Multiscale parameters in computational topology

Vin de Silva,
Stanford University

Acknowledgements

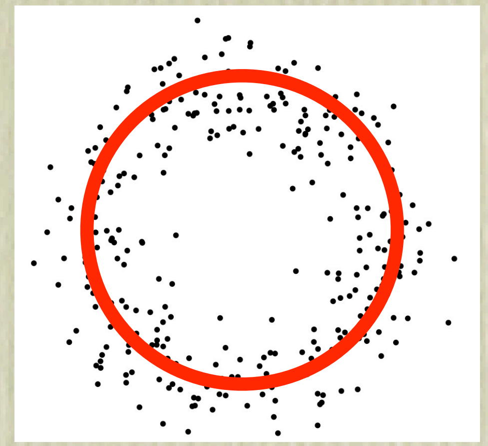
- Gunnar Carlsson (Mathematics, Stanford)
  —principal collaborator
- Afra Zomorodian (CS/Robotics, Stanford)
  —persistent homology software
- Josh Tenenbaum (Brain & CogSci, MIT)
  —‘landmarks’ philosophy
- David Mumford (Mathematics, Brown)
  —visual image data

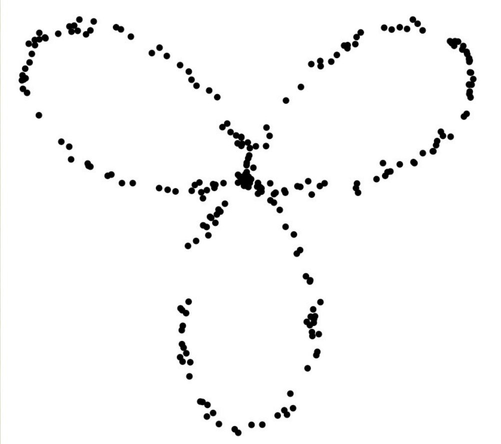
High-dimensional data

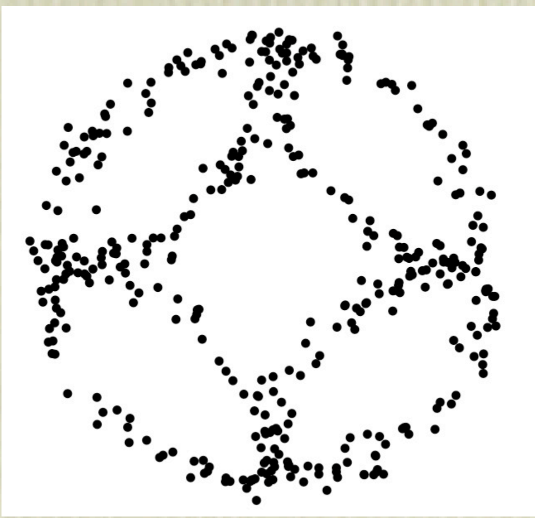
- Modern scientists are often confronted with very large high-dimensional data sets.
  - lots of test subjects
  - lots of observed variables
  - observed phenomenon may still be simple
- How do you extract low-dimensional structure from a high-dimensional data set?

Topological structure

- "Identify topological features of a point-cloud dataset."
- Perhaps the data are sampled finely from some unknown object.
- Can we describe the topological properties of the object?







Present goal

- Develop robust methods for extracting topological features from point-cloud data.
- Develop an accompanying theory of "point-cloud topology".
- Address geometrical questions such as localisation of features.

Applications

- Shape descriptors from tangent-space topology.
  [Collins, Zomorodian, Carlsson, Guibas, 2004]

- Locating singular points in a data set.
  [Carlsson, Carlsson, de Silva, 2003]

- Estimating the fractal dimension of dynamical system attractors.
  [Robins, Meiss, Bradley, 2000]

- Dimension estimation, hole detection, ...

1. Topology of spaces

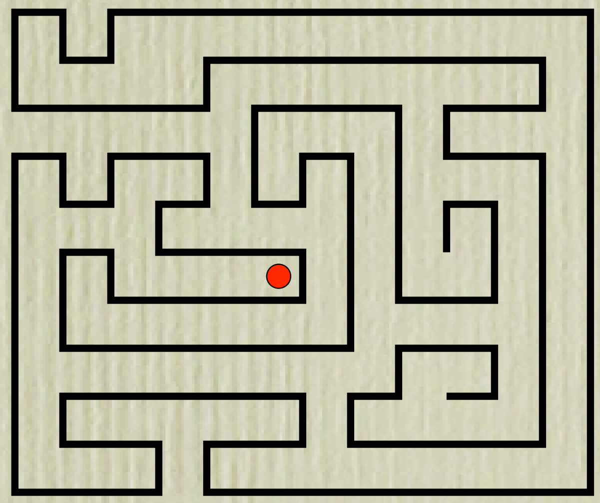
What is topology?

- It is the branch of mathematics which cannot distinguish between a teacup and a bagel.

=

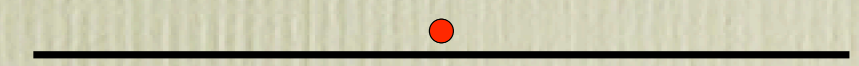
Why topology?

- It strips away irrelevant geometrical details and identifies the essential structure of a space.



Why topology?

- It strips away irrelevant geometrical details and identifies the essential structure of a space.



Why topology?

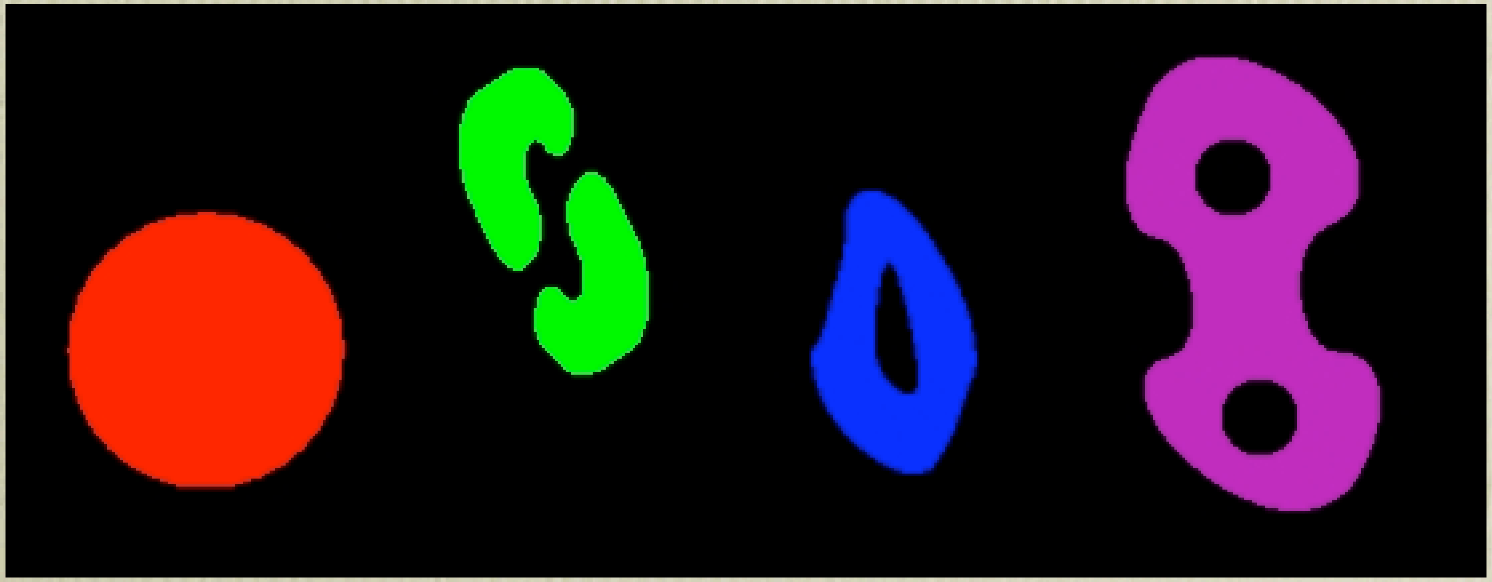
- It gives answers to qualitative questions.

Why topology?

- It gives answers to qualitative questions. [Carlsson, Collins, Guibas, Zomorodian]

Betti numbers

- Betti numbers give a count of basic topological features: components, holes, etc.
- Sensible goal: estimate Betti numbers.

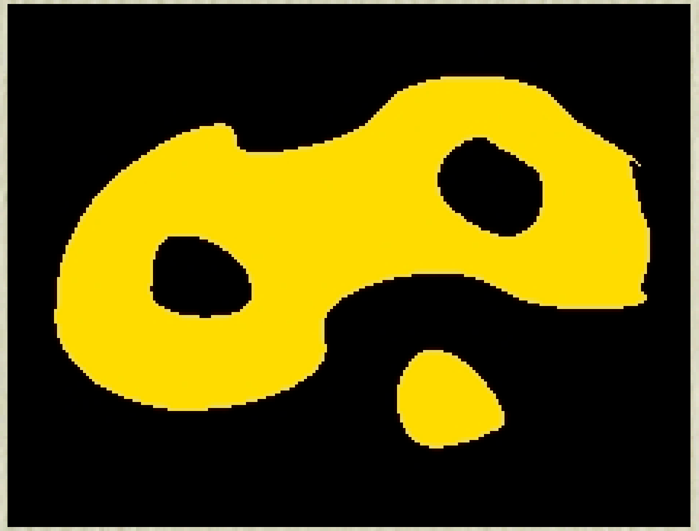


Betti numbers

- The k-th Betti number bk(X) is a non-negative integer which measures the k-dimensional connectivity of a space X.

For a 2-dimensional object

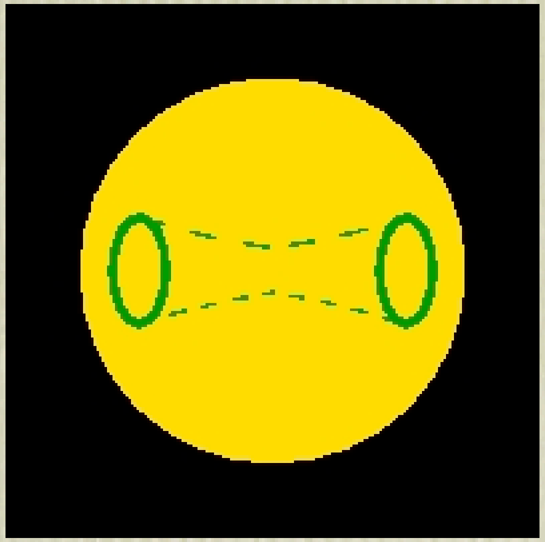
- b_0 = # connected components
- b_1 = # holes



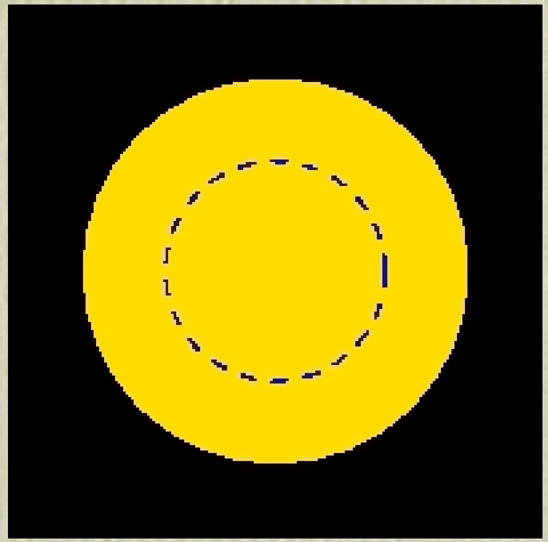
$b_0 = 2$, $b_1 = 2$

For a 3-dimensional object

- b_0 = # connected components
- b_1 = # tunnels or handles
- b_2 = # voids



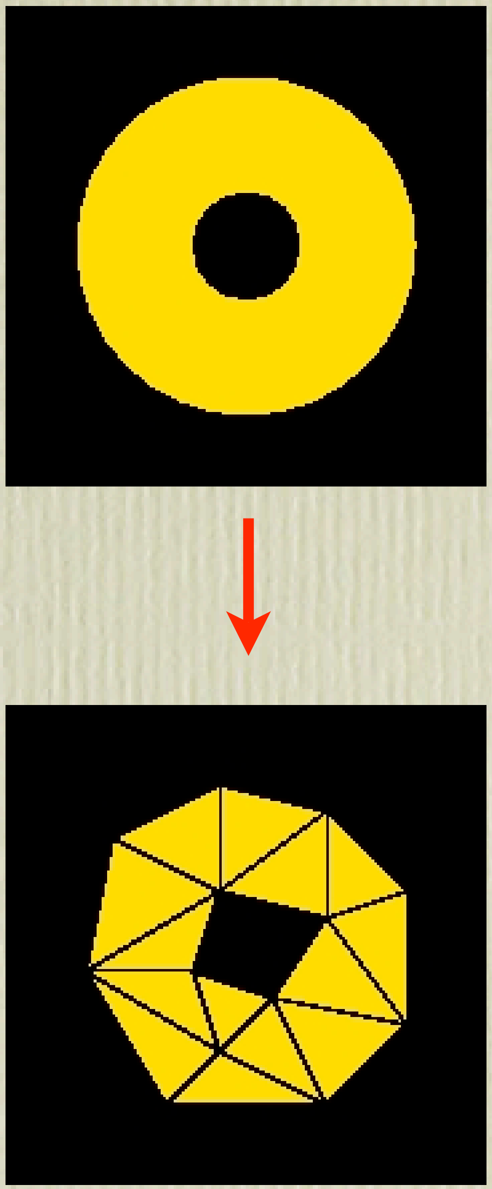
$b_0 = 1, b_1 = 1, b_2 = 0$



$b_0 = 1, b_1 = 0, b_2 = 1$

Calculating Betti numbers

- Betti numbers are defined abstractly for topological spaces.
- (This uses infinite-dimensional linear algebra.)
- Often we can represent the space by a finite simplicial complex.
- This reduces the problem to finite-dimensional linear algebra.



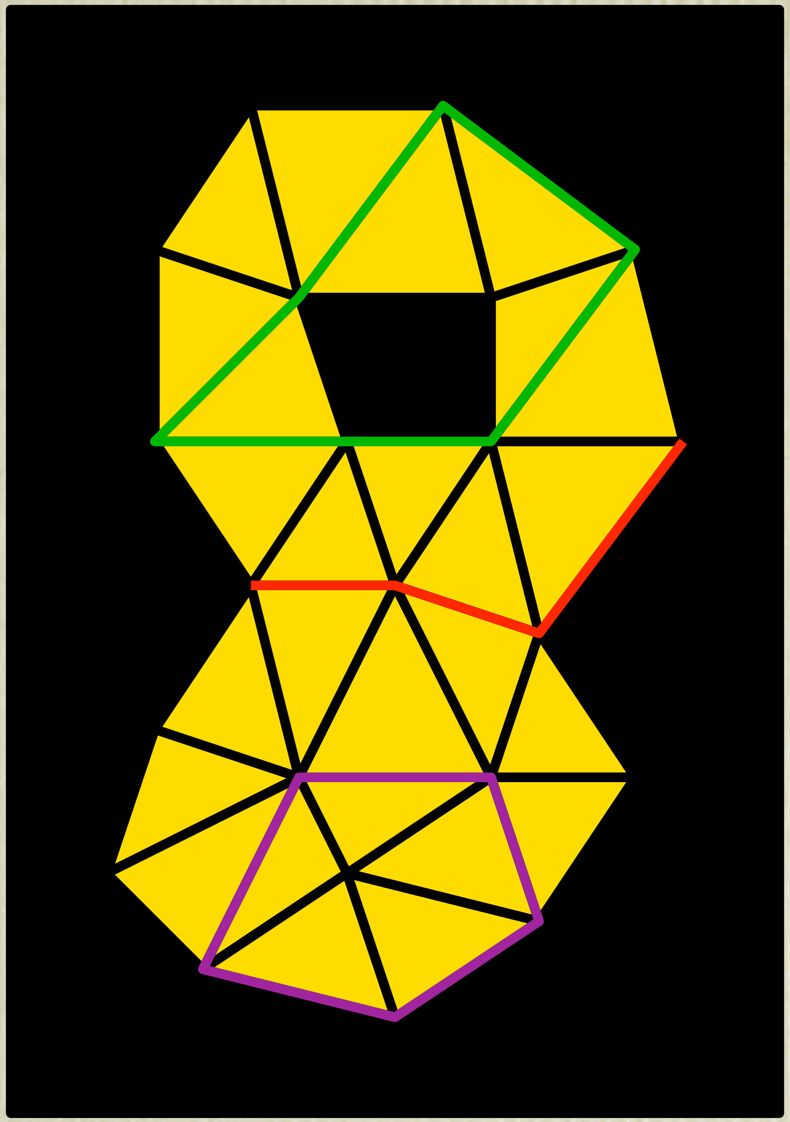
Simplicial homology

- C_k = vector space with a generator $\underline{\alpha}$ for each k-simplex α of simplicial complex
- α_i = $(k-1)$-simplex obtained by deleting the i-th vertex of α.
- Boundary map $\partial : C_k \to C_{k-1}$ defined:

$$\partial \underline{\alpha} = \sum_{i=0}^{k} (-1)^i \underline{\alpha}_i$$

Simplicial homology

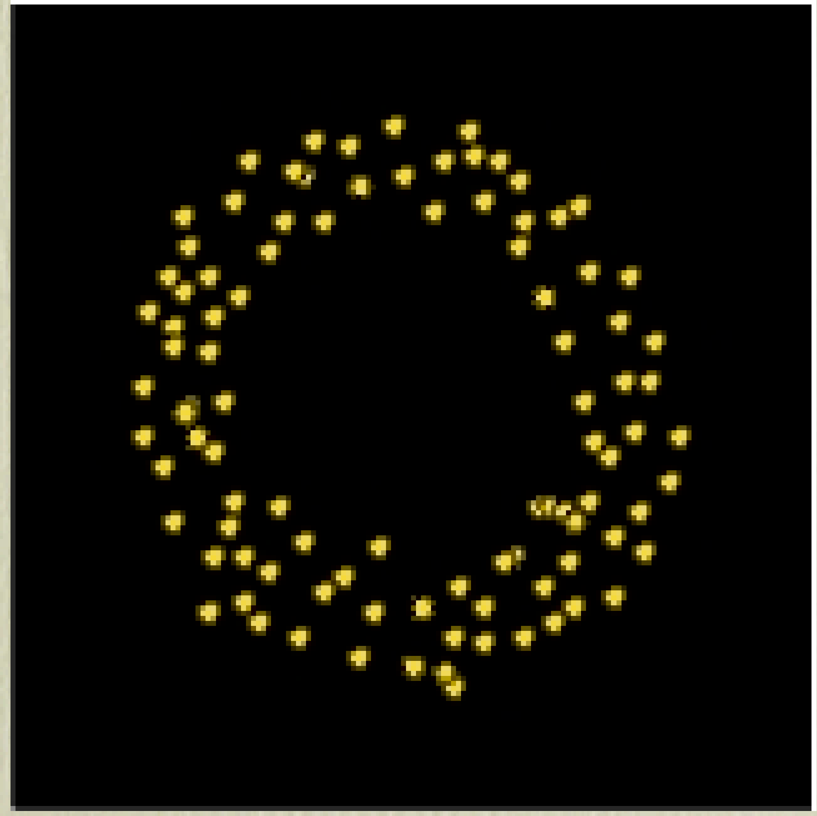
- $\partial^2 = 0$ (a boundary has no boundary)
- $\mathrm{Ker}(\partial)$ = cycle-space
- $\mathrm{Im}(\partial)$ = boundary-space
- $H_* = \mathrm{Ker}(\partial)/\mathrm{Im}(\partial)$ = homology
- $b_k = \dim(H_k)$



2. Topology of point-clouds

Point-cloud data

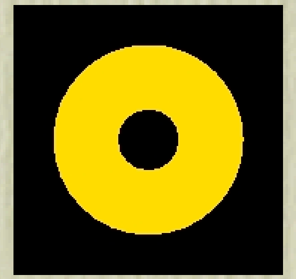
- Rather than a topological space, we have a cloud of data points.



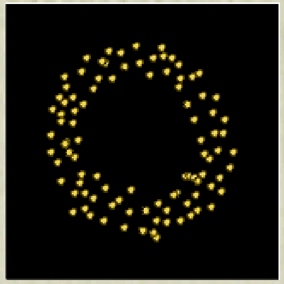
Space reconstruction...

Start here

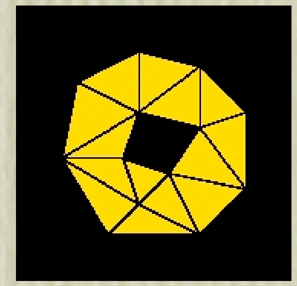
topological space



point-cloud dataset



simplicial complex

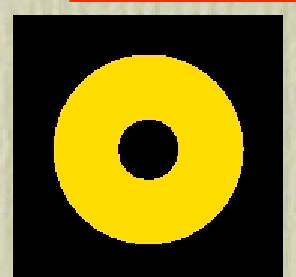


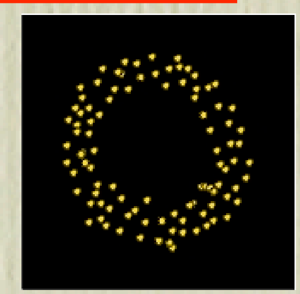
...or point-cloud topology?

Start here

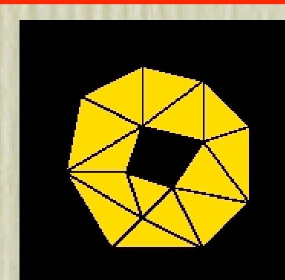
topological space

point-cloud dataset





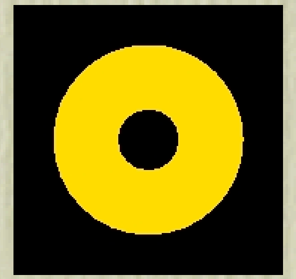
simplicial complex



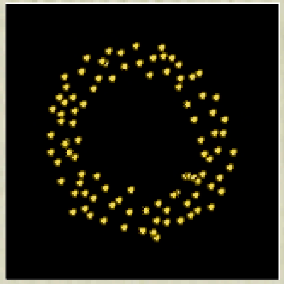
Space reconstruction

Start here

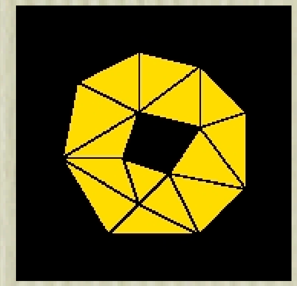
topological space



point-cloud dataset



simplicial complex

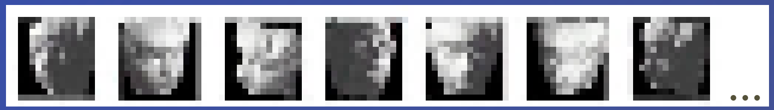


Vin de Silva
Stanford University

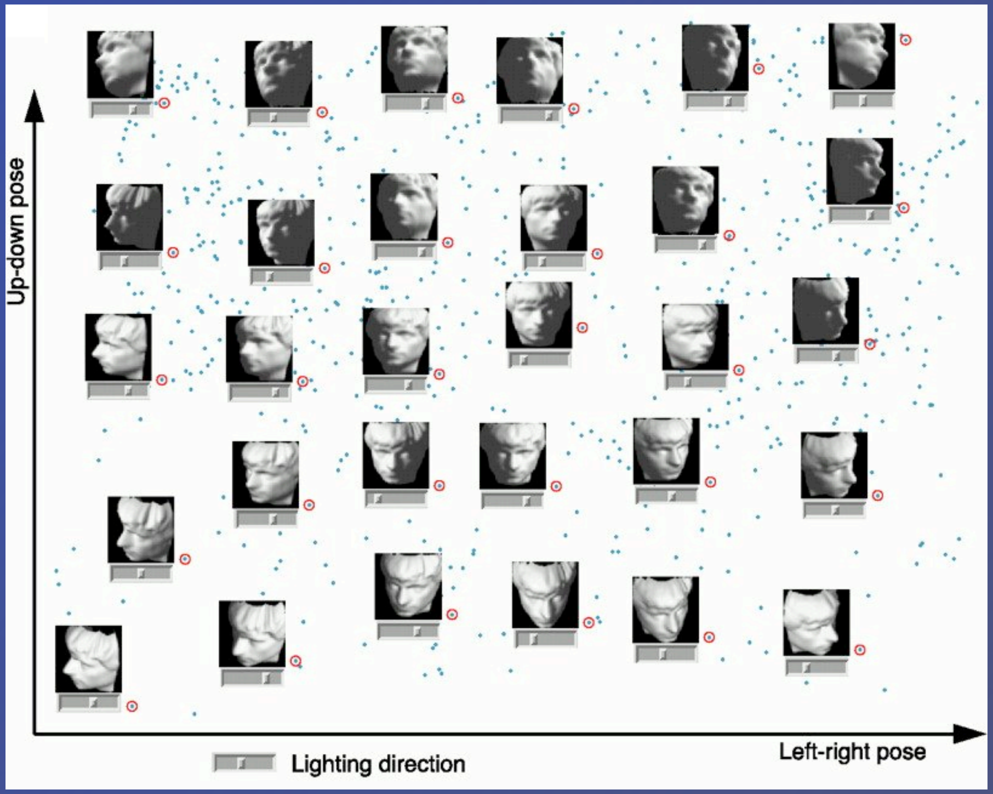
MGA Workshop III
IPAM, UCLA, October 25-29, 2004

NLDR (e.g. Isomap)

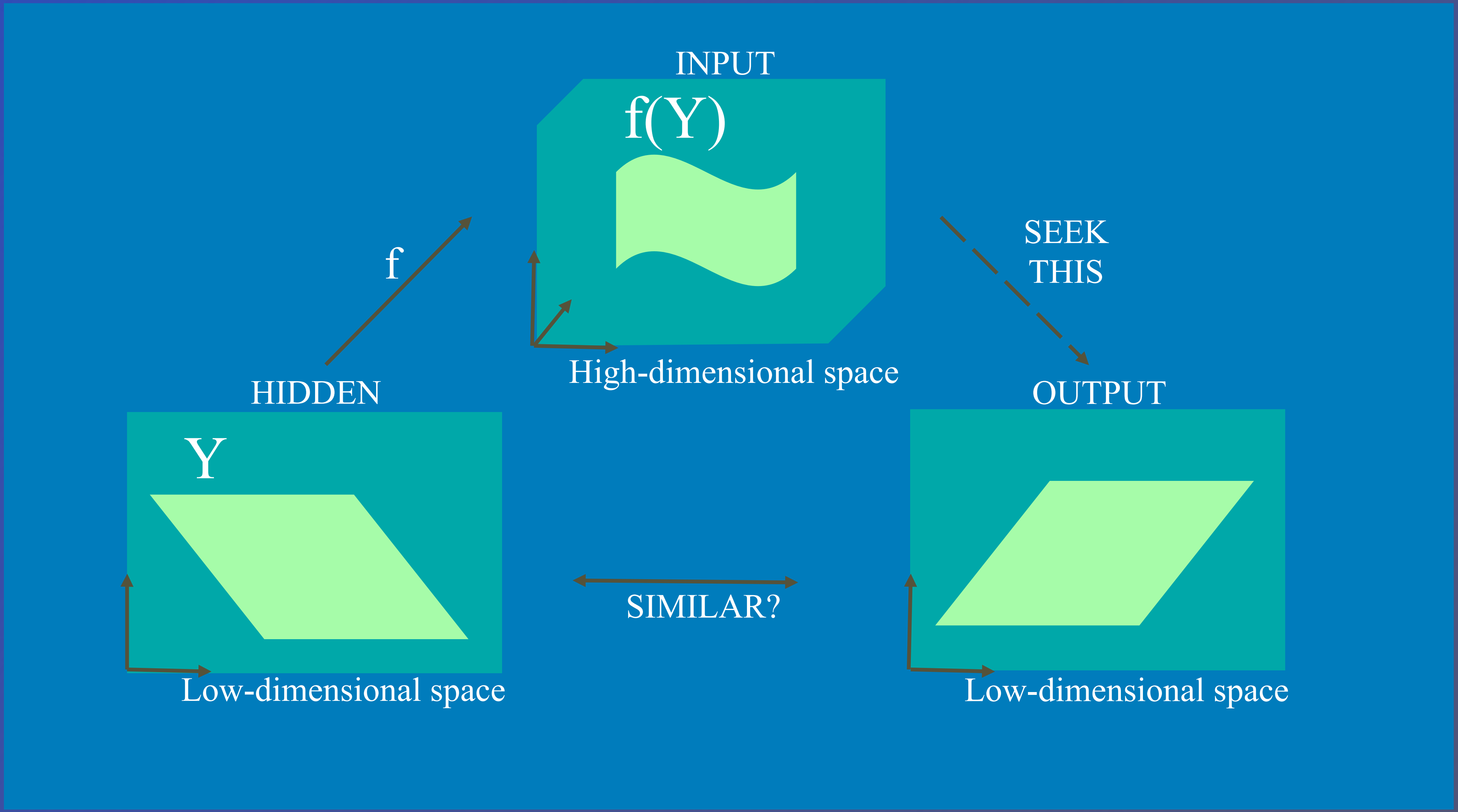
Input: randomly ordered sequence of images varied in pose and lighting.



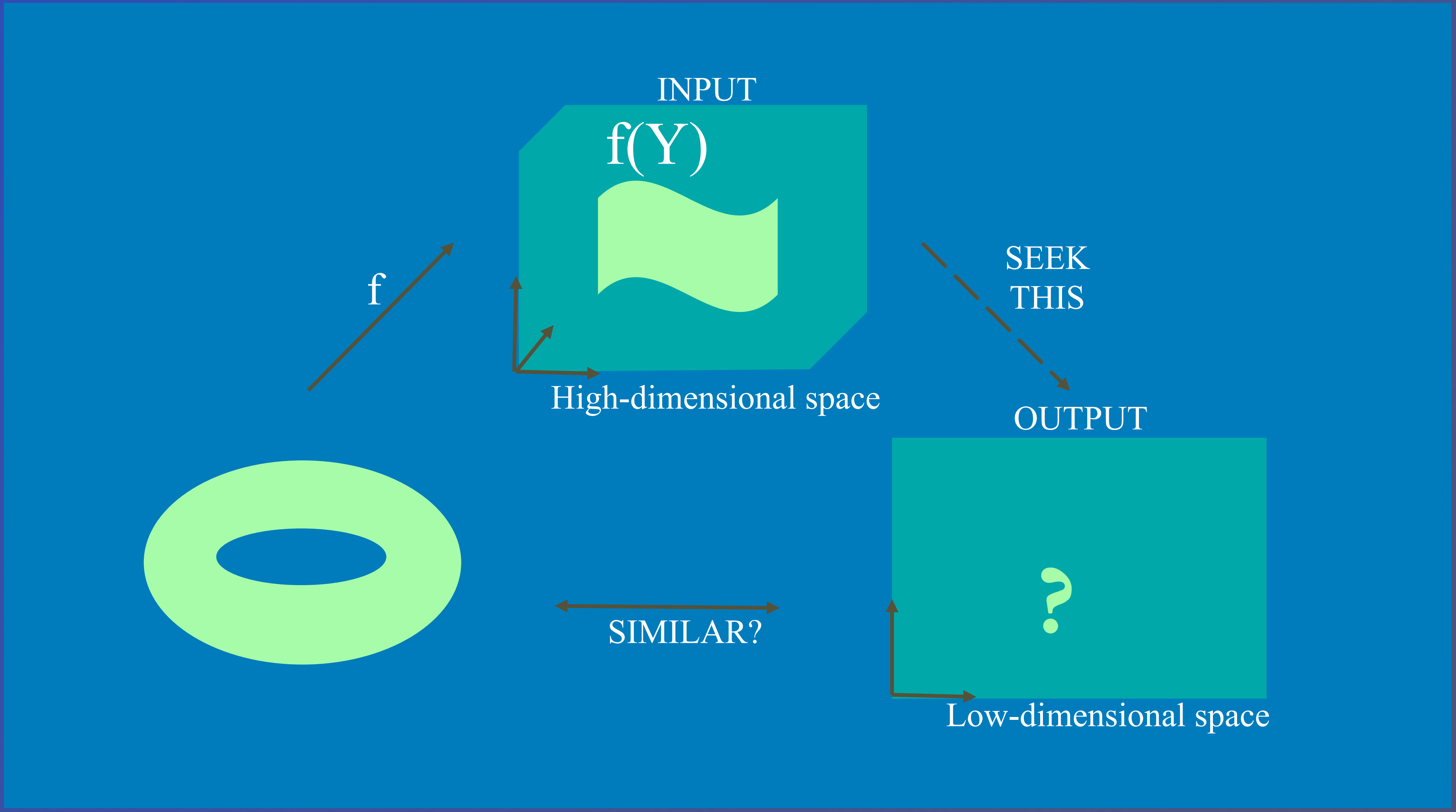
Output: low-dimensional embedding.



Generative model for NLDR



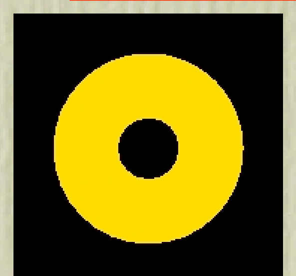
Non-euclidean topology

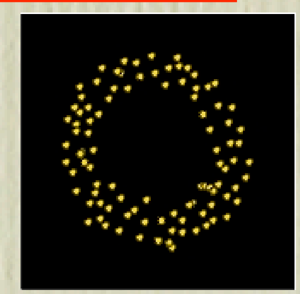


Space reconstruction...

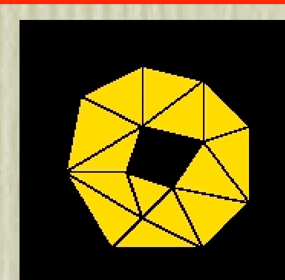
Start here

topological space → point-cloud dataset





simplicial complex



Reconstruction criterion

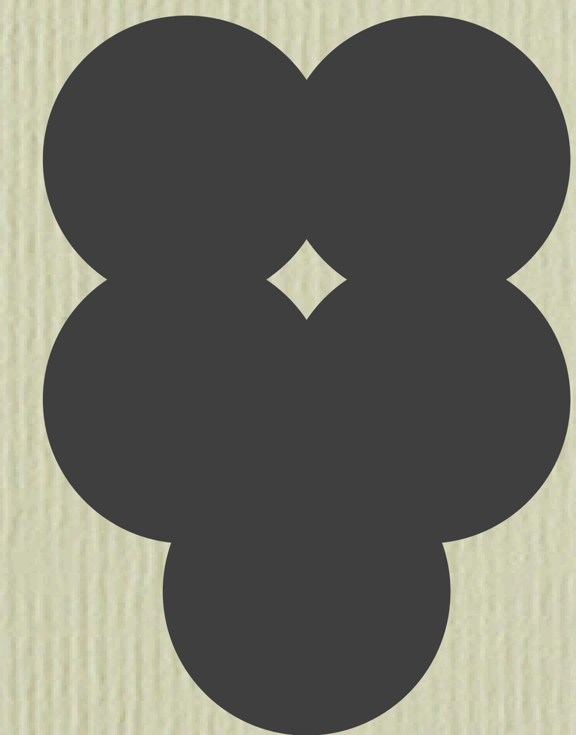
- In surface/manifold reconstruction, we ask that the simplicial complex and the hidden space be homeomorphic to each other.
- If the goal is to estimate Betti numbers, it is enough for them to be homotopy equivalent.
- For example, "nerve complexes" are amenable to proofs of homotopy equivalence.

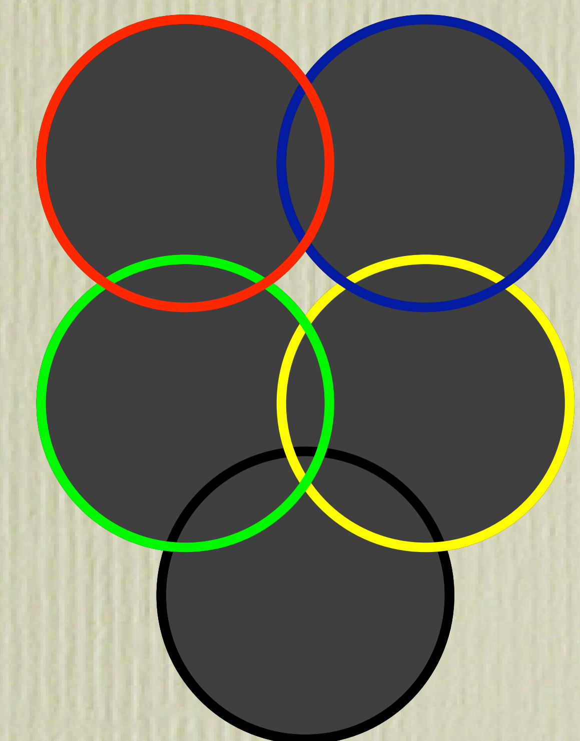
Nerve complexes

- Let $X = U_1 \cup \ldots \cup U_n$ be a space (or set) expressed as union of subspaces (or subsets). The Nerve complex is defined to have:
  - a vertex {i} for every i such that $U_i \neq \emptyset$;
  - an edge {ij} whenever $U_i \cap U_j \neq \emptyset$;
  - a triangle {ijk} whenever $U_i \cap U_j \cap U_k \neq \emptyset$;
  - and so on.

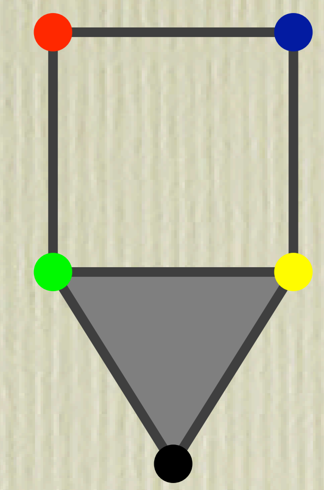
Example

X





Nerve



$U_1 \cup U_2 \cup U_3 \cup U_4 \cup U_5$

The Čech nerve theorem

- [This is the basis of Čech (co-)homology]
  If every finite intersection of the sets Ui is empty or contractible, then the Nerve complex and X have the same homotopy type.

Example: Čech complex

- Let $R > 0$. Define Čech(X,R) has:
  - a vertex $[x]$ for every data point x in X;
  - an edge $[xy]$ if $|x-y| < 2R$;
  - a triangle $[xyz]$ if the three balls with centres x,y,z and radius R have a non-empty common intersection;
  - and so on, for higher dimensional cells.

Čech complex as a nerve

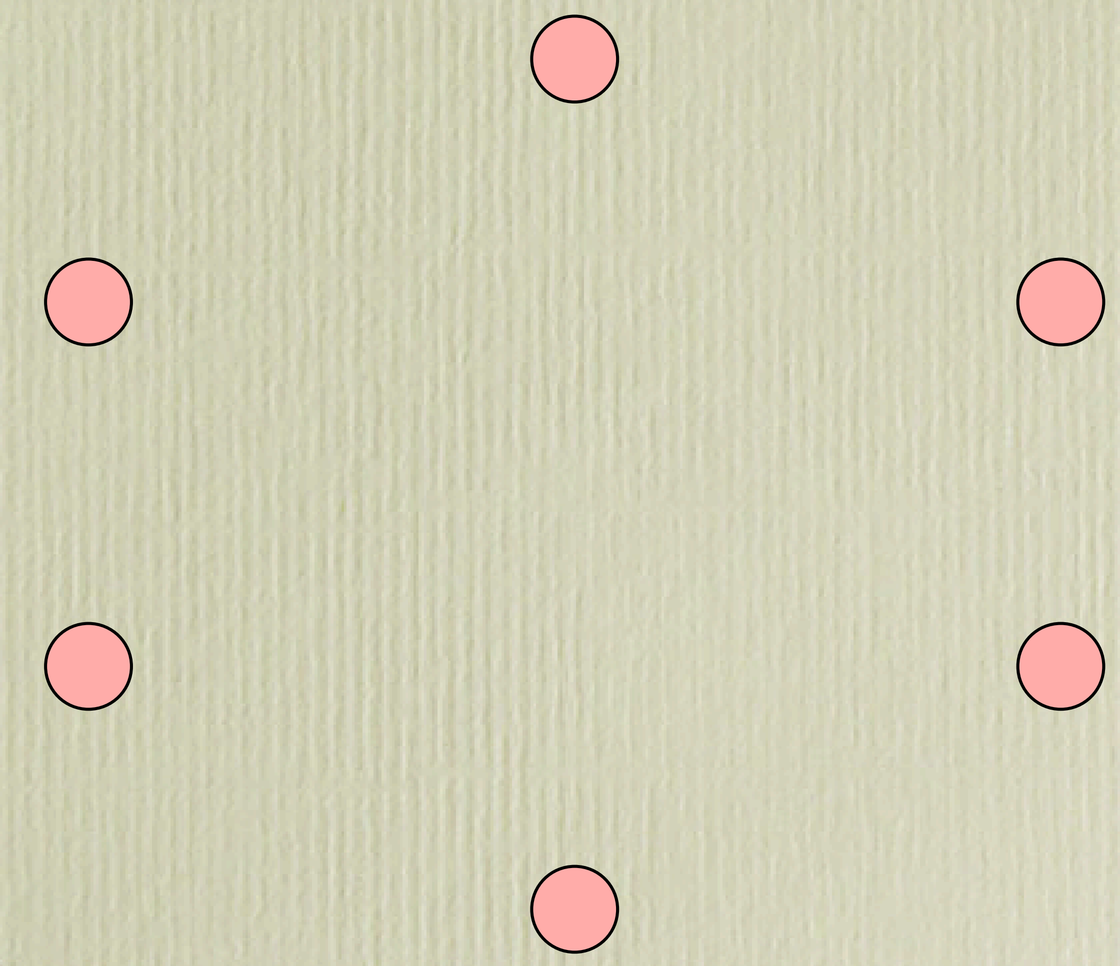
- Čech(X,R) is the nerve of the union of balls of radius R centered at the points x of X.

Reconstruction Theorem

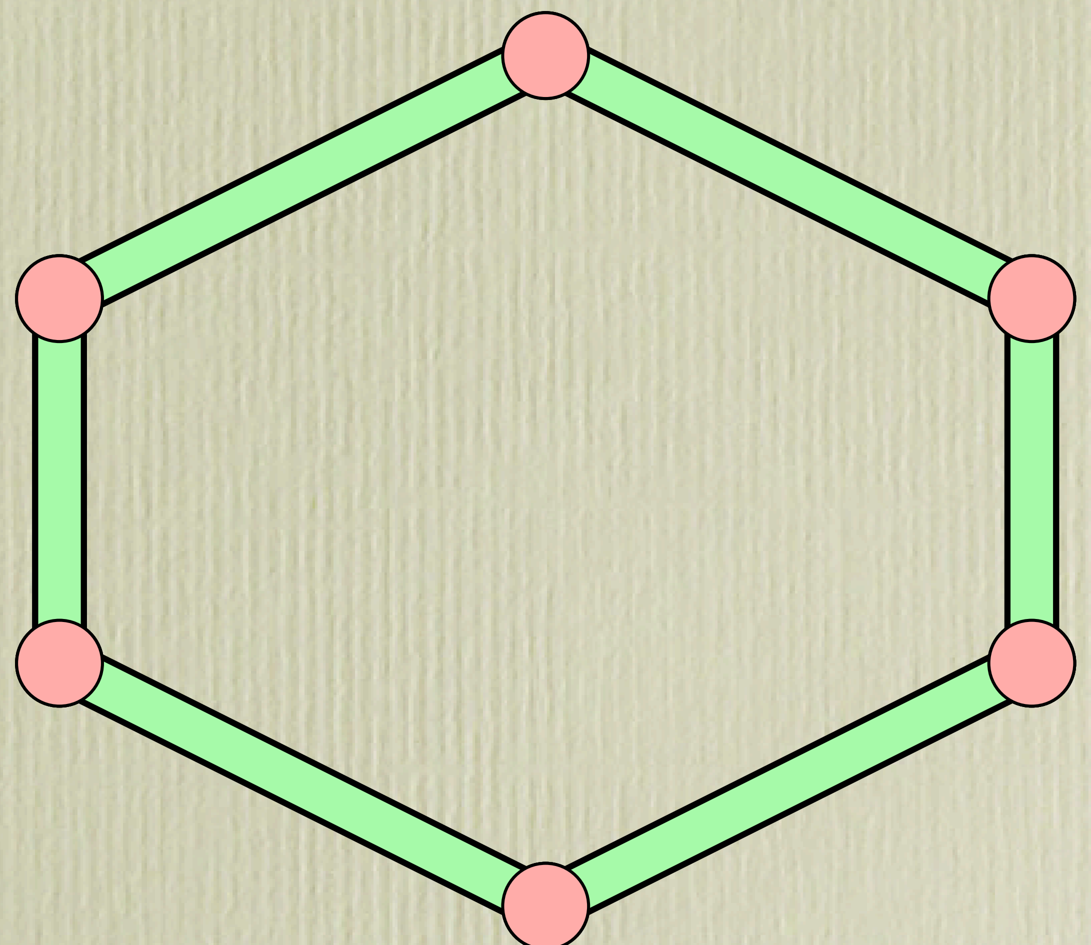
- [Niyogi, Smale, Weinberger, 2004]

  Let $M \subset \mathbf{R}^n$ be a smooth submanifold with feature size τ. For any $0 < R < \tau\sqrt{(3/5)}$, suppose $X \subset M$ is a finite sample which is $(R/2)$-dense in M. Then Čech(X,R) has the same homotopy type as M.

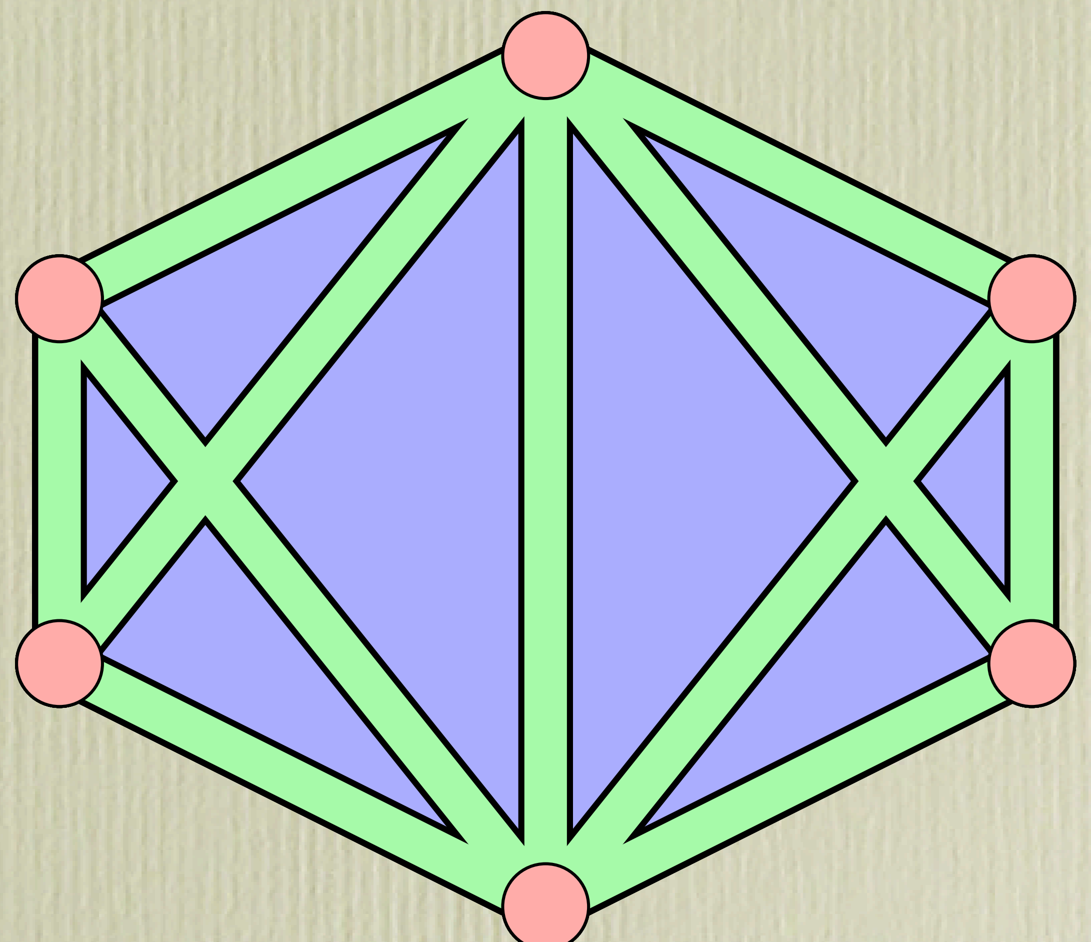
How to choose R?



How to choose R?



How to choose R?



3. Persistent homology

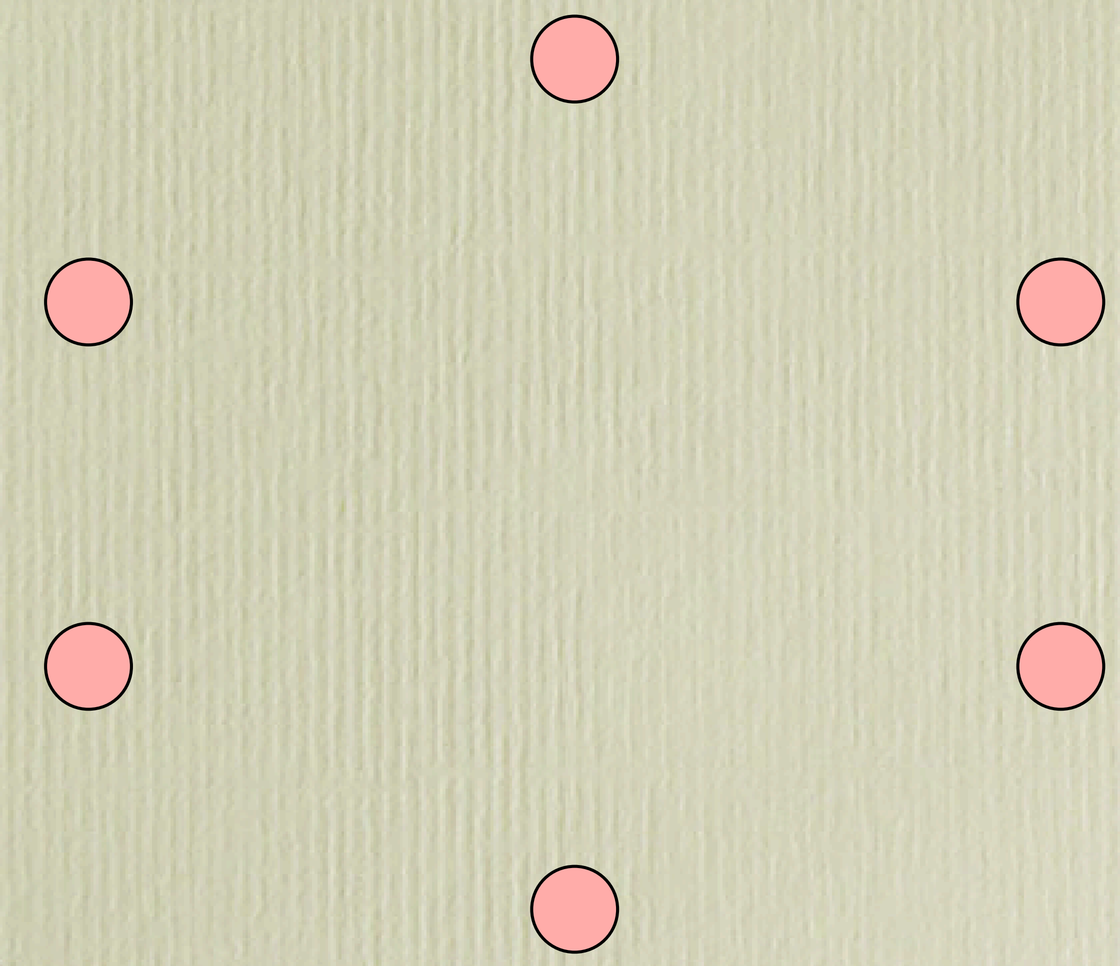
Dichotomy

- Homology groups and Betti numbers are discrete quantities.
- The world of data sets is continuous.
- How can we maintain the (useful, qualitative) discrete flavour of homology, while taking into account the continuous flavour of real data?

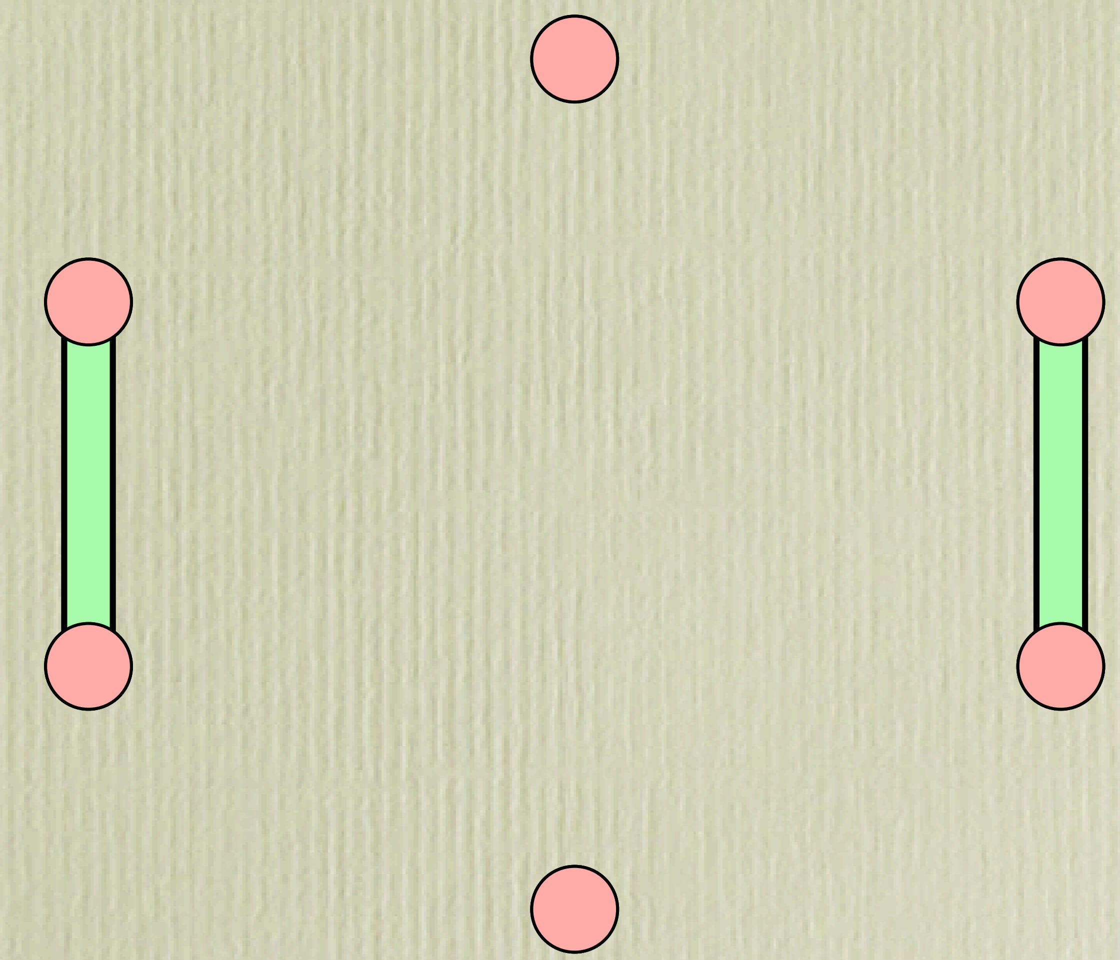
Persistent homology

- Instead of computing Betti numbers for each value of R, combine the calculations for all values of R simultaneously.
- Edelsbrunner, Delfinado, Zomorodian (2000) give a strikingly effective algorithm for computing persistent homology.
- The output takes the form of an "interval graph", where each interval represents the lifetime of a feature.

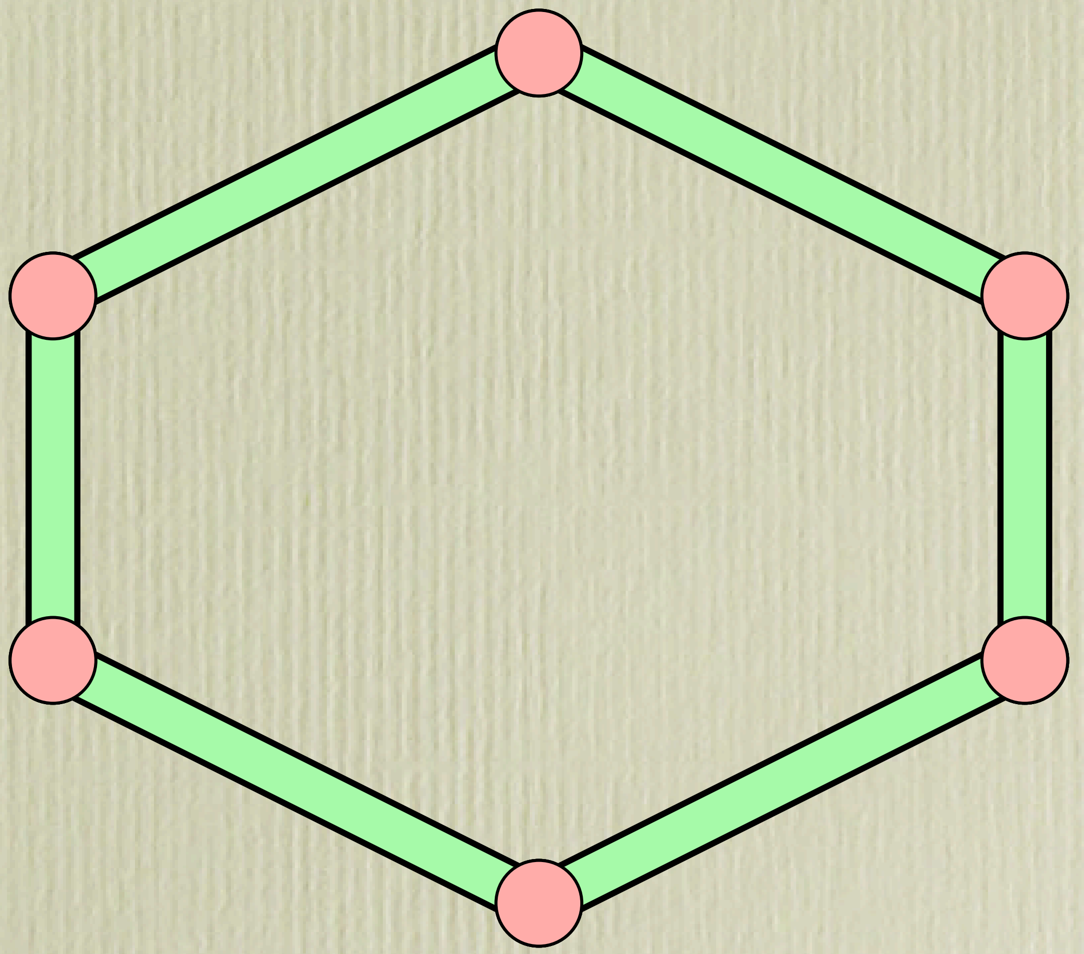
Filtered complex parametrised by R



Filtered complex parametrised by $\mathbf{R}$

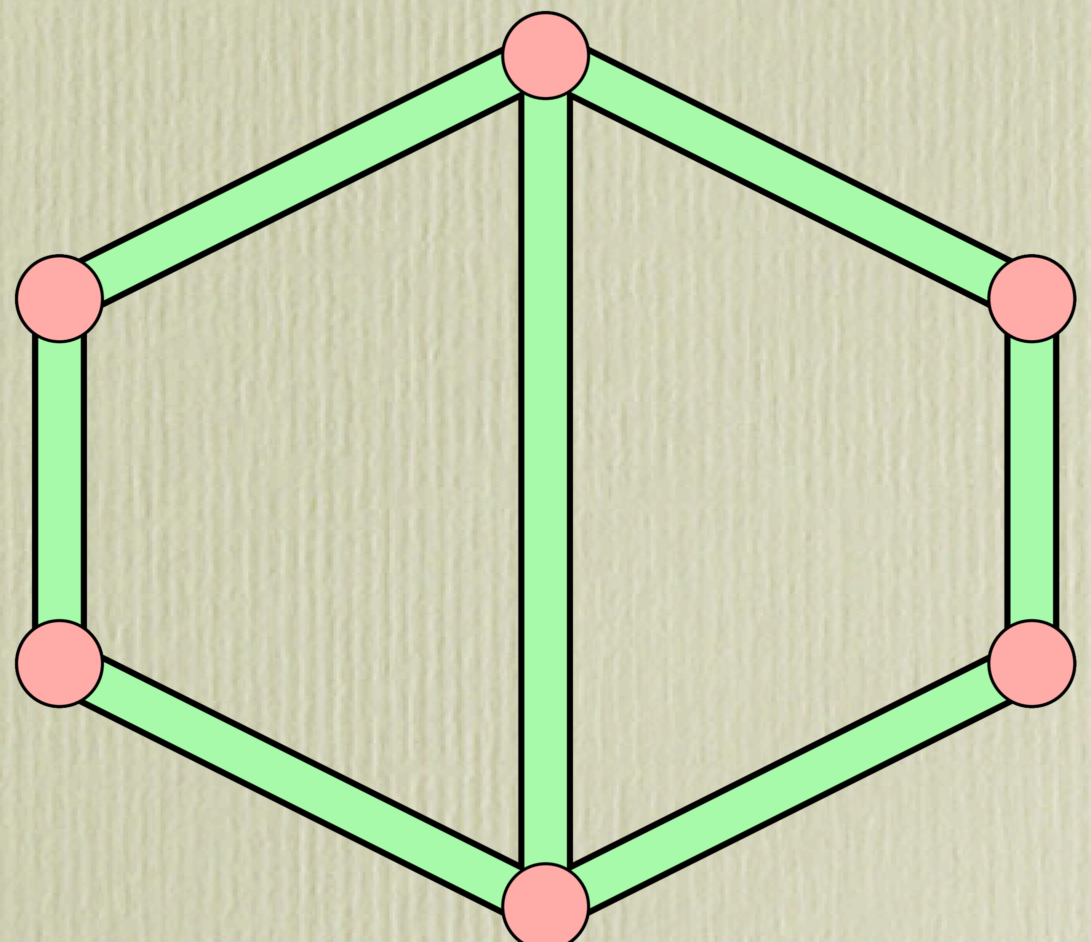


Filtered complex parametrised by R

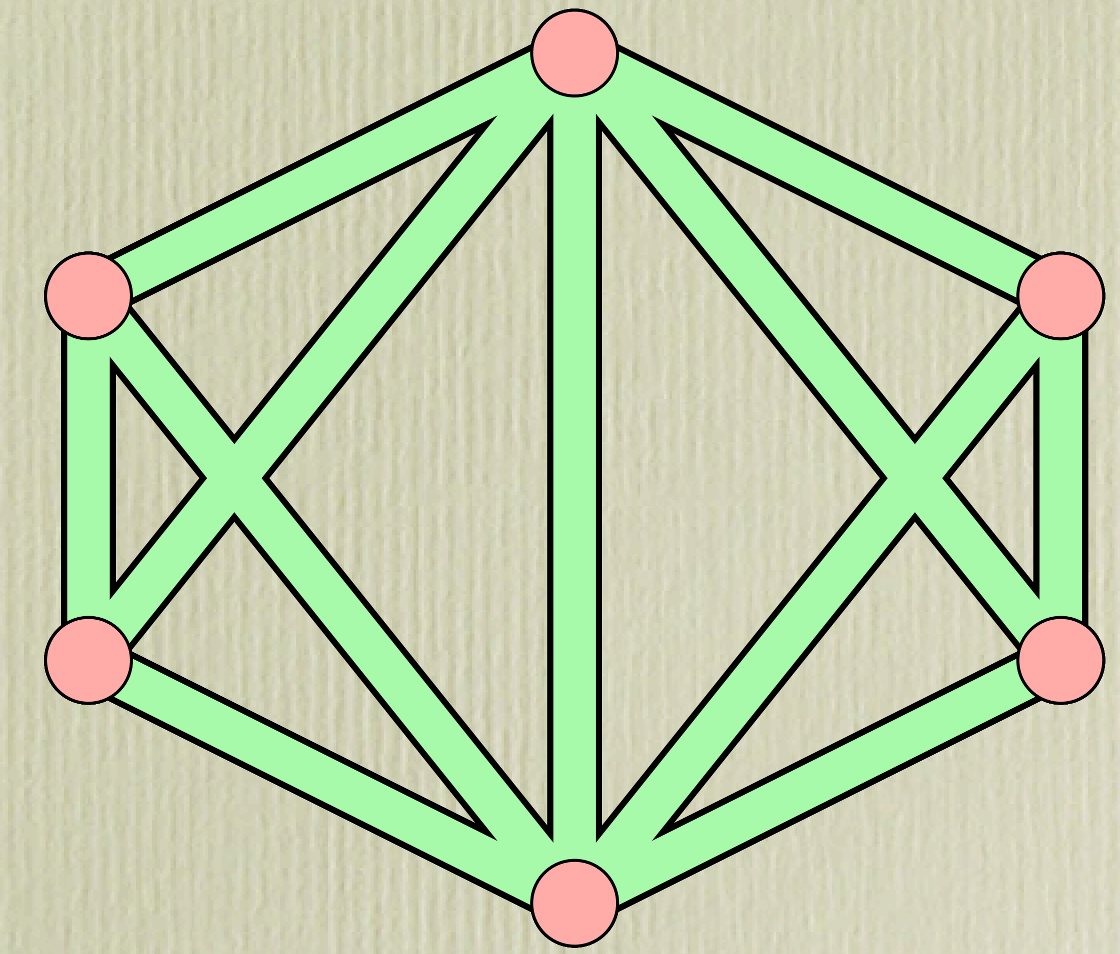


Vin de Silva
Stanford University

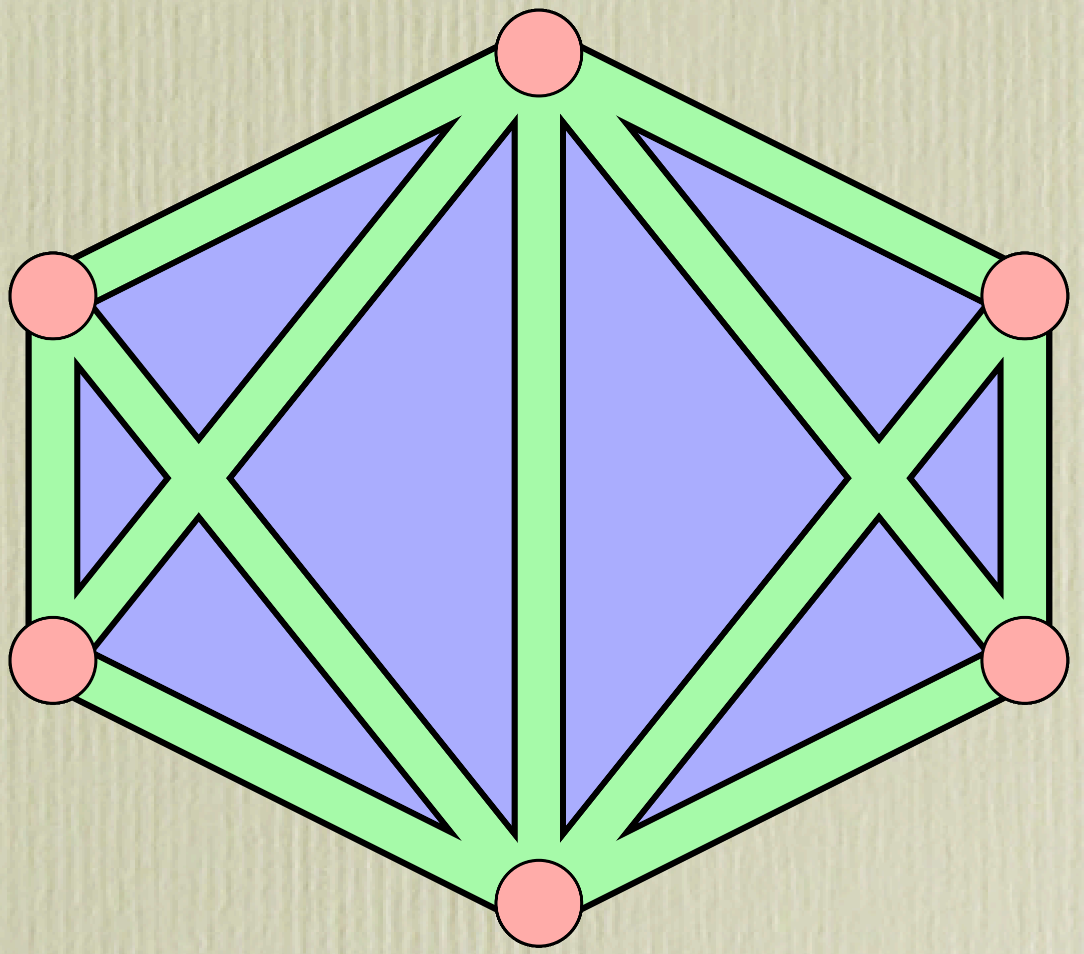
Filtered complex parametrised by $\mathbf{R}$



Filtered complex parametrised by R

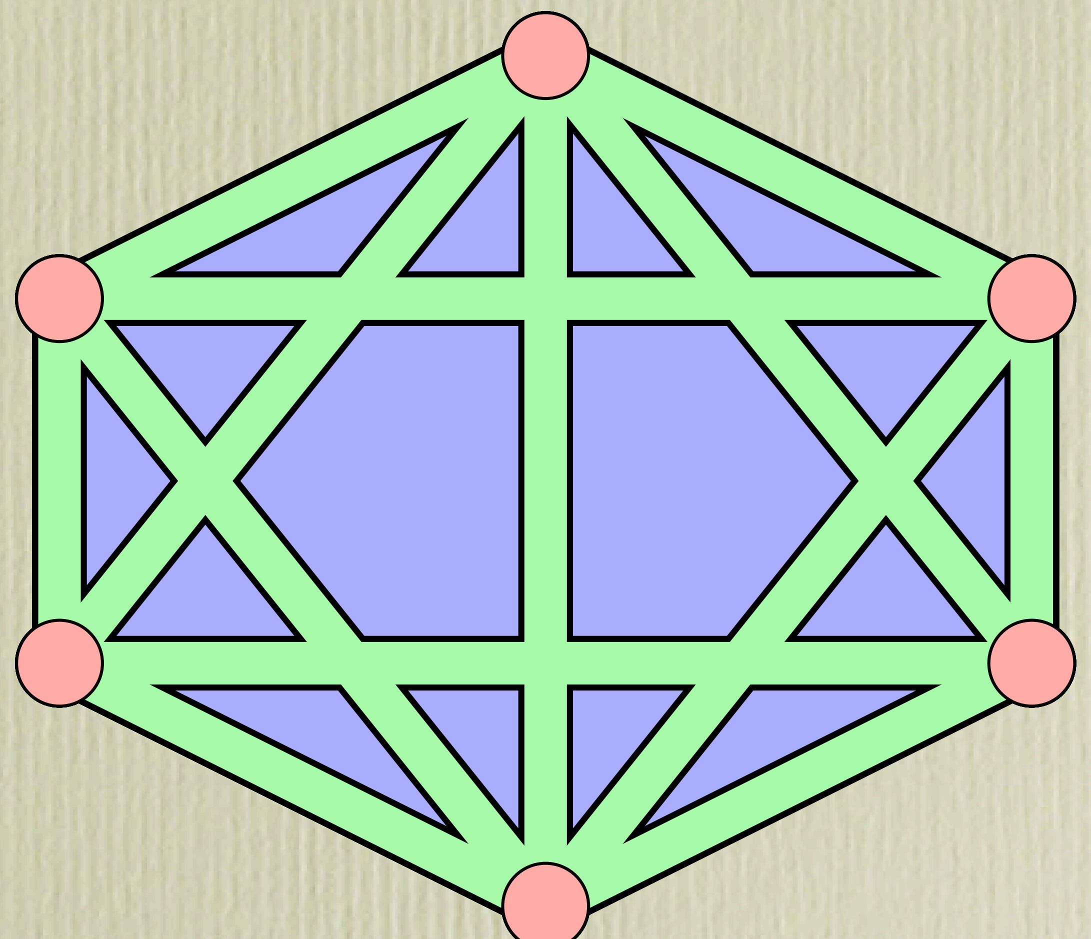


Filtered complex parametrised by R



Vin de Silva
Stanford University

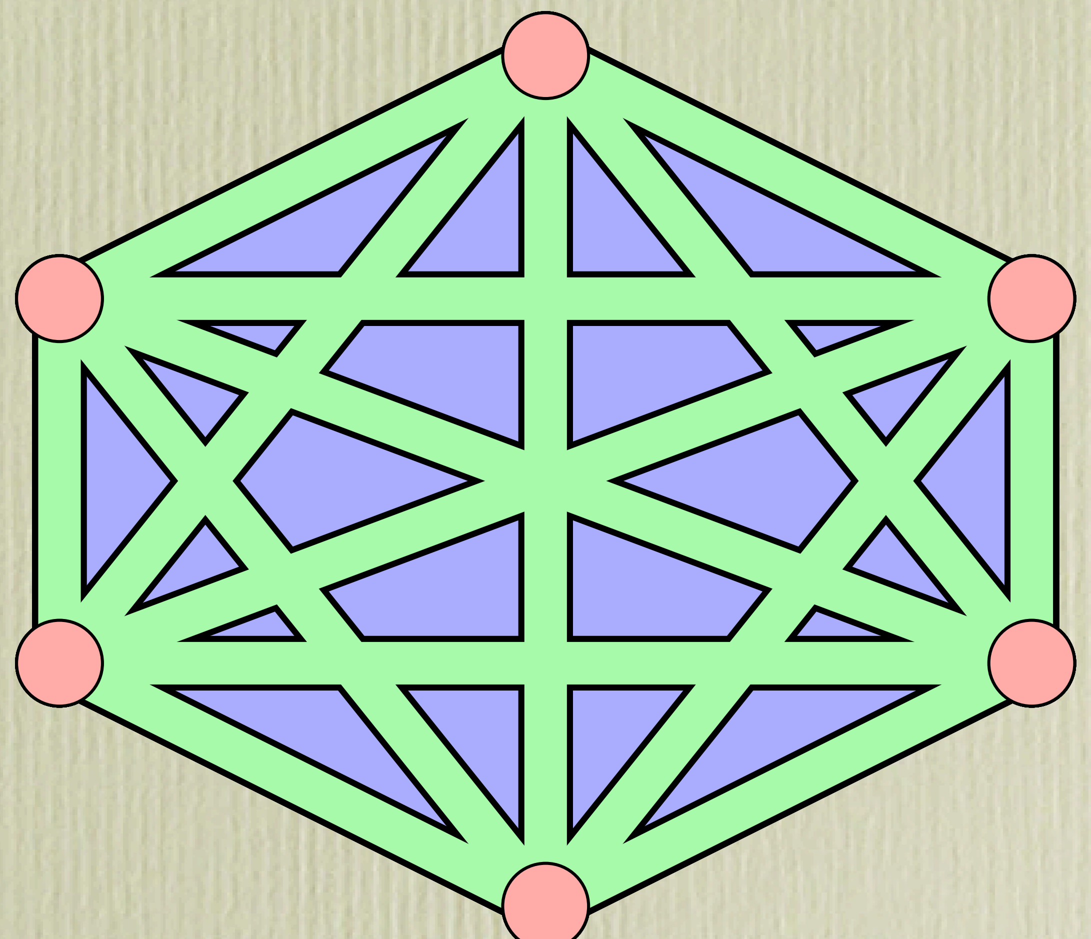
Filtered complex parametrised by $\mathbf{R}$



Vin de Silva
Stanford University

MGA Workshop III
IPAM, UCLA, October 25-29, 2004

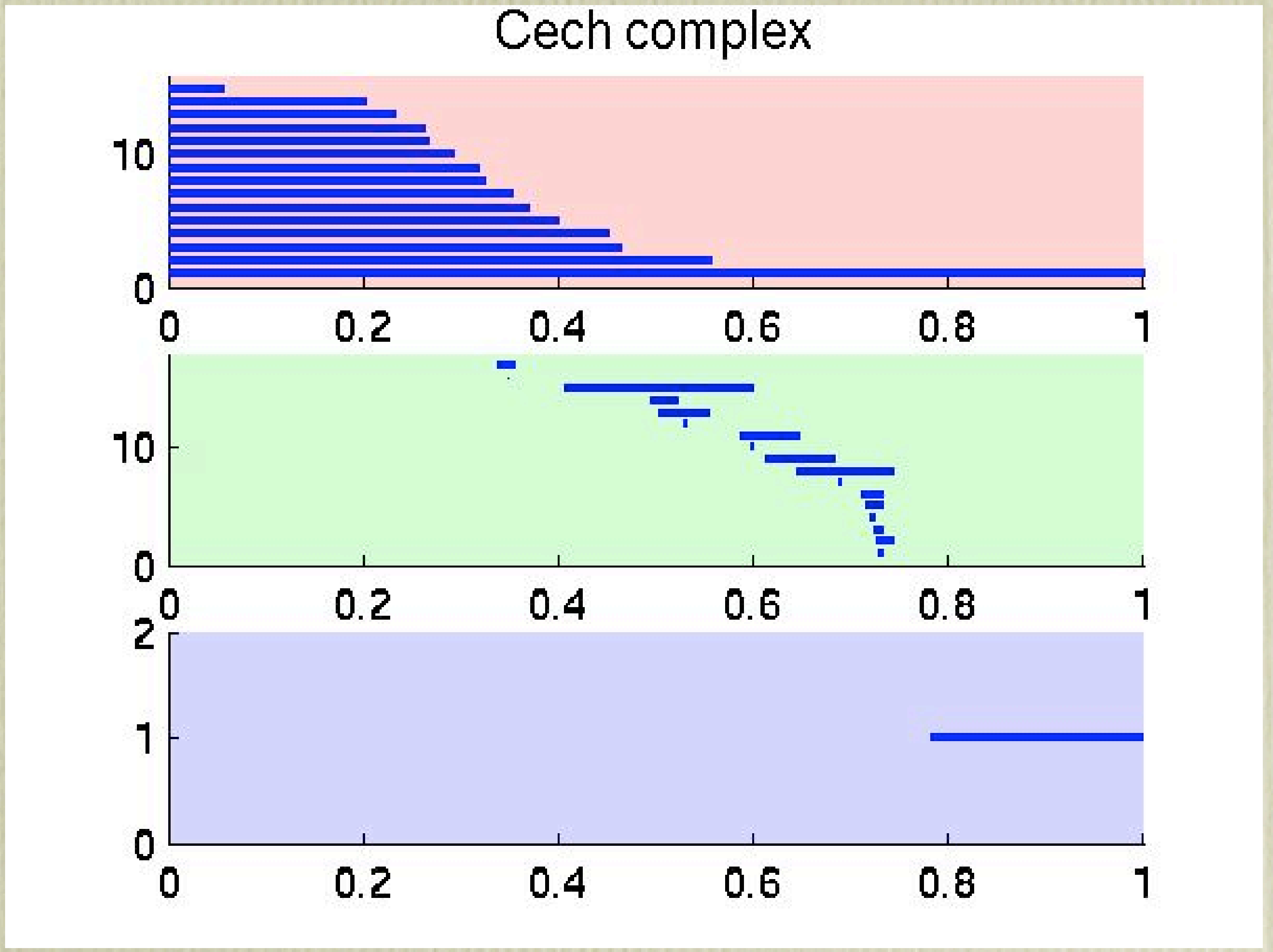
Filtered complex parametrised by $\mathbb{R}$



Vin de Silva
Stanford University

MGA Workshop III
IPAM, UCLA, October 25-29, 2004

Example of an interval graph



b0

b1

b2

Theoretical interpretation

- [Carlsson, Zomorodian, 2003]
  This kind of interval graph structure occurs whenever you have a sequence of complexes with maps $S_1 \to S_2 \to \dots \to S_n$.

- Ordinary homology uses coefficients over the field $\mathbf{Z}_2$ (for example).

- Persistent homology uses coefficients over the polynomial ring $\mathbf{Z}_2[t]$. This has a well-behaved module theory.

Outstanding open problem

- With a single filtration parameter, persistent homology works beautifully.
- With two independent filtration parameters, the corresponding polynomial ring $\mathbf{Z}_2[s,t]$ has a horribly complicated module theory.
- How should one handle these situations?

4. Witness complexes

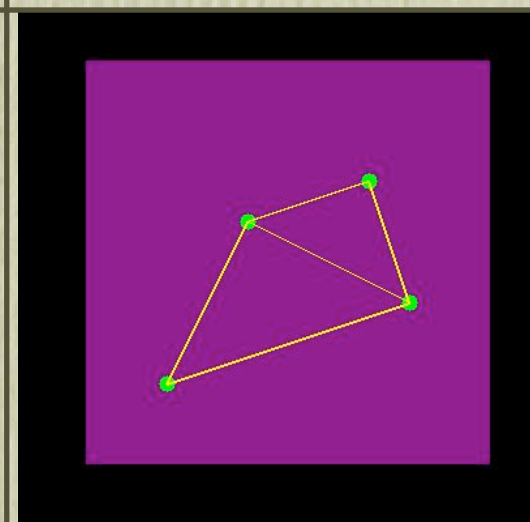
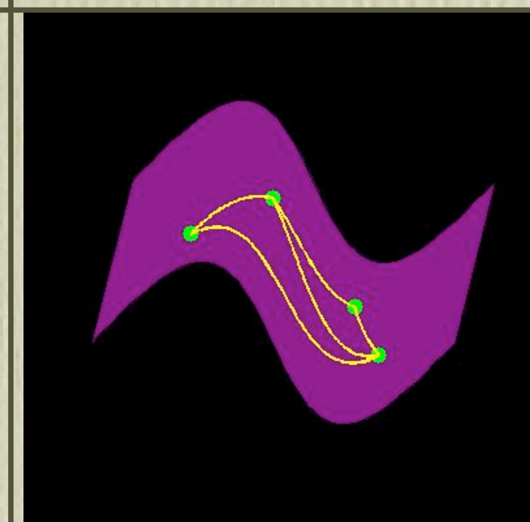
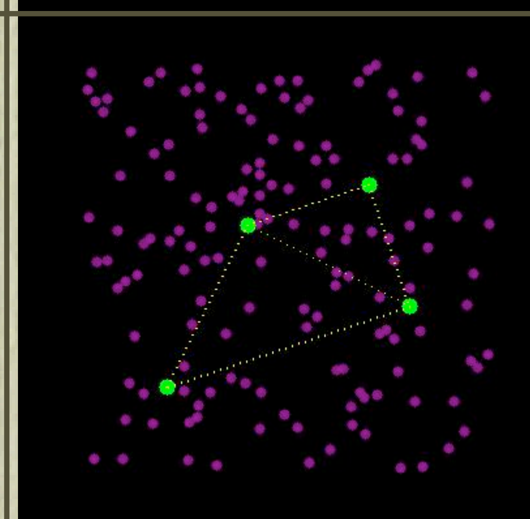
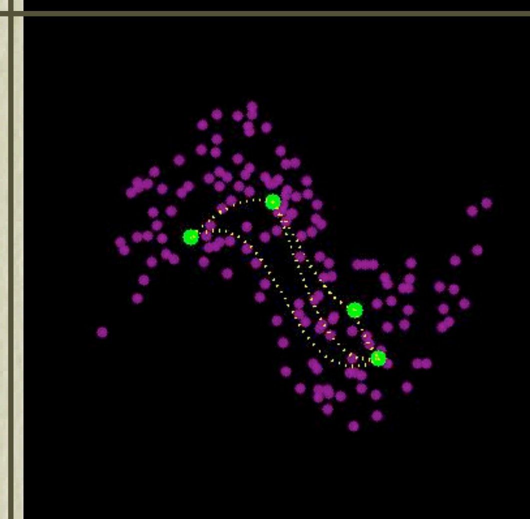
In search of efficiency

- The Čech complex has good homotopy properties. However, the number of cells becomes huge as R grows.
- The Alpha-shape complex [Edelsbrunner, 1995] has the same homotopy type with far fewer cells. Based on Delaunay triangulation: curse of dimensionality (extrinsic).
- Can we avoid this trade-off?

Strategy

- [Carlsson, VdS, 2003]
  Strong & Weak witness complexes.
- Use a small subset of the data as the vertex set.
- Simplices should lie close to existing data points (rather than cutting across chasms).
- (Cheaply) mimic the restricted Delaunay triangulation, in a point-cloud data setting.

4 paradigms

	flat	curved
continuous		
point cloud		

4 paradigms

	flat	curved
manifold	Delaunay triangulation	restricted Delaunay triangulation
point cloud	?	?

4 paradigms

	flat	curved
manifold	Delaunay triangulation	restricted Delaunay triangulation
point cloud	weak/strong witness complex	weak/strong witness complex

Strategy

- Given large point-cloud data set X, choose a much smaller set L of vertices.
- L can be chosen randomly or using a greedy optimisation strategy for good coverage.
- The number of landmark points constrains the complexity of the detectable topology. Fewer may be better.

Delaunay triangulation

- Let $L \subset \mathbf{R}^n$ be a finite set of points and let $x_0, x_1, \ldots, x_k \in L$. Then TFAE:
  - $x_0, x_1, \ldots, x_k$ span a Delaunay k-cell;
  - the Voronoi cells for $x_0, x_1, \ldots, x_k$ meet;
  - there is a point $w \in \mathbf{R}^n$, whose k+1 nearest neighbours in L are $x_0, x_1, \ldots, x_k$, and which is equidistant from them.

Restricted Delaunay triangulation

- Let L be a set of points in a manifold $M \subset \mathbf{R}^n$ and let $x_0, x_1, \ldots, x_k \in L$. Then TFAE:
  - $x_0, x_1, \ldots, x_k$ span a restricted Delaunay k-cell;
  - the Voronoi cells for $x_0, x_1, \ldots, x_k$ meet in M;
  - there is a point $w \in M$, whose $k+1$ nearest neighbours in L are $x_0, x_1, \ldots, x_k$, and which is equidistant from them.

Strong witness complex

- Let L be a set of points taken from a finite set $X \subset M \subset \mathbf{R}^n$ and let $x_0, x_1, \ldots, x_k \in L$. We decree that $x_0, x_1, \ldots, x_k$ span a k-cell in the strong witness complex if and only if:
  - There is a point $w \in X$, whose k+1 nearest neighbours in L are $x_0, x_1, \ldots, x_k$; and
  - w is equidistant from $x_0, x_1, \ldots, x_k$.

Immediate disaster!

- The existence of the point w in the finite set X is a ‘probability zero’ event.
- Need to introduce a tolerance parameter R, and interpret the definition “up to error R”.

Strong nerves (1)

- Strong(X,L) can be defined as follows.
  - Let $f : X \to \mathbf{R}^n$ map x in X to the vector of its distances to the n landmarks.
  - Partition the positive quadrant of $\mathbf{R}^n$ into sets $V_i = \{\mathbf{v} : v_i$ is the smallest coordinate$\}$.
  - Let $U_i = f^{-1}(V_i)$ and construct the nerve.

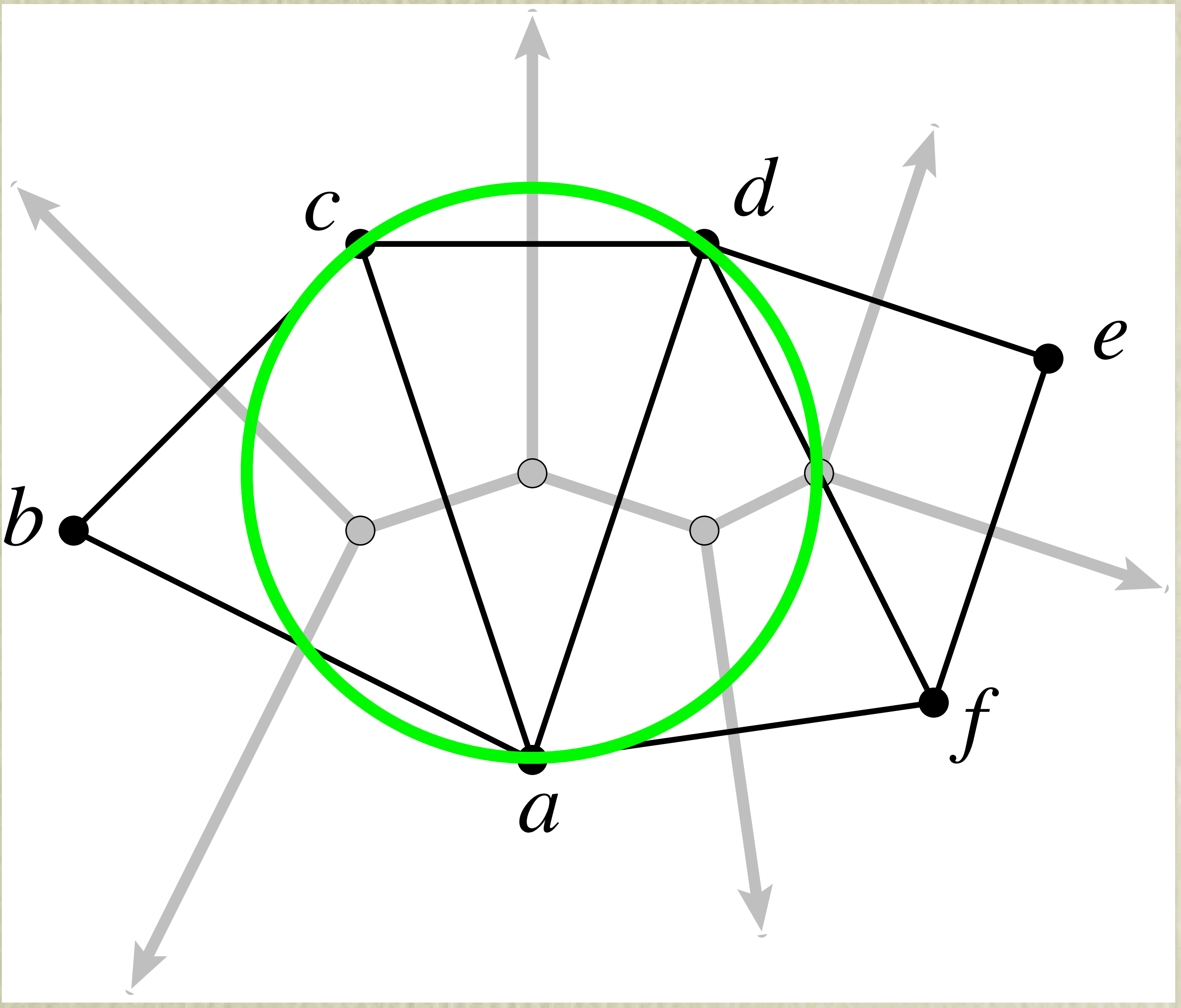
Strong nerves (2)

- Strong(X,L,R) can be defined as follows.
  - Thicken $V_i \subset \mathbf{R}^n$ to its R-neighbourhood $V_i(R)$ with respect to a suitable metric on $\mathbf{R}^n$.
  - The l_∞ (supremum) norm is convenient.
  - Construct the nerve with $U_i(R) = f^{-1}(V_i(R))$.

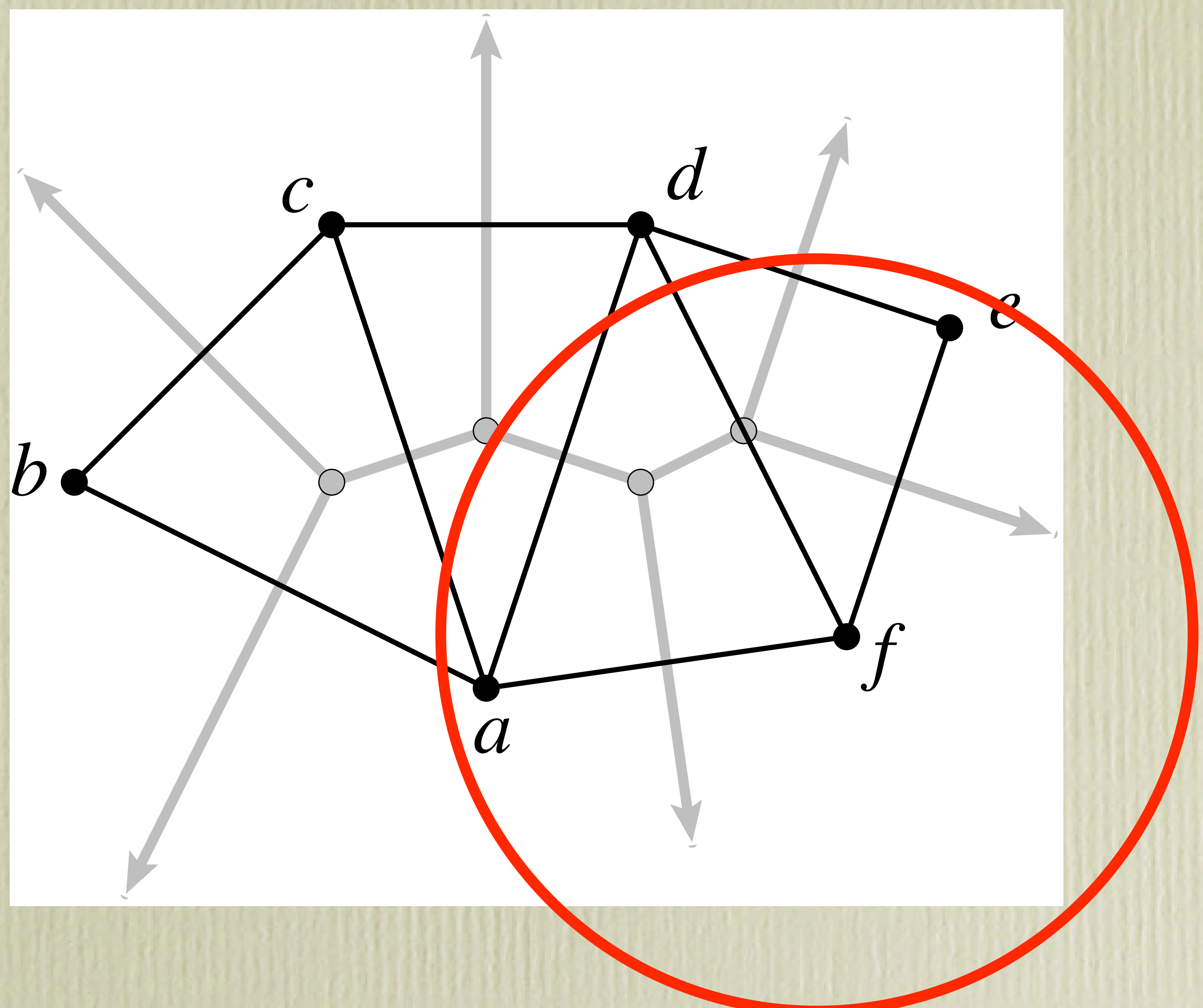
Strong and weak witnesses

- Consider again the following statement:
  - there is a point $w \in R^n$, whose $k+1$ nearest neighbours in L are $x_0, x_1, \ldots, x_k$, and <u>which is equidistant from them</u>.
- Such a point w is called a strong witness for the simplex $\{x_0, x_1, \ldots, x_k\}$. If we drop the equidistance condition, we say that w is a weak witness for $\{x_0, x_1, \ldots, x_k\}$.

Example



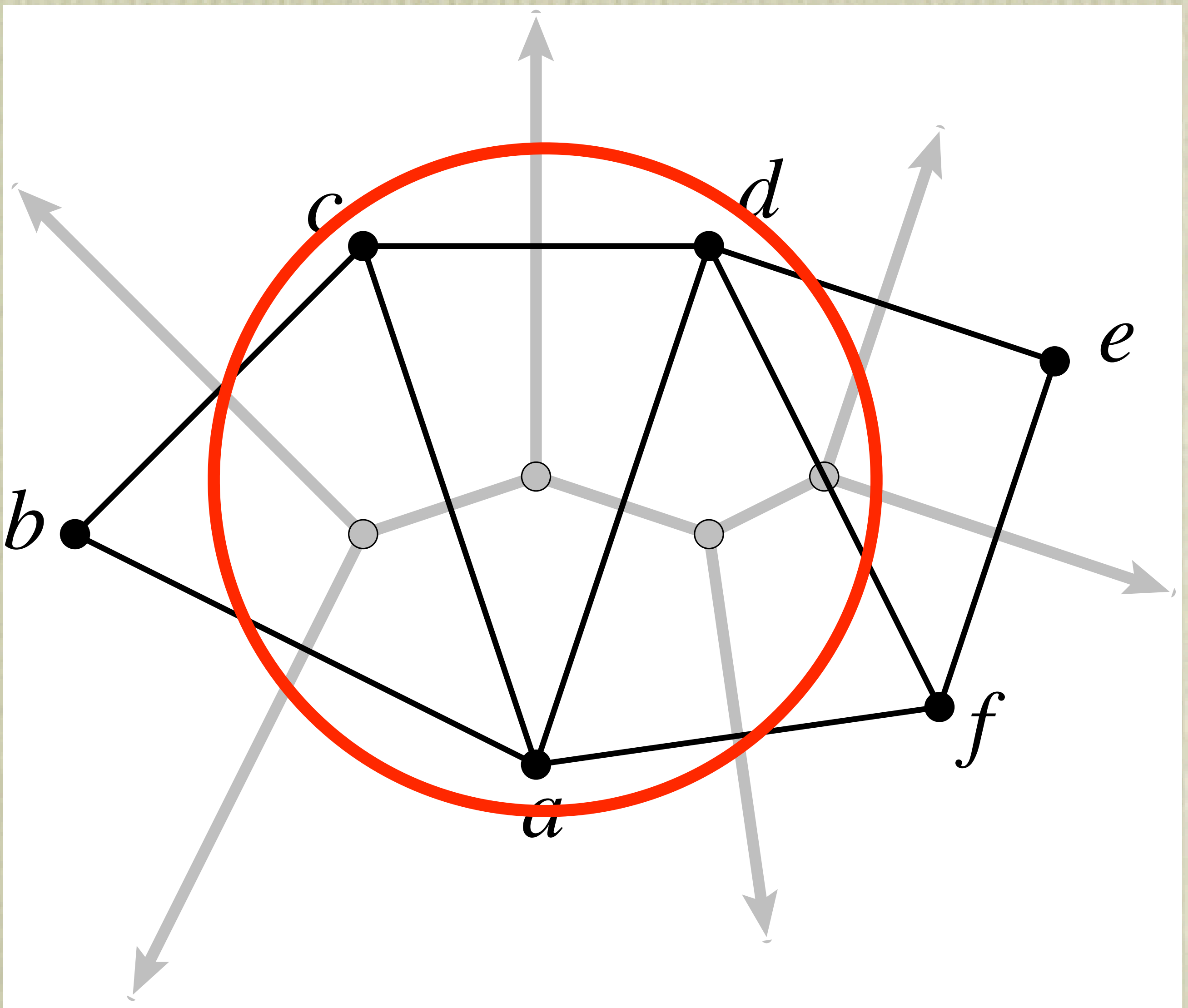
Example



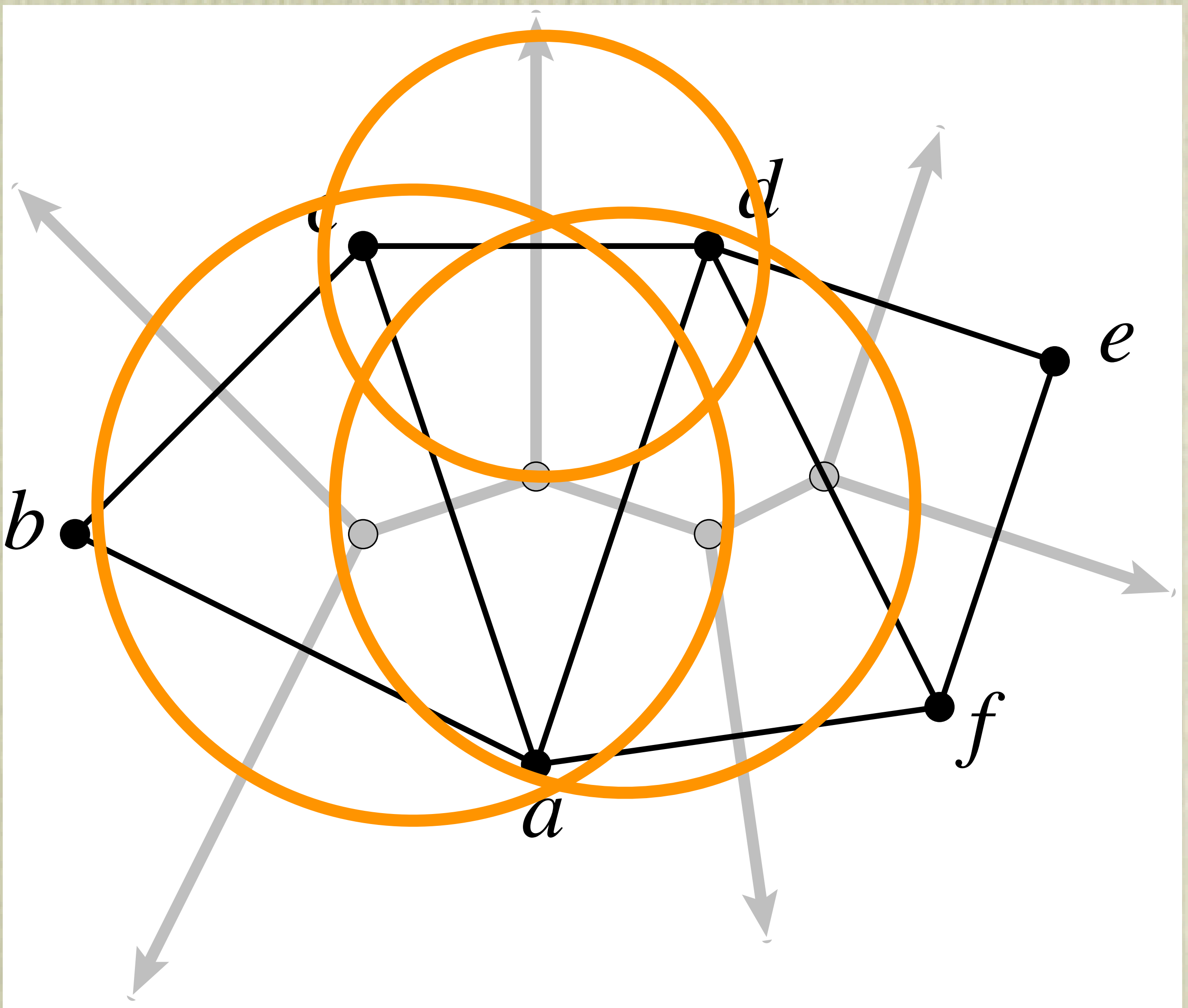
The weak witnesses theorem

- [VdS, 2003] Let $L \subset \mathbf{R}^n$ be a finite set of points and let $x_0,x_1,\ldots,x_k \in L$. Then $[x_0,x_1,\ldots,x_k]$ has a strong witness in $\mathbf{R}^n \Leftrightarrow [x_0,x_1,\ldots,x_k]$ and all of its subsimplices have weak witnesses in $\mathbf{R}^n$.
- For edges, this is well known. Exploited by Martinetz & Schulten (1994) to build topology-representing graphs.

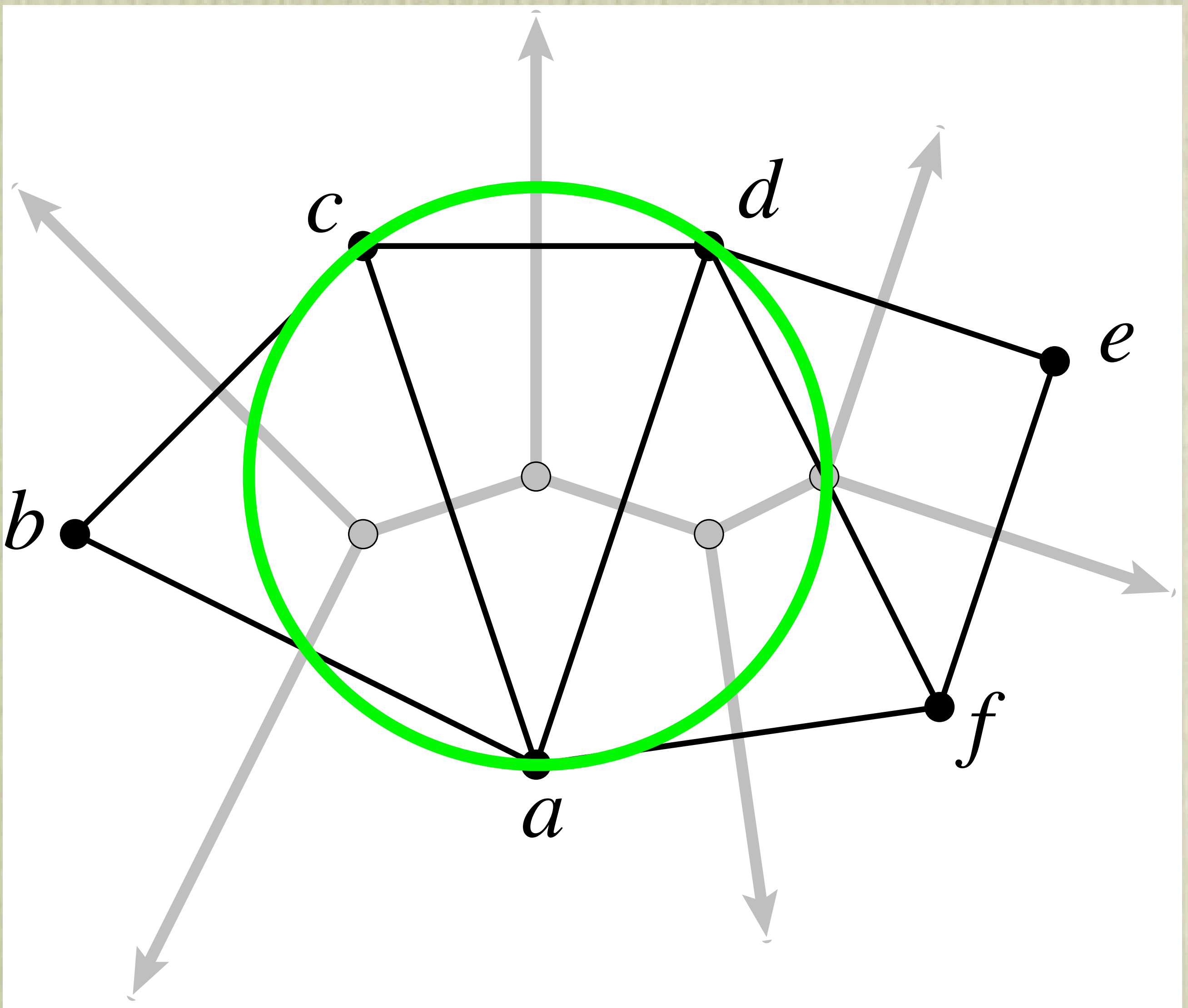
Example (continued)



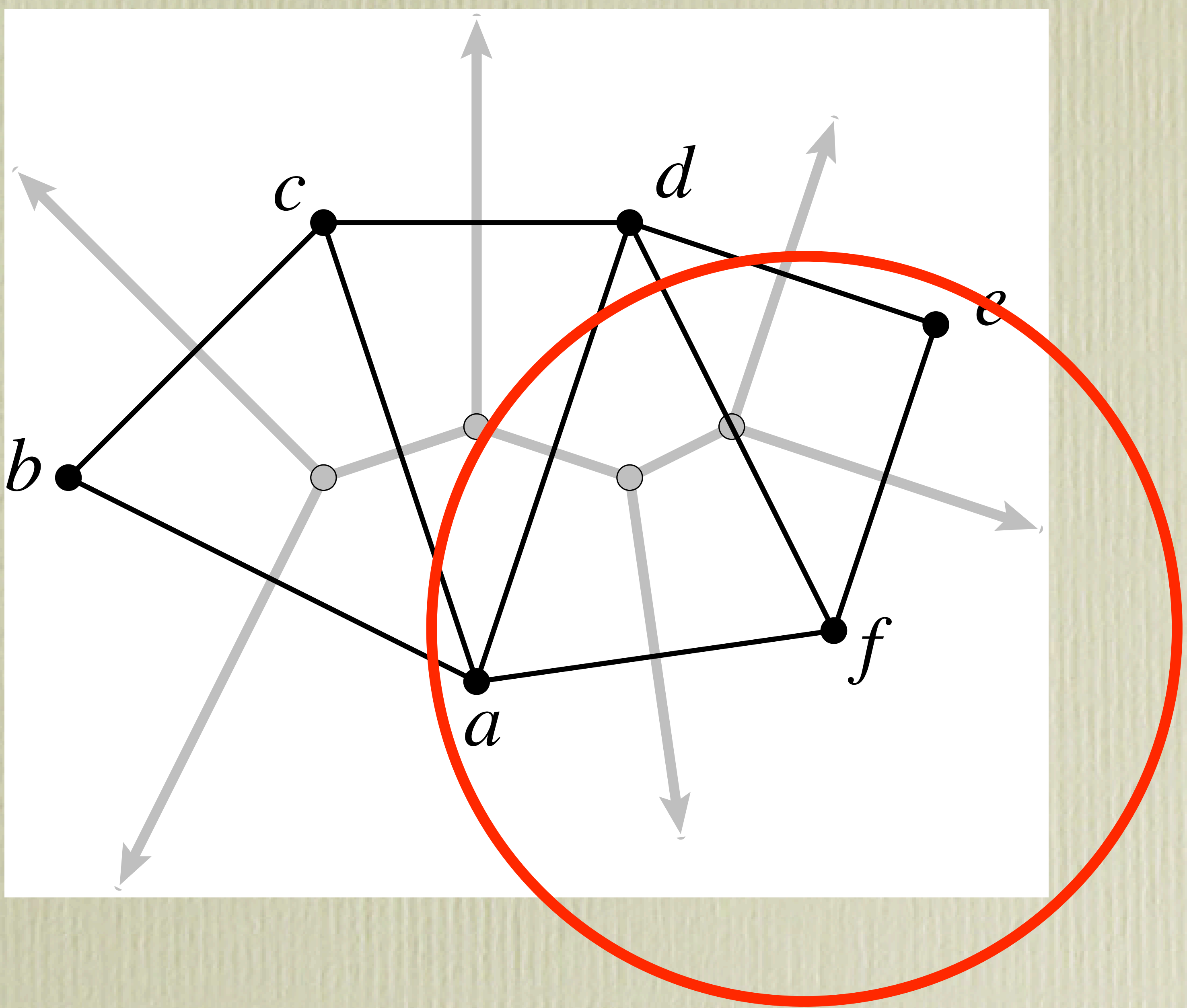
Example (continued)



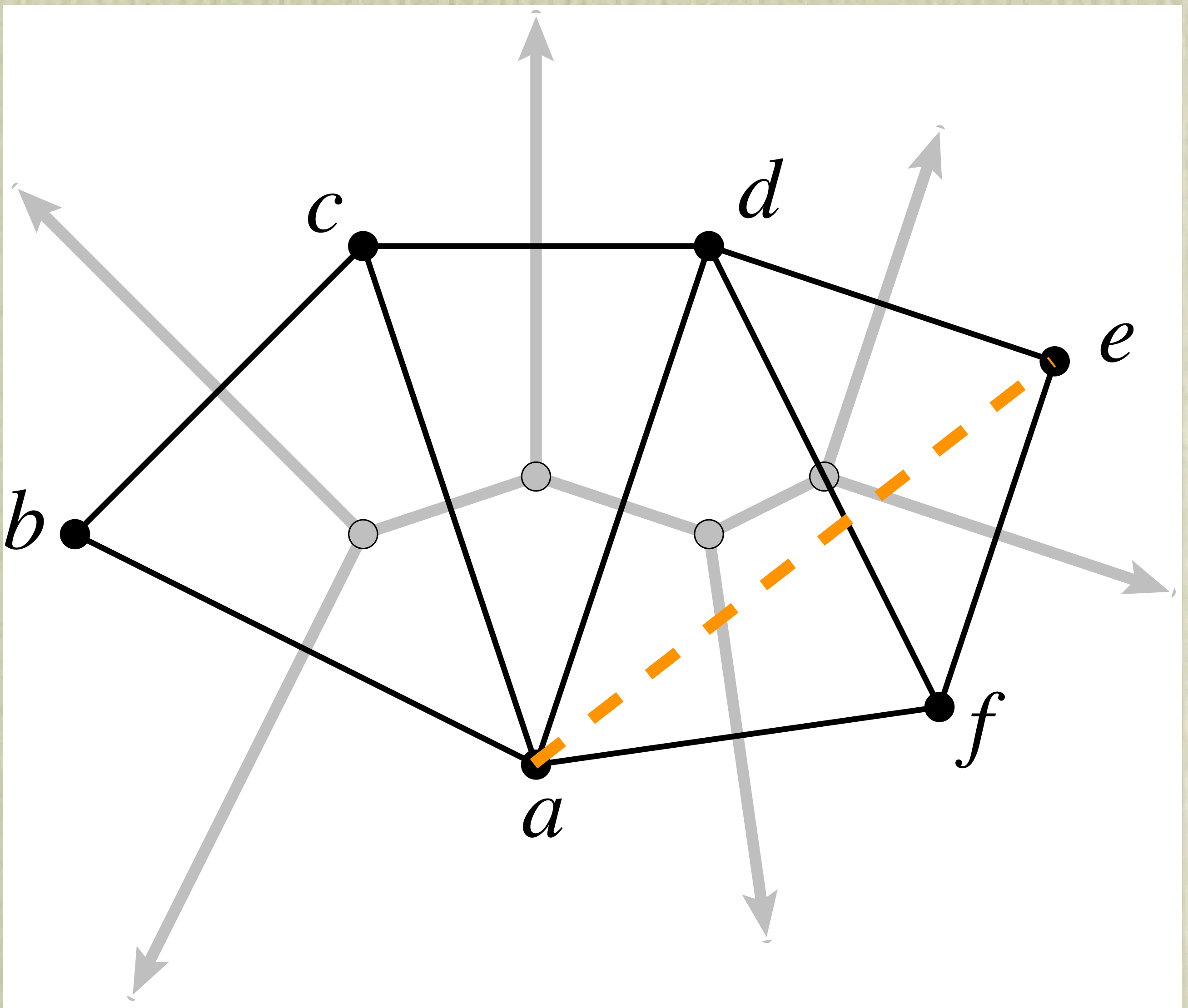
Example (continued)



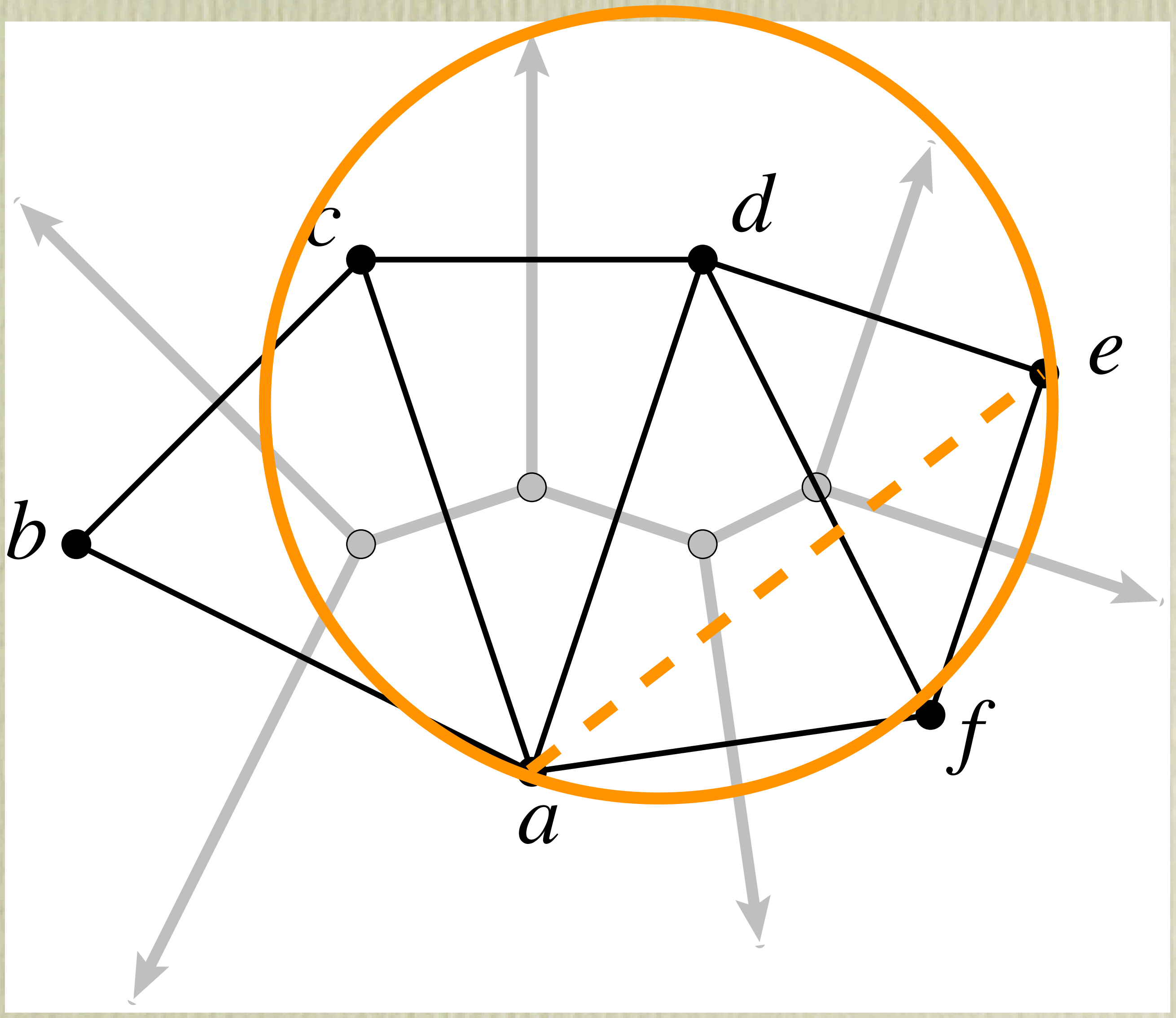
Example (continued)



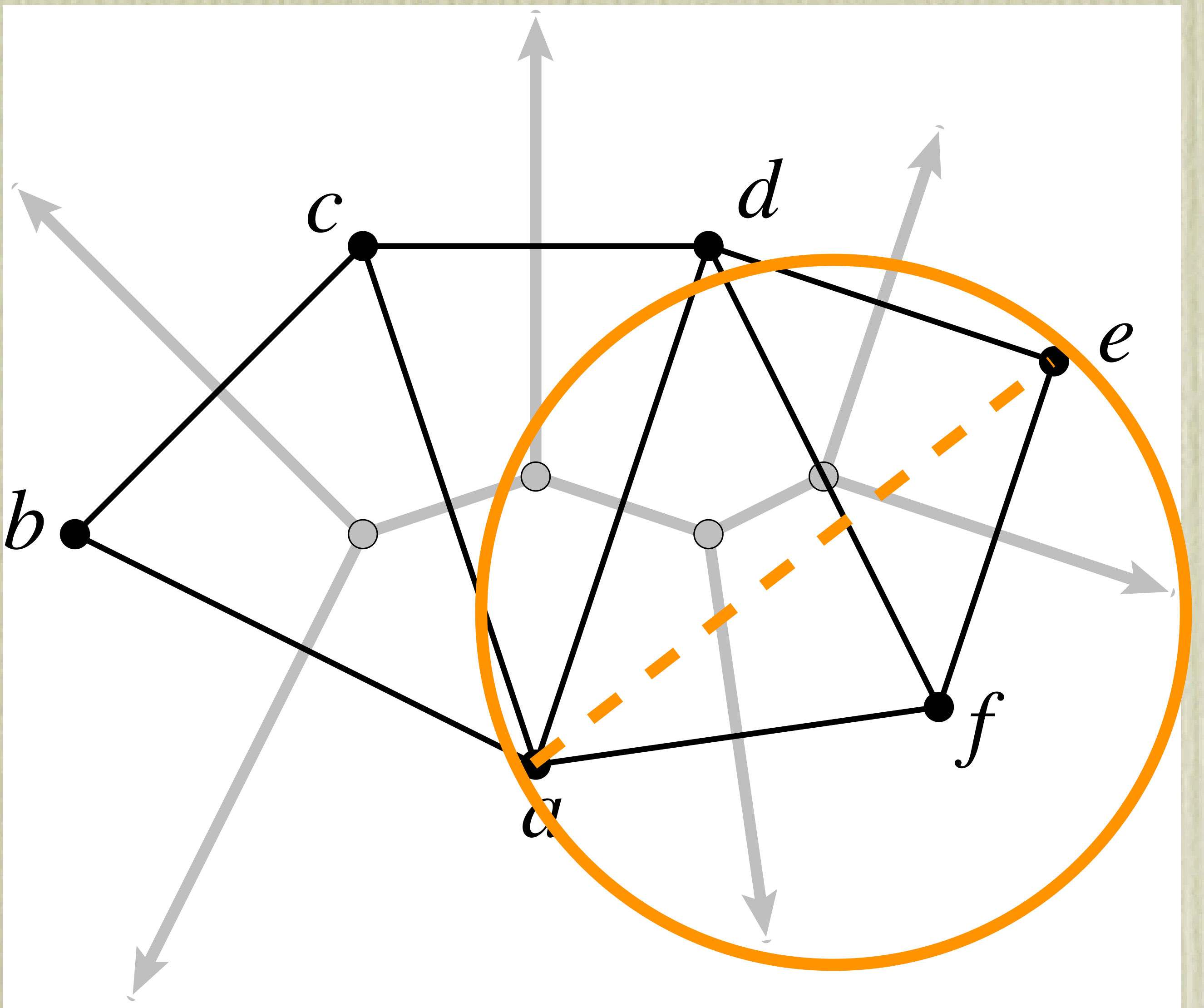
Example (continued)



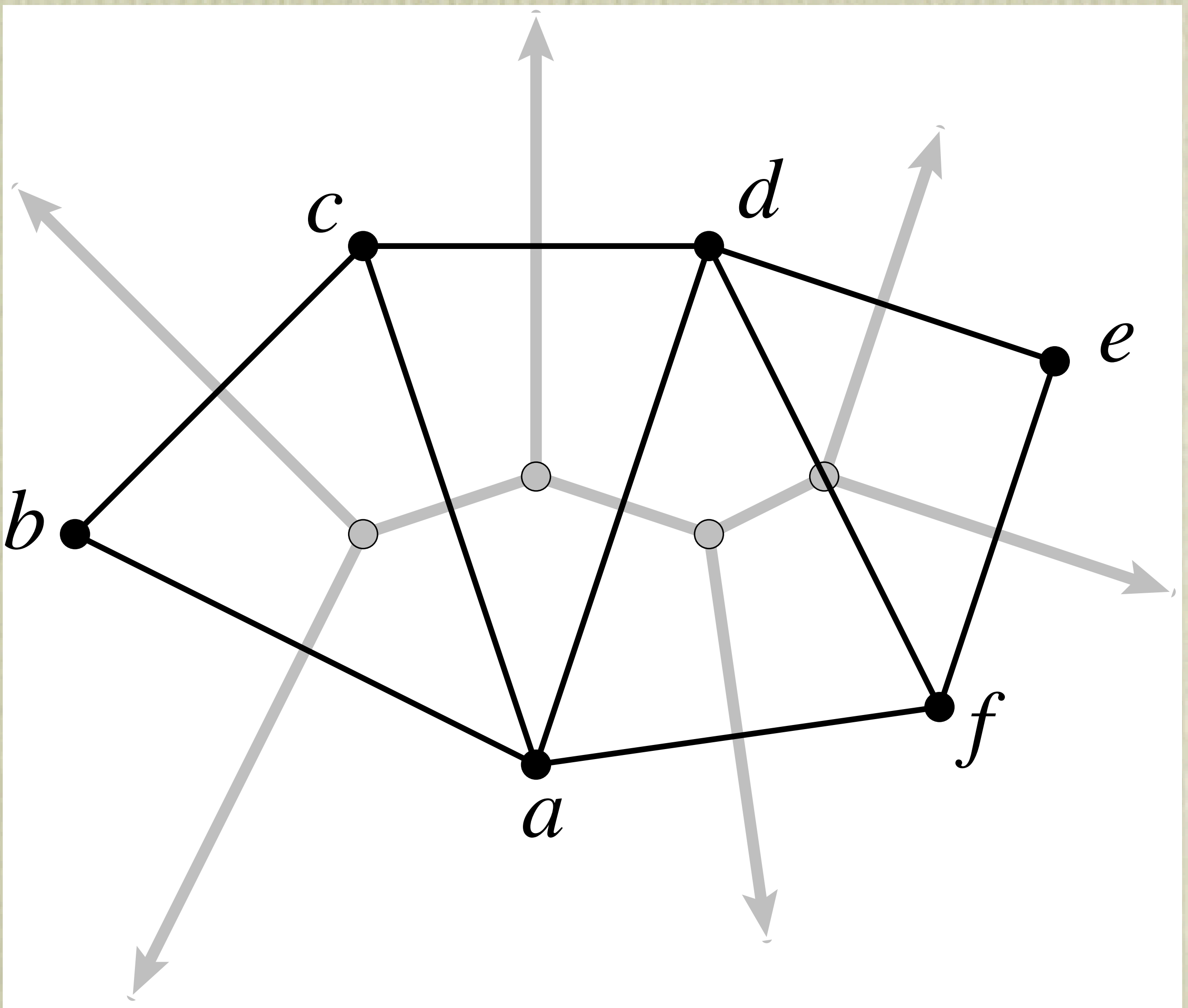
Example (continued)



Example (continued)



Example (continued)



Weak witness complex

- Let L be a set of points taken from a finite set $X \subset M \subset \mathbf{R}^n$ and let $x_0, x_1, \ldots, x_k \in L$. We decree that $x_0, x_1, \ldots, x_k$ span a k-cell in the weak witness complex if and only if:
  - There is a point $w \in X$, whose k+1 nearest neighbours in L are $x_0, x_1, \ldots, x_k$; and
  - all the faces of $[x_0, x_1, \ldots, x_k]$ belong to the weak witness complex.

Comments

- Weak witnesses exist with positive probability (though sometimes positive = small).
- We also define a version of the weak witness complex with a tolerance parameter R.
- Heuristically, weak witness complexes ought to give good results even when R is very small.

Strong vs. Weak

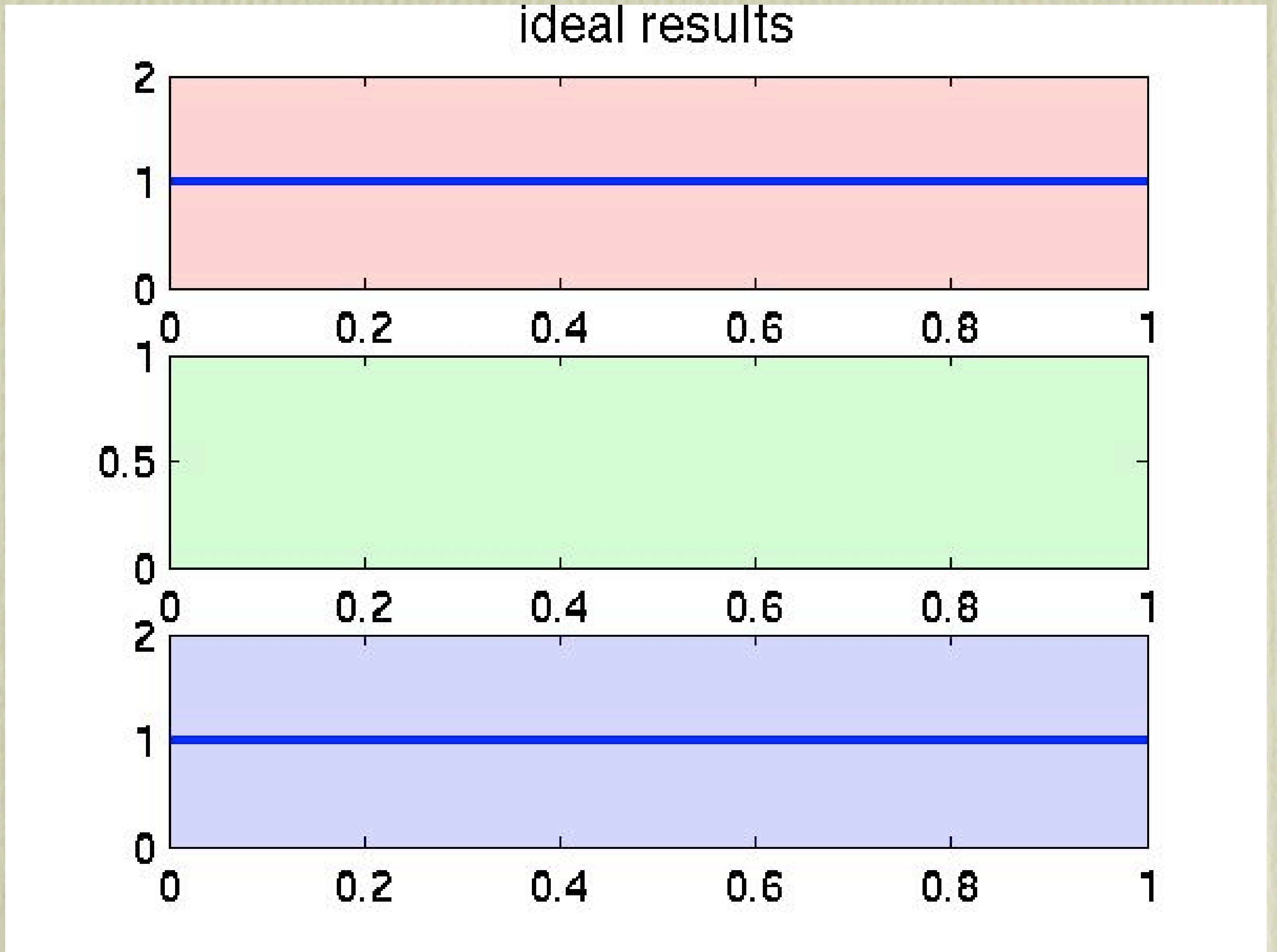
- Empirical evidence and heuristic arguments suggest:
  - Strong: noisy for small values of R; the "correct" stable realm begins later.
  - Weak: stable realm begins at (or near) R=0.
  - ✔ Weak: overcomes sampling irregularity.
  - ✘ Weak: ignores small features.

6. Example: the 2-sphere

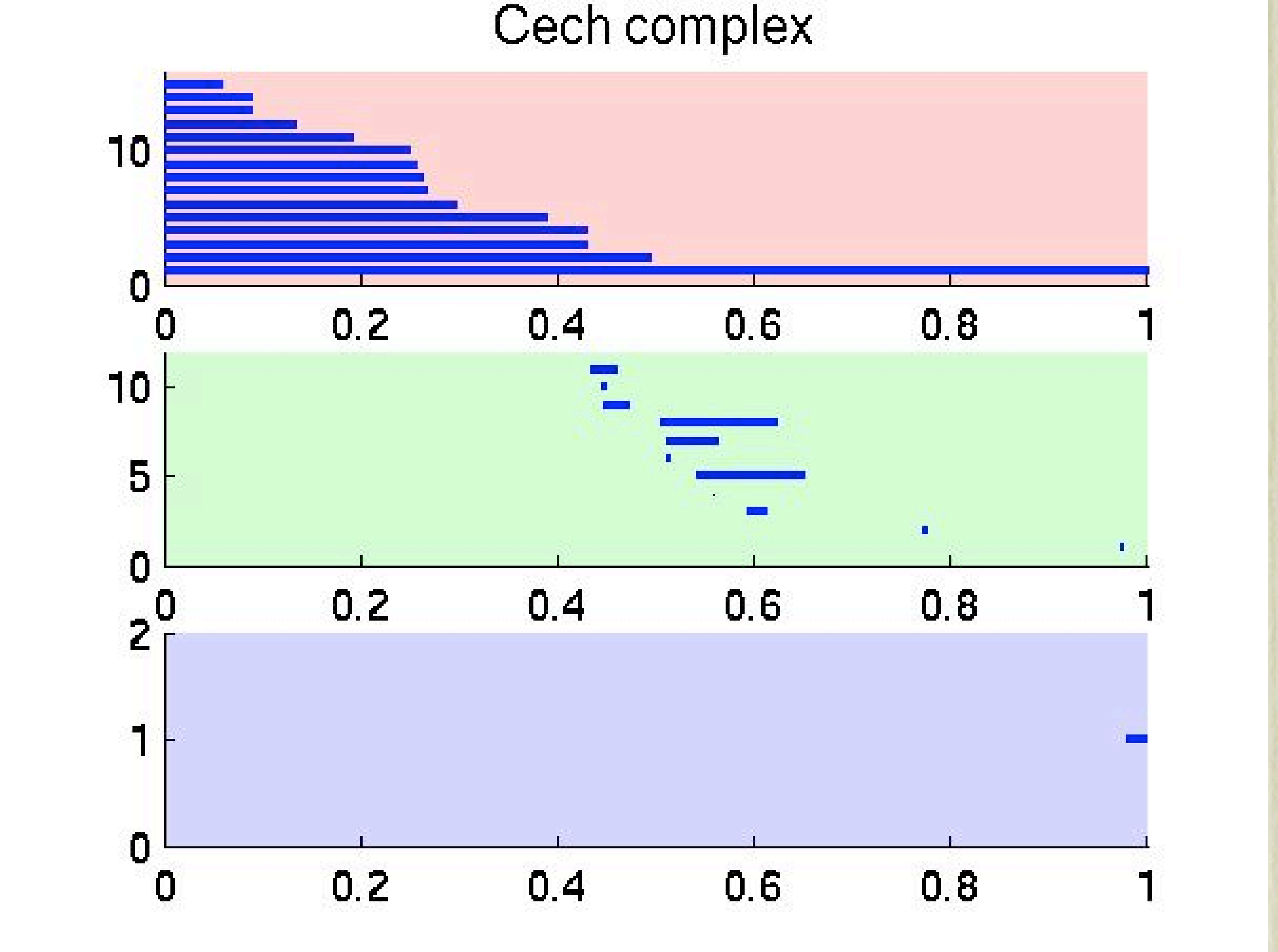
The 2-sphere

- Toy example (to check that everything works).
- 1000 points sampled uniformly randomly on the unit sphere in 3-space.
- 15 landmark points chosen randomly or by greedy separation maximisation.
- Compare Čech/Alpha, strong witness, weak witness complexes.

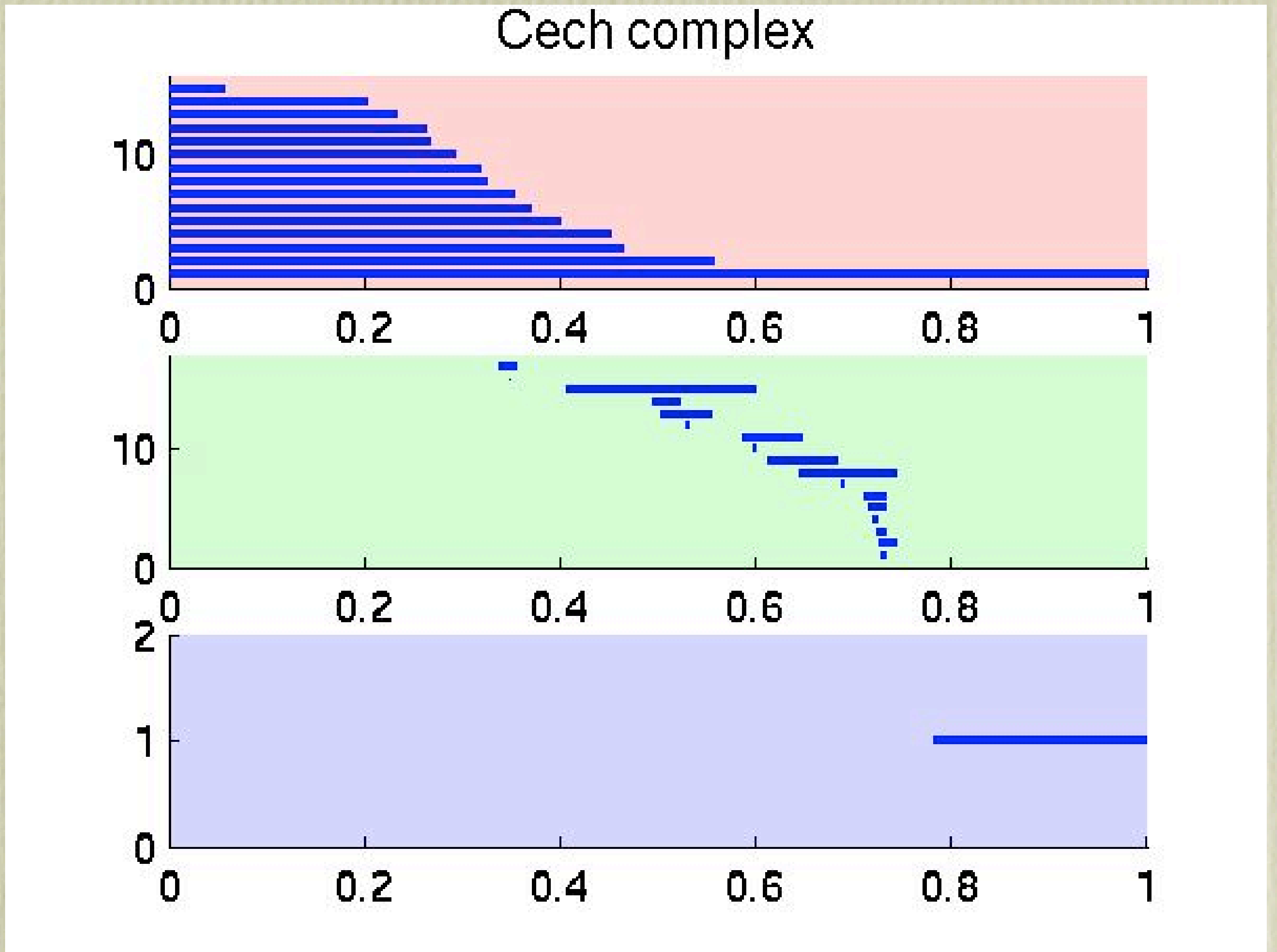
"true" Betti number profile for 2-sphere



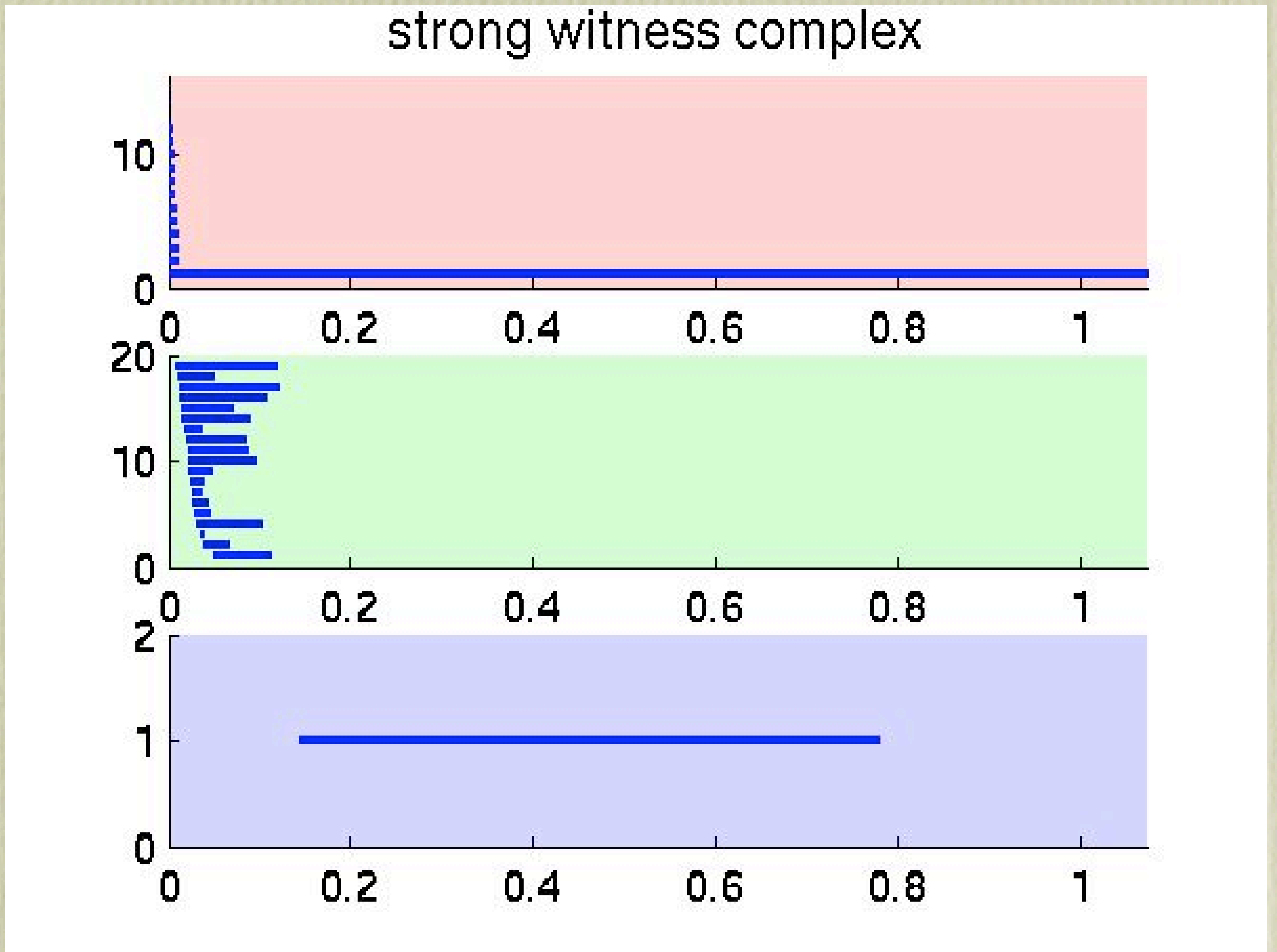
Čech/Alpha complex
15 random landmarks



Čech/Alpha complex
15 separated landmarks

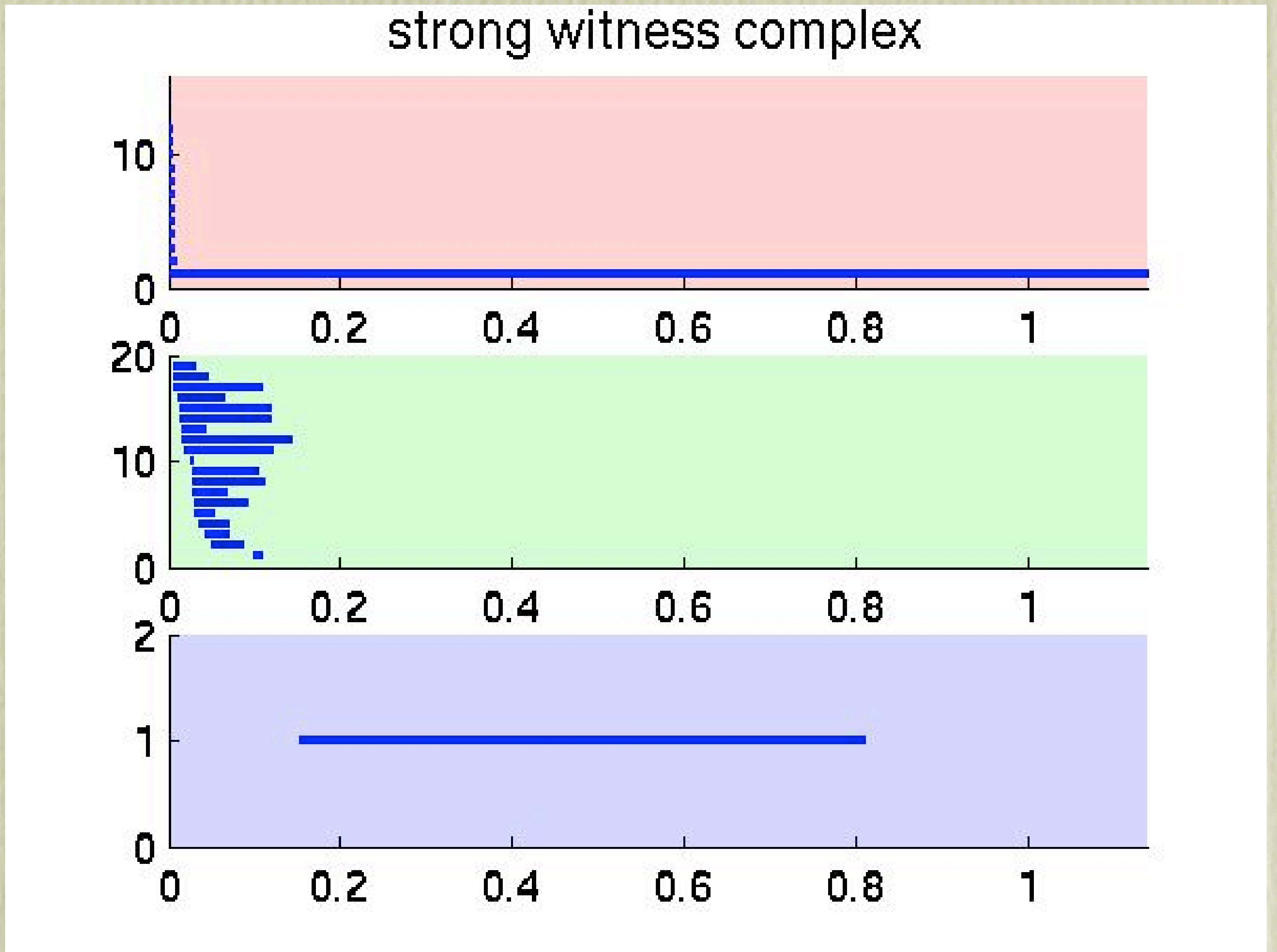


Strong witness complex
15 random landmarks

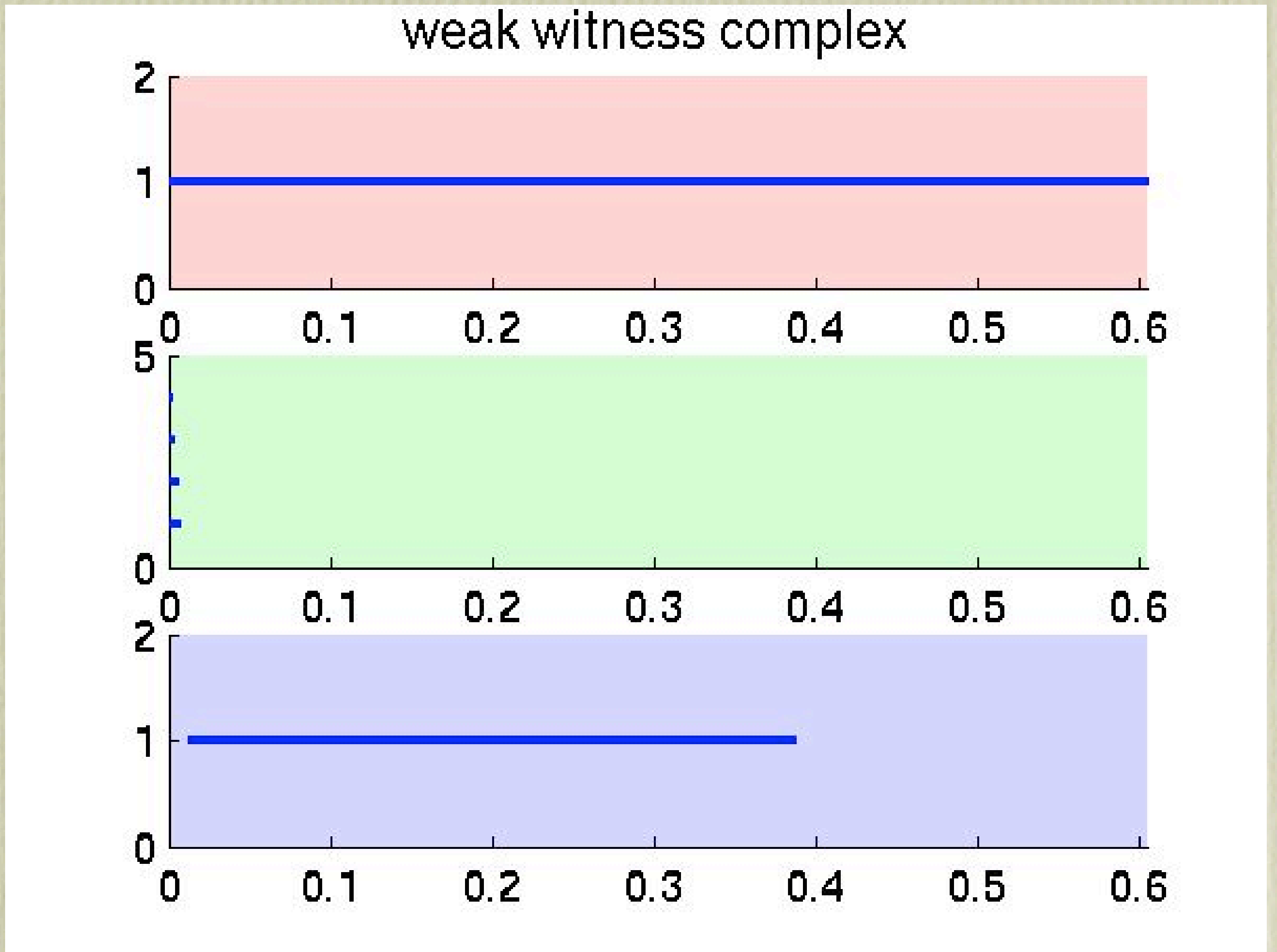


Vin de Silva
Stanford University

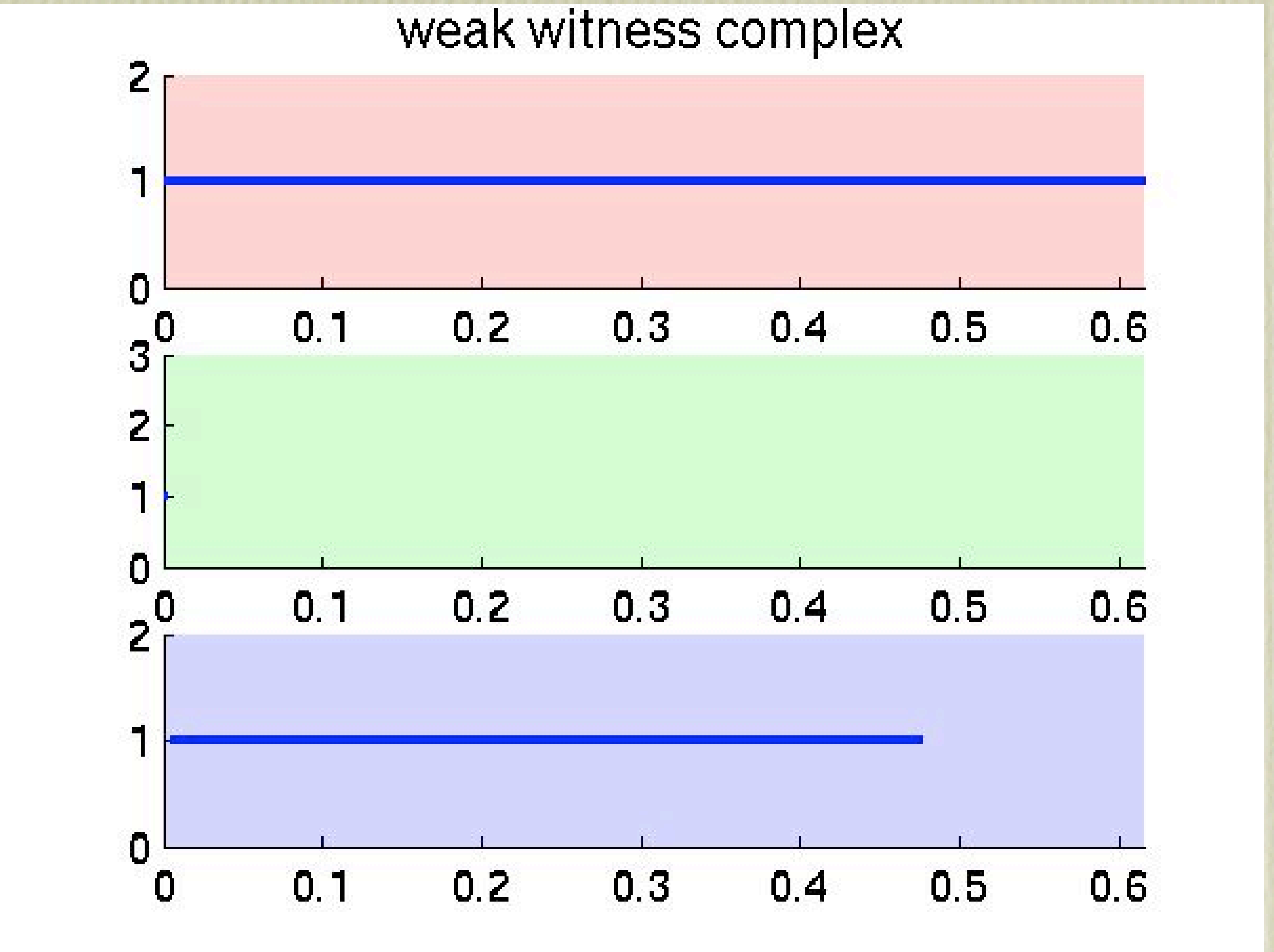
Strong witness complex
15 separated landmarks



Weak witness complex
15 random landmarks



Weak witness complex
15 separated landmarks



Landmark choice

- Theorems wanted! Analogous to the question of how to sample points from manifold.
- [NSW]: (R/2)-denseness is enough for Čech.
- Manifold reconstruction literature: ensure that sample points are separated.
- "Greedy furthest point" satisfies both.
- Coverage scale normalises R-truncation.

7. Example: high-contrast image patches

High-contrast visual image patches

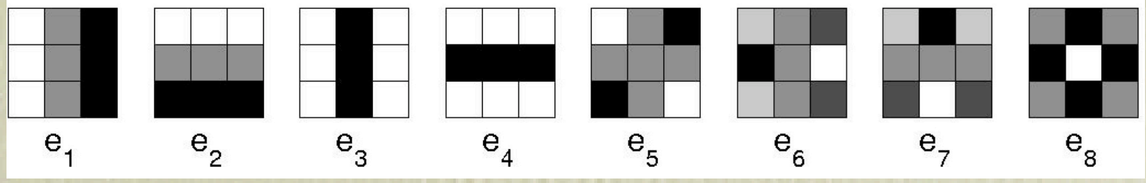
- Ann Lee, Kim Pedersen, David Mumford (2003) studied the local statistical properties of natural images (from Van Hateren's database).

- Restrict attention to 3-by-3 pixel patches with high contrast between pixels: are some patterns more likely than others?

- We investigated the topological properties of high-density regions in pixel-patch space.

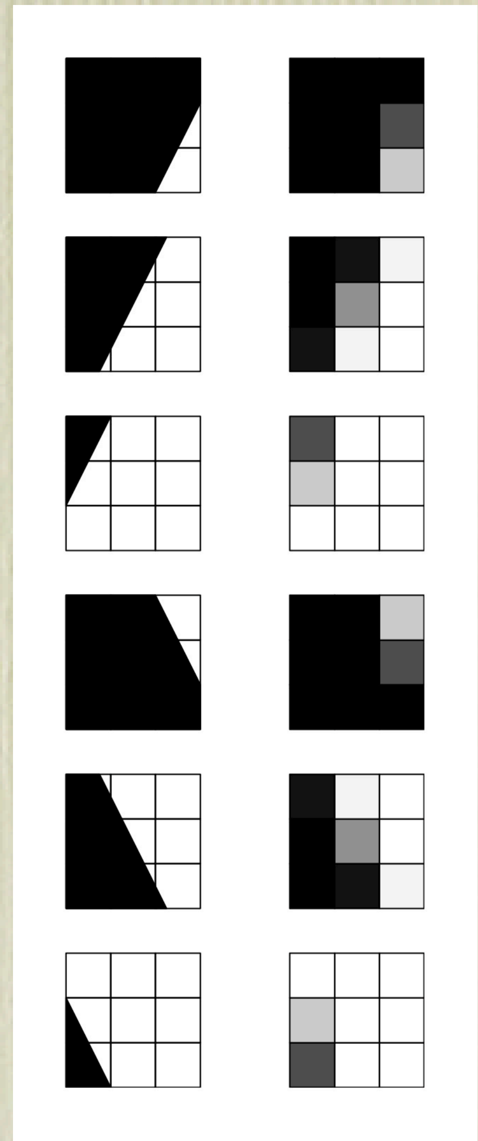
The space of image patches

- ~4.2 million high-contrast 3-by-3 patches selected randomly from images in database.
- Normalise each patch twice: subtract mean intensity, then rescale to unit norm.
- Normalised patches live on a unit 7-sphere in 8-dimensional space with the following basis:



High-density regions

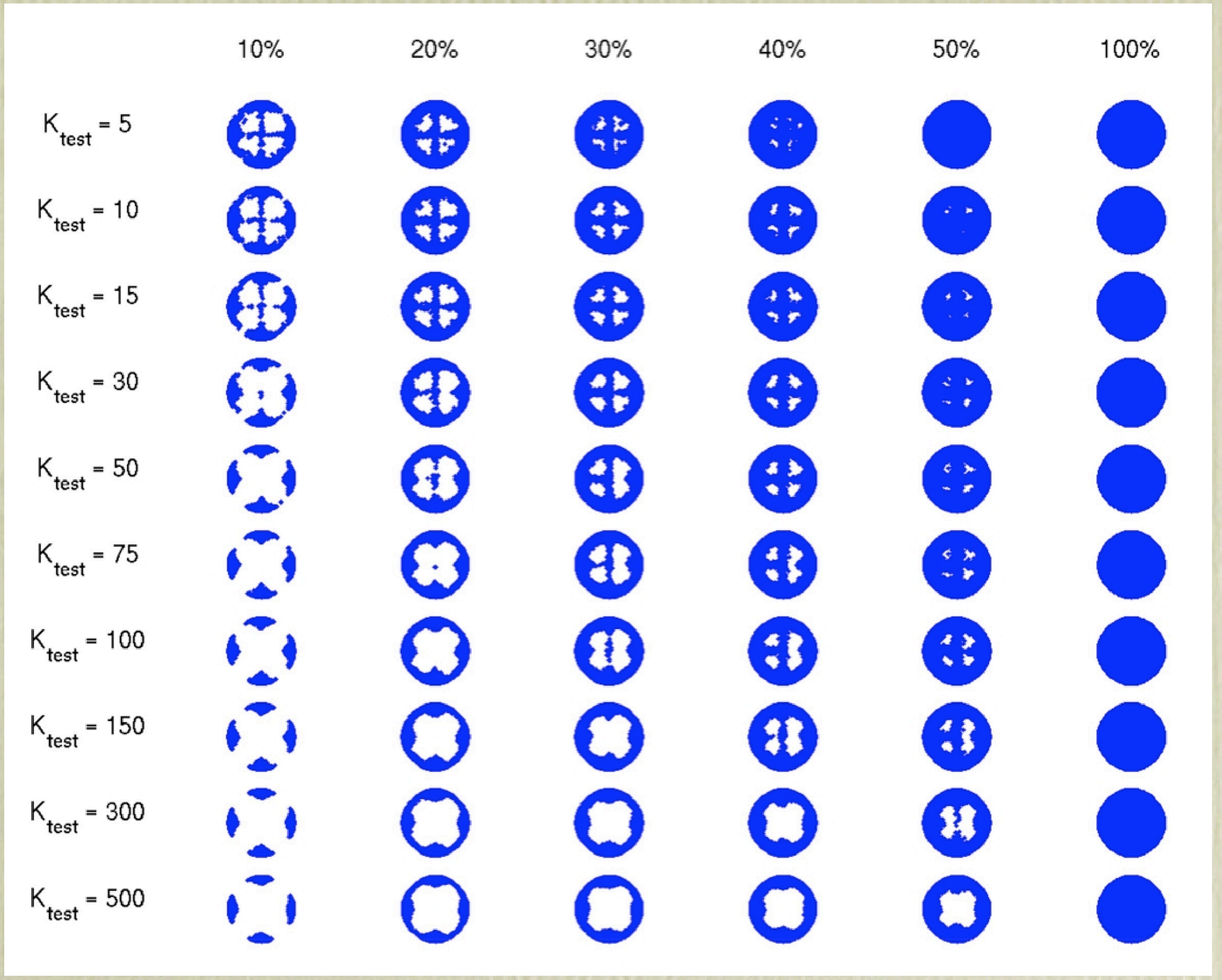
- The distribution of patches is dense in the 7-sphere (it turns out).
- There are high-density regions: for example, edge features are prevalent in natural images.
- Can we describe the topology of the high-density regions?



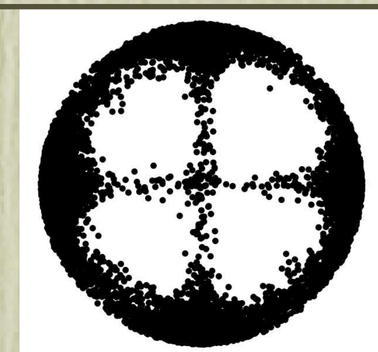
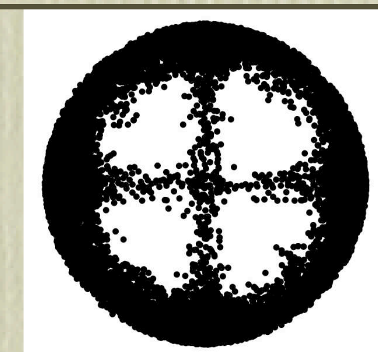
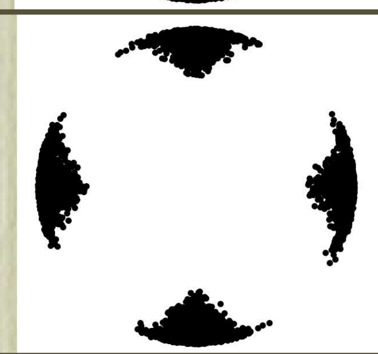
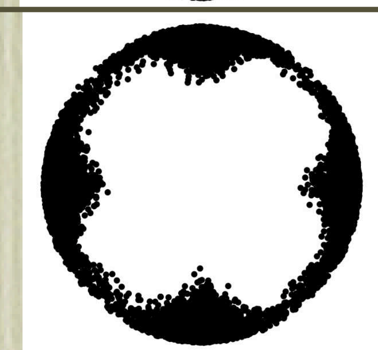
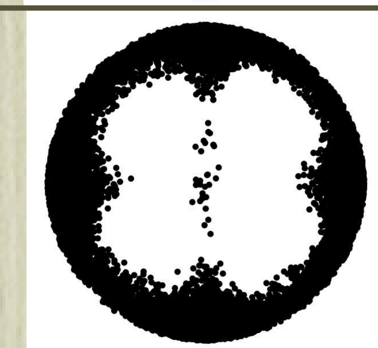
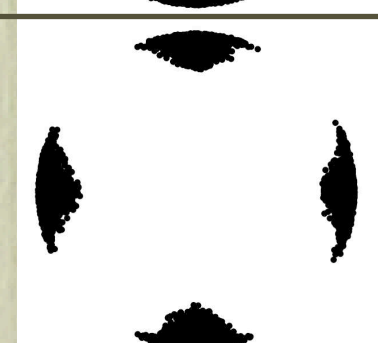
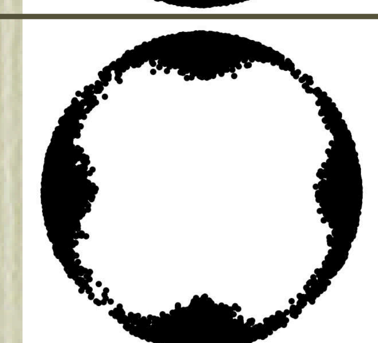
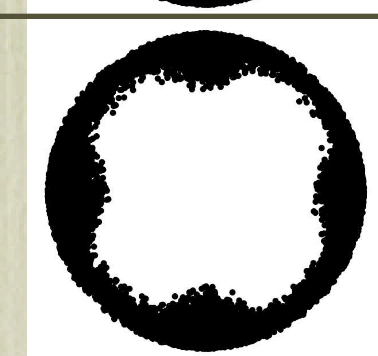
Defining “high-density”

- When does a point belong to a high-density region? There is no single answer to this.
- Select a positive integer K.
- For each data point x, let r(x,K) denote the distance between x and its K-th nearest neighbour.
- Threshold on r(x,K):
  x is a high-density point $\leftrightarrow$ r(x,K) is small

Different high-density cuts

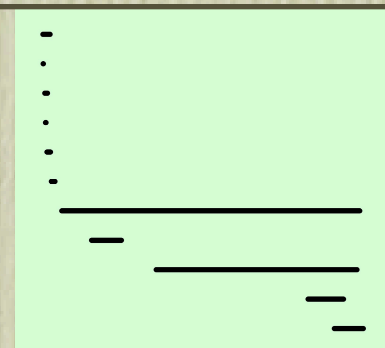
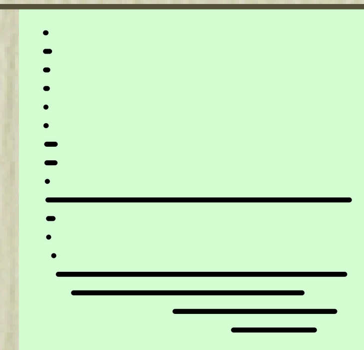
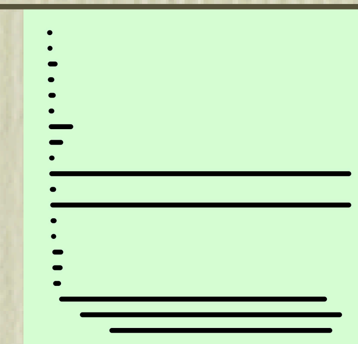
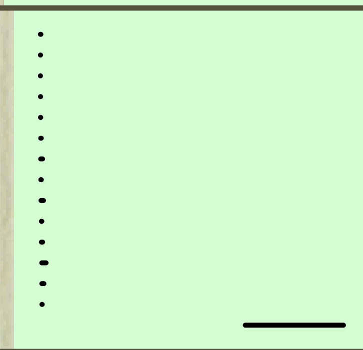
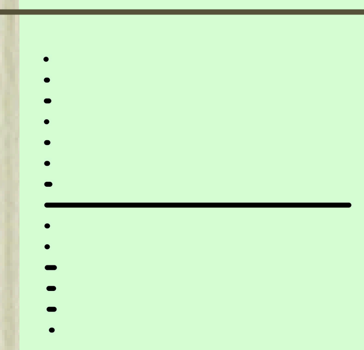
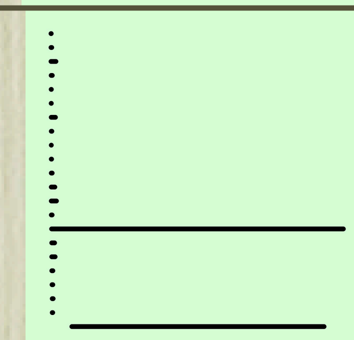
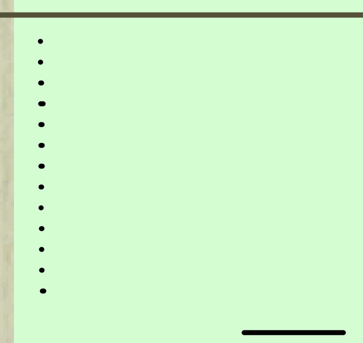
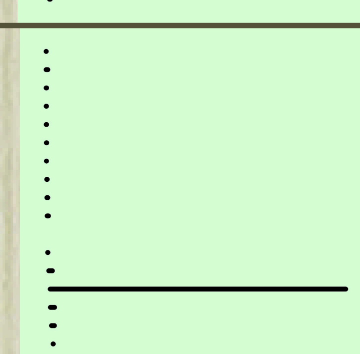
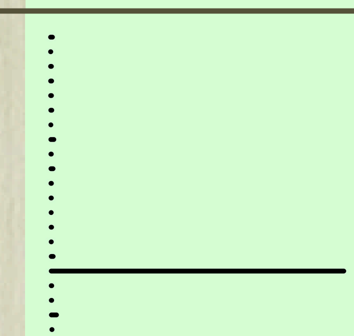


A small platter of cuts

	10%	20%	30%
K=15			
K=100			
K=300			

Vin de Silva
Stanford University

Persistent homology: Betti 1

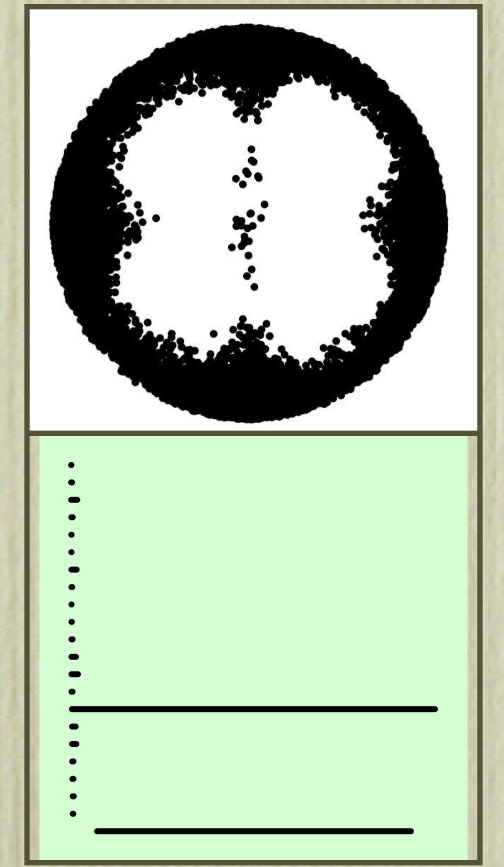
	10%	20%	30%
K=15			
K=100			
K=300			

Vin de Silva
Stanford University

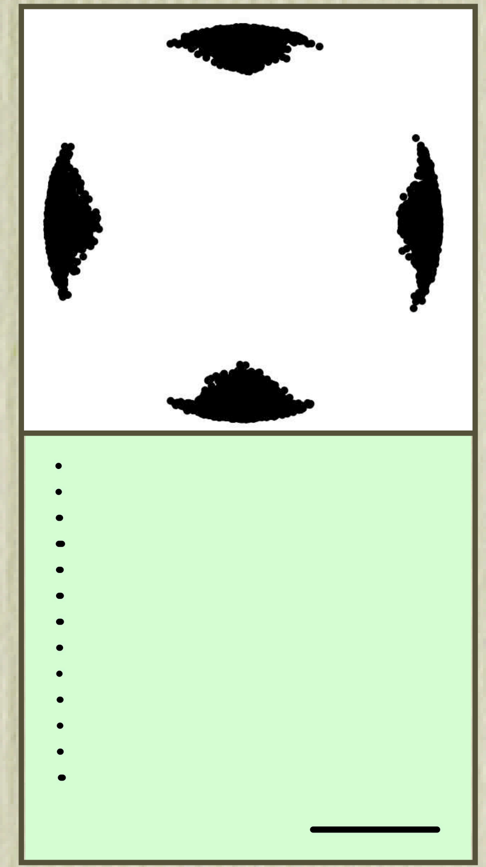
Obvious patterns

- Certain results are easy to interpret.

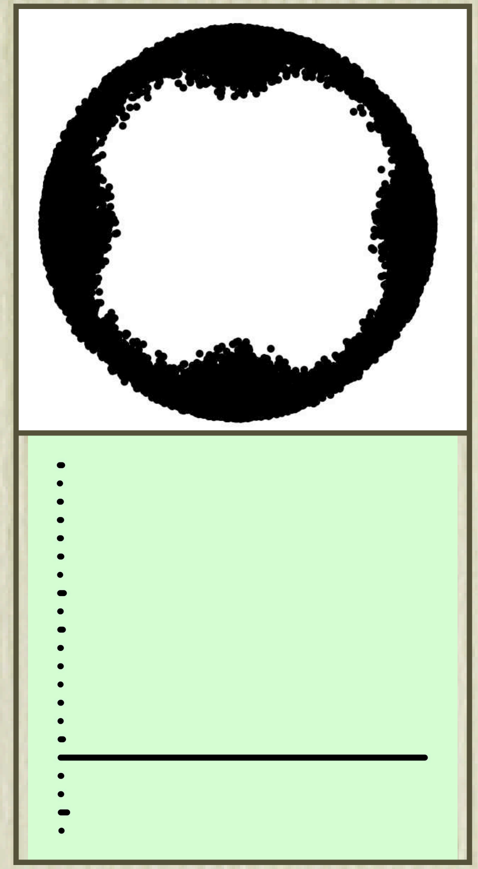
K = 100; 30%



K = 300; 10%

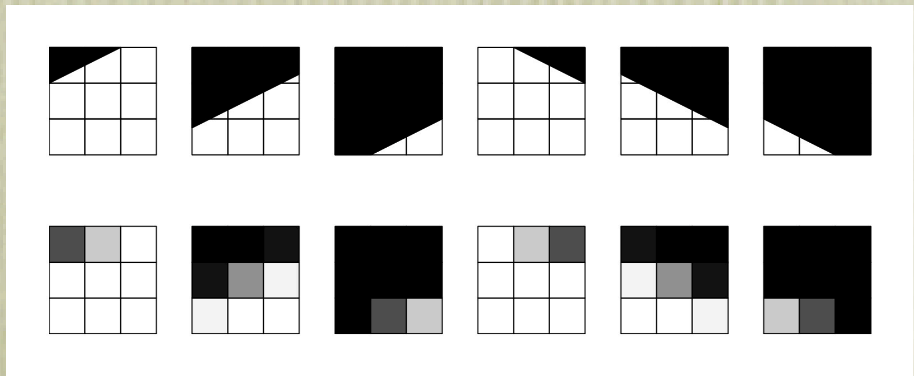


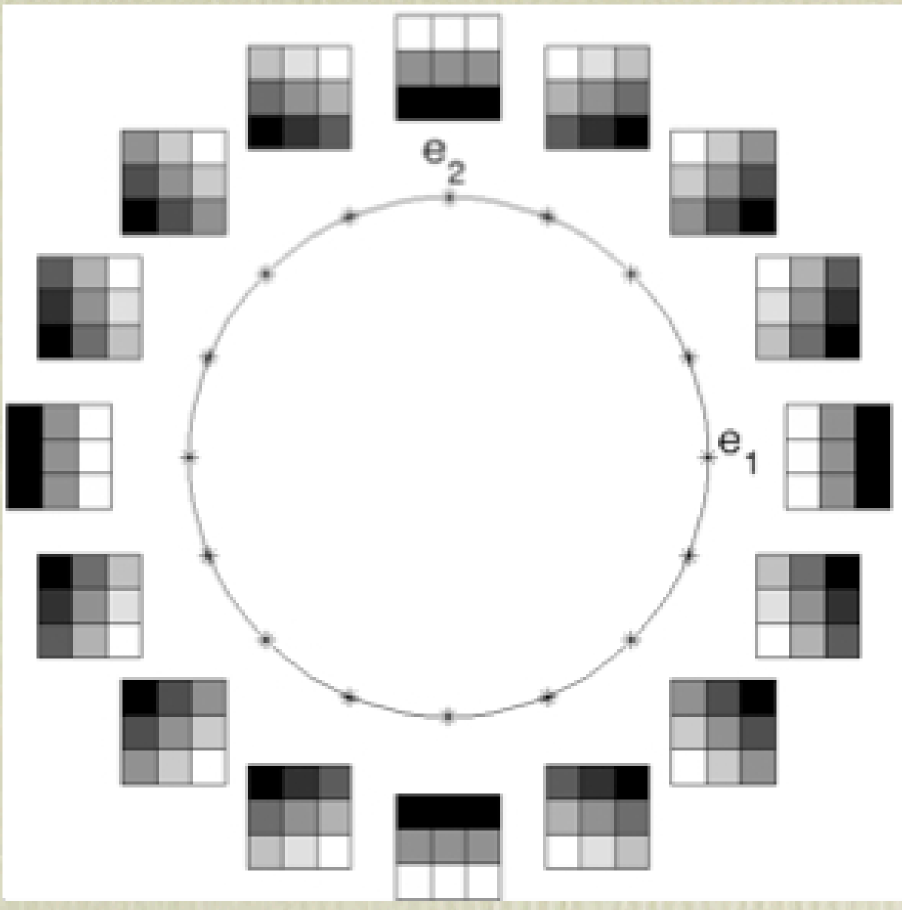
K = 300; 30%



The primary circle

- The thick e1–e2 circle consists of linear gradient patches and their nearby edge feature patches.

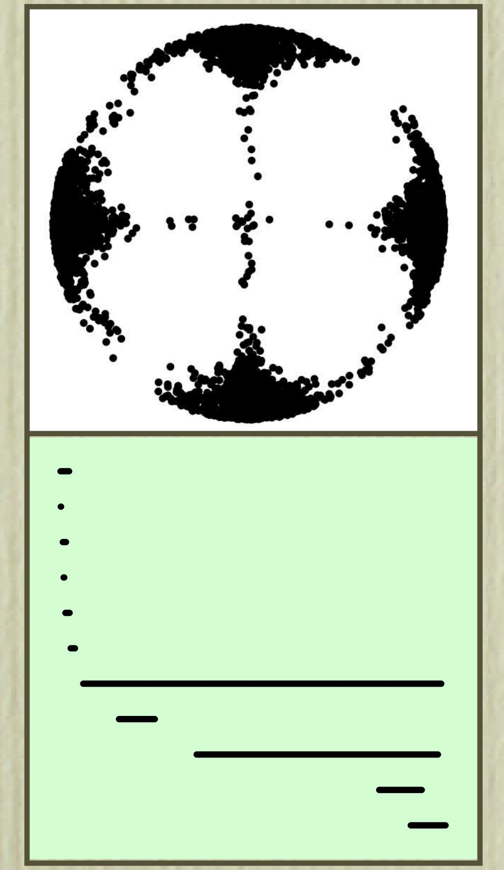




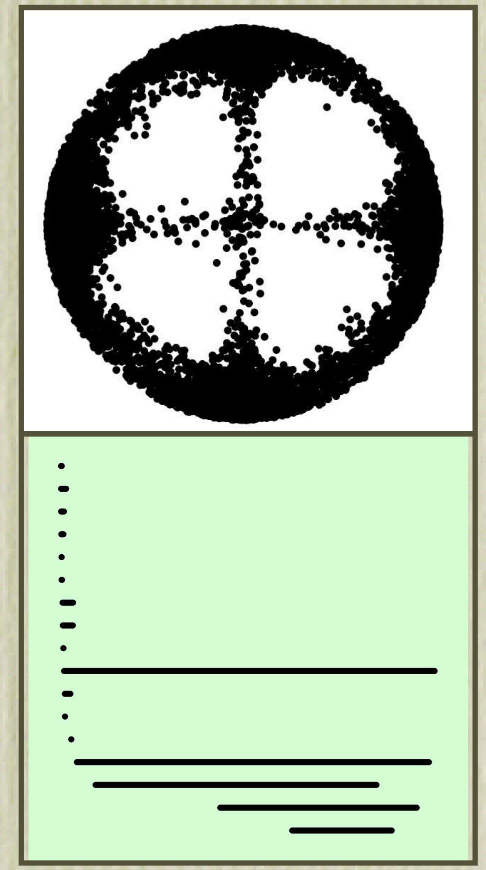
Less obvious

- The K = 15 row is initially more mysterious.

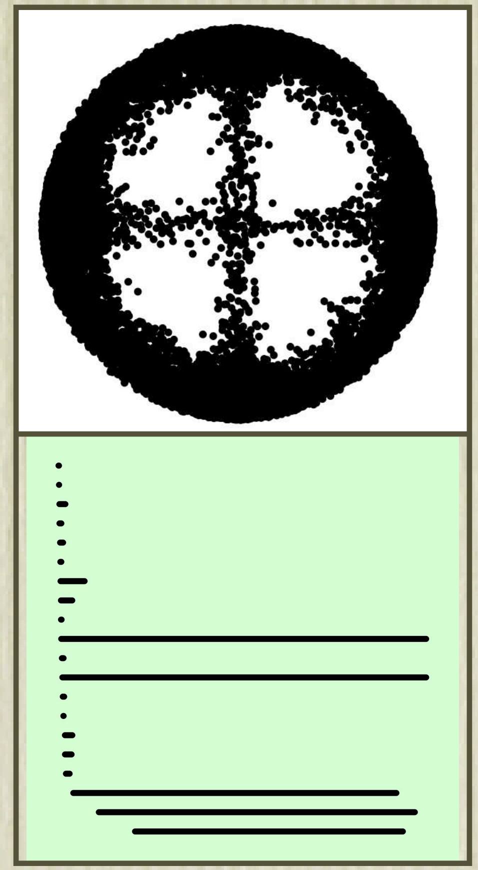
K = 15; 10%



K = 15; 20%

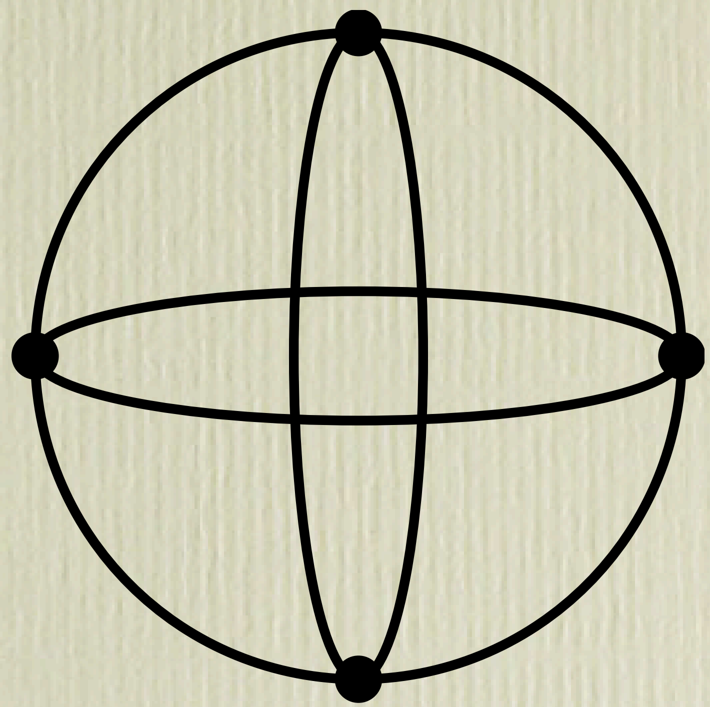


K = 15; 30%



Three circles model

- Answer: three circles in $\mathbf{R}^4$ (projected into $\mathbf{R}^2$).
- The primary circle in the e1–e2 plane meets two secondary circles (e1-e3 and e2-e4) twice each.
- The two secondary circles are disjoint.



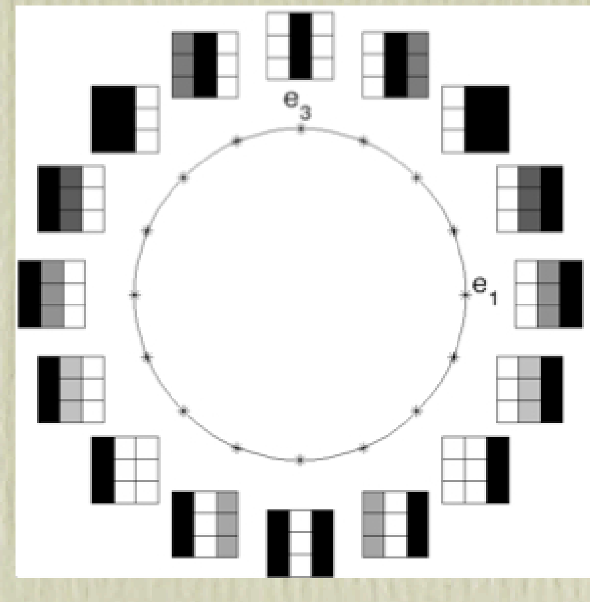
Movie (by Afra Zomorodian)

Movie (by Afra Zomorodian)

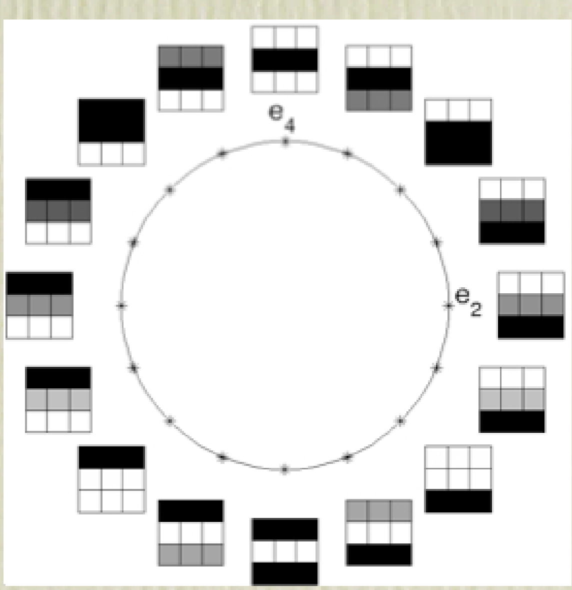
Movie (by Afra Zomorodian)

The secondary circles

- The thin circles in the e1–e3 and e2–e4 planes consist of vertically symmetric and horizontally symmetric patches.



- Why is there a greater concentration of these patches? Two answers.



Closing remarks

Closing remarks

- Persistent homology + witness complexes: make topological measurements robustly, reasonably cheaply, with hardly any arbitrary parameters.
- “Continuisation” and parameter elimination are both based on “integrating” over R. Calculations over $\mathbf{Z}_2$ are still discrete.
- Working in a more analytic framework over $\mathbf{R}$ leads to other approaches. (Laplacians etc.)