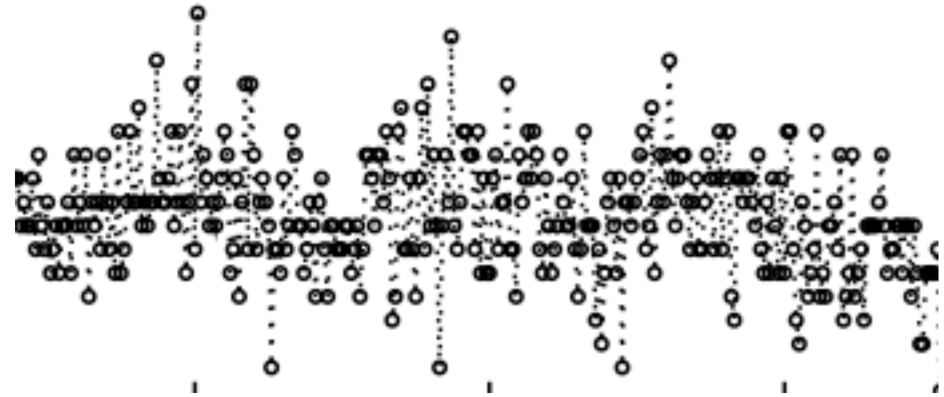
Tools from Computational Topology

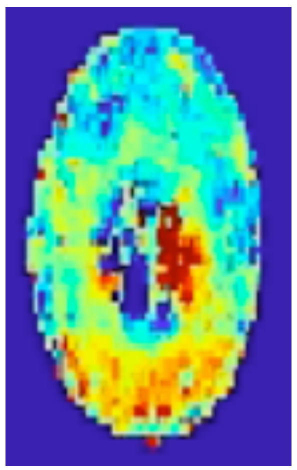
with applications in the life sciences

Sarah Day
Department of Mathematics
College of William & Mary

February 10, 2016

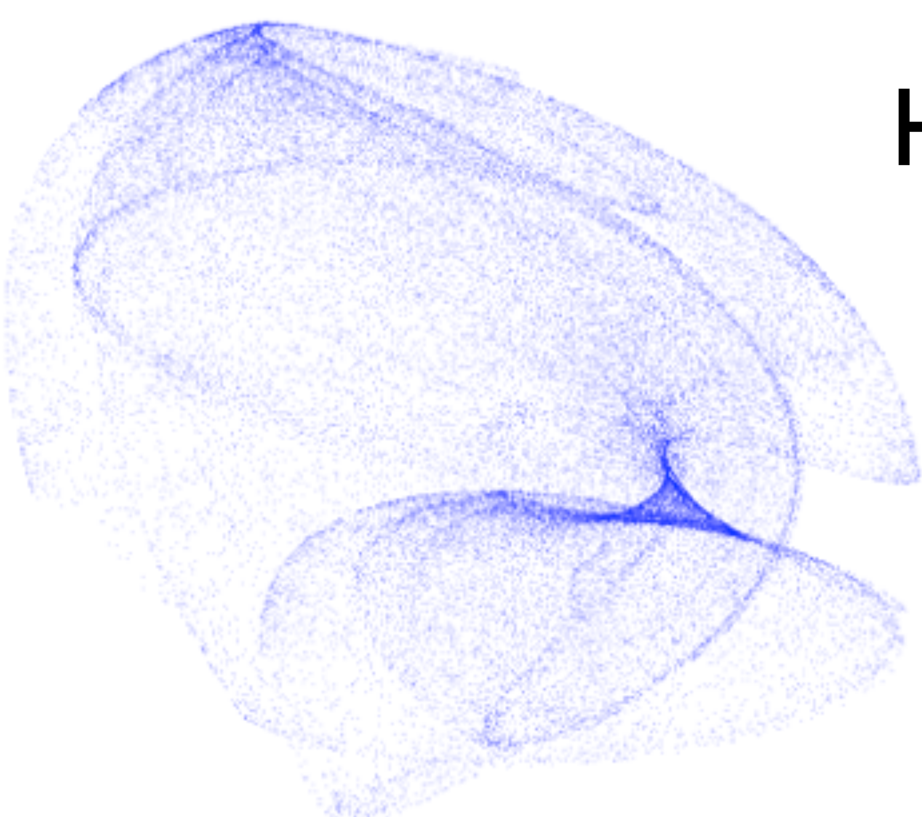
Data can be





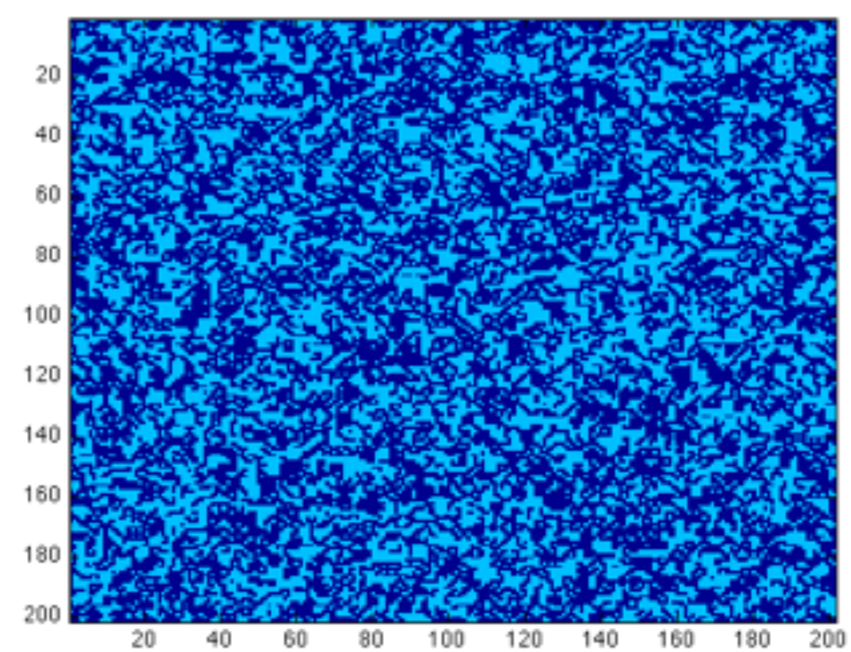
Messy (recurrent dynamics, chaos)

Noisy (measurement error, stochastity)



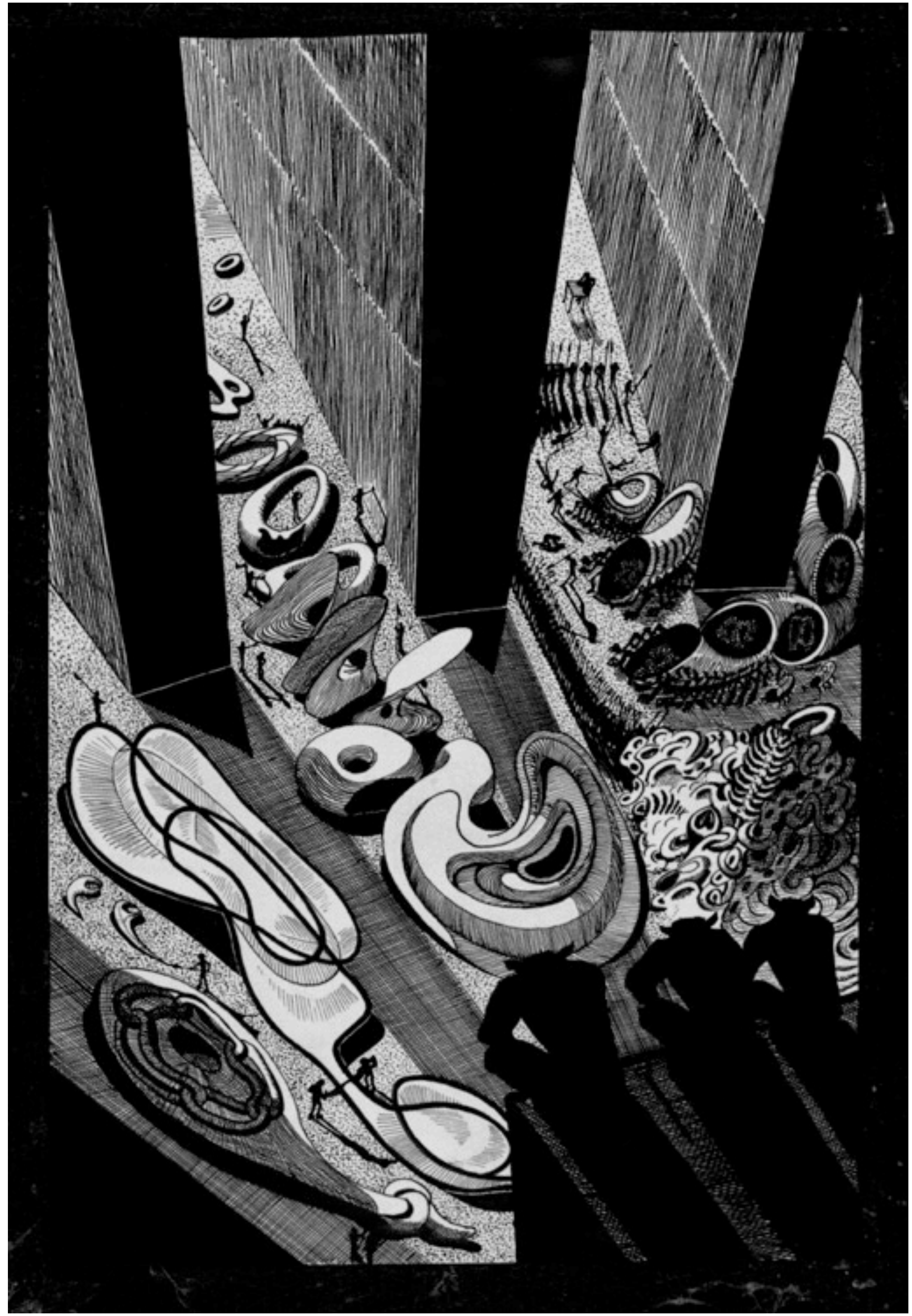
High dimensional

Sparse



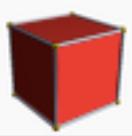
(Topology, Wikipedia)

Topological Zoo
Anatoly Fomenko
1967



Euler Characteristic

$$\chi = V - E + F$$

Name	Image	Vertices V	Edges E	Faces F	Euler characteristic: $V - E + F$
Tetrahedron		4	6	4	**2**
Hexahedron or cube		8	12	6	**2**
Octahedron		6	12	8	**2**
Dodecahedron		20	30	12	**2**
Icosahedron		12	30	20	**2**

Wikipedia

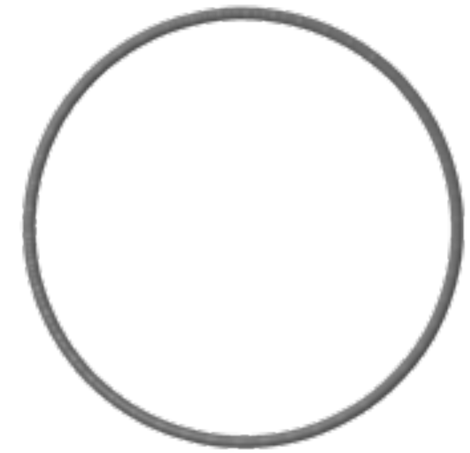
surface	χ
cylinder	0
double torus	−2
Klein bottle	0
Möbius strip	0
projective plane	1
sphere	2
torus	0

Wolfram

$$\chi(g) = 2 - 2g$$

Homology

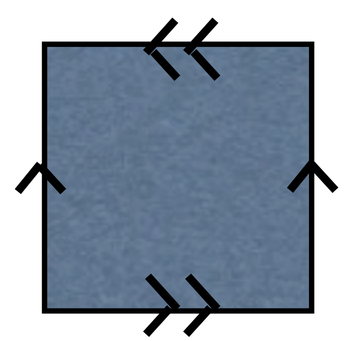
$$H_k(X) \cong \begin{cases} \mathbb{Z} & k = 0, 1 \\ 0 & \text{otherwise.} \end{cases}$$



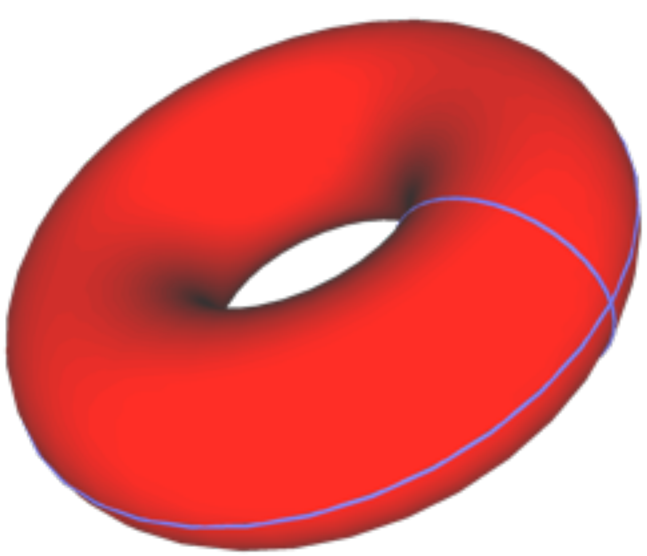
$$H_k(S^n) \cong \begin{cases} \mathbb{Z} & k = 0, n \\ 0 & \text{otherwise.} \end{cases}$$

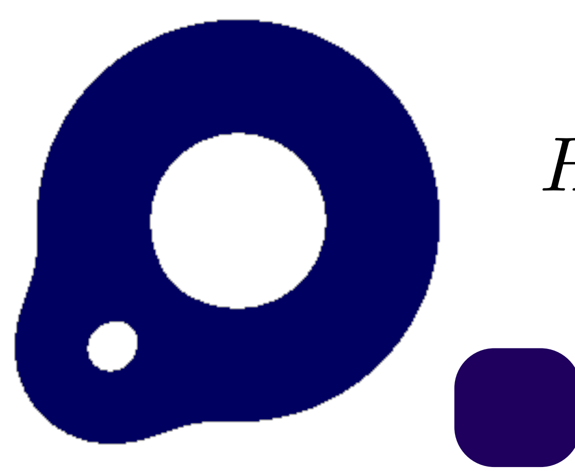
or

$$H_k(X) \cong \begin{cases} \mathbb{Z} & k = 0, 2 \\ \mathbb{Z}^2 & k = 1 \\ 0 & \text{otherwise.} \end{cases}$$



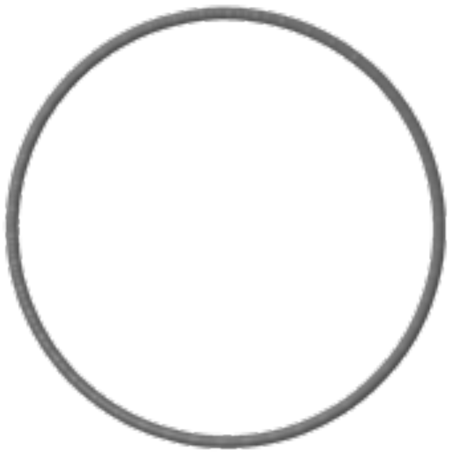
$$H_k(K) \cong \begin{cases} \mathbb{Z} & k = 0, \\ \mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z} & k = 1, \\ 0 & \text{otherwise.} \end{cases}$$





$$H_k(X) \cong \begin{cases} \mathbb{Z}^2 & k = 0, 1 \\ 0 & \text{otherwise.} \end{cases}$$

Betti numbers



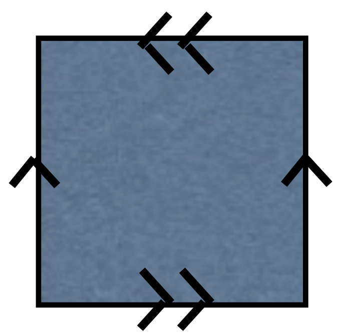
$$H_k(X) \cong \begin{cases} \mathbb{Z} & k = 0, 1 \\ 0 & \text{otherwise.} \end{cases}$$

$$\boxed{\begin{array}{l} \beta_0 = 1 \\ \beta_1 = 1 \end{array}}$$

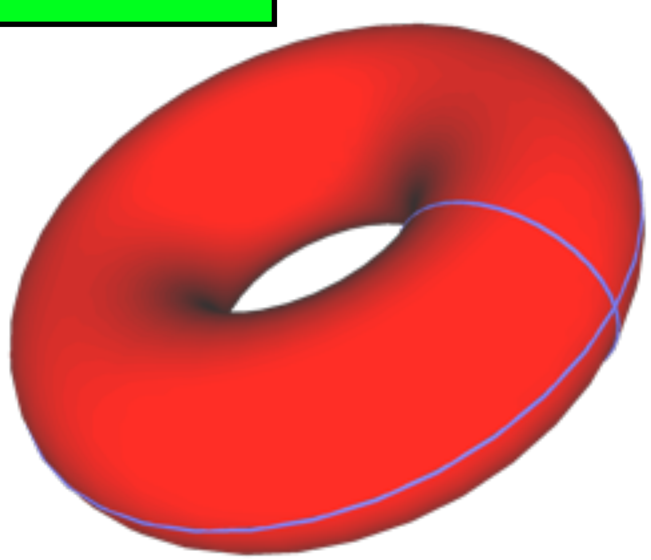
or

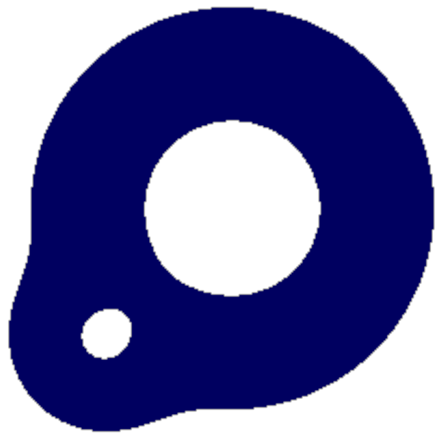
$$\boxed{\begin{array}{l} \beta_0 = 1 \\ \beta_1 = 2 \\ \beta_2 = 1 \end{array}}$$

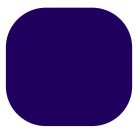
$$H_k(X) \cong \begin{cases} \mathbb{Z} & k = 0, 2 \\ \mathbb{Z}^2 & k = 1 \\ 0 & \text{otherwise.} \end{cases}$$



$$H_k(K) \cong \begin{cases} \mathbb{Z} & k = 0, \\ \mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z} & k = 1, \\ 0 & \text{otherwise.} \end{cases}$$



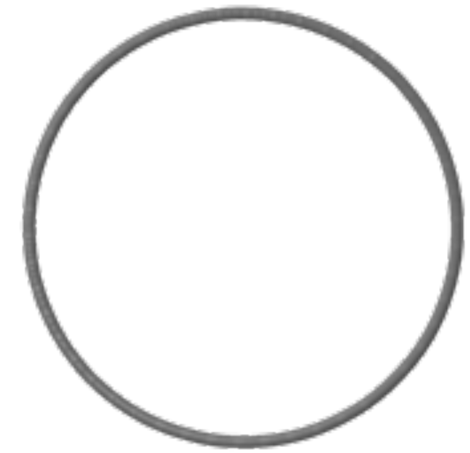




$$\boxed{\begin{array}{l} \beta_0 = 2 \\ \beta_1 = 2 \end{array}}$$

$$H_k(X) \cong \begin{cases} \mathbb{Z}^2 & k = 0, 1 \\ 0 & \text{otherwise.} \end{cases}$$

Betti numbers and Euler Characteristic



$$H_k(X) \cong \begin{cases} \mathbb{Z} & k = 0, 1 \\ 0 & \text{otherwise.} \end{cases}$$

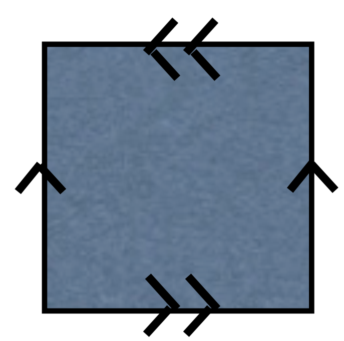
$$\boxed{\begin{aligned} \beta_0 &= 1 \\ \beta_1 &= 1 \end{aligned}}$$

$$\boxed{\chi = \sum (-1)^n k_n = \sum (-1)^n \beta_n}$$

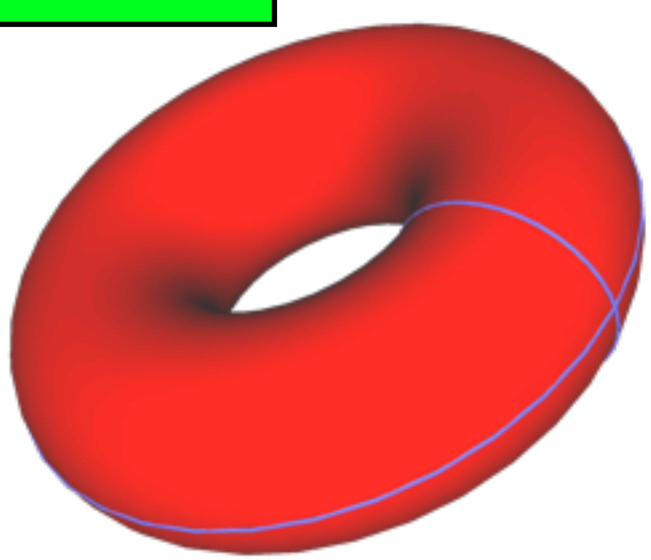
or

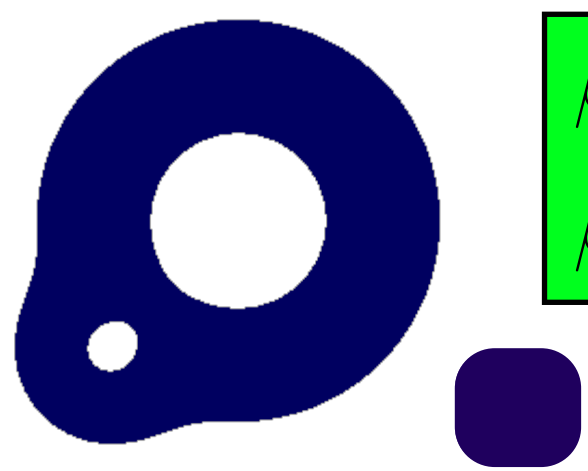
$$\boxed{\begin{aligned} \beta_0 &= 1 \\ \beta_1 &= 2 \\ \beta_2 &= 1 \end{aligned}}$$

$$H_k(X) \cong \begin{cases} \mathbb{Z} & k = 0, 2 \\ \mathbb{Z}^2 & k = 1 \\ 0 & \text{otherwise.} \end{cases}$$



$$H_k(K) \cong \begin{cases} \mathbb{Z} & k = 0, \\ \mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z} & k = 1, \\ 0 & \text{otherwise.} \end{cases}$$

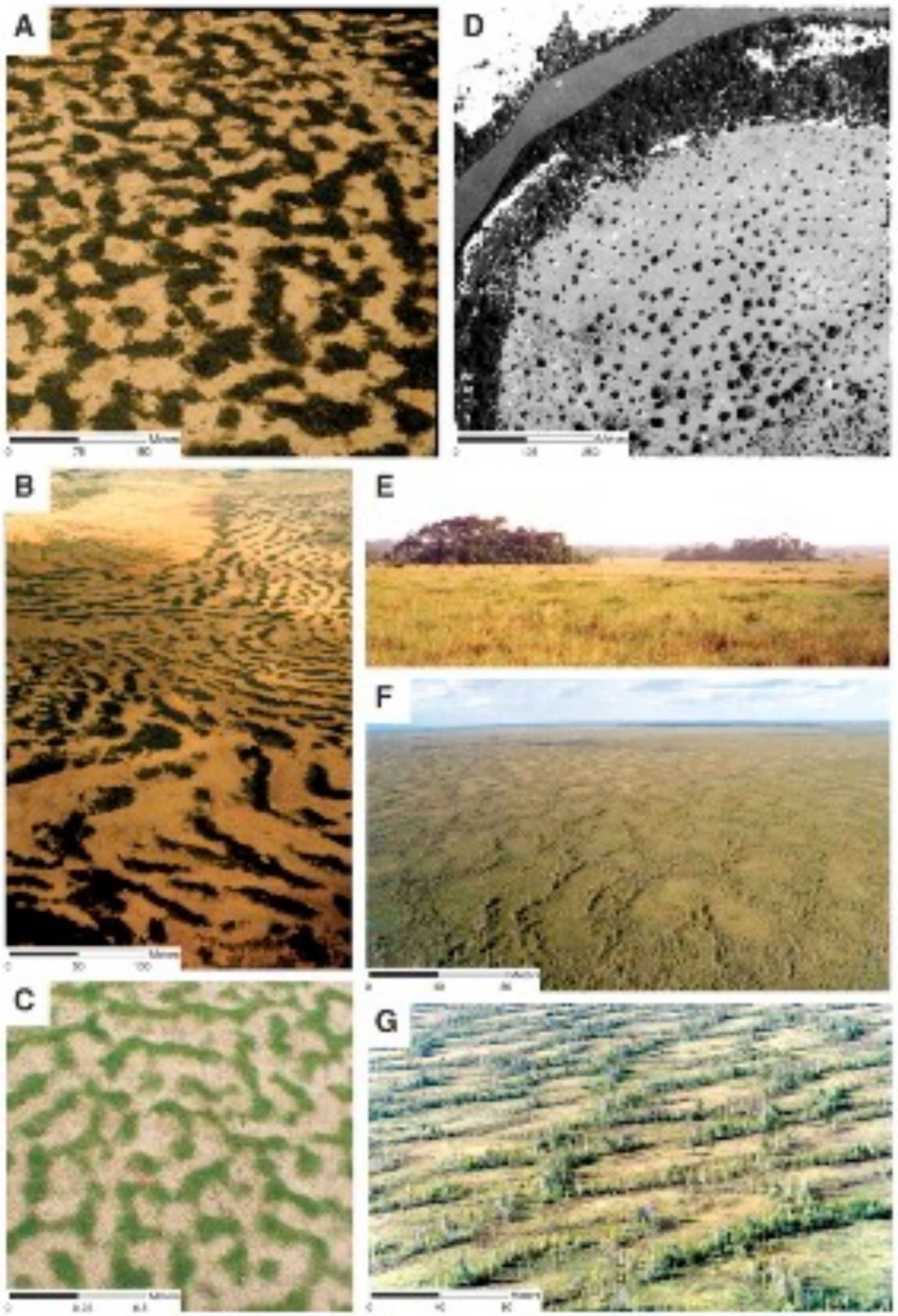




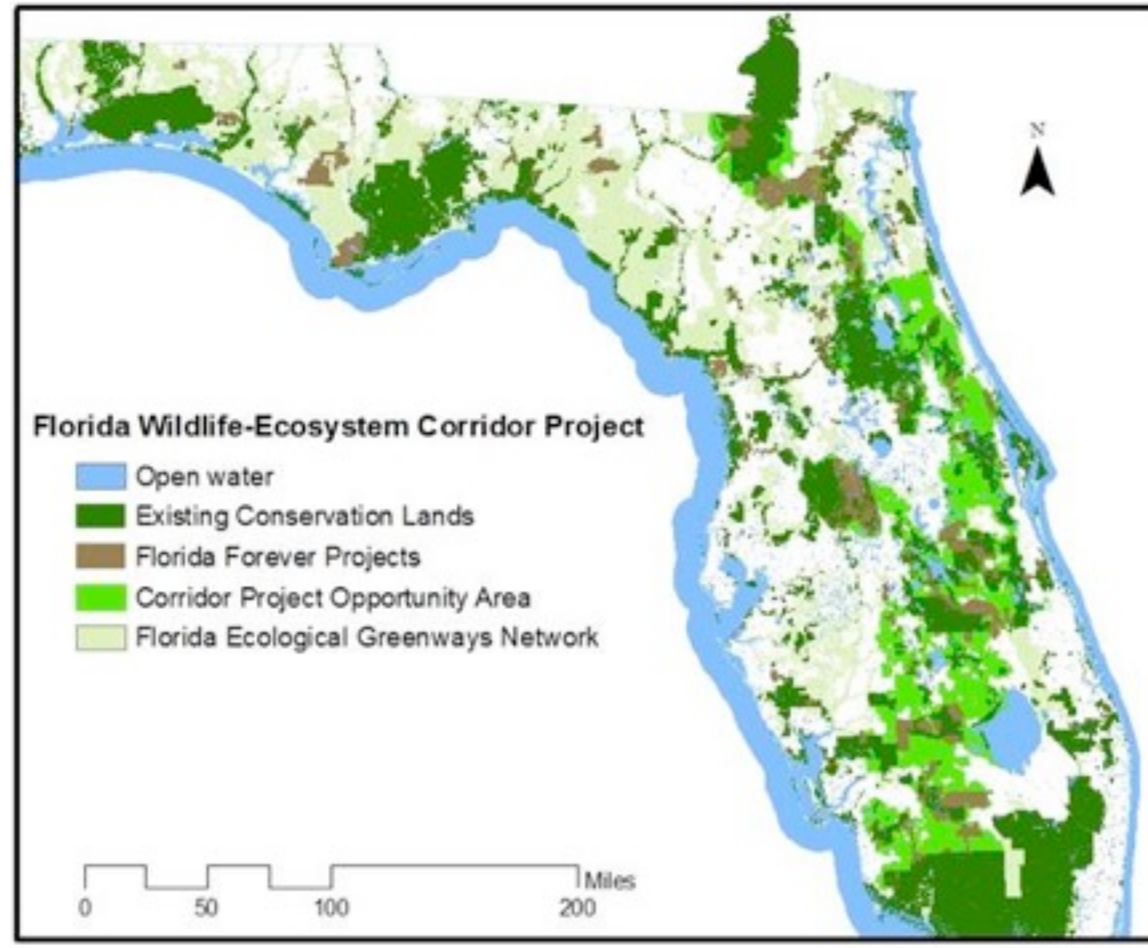
$$\boxed{\begin{aligned} \beta_0 &= 2 \\ \beta_1 &= 2 \end{aligned}}$$

$$H_k(X) \cong \begin{cases} \mathbb{Z}^2 & k = 0, 1 \\ 0 & \text{otherwise.} \end{cases}$$

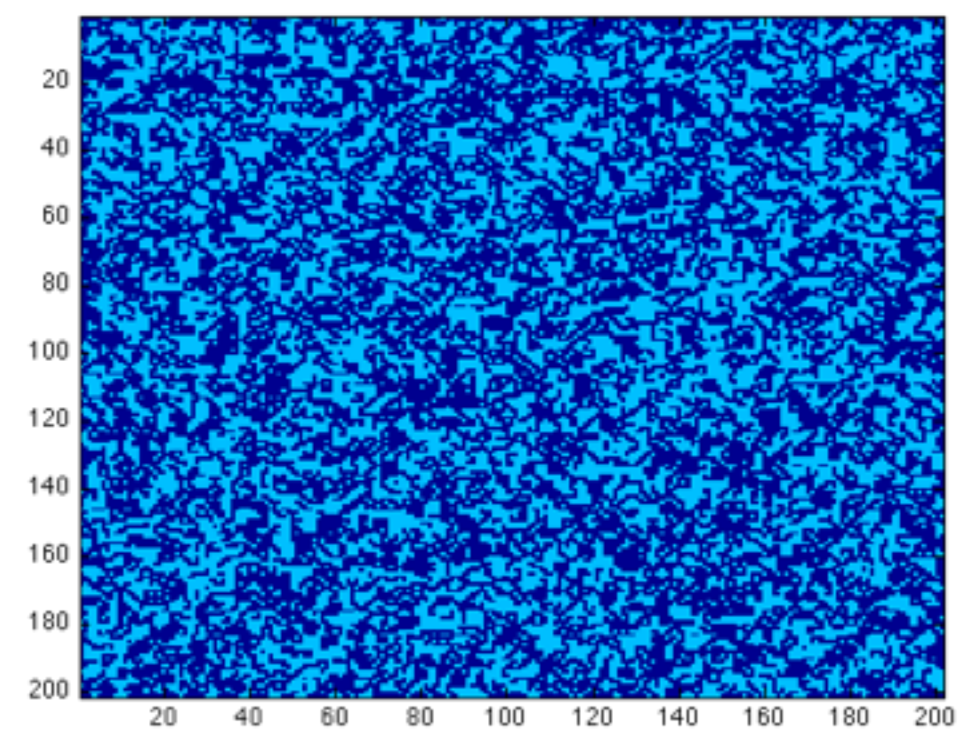
population density patterns



M Rietkerk et al. Science 2004;305:1926-1929

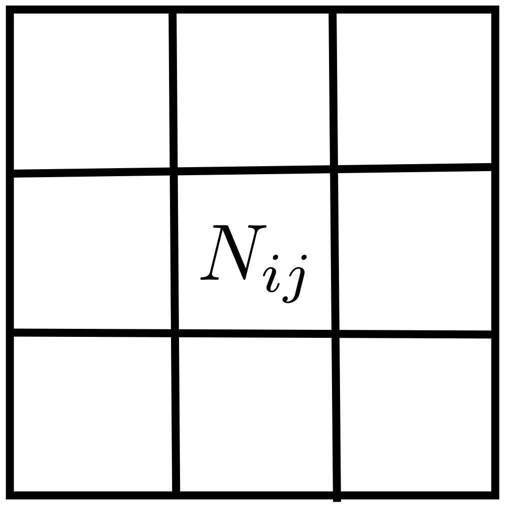


http://www.linc.us/FloridaWidlifeCorridor_Info.html



Project I: Coupled-Patch Model

(with Ben Holman)



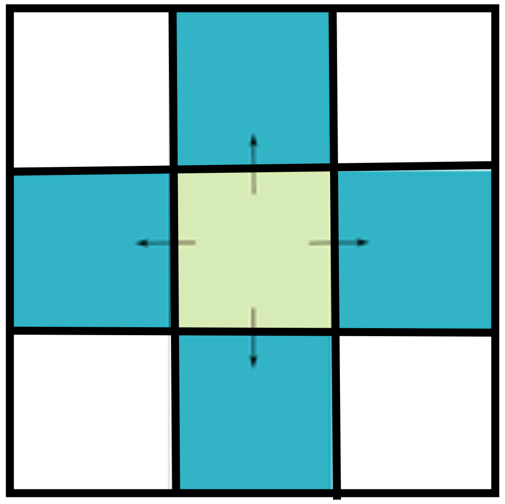
Ricker map $f(N) = rNe^{-N}$

fitness parameter

growth phase

$$\bar{N}_{ij}(t+1) = f(N_{ij}(t))$$

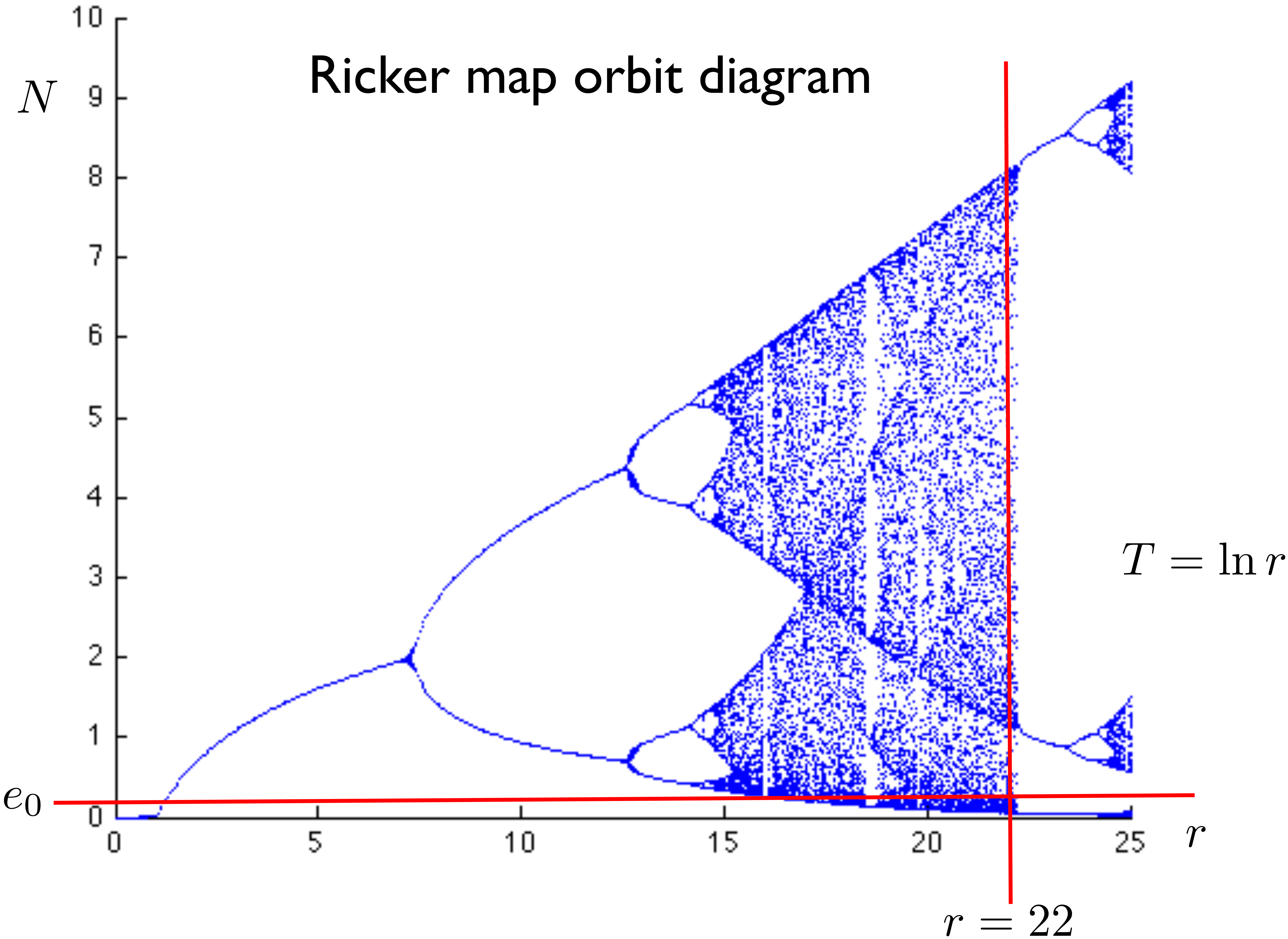
dispersal phase



$$N_{ij}(t+1) = (1-d)\bar{N}_{ij}(t) + \frac{d}{4} \sum_{|i-i'|+|j-j'|=1} \bar{N}_{i'j'}(t)$$

dispersal parameter

Ricker map orbit diagram



Example 1: dispersal and smoothing

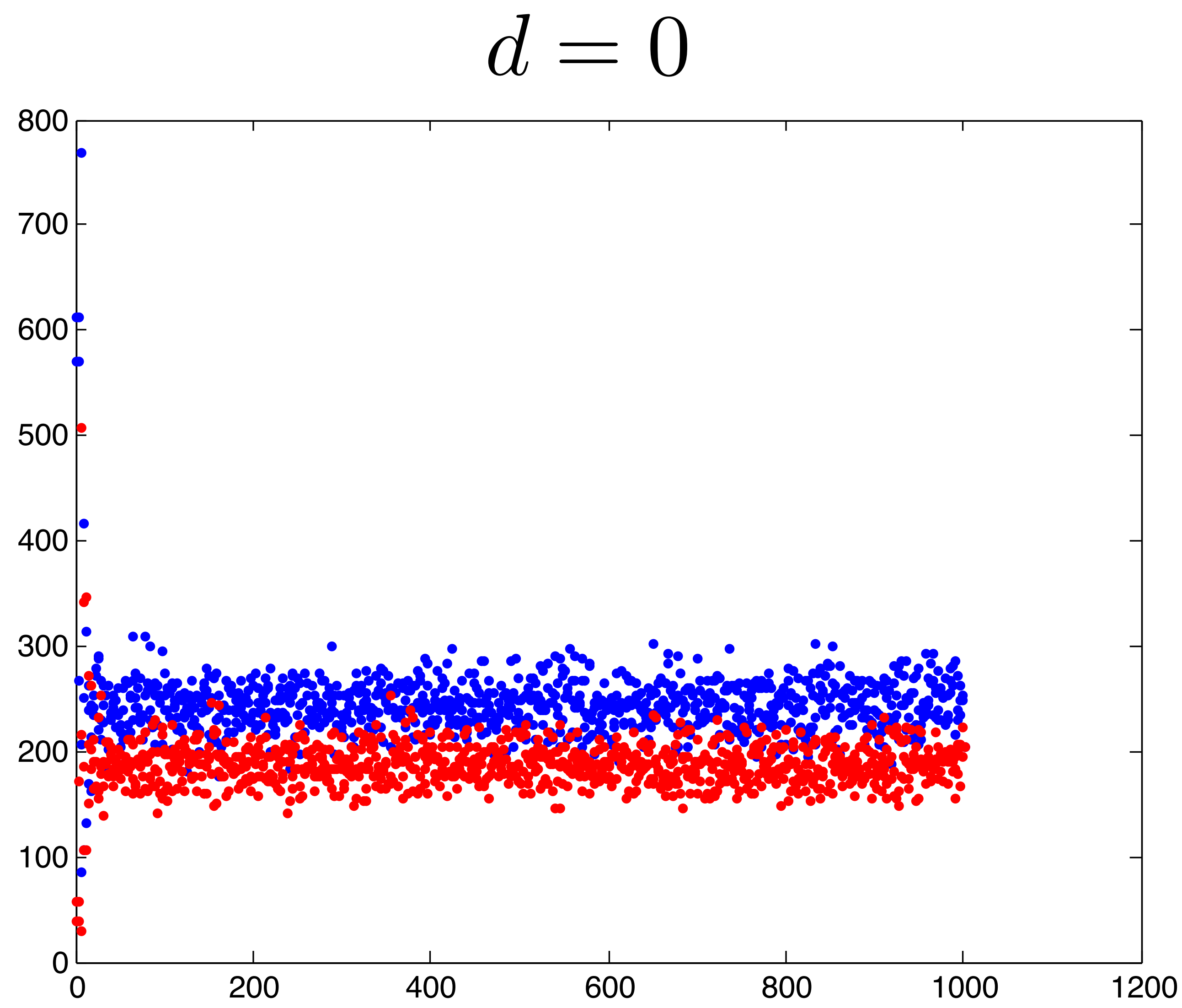
$d = 0$

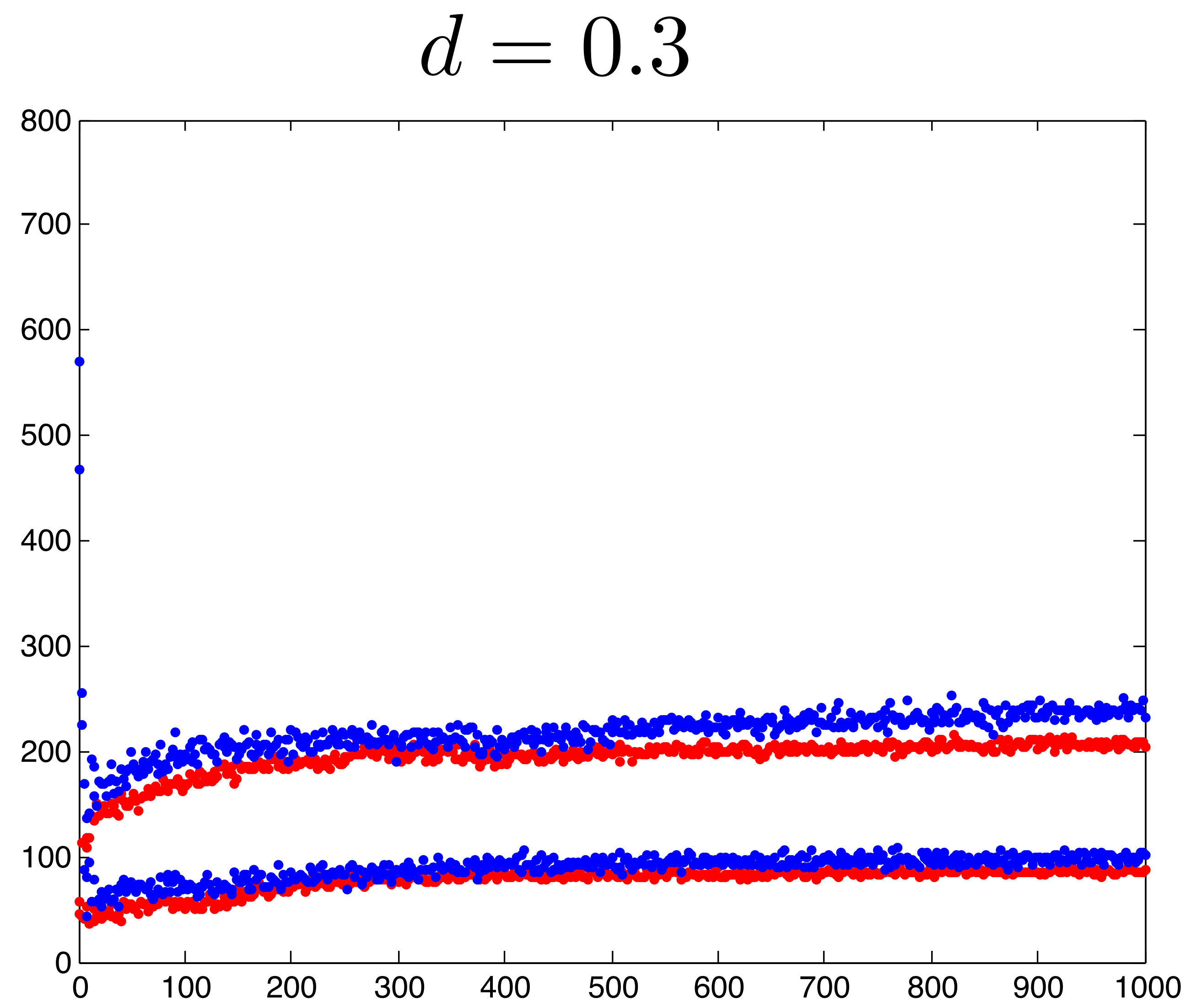
$d = 0.3$

$$n = 100, r = 22$$

Example 1: dispersal and smoothing

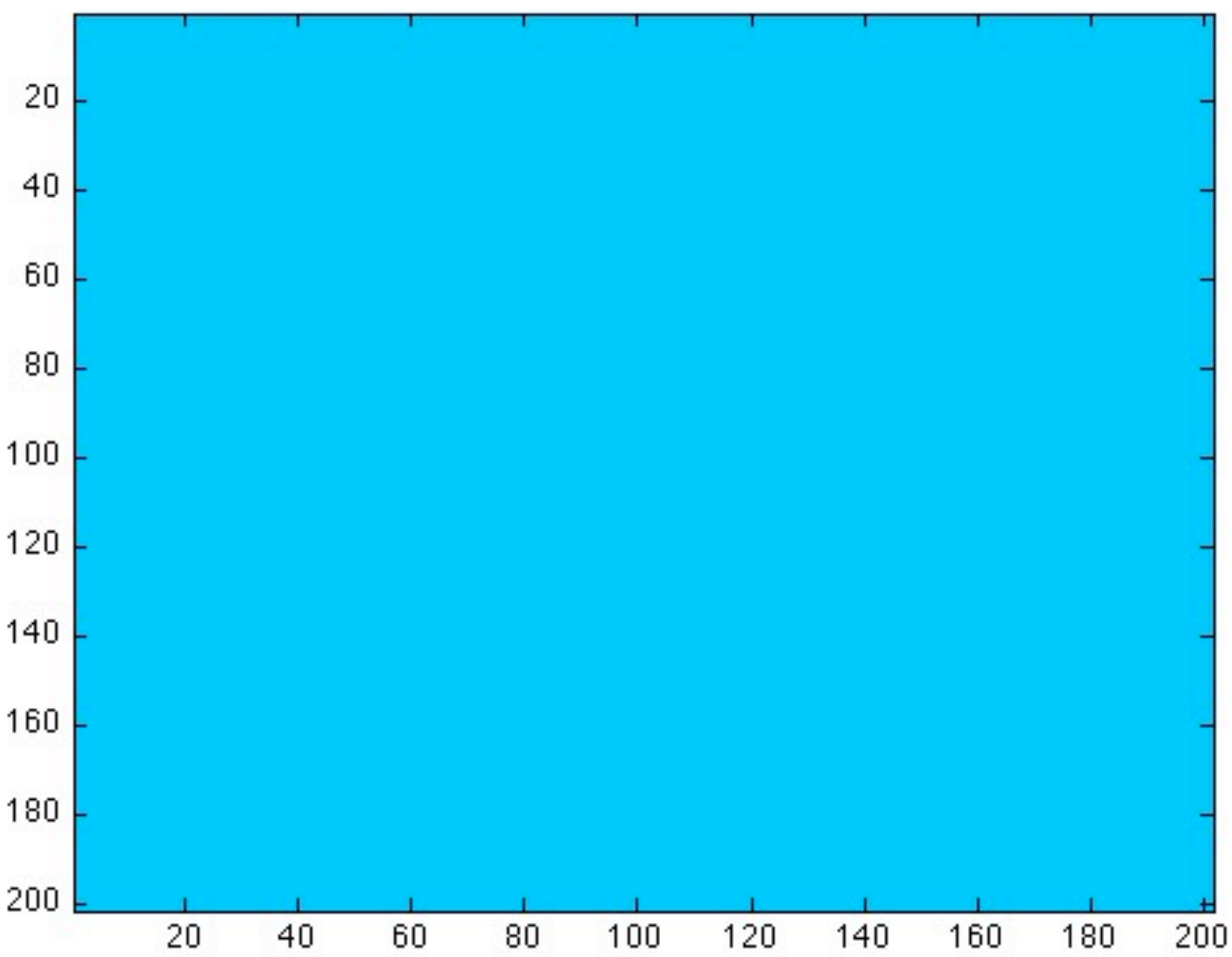
β_0, β_1 vs t





Example 2: global extinction event

$t = 0$



total abundance:

$$\|N\|_1 = 142{,}380$$

decoupled subsystems:

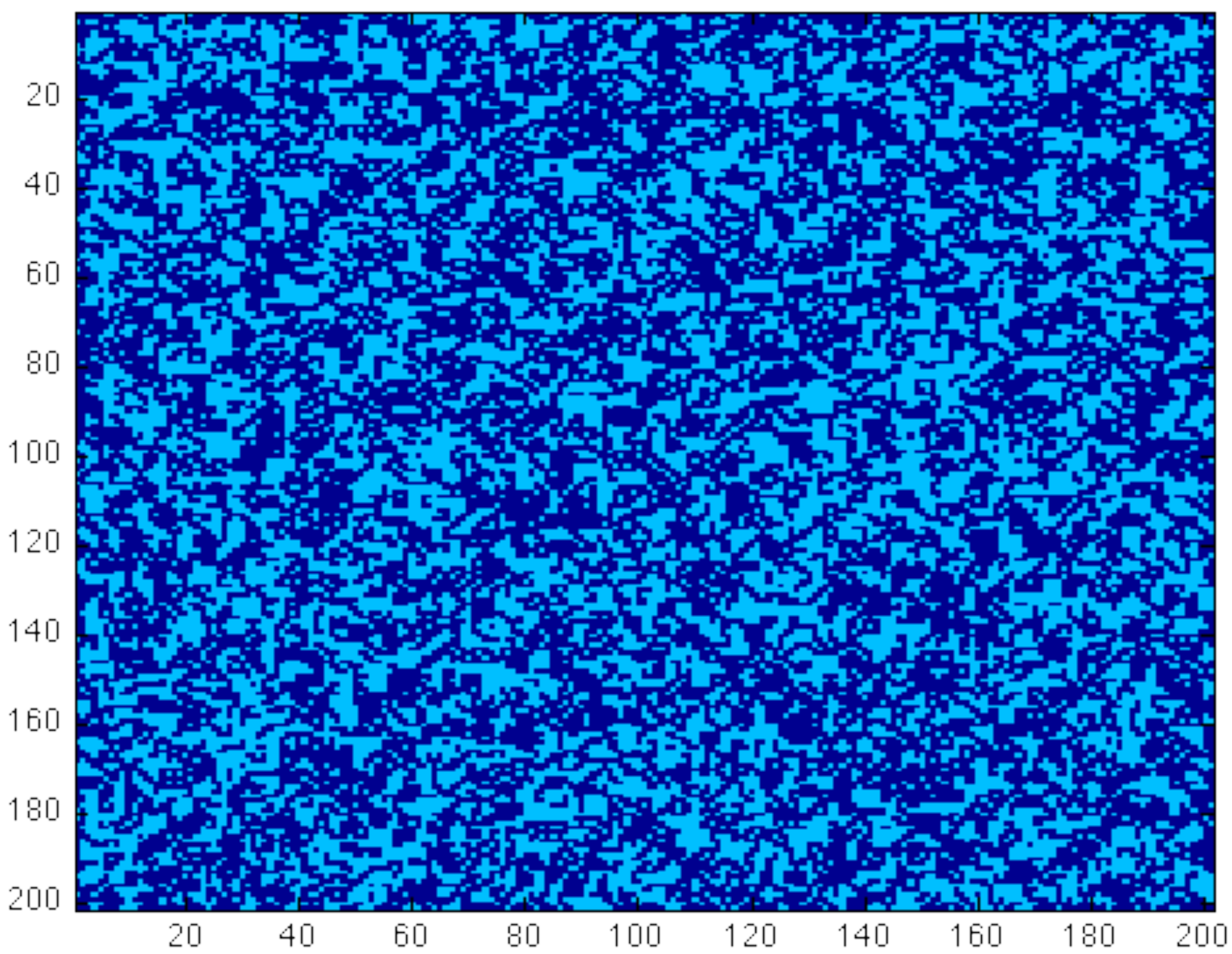
$$\beta_0 = 1$$

enclosed extinct regions:

$$\beta_1 = 0$$

Example 2: global extinction event

$$t = 1$$



total abundance:

$$\|N\|_1 = 133{,}186$$

decoupled subsystems:

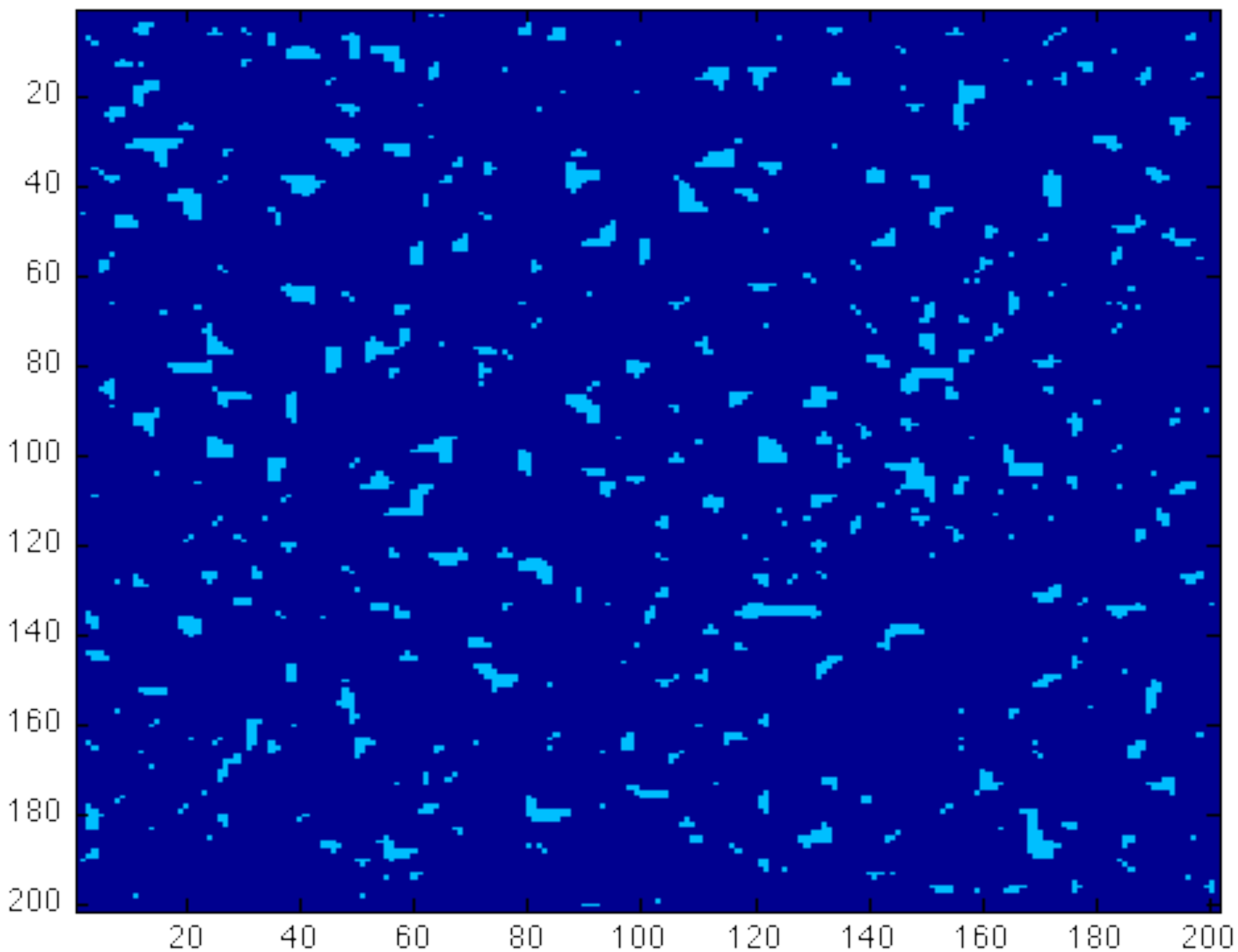
$$\beta_0 = 222$$

enclosed extinct regions:

$$\beta_1 = 1{,}115$$

Example 2: global extinction event

$t = 5$



total abundance:

$$\|N\|_1 = 10{,}339$$

decoupled subsystems:

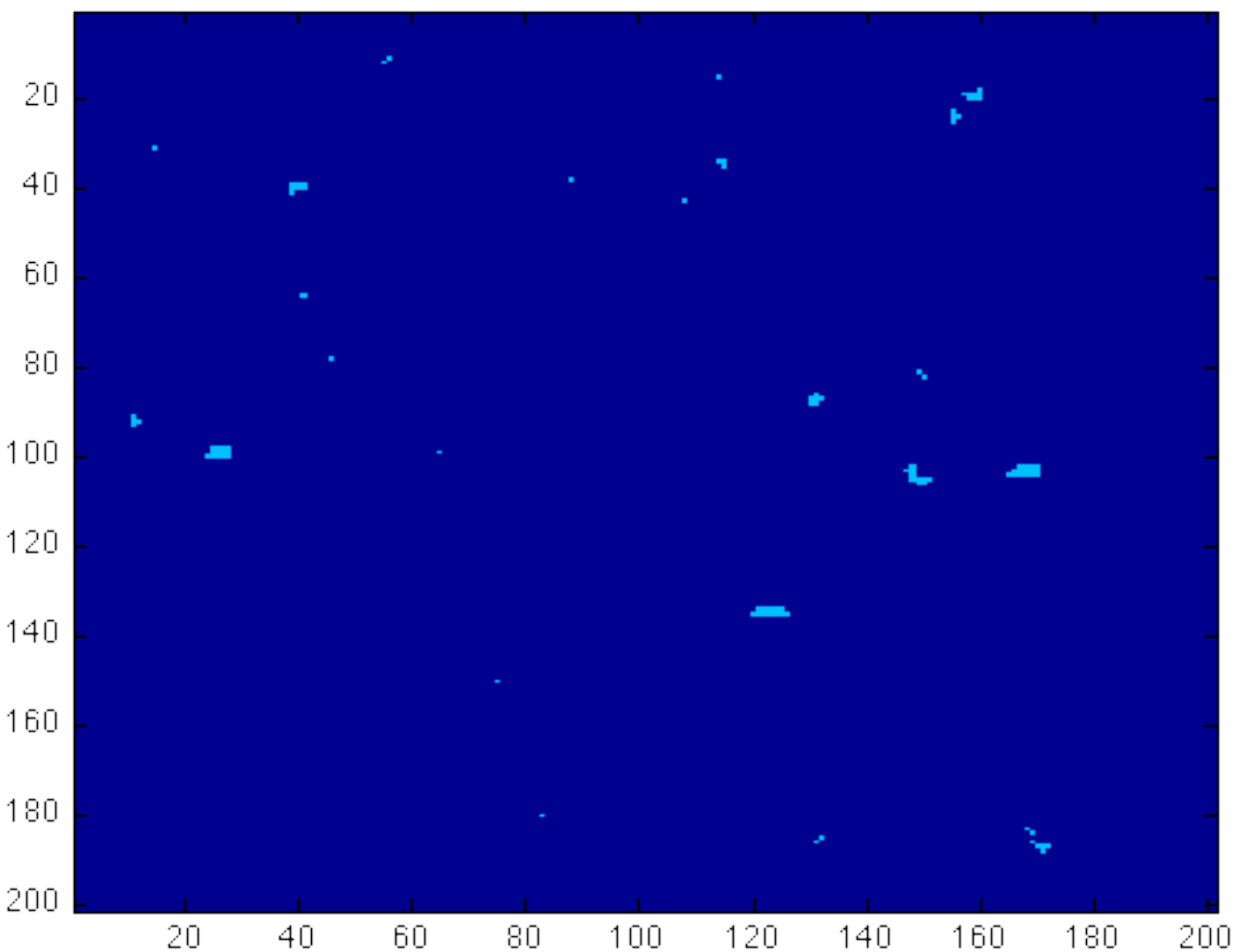
$$\beta_0 = 388$$

enclosed extinct regions:

$$\beta_1 = 0$$

Example 2: global extinction event

$t = 10$



total abundance:

$$\|N\|_1 = 10$$

decoupled subsystems:

$$\beta_0 = 22$$

enclosed extinct regions:

$$\beta_1 = 0$$

Three scenarios

The system decouples into a few small subsystems before rebounding and re-coupling.

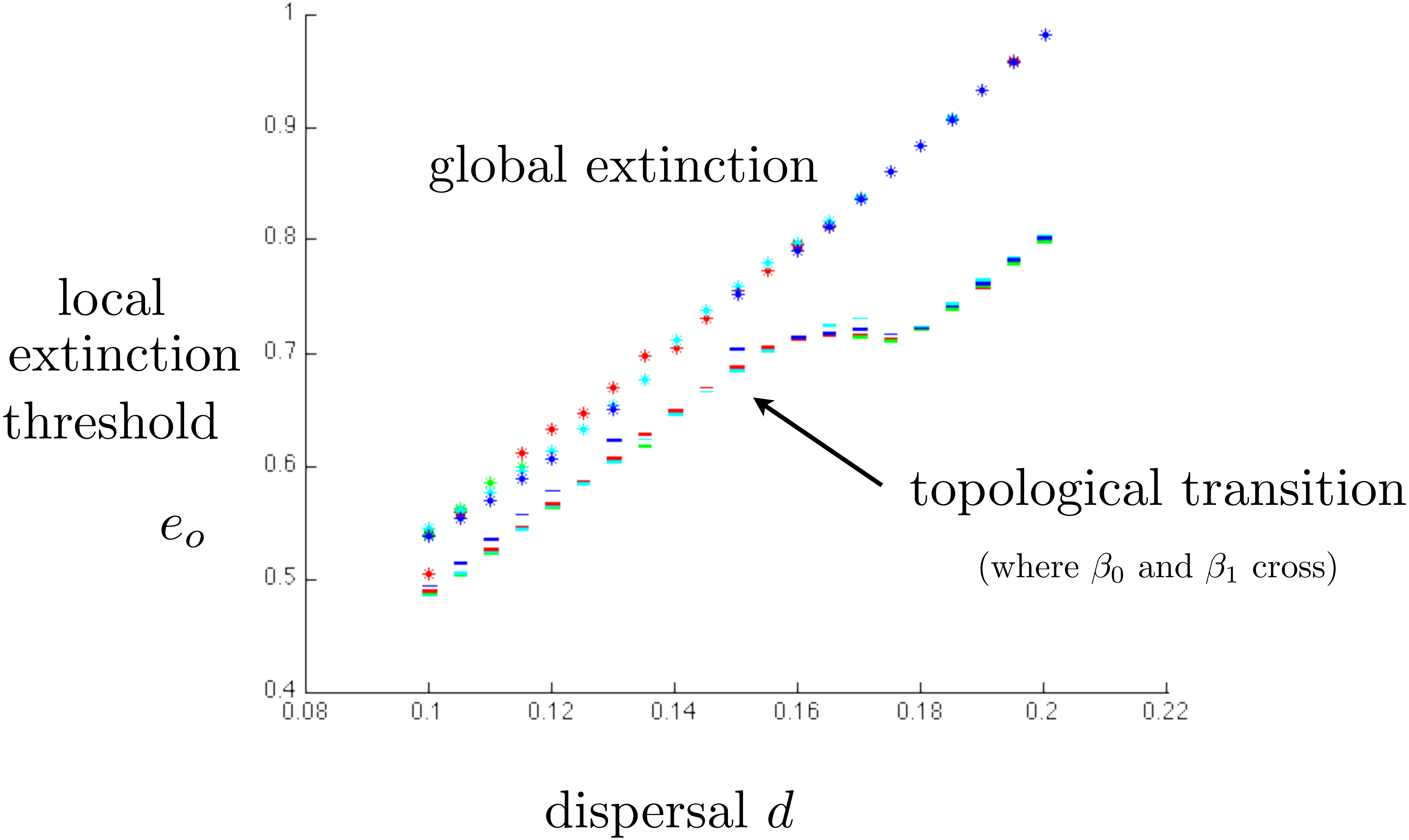
$$(r = 22, d = 0.15, e_o = 0.2)$$

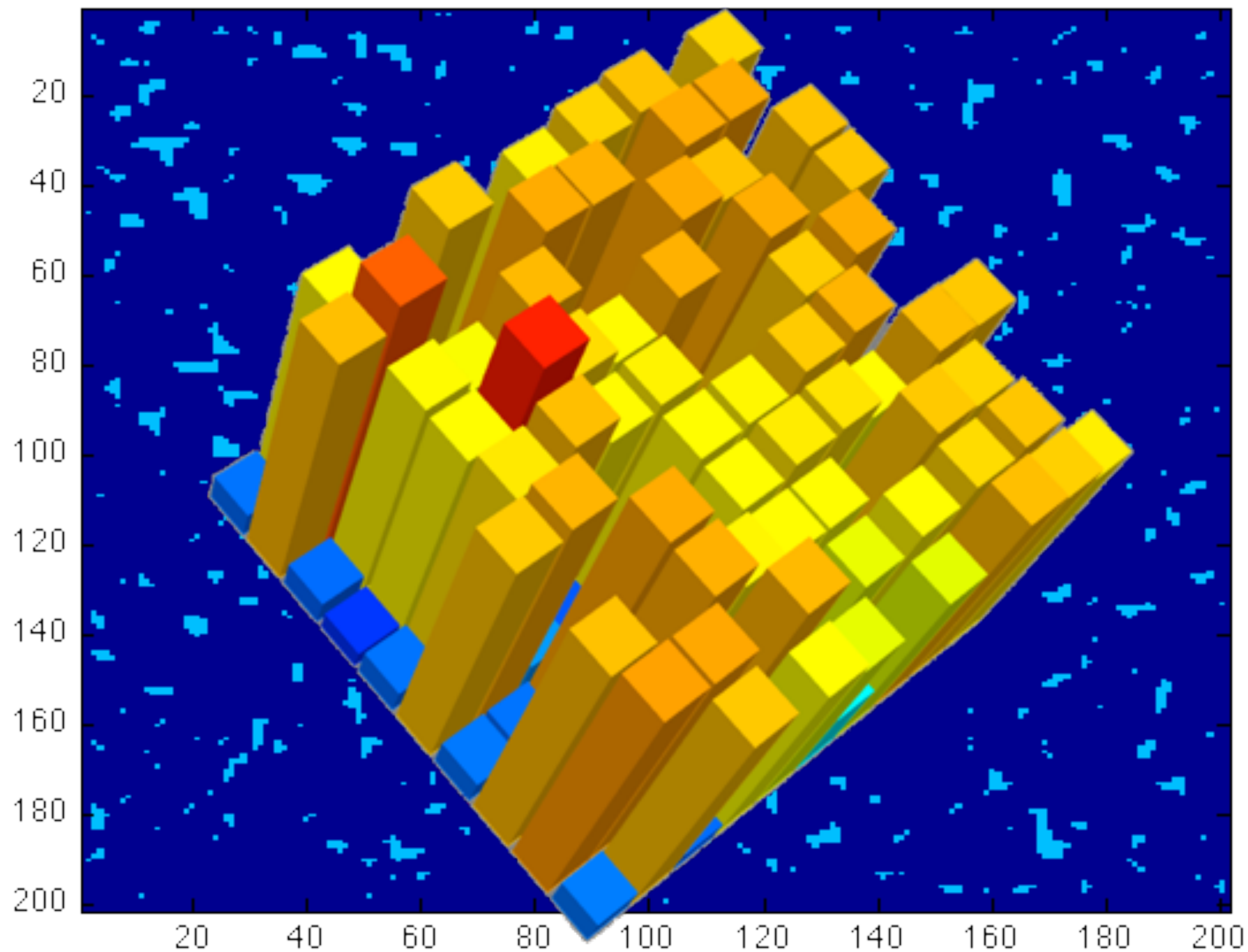
The system decouples into persistent subsystems.

$$(r = 22, d = 0.15, e_o = 0.6)$$

The system decouples into subsystems but only as a transient stage prior to extinction.

$$(r = 22, d = 0.15, e_o = 0.77)$$


global extinction
local extinction threshold
e_o
topological transition
(where β_0 and β_1 cross)
dispersal d
1
0.9
0.8
0.7
0.6
0.5
0.4
0.08
0.1
0.12
0.14
0.16
0.18
0.2
0.22

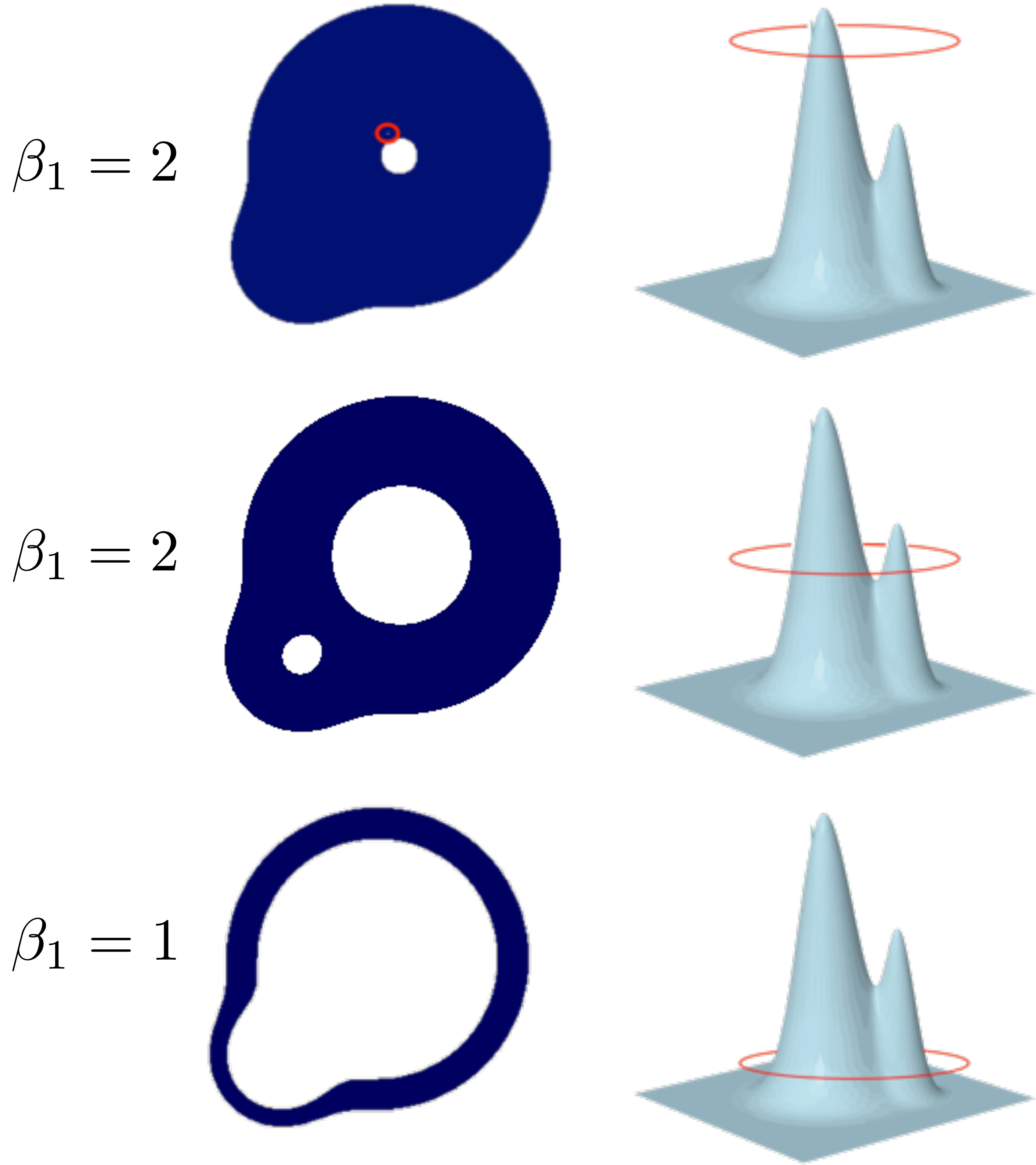


What are the effects of threshold choice, noise, and measurement error?

From Homology to Persistent Homology

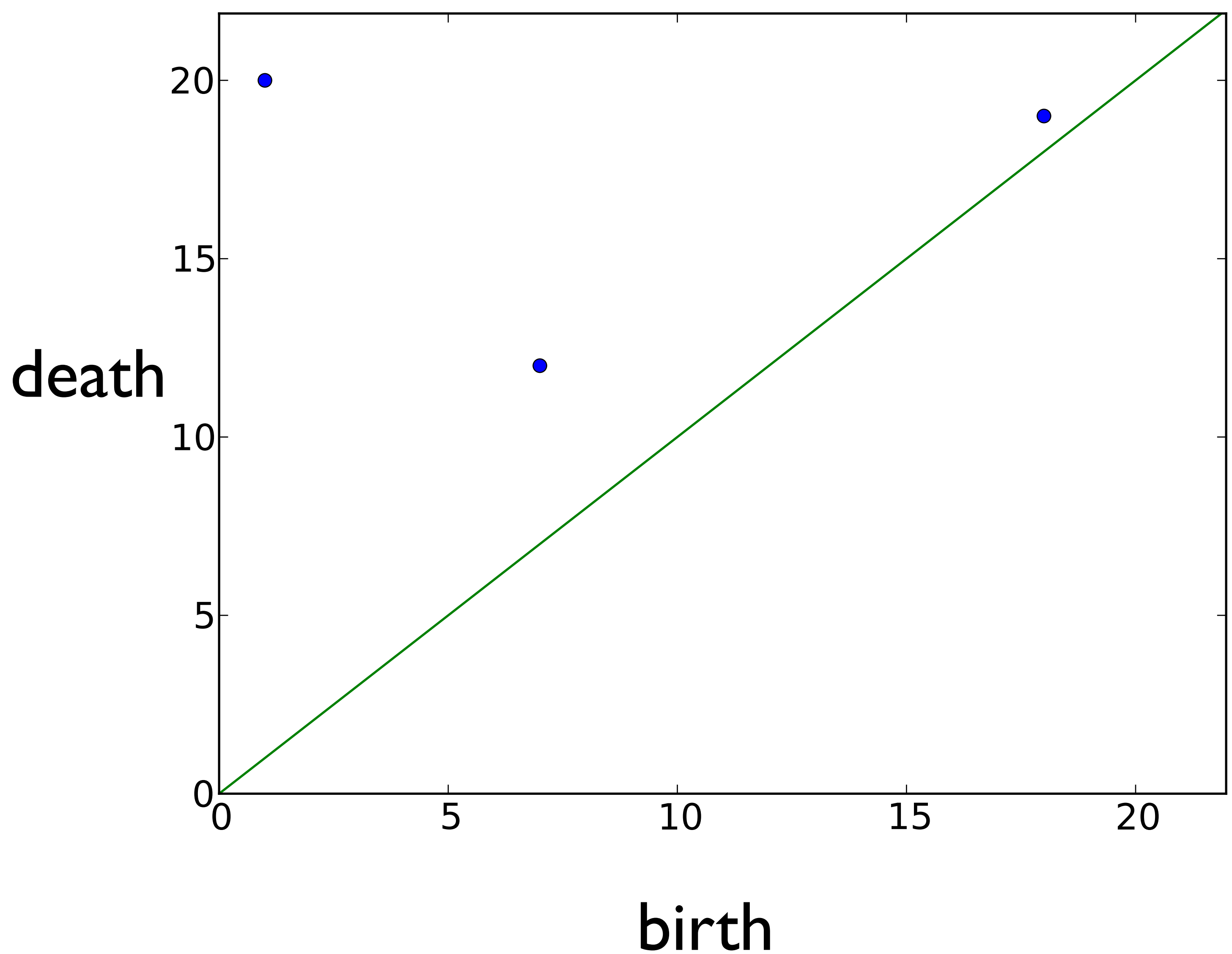
$$X_\tau := \{x \,|\, f(x) \leq \tau\}$$

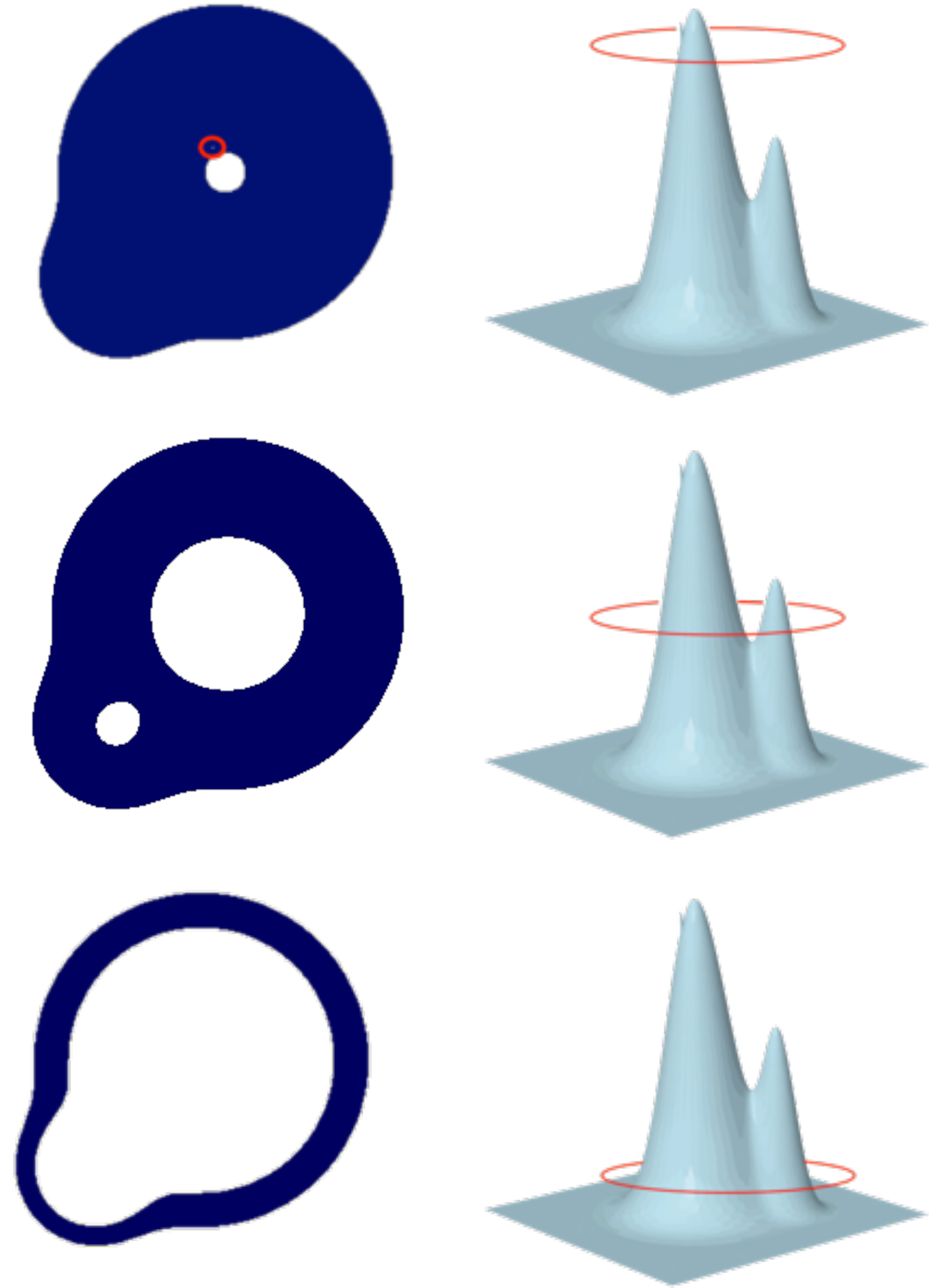
- We will focus on computing the homology of sublevel sets X_τ over a continuous range of thresholds.
- For each generator (hole) we record its birth threshold b and death threshold d.
- The importance of a homology generator (topological feature) is correlated to the generator's *lifespan* $(d - b)$.



Persistent Homology

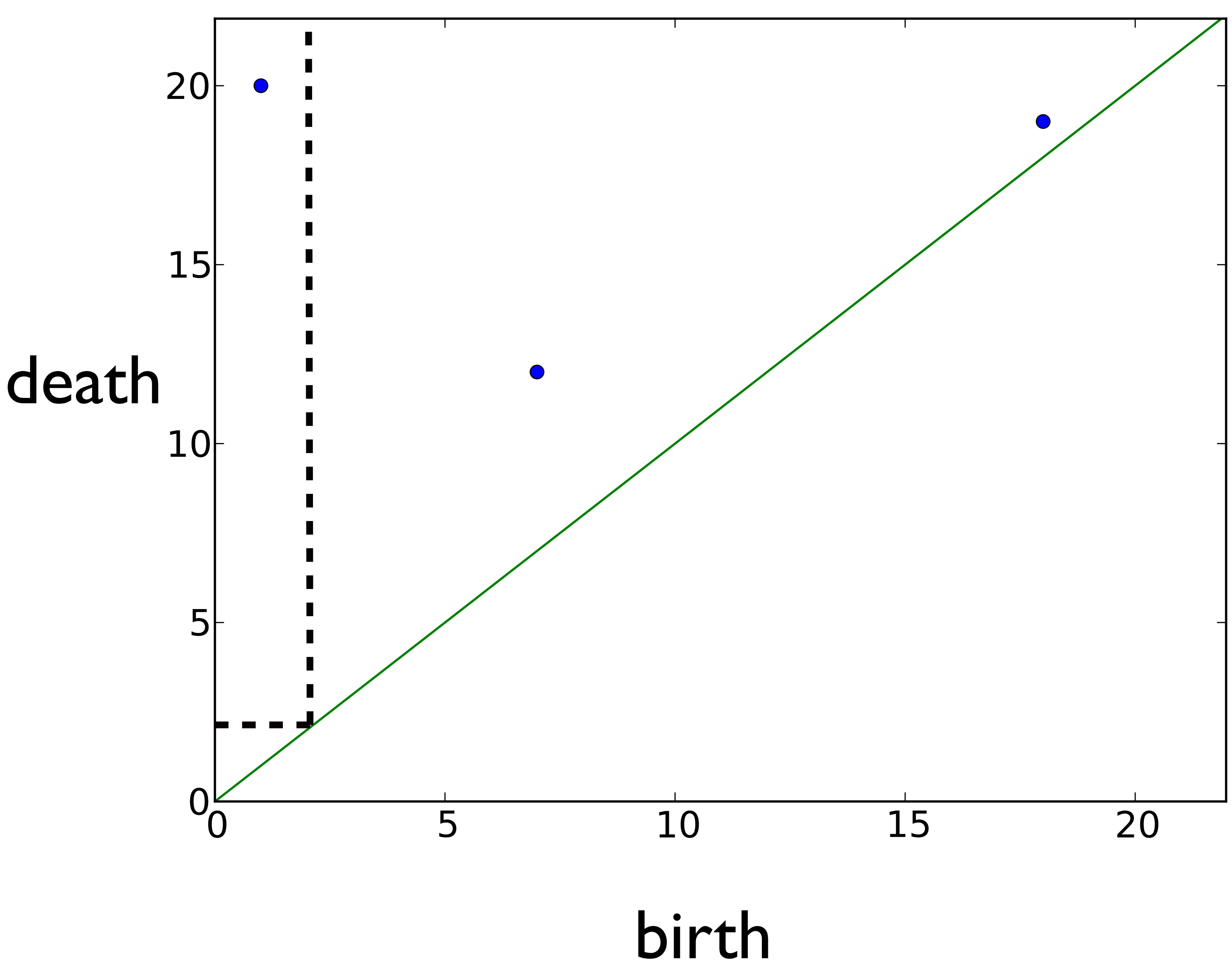
H_1 Persistence Diagram

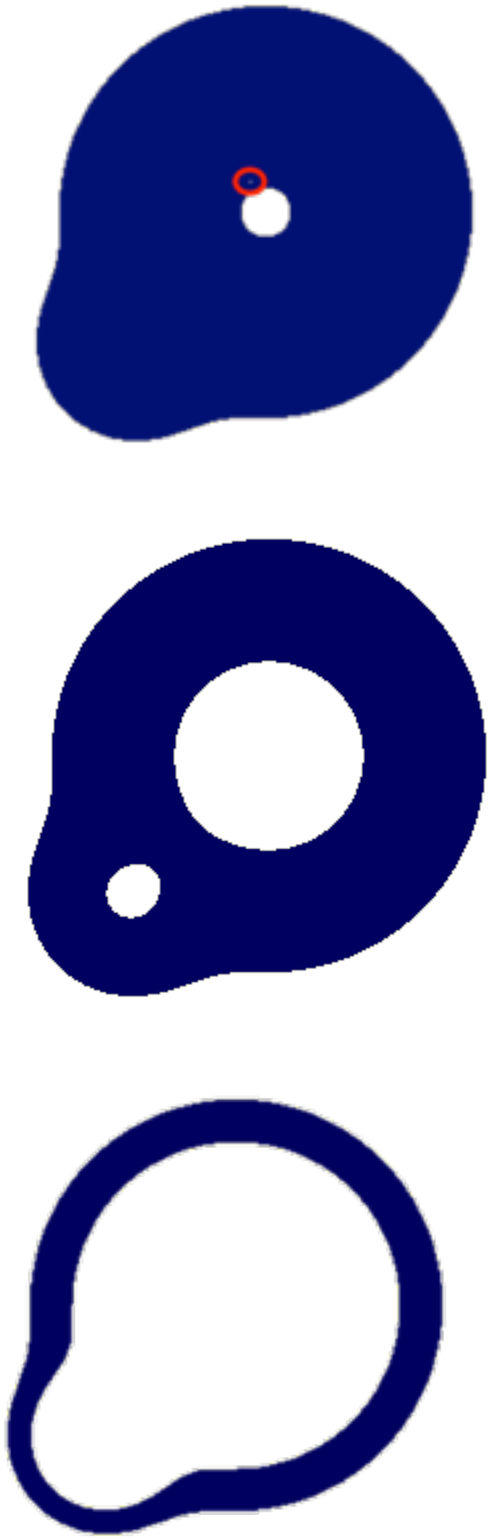


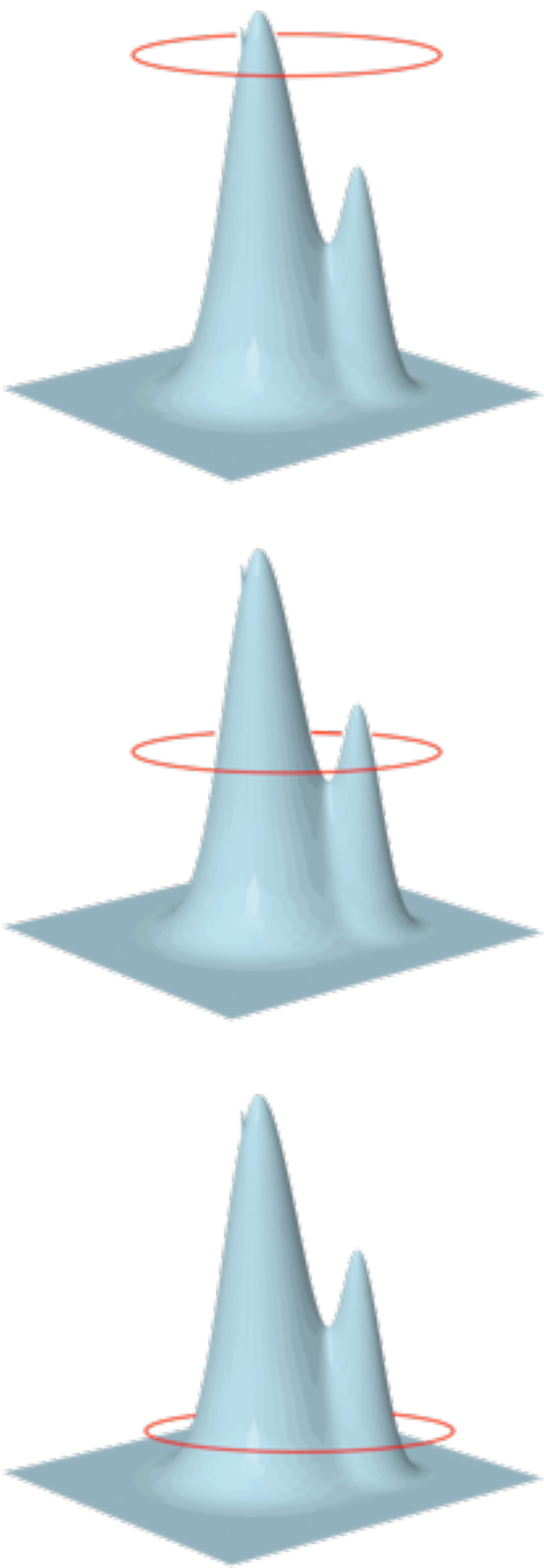


Persistent Homology

H_1 Persistence Diagram

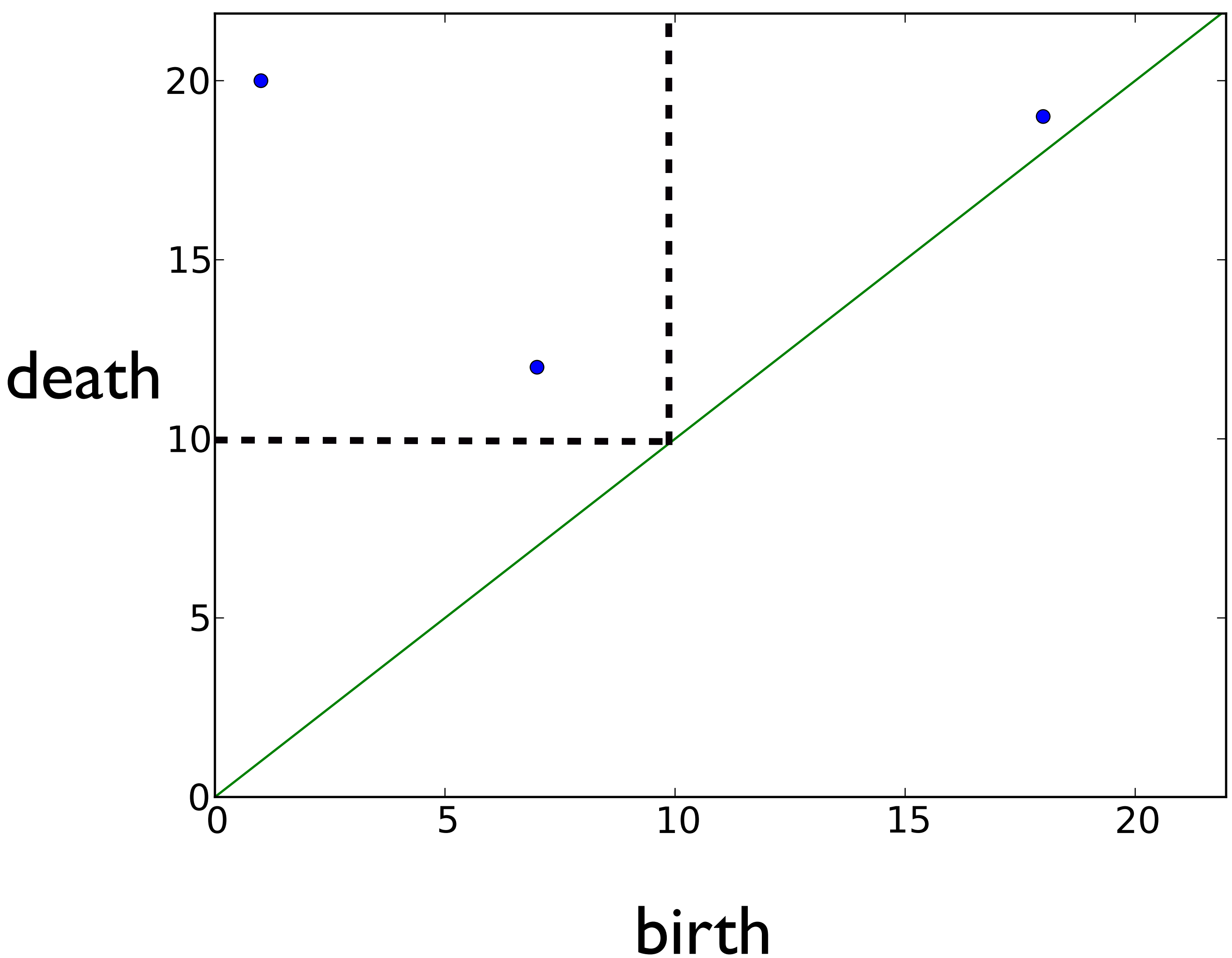


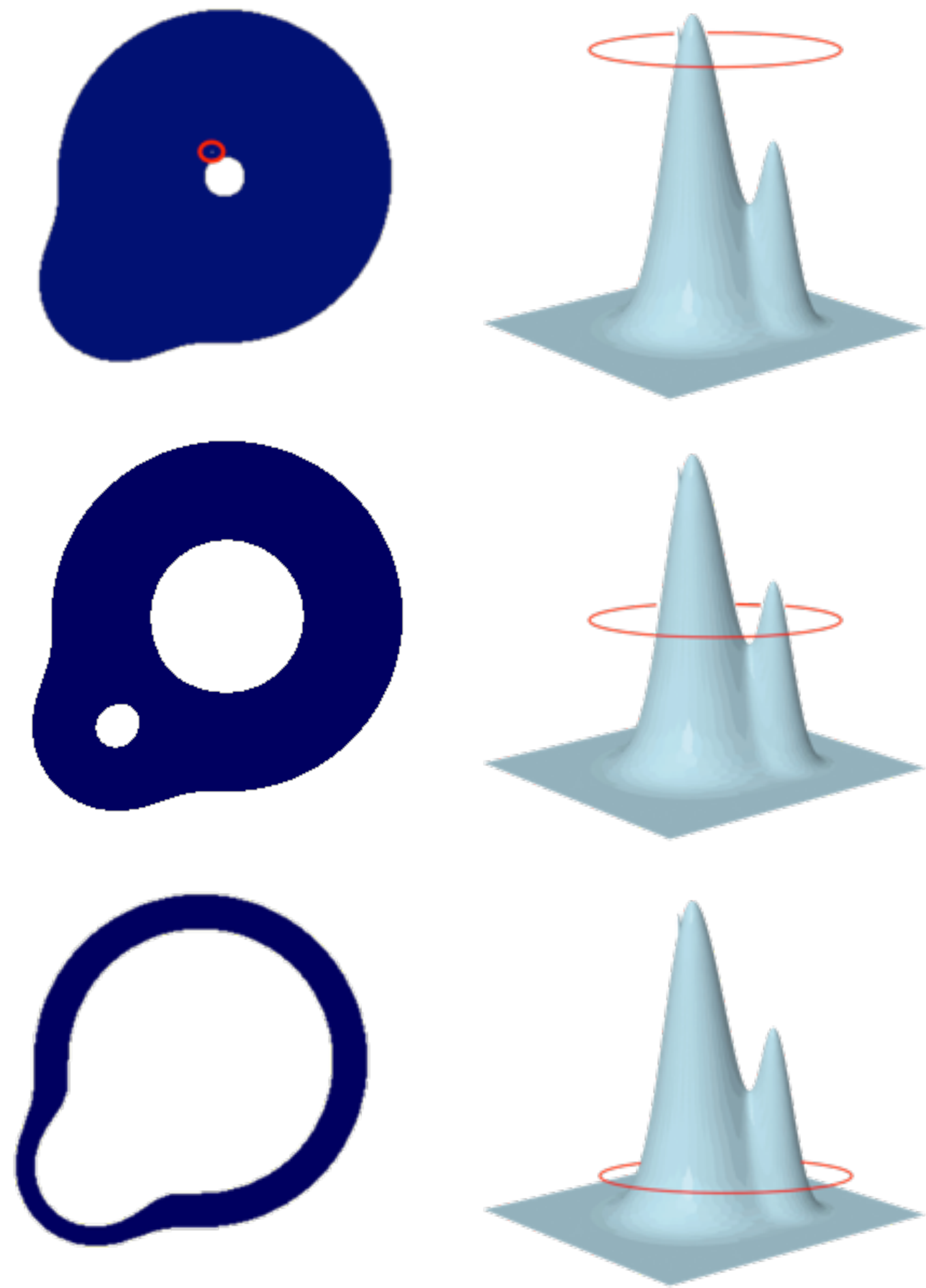




Persistent Homology

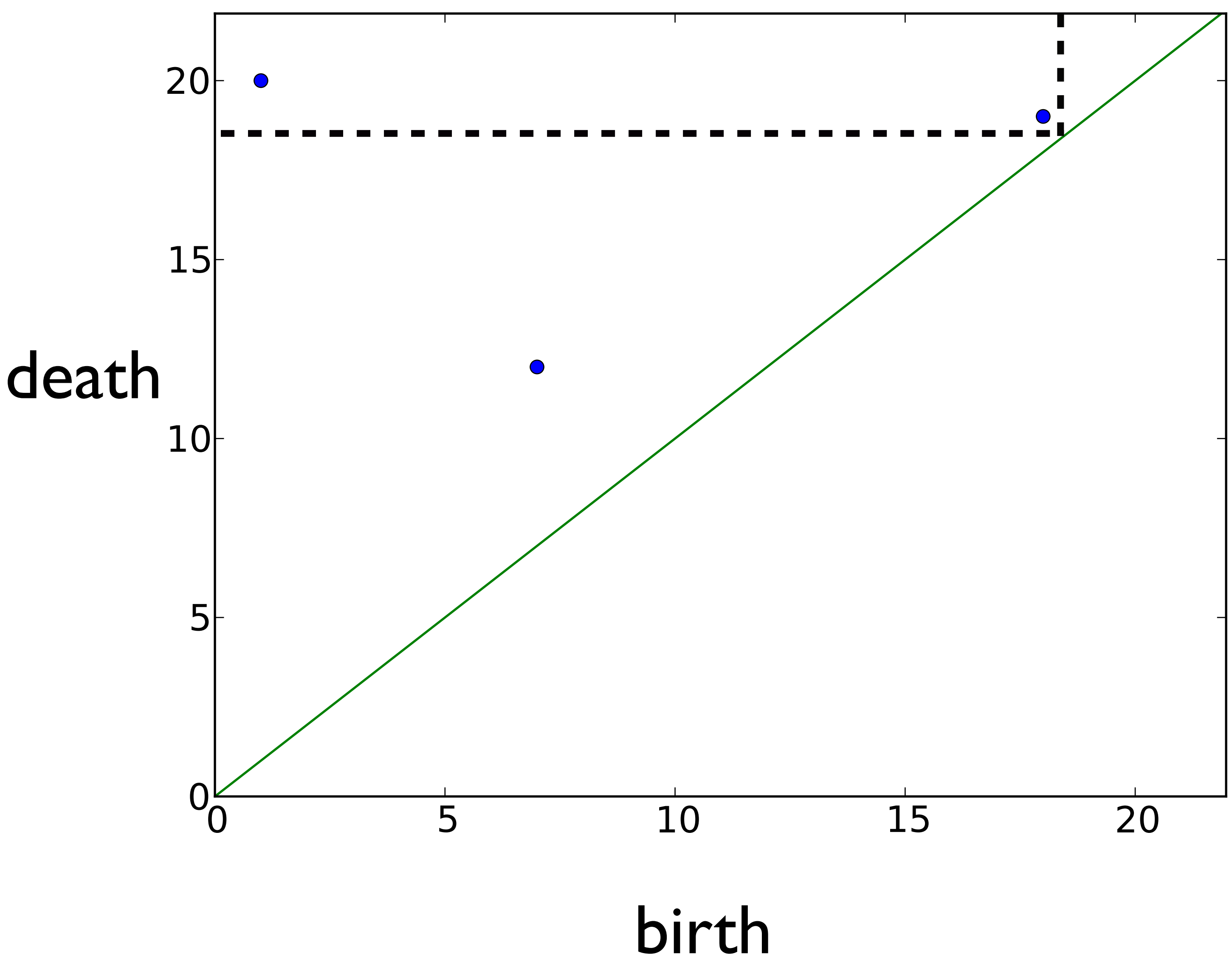
H_1 Persistence Diagram

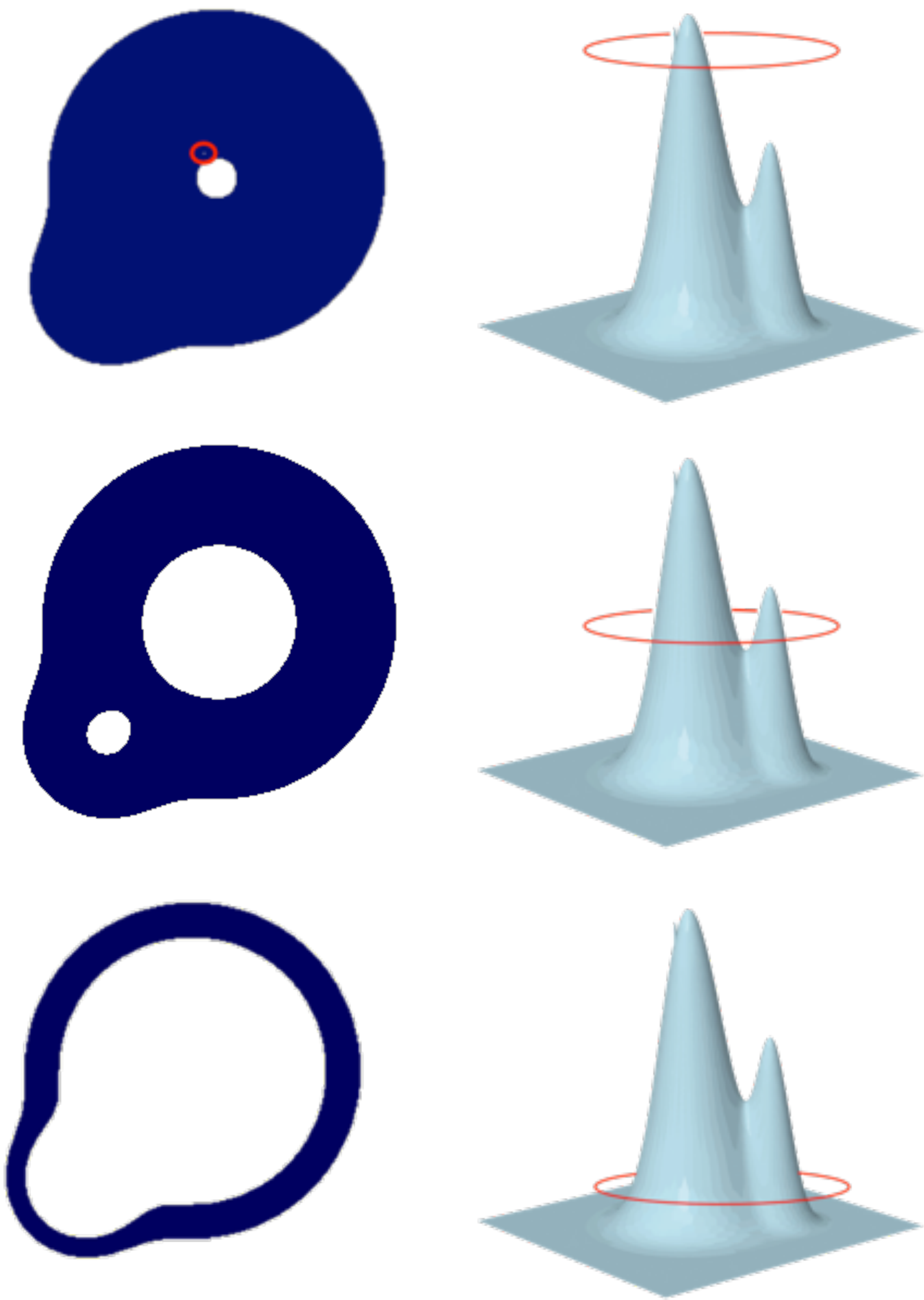




Persistent Homology

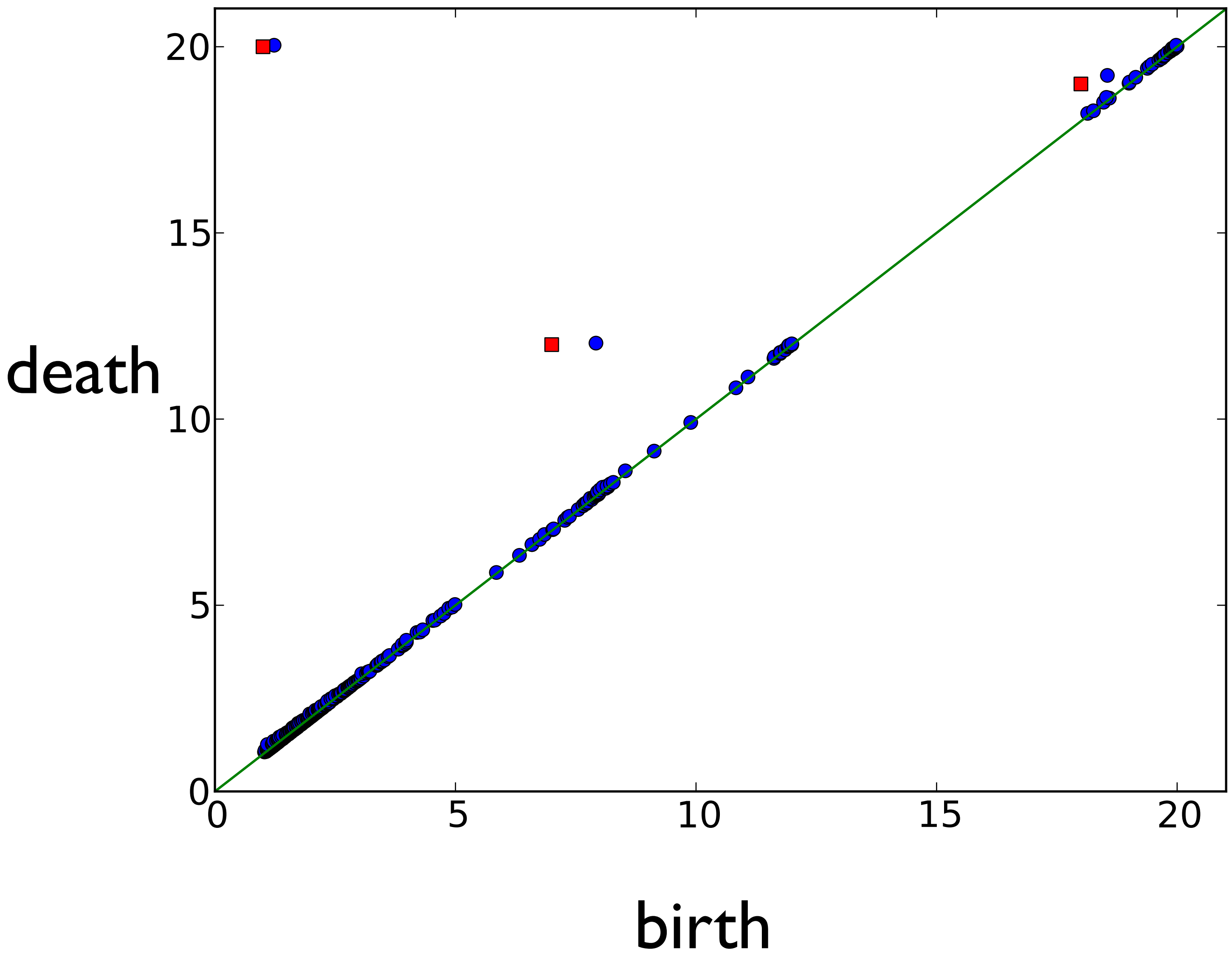
H_1 Persistence Diagram





Adding noise

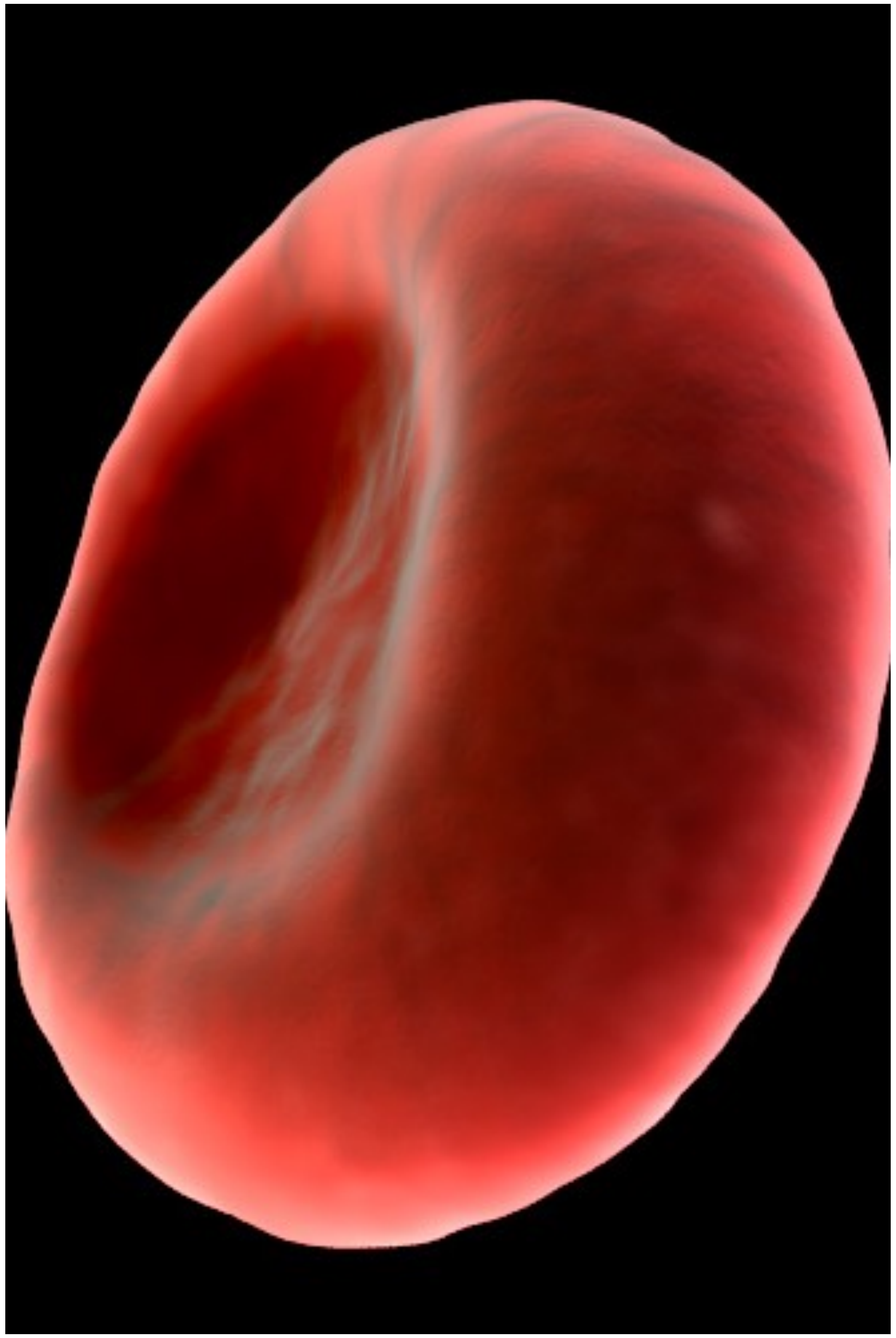
H_1 Persistence Diagram



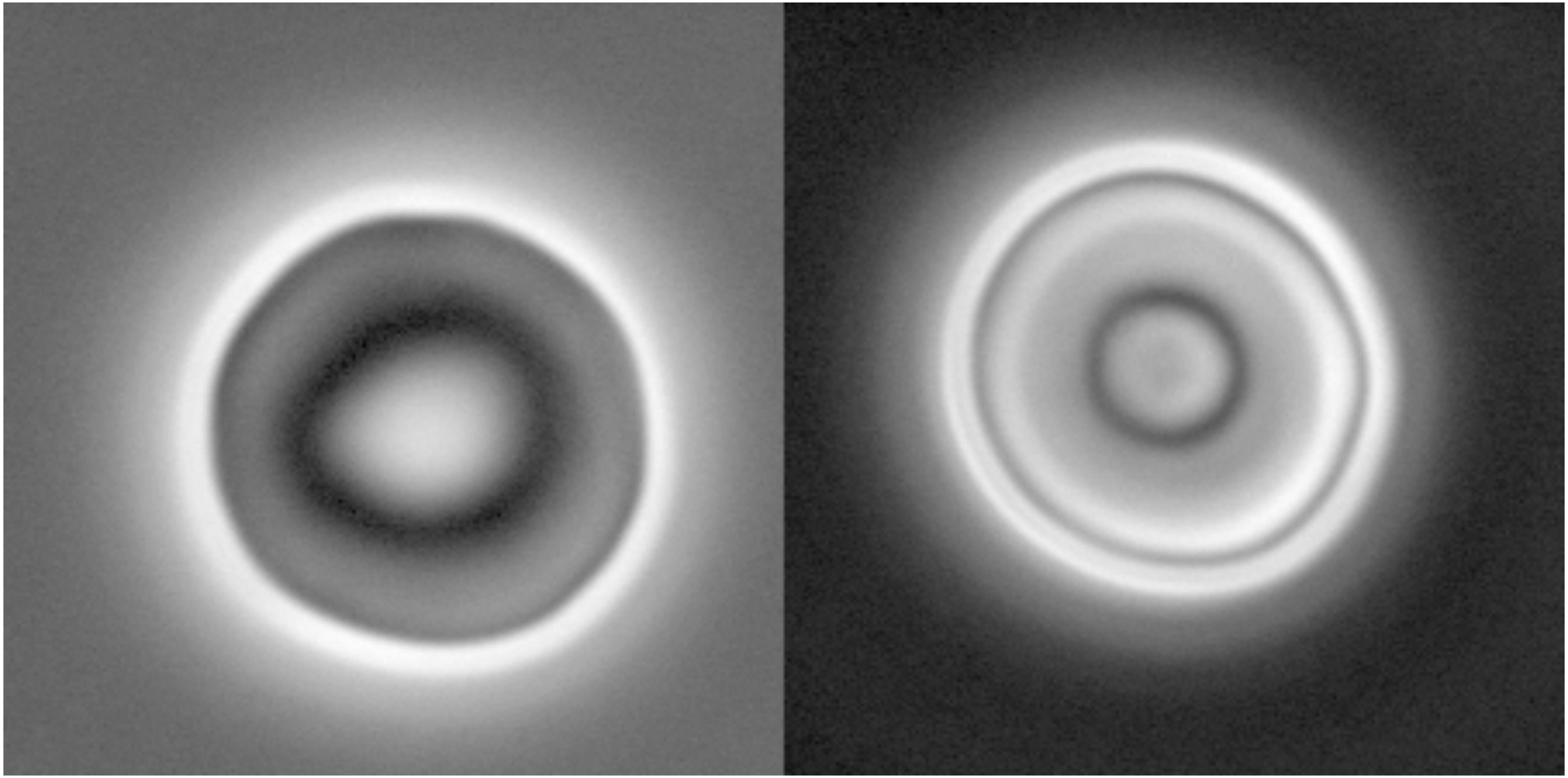
Project II: Red Blood Cells and Flickering

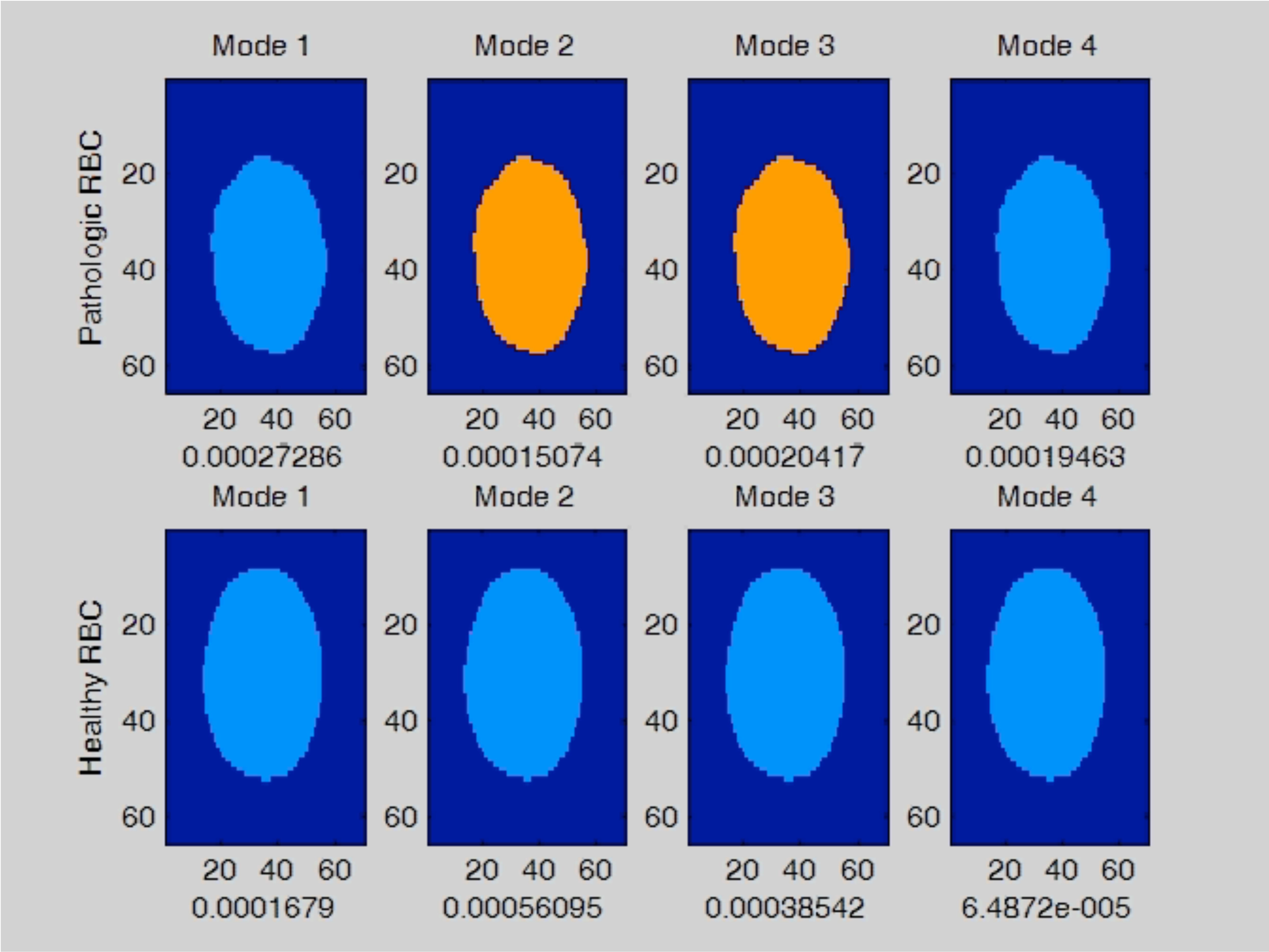
with Jesse Berwald, Kelly Spendlove, Madalena Costa, Ary Goldberger

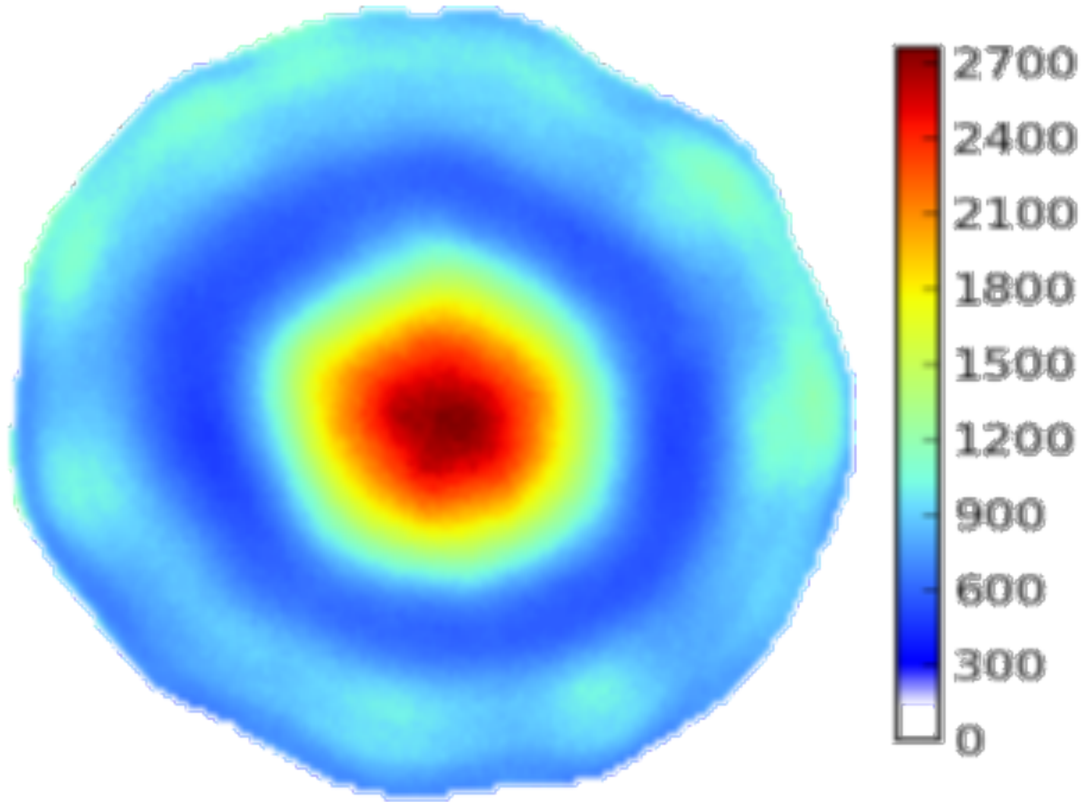
- Red blood cells "flicker".
- As red blood cells age their membranes lose plasticity and this flickering changes (as noted in Costa, et al).
  - Membrane changes have implications for oxygen transport.
  - Blood banks want to know RBC age for this reason.
- We study the change in membrane structure using persistent homology.

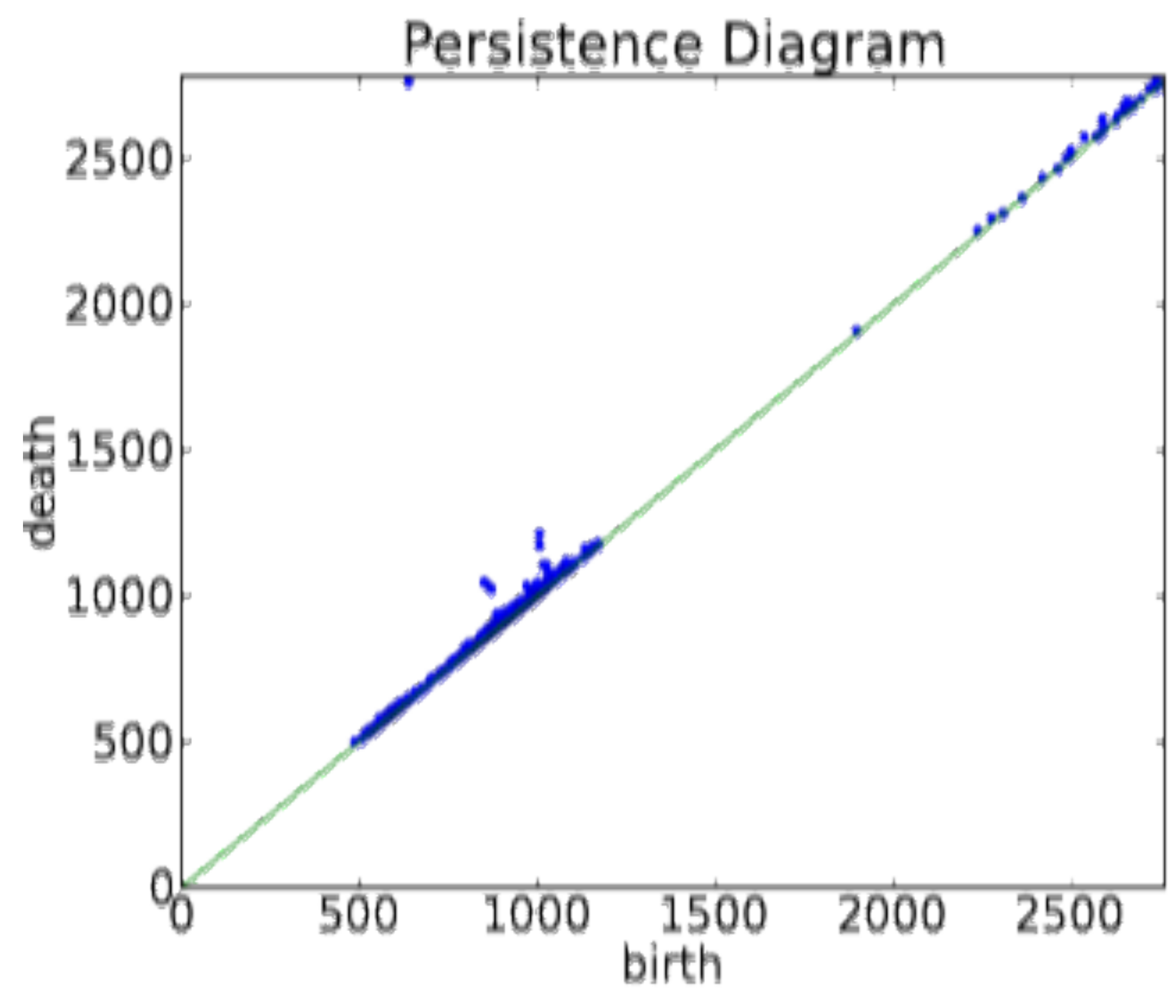


Phase Contrast Microscopy of RBCs

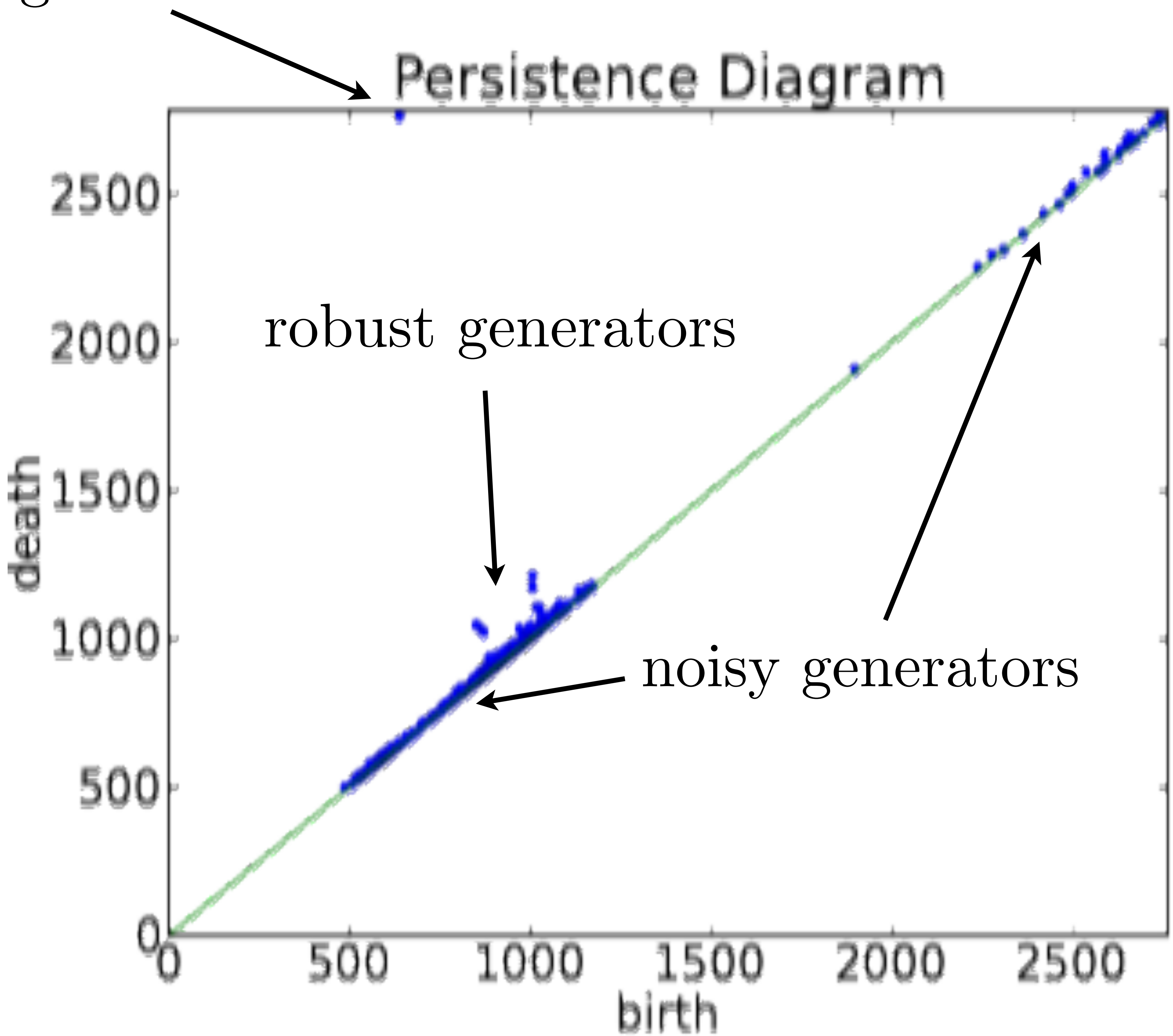


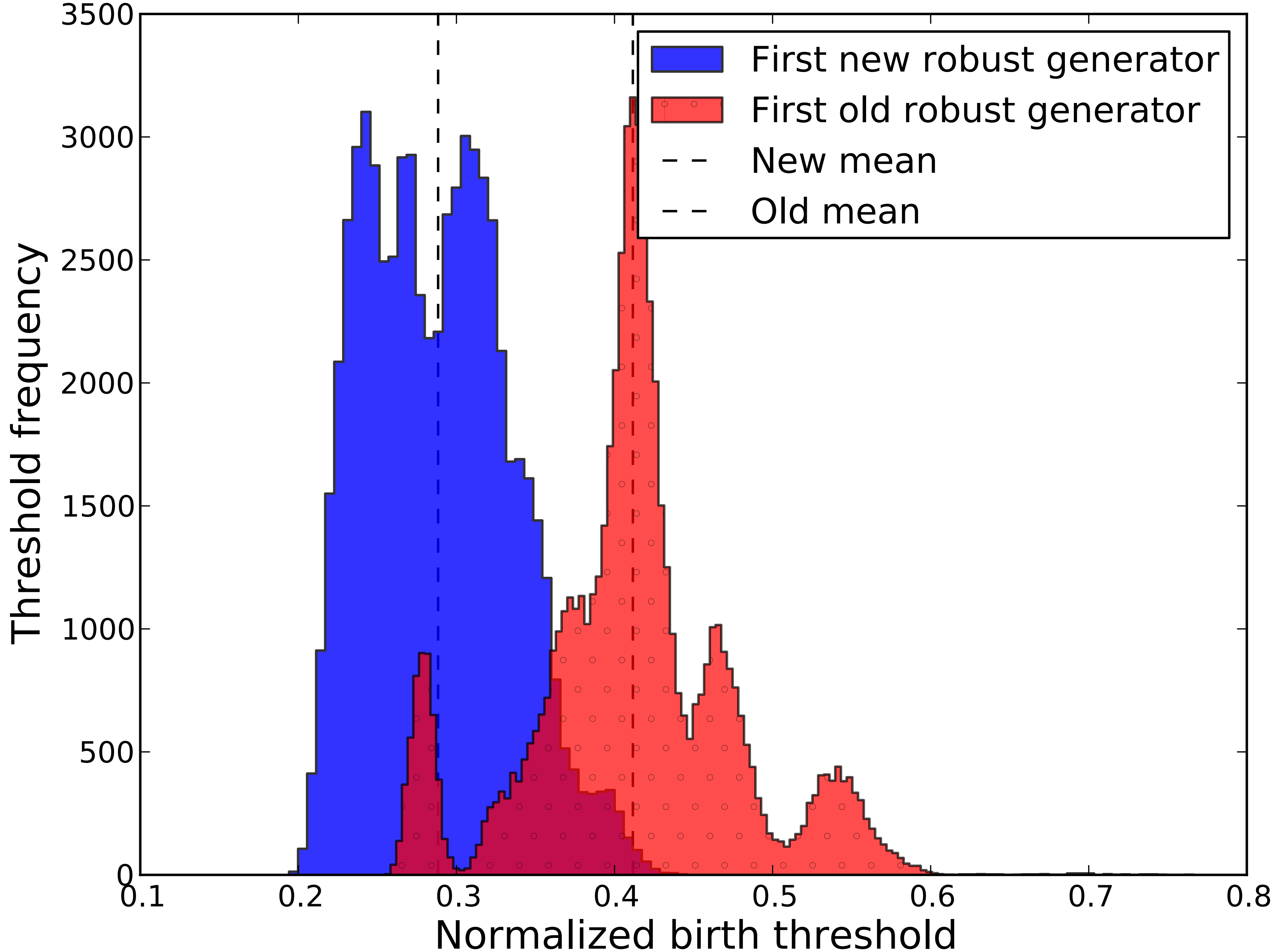

Mode 1
Mode 2
Mode 3
Mode 4
Pathologic RBC
20
40
60
20 40 60
0.00027286
0.00015074
0.00020417
0.00019463
Mode 1
Mode 2
Mode 3
Mode 4
Healthy RBC
20
40
60
20 40 60
0.0001679
0.00056095
0.00038542
6.4872e-005


2700
2400
2100
1800
1500
1200
900
600
300
0

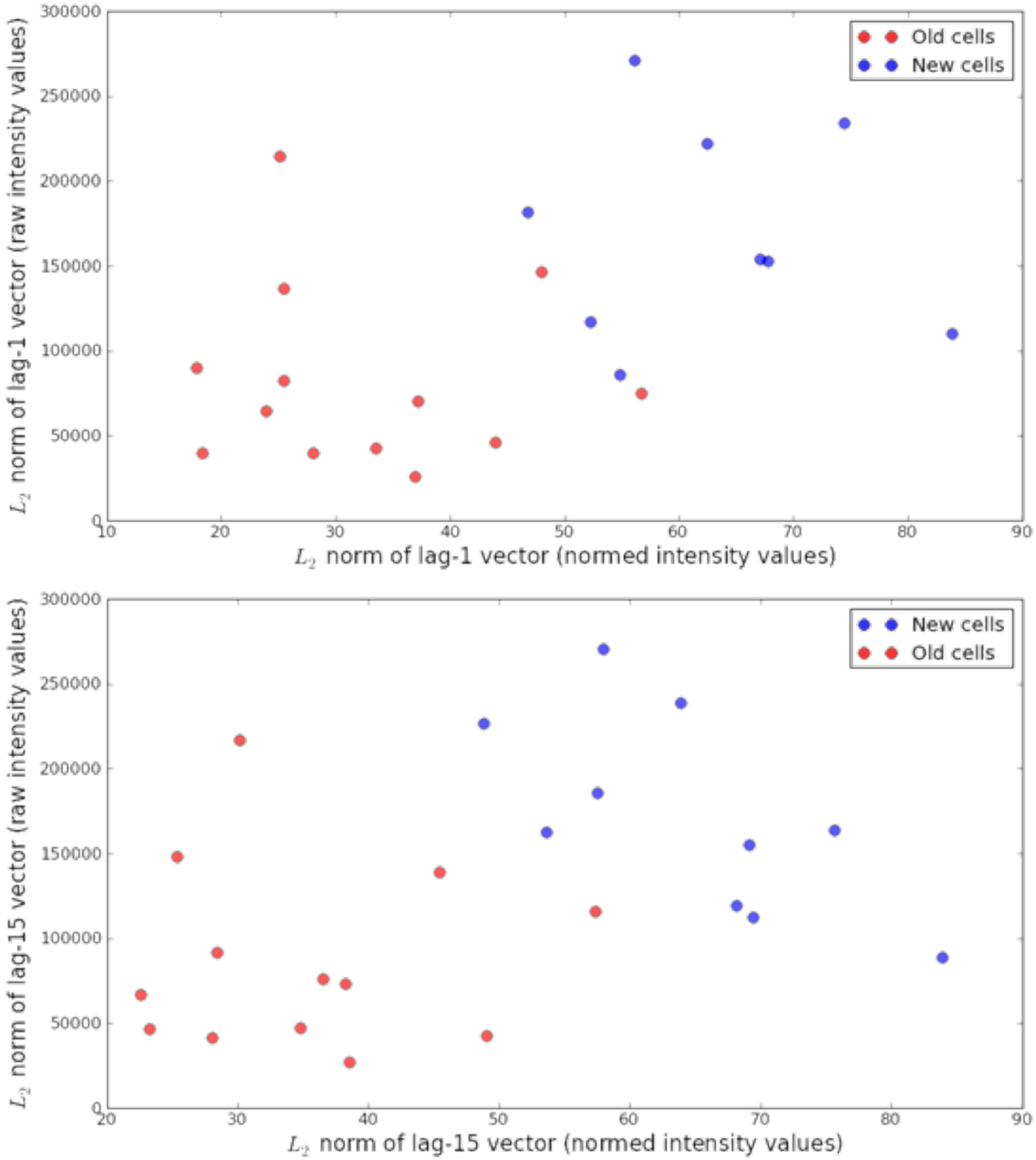

Persistence Diagram
death
birth
0
500
1000
1500
2000
2500

infinite generator

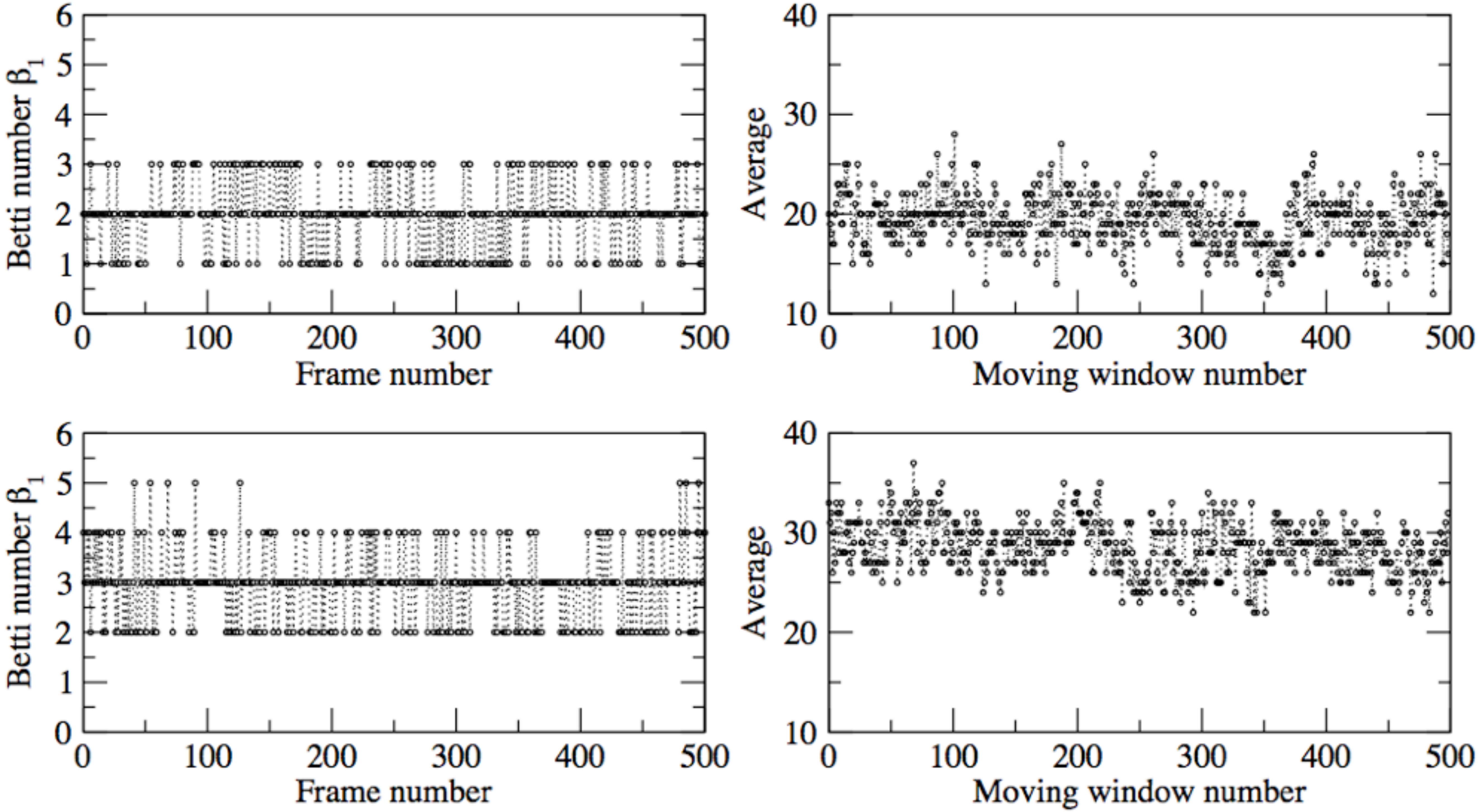




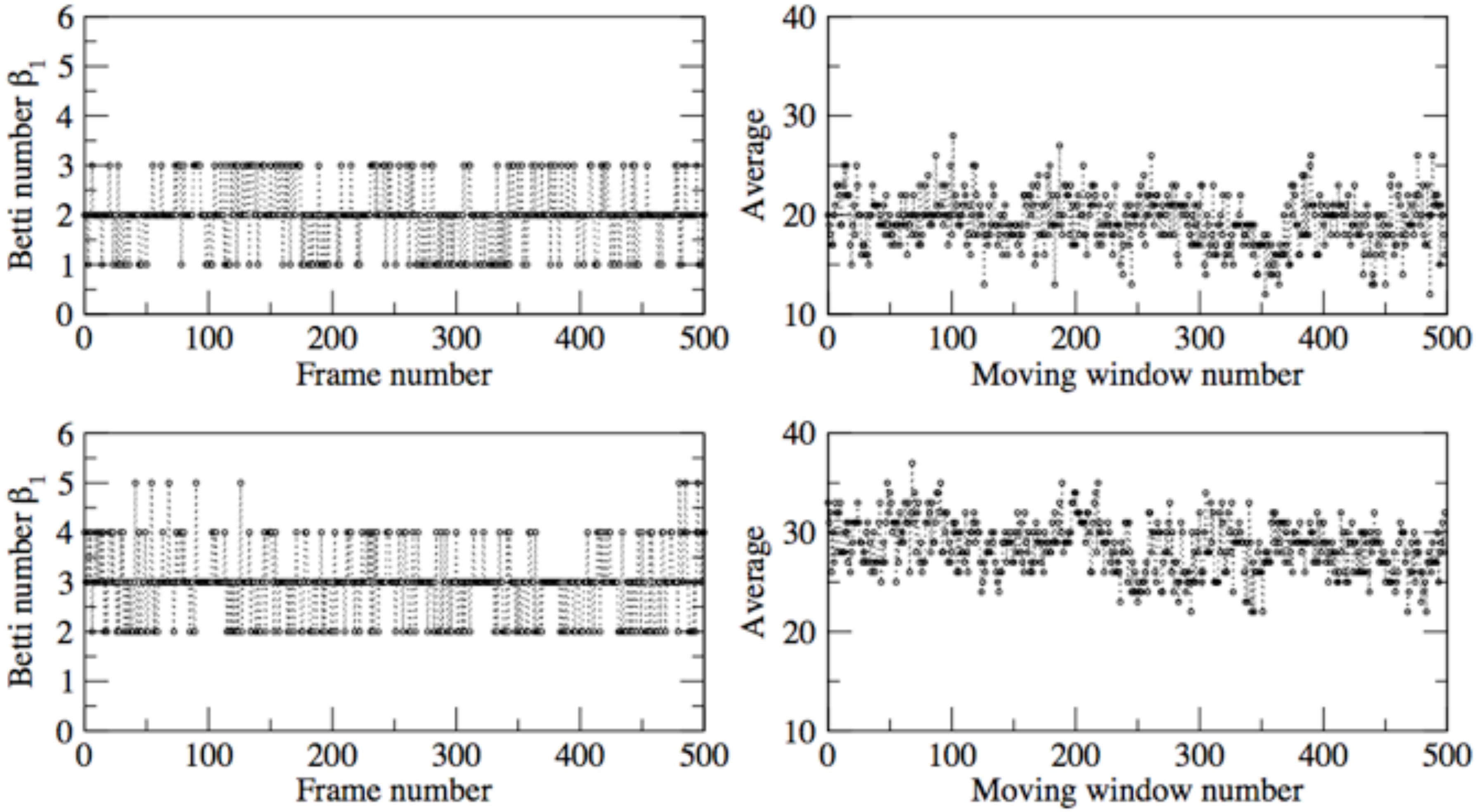
The largest feature of interest represents a deeper (intensity) depression in young cells.



Initial experiments:
Topological features change more rapidly for young cells.

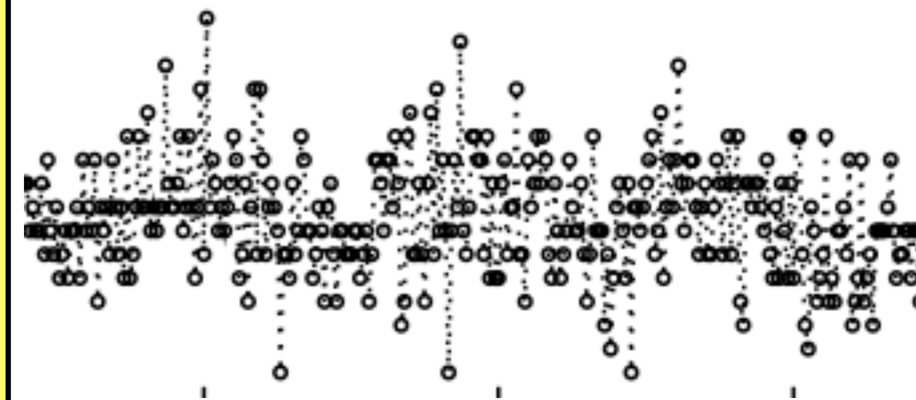


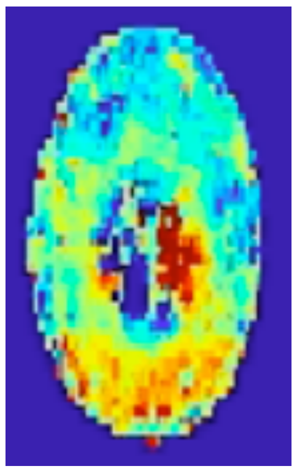
The number of robust generators varies in a more complex way for young cells.



Is there a way to measure dynamics without passing to time series analysis?

Computational topology can extract information from data that is



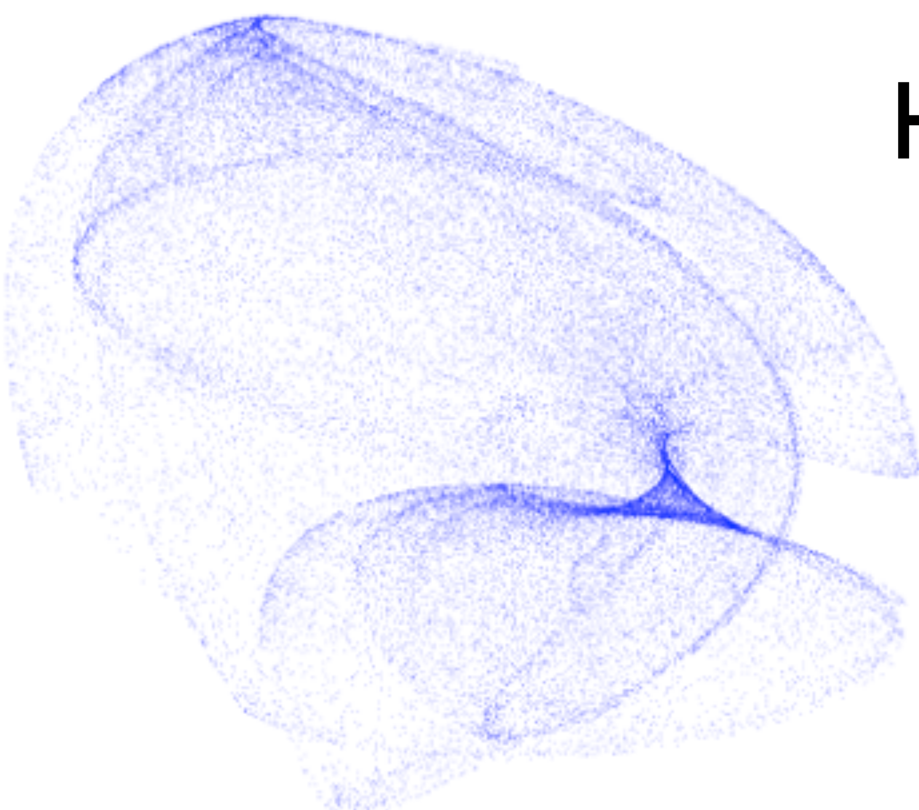


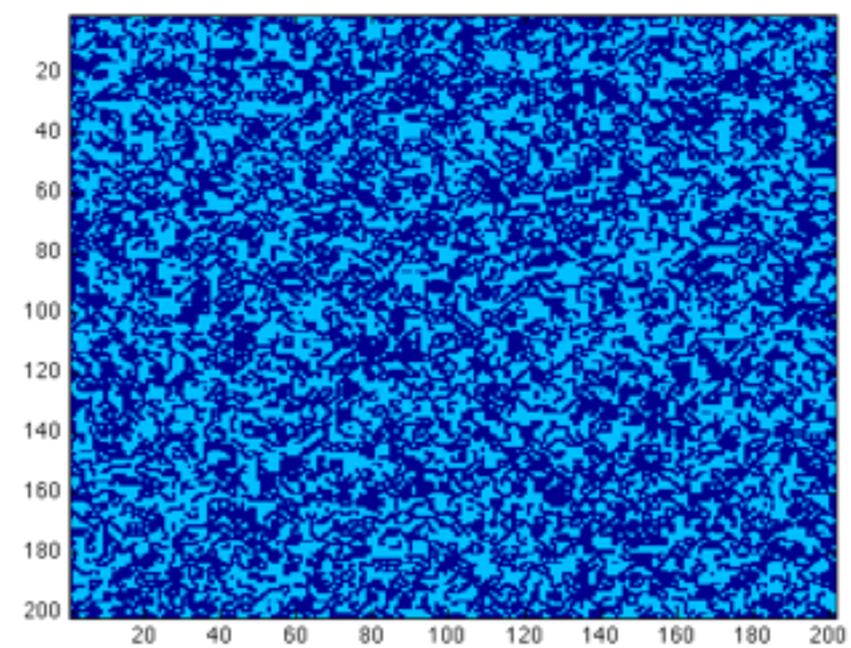
Messy (recurrent dynamics, chaos)

Noisy (measurement error, stochastity)

High dimensional

Sparse





students currently working on projects in dynamics and computational topology:

Martin Salgado-Flores, Liam Bench, Matthew Andriotty

Thank you!