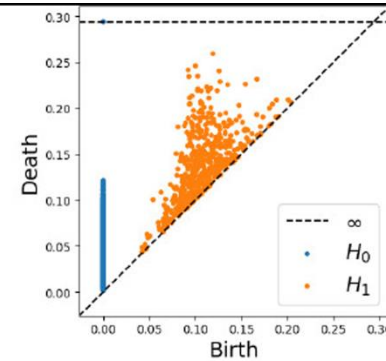


(f) Corresponding diagram.

Corresponds to the
“crust” of the bunny
which is a 2D hole

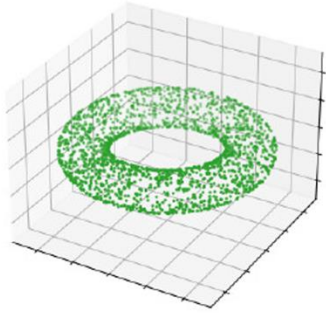


(a) 2000 points on a 3D sphere.

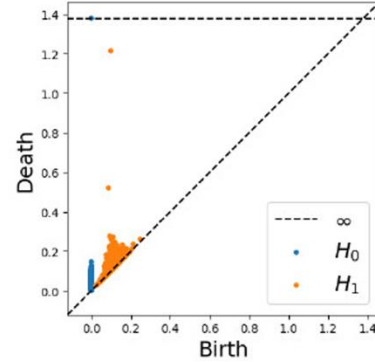


(b) Corresponding diagram.

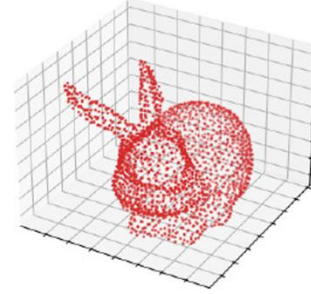
Persistence diagrams for differentiating point clouds



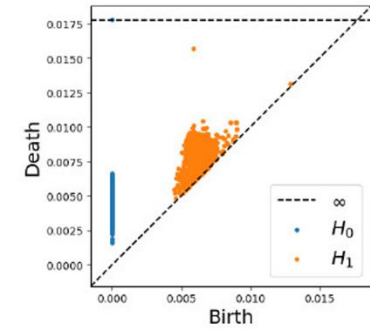
(c) 2000 points on a 3D torus.



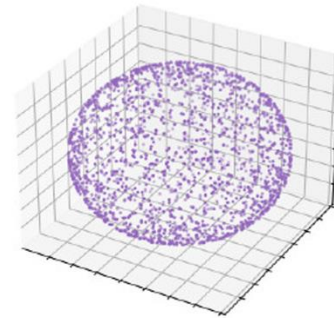
(d) Corresponding diagram.



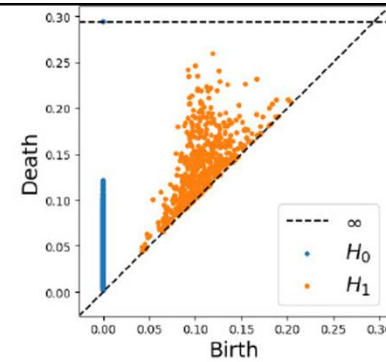
(e) Stanford bunny with 1889 points.



(f) Corresponding diagram.



(a) 2000 points on a 3D sphere.



(b) Corresponding diagram.

This is a solid ball
which has no
interesting holes

Persistence barcode

Recall:

- **Definition:** A **persistence diagram** (PD) is a set of points on the 2D plane above the diagonal such that for each point (b, d) :
 - b indicates birth value (the α value in which the feature is born)
 - d indicates death value (the α value in which the feature is dies)

Persistence barcode

Recall:

- **Definition:** A **persistence diagram** (PD) is a set of points on the 2D plane above the diagonal such that for each point (b, d) :
 - b indicates birth value (the α value in which the feature is born)
 - d indicates death value (the α value in which the feature is dies)
- **Definition:** If we draw each point (b, d) as a (left-closed, right open) interval $[b, d)$ on the real line, then what we get is a **persistence barcode** (so it's just persistence diagram interpreted differently)

Persistence barcode

Recall:

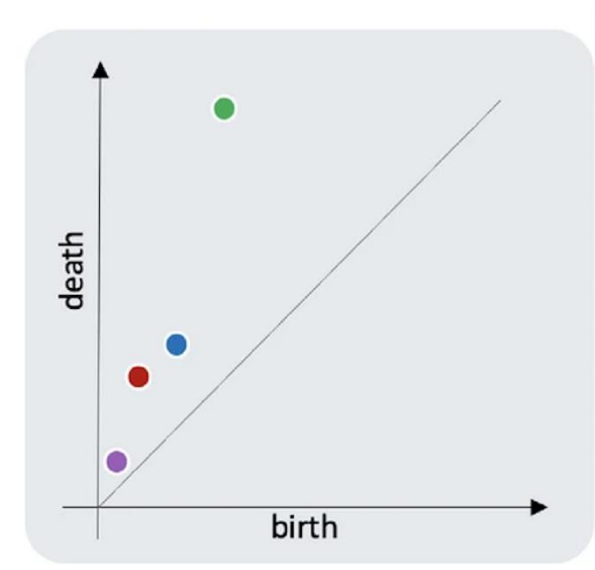
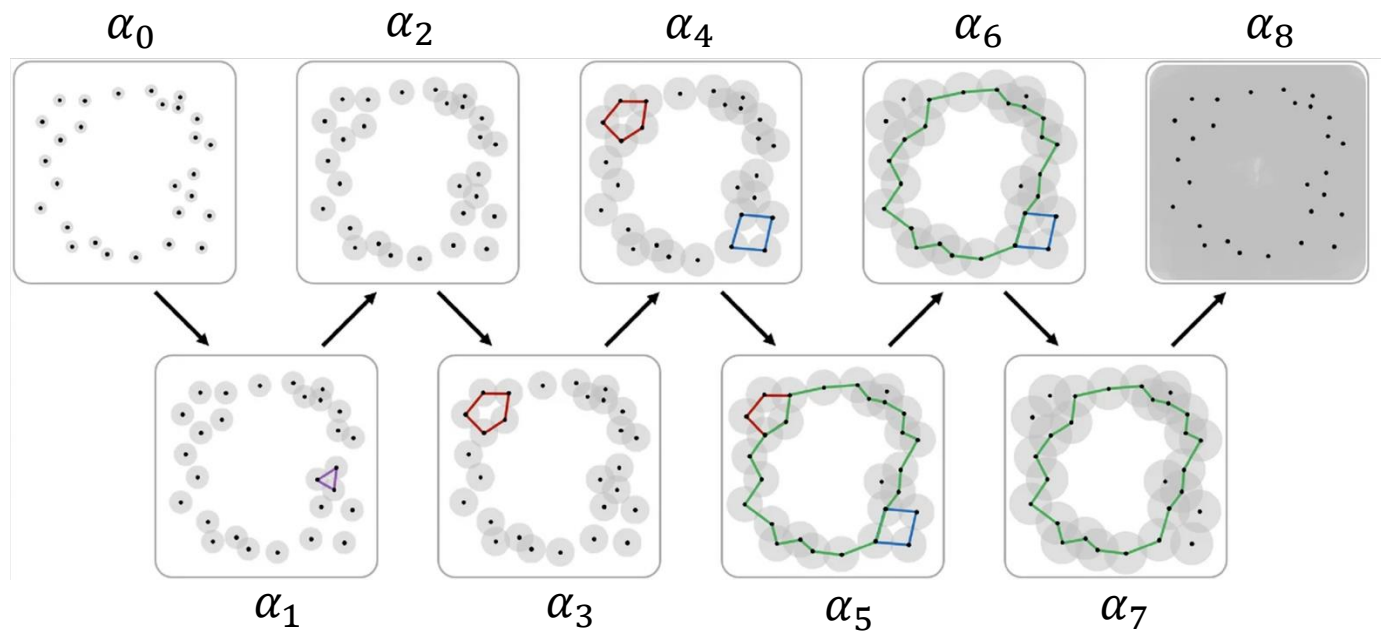
- **Definition:** A **persistence diagram** (PD) is a set of points on the 2D plane above the diagonal such that for each point (b, d) :
 - b indicates birth value (the α value in which the feature is born)
 - d indicates death value (the α value in which the feature is dies)
- **Definition:** If we draw each point (b, d) as a (left-closed, right open) interval $[b, d)$ on the real line, then what we get is a **persistence barcode** (so it's just persistence diagram interpreted differently)
- Sometimes drawing points in PD as interval is a very helpful for visualizing the change of homological features in your data

Persistence barcode

Recall:

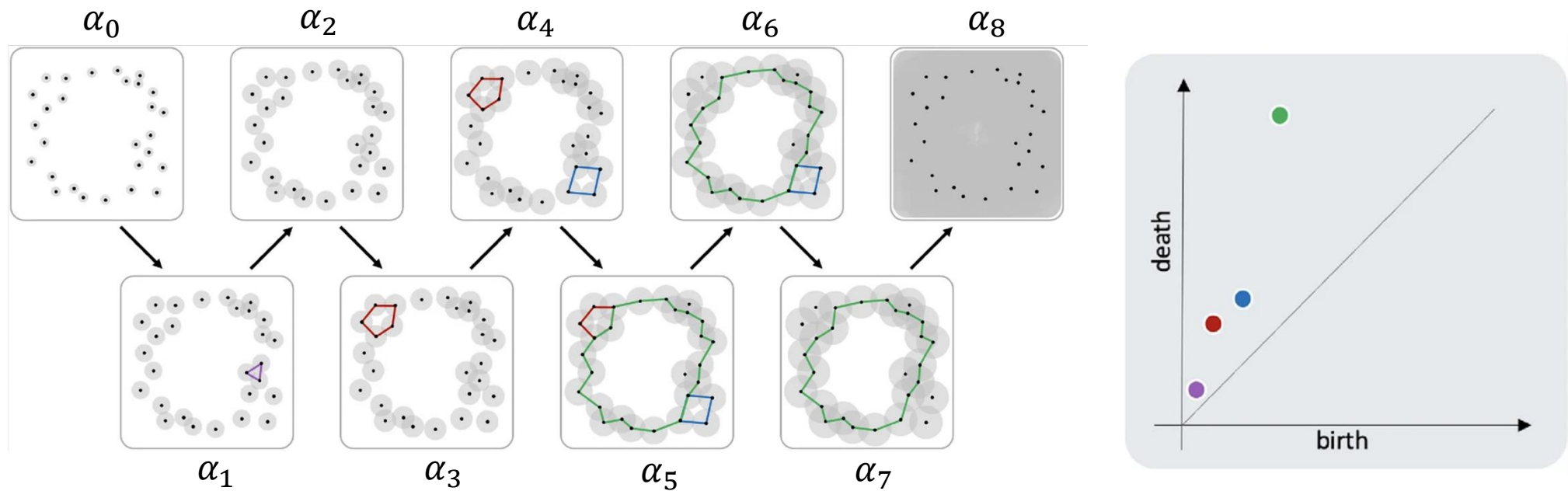
- **Definition:** A **persistence diagram** (PD) is a set of points on the 2D plane above the diagonal such that for each point (b, d) :
 - b indicates birth value (the α value in which the feature is born)
 - d indicates death value (the α value in which the feature is dies)
- **Definition:** If we draw each point (b, d) as a (left-closed, right open) interval $[b, d)$ on the real line, then what we get is a **persistence barcode** (so it's just persistence diagram interpreted differently)
- Sometimes drawing points in PD as interval is a very helpful for visualizing the change of homological features in your data
- Notice that we sometimes use the terms “persistence diagram” and “persistence barcode” interchangeable, i.e., we may call a point in a PD also an interval.

Example

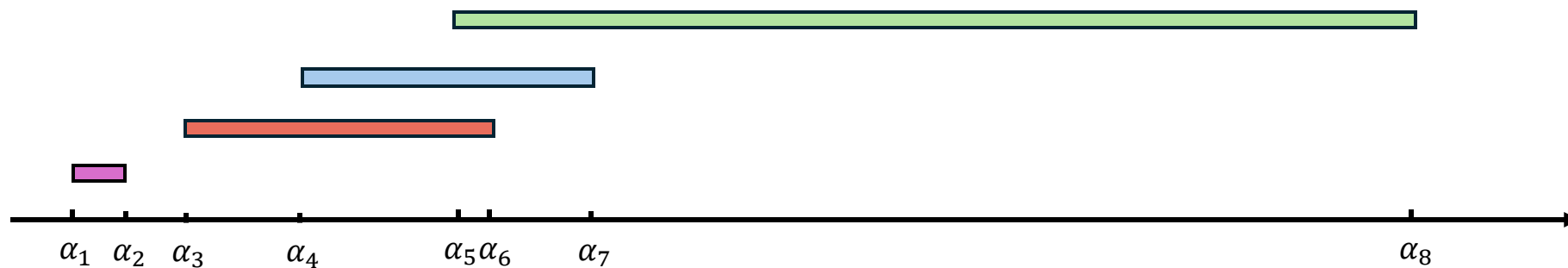


- Corresponding barcode:

Example



- Corresponding barcode:



Another example

