Case study: Image segmentation for accurate topology inference

Tao Hou, University of Oregon

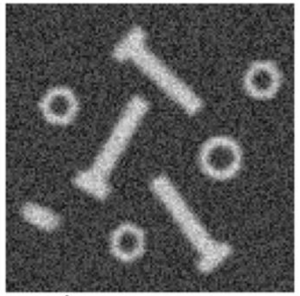
Image segmentation for correct topology inference

This application is documented in the following paper

- Chung, Yu-Min, and Sarah Day. "Topological fidelity and image thresholding: A persistent homology approach." Journal of Mathematical Imaging and Vision 60 (2018): 1167-1179.

Persistent homology pipeline

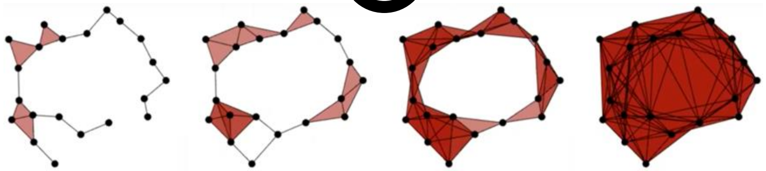
1



Gray-scale Images

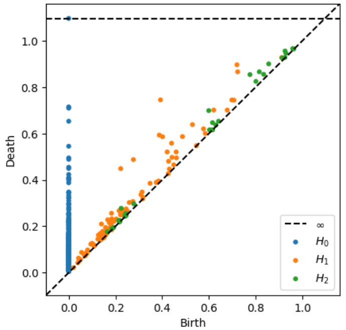
Build Filtration

2



Sublevelset Filtration

Persistent Homology



3

PD / barcode

Automatic algorithm for determining threshold

4

Segmented image for homology inference

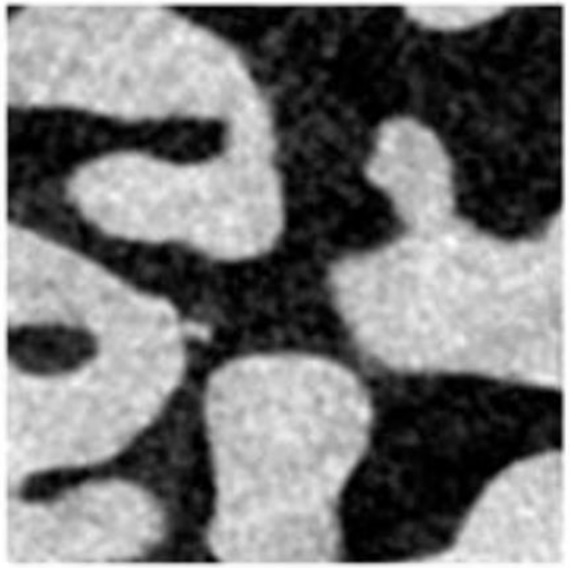
Some img from: AATRN; https://quantdare.com/understanding-the-shape-of-data-ii/

Background

- Primary motivation / goal: extract useful and accurate topological information from noisy digital images of structures combining two materials or subdomains of interest

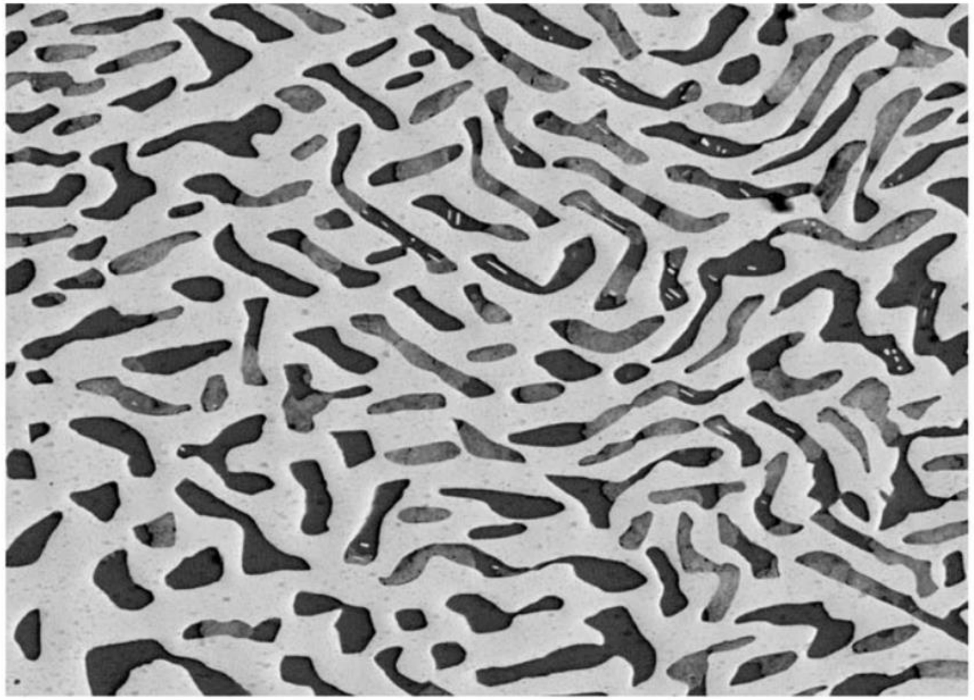
Background

- Primary motivation / goal: extract useful and accurate topological information from noisy digital images of structures combining two materials or subdomains of interest
- Ex1: Firn (Ice and Pore)



Background

- Primary motivation / goal: extract useful and accurate topological information from noisy digital images of structures combining two materials or subdomains of interest
- Ex2: Bi–Sn (Bismuth Tin) binary alloy



Background

- They focus on using persistent homology to infer the Betti number β_i for each applicable i, which counts holes of various dimensions

Background

- They focus on using persistent homology to infer the Betti number β_i for each applicable i, which counts holes of various dimensions
- While the demonstrated examples (images) are in 2D, the image datasets are in fact in 3D

Background

- They focus on using persistent homology to infer the Betti number β_i for each applicable i, which counts holes of various dimensions
- While the demonstrated examples (images) are in 2D, the image datasets are in fact in 3D
- For 3D images, holes in the various dimensions correspond to the following:
  - β_0, # of 0-dimensional holes: connected components

Background

- They focus on using persistent homology to infer the Betti number β_i for each applicable i, which counts holes of various dimensions
- While the demonstrated examples (images) are in 2D, the image datasets are in fact in 3D
- For 3D images, holes in the various dimensions correspond to the following:
  - β_0, # of 0-dimensional holes: connected components
  - β_1, # of 1-dimensional holes: “tunnels”

Background

- They focus on using persistent homology to infer the Betti number β_i for each applicable i, which counts holes of various dimensions
- While the demonstrated examples (images) are in 2D, the image datasets are in fact in 3D
- For 3D images, holes in the various dimensions correspond to the following:
  - β_0, # of 0-dimensional holes: connected components
  - β_1, # of 1-dimensional holes: "tunnels"
  - β_2, # of 2-dimensional holes: "cavities"

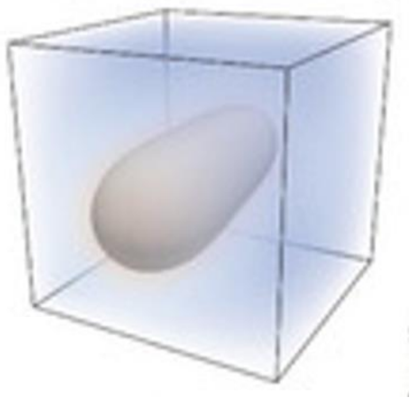


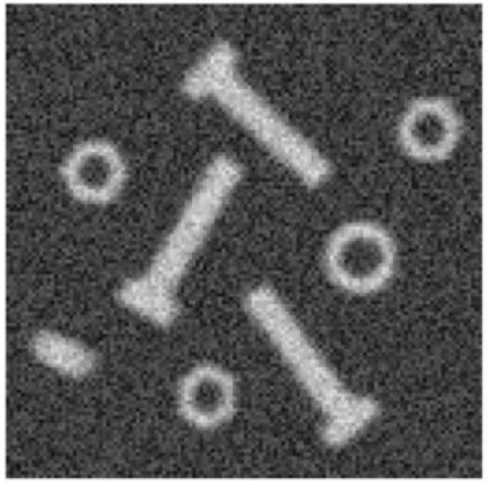
Image source: Eckert et al. The Topology of Pediatric Structural Asymmetries in Language-Related Cortex

Binarization

- What they aim to do is to segment (binarize) each pixel of image into 0 or 1, each representing one material (subdomain) of the two

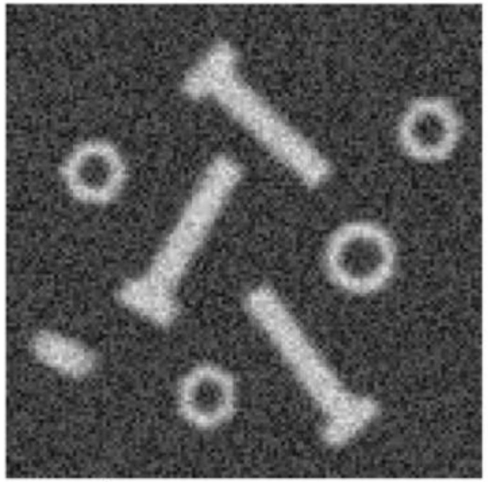
Binarization

- What they aim to do is to segment (binarize) each pixel of image into 0 or 1, each representing one material (subdomain) of the two



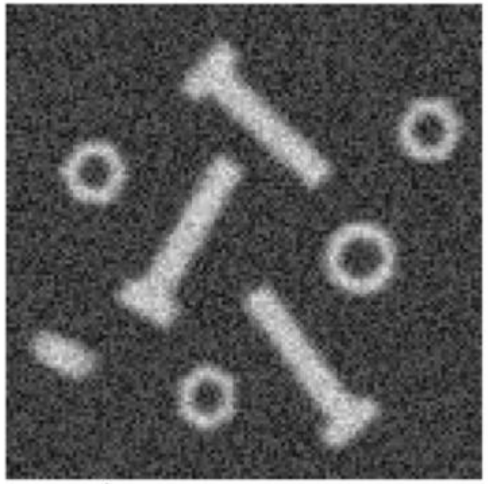
Binarization

- What they aim to do is to segment (binarize) each pixel of image into 0 or 1, each representing one material (subdomain) of the two



Binarization

- What they aim to do is to segment (binarize) each pixel of image into 0 or 1, each representing one material (subdomain) of the two
- After the binarization, we can then study the property of each region



Why topology matters here?

- In above-mentioned example applications (firn or alloy), topological information can serve as a measurement of complexity of the structure in addition to offering more detailed information of interest

Why topology matters here?

- In above-mentioned example applications (firn or alloy), topological information can serve as a measurement of complexity of the structure in addition to offering more detailed information of interest
- For firn, understanding the connectivity of the pore space and ensuring topological accuracy of the pore directly impacts the accuracy of model simulations of gas transfer in the firn column

Why topology matters here?

- In above-mentioned example applications (firn or alloy), topological information can serve as a measurement of complexity of the structure in addition to offering more detailed information of interest
- For firn, understanding the connectivity of the pore space and ensuring topological accuracy of the pore directly impacts the accuracy of model simulations of gas transfer in the firn column
    - The pore consists of tunnels (1-dimensional holes in the ice) and cavities or air pockets (2-dimensional holes in the ice).

Why topology matters here?

- In above-mentioned example applications (firn or alloy), topological information can serve as a measurement of complexity of the structure in addition to offering more detailed information of interest
- For firn, understanding the connectivity of the pore space and ensuring topological accuracy of the pore directly impacts the accuracy of model simulations of gas transfer in the firn column
  - The pore consists of tunnels (1-dimensional holes in the ice) and cavities or air pockets (2-dimensional holes in the ice).
  - Gas molecules travel through tunnels

Why topology matters here?

- In above-mentioned example applications (firn or alloy), topological information can serve as a measurement of complexity of the structure in addition to offering more detailed information of interest
- For firn, understanding the connectivity of the pore space and ensuring topological accuracy of the pore directly impacts the accuracy of model simulations of gas transfer in the firn column
  - The pore consists of tunnels (1-dimensional holes in the ice) and cavities or air pockets (2-dimensional holes in the ice).
  - Gas molecules travel through tunnels
  - Tunnel can pinch off into cavities under the pressure of accumulated snow and thus the gas may get trapped

Why topology matters here?

- In above-mentioned example applications (firn or alloy), topological information can serve as a measurement of complexity of the structure in addition to offering more detailed information of interest
- For firn, understanding the connectivity of the pore space and ensuring topological accuracy of the pore directly impacts the accuracy of model simulations of gas transfer in the firn column
  - The pore consists of tunnels (1-dimensional holes in the ice) and cavities or air pockets (2-dimensional holes in the ice).
  - Gas molecules travel through tunnels
  - Tunnel can pinch off into cavities under the pressure of accumulated snow and thus the gas may get trapped
  - If tunnels are not correctly detected due to topological inaccuracies, then gas transfer simulations will underestimate the propagation of gases through the firn

Why topology matters here?

- In above-mentioned example applications (firn or alloy), topological information can serve as a measurement of complexity of the structure in addition to offering more detailed information of interest
- For firn, understanding the connectivity of the pore space and ensuring topological accuracy of the pore directly impacts the accuracy of model simulations of gas transfer in the firn column
    - The pore consists of tunnels (1-dimensional holes in the ice) and cavities or air pockets (2-dimensional holes in the ice).
    - Gas molecules travel through tunnels
    - Tunnel can pinch off into cavities under the pressure of accumulated snow and thus the gas may get trapped
    - If tunnels are not correctly detected due to topological inaccuracies, then gas transfer simulations will underestimate the propagation of gases through the firn
    - Accurate estimates of gas transfer are of interest to climate scientists seeking to understand the age of atmospheric gases that have been trapped in air pockets within firn core samples

Thresholding

- While there have been many image segmentation / binarization methods developed, the work focuses on a simple approach using a threshold, which has proven powerful in practice

Thresholding

- While there have been many image segmentation / binarization methods developed, the work focuses on a simple approach using a threshold, which has proven powerful in practice
- Given a threshold θ, for each pixel:
  - If the gray scale value is less than or equal to θ, then assign material 0 to the pixel
  - Otherwise, assign material 1 (color 255)

Thresholding

- While there have been many image segmentation / binarization methods developed, the work focuses on a simple approach using a threshold, which has proven powerful in practice
- Given a threshold θ, for each pixel:
  - If the gray scale value is less than or equal to θ, then assign material 0 to the pixel
  - Otherwise, assign material 1 (color 255)
- A very famous and powerful thresholding approach for gray-scale image is the **Otsu's method**
  - It is a statistics-base method by inspecting the histogram of the image

Thresholding

- While there have been many image segmentation / binarization methods developed, the work focuses on a simple approach using a threshold, which has proven powerful in practice
- Given a threshold θ, for each pixel:
  - If the gray scale value is less than or equal to θ, then assign material 0 to the pixel
  - Otherwise, assign material 1 (color 255)
- A very famous and powerful thresholding approach for gray-scale image is the **Otsu's method**
  - It is a statistics-base method by inspecting the histogram of the image
  - This is also the method of comparison in the work

Otsu’s method

- A histogram is just a counting of the number of pixels falling in each of the gray-scale bin (0-255)

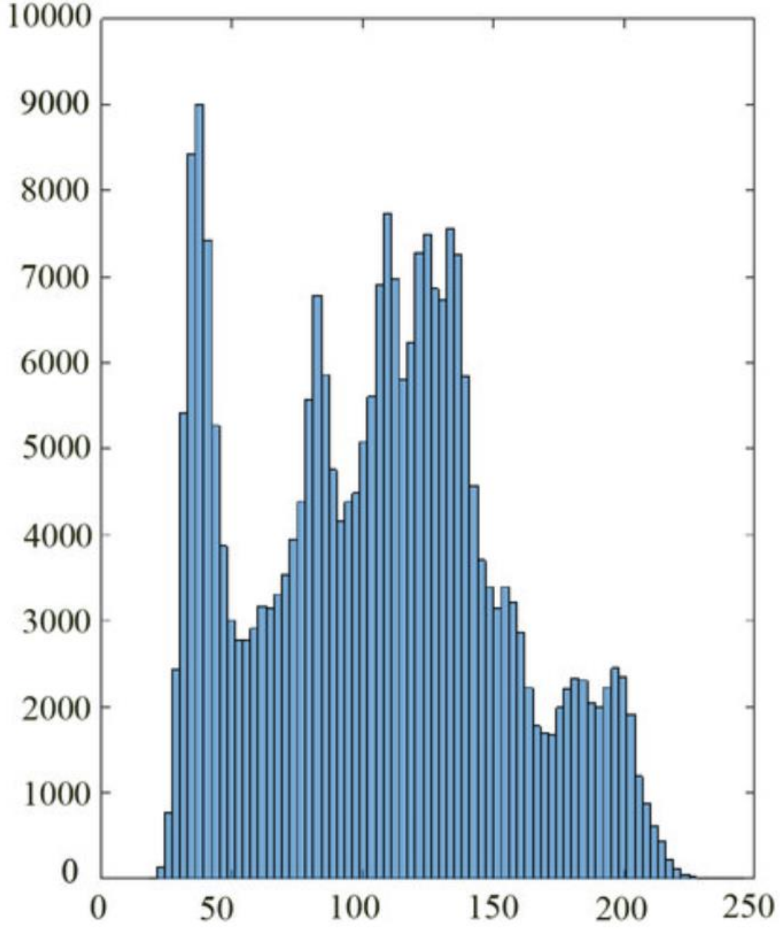
Otsu’s method

- A histogram is just a counting of the number of pixels falling in each of the gray-scale bin (0-255)

Heidari et al. A new general model for quantum image histogram (QIH)

Otsu’s method

- A histogram is just a counting of the number of pixels falling in each of the gray-scale bin (0-255)



Heidari et al. A new general model for quantum image histogram (QIH)

Otsu’s method

- The algorithm exhaustively searches for the threshold that minimizes the intra-class variance, defined as a weighted sum of variances of the two classes:

$$\sigma_w^2(t) = \omega_0(t)\sigma_0^2(t) + \omega_1(t)\sigma_1^2(t)$$

Otsu's method

- The algorithm exhaustively searches for the threshold that minimizes the intra-class variance, defined as a weighted sum of variances of the two classes:

$$\sigma_w^2(t) = \omega_0(t)\sigma_0^2(t) + \omega_1(t)\sigma_1^2(t)$$

- Essentially, given a histogram, it tries to find the "basin" in between two peaks (if this is true)
- See: https://en.wikipedia.org/wiki/Otsu%27s_method

Otsu's method

- The algorithm exhaustively searches for the threshold that minimizes the intra-class variance, defined as a weighted sum of variances of the two classes:

$$\sigma_w^2(t) = \omega_0(t)\sigma_0^2(t) + \omega_1(t)\sigma_1^2(t)$$

- Essentially, given a histogram, it tries to find the "basin" in between two peaks (if this is true)
- See: https://en.wikipedia.org/wiki/Otsu%27s_method
- While being widely successfully, Otsu's method is "topology-agnostic", making it unsuitable for applications where the accuracy of topology is of great importance, such as ones we are considering here

Otsu’s method

- The algorithm exhaustively searches for the threshold that minimizes the intra-class variance, defined as a weighted sum of variances of the two classes:

$$\sigma_w^2(t) = \omega_0(t)\sigma_0^2(t) + \omega_1(t)\sigma_1^2(t)$$

- Essentially, given a histogram, it tries to find the “basin” in between two peaks (if this is true)
- See: https://en.wikipedia.org/wiki/Otsu%27s_method
- While being widely successfully, Otsu’s method is “topology-agnostic”, making it unsuitable for applications where the accuracy of topology is of great importance, such as ones we are considering here
- We will later see examples where the topology inferences based on Otsu’s is far off

Achievements

- Given a digital image of two or more important subdomains, proposed an automated thresholding approach based on the PD

Achievements

- Given a digital image of two or more important subdomains, proposed an automated thresholding approach based on the PD
- It utilizes an optimization scheme so that important, robust topological features are present, but the number of features (holes) that are (most likely) due to noise is minimized

Achievements

- Given a digital image of two or more important subdomains, proposed an automated thresholding approach based on the PD
- It utilizes an optimization scheme so that important, robust topological features are present, but the number of features (holes) that are (most likely) due to noise is minimized
- As shown in the paper, the proposed method often leads to improved topological accuracy in the segmented image and in the resulting Betti numbers

Setting

- A grayscale image may be viewed as a function

$$f : P \to \mathbb{R}$$

where P is the set of pixels or voxels and $f(x)$ is the greyscale value at $x \in P$.

Setting

- A grayscale image may be viewed as a function

$$f : P \to \mathbb{R}$$

where P is the set of pixels or voxels and $f(x)$ is the greyscale value at $x \in P$.

- In practice, we may assume

$$f : P \to [0, I]$$

for some value $I \geq 0$. For example, $I = 255$ where f is an 8-bit greyscale image.

Setting

- A grayscale image may be viewed as a function

$$f : P \to \mathbb{R}$$

where P is the set of pixels or voxels and $f(x)$ is the greyscale value at $x \in P$.

- In practice, we may assume

$$f : P \to [0, I]$$

for some value $I \geq 0$. For example, $I = 255$ where f is an 8-bit greyscale image.

- For a threshold t, the work focuses on the black/dark region corresponding to this threshold, which is indeed the sublevel set of t:

$$f_t^- := \{x \in P \mid f(x) \leq t\}$$

Setting

- We then build the sublevelset filtration of *f*:

$$f^-_{t_0} \subseteq f^-_{t_1} \subseteq \cdots \subseteq f^-_{t_n}$$

- and compute the PD/barcode denoted $\mathscr{P}$

Setting

- We then build the sublevelset filtration of *f*:

$$f_{t_0}^{-} \subseteq f_{t_1}^{-} \subseteq \cdots \subseteq f_{t_n}^{-}$$

- and compute the PD/barcode denoted $\mathscr{P}$
- Recall: computing the number of intervals in the barcode $\mathscr{P}$ intersecting a threshold *t* gives the Betti number

Setting

- We then build the sublevelset filtration of *f*:

$$f_{t_0}^{-} \subseteq f_{t_1}^{-} \subseteq \cdots \subseteq f_{t_n}^{-}$$

- and compute the PD/barcode denoted $\mathscr{P}$
- Recall: computing the number of intervals in the barcode $\mathscr{P}$ intersecting a threshold *t* gives the Betti number
- which in turn is equivalent to counting the number of number of points in the upper-left quadrant of the point (*t*,*t*) in the PD $\mathscr{P}$

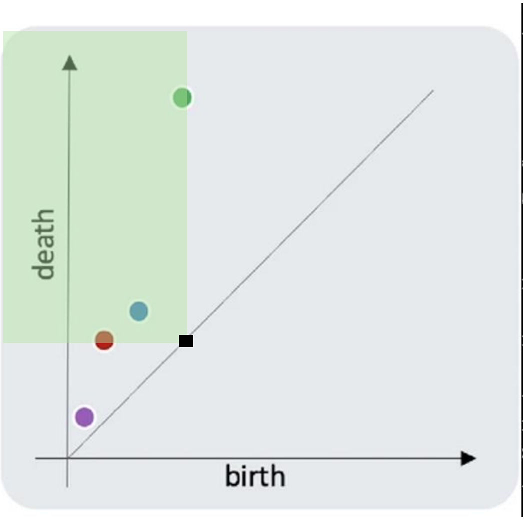


Image source: Bobrowski O, Skraba P. A universal null-distribution for topological data analysis.

Setting

- So, as we move the lower right corner of the quadrant along the diagonal and count the number of points falling within it, we track the variation of the Betti number $\beta(\hat{f}_t^-)$ of the sublevel sets in the filtration

$$f_{t_0}^- \subseteq f_{t_1}^- \subseteq \cdots \subseteq f_{t_n}^-$$

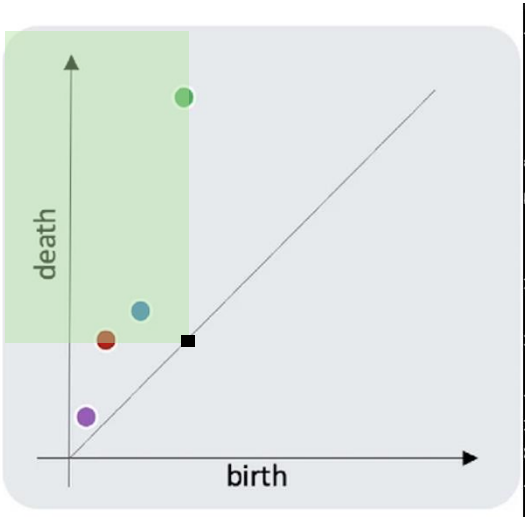


Image source: Bobrowski O, Skraba P. A universal null-distribution for topological data analysis.

Setting

- So, as we move the lower right corner of the quadrant along the diagonal and count the number of points falling within it, we track the variation of the Betti number $\beta(\hat{f}_t^-)$ of the sublevel sets in the filtration

$$f_{t_0}^- \subseteq f_{t_1}^- \subseteq \cdots \subseteq f_{t_n}^-$$

- Now the problem boils down to finding a *t* such that $\beta(\hat{f}_t^-)$ is the "ground truth"

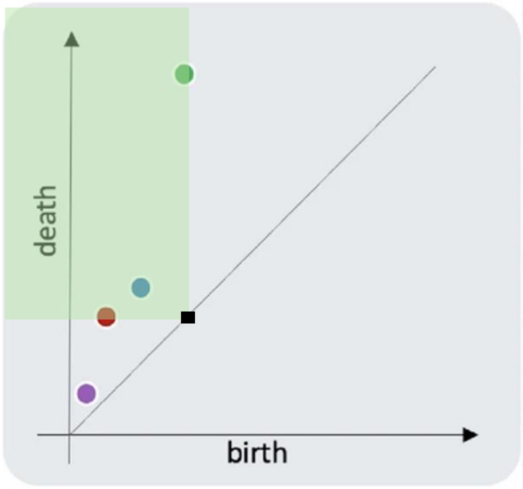
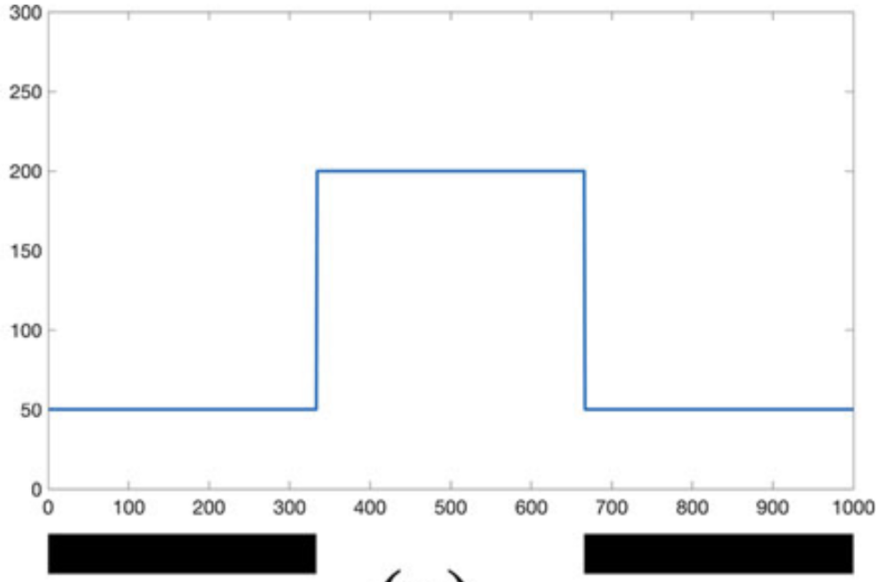


Image source: Bobrowski O, Skraba P. A universal null-distribution for topological data analysis.

Motivational example

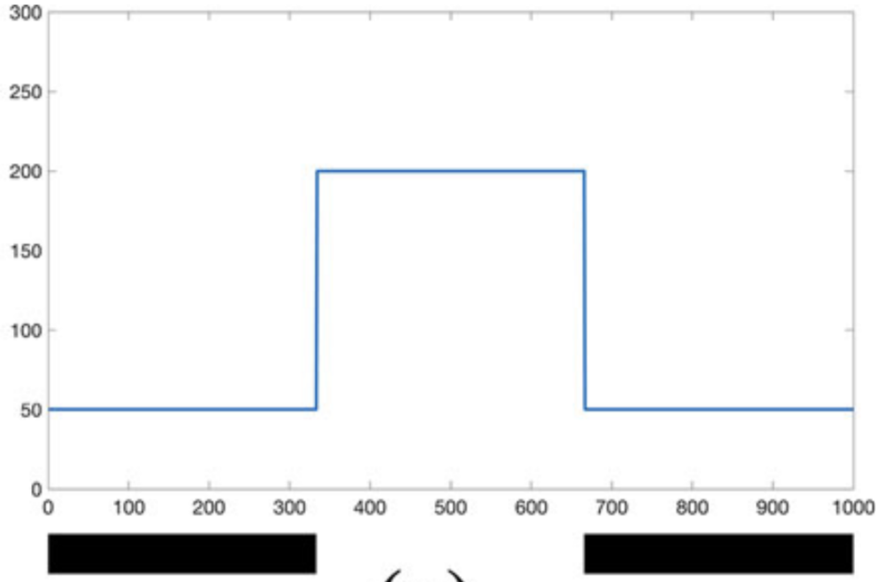
- Let's assume a very simple scenario for the motivation of the idea, aka. a "1D" image (function) $\hat{f}$ with only two values (so already binarized)



(a)

Motivational example

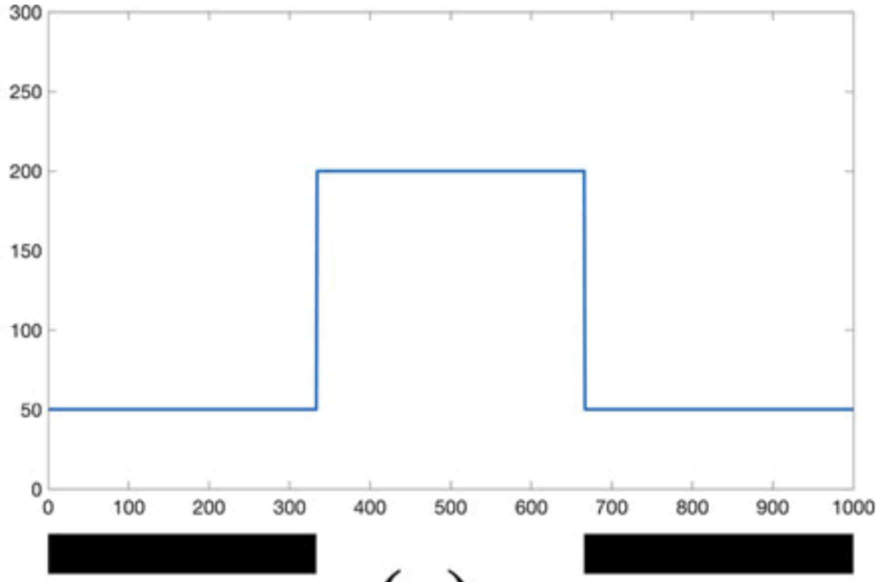
- Let's assume a very simple scenario for the motivation of the idea, aka. a "1D" image (function) $\hat{f}$ with only two values (so already binarized)
- We take $\hat{f}$ to be the "ground truth"



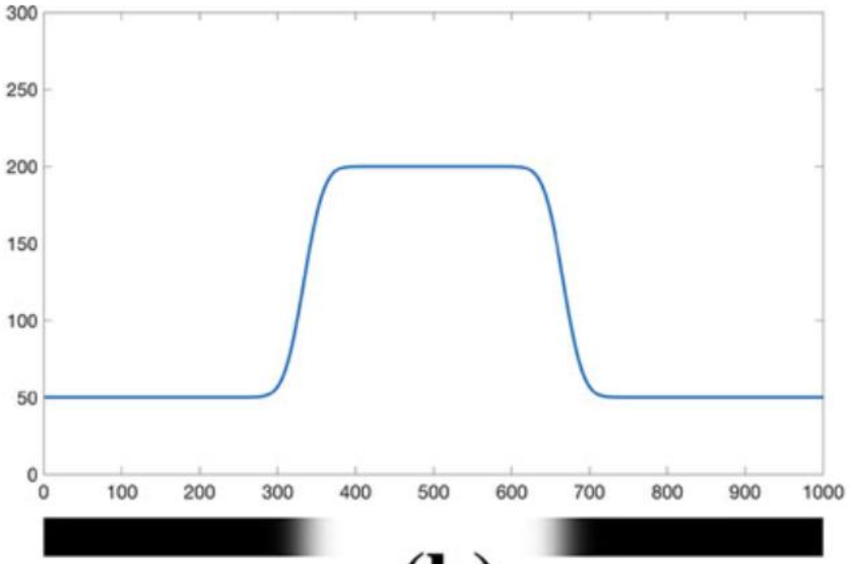
(a)

Motivational example

- We also allow for the case where $\hat{f}$ is a function with monotone transitions between the extreme values



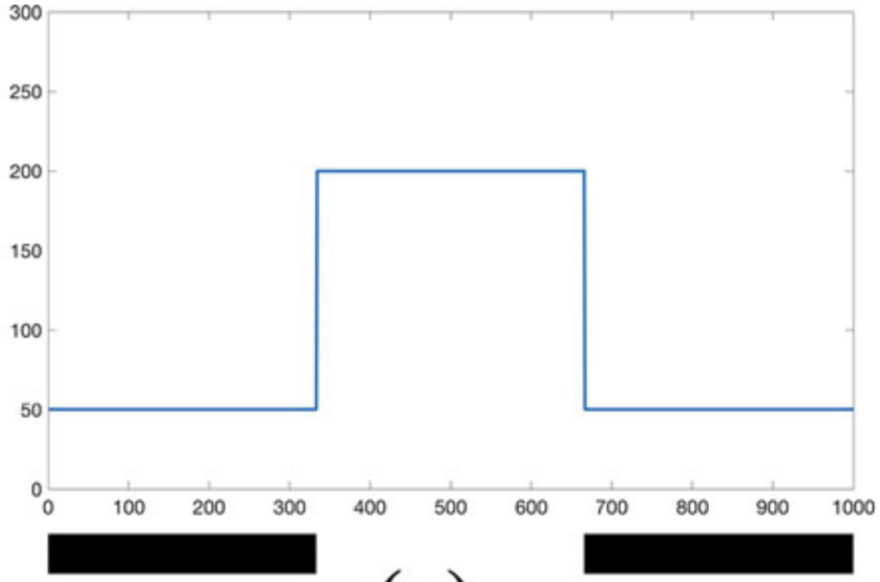
(a)



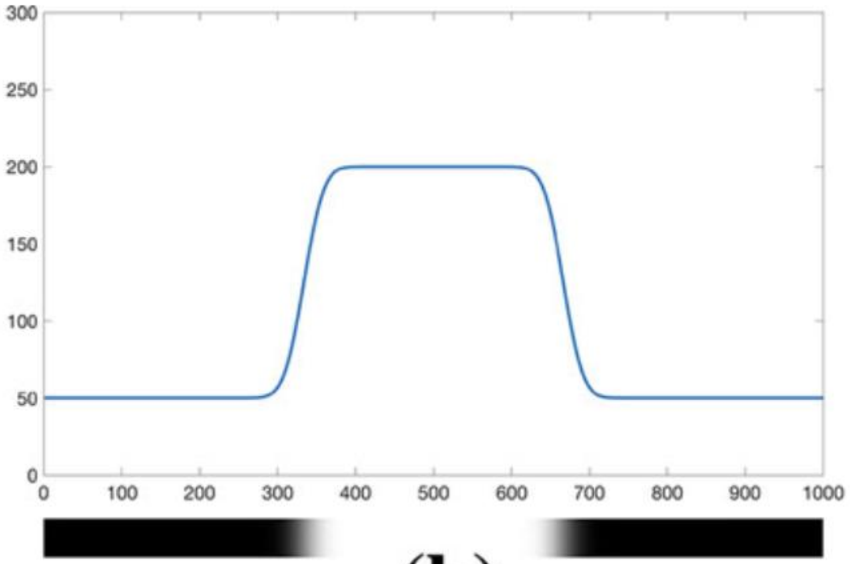
(b)

Motivational example

- We also allow for the case where $\hat{f}$ is a function with monotone transitions between the extreme values
- In either case, a threshold value between 50 (min value) and 200 (max value) would give you a correct binarization (in terms of Betti number)



(a)



(b)

Motivational example

- The barcode of the previous two functions would have two intervals, $[50, \infty)$ and $[50{,}200)$

Motivational example

- The barcode of the previous two functions would have two intervals, $[50, \infty)$ and $[50,200)$
- Inspecting the barcode/PD, any value between 50 and 200 would give you the correct $\beta_0 = 2$

Motivational example

- The barcode of the previous two functions would have two intervals, $[50, \infty)$ and $[50,200)$
- Inspecting the barcode/PD, any value between 50 and 200 would give you the correct $\beta_0 = 2$
  - (Notice here since the domain of the function is only 1-dimensional, we could only count β_0 and all other Betti numbers are 0)

Motivational example

- The barcode of the previous two functions would have two intervals, $[50, \infty)$ and $[50,200)$
- Inspecting the barcode/PD, any value between 50 and 200 would give you the correct $\beta_0 = 2$
  - (Notice here since the domain of the function is only 1-dimensional, we could only count β_0 and all other Betti numbers are 0)
- So for this problem, the homology inference is a done deal once you have the barcode

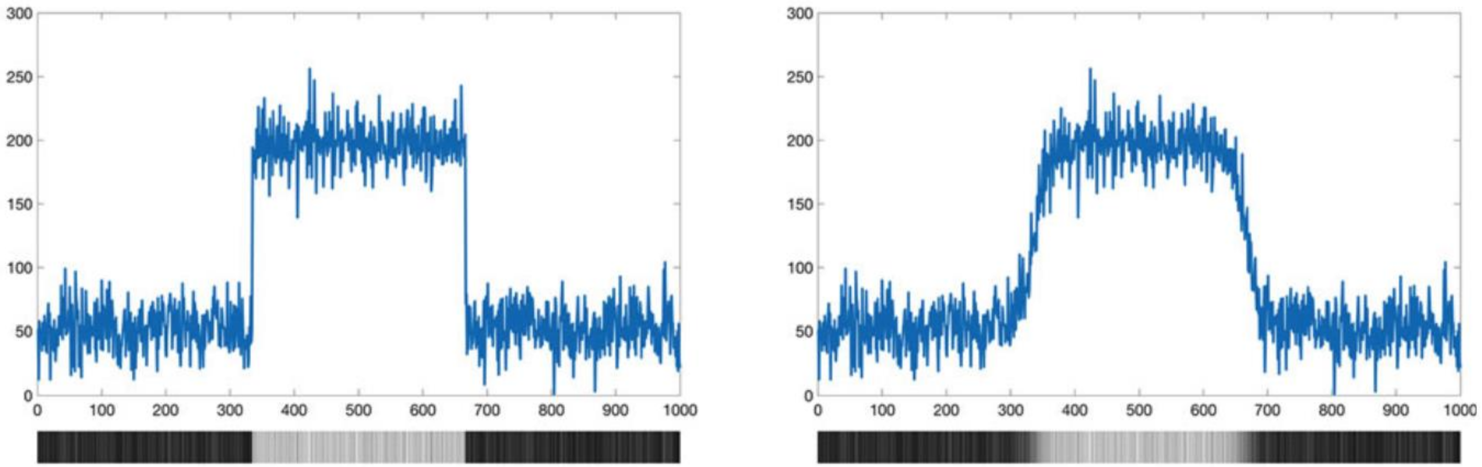
Motivational example

- In reality, we don’t have such clean data and we don’t usually know the ground truth
- Data often comes with noise

Motivational example

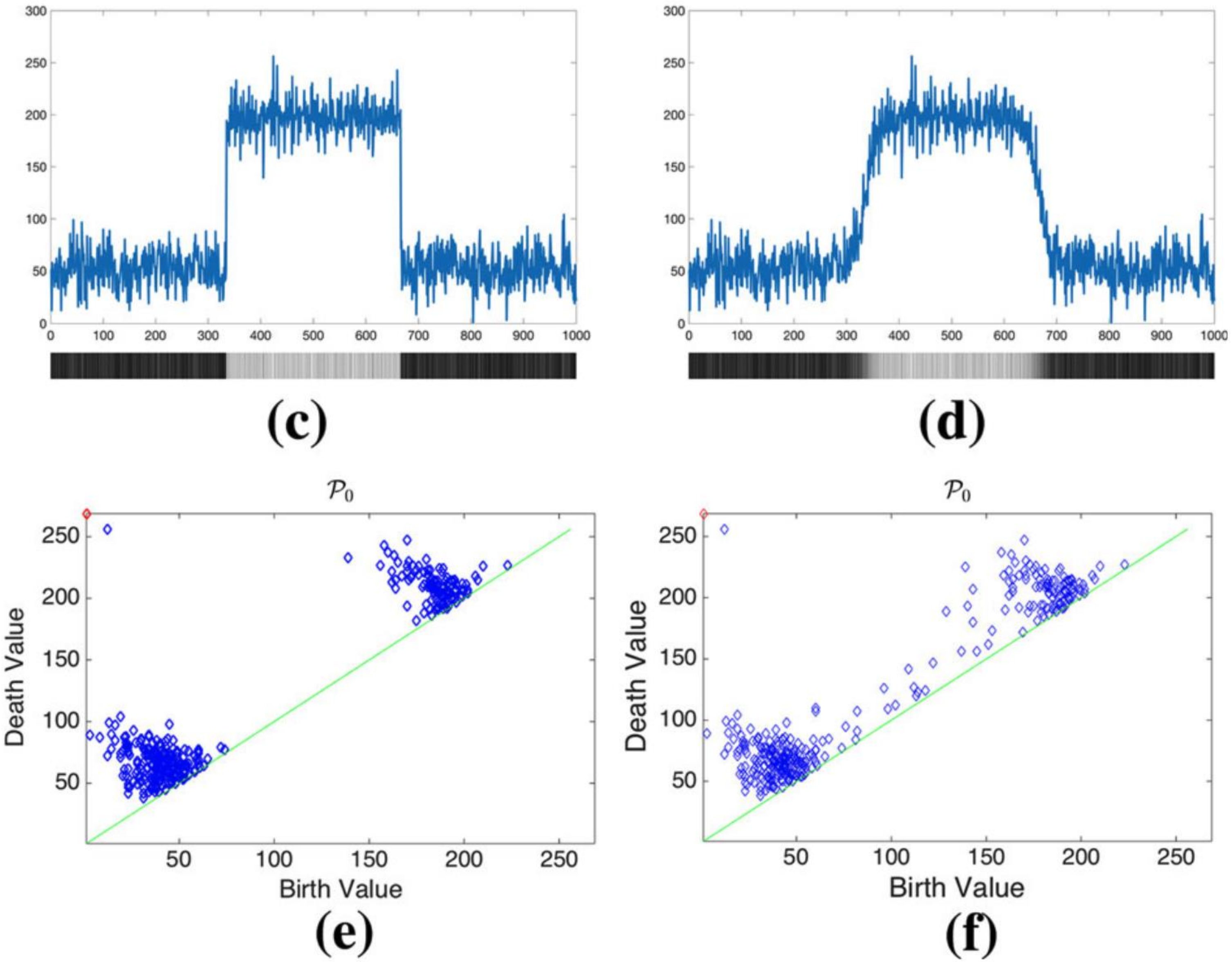
- In reality, we don't have such clean data and we don't usually know the ground truth
- Data often comes with noise
- Now, to make things closer to reality, we add some noise to the function $\hat{f}$, and get a more realistic function f that one could encounter

Motivational example



(c) (d)

Motivational example



Motivational example

- By looking at the intersection of the quadrant as we move the lower-right corner along the diagonal, we observe:
  - The range of values where the Betti number equals 2 is greatly reduced for the left function, now become [105, 140)

Motivational example

- By looking at the intersection of the quadrant as we move the lower-right corner along the diagonal, we observe:
  - The range of values where the Betti number equals 2 is greatly reduced for the left function, now become [105, 140)
  - What's more, for the right function, there is now not a single value whose Betti number evaluates to 2 (we can never find the ground truth by simply using a threshold)

Motivational example

- By looking at the intersection of the quadrant as we move the lower-right corner along the diagonal, we observe:
  - The range of values where the Betti number equals 2 is greatly reduced for the left function, now become [105, 140)
  - What's more, for the right function, there is now not a single value whose Betti number evaluates to 2 (we can never find the ground truth by simply using a threshold)
- BTW, to solve the problem of the right function, one could use a cut-off for the length of the intervals (distance to the diagonal) and keep only those points in PD above a $45°$ degree line (we could then count to 2 with remaining points)

Design idea

- The idea for the algorithm proposed is better illustrated by drawing the barcode for the left function which contains two long intervals $[50, \infty)$, $[50,200)$, and some small noisy intervals on the two sides (left of 105 and right of 140)

Design idea

- The idea for the algorithm proposed is better illustrated by drawing the barcode for the left function which contains two long intervals $[50, \infty)$, $[50,200)$, and some small noisy intervals on the two sides (left of 105 and right of 140)
- The "best values" would probably be the value in the middle of 105 and 140, around 72, for two reasons:

Design idea

- The idea for the algorithm proposed is better illustrated by drawing the barcode for the left function which contains two long intervals $[50, \infty)$, $[50{,}200)$, and some small noisy intervals on the two sides (left of 105 and right of 140)
- The "best values" would probably be the value in the middle of 105 and 140, around 72, for two reasons:
  1. In the middle of long features: This threshold tends to be in the middle of the life of long lifespan features so that these features tend not to have just appeared (been born) or be on the verge of disappearing (dying)

Design idea

- The idea for the algorithm proposed is better illustrated by drawing the barcode for the left function which contains two long intervals $[50, \infty)$, $[50,200)$, and some small noisy intervals on the two sides (left of 105 and right of 140)
- The "best values" would probably be the value in the middle of 105 and 140, around 72, for two reasons:
  1. In the middle of long features: This threshold tends to be in the middle of the life of long lifespan features so that these features tend not to have just appeared (been born) or be on the verge of disappearing (dying)
  2. Far from short features: It is also far enough from the short-lived features (most probably noise) from the both sides so that small perturbations of the threshold do not produce large changes in the noisy features included

Objective function

- To achieve the first goal (threshold sitting in middle of long features), they try to find a threshold t maximizing the following objective function:

$$\Psi(t) = \frac{1}{\#\mathscr{P}(t) + 1} \sum_{(b,d) \in \mathscr{P}(t)} (d - t)(t - b)$$

Objective function

- To achieve the first goal (threshold sitting in middle of long features), they try to find a threshold t maximizing the following objective function:

$$\Psi(t) = \frac{1}{\#\mathscr{P}(t) + 1} \sum_{(b,d) \in \mathscr{P}(t)} (d - t)(t - b)$$

- $\mathscr{P}(t)$: set of intervals intersecting (present at) threshold t

Objective function

- To achieve the first goal (threshold sitting in middle of long features), they try to find a threshold t maximizing the following objective function:

$$\Psi(t) = \frac{1}{\#\mathscr{P}(t) + 1} \sum_{(b,d)\in\mathscr{P}(t)} (d - t)(t - b)$$

- $\mathscr{P}(t)$: set of intervals intersecting (present at) threshold t
- $(d - t)(t - b)$ measures how "in the middle" t is sitting within interval $[b, d)$:

Objective function

- To achieve the first goal (threshold sitting in middle of long features), they try to find a threshold t maximizing the following objective function:

$$\Psi(t) = \frac{1}{\#\mathscr{P}(t) + 1} \sum_{(b,d)\in\mathscr{P}(t)} (d-t)(t-b)$$

- $\mathscr{P}(t)$: set of intervals intersecting (present at) threshold t
- $(d-t)(t-b)$ measures how “in the middle” t is sitting within interval $[b,d)$:
  - If $[b,d)$ is short, then $(d-t)(t-b)$ is small anyways

Objective function

- To achieve the first goal (threshold sitting in middle of long features), they try to find a threshold t maximizing the following objective function:

$$\Psi(t) = \frac{1}{\#\mathscr{P}(t) + 1} \sum_{(b,d) \in \mathscr{P}(t)} (d - t)(t - b)$$

- $\mathscr{P}(t)$: set of intervals intersecting (present at) threshold t
- $(d - t)(t - b)$ measures how "in the middle" t is sitting within interval $[b, d)$:
  - If $[b, d)$ is short, then $(d - t)(t - b)$ is small anyways
  - If $[b, d)$ is long, then $(d - t)(t - b)$ achieves its highest value when t sits in the middle

Objective function

- To achieve the first goal (threshold sitting in middle of long features), they try to find a threshold t maximizing the following objective function:

$$\Psi(t) = \frac{1}{\#\mathscr{P}(t) + 1} \sum_{(b,d)\in\mathscr{P}(t)} (d - t)(t - b)$$

- $\mathscr{P}(t)$: set of intervals intersecting (present at) threshold t
- $(d - t)(t - b)$ measures how "in the middle" t is sitting within interval $[b, d)$:
  - If $[b, d)$ is short, then $(d - t)(t - b)$ is small anyways
  - If $[b, d)$ is long, then $(d - t)(t - b)$ achieves its highest value when t sits in the middle
- We divide the whole function by $\#\mathscr{P}(t) + 1$ because we want to measure the average contribution of the term for the intervals (so for we do not penalize a t intersecting less long intervals)

Objective function

- To achieve the first goal (threshold sitting in middle of long features), they try to find a threshold t maximizing the following objective function:

$$\Psi(t) = \frac{1}{\#\mathscr{P}(t) + 1} \sum_{(b,d)\in\mathscr{P}(t)} (d-t)(t-b)$$

- $\mathscr{P}(t)$: set of intervals intersecting (present at) threshold t
- $(d-t)(t-b)$ measures how "in the middle" t is sitting within interval $[b,d)$:
  - If $[b,d)$ is short, then $(d-t)(t-b)$ is small anyways
  - If $[b,d)$ is long, then $(d-t)(t-b)$ achieves its highest value when t sits in the middle
- The above objective function penalizes a t intersecting a lot of short intervals, and favors a t intersecting long intervals in the middle

Objective function

- To account for topological features that are missed at t, penalizing exclusion of long lifespan features, they also maximize the following:

Objective function

- To account for topological features that are missed at t, penalizing exclusion of long lifespan features, they also maximize the following:

$$\Psi^-(t) = \sum_{(b,d)\in\mathscr{P}^-(t)} \frac{t-d}{d-b}, \qquad \Psi^+(t) = \sum_{(b,d)\in\mathscr{P}^+(t)} \frac{b-t}{d-b},$$

- $\mathscr{P}^-(t) = \{(b,d) \in \mathscr{P} \mid \ d \leq t\}$: set of intervals to left of t
- $\mathscr{P}^+(t) = \{(b,d) \in \mathscr{P} \mid \ b > t\}$: set of intervals to right of t

Objective function

- To account for topological features that are missed at t, penalizing exclusion of long lifespan features, they also maximize the following:

$$\Psi^-(t) = \sum_{(b,d)\in\mathscr{P}^-(t)} \frac{t-d}{d-b}, \qquad \Psi^+(t) = \sum_{(b,d)\in\mathscr{P}^+(t)} \frac{b-t}{d-b},$$

- $\mathscr{P}^-(t) = \{(b,d) \in \mathscr{P} \mid \ d \leq t\}$: set of intervals to left of t
- $\mathscr{P}^+(t) = \{(b,d) \in \mathscr{P} \mid \ b > t\}$: set of intervals to right of t
- For each term in $\Psi^-(t)$, if $[b,d)$ is long, then the denominator is large making the whole term small; also, if t is far from $[b,d)$, $t-d$ (distance of t to the interval) is large

Objective function

- To account for topological features that are missed at t, penalizing exclusion of long lifespan features, they also maximize the following:

$$\Psi^{-}(t) = \sum_{(b,d)\in\mathscr{P}^{-}(t)} \frac{t-d}{d-b}, \qquad \Psi^{+}(t) = \sum_{(b,d)\in\mathscr{P}^{+}(t)} \frac{b-t}{d-b},$$

- $\mathscr{P}^{-}(t) = \{(b,d) \in \mathscr{P} \mid \ d \leq t\}$: set of intervals to left of t
- $\mathscr{P}^{+}(t) = \{(b,d) \in \mathscr{P} \mid \ b > t\}$: set of intervals to right of t
- For each term in $\Psi^{-}(t)$, if $[b,d)$ is long, then the denominator is large making the whole term small; also, if t is far from $[b,d)$, $t-d$ (distance of t to the interval) is large
- This means that it would penalize the situation where t is very close to some long intervals

Objective function

- To maximize all three objective functions, they maximize the following which is the multiplication of the three:

$$\Phi(t) = \Psi(t) \cdot \Psi^{-}(t) \cdot \Psi^{+}(t),$$

Objective function

- To maximize all three objective functions, they maximize the following which is the multiplication of the three:

$$\Phi(t) = \Psi(t) \cdot \Psi^{-}(t) \cdot \Psi^{+}(t),$$

- The threshold T they look for is then the following:

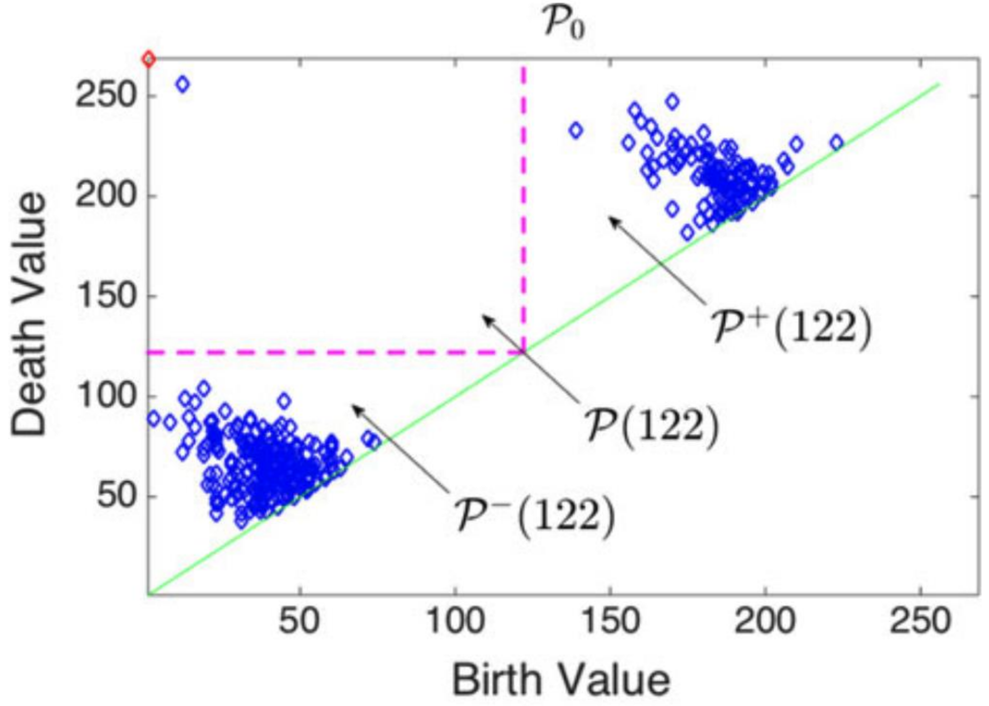
$$T = \underset{t \in [0, I]}{\arg\max} \; \Phi(t).$$

A side note

- One interesting thing to notice: for a threshold t, the quadrant with corner on (t, t) separates a PD into the three parts we were considering previously:
- $\mathscr{P}(t)$: the quadrant
- $\mathscr{P}^{-}(t)$: the lower-left part
- $\mathscr{P}^{+}(t)$: the upper-right part

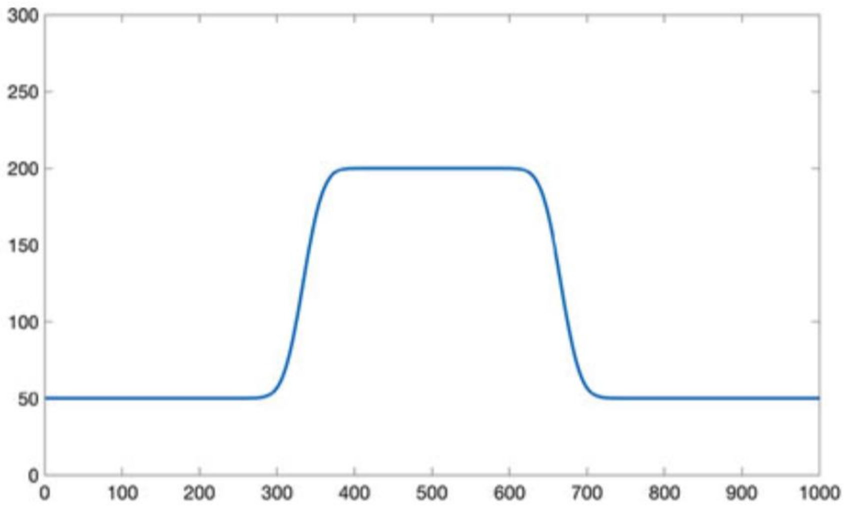
A side note

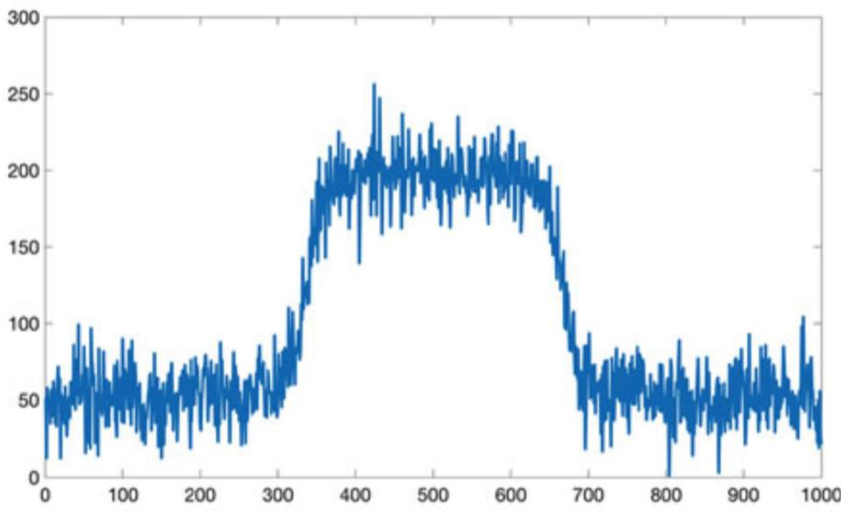
- One interesting thing to notice: for a threshold t, the quadrant with corner on (t, t) separates a PD into the three parts we were considering previously:
- $\mathscr{P}(t)$: the quadrant
- $\mathscr{P}^-(t)$: the lower-left part
- $\mathscr{P}^+(t)$: the upper-right part



Result

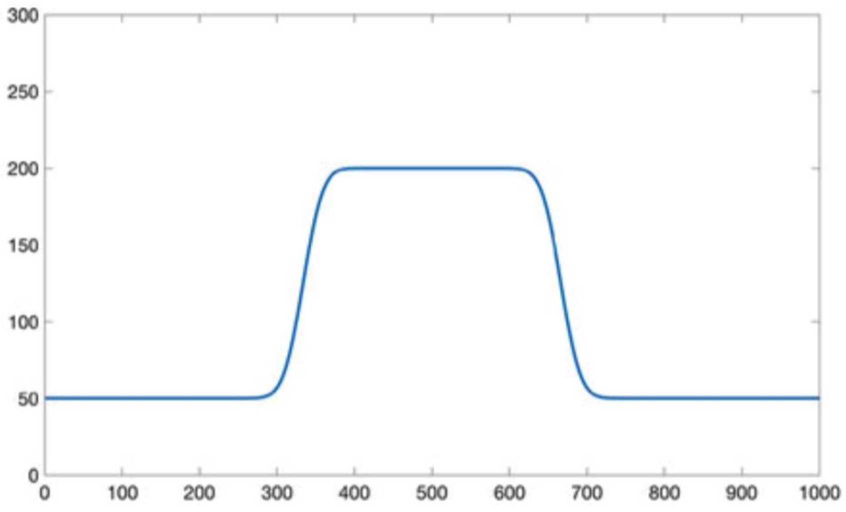
- They performed 1000 trials applying thresholding to previous 1D function each time with an additive noise generated from a Gaussian distribution with mean 0 and standard deviation 2

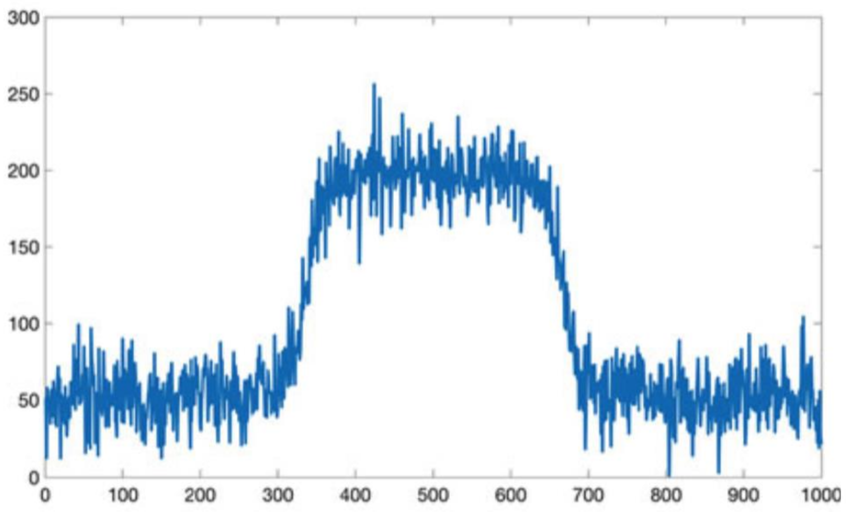




Result

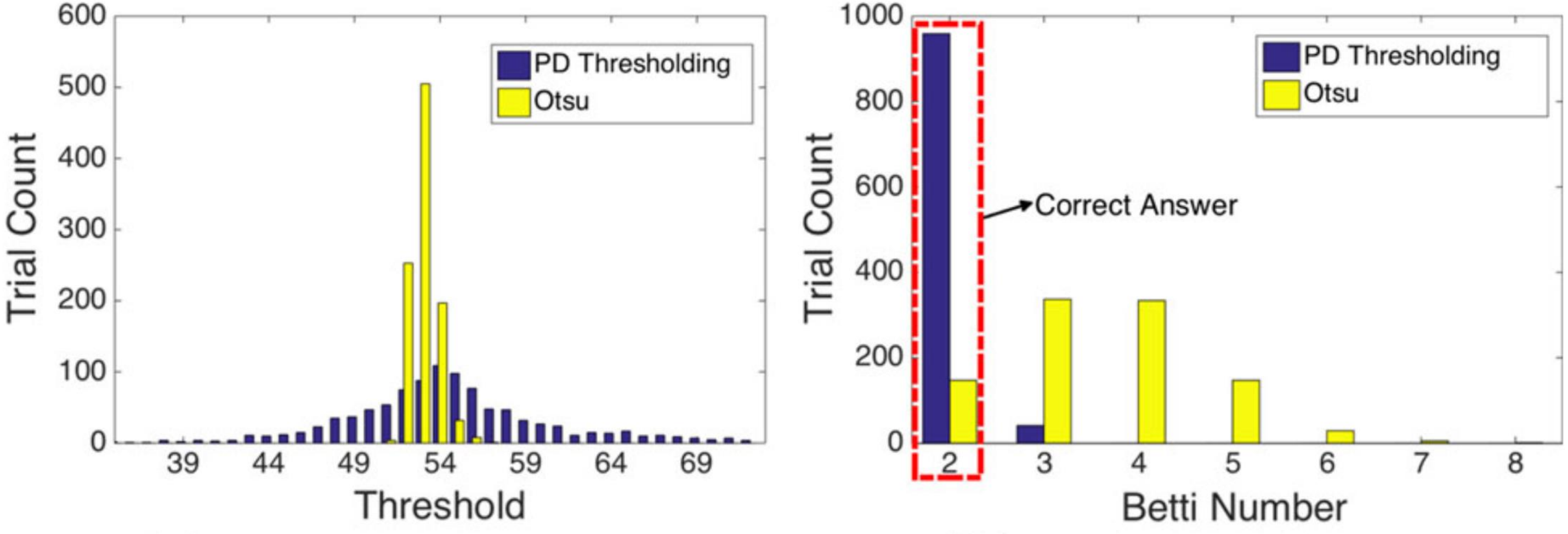
- They performed 1000 trials applying thresholding to previous 1D function each time with an additive noise generated from a Gaussian distribution with mean 0 and standard deviation 2
- They then use Otsu’s method and their method to generate threshold and compute the Betti number





Result

- Their proposed method constantly outperforms Otsu's in terms of the inference of Betti numbers



(a) PD Thresholding Method (b) Betti numbers for each trial

Result

$\hat{\beta} = (\beta_0, \beta_1)$

Original Otsu PD Thresholding

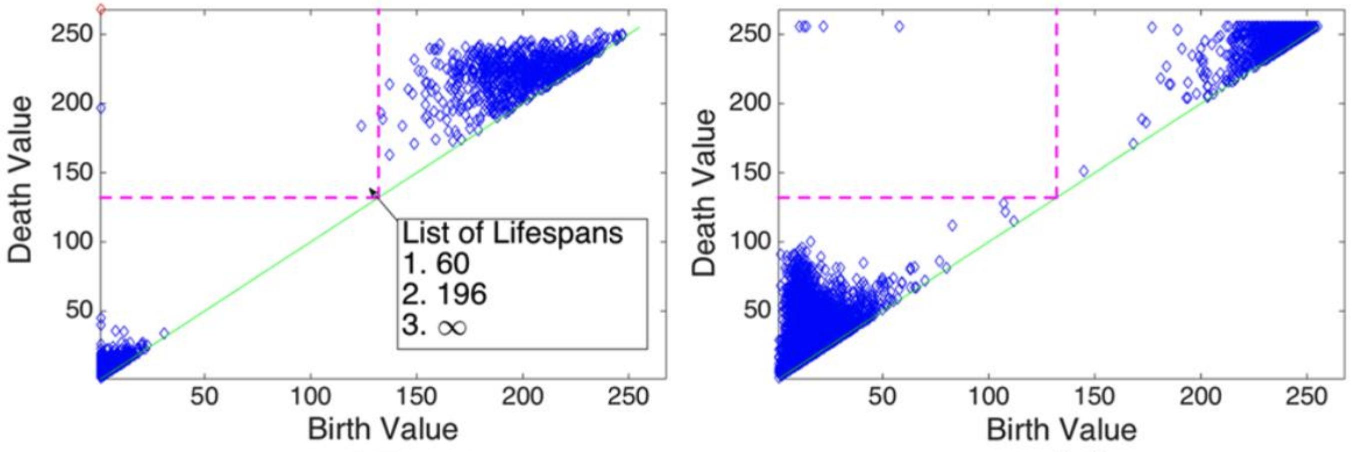
(a) $\hat{\beta} = (2,6)$ **(b)** $\beta(f_{120}^{-}) = (2,8)$ **(c)** $\beta(f_{132}^{-}) = (3,6)$ and $\beta(f_{132}^{-}, L) = (2,6)$

Result

Original | Otsu | PD Thresholding

$\hat{\beta} = (\beta_0, \beta_1)$

(a) $\hat{\beta} = (2,6)$ **(b)** $\beta(f_{120}^{-}) = (2,8)$ **(c)** $\beta(f_{132}^{-}) = (3,6)$ and $\beta(f_{132}^{-}, L) = (2,6)$



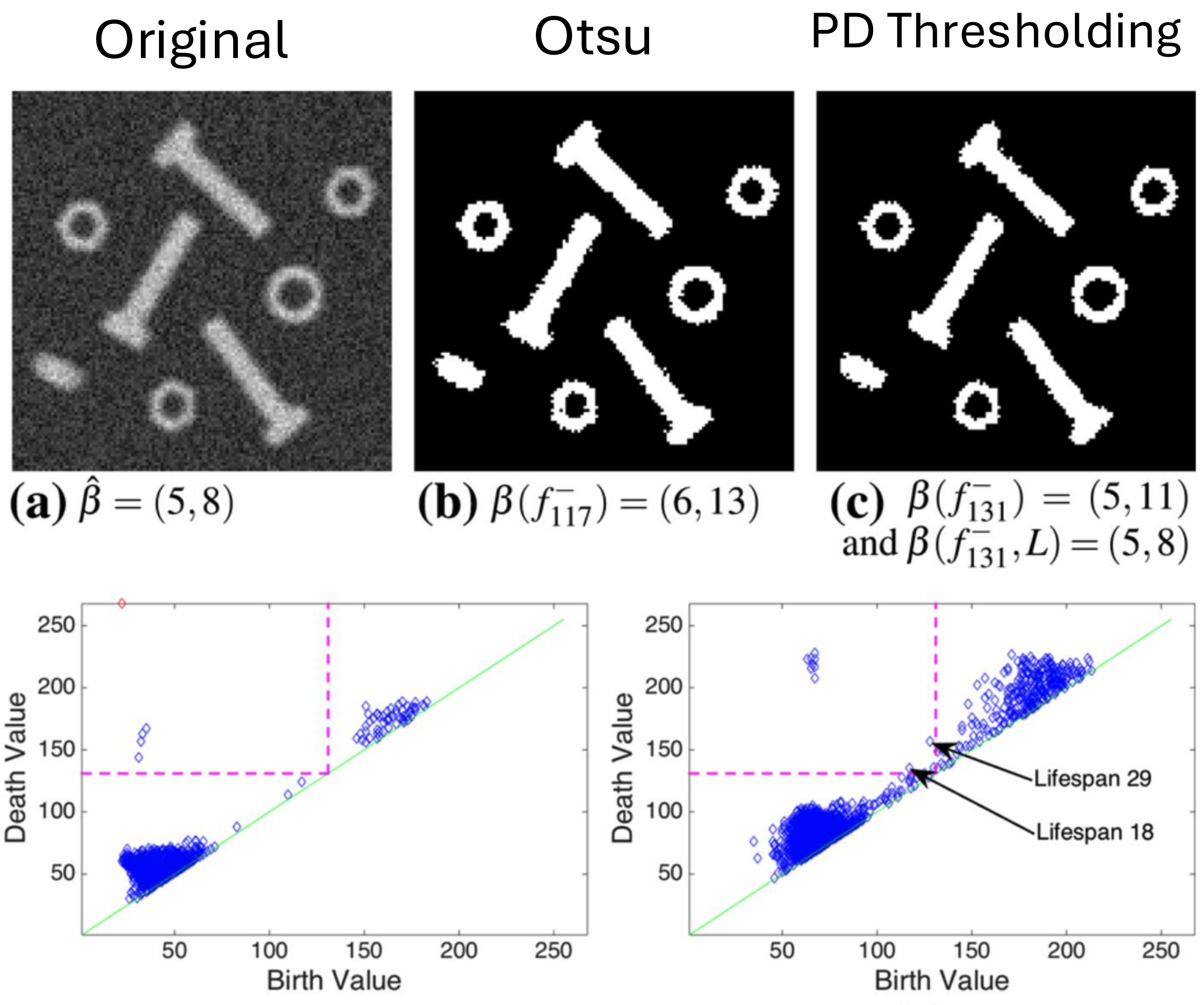
(d) $\mathscr{P}_0$ **(e)** $\mathscr{P}_1$

Result

Original | Otsu | PD Thresholding

(a) $\hat{\beta} = (5, 8)$ **(b)** $\beta(f^{-}_{117}) = (6, 13)$ **(c)** $\beta(f^{-}_{131}) = (5, 11)$ and $\beta(f^{-}_{131}, L) = (5, 8)$

Result



(a) $\hat{\beta} = (5,8)$ (b) $\beta(f^{-}_{117}) = (6,13)$ (c) $\beta(f^{-}_{131}) = (5,11)$ and $\beta(f^{-}_{131}, L) = (5,8)$

(d) $\mathscr{P}_0$ (e) $\mathscr{P}_1$

Result

Original | Otsu | PD Thresholding

(a) $\hat{\beta} = (1, 52)$ **(b)** $\beta(f^{-}_{110}) = (4, 64)$ **(c)** $\beta(f^{-}_{126}) = (4, 53)$ and $\beta(f^{-}_{126}, L) = (1, 51)$

Result

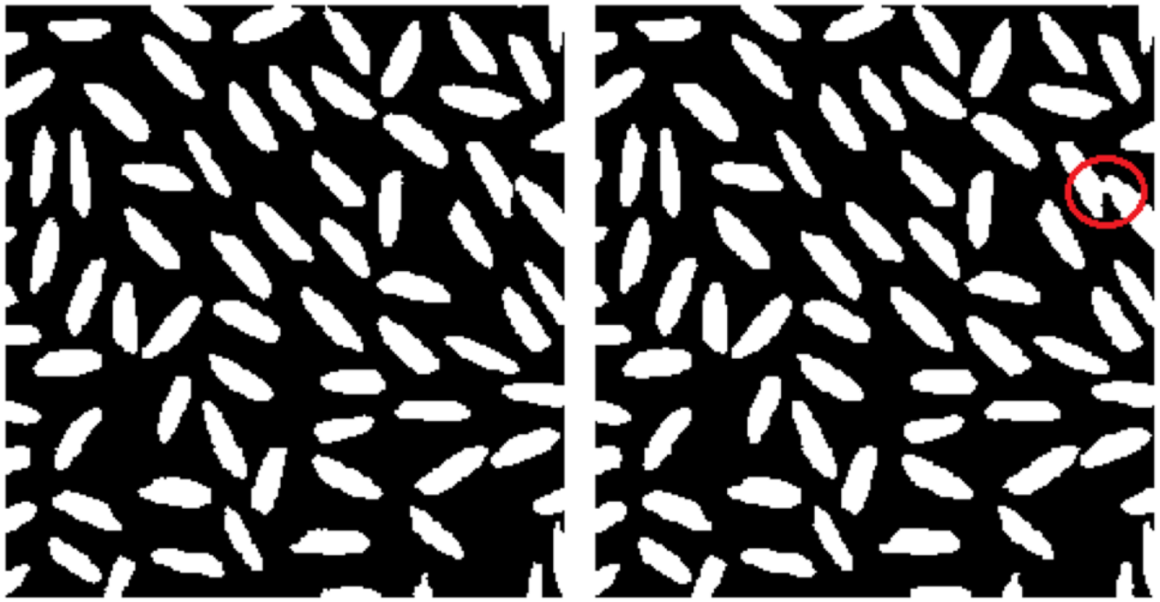
Original | Otsu | PD Thresholding

(a) $\hat{\beta} = (1,52)$ (b) $\beta(f^{-}_{110}) = (4,64)$ (c) $\beta(f^{-}_{126}) = (4,53)$ and $\beta(f^{-}_{126}, L) = (1,51)$

- Although lifespan cutoff gives accurate Betti number, binary image in (c) produced by their method does not represent the original image well, in particular, in the lower region. This appears to be due to the inconsistent illumination in this image (difference in brightness)

Result

PD Thresholding Otsu

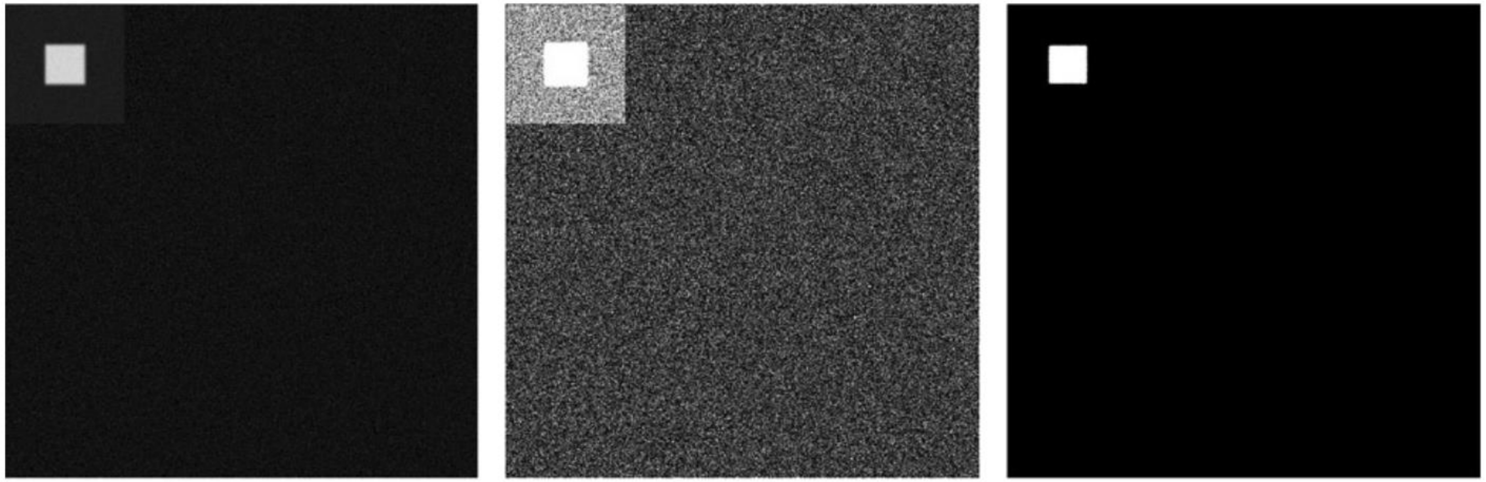


(d) $\beta = (1, 52)$ **(e)** $\beta = (1, 51)$

- By manually partitioning the image into a 4×4 grid of subimages and segmenting each separately, both results are good. But Otsu’s still produces a topological error (red)

Another synthetic dataset

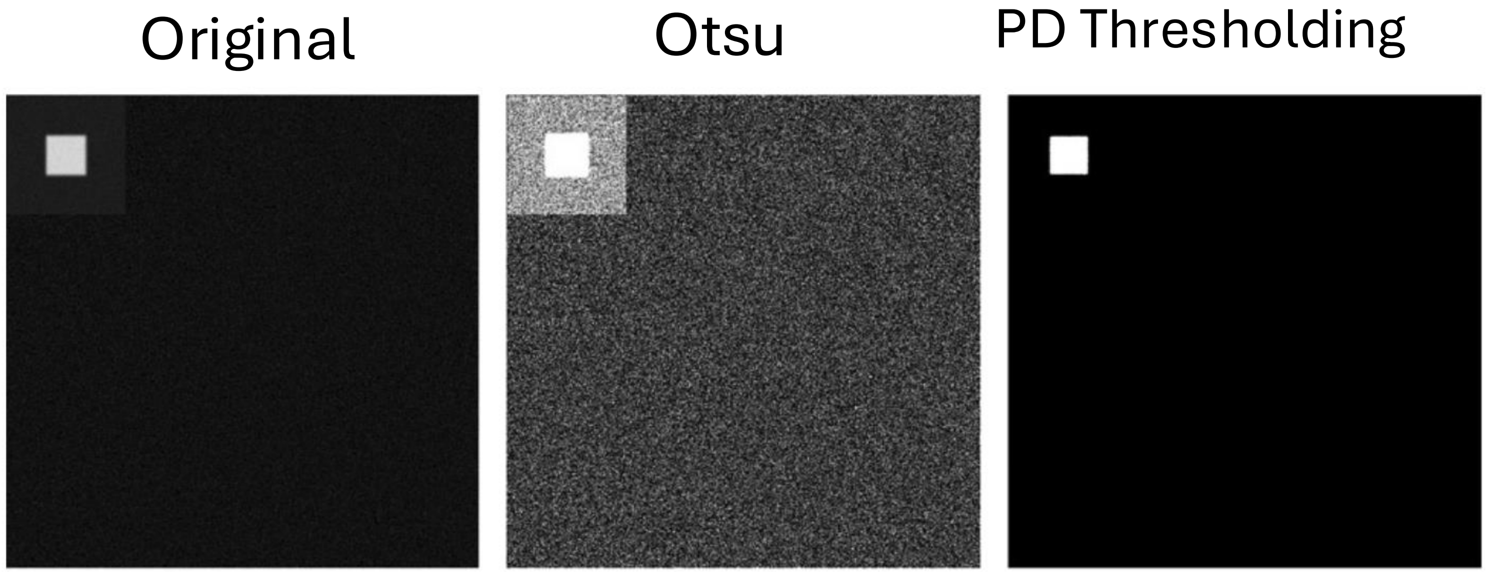
Original | Otsu | PD Thresholding



(a) $\hat{\beta} = (1,1)$ **(b)** $\beta(f_{33}^{-}) = (2072, 127084)$ **(c)** $\beta(f_{142}^{-}) = (1,3)$

1. Similarly to the 1D function, they start with a black / white image with a white square on the upper-left corner.
2. They then add some noise within a (slightly larger) square around the white square
3. Finally, the add noise on the whole image

Another synthetic dataset

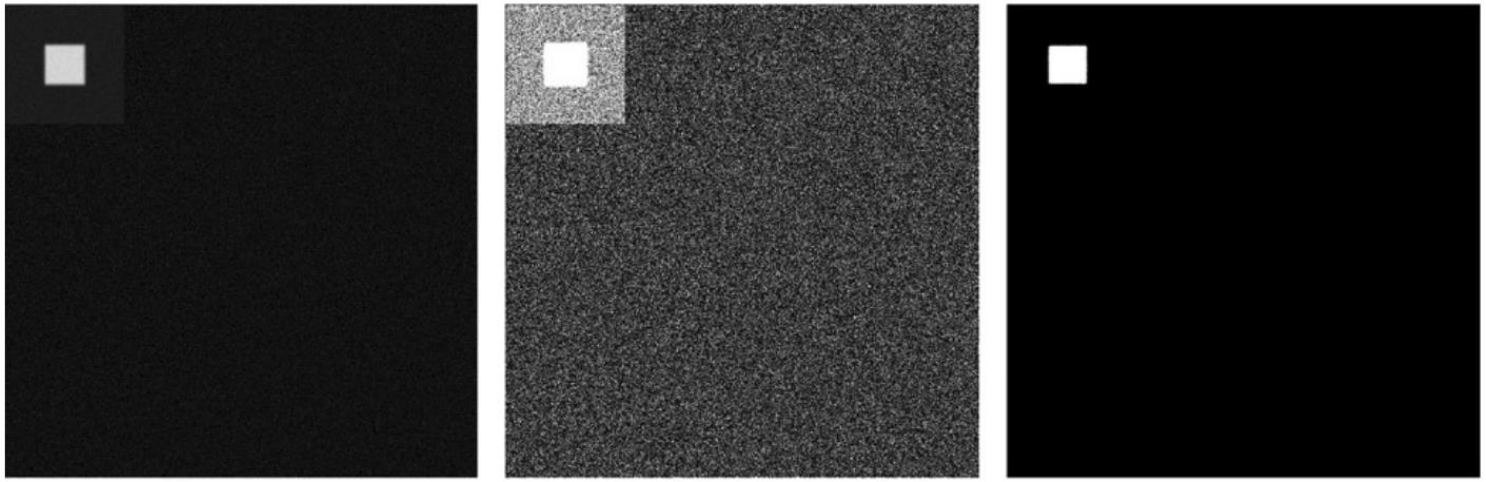


(a) $\hat{\beta} = (1,1)$ **(b)** $\beta(f_{33}^{-}) = (2072, 127084)$ **(c)** $\beta(f_{142}^{-}) = (1,3)$

- Histogram-base segmentation methods (e.g., Otsu) are strongly influenced by the large dark region
- This may result in losing the essential topological features of one connected component and a single hole ($\hat{\beta} = (1,1)$)

Another synthetic dataset

Original Otsu PD Thresholding

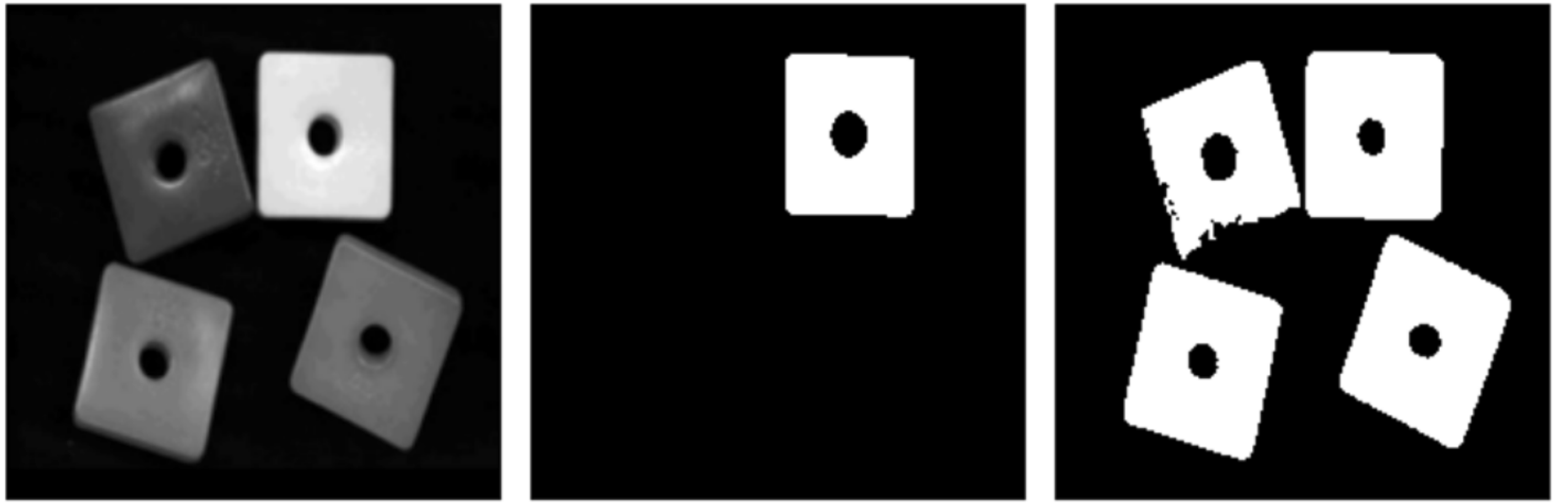


(a) $\hat{\beta} = (1,1)$ **(b)** $\beta(f_{33}^{-}) = (2072, 127084)$ **(c)** $\beta(f_{142}^{-}) = (1,3)$

- And of course, with lifespan cut-off, PD thresholding produces correct result

An example with disparate brightness

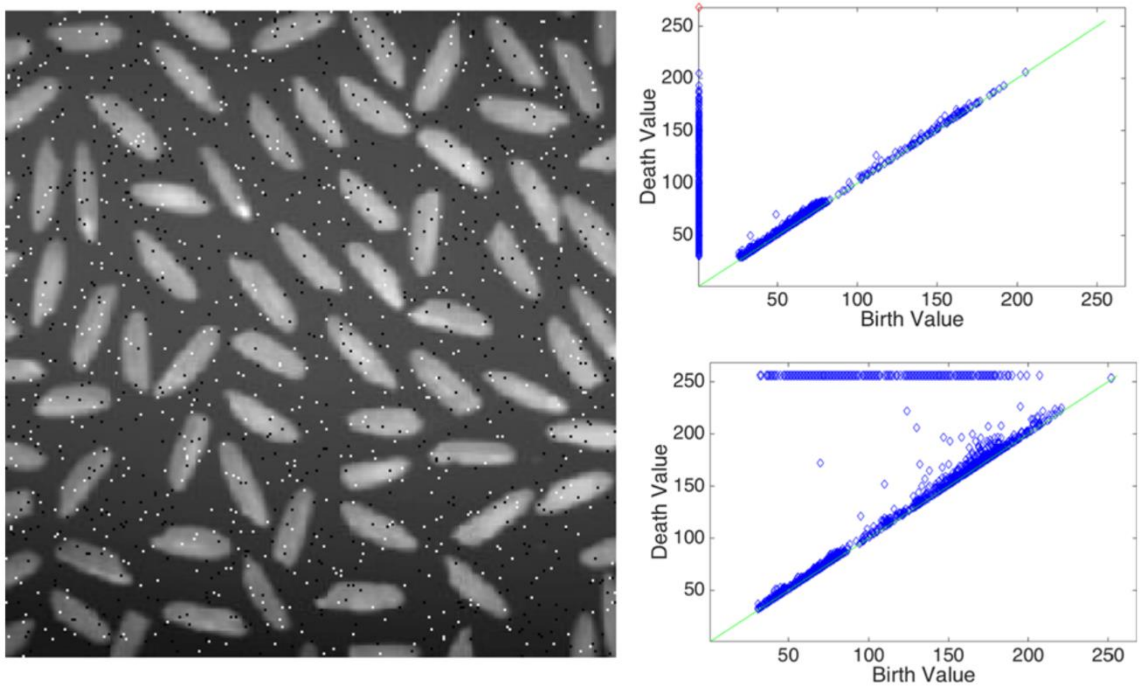
Original Otsu PD Thresholding



(a) $\hat{\beta} = (5,4)$ **(b)** $\beta(f^{-}_{185}) = (2,1)$ **(c)** $\beta(f^{-}_{66}) = (5,4)$

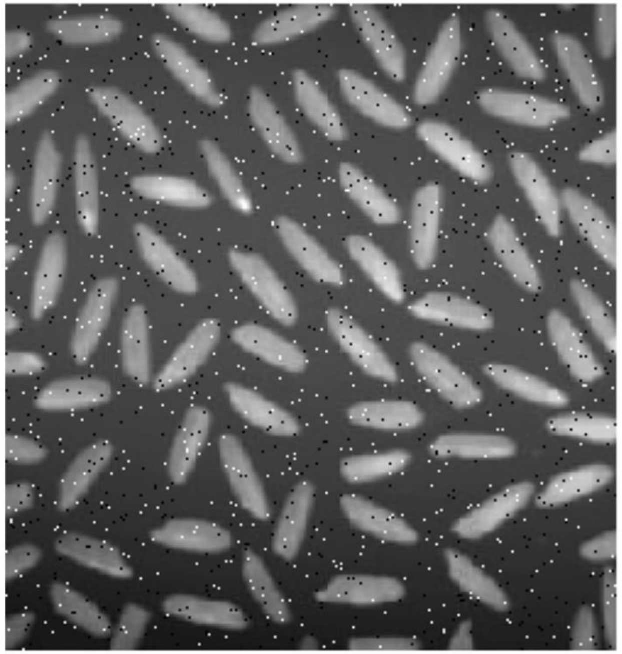
- And of course, with lifespan cut-off, PD thresholding produces correct result

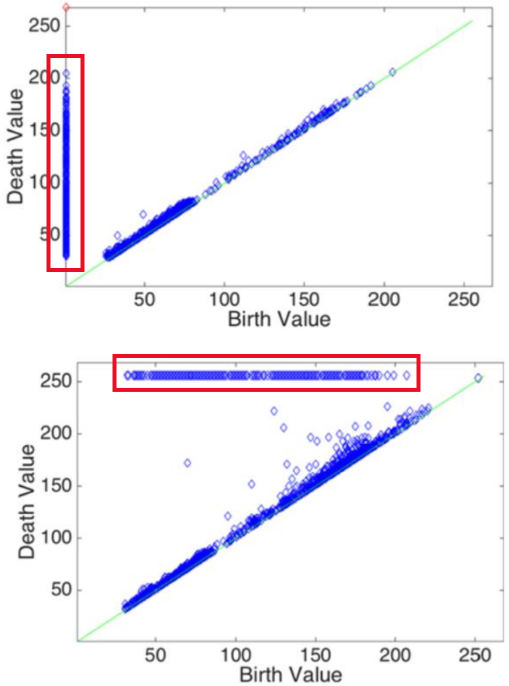
An example that PD thresholding fails



- Add “salt and pepper” noise which randomly resets some pixels to white or black

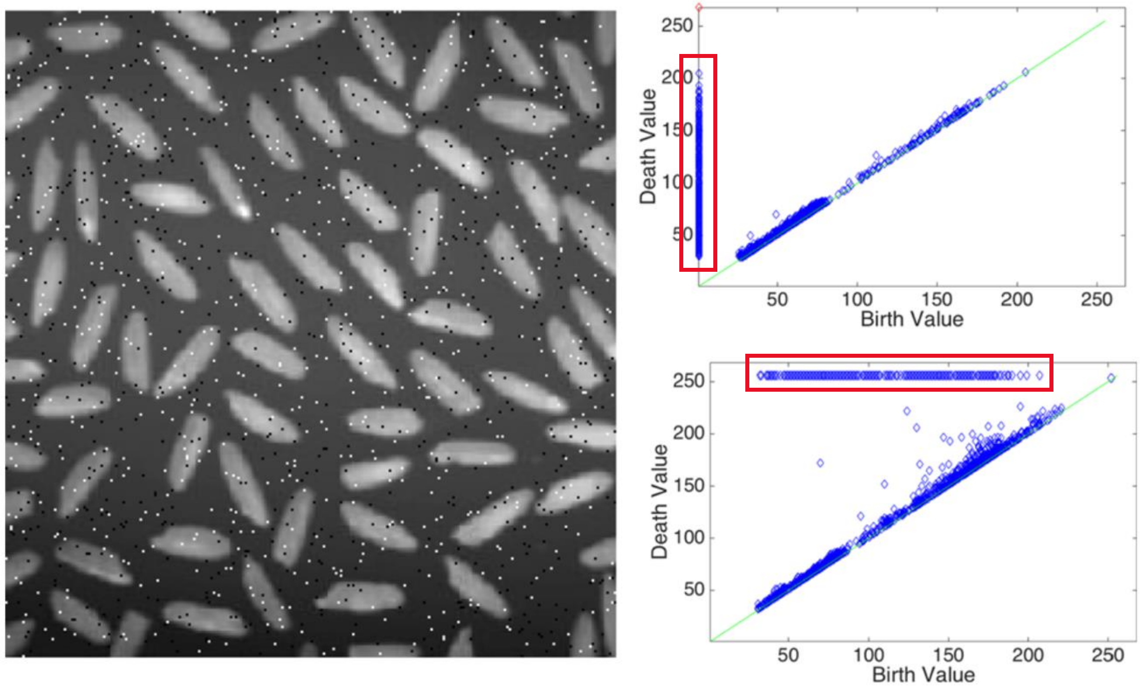
An example that PD thresholding fails





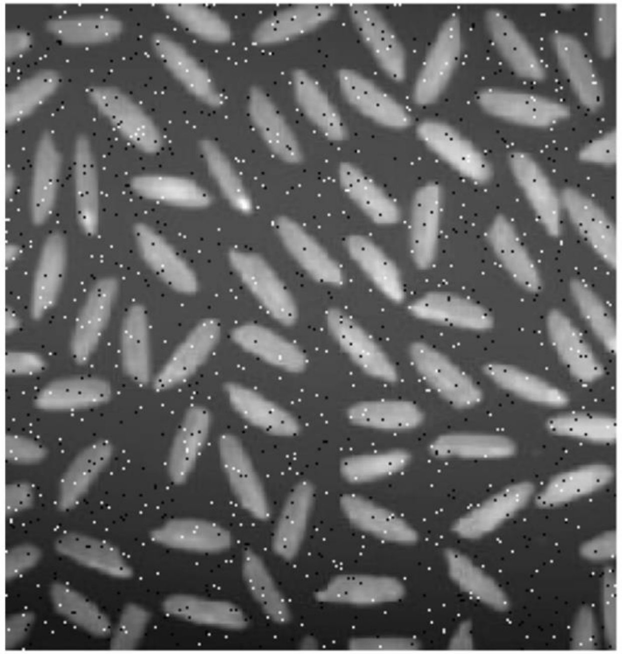
- Add “salt and pepper” noise which randomly resets some pixels to white or black
- The produced PD contain a lot of long-lived intervals due to the small white or black pixels

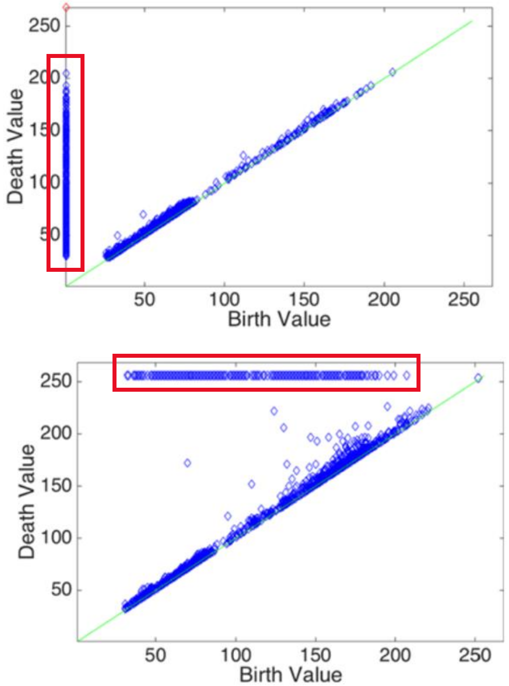
More explanation on why it fails in this case



- Recall the stability of PD: $d_{\mathrm{B}}(PD(f), PD(g)) \leq \| f - g \|_{\infty}$

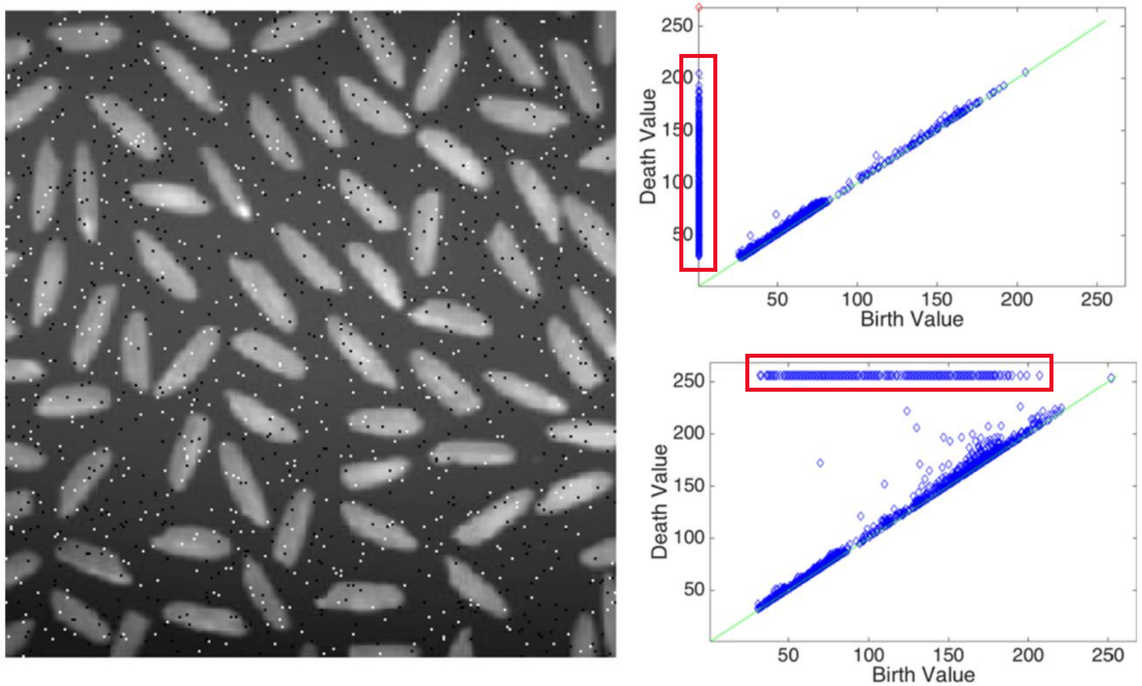
More explanation on why it fails in this case





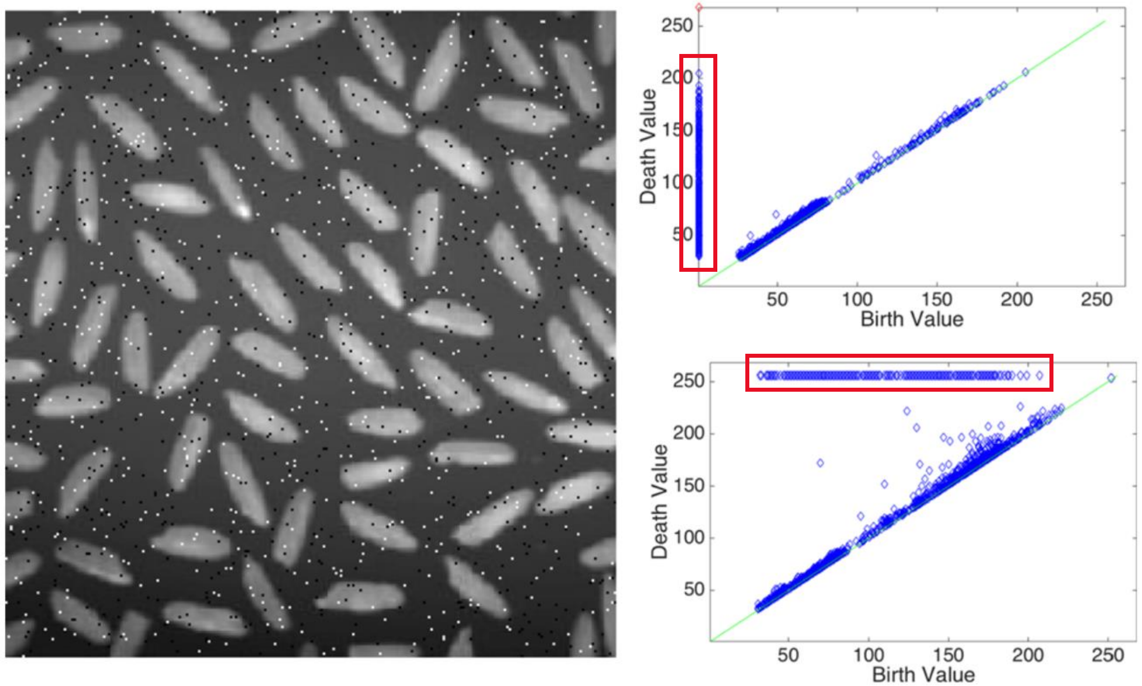
- Recall the stability of PD: $d_{\mathrm{B}}(PD(f), PD(g)) \leq \| f - g \|_\infty$
- The bound $\| f - g \|_\infty$ is determined by the maximum value difference on a single pixel

More explanation on why it fails in this case



- Recall the stability of PD: $d_B(PD(f), PD(g)) \leq \| f - g \|_\infty$
- The bound $\| f - g \|_\infty$ is determined by the maximum value difference on a single pixel
- With the “salt-and-pepper” noise, distance of the two functions (before and after adding noise) can be arbitrarily large

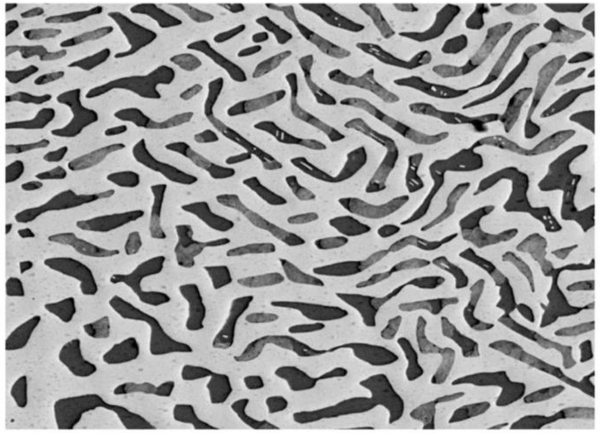
More explanation on why it fails in this case



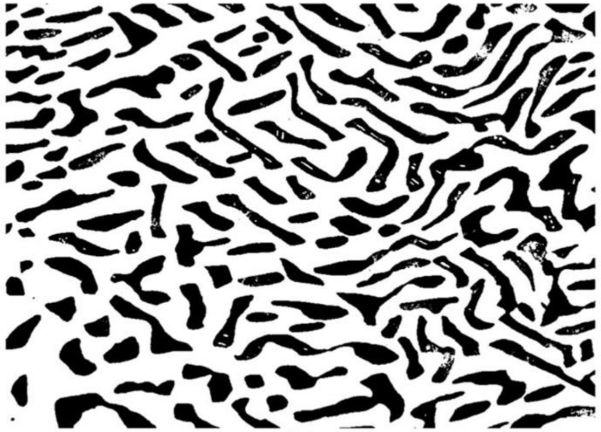
- Recall the stability of PD: $d_{\mathrm{B}}(PD(f), PD(g)) \leq \| f - g \|_{\infty}$
- The bound $\| f - g \|_{\infty}$ is determined by the maximum value difference on a single pixel
- With the "salt-and-pepper" noise, distance of the two functions (before and after adding noise) can be arbitrarily large
- So the two PDs are not stable anymore

Bi–Sn (Bismuth Tin) binary alloy

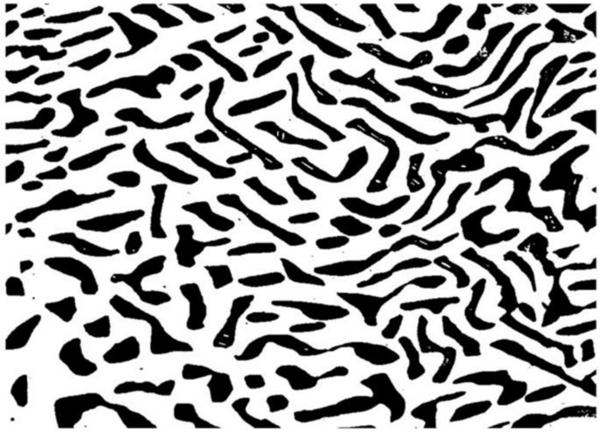
Original



Otsu



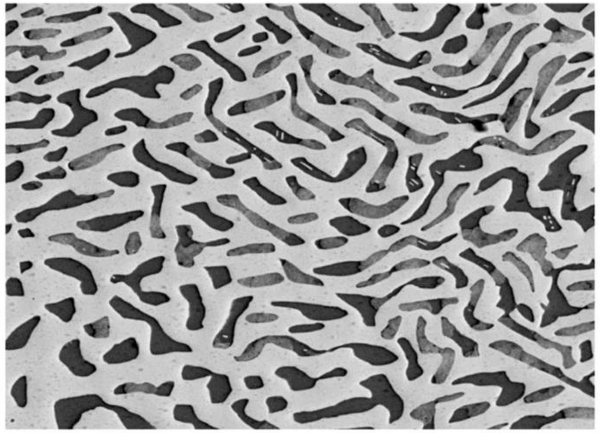
PD Thresholding



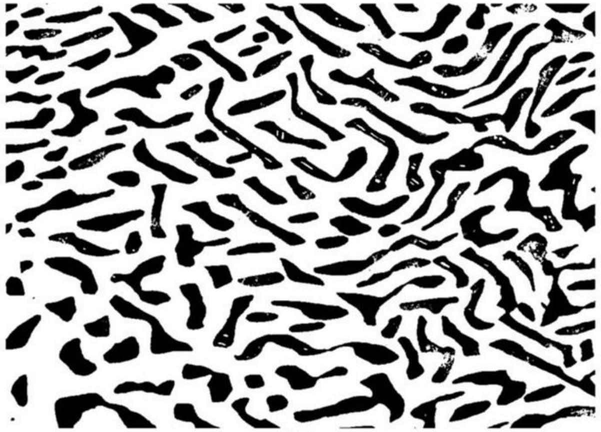
- Observe that the black region contains a lot of connected components which does not form two many tunnels (1-holes), so $\beta_0 \gg \beta_1$

Bi–Sn (Bismuth Tin) binary alloy

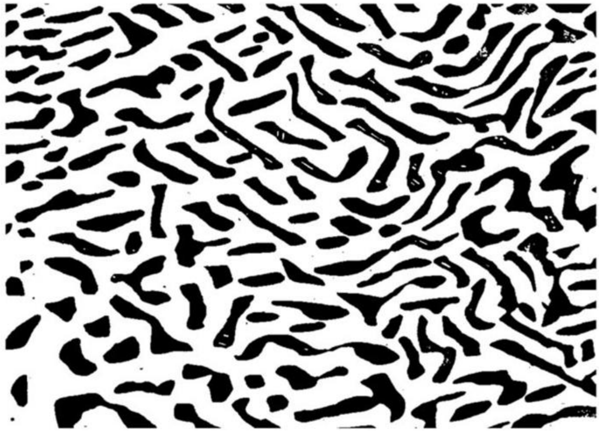
Original



Otsu



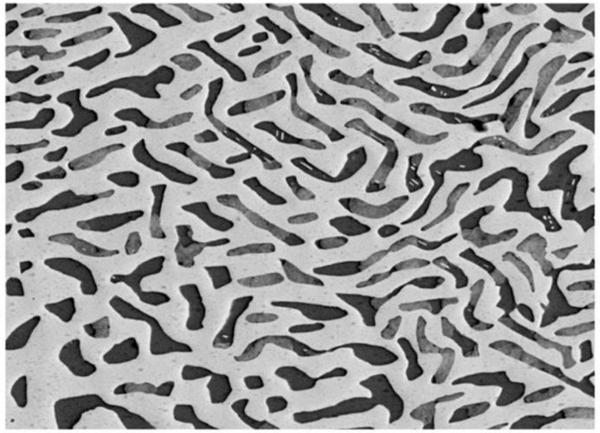
PD Thresholding



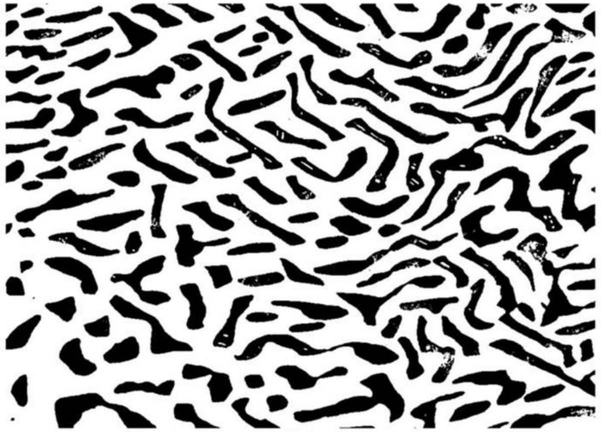
- Observe that the black region contains a lot of connected components which does not form two many tunnels (1-holes), so $\beta_0 \gg \beta_1$
- Otsu : $\beta(f^-) = (206, 432)$ with threshold 134
- PD-based: $\beta(f^-) = (194,143)$ with threshold 151

Bi–Sn (Bismuth Tin) binary alloy

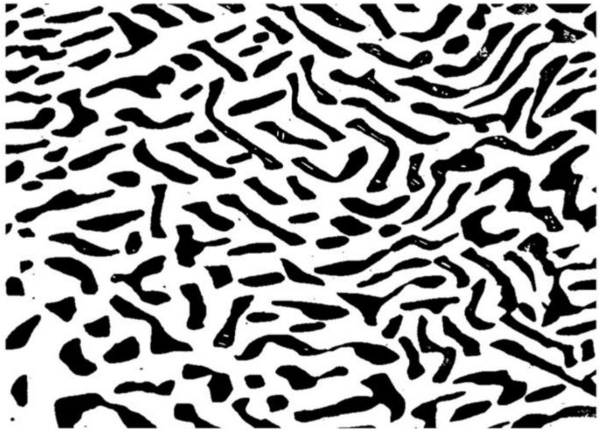
Original



Otsu



PD Thresholding



- Observe that the black region contains a lot of connected components which does not form two many tunnels (1-holes), so $\beta_0 \gg \beta_1$
- Otsu : β(f^-) = (206, 432) with threshold 134
- PD-based: β(f^-) = (194,143) with threshold 151
- With length cut-off, β(f^-) = (149, 33)

Variations

- One variation of the objective function is to replace $\Phi(t)$ with $\Phi_i(t)$ which is the same as $\Phi(t)$ by considering only the i-th PD
  - So $\Phi_i(t)$ only tries to optimize over the i-dimensional homology features

Variations

- One variation of the objective function is to replace $\Phi(t)$ with $\Phi_i(t)$ which is the same as $\Phi(t)$ by considering only the i-th PD
  - So $\Phi_i(t)$ only tries to optimize over the i-dimensional homology features
- Then, we can assign a weight w_i to each dimension p indicating how one wants to stress this dimension

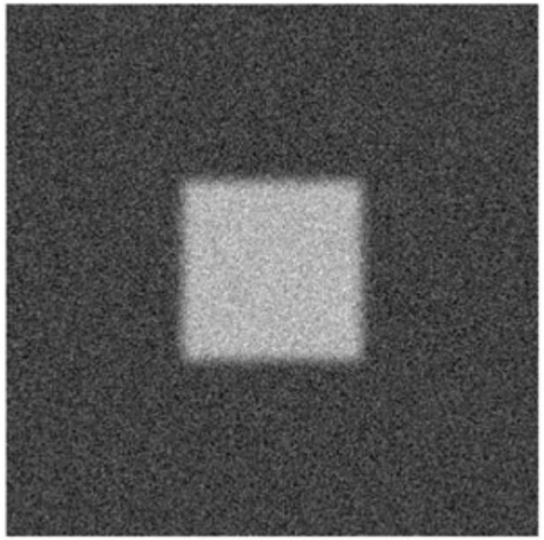
Variations

- One variation of the objective function is to replace $\Phi(t)$ with $\Phi_i(t)$ which is the same as $\Phi(t)$ by considering only the i-th PD
  - So $\Phi_i(t)$ only tries to optimize over the i-dimensional homology features
- Then, we can assign a weight w_i to each dimension p indicating how one wants to stress this dimension
- Then final objective function is a weighted sum of $\Phi_i(t)$:

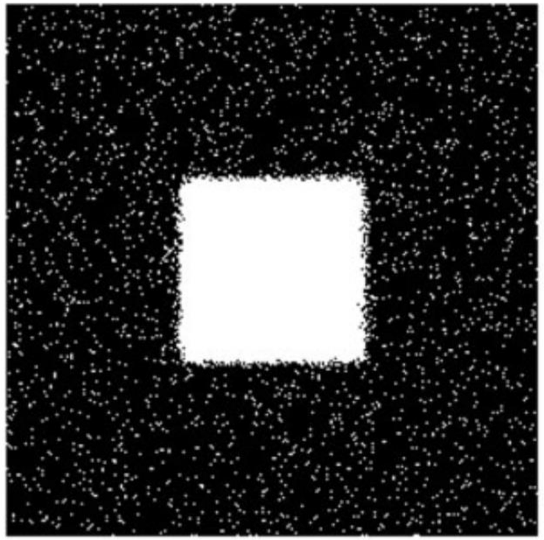
$$T = \arg\max \sum_{i=0}^{\dim -1} w_i \Phi_i(t),$$

Variations

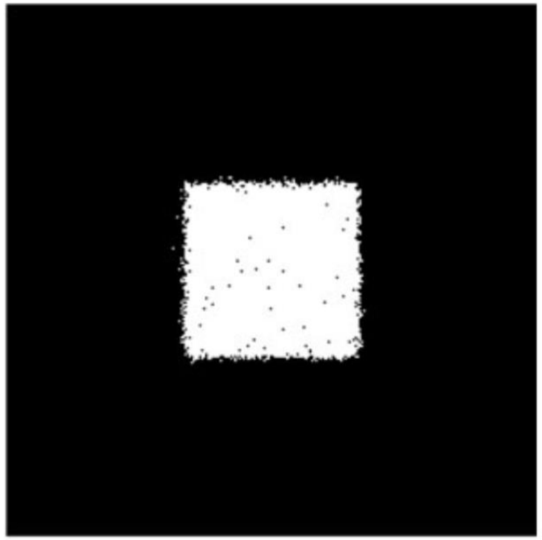
Original



$w_0 = 1, w_1 = 0$



$w_0 = 0, w_1 = 1$



Variations

$$\Psi(t) = \frac{1}{\#\mathscr{P}(t) + 1} \sum_{(b,d)\in\mathscr{P}(t)} [(d - t)(t - b)]^p$$

Variations

- Take the p-th power for each term in $\Psi(t)$, to put greater emphasis on long intervals

$$\Psi(t) = \frac{1}{\#\mathscr{P}(t) + 1} \sum_{(b,d)\in\mathscr{P}(t)} [(d-t)(t-b)]^p$$