On Topological Tools for Network Analysis

情報セキュリティにおける数学的方法とその実践

Dec 20, 2016

荒井迅
Zin ARAI
Hokkaido University

Contents

Zin Arai

Research fields: Dynamical Systems,
Computational Topology

A. Topological Tools for Sensor Network

Distributed homology computation via Mayer-Vietoris sequences and the graph Laplacian

B. Graph Theoretical Tools for Dynamical Systems

Discrete Hodge decomposition (HodgeRank) for the study of gradient structure of dynamical systems

Sensor Networks

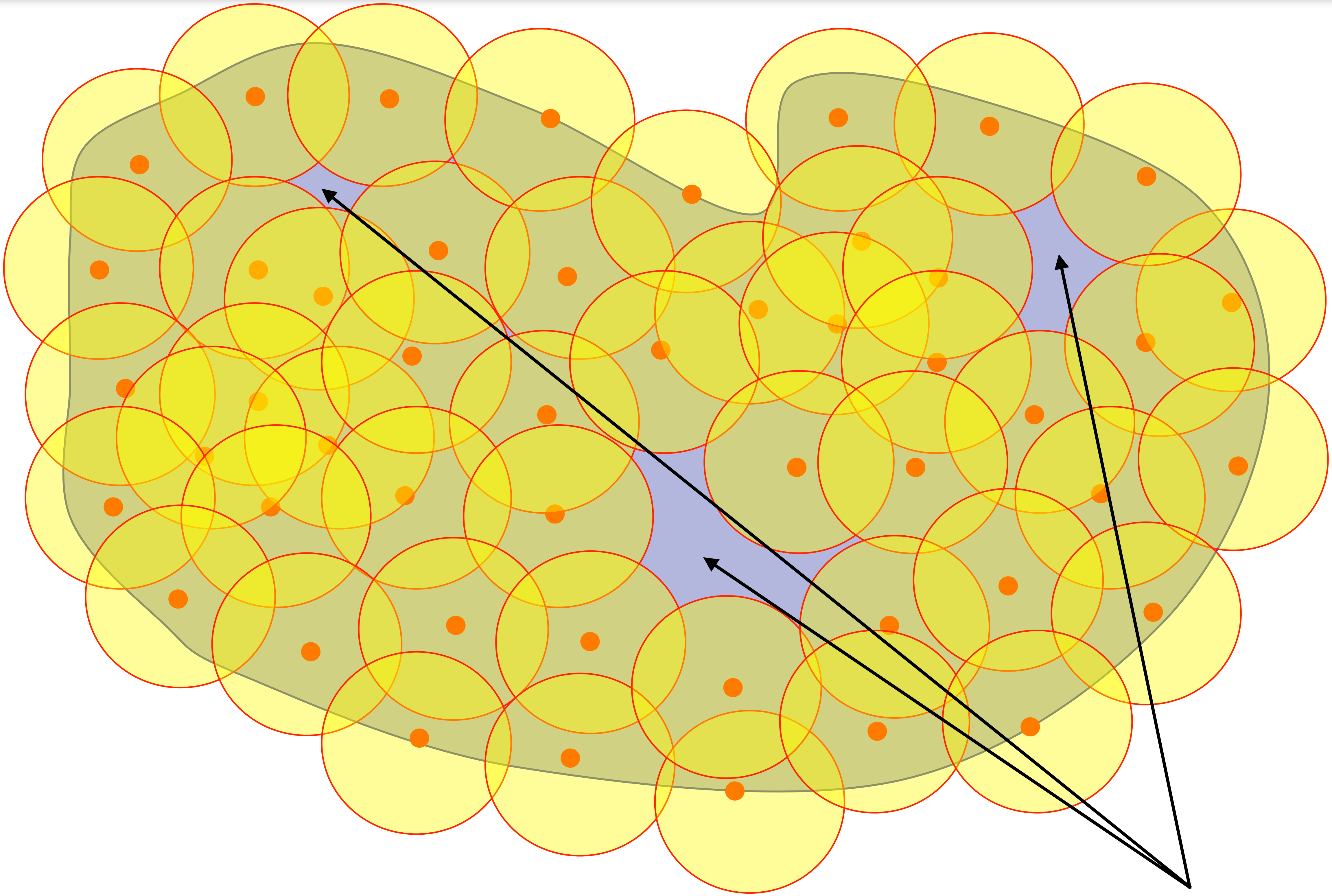
Ubiquitous Sensors


Sensor network consists of spatially distributed sensors to monitor physical or environmental conditions.

Sensors today: Single, powerful sensor

Future: many, cheap sensors working cooperatively

Coverage Problem



We want to guarantee the non-existence of "holes" in the coverage region.

Conventional Approaches

Computational Geometry

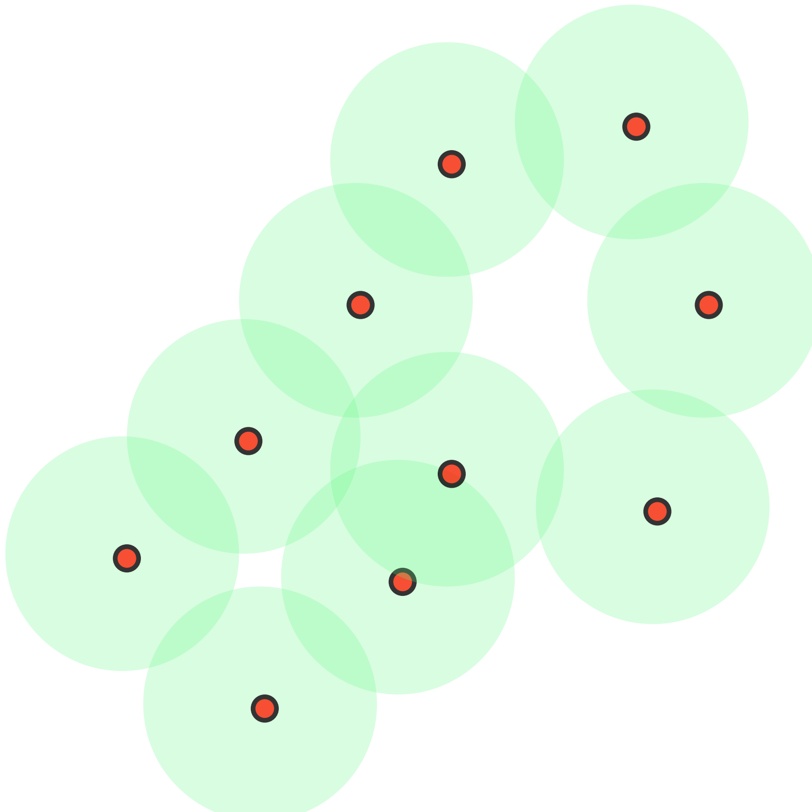
Need to know the absolute position of sensors.
We can use GPS, etc, to know the position,
but this is battery consuming.

Probability theoretical approach

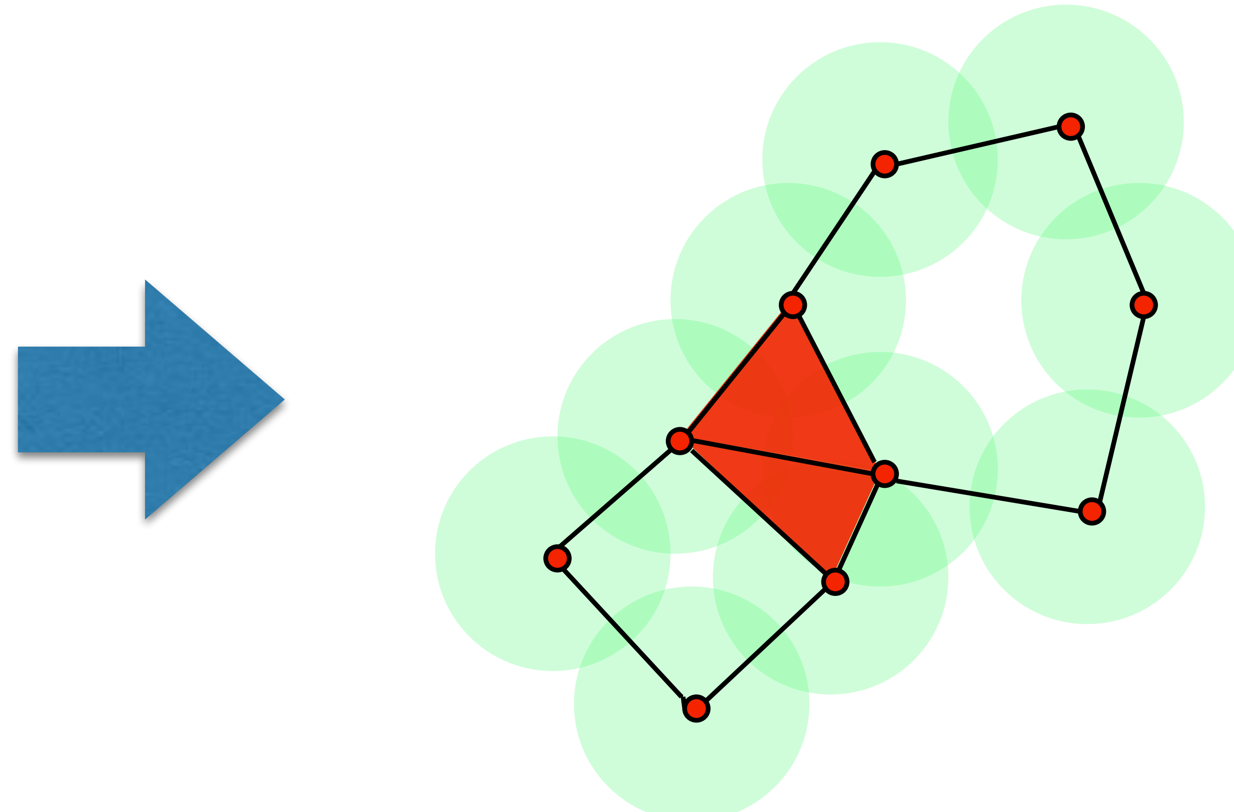
Compute the probability of coverage failures.

Need to assume that sensors are distributed
uniformly, which is too strong in practice.

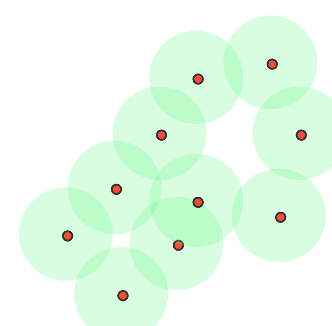
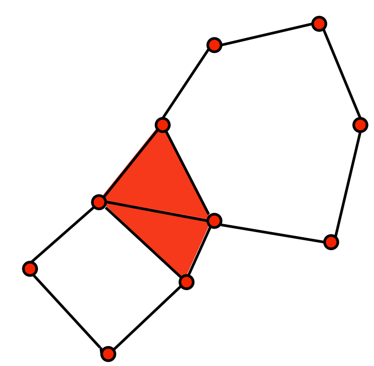
Algebro-Topological Approach



Network Data



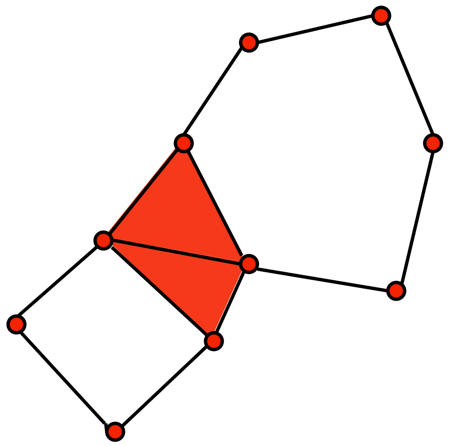
Polyhedron (Čech Complex)

"Nerve Theorem" tells us that  and  are topologically the same.

We can easily count the number of holes of the latter with the help of linear algebra, aka "Homology theory".

Rapid Course in Homology (1/2)

K



vertices: $v_1, v_2, \ldots, v_k$

edges: $e_1, e_2, \ldots, e_\ell$

faces: $f_1, f_2, \ldots, f_m$

Chain Groups

$$C_0(K) := \{\lambda_1 v_1 + \cdots + \lambda_k v_k \mid \lambda_i \in \mathbb{Z}\} \cong \mathbb{Z}^k$$

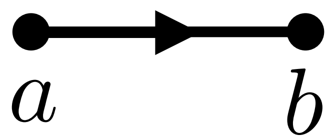
$$C_1(K) := \{\lambda_1 e_1 + \cdots + \lambda_\ell e_\ell \mid \lambda_i \in \mathbb{Z}\} \cong \mathbb{Z}^\ell$$

$$C_2(K) := \{\lambda_1 e_1 + \cdots + \lambda_m e_m \mid \lambda_i \in \mathbb{Z}\} \cong \mathbb{Z}^m$$

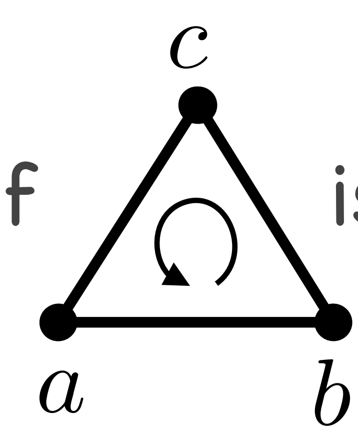
Boundary map (given arbitrary orientations to edges and faces)

$$\partial_n : C_n(K) \to C_{n-1}(K)$$

The boundary of $\vec{ab} =$



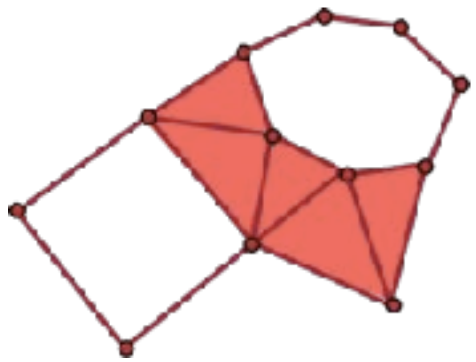
is $b - a$, of



is $\vec{ab} + \vec{bc} + \vec{ca}$

Rapid Course in Homology (2/2)

K



vertices: $v_1, v_2, \ldots, v_k$

edges: $e_1, e_2, \ldots, e_\ell$

faces: $f_1, f_2, \ldots, f_m$

We can show that **"the boundary of a boundary is empty"**, that is,

$$\partial_n \circ \partial_{n+1} = 0.$$

Therefore, we have $\mathrm{im}\, \partial_{n+1} \subset \ker \partial_n$ and define the n-th homology group by

$$H_n(K) := \ker \partial_n / \mathrm{im}\, \partial_{n+1}.$$

To compute homology groups, we need the Smith normal forms of the boundary maps.

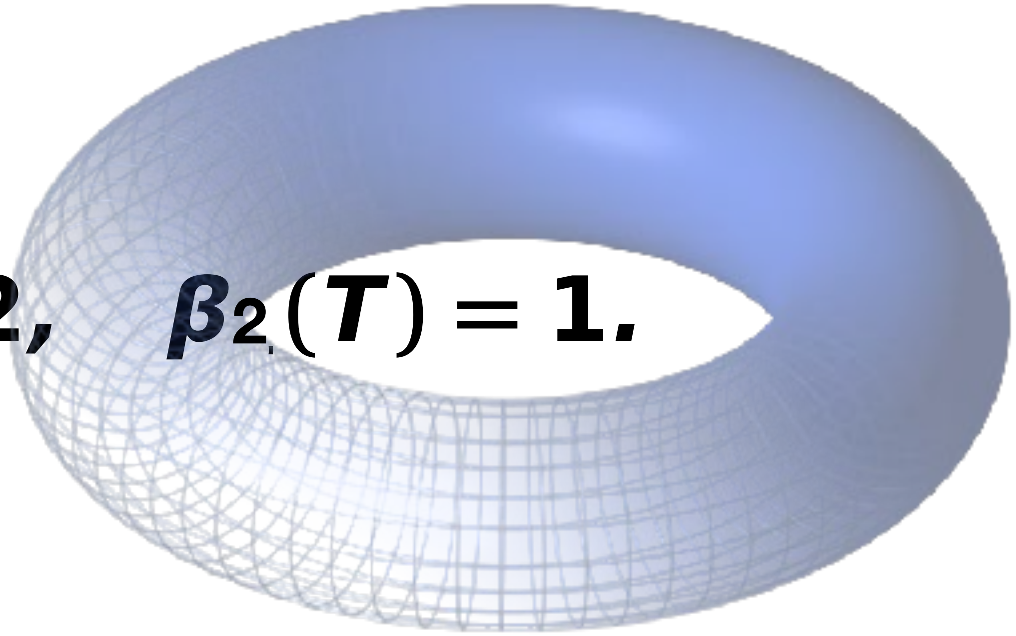
Betti Numbers

The betti number is the rank of the homology group.
That is, $\boldsymbol{\beta_i(X) = \text{rank}\, H_i(X)}$

Informally, the betti number $\boldsymbol{\beta_i(X)}$ of a space $\boldsymbol{X}$ counts the number of independent $\boldsymbol{i}$-dimensional objects in $\boldsymbol{X}$.
For example,

- $\boldsymbol{\beta_0(X)}$ is the number of connected components
- $\boldsymbol{\beta_1(X)}$ is the number of two-dimensional or ”circular” holes
- $\boldsymbol{\beta_2(X)}$ is the number of three-dimensional holes or ”voids”

In the case of the torus $\boldsymbol{T}$,



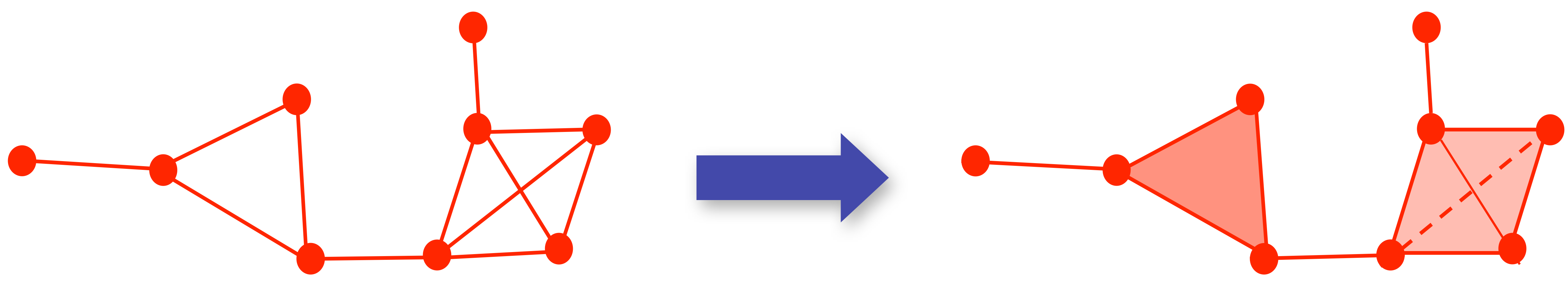
$$\boldsymbol{\beta_0(T) = 1, \quad \beta_1(T) = 2, \quad \beta_2(T) = 1.}$$

Rips Complex

Unfortunately, it's computationally very hard to build the Cech complex.

Therefore we introduce the Rips complex $\mathcal{R}$:

- 0-simplexes: sensor nodes $x_1, x_2, \ldots, x_N$
- $|x_{i_0}, x_{i_1}, \ldots, x_{i_k}|$ is a k-simplex $\Leftrightarrow$ Every pair of Nodes in $\{x_{i_0}, x_{i_1}, \ldots, x_{i_k}\}$ are connected



Note that $\mathcal{R}$ is determined by the underlying graph.

Problem Setting

A1
|
A4

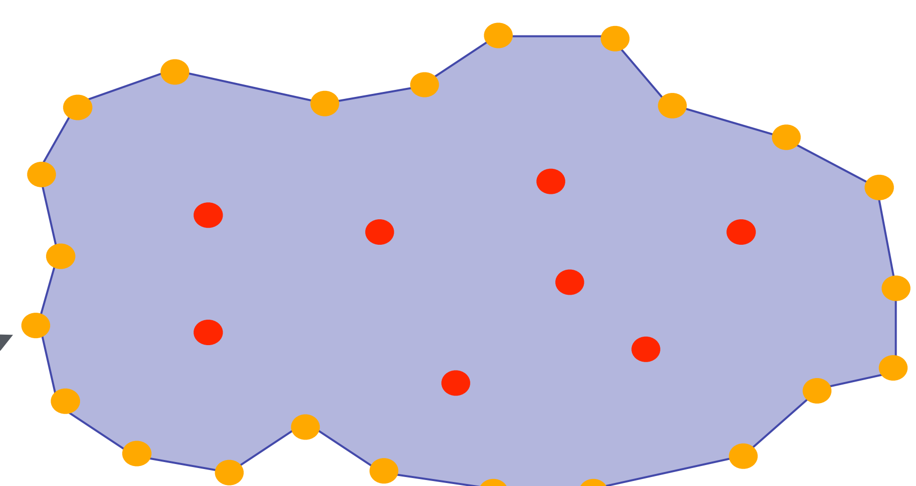
- $D \subset \mathbb{R}^2$ is compact and connected.
- The boundary ∂D is connected.
- $\mathcal{X} := \{x_i \in D \mid i = 1, \cdots, N\}$ sensors
- $\mathcal{X}_f$: sensors on the boundary ∂D, we call th n “fence nodes”
- r_b : a sensor can communicate with another s or within the distance r_b
- r_c : the radius of monitoring region.

We need this since the Rips complex is not a nerve

We assume $r_c \geq \frac{r_b}{\sqrt{3}}$

$\mathcal{U}$: the union of all monitoring region.

Problem: $D \subset \mathcal{U}$?



fence nodes

Sufficient Conditions

$\mathcal{F}$: the Rips complex of the "fence" nodes.

Theorem [Robert Ghrist]

If there exists $[\sigma] \in H_2(\mathcal{R}, \mathcal{F})$ such that $\delta_2[\sigma] \neq 0 \in H_1(\mathcal{F})$ then $D \subset \mathcal{U}$.

We can simplify this condition as follows:

Theorem [ZA-Hayashi-Hiraoka]

If $H_1(\mathcal{R}) = 0$ then $D \subset \mathcal{U}$.

Distributed Homology Computation

Practical Problems

Who compute the homology?

We do not want to set a central node in the network, for security reasons and robustness.

Homology computation is slow:

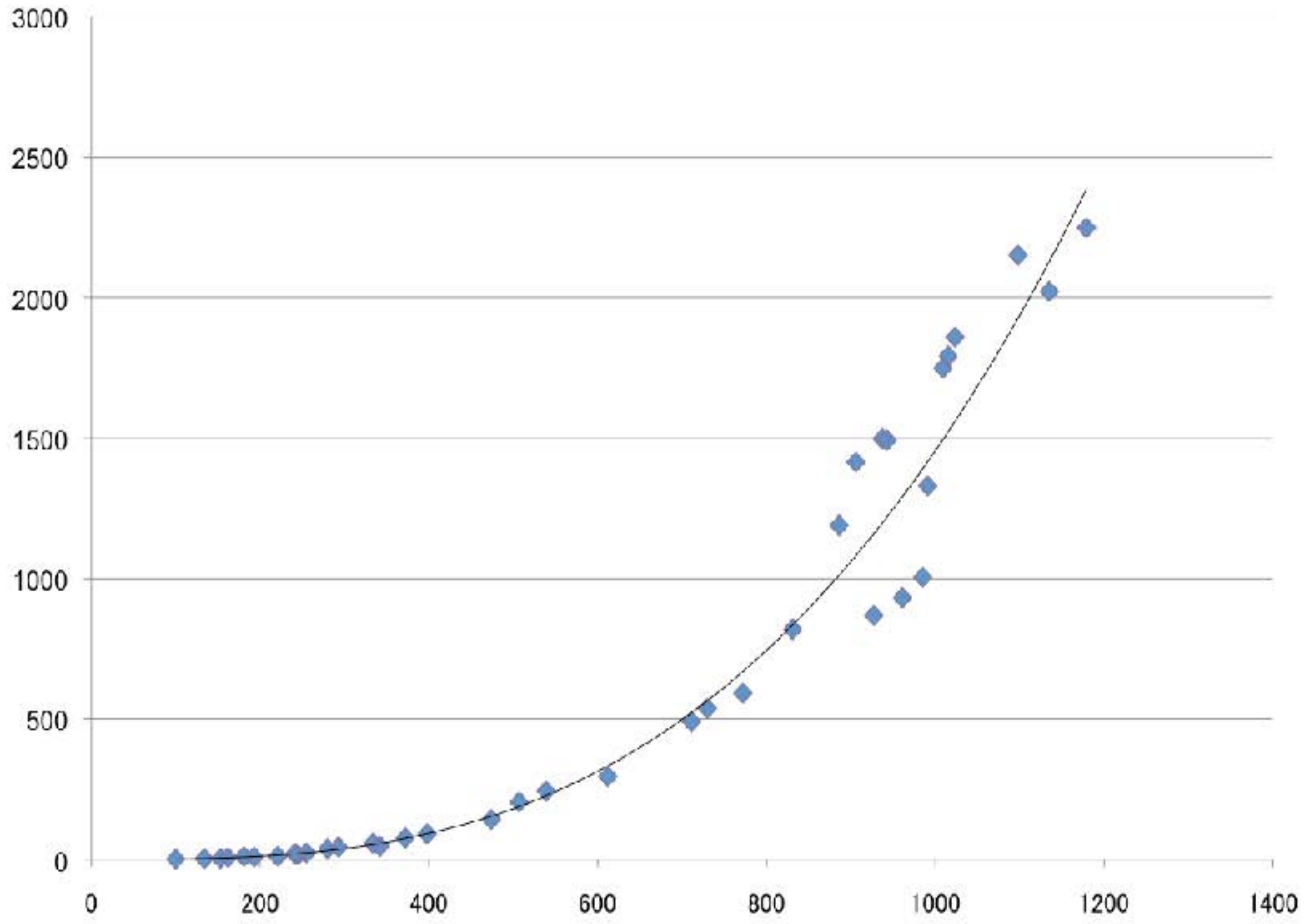
Homology computation is $O(n^3)$, therefore requires a lot of CPU power when the network is large.

We want to distribute the homology computation!

Idea 1: Mayer-Vietoris sequence (ZA-Hayashi-Hiraoka)

Idea 2: Discrete Laplacian and the gossip algorithm

Computational Cost

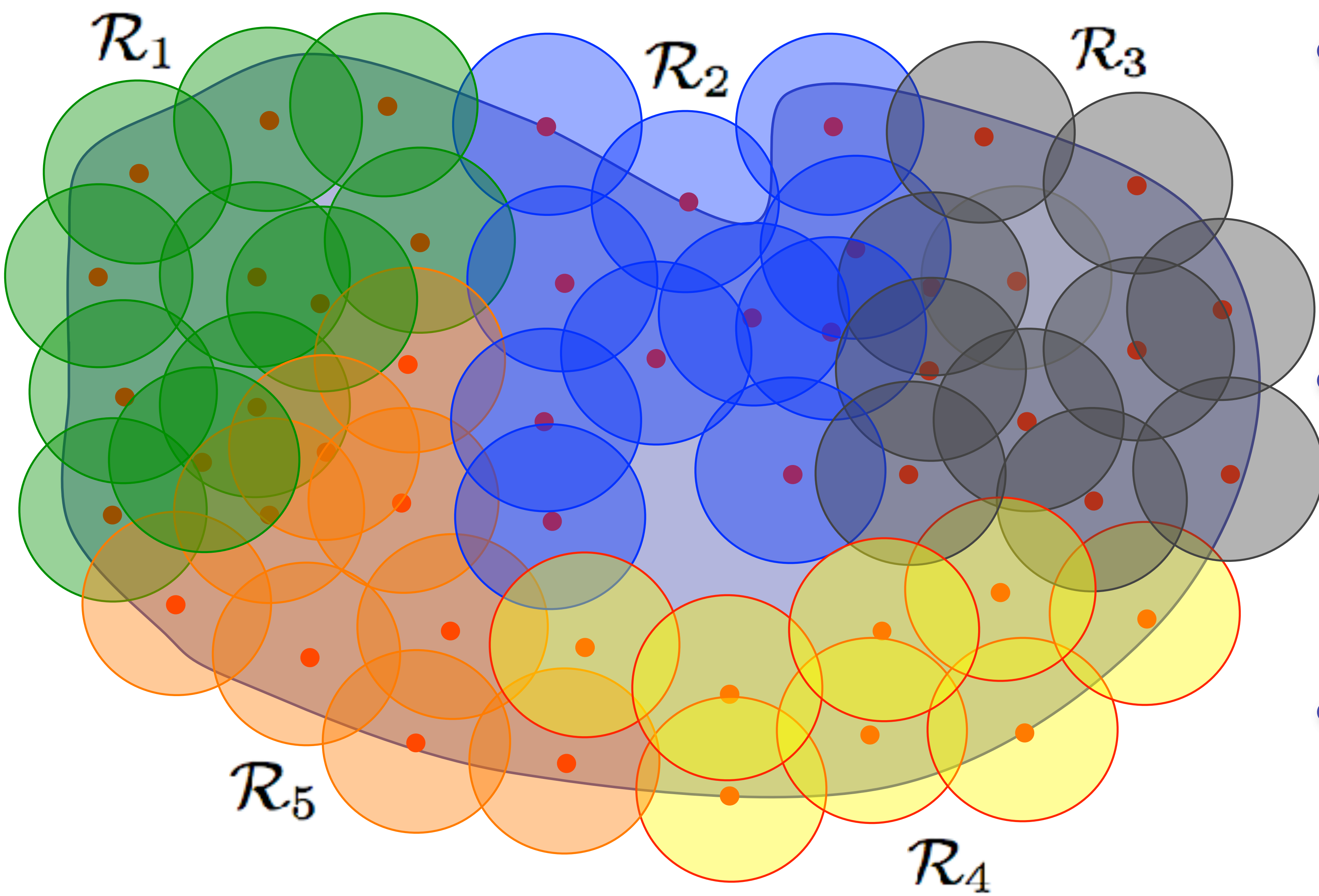


Mayer-Vietoris Sequence

The Mayer-Vietoris long exact sequence:

$$\rightarrow H_k(\mathcal{R}_1 \cap \mathcal{R}_2) \rightarrow H_k(\mathcal{R}_1) \oplus H_k(\mathcal{R}_2) \rightarrow H_k(\mathcal{R}_1 \cup \mathcal{R}_2) \rightarrow H_{k-1}(\mathcal{R}_1 \cap \mathcal{R}_2) \rightarrow$$

Local information $\Rightarrow$ Global Information



- 1. Decompose $\mathcal{R}$

$$\mathcal{R} = \bigcup_{i=1}^{K} \mathcal{R}_i$$

- 2. Compute

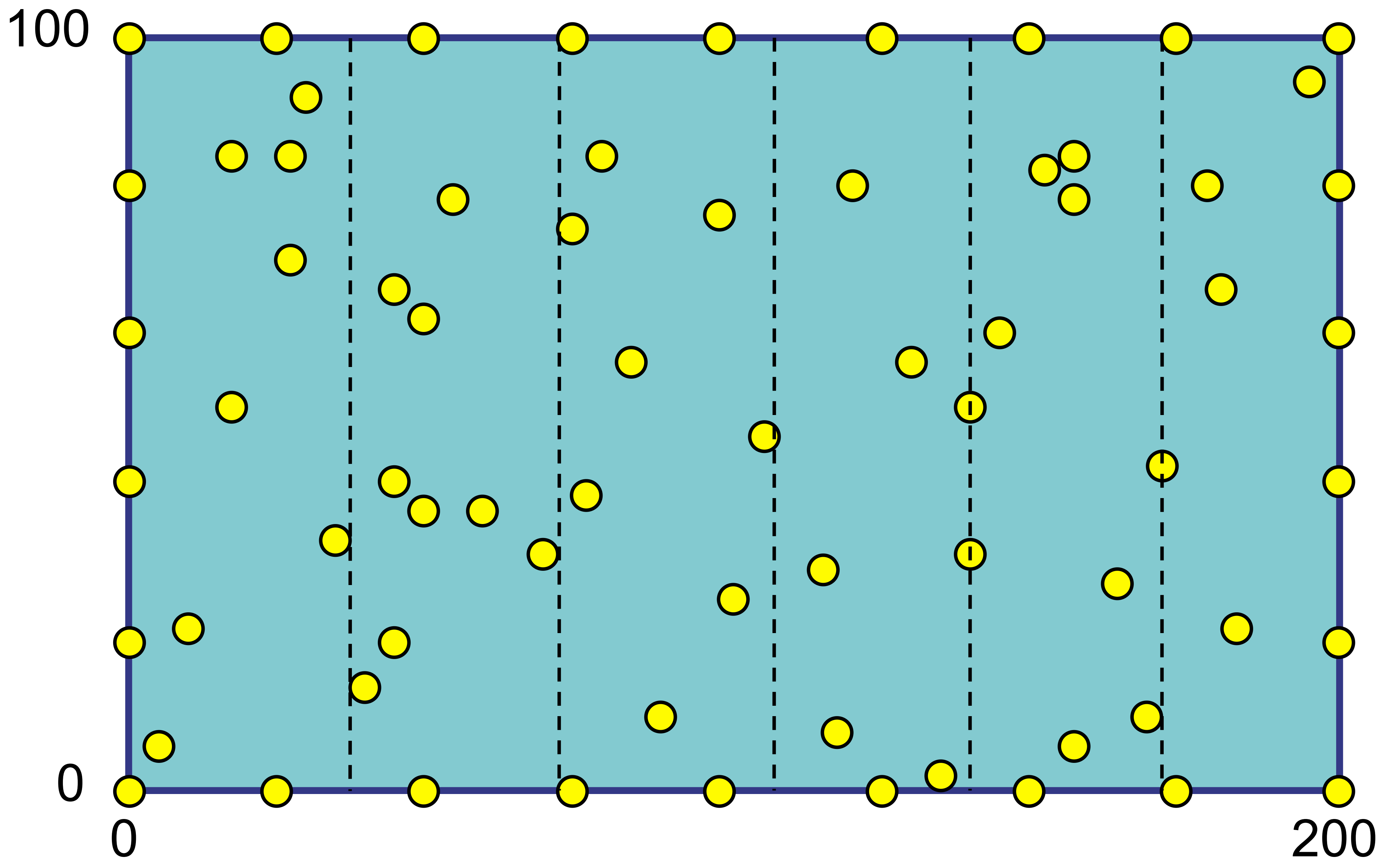
$H_*(\mathcal{R}_i), H_*(\mathcal{R}_i \cap \mathcal{R}_j)$

in parallel.

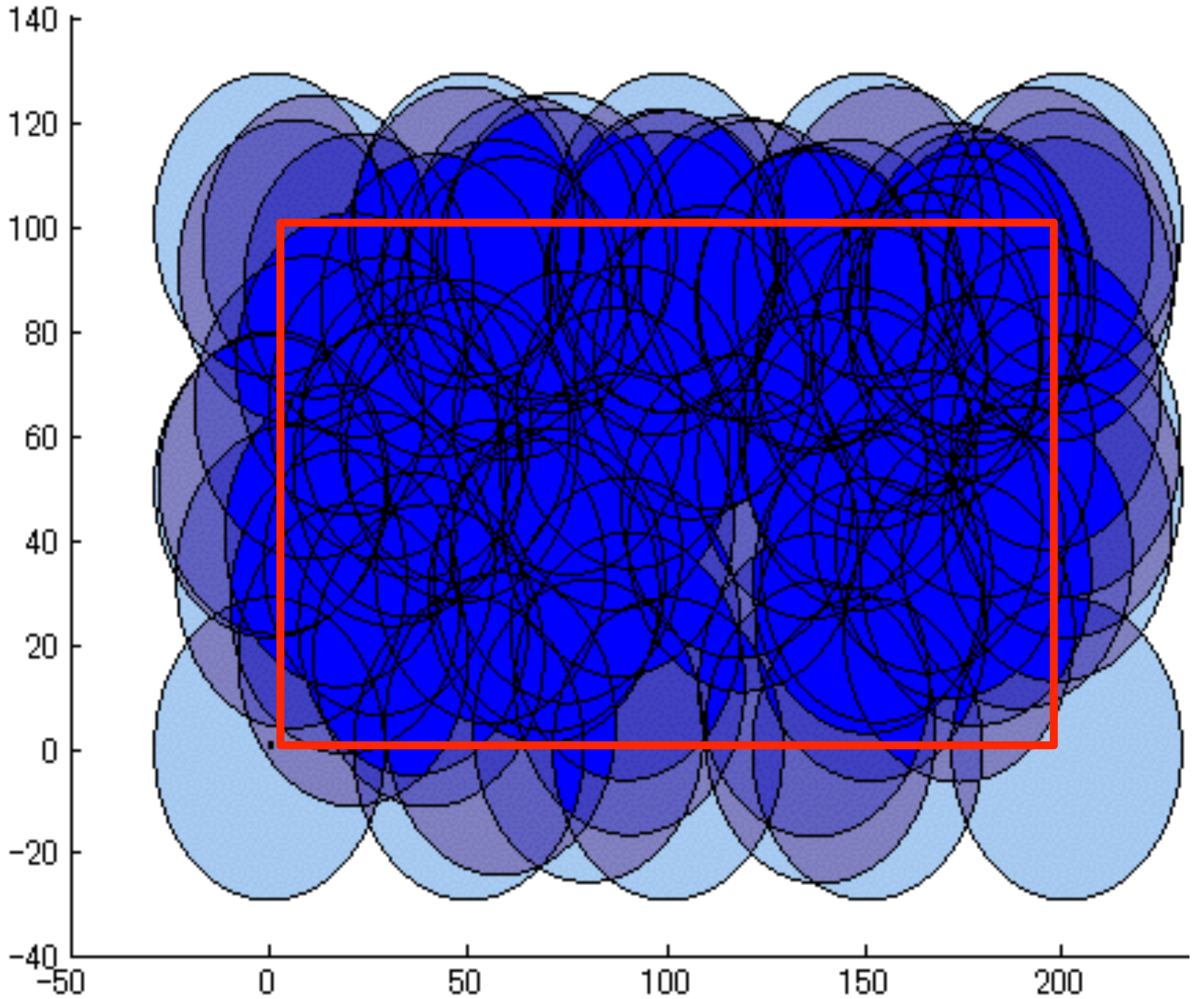
- Use Mayer-Vietoris to compute $H_*(\mathcal{R})$

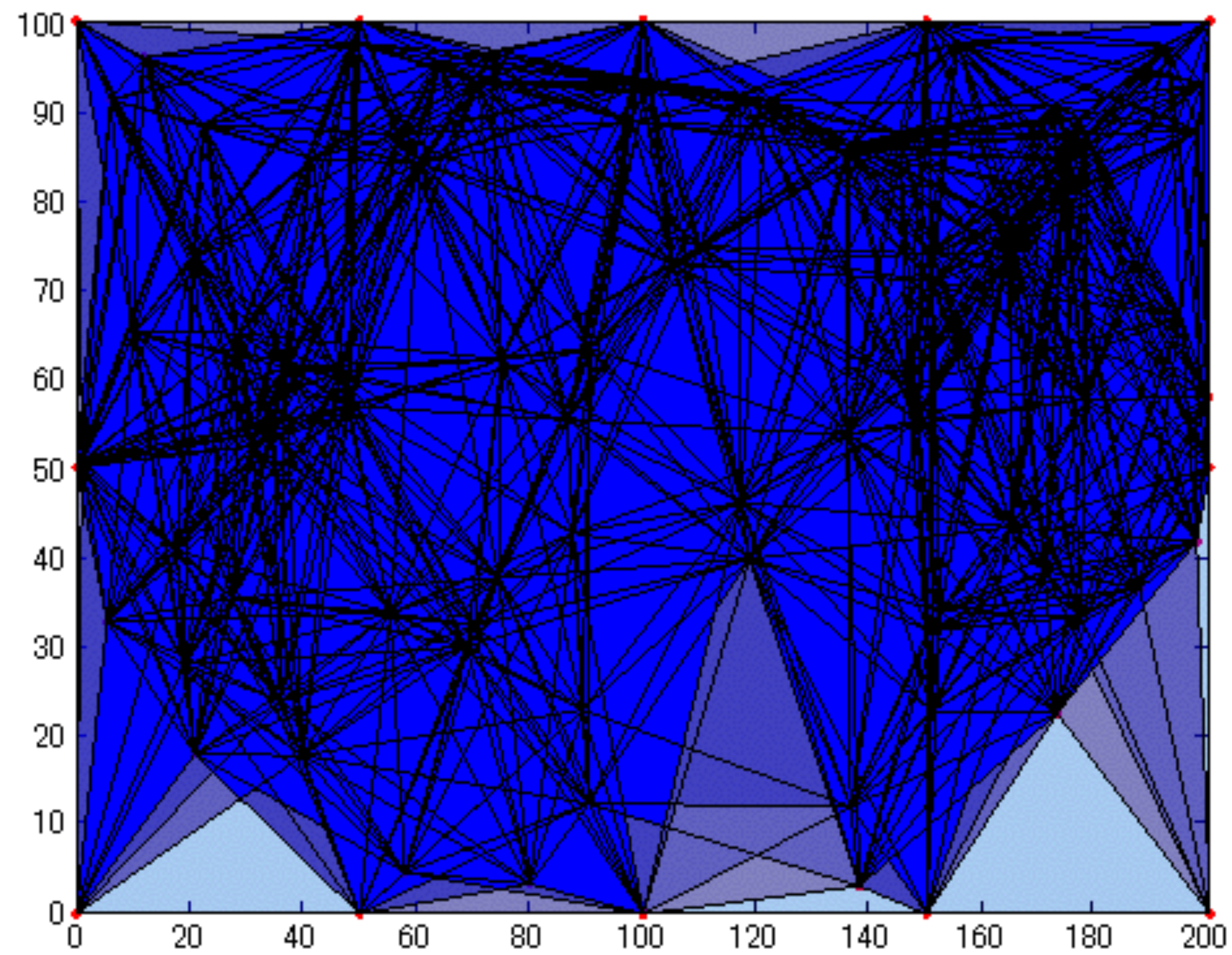
Experiment

Decompose D into vertical rectangles



Experimental Result



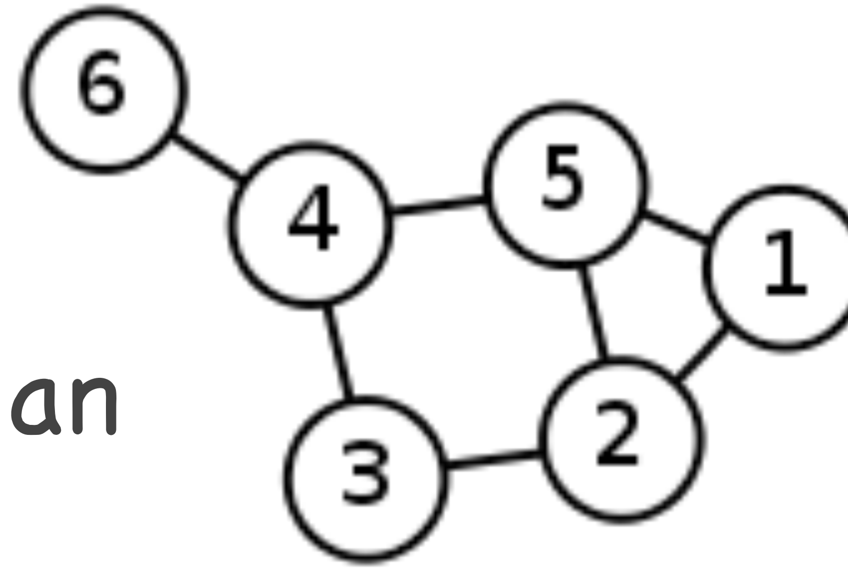


#{2-simplex}=519, #{1-simplexes}=184, #{0-simplexes}=30

# of partition elements	1	2	5
computational time	81.0	49.2	1.0

Discrete Laplacian and Homology Computation

Graph Laplacian



For a graph G, the matrix

$L = D - A$ is called its Laplacian

where

D : degree matrix (diagonal)

A : adjacency matrix (off-diagonal)

$$\begin{pmatrix} 2 & -1 & 0 & 0 & -1 & 0 \\ -1 & 3 & -1 & 0 & -1 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 3 & -1 & -1 \\ -1 & -1 & 0 & -1 & 3 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \end{pmatrix}$$

Using the directed Incidence matrix B, we can write $L = BB^T$

Laplacian for Simplicial Complexes

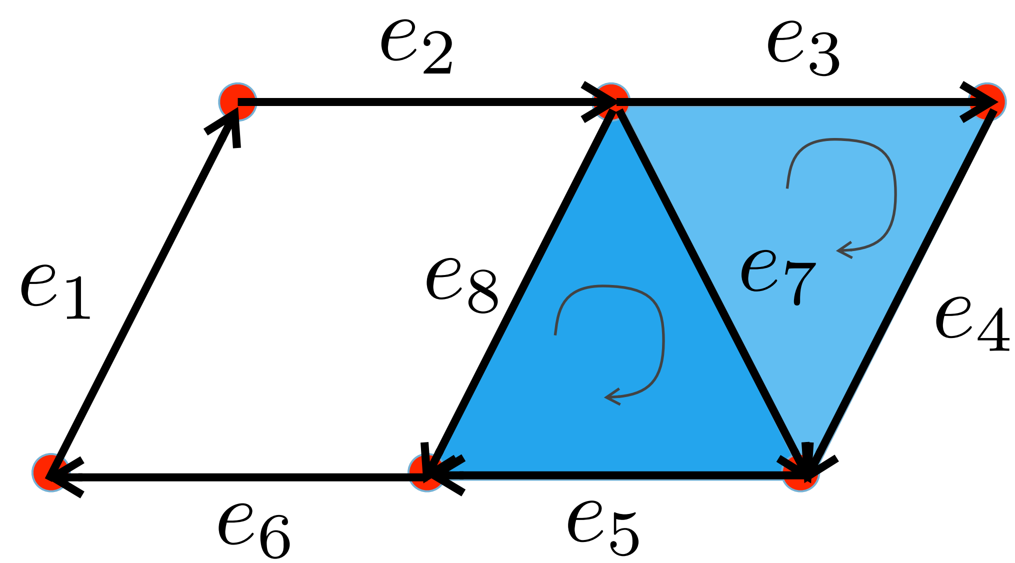
Recall that the incidence matrix of a graph is just the boundary operator of the graph considered as a simplicial complex.

For a simplicial complex X, define its k-th Laplacian by

$$L_k : C_k(X) \to C_k(X)$$
$$L_k := \partial_k^* \circ \partial_k + \partial_{k+1} \circ \partial_{k+1}^*$$

(L_0 is the usual graph Laplacian)

An Example



$$L_1 = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & -1 & -1 \\ 0 & -1 & 3 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 3 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 3 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 & -1 & 2 & 0 & -1 \\ 0 & -1 & 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & -1 & 1 & 0 & 0 & -1 & 0 & 3 \end{pmatrix}$$

A Little bit of Harmonic Analysis

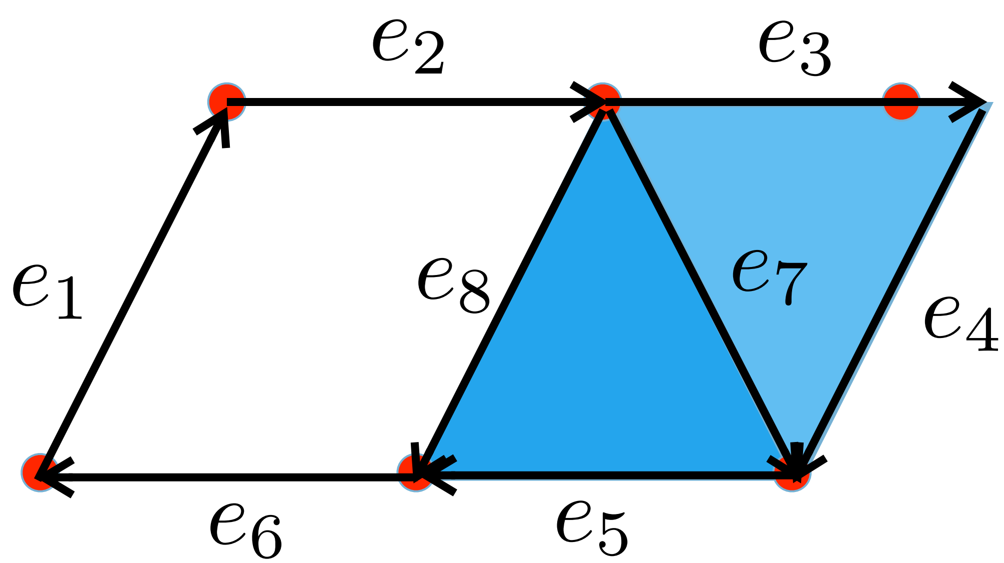
Theorem.

$$C_k(X) = \operatorname{im}\partial_{k+1} \oplus \ker L_k \oplus \operatorname{im}\partial_k^*$$

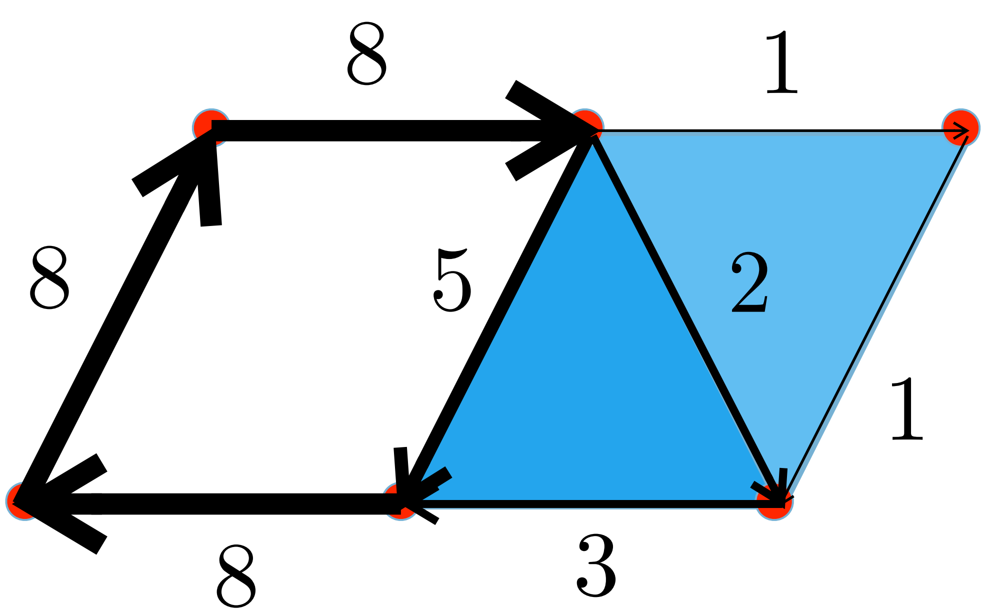
Therefore, $H_k(X) \cong \ker\ L_k$

A discrete analogue of the Hodge decomposition.

An Example



$$\ker L_1 = \langle (8,8,1,1,3,8,2,5) \rangle \cong \mathbb{R}$$



This recovers $H_1(X, \mathbb{R}) = \mathbb{R}$

Note that at each vertex, we have inflow = outflow.

Weights on edges represents the corresponding generator.

Application to Coverage Problem

If $\ker L_1 = \{0\}$ then $D \subset \mathcal{U}$

That is, if the kernel of L_1 is trivial, there is no hole in the coverage.

How can we check $\ker L_1 = \{0\}$?

Heat Equation

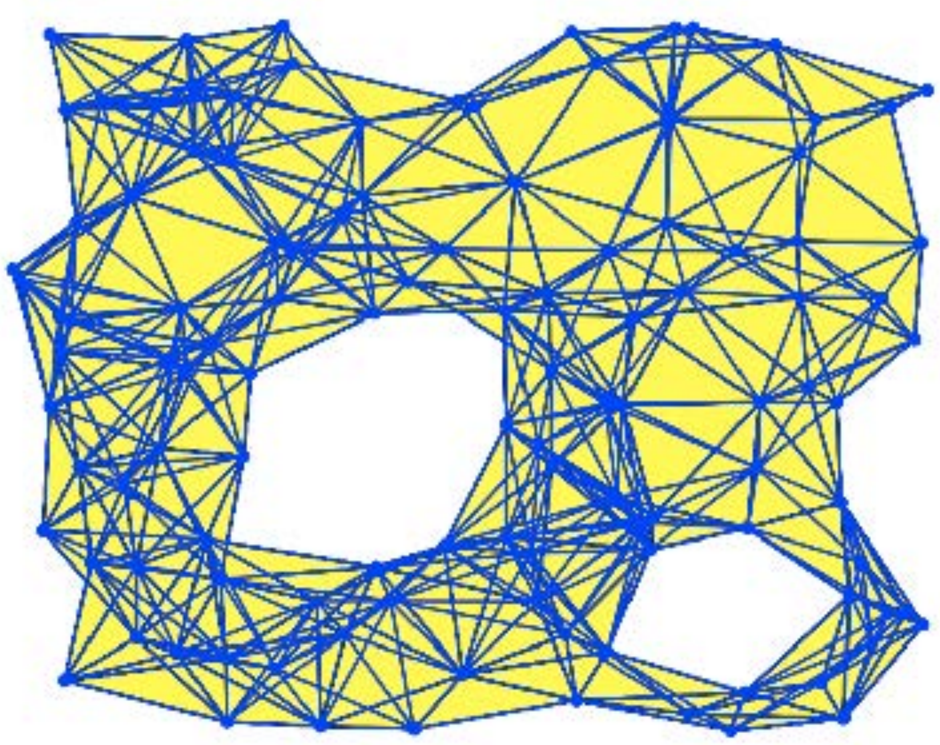
Consider the "heat" ODE $$\frac{dx}{dt} = -L_1 x$$

0-solution is globally stable $\Longleftrightarrow$ $\ker L_1 = \{0\}$

(Note that L_1 is positive semi-definite)

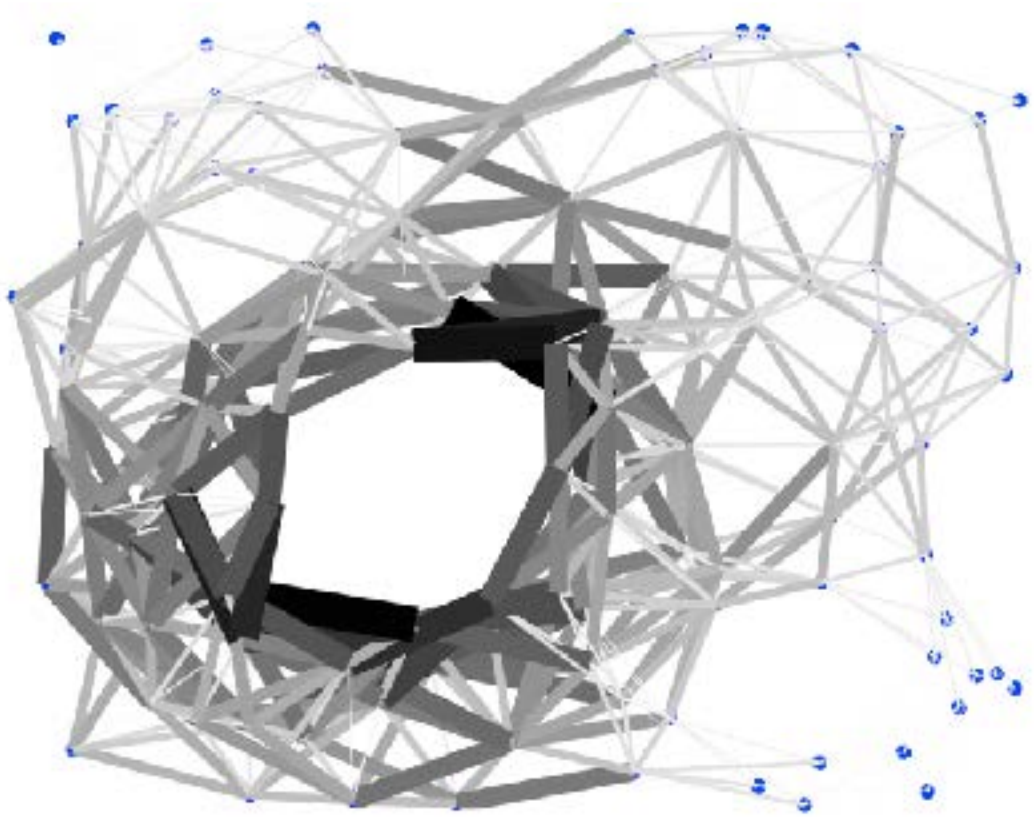
Therefore, if the solutions of this ODE with several distinct initial conditions always converges to 0, then we can conclude $D \subset \mathcal{U}$.

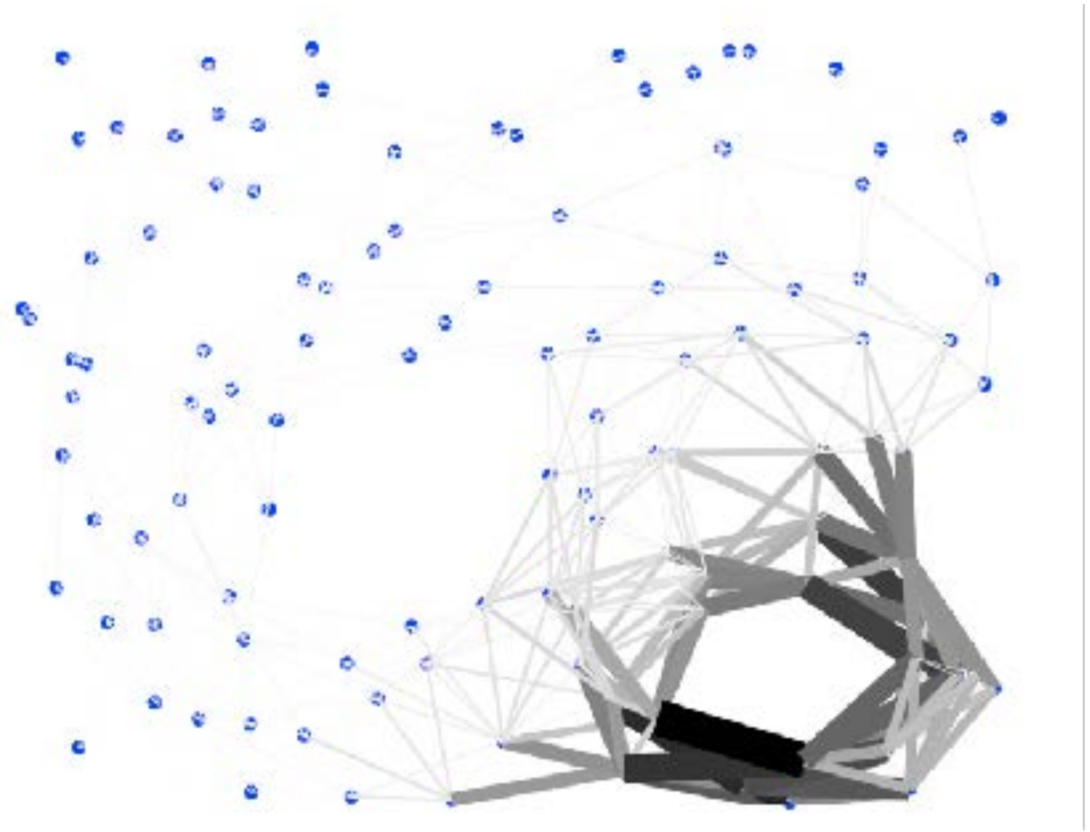
An Example



$$H_1(\mathcal{R}, \mathbb{R}) = \mathbb{R}^2$$

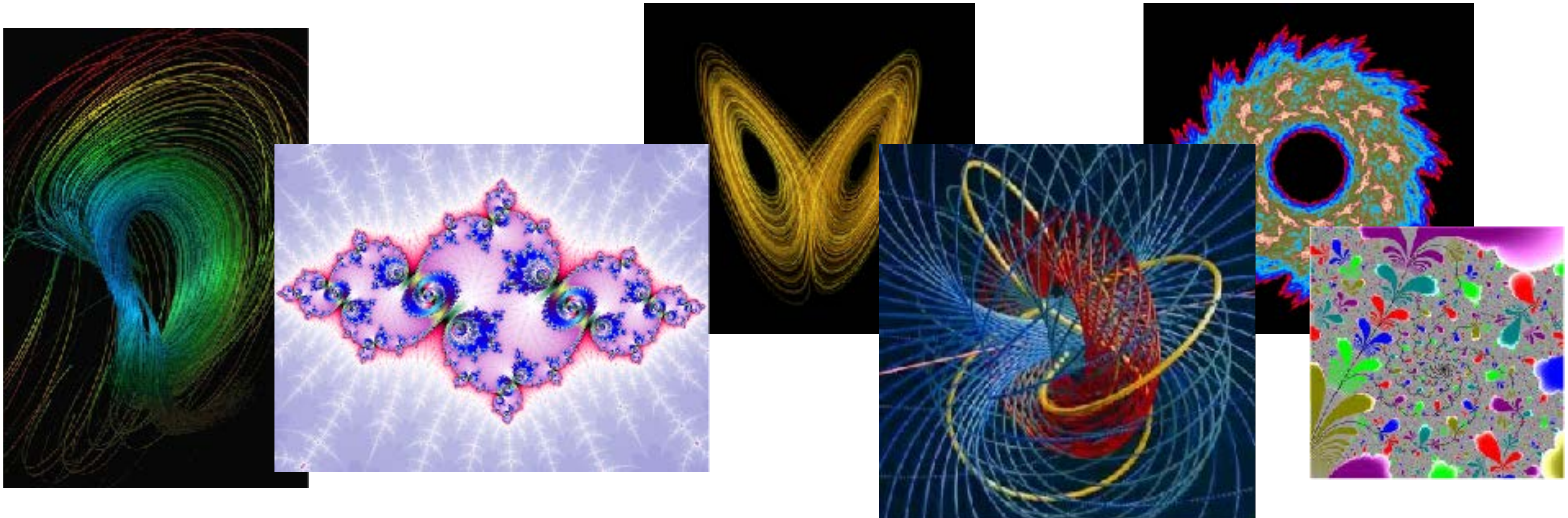
There should be two independent solutions





Graph Theoretical Tools for Dynamical Systems

Dynamical Systems



Dynamical Systems

= phase space + time evolution law

Typically, the phase space is a manifold M.

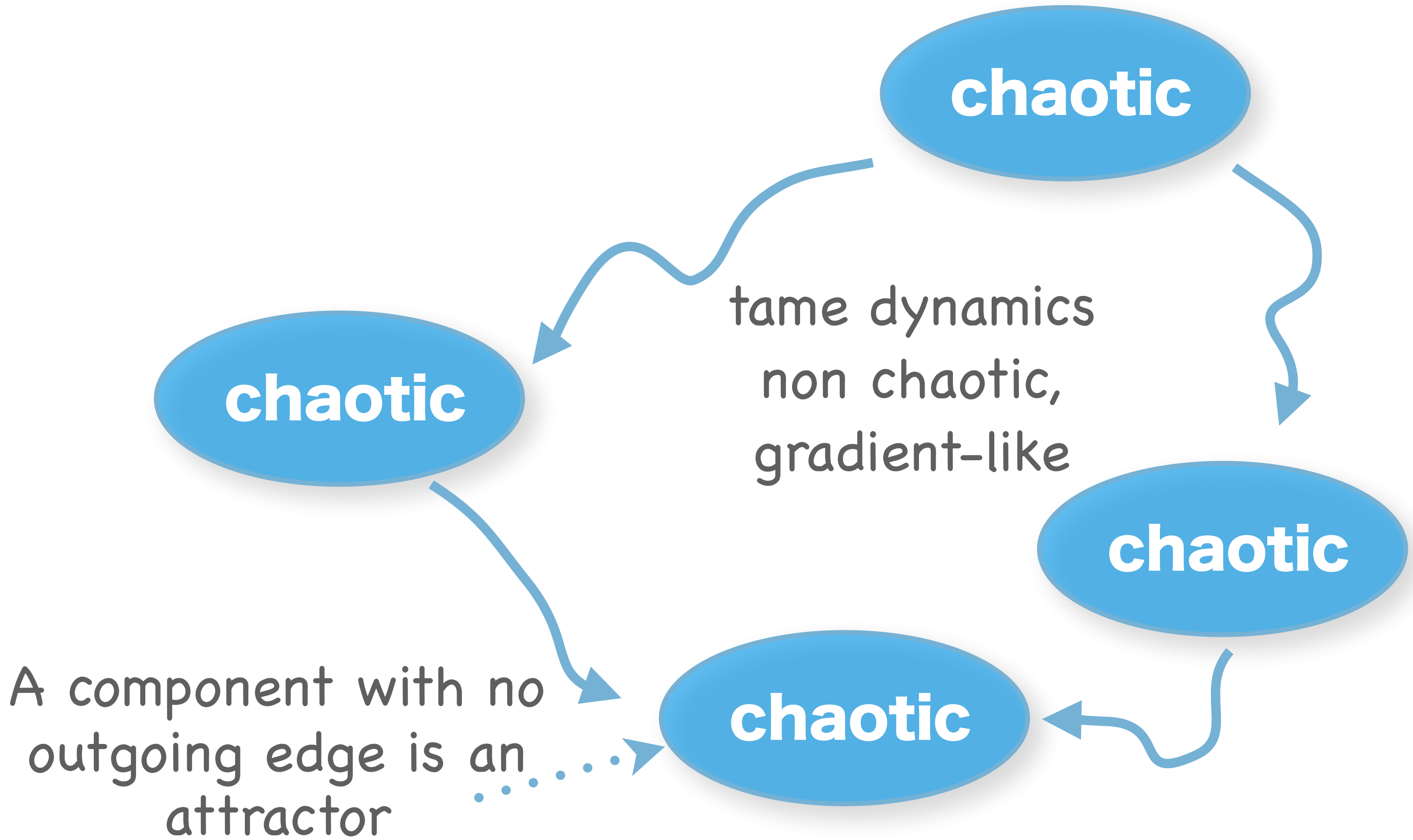
The time evolution law would be given by:

a differential equation on M (continuous time);

a map $f : M \to M$ (discrete time) .

Conley's Fundamental Theorem

Theorem (Conley). Any dynamical system can be decomposed into **"chain recurrent sets"** and **"connecting orbits"** between them. All the chaotic orbits are contained in chain recurrent sets.

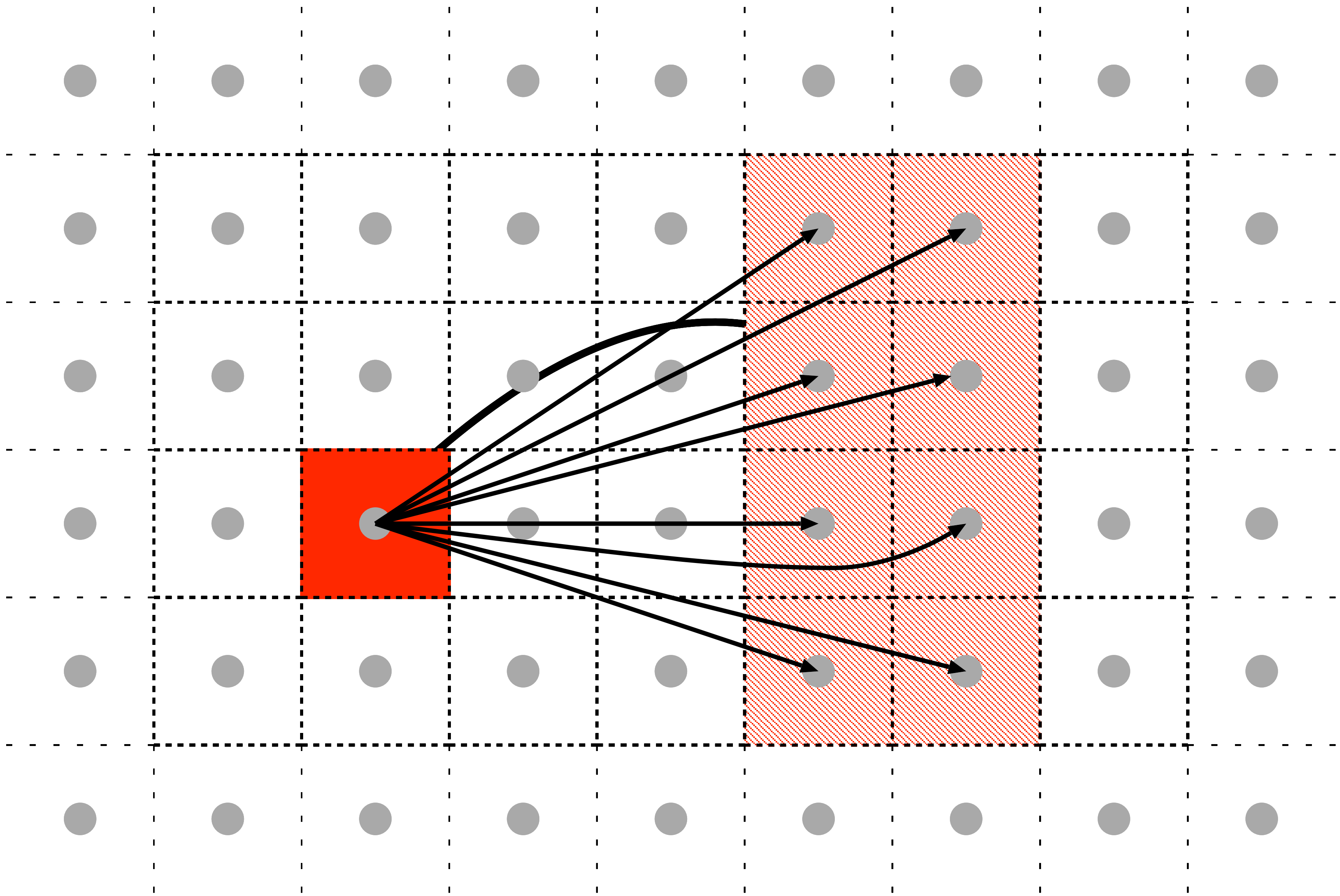


Wanted:

a computational approach to this decomposition

From Dynamics to Di-Graph

Given a dynamical systems $f : X \to X$ we construct a directed graph $G(f)$ which imitates the dynamics of f



Automated Analysis for DS

Periodic orbits:

Theorem. The set of periodic points of f is contained in the cubes corresponding to cycles of $G(f)$.

Note that this theorem holds regardless of the stability of the periodic orbits.

The maximal invariant set and the chain recurrent set:

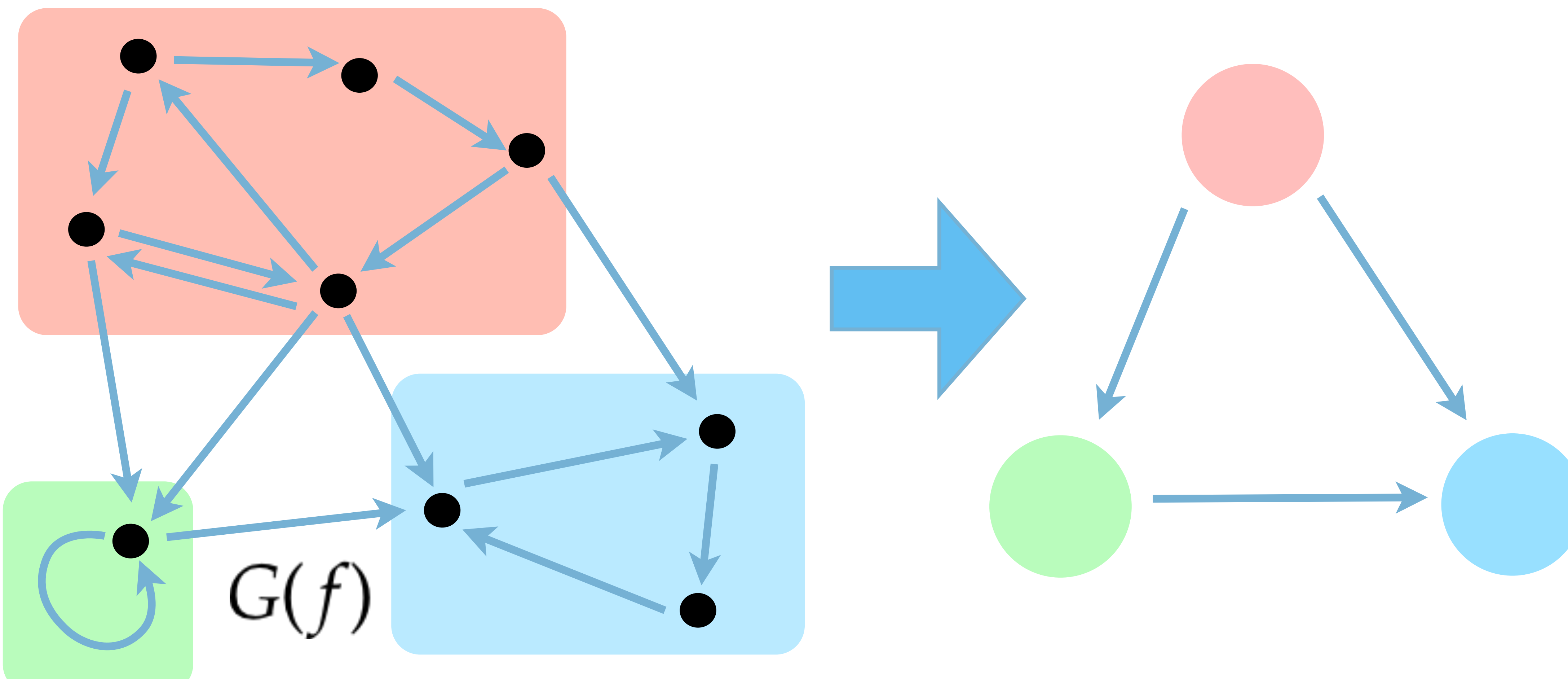
$\mathrm{Inv}\, G(f) := \{v \in G(f) \mid \exists \text{bi-infinitely long path through } v\}$

$\mathrm{Scc}\, G(f) := \{v \in G(f) \mid \exists \text{cycle through } v\}$

Theorem. The maximal invariant set of f is contained in $|\,\mathrm{Inv}\, G(f)\,|$. The chain recurrent set of f is contained in $|\,\mathrm{Scc}\, G(f)\,|$.

Collapsing G(f)

Definition. A Conley-Morse graph of a dynamical system f is the directed graph obtained from $G(f)$ by collapsing each strongly connected component into a single vertex and defining the edges by the transitivity among each strongly connected components.



Example:
Nonlinear Leslie Model

Linear Leslie Population Model

$$f : \mathbb{R}^k \to \mathbb{R}^k : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_k \end{pmatrix} \mapsto \begin{pmatrix} \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k \\ p_1 x_1 \\ p_2 x_2 \\ \vdots \\ p_{k-1} x_{k-1} \end{pmatrix}$$

Consider a population divided into k generations:

x_i : the number of individuals in the i-th generation

p_i : the probability that an individual in the i-th generation survives for one generation

θ_i : the fertility of the i-th generation

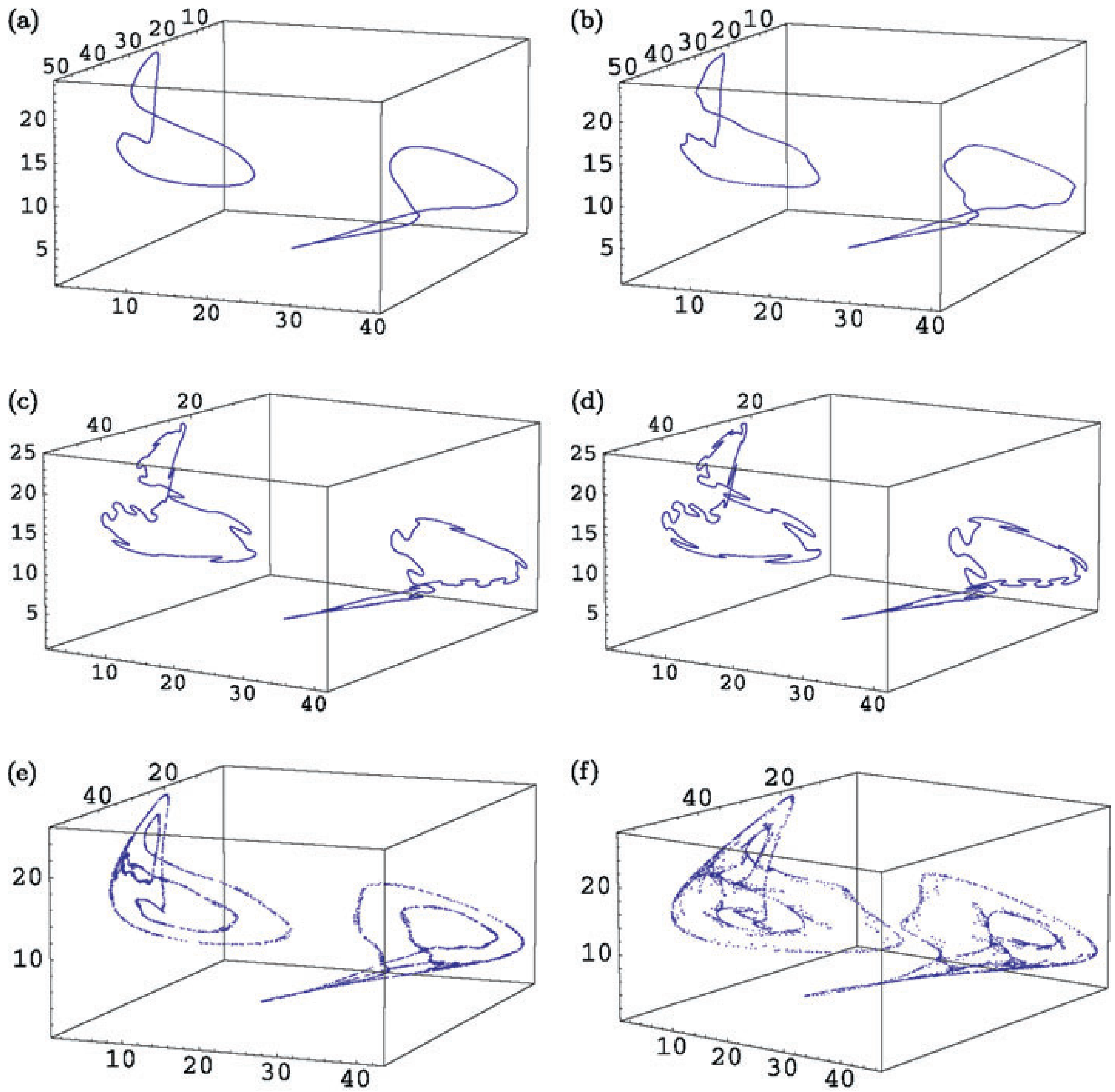
Nonlinear Leslie Model

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_k \end{pmatrix} \mapsto \begin{pmatrix} (\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k) \cdot e^{-\lambda(x_1+x_2+\cdots+x_k)} \\ p_1 x_1 \\ p_2 x_2 \\ \vdots \\ p_{k-1} x_{k-1} \end{pmatrix}$$

Consider a population divided into *k* generations:

- x_i : the number of individuals in the *i*-th generation
- p_i : the probability that an individual in the *i*-th generation survives for one generation
- θ_i : the fertility of the *i*-th generation

Multiple Coexisting Attractors



Ugarcovich-Weiss, Nonlinearity 17 (2004)

Qualitative Questions

Q. When is a dynamical system has multiple attractors?

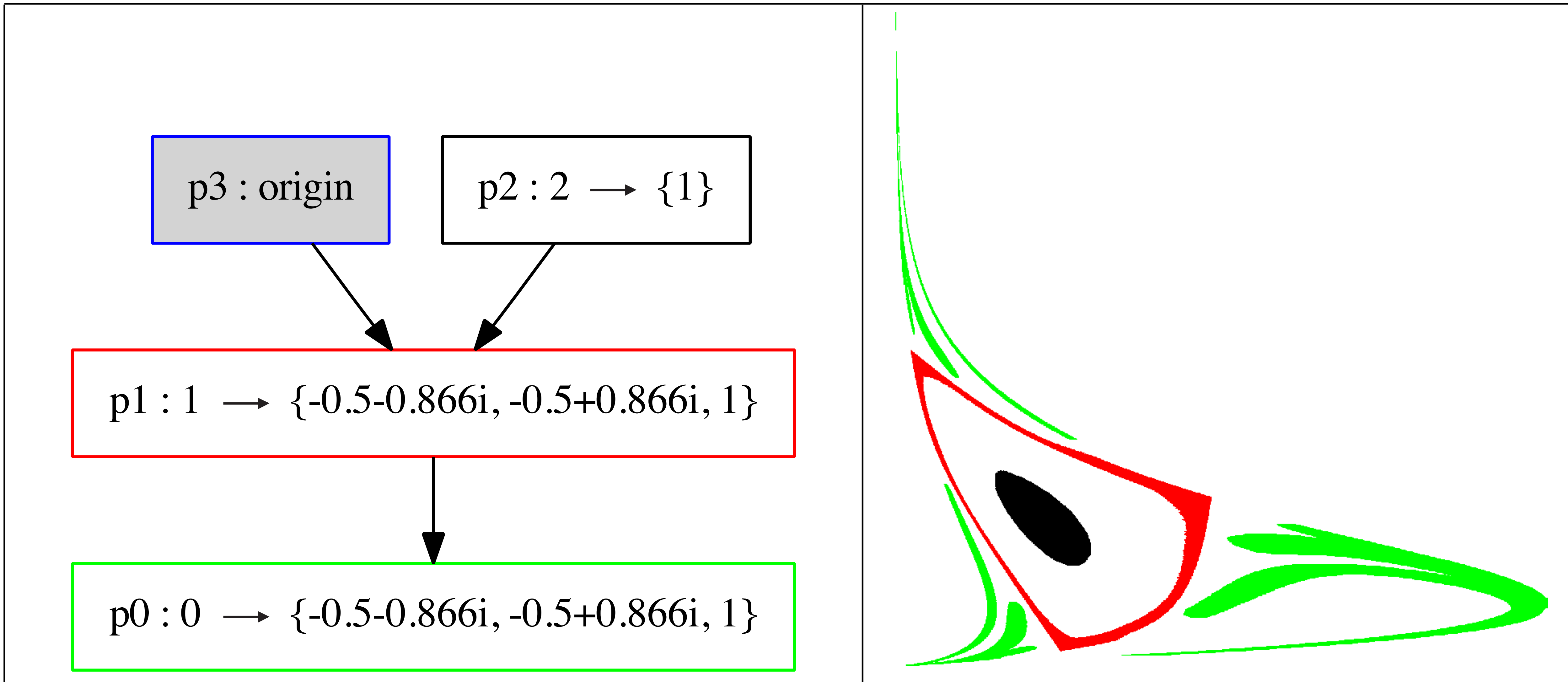
Q. When is a dynamical system chaotic?

Conventional Method: bifurcation analysis

too many parameters
too many ways to chaos
few (and sometimes poor) postdocs…

Want. We want to construct automatic and rigorous methods which can answer these "qualitative questions".

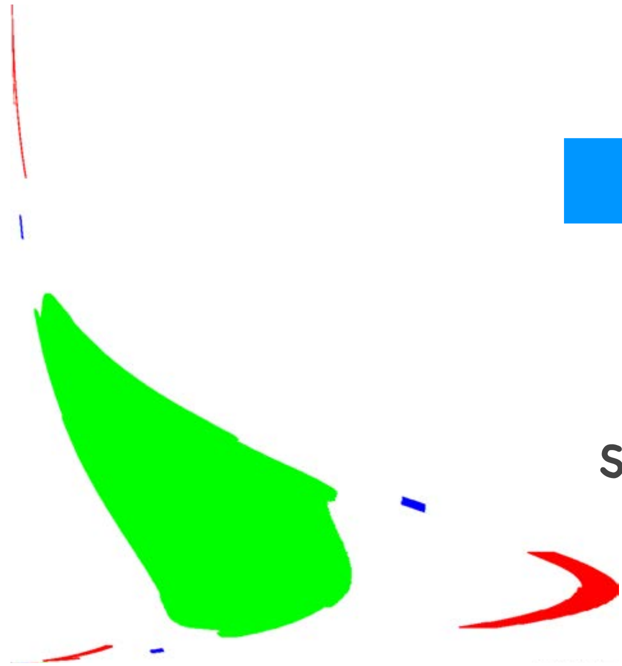
Conley-Mose Graph



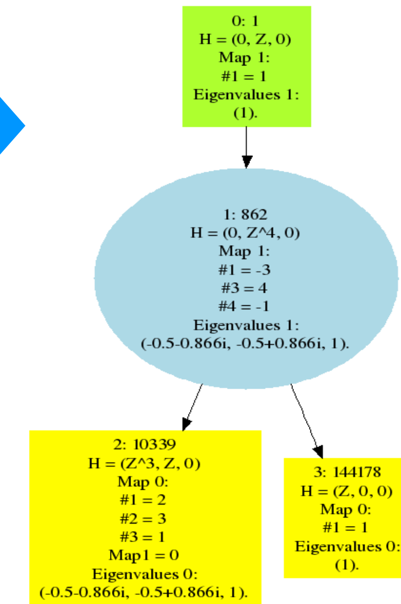
7. The Conley-Morse graph CMG(7) and the sets $\mathcal{M}(p)$ at the box $(31, 22)$.

Conley Morse Graph

chaotic invariant set

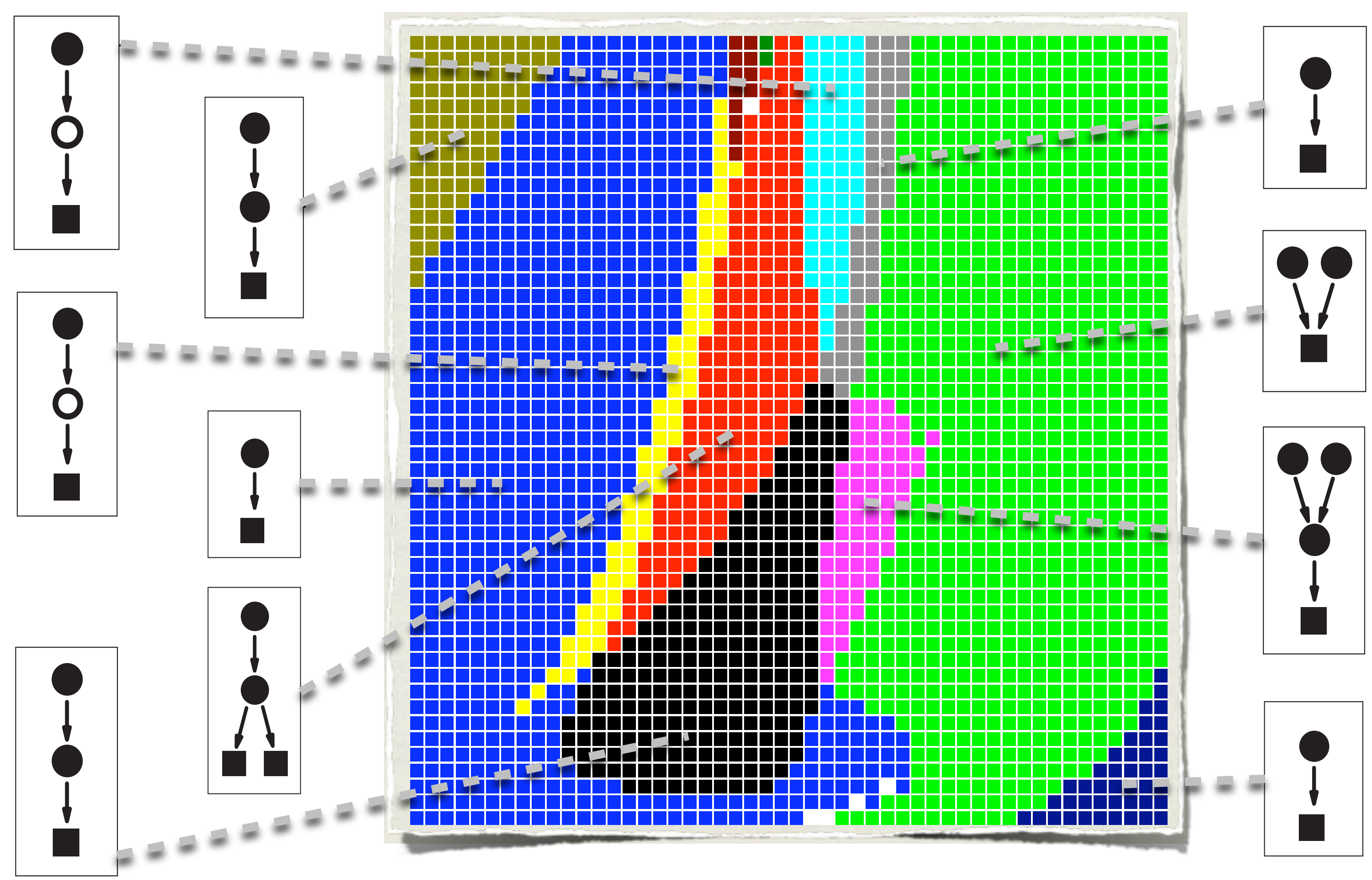


Represent the phase space structure using directed graph called Conley-Morse Graph



discretized algebraic information

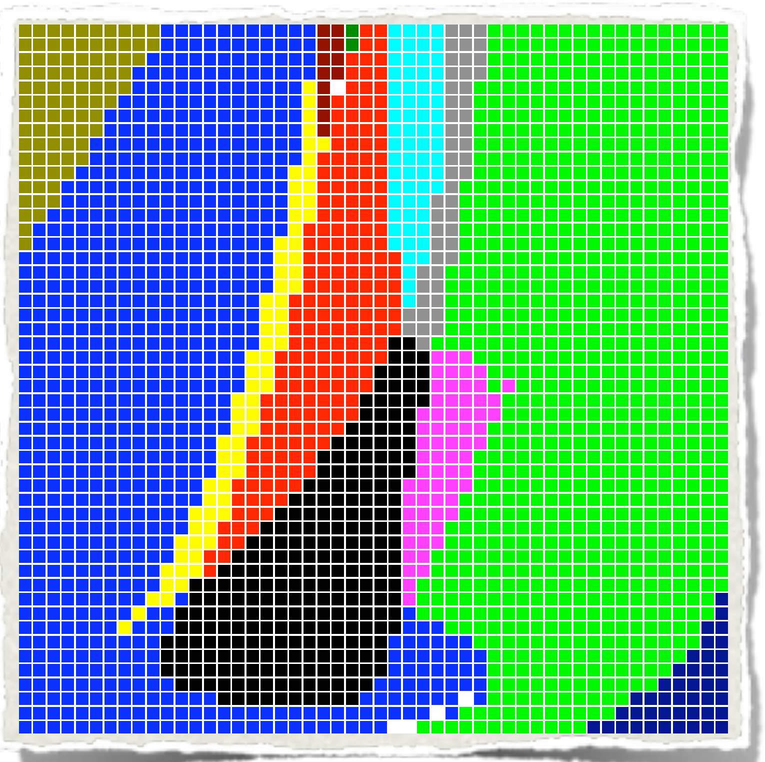
Classify the parameter values according to the discretized dynamical information to obtain "bifurcation diagram"

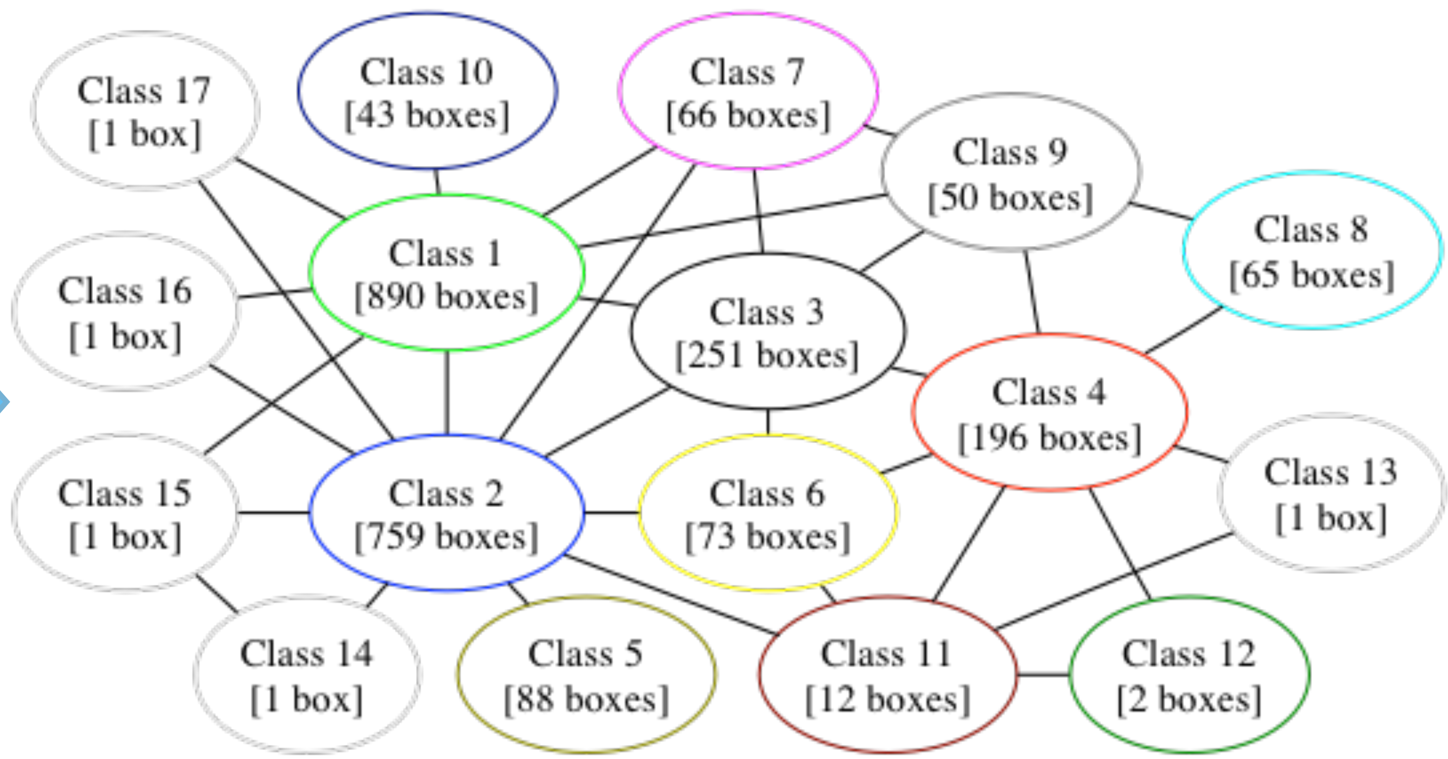


Z. Arai, W. Kalies, H. Kokubu, K. Mischaikow, H. Oka and P. Pilarczyk
"A database schemat for the analysis of global dynamics of multiparameter systems"
SIAM J. Appl. Dyn. Sys. 8 (2009)

Continuation Graph (2D)

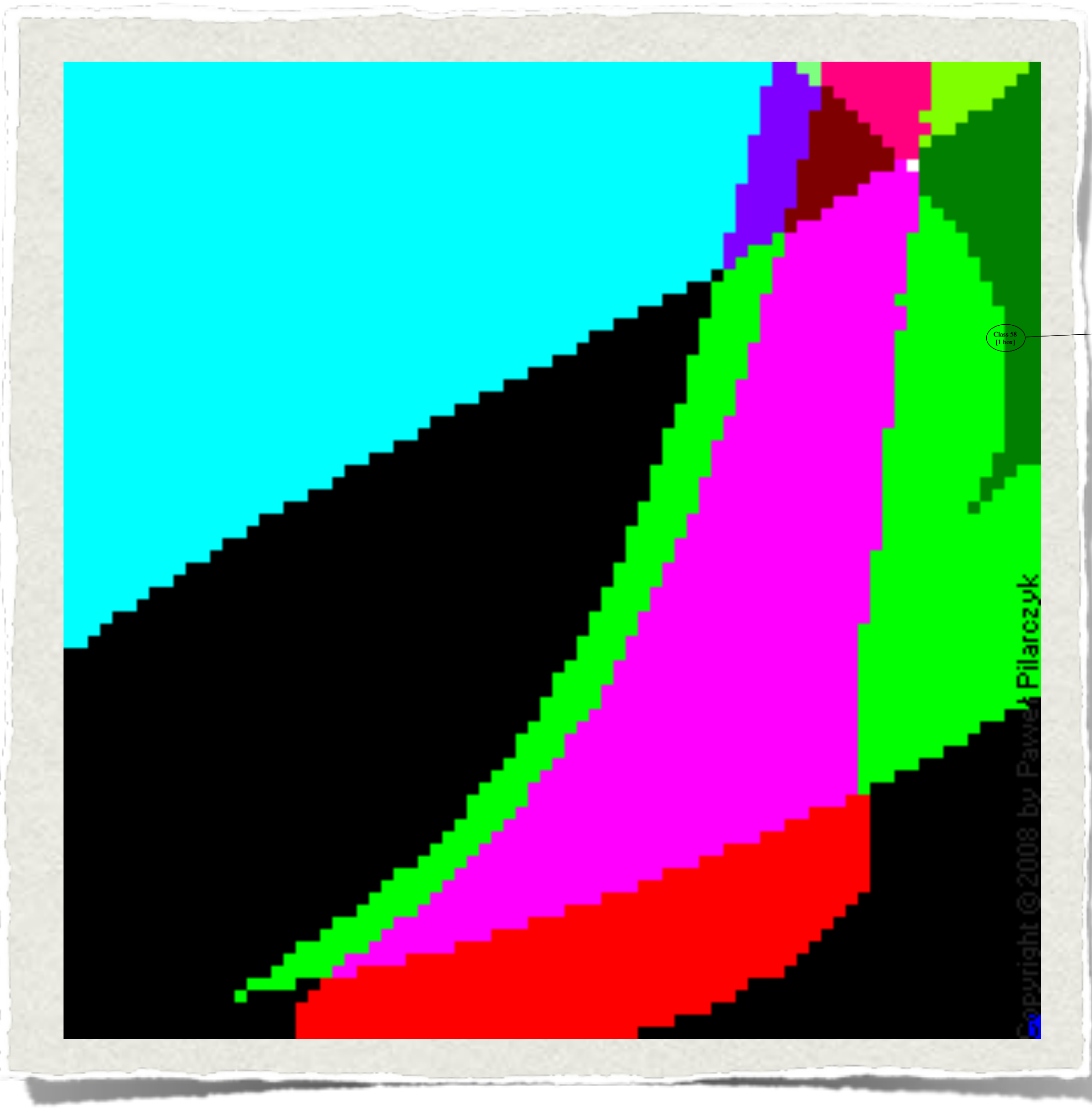
Adjacency data of the decomposition of the parameter space can be expressed by continuation graphs.

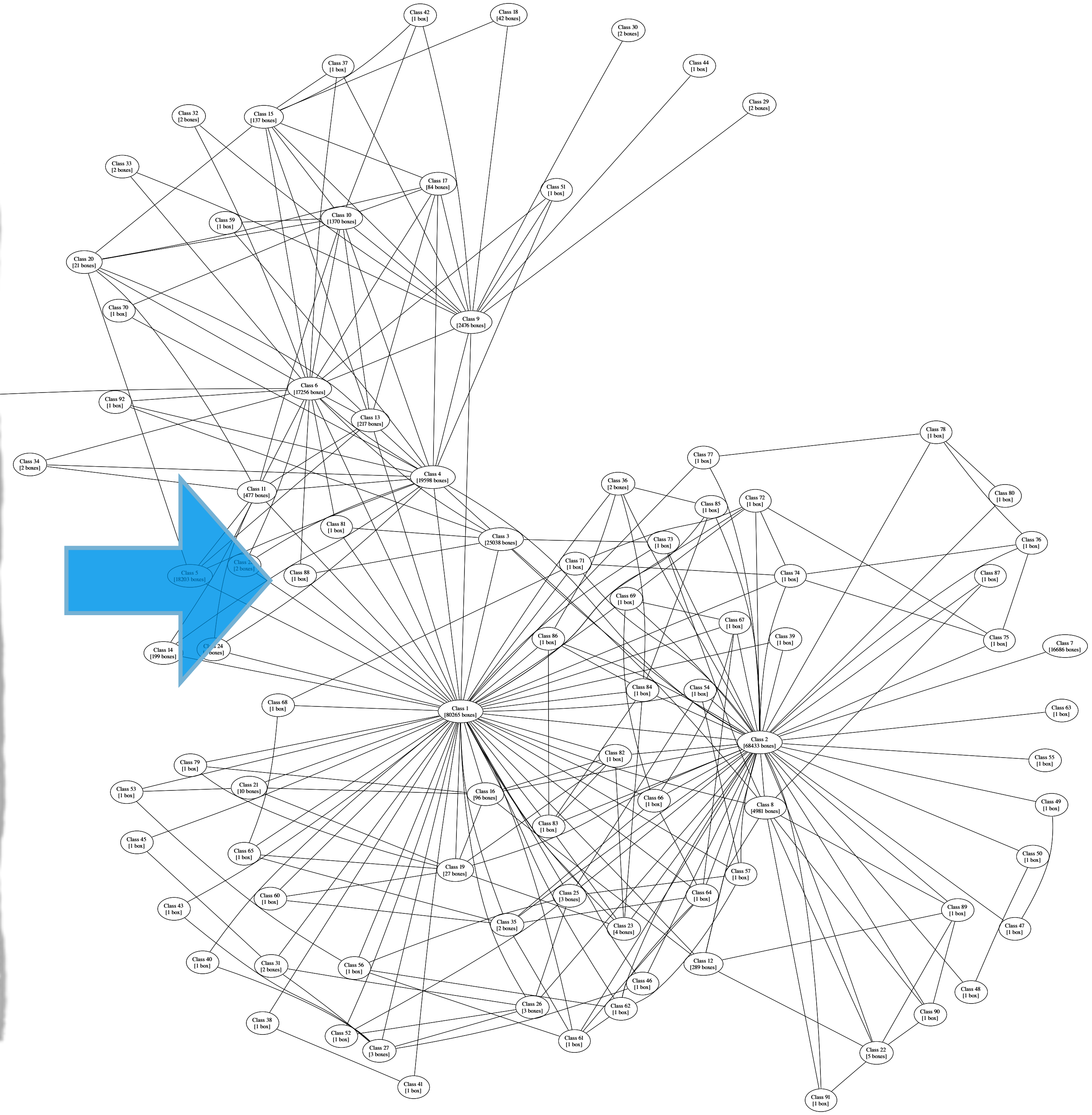




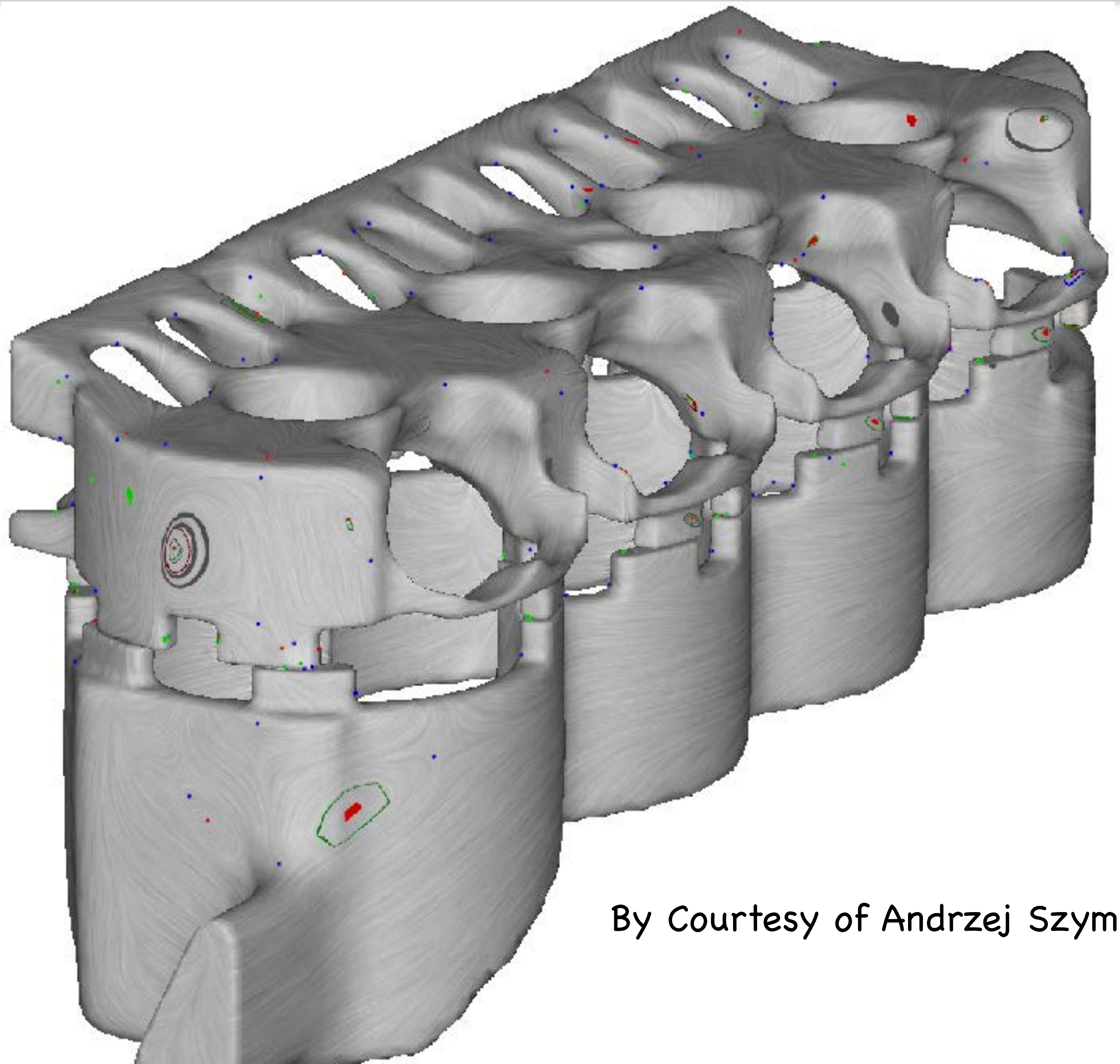
Edges of the continuation graph correspond to "bifurcation curves" in the parameter space.

Continuation Graph (3D)



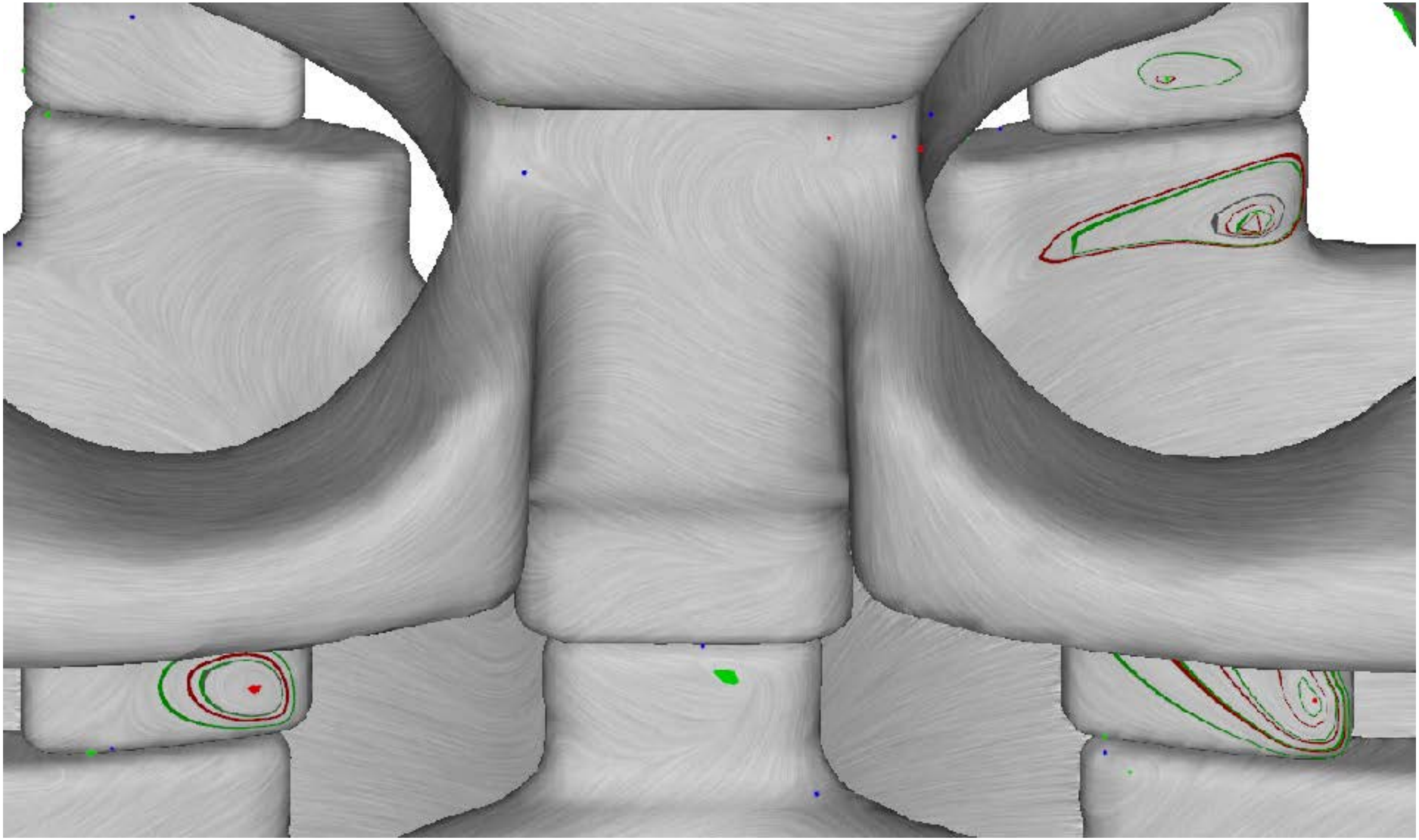


Example: Engine cooling jacket 1



By Courtesy of Andrzej Szymczak

Example: Engine cooling jacket 2



By Courtesy of Andrzej Szymczak

Pseudo Lyapunov Function

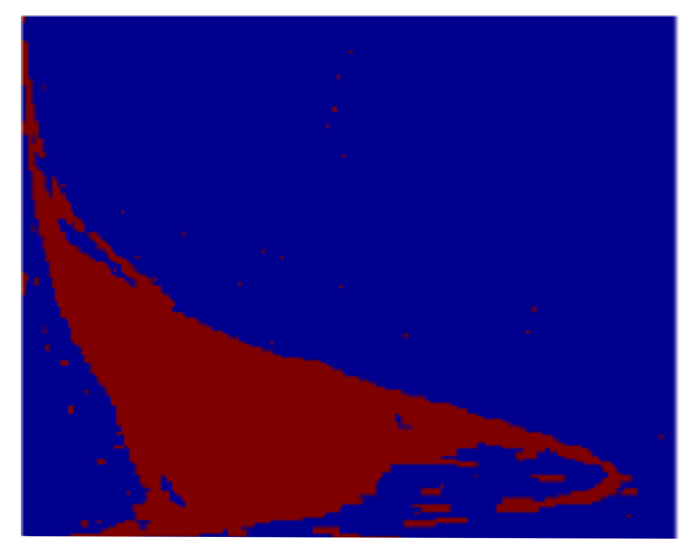
Failure of CM Decomposition

CM decomposition might be trivial when

* the resolution is not fine enough...
* the noise is too large...
* there is no gradient structure

A good decomposition

low resolution



unbounded noise

Pseudo Lyapunov Function

We want to construct a combinatorial analog of Lyapunov function, which is a map $s : V \to \mathbb{R}$ such that for each edge $e = (i, j) \in E$, the value of the map is very likely to increase along e.

Precisely, we want to find s that minimizes

$$\sum_{(i,j)\in E} ((\mathrm{grad}\ s)(i, j) - A_{ij})^2.$$

where $A = (A_{ij})$ is the adjacent matrix and $\mathrm{grad}\ s : V \times V \to \mathbb{R}$ is defined by

$$(\mathrm{grad}\ s)(i, j) := s(j) - s(i).$$

Combinatorial Hodge Decomposition

Using the combinatorial Hodge decomposition (Jiang-Lim-Yao 2011), we can show that the normal equation for this minimizing problem is

$$\Delta_0 s = -\text{div}\, A$$

where Δ_0 is the graph Laplacian and

$$(\text{div}\, A)(i) = \sum_{j:(i,j)\in E} A_{ij}\,.$$

Thus our minimizer can be written as

$$s = -\Delta_0^{\dagger}\text{div}\, A$$

where $\dagger$ indicates Moore-Penrose pseudoinverse.

Results