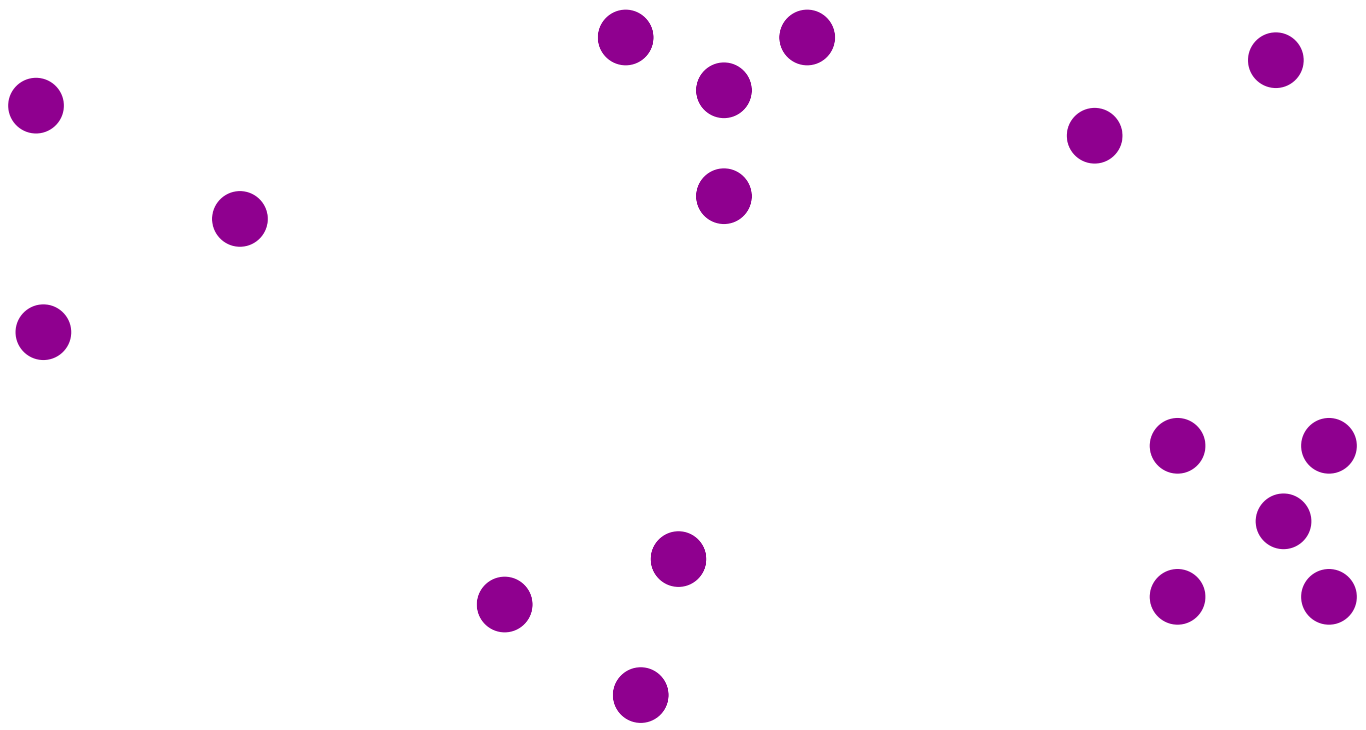
MATH:7450 (22M:305) Topics in Topology: Scientific and Engineering Applications of Algebraic Topology

Sept 4, 2013

Clustering Via Persistent Homology

Creating a simplicial complex



Step 0.) Start by adding 0-dimensional vertices (0-simplices)

Creating a simplicial complex

1.) Adding 1-dimensional edges (1-simplices)

Add an edge between data points that are "close"

Creating a simplicial complex

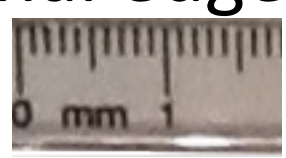
1.) Adding 1-dimensional edges (1-simplices)

Let T = Threshold

Connect vertices v and w with an edge iff the distance between v and w is less than T

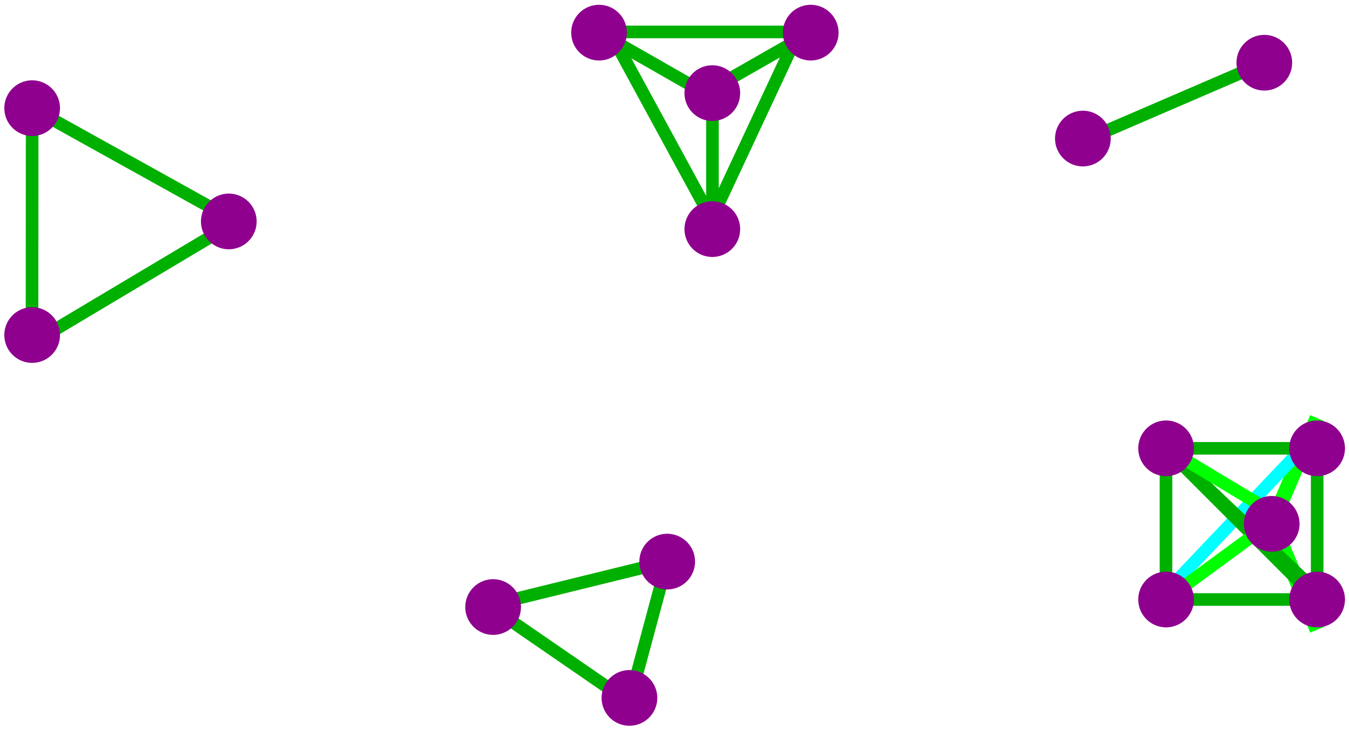
Creating a simplicial complex

1.) Adding 1-dimensional edges (1-simplices)

Let T = Threshold = 

Connect vertices v and w with an edge iff the distance between v and w is less than T

Creating a simplicial complex

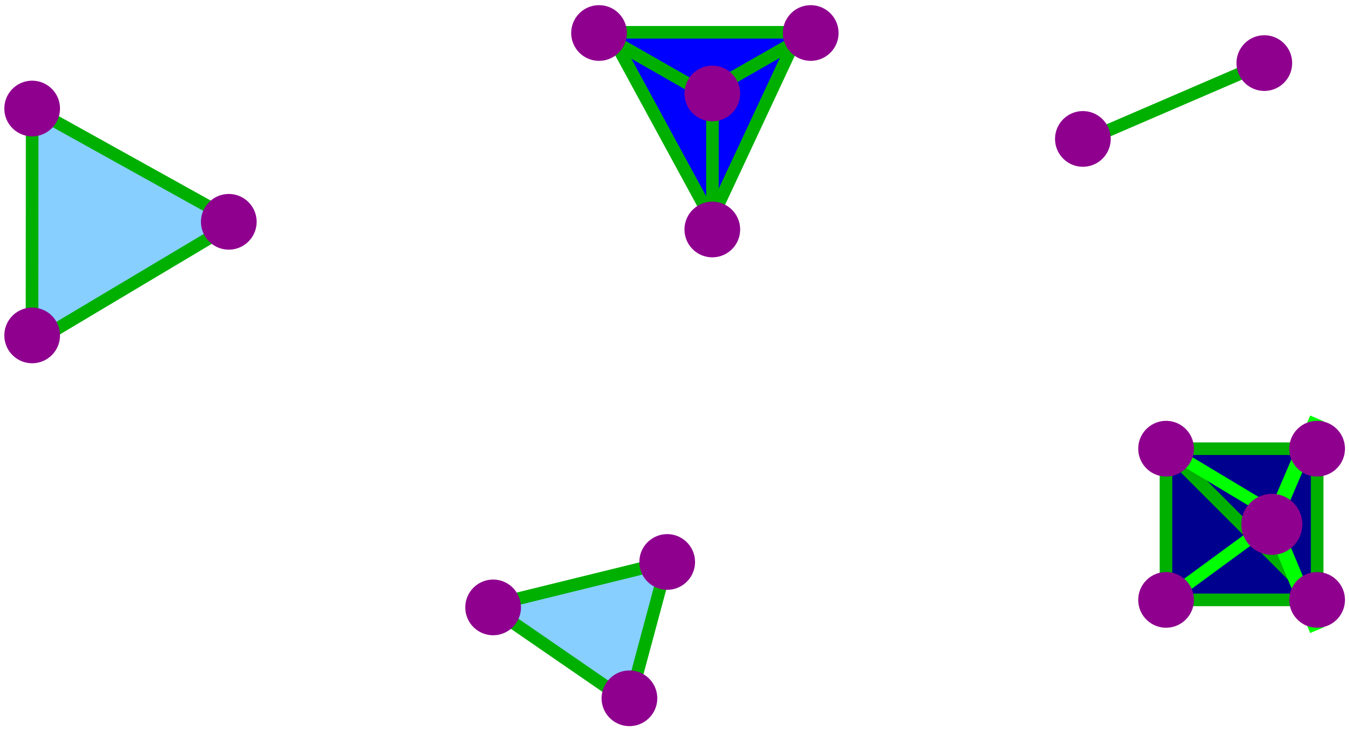


1.) Adding 1-dimensional edges (1-simplices)

Add an edge between data points that are "close"

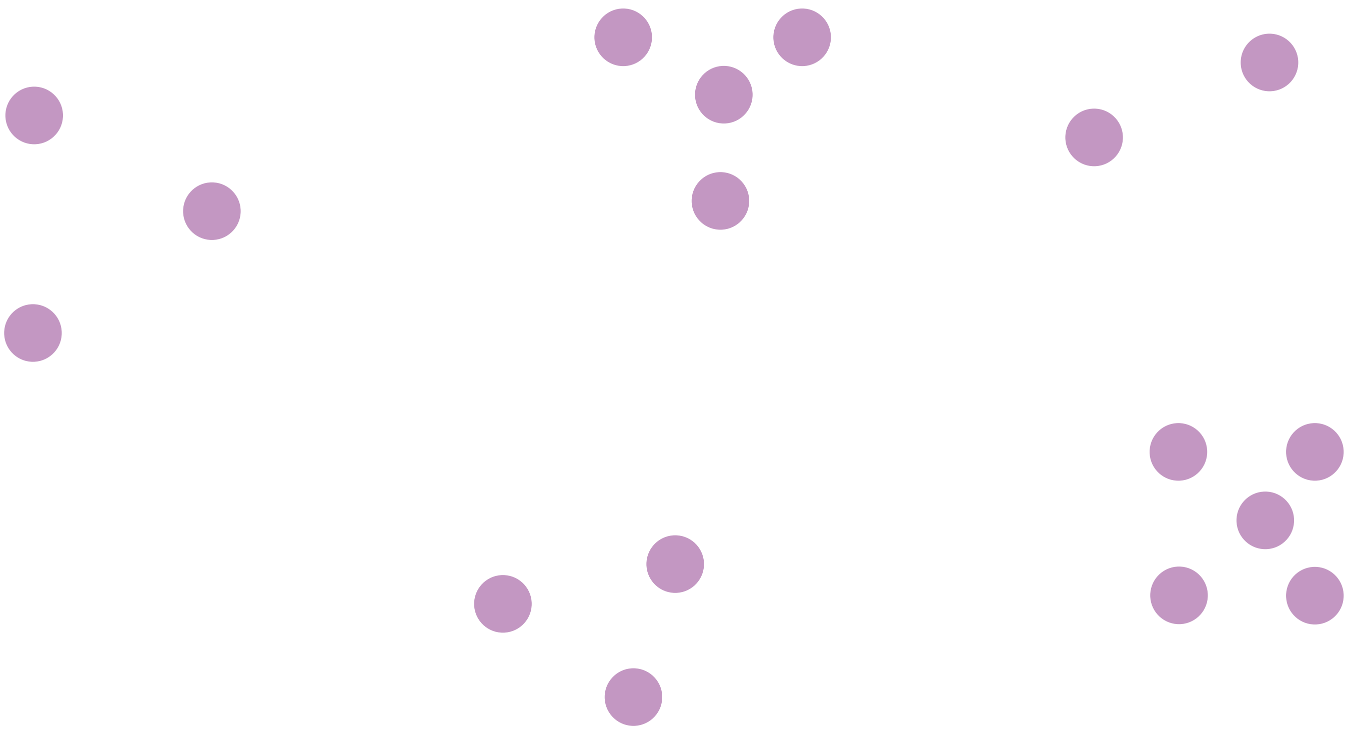
Note: we only need a definition of closeness between data points. The data points do not need to be actual points in R^n

Creating the Vietoris Rips simplicial complex



2.) Add all possible simplices of dimensional > 1.

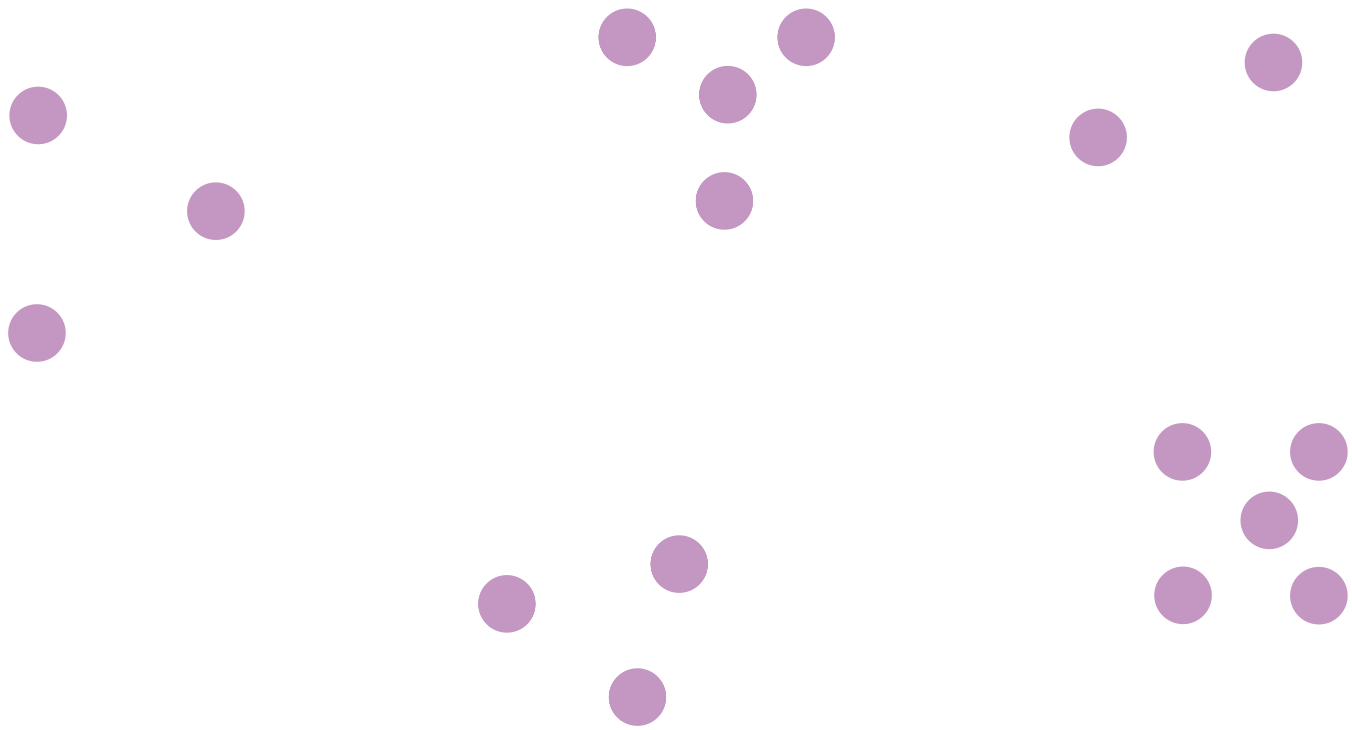
Creating the Vietoris Rips simplicial complex



0.) Start by adding 0-dimensional data points

Use balls to measure distance (Threshold = diameter). Two points are close if their corresponding balls intersect

Creating the Vietoris Rips simplicial complex



0.) Start by adding 0-dimensional data points

Use balls of varying radii to determine persistence of clusters.

H. Edelsbrunner, D. Letscher, and A. Zomorodian, Topological persistence and simplication, Discrete and Computational Geometry 28, 2002, 511-533.

Discriminative persistent homology of brain networks, 2011

Hyekyoung Lee Chung, M.K.; Hyejin Kang; Bung-Nyun Kim;Dong Soo Lee

Constructing functional brain networks with 97 regions of interest (ROIs)
extracted from FDG-PET data for
24 attention-deficit hyperactivity disorder (ADHD),
26 autism spectrum disorder (ASD) and
11 pediatric control (PedCon).

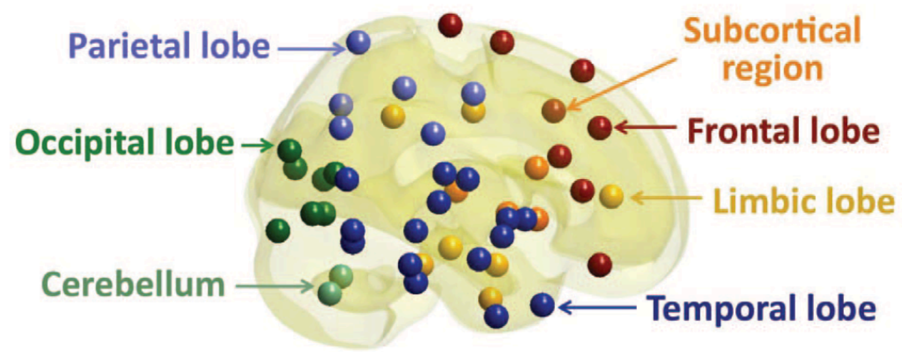


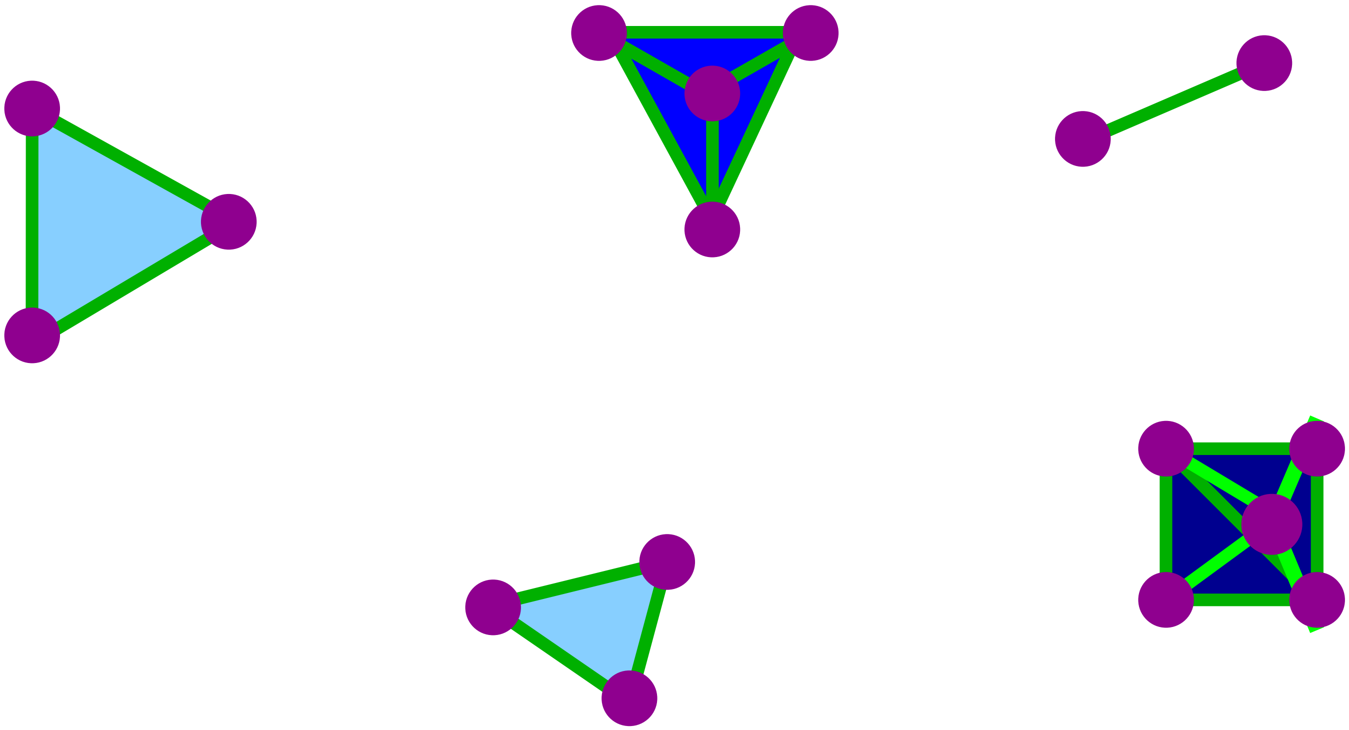
Fig. 3. Location of ROIs

Data = measurement f_j taken at region j

Graph: 97 vertices representing 97 regions of interest
edge exists between two vertices i,j if correlation between f_i and $f_j \geq$ threshold

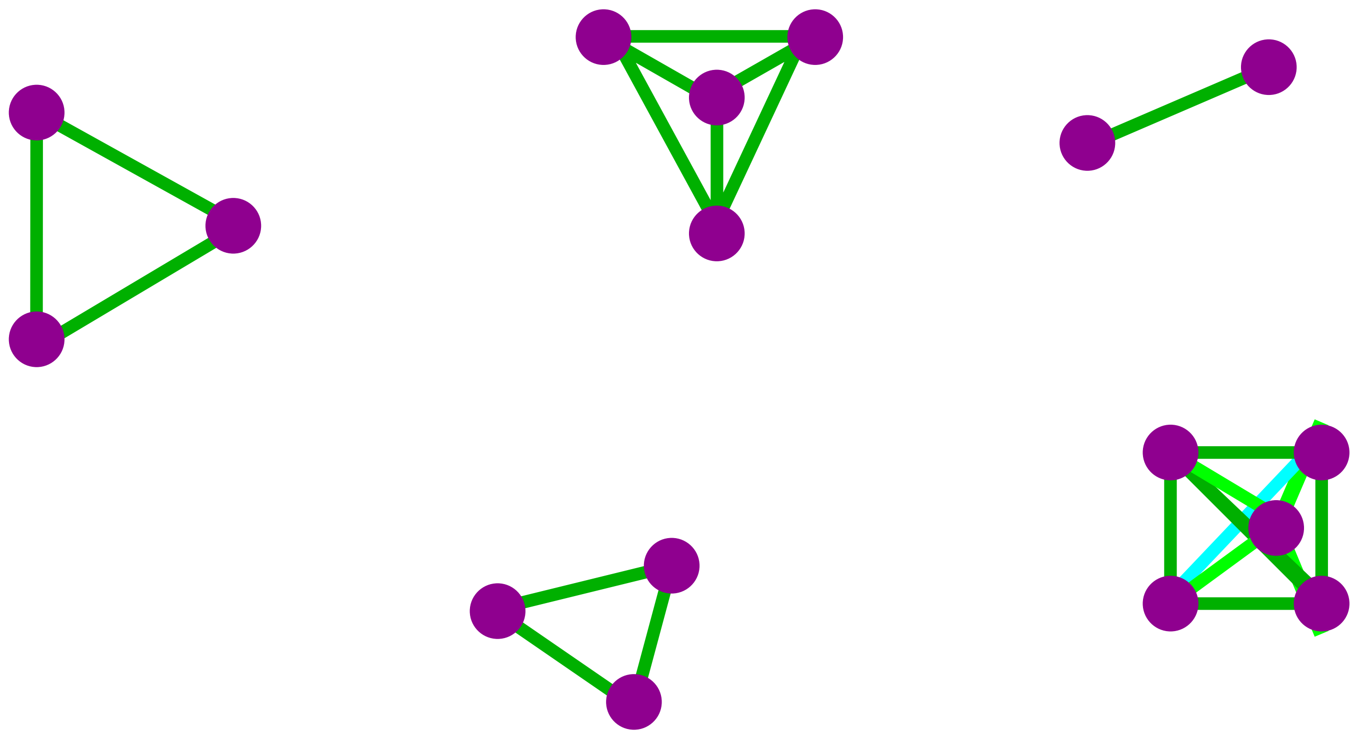
How to choose the threshold? Don't, instead use persistent homology

Creating the Vietoris Rips simplicial complex

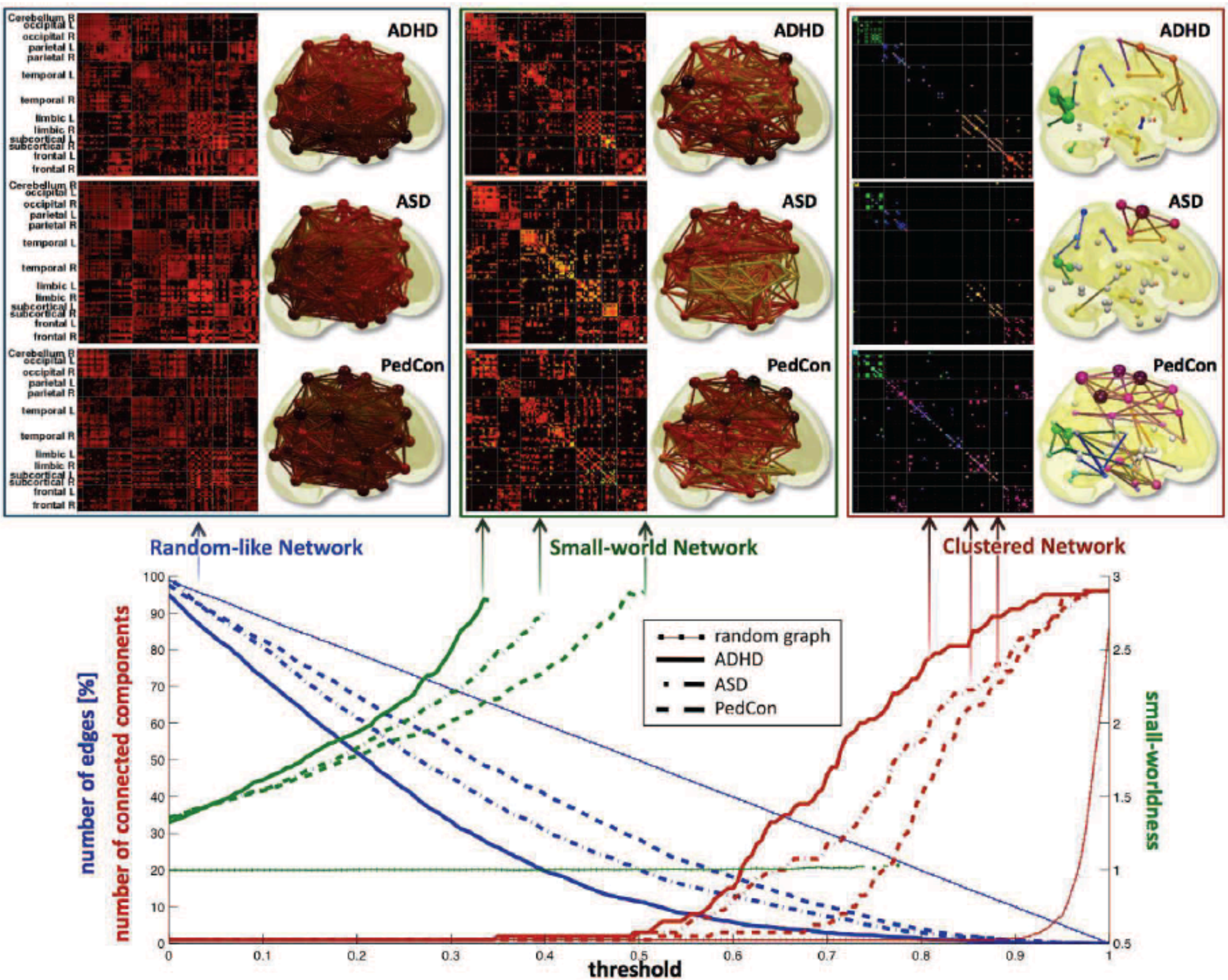


Note to determine clusters, only need vertices and edges (but we will use higher dimensional simplices Friday to look at holes in data).

Creating a simplicial complex



Note to determine clusters, only need vertices and edges.



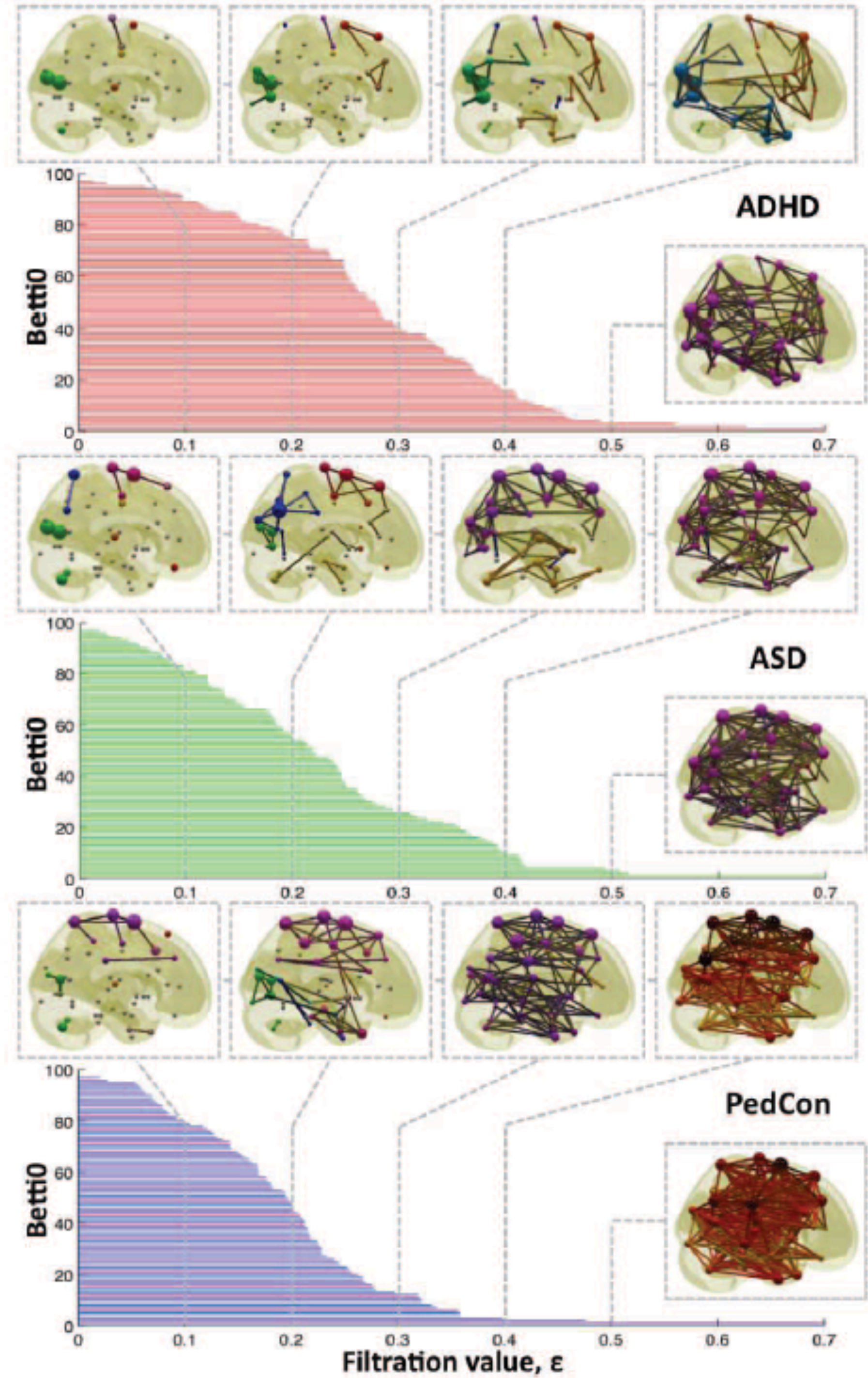
Cerebellum R
occipital L
occipital R
parietal L
parietal R
temporal L
temporal R
limbic L
limbic R
subcortical L
subcortical R
frontal L
frontal R
ADHD
ASD
PedCon
Random-like Network
Small-world Network
Clustered Network
random graph
ADHD
ASD
PedCon
number of edges [%]
number of connected components
small-worldness
threshold

Vertices = Regions of Interest

Create Rips complex by growing epsilon balls (i.e. decreasing threshold) where distance between two vertices is given by

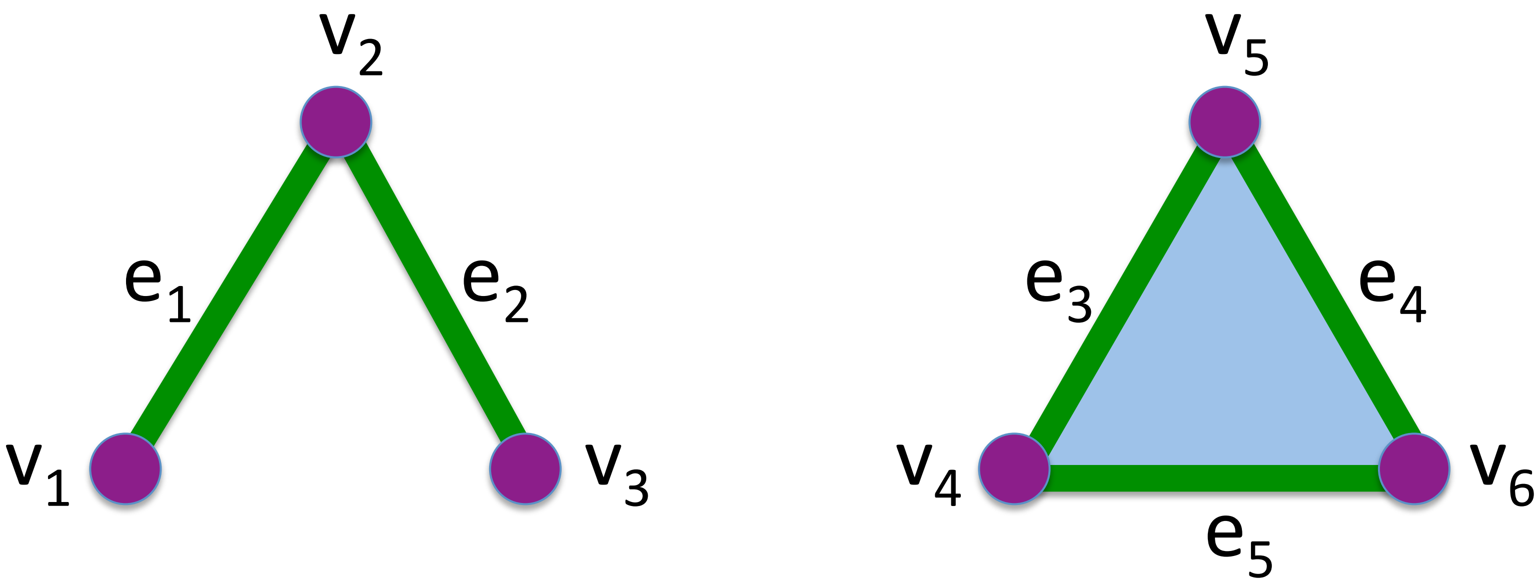
$$d(\boldsymbol{f}_i, \boldsymbol{f}_j) = 1 - \mathrm{corr}(\boldsymbol{f}_i, \boldsymbol{f}_j)$$

where $\boldsymbol{f}_i$ = measurement at location *i*



How many connected components?

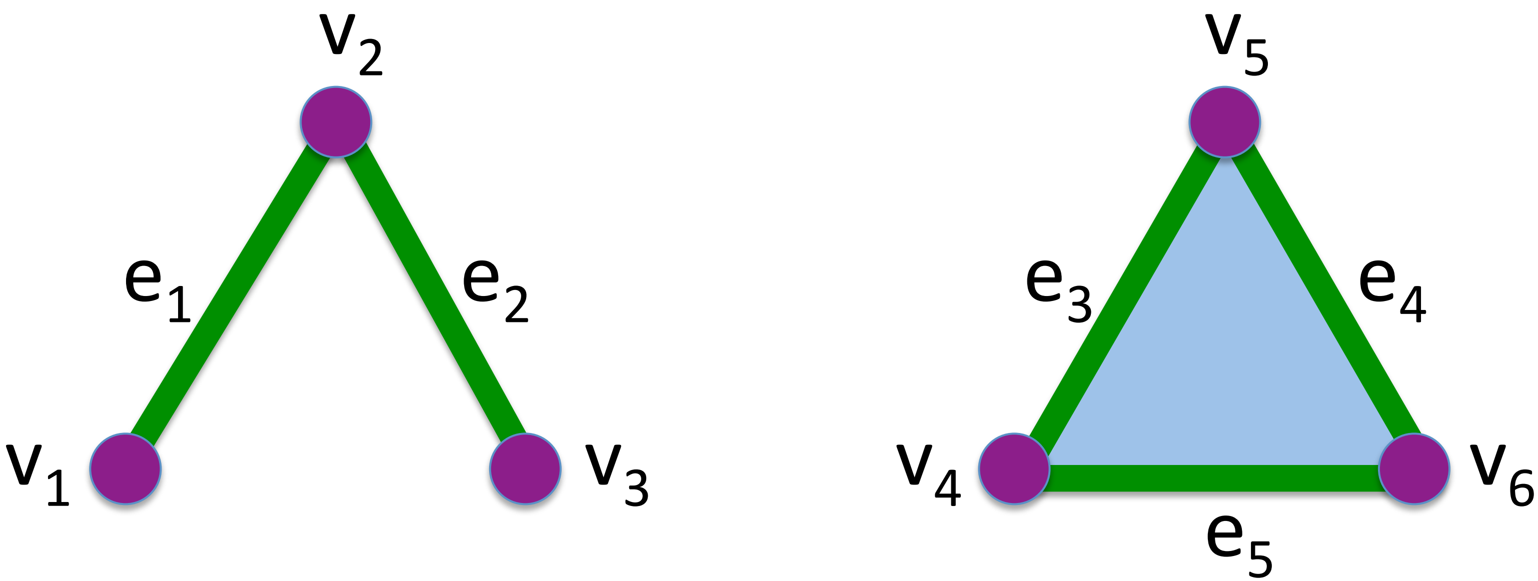
Counting number of connected components using homology



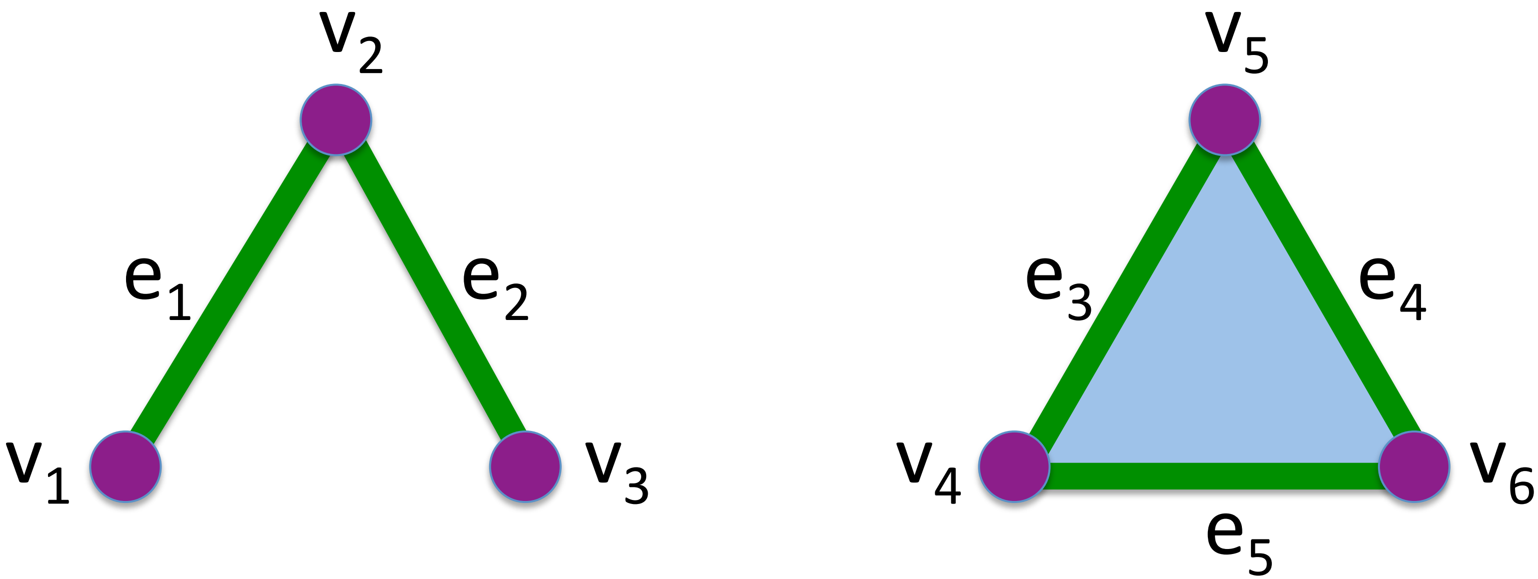
$C_0 = \mathbf{Z_2}[v_1, v_2, v_3, v_4, v_5, v_6]$ = set of 0-chains

$C_1 = \mathbf{Z_2}[e_1, e_2, e_3, e_4, e_5]$ = set of 1-chains

Counting number of connected components using homology



Counting number of connected components using homology



$\partial(e_1) = v_1 + v_2$

$\partial(e_2) = v_2 + v_3$

$\partial(e_3) = v_4 + v_5$

$\partial(e_4) = v_5 + v_6$

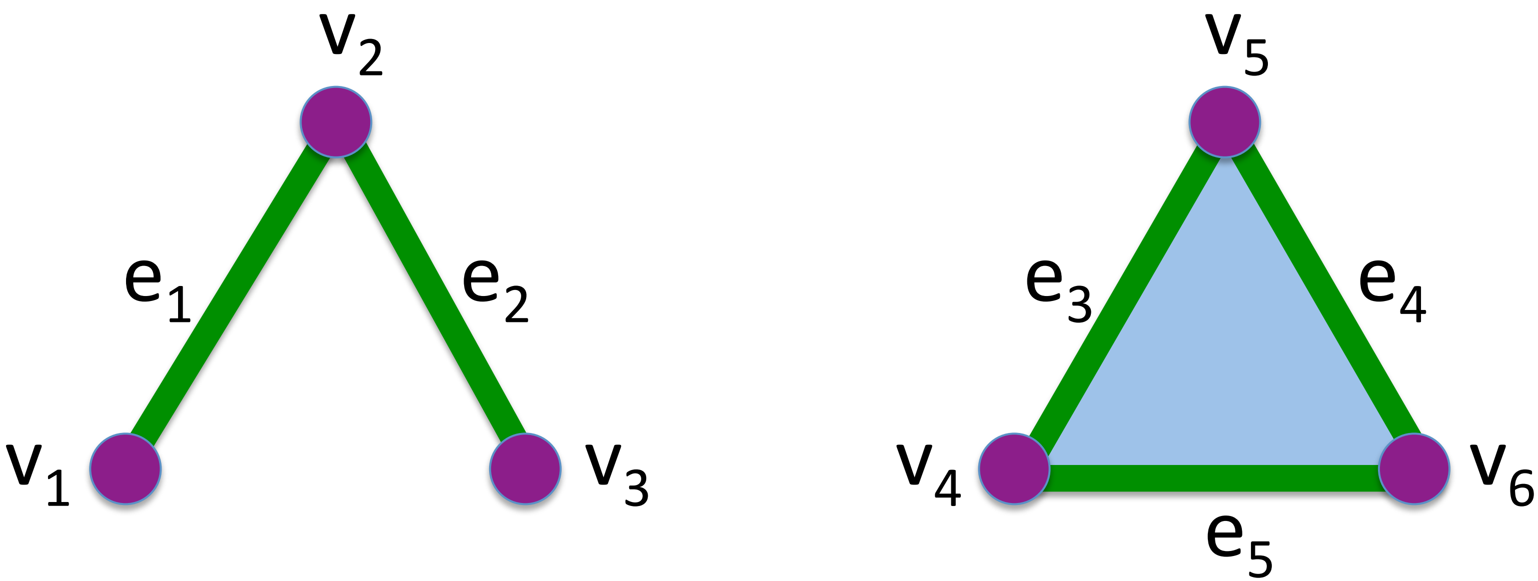
$\partial(e_5) = v_4 + v_6$

Components are

$\{v_1, v_2, v_3\}$ and

$\{v_4, v_5, v_6\}$

Counting number of connected components using homology

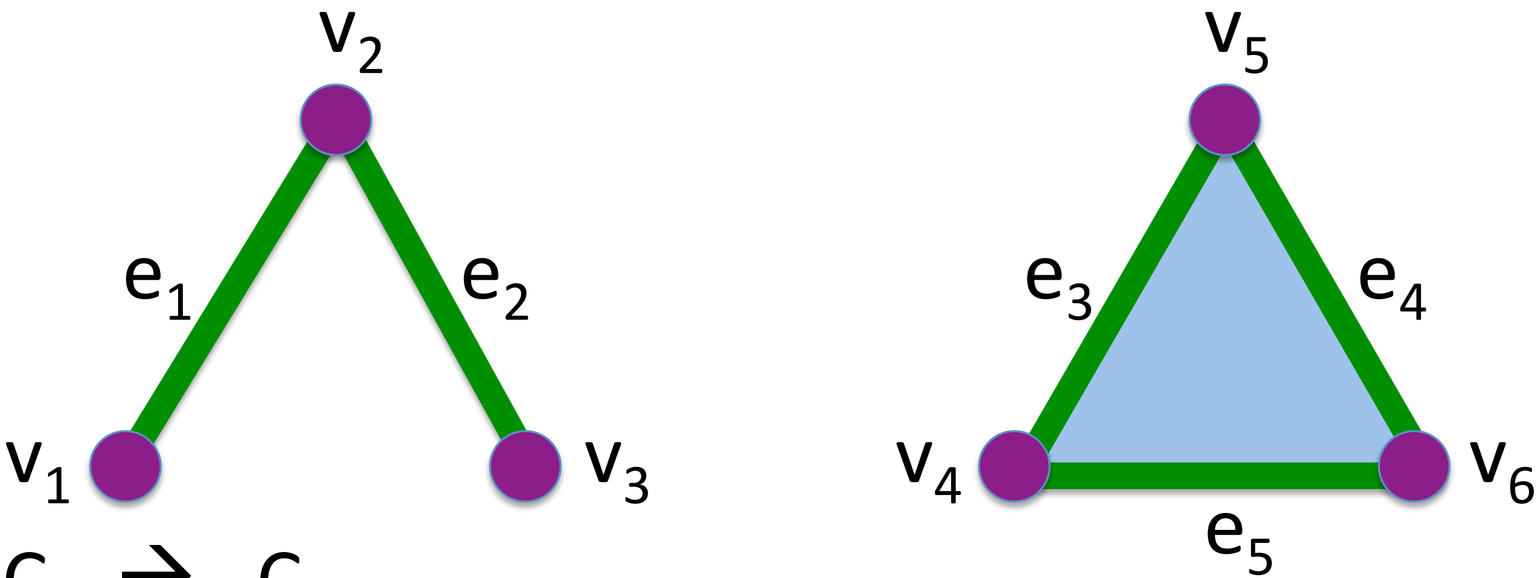


$C_0 = \mathbf{Z_2}[v_1, v_2, v_3, v_4, v_5, v_6]$ = set of 0-chains

$C_1 = \mathbf{Z_2}[e_1, e_2, e_3, e_4, e_5]$ = set of 1-chains

$\partial : C_1 \rightarrow C_0$

Counting number of connected components using homology



$\partial : C_1 \rightarrow C_0$

$$\partial(e_1) = v_1 + v_2$$

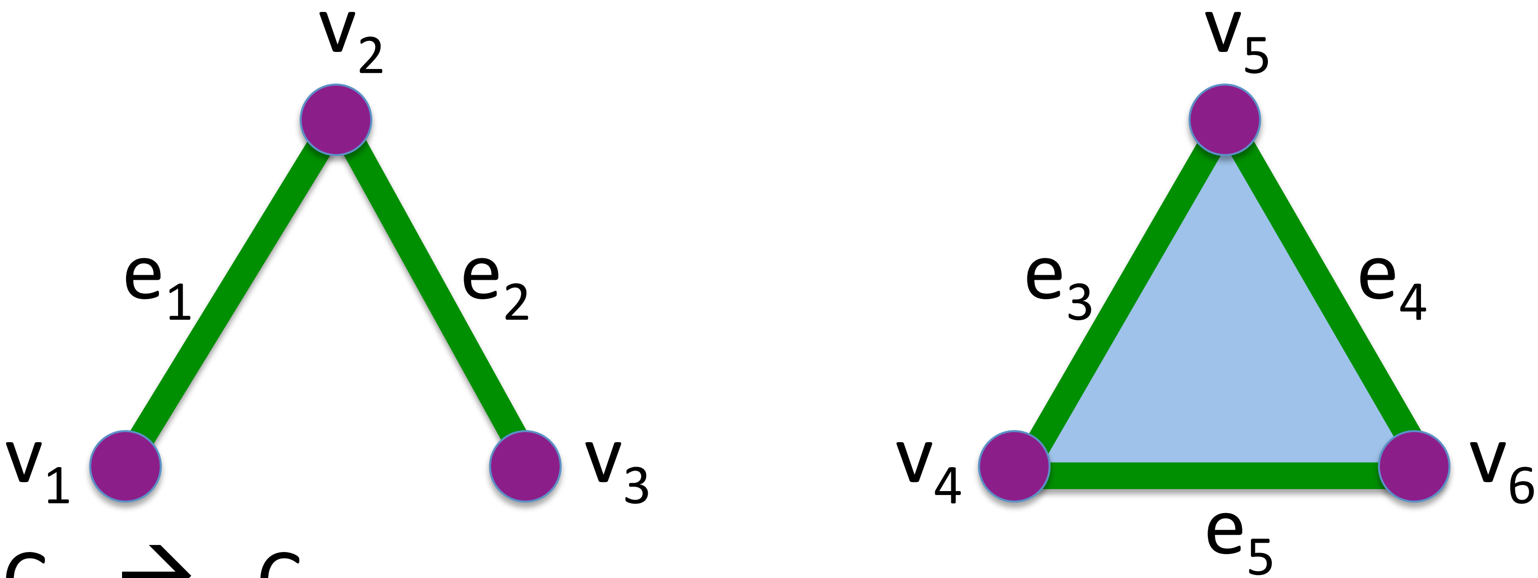
$$\partial(e_2) = v_2 + v_3$$

$$\partial(e_3) = v_4 + v_5$$

$$\partial(e_4) = v_5 + v_6$$

$$\partial(e_5) = v_4 + v_6$$

Counting number of connected components using homology



$\partial : C_1 \rightarrow C_0$

$\partial(e_1) = v_1 + v_2$

$\partial(e_2) = v_2 + v_3$

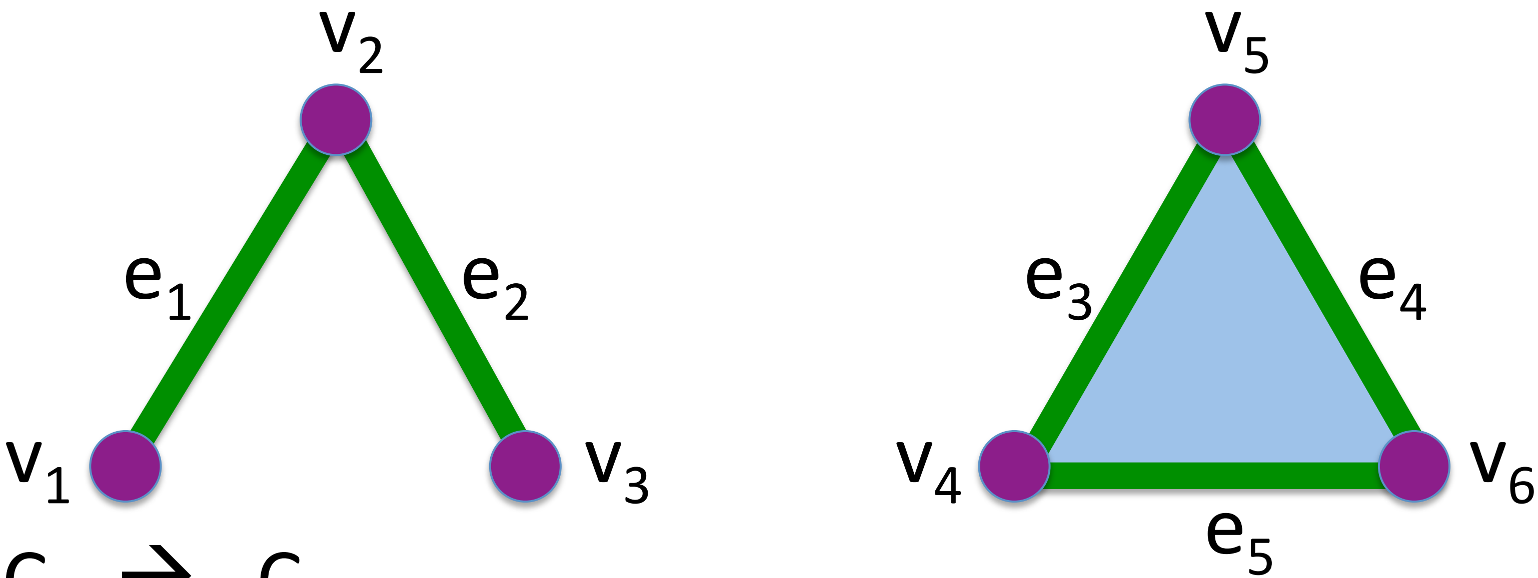
$\partial(e_3) = v_4 + v_5$

$\partial(e_4) = v_5 + v_6$

$\partial(e_5) = v_4 + v_6$

Extend linearly:

Counting number of connected components using homology



$\partial : C_1 \rightarrow C_0$

$\partial(e_1) = v_1 + v_2$

$\partial(e_2) = v_2 + v_3$

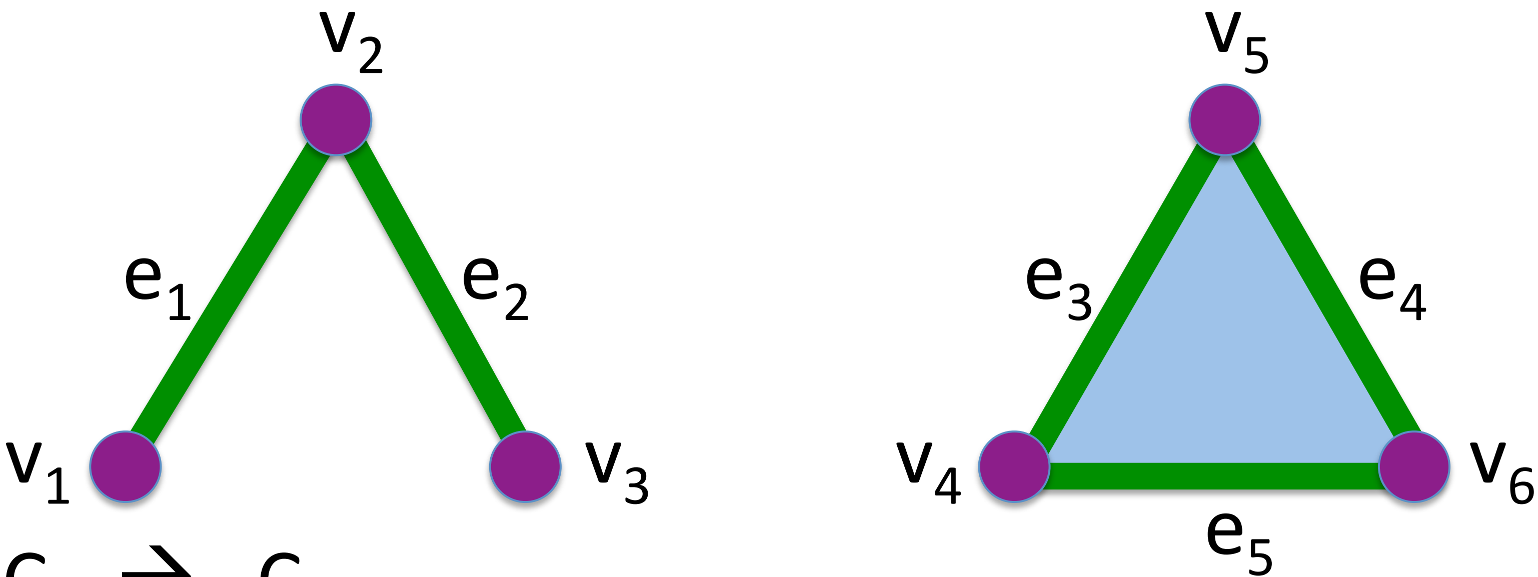
$\partial(e_3) = v_4 + v_5$

$\partial(e_4) = v_5 + v_6$

$\partial(e_5) = v_4 + v_6$

Extend linearly:

Counting number of connected components using homology



$\partial : C_1 \rightarrow C_0$

Extend linearly:

$$\partial(e_1) = v_1 + v_2$$

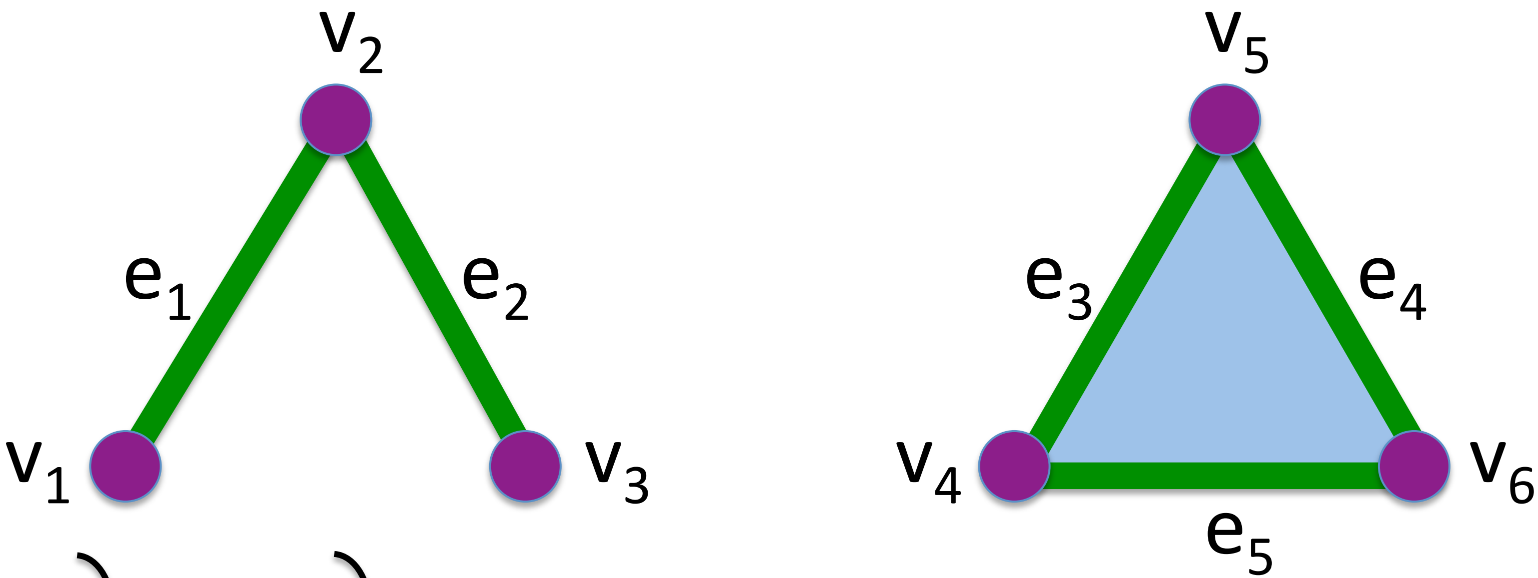
$$\partial(e_2) = v_2 + v_3$$

$$\partial(e_3) = v_4 + v_5$$

$$\partial(e_4) = v_5 + v_6$$

$$\partial(e_5) = v_4 + v_6$$

Counting number of connected components using homology

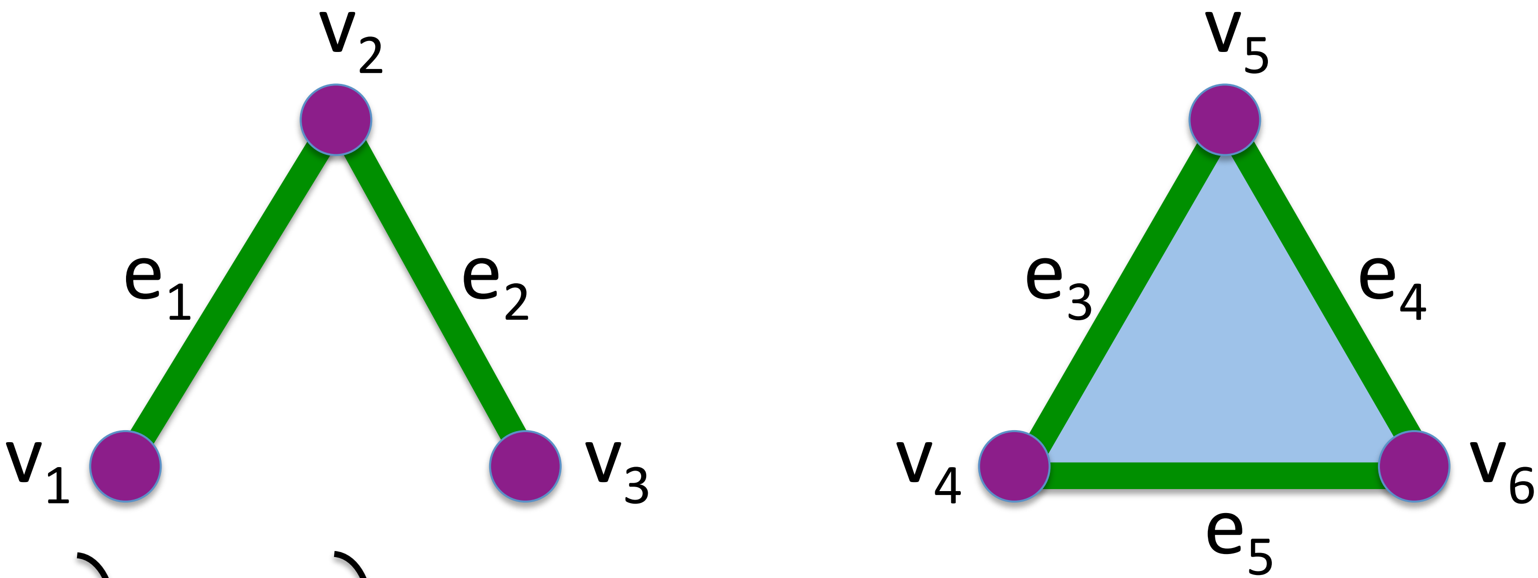


$$C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} 0$$

Z_0 = kernal of ∂_0

$= \{x : \partial_0(x) = 0\}$

Counting number of connected components using homology



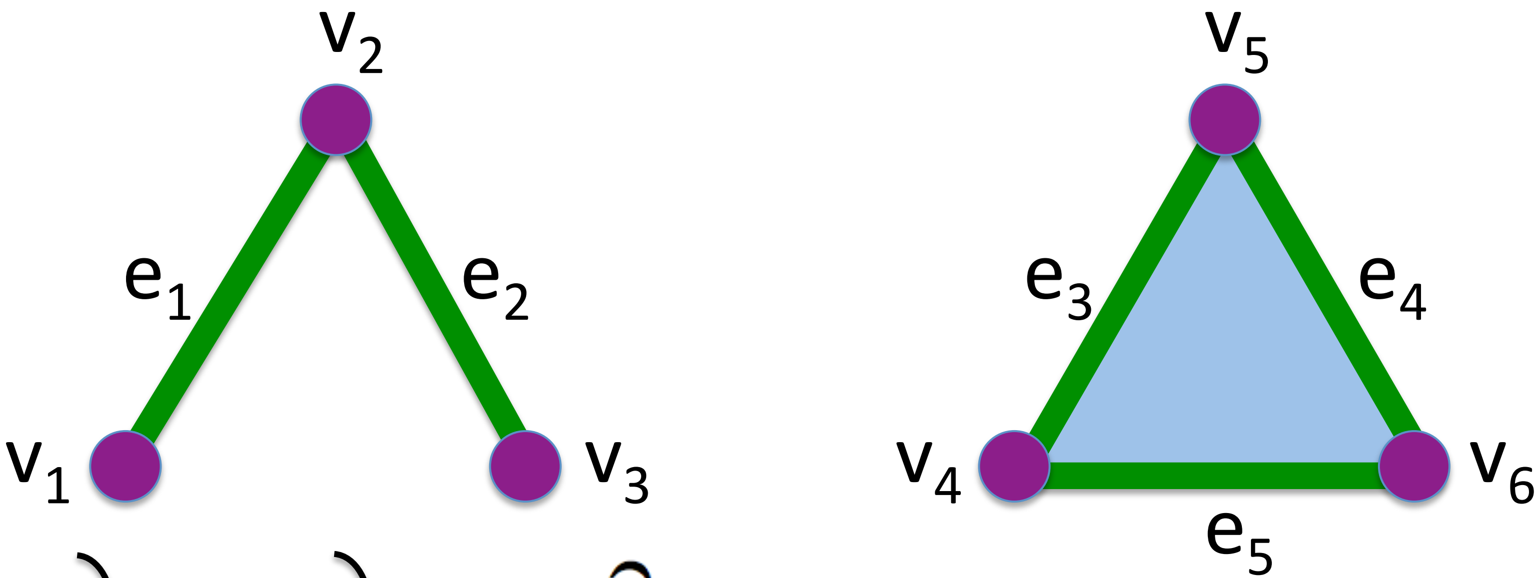
$$C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} 0$$

Z_0 = kernal of ∂_0

$= \{x : \partial_0(x) = 0\}$

$= C_0 = \mathbf{Z_2}[v_1, v_2, v_3, v_4, v_5, v_6]$

Counting number of connected components using homology



$$C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} 0 \qquad \partial$$

$\partial(e_1) = v_1 + v_2$

$\partial(e_2) = v_2 + v_3$

$\partial(e_3) = v_4 + v_5$

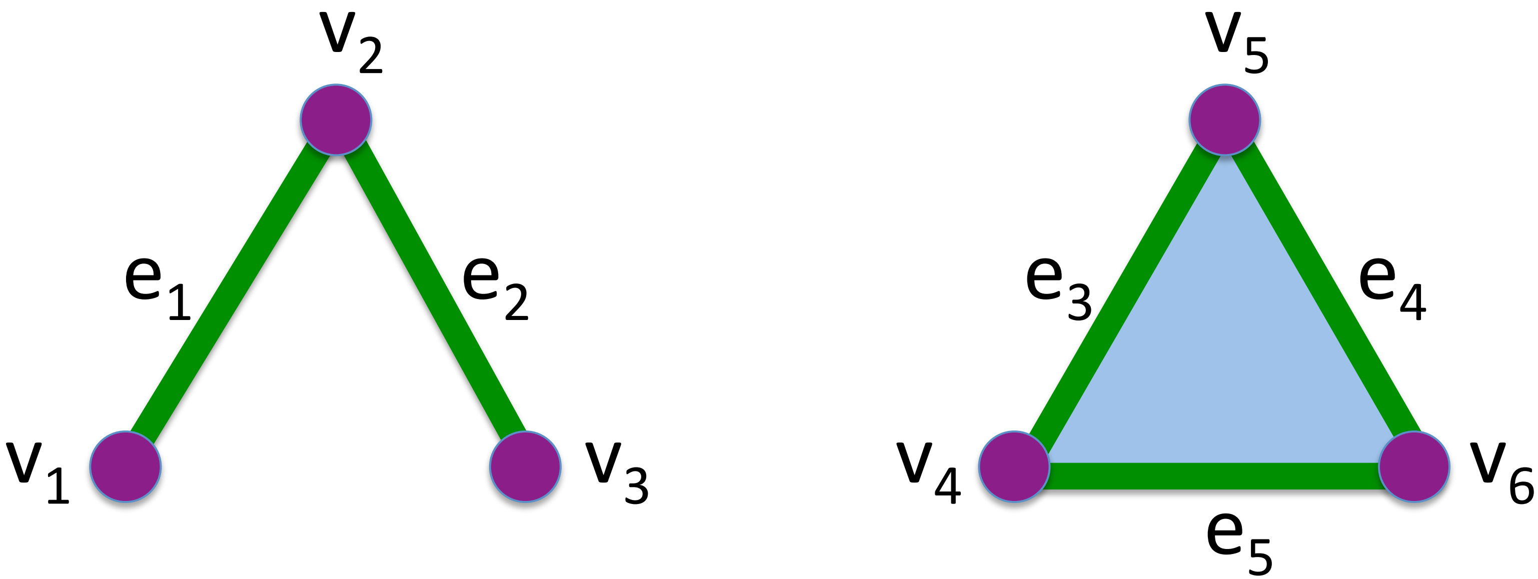
$\partial(e_4) = v_5 + v_6$

$\partial(e_5) = v_4 + v_6$

$Z_0 = \mathbf{Z_2}[v_1, v_2, v_3, v_4, v_5, v_6]$

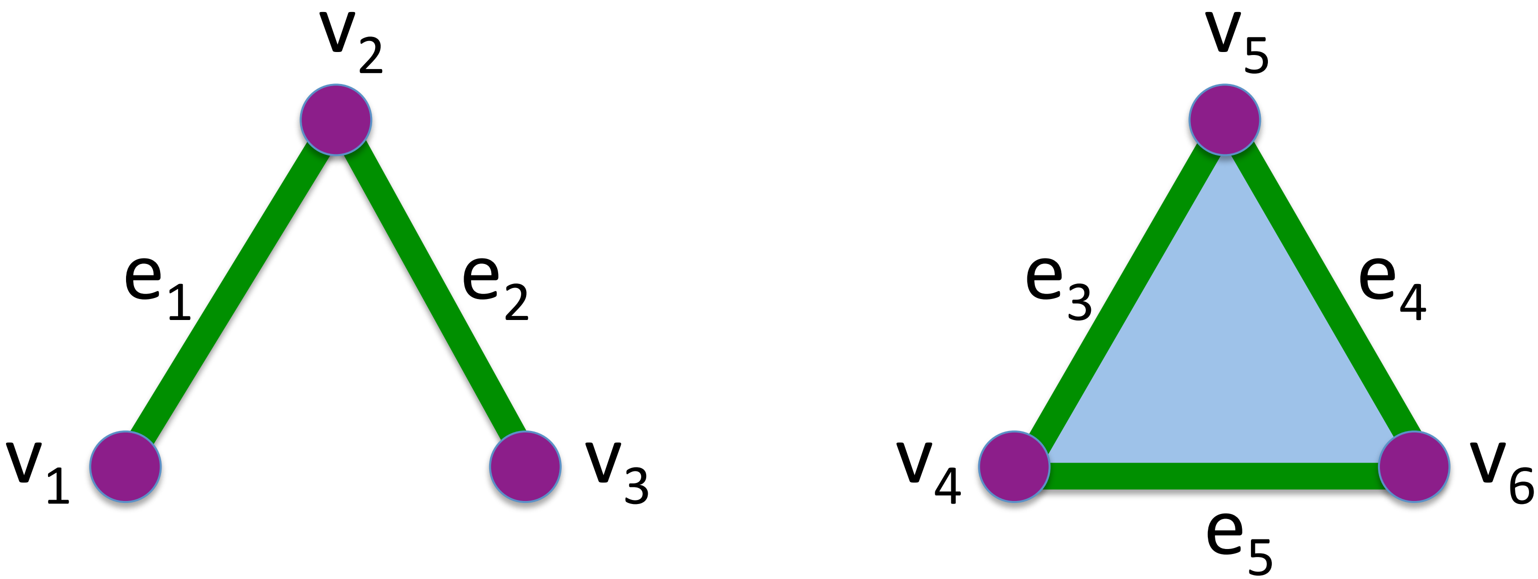
B_0 = image of ∂_1

Counting number of connected components using homology



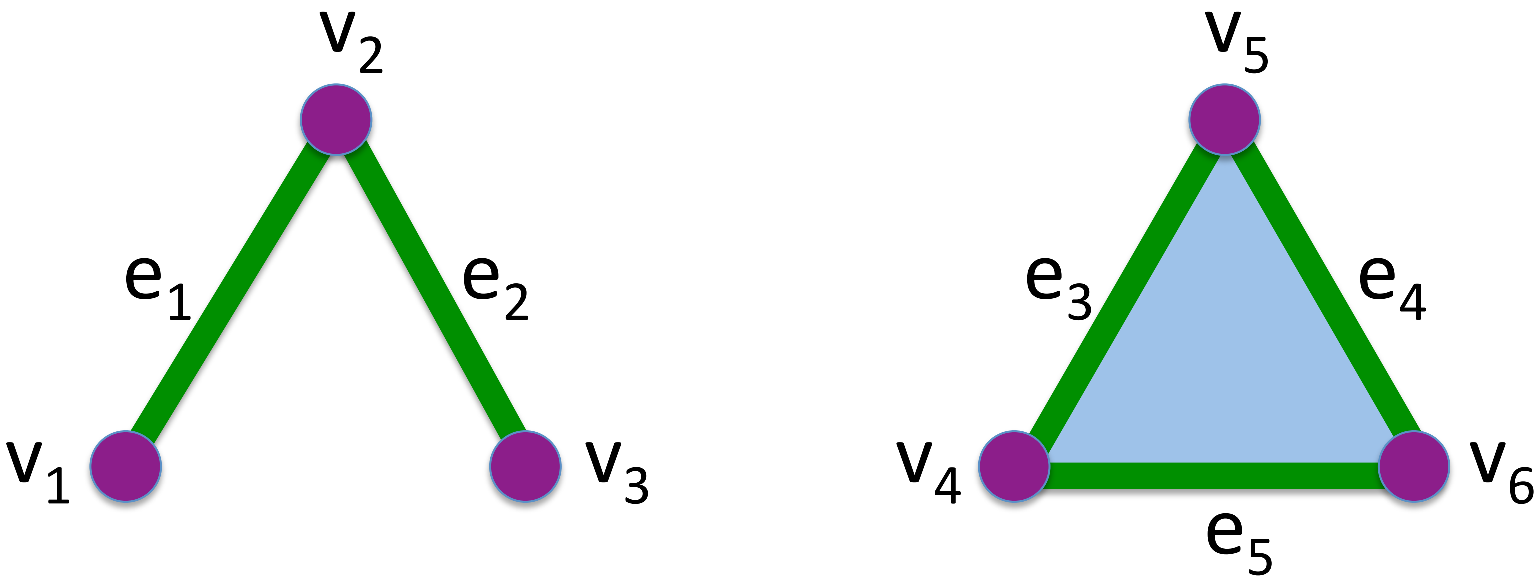
$$H_0 = Z_0/B_0 = \langle v_1, v_2, v_3, v_4, v_5, v_6 : v_1 + v_2 = 0,$$
$$v_2 + v_3 = 0,$$
$$v_4 + v_5 = 0,$$
$$v_5 + v_6 = 0,$$
$$v_4 + v_6 = 0\rangle$$

Counting number of connected components using homology



$Z_0/B_0 = \langle v_1, v_2, v_3, v_4, v_5, v_6 : v_1 + v_2 = 0, v_2 + v_3 = 0,$
$v_4 + v_5 = 0, v_5 + v_6 = 0, v_4 + v_6 = 0\rangle$

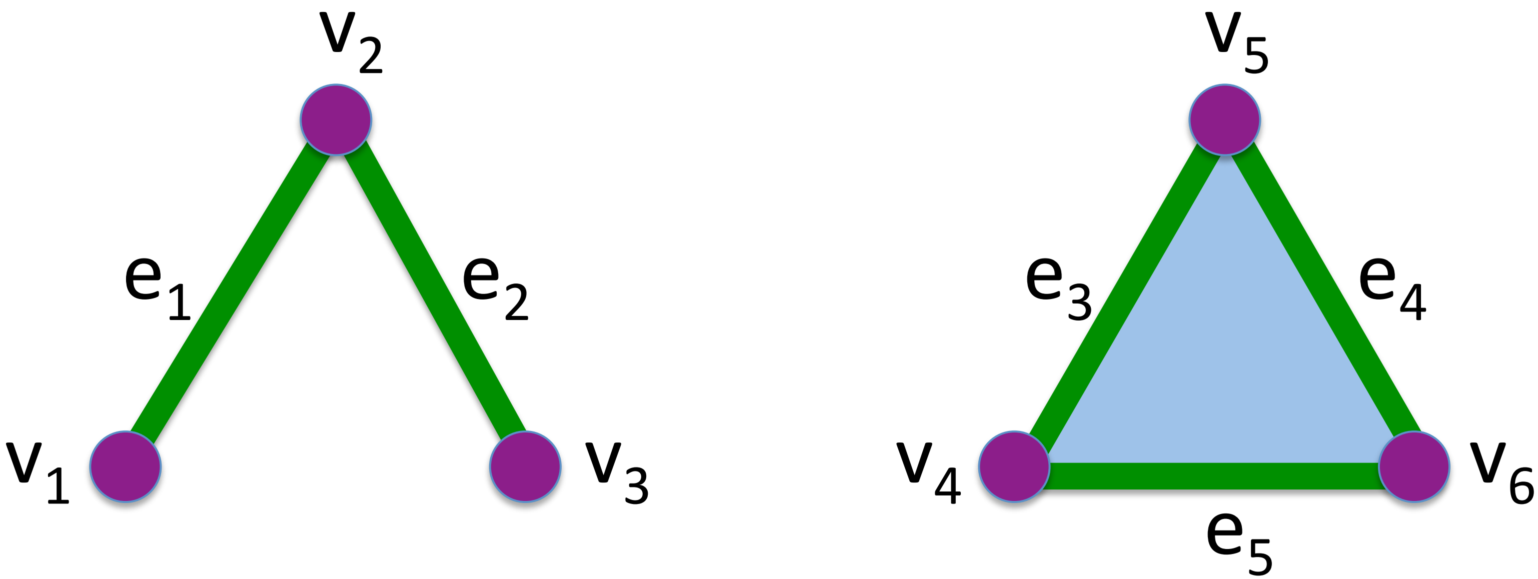
Counting number of connected components using homology



$Z_0/B_0 = \langle v_1, v_2, v_3, v_4, v_5, v_6 : v_1 + v_2 = 0,\ v_2 + v_3 = 0,$
$v_4 + v_5 = 0,\ v_5 + v_6 = 0,\ v_4 + v_6 = 0 \rangle$

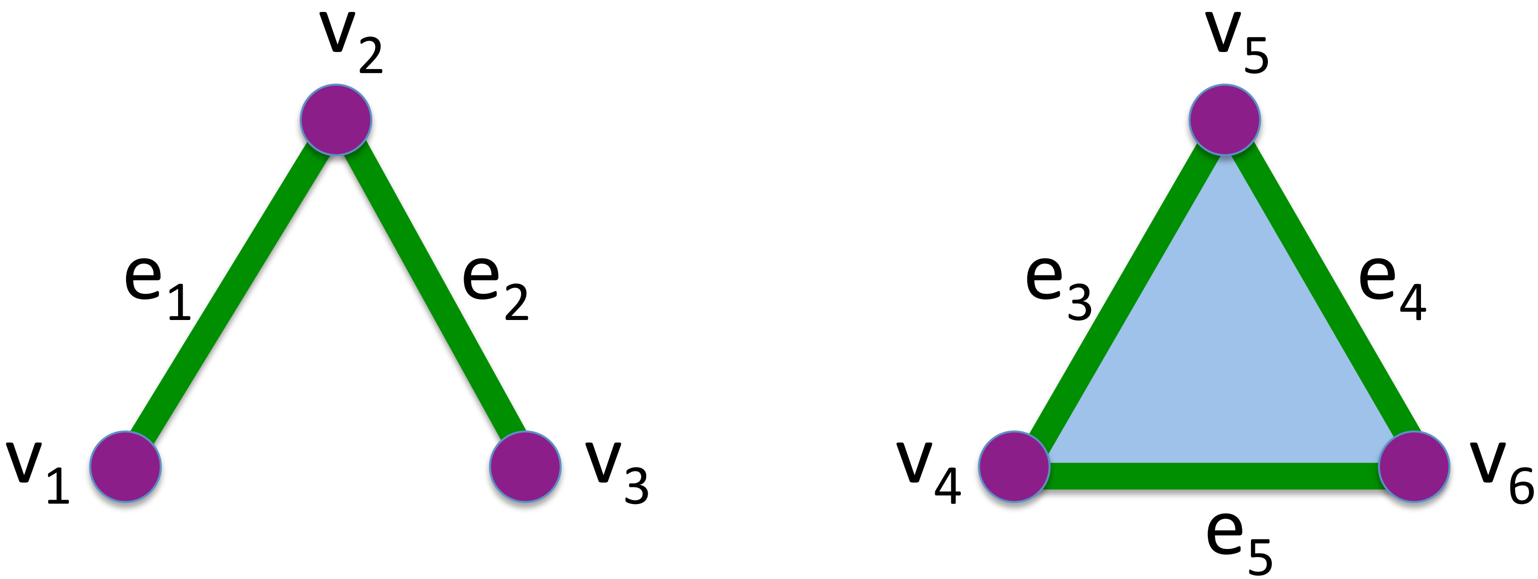
$= \langle v_1, v_2, v_3, v_4, v_5, v_6 : v_1 + v_2,\ v_2 + v_3,\ v_4 + v_5,\ v_5 + v_6,$
$v_4 + v_6 \rangle$

Counting number of connected components using homology



$Z_0/B_0 = \langle v_1, v_2, v_3, v_4, v_5, v_6 : v_1 + v_2 = 0, v_2 + v_3 = 0,$
$v_4 + v_5 = 0, v_5 + v_6 = 0, v_4 + v_6 = 0\rangle$

Counting number of connected components using homology

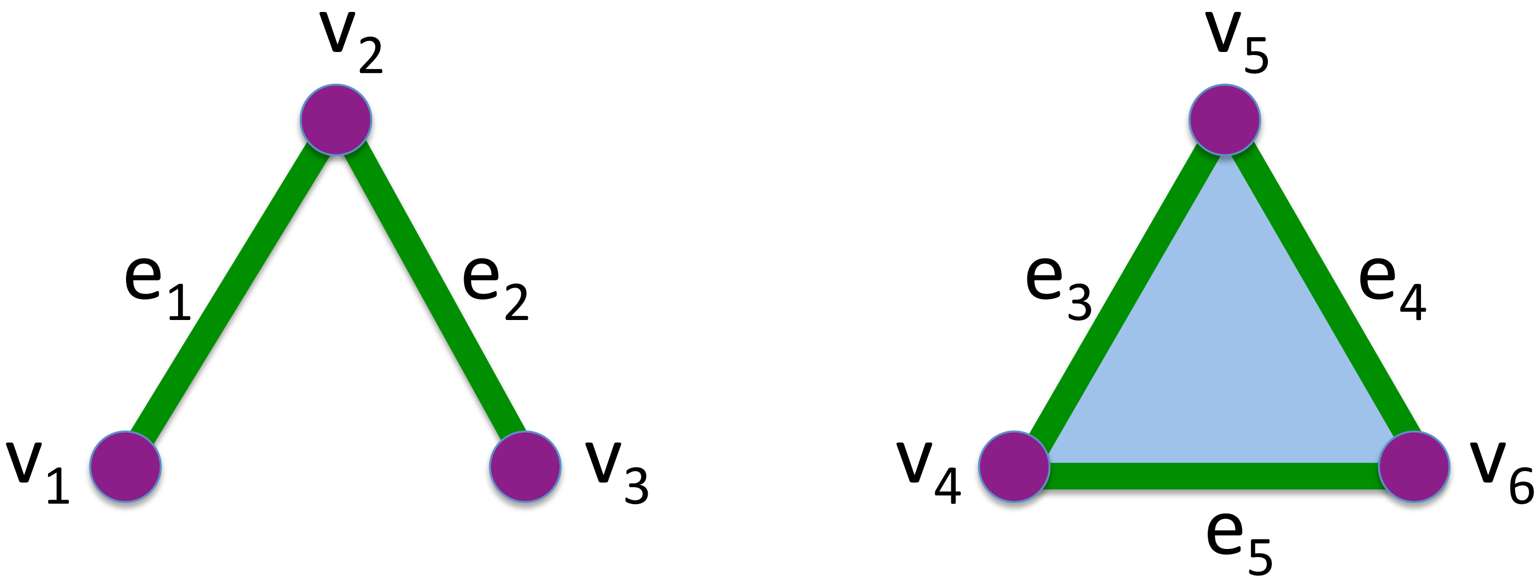


Z_0/B_0 = < $v_1, v_2, v_3, v_4, v_5, v_6 : v_1 + v_2 = 0, v_2 + v_3 = 0, v_4 + v_5 = 0, v_5 + v_6 = 0, v_4 + v_6 = 0$ >

$v_1 + v_2 = 0$ implies $v_1 = v_2$

Z_0/B_0 = < $v_1, v_3, v_4, v_5, v_6 : v_2 + v_3 = 0, v_4 + v_5 = 0, v_5 + v_6 = 0, v_4 + v_6 = 0$ >

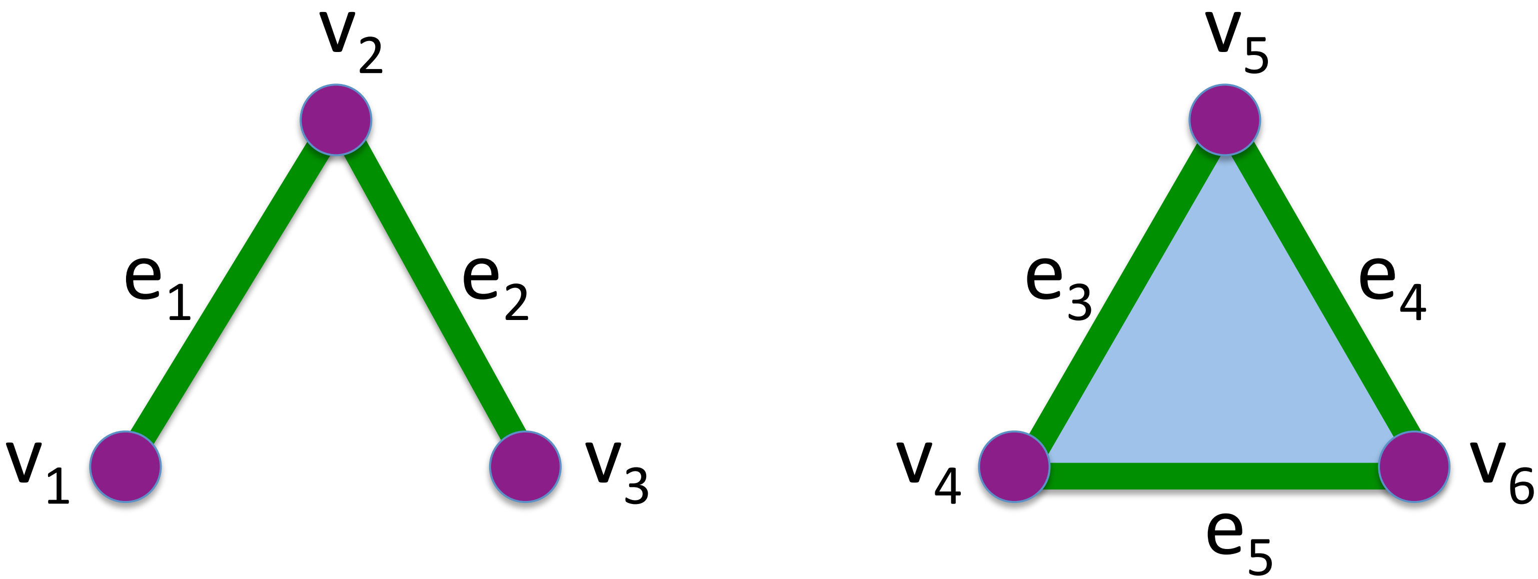
Counting number of connected components using homology



$Z_0/B_0 = \langle v_1, v_3, v_4, v_5, v_6 : v_2 + v_3 = 0, v_4 + v_5 = 0,$
$v_5 + v_6 = 0, v_4 + v_6 = 0\rangle$

$Z_0/B_0 = \langle v_1, v_4\rangle$

Counting number of connected components using homology



$Z_0/B_0 = \langle v_1, v_2, v_3, v_4, v_5, v_6 : v_1 + v_2 = 0, v_2 + v_3 = 0,$
$v_4 + v_5 = 0, v_5 + v_6 = 0, v_4 + v_6 = 0\rangle$

$Z_0/B_0 = \langle [v_1], [v_4]\rangle$ where $[v_1] = \{v_1, v_2, v_3\}$

and $[v_4] = \{v_4, v_5, v_6\}$

Counting number of connected components using homology

$Z_0/B_0 = \langle v_1, v_2, v_3, v_4, v_5, v_6 : v_1 + v_2 = 0, v_2 + v_3 = 0, v_4 + v_5 = 0, v_5 + v_6 = 0, v_4 + v_6 = 0 \rangle$

Use matrices:

See Computing Persistent Homology by Afra Zomorodian, Gunnar Carlsson