Topological Data Analysis

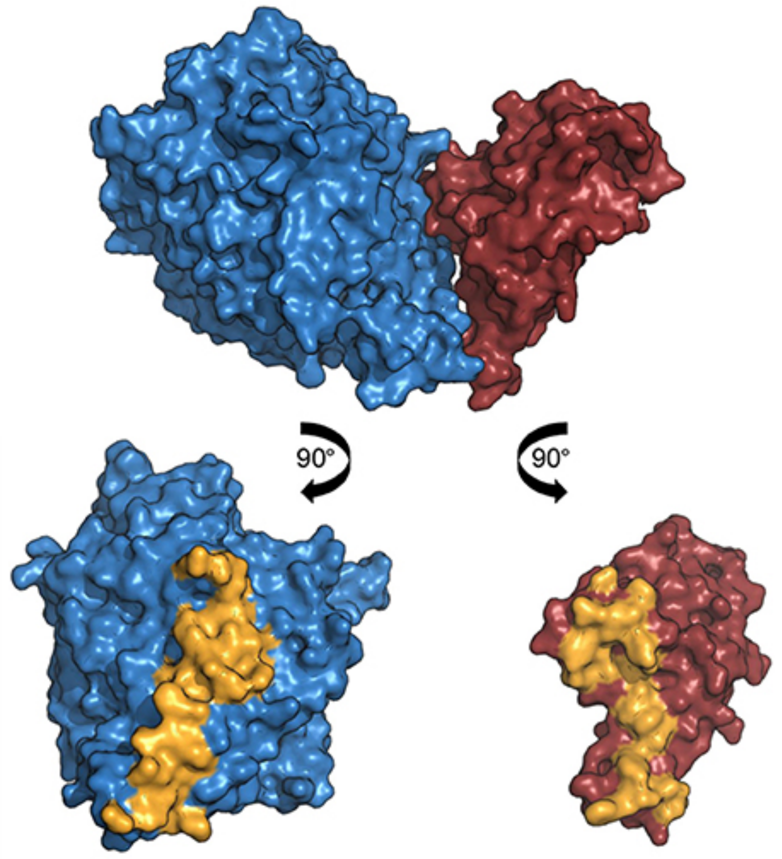
Lecture 6

Higher-order Network Science

Oleg Kachan

Complex networks

Social and collaboration networks



Protein interaction networks

Functional brain networks

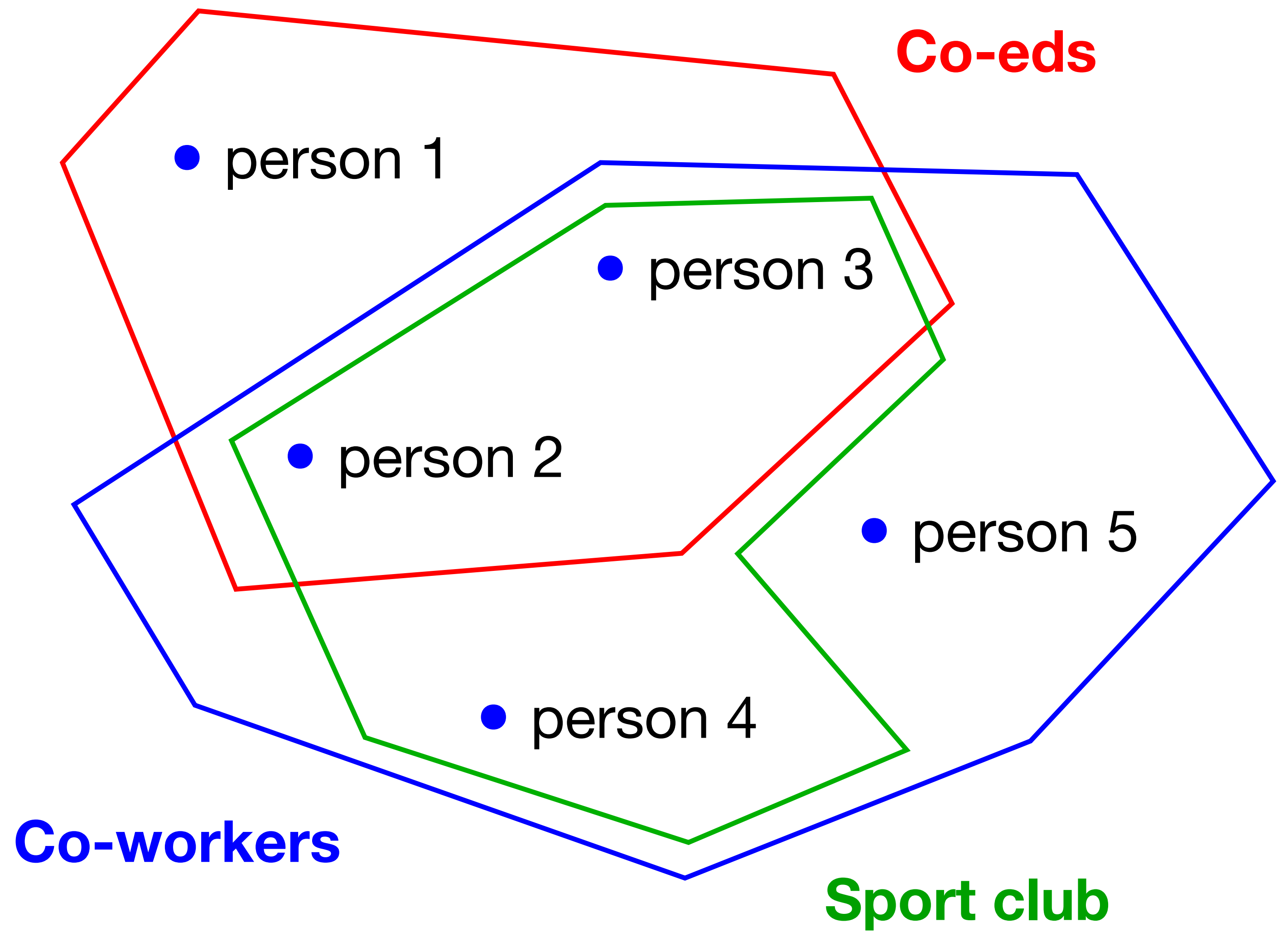
Financial networks

Networks Beyond Pairwise Interactions: Structure and Dynamics
Battiston et al. *Physics Reports (2020)*

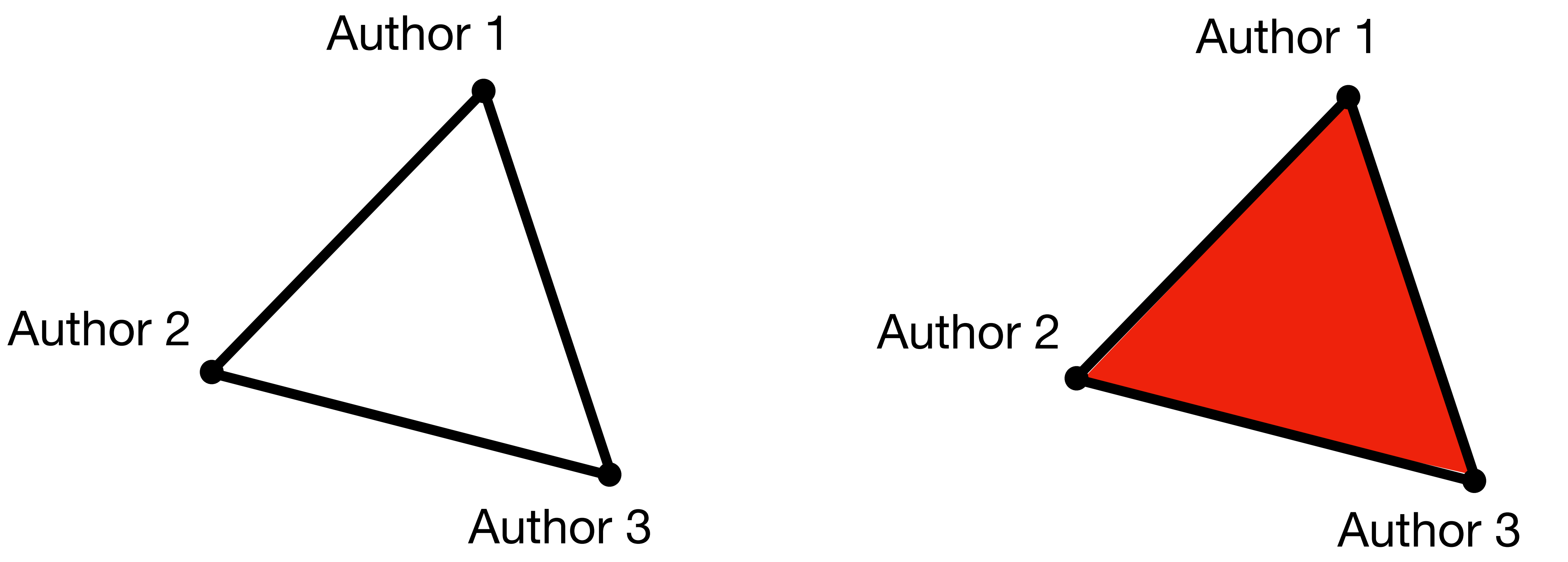
Higher-order complex networks

Collaboration networks

Social network



Collaboration network



- paper 1 (authors 1-2)
- paper 2 (authors 1-3)
- paper 3 (authors 2-3)

- paper 4 (Authors 1-2-3)

Higher-order complex networks

Co-occurence networks

Document 1

- keyword 1
- keyword 2
- keyword 3

Word-document matrix

	Word 1	Word 2	Word 3
Doc 1		1	
Doc 2	1	1	
Doc 3			1
Doc 4		1	

Basket 1

- item 1, item 2, item 3

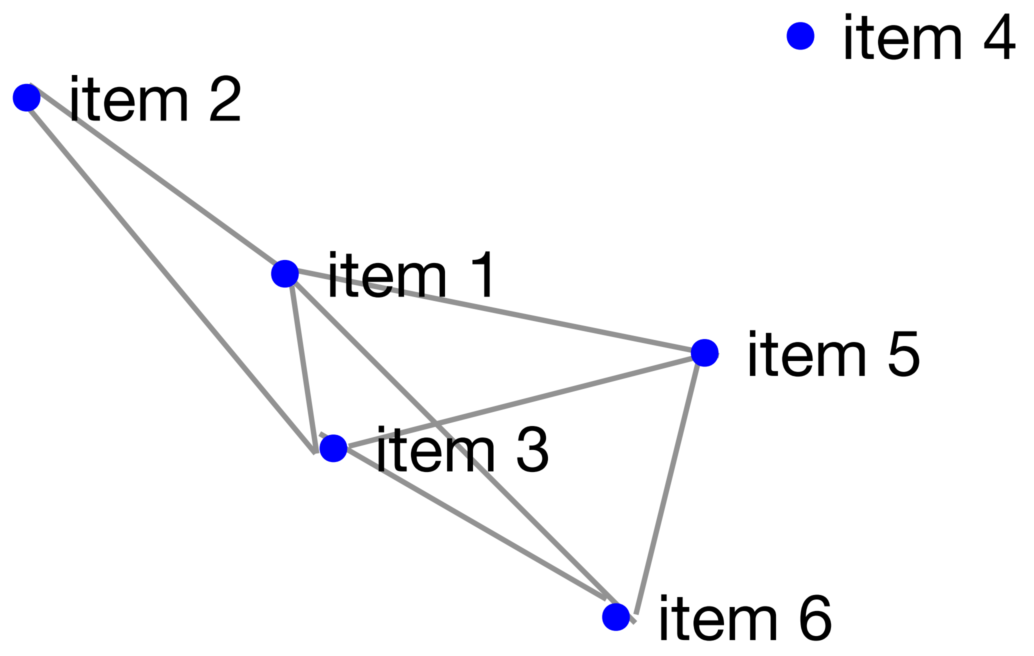
Basket 2

- item 1, item 3, item 5, item 6

Item-basket matrix

	Item 1	…	Item 6
Basket 1	1	…	
Basket 2	1	…	1

Graph
Homogenous



Context 1

quick brown **fox** jumps over

Context 2

brown fox **jumps** over lazy

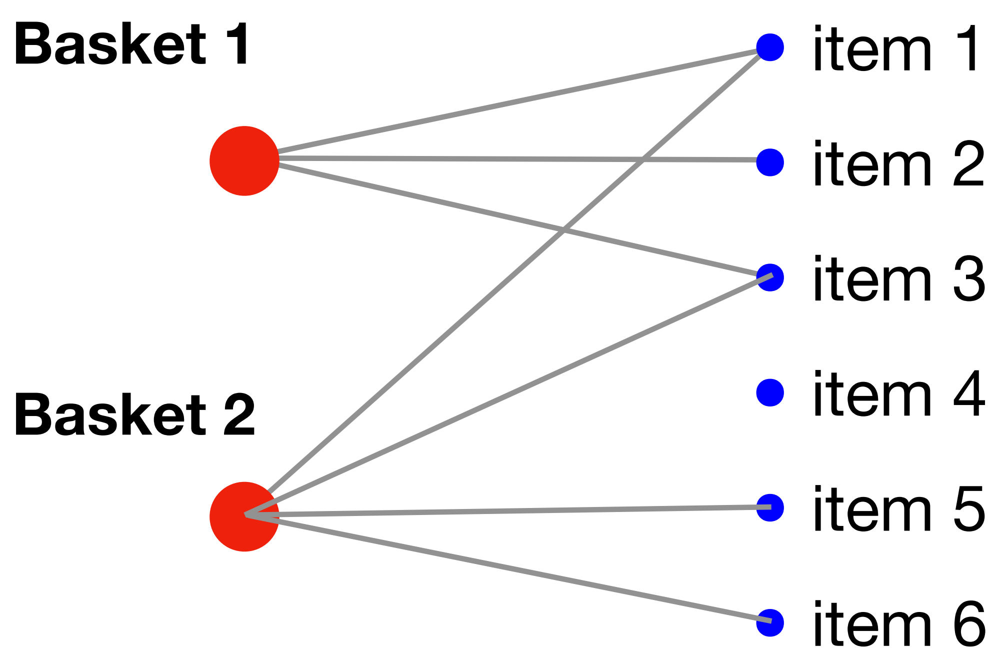
Context 3

fox jumps **over** lazy dog

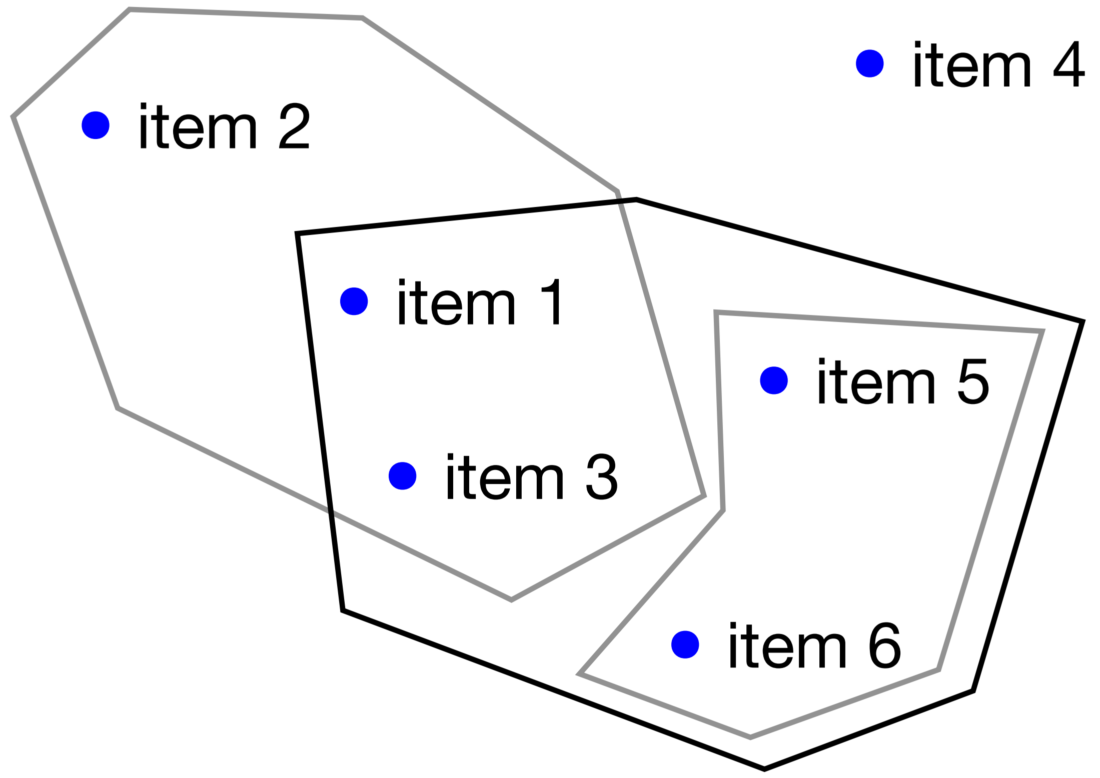
Word-context matrix

	Word 1	Word 2	Word 3
Context 1	1		
Context 2	1	1	
Context 3		1	
Context 4	1		1

Bipartite graph
Heterogenous

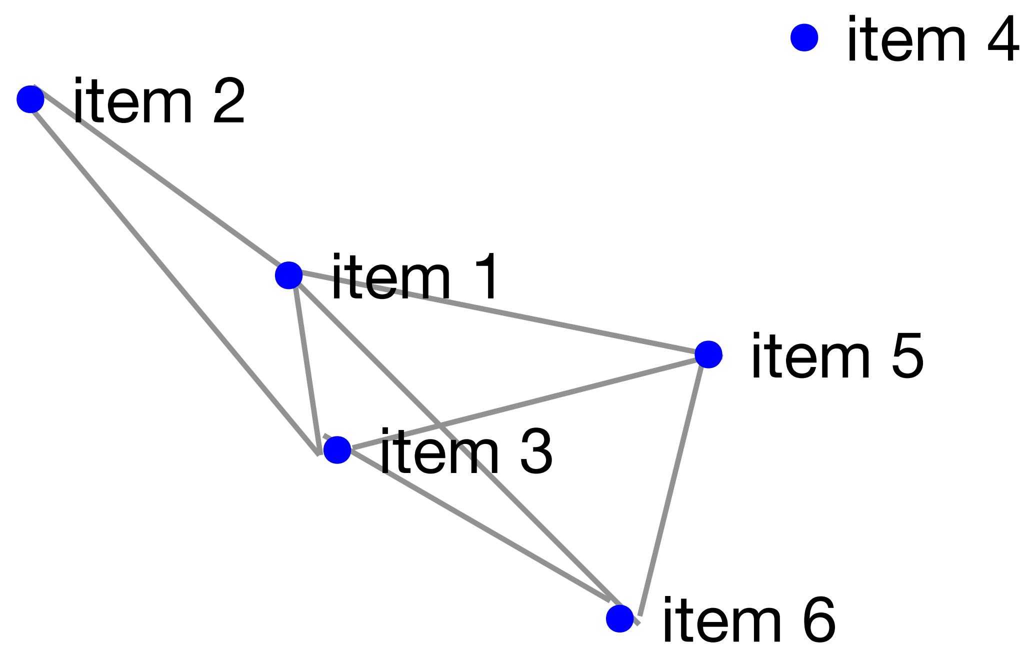


Hypergraph
Homogenous

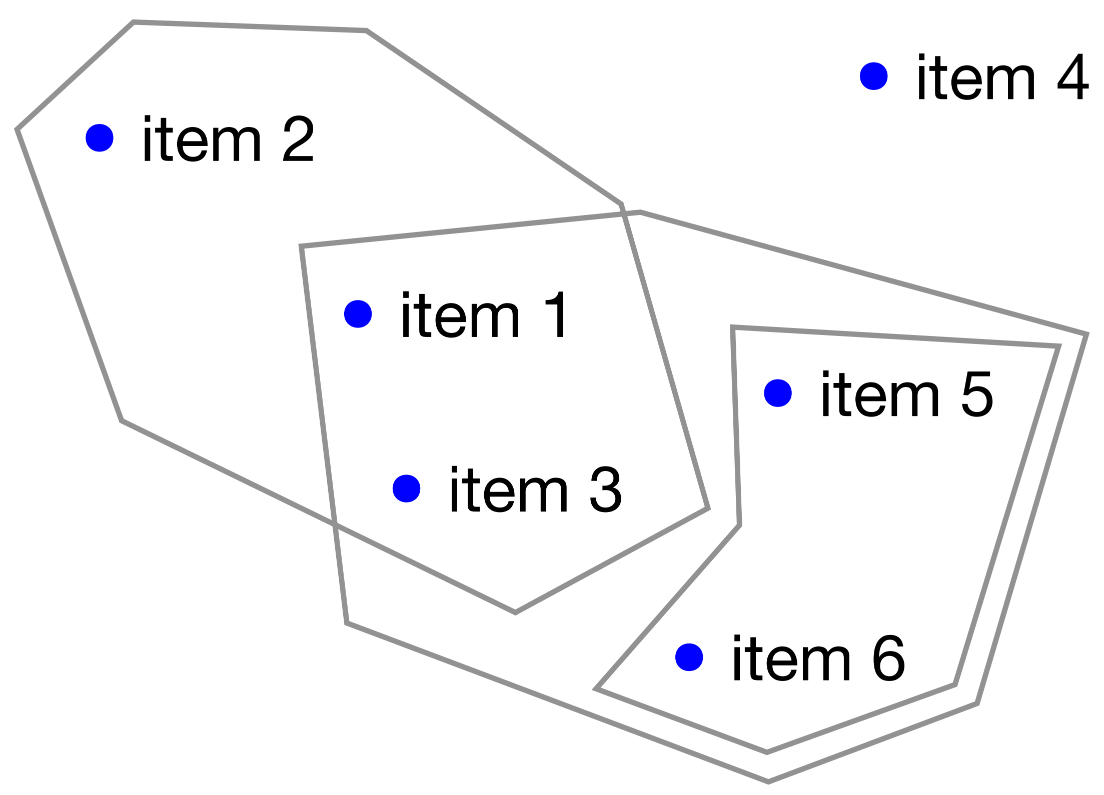


Higher-order complex networks

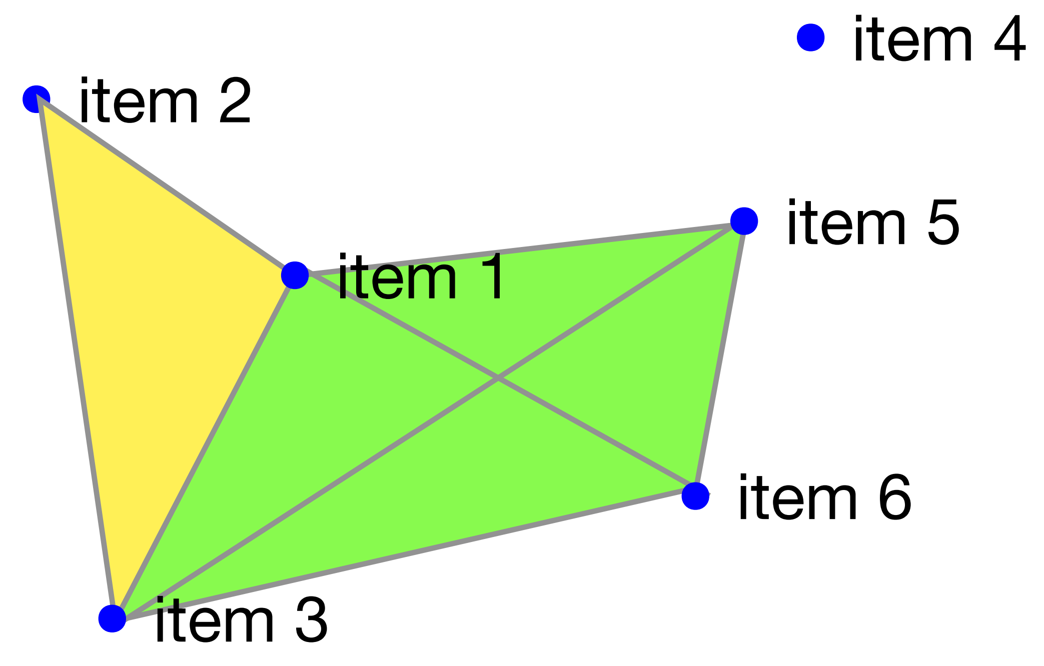
Models



Graph



Hypergraph

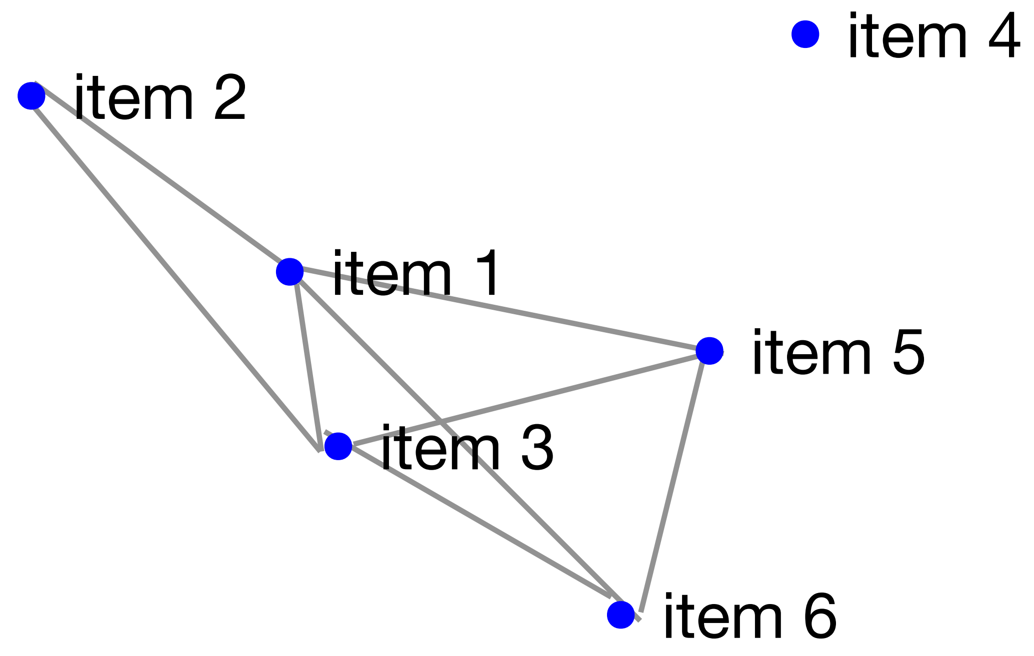


Simplicial complex

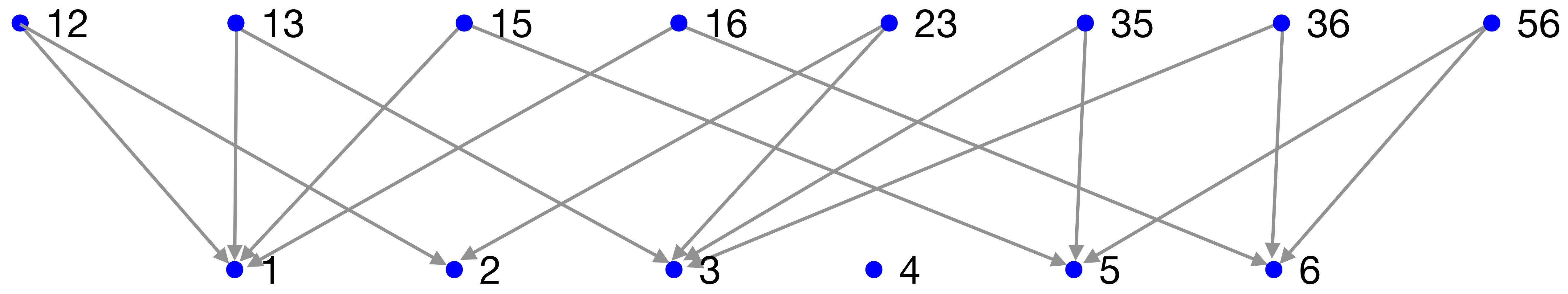
Closed under inclusion
Edges of different dimensions

Hasse diagram

Graph



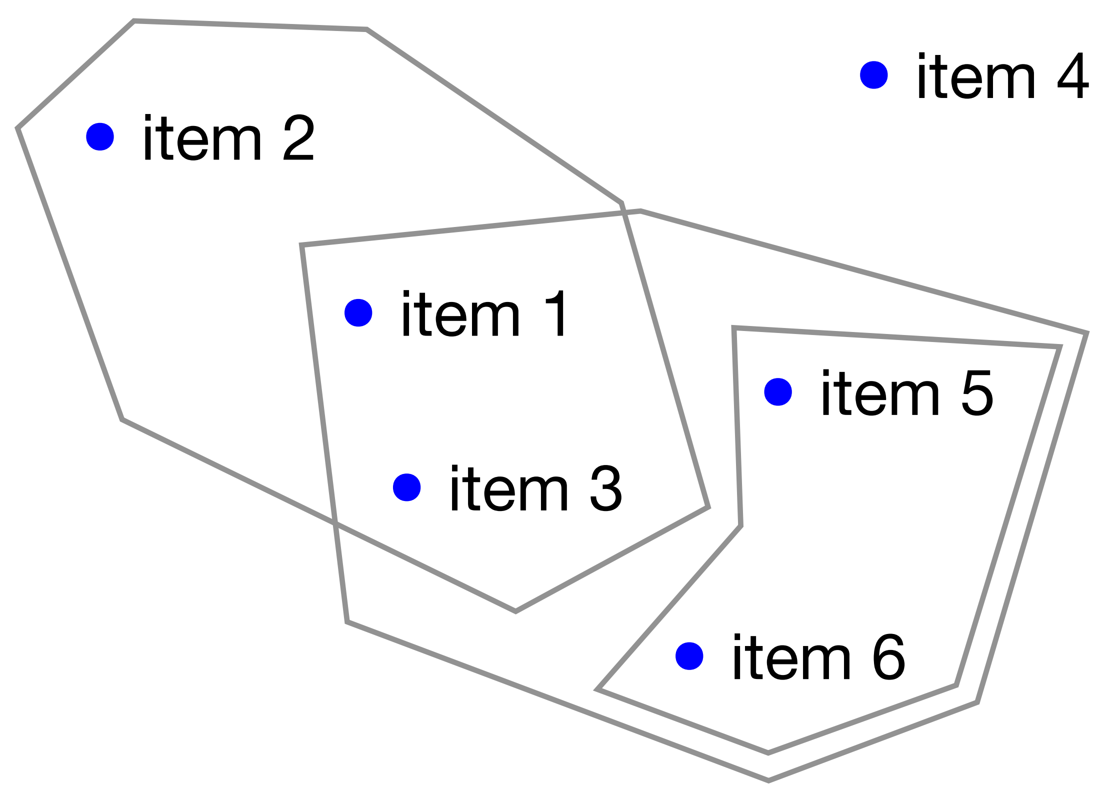
Graph



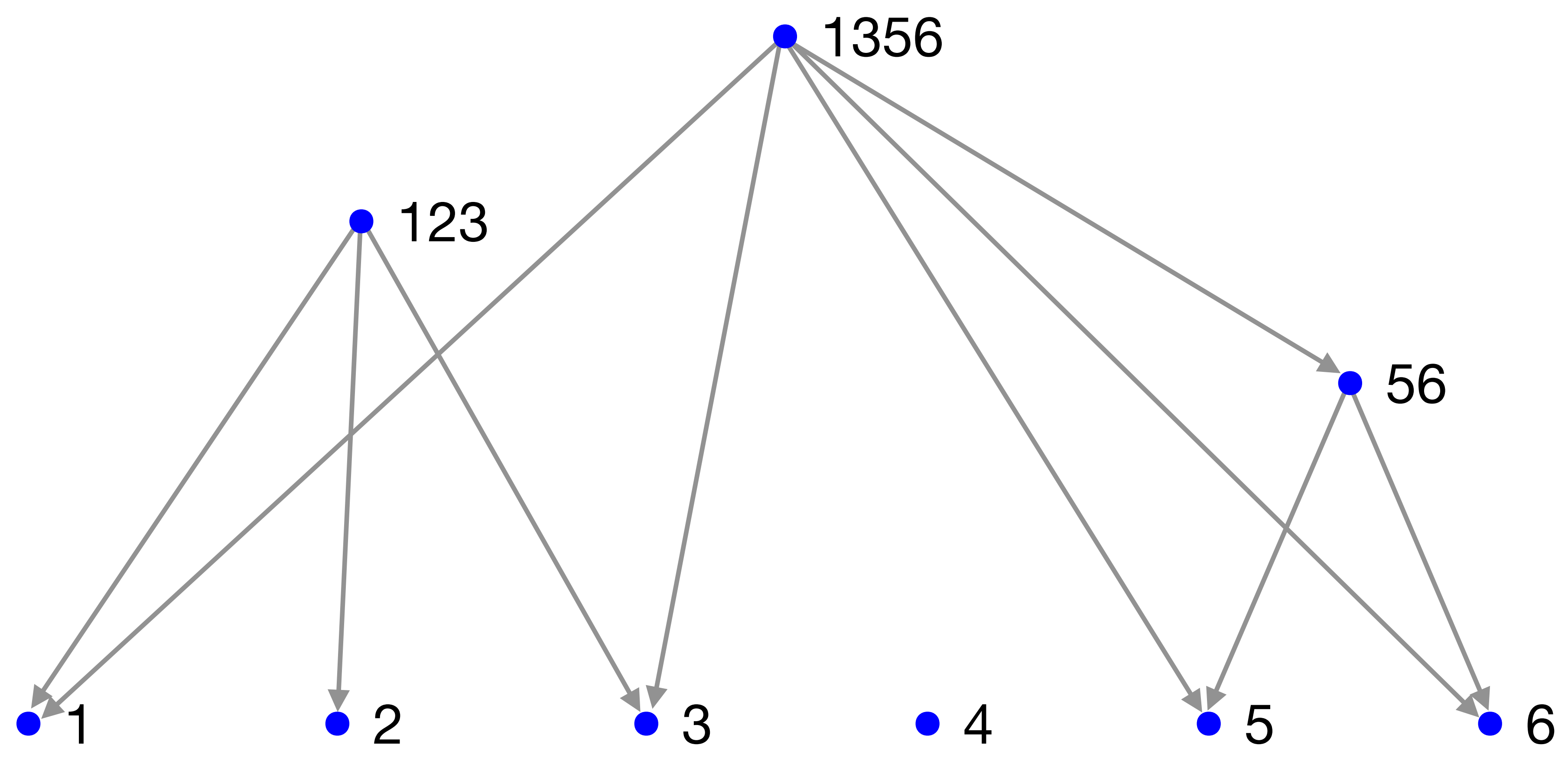
Hasse diagram

Hasse diagram

Hypergraph

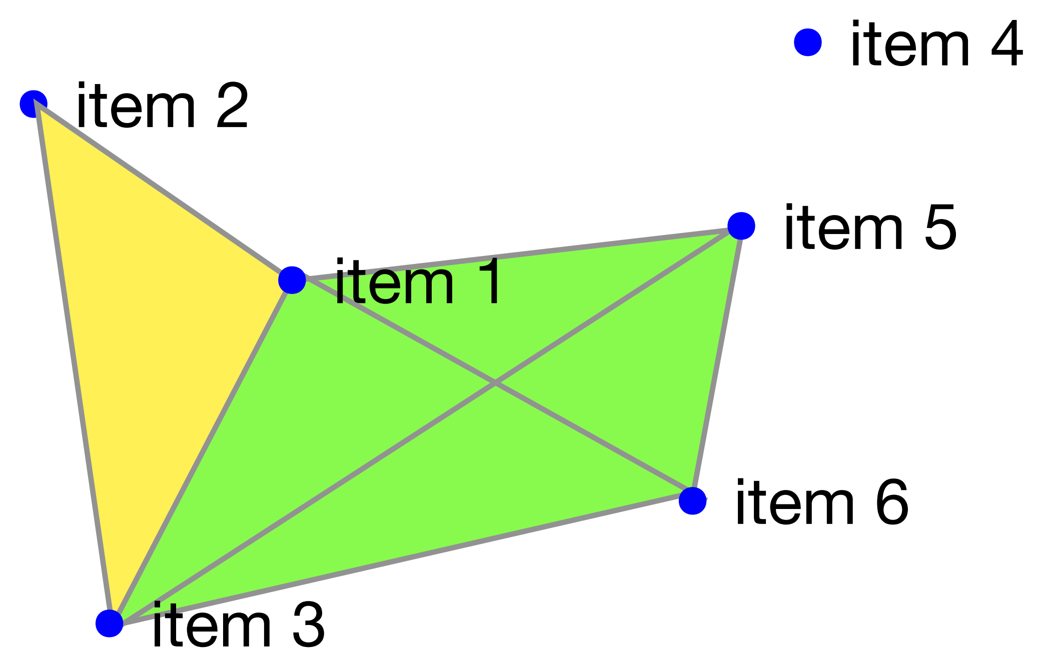


Hypergraph



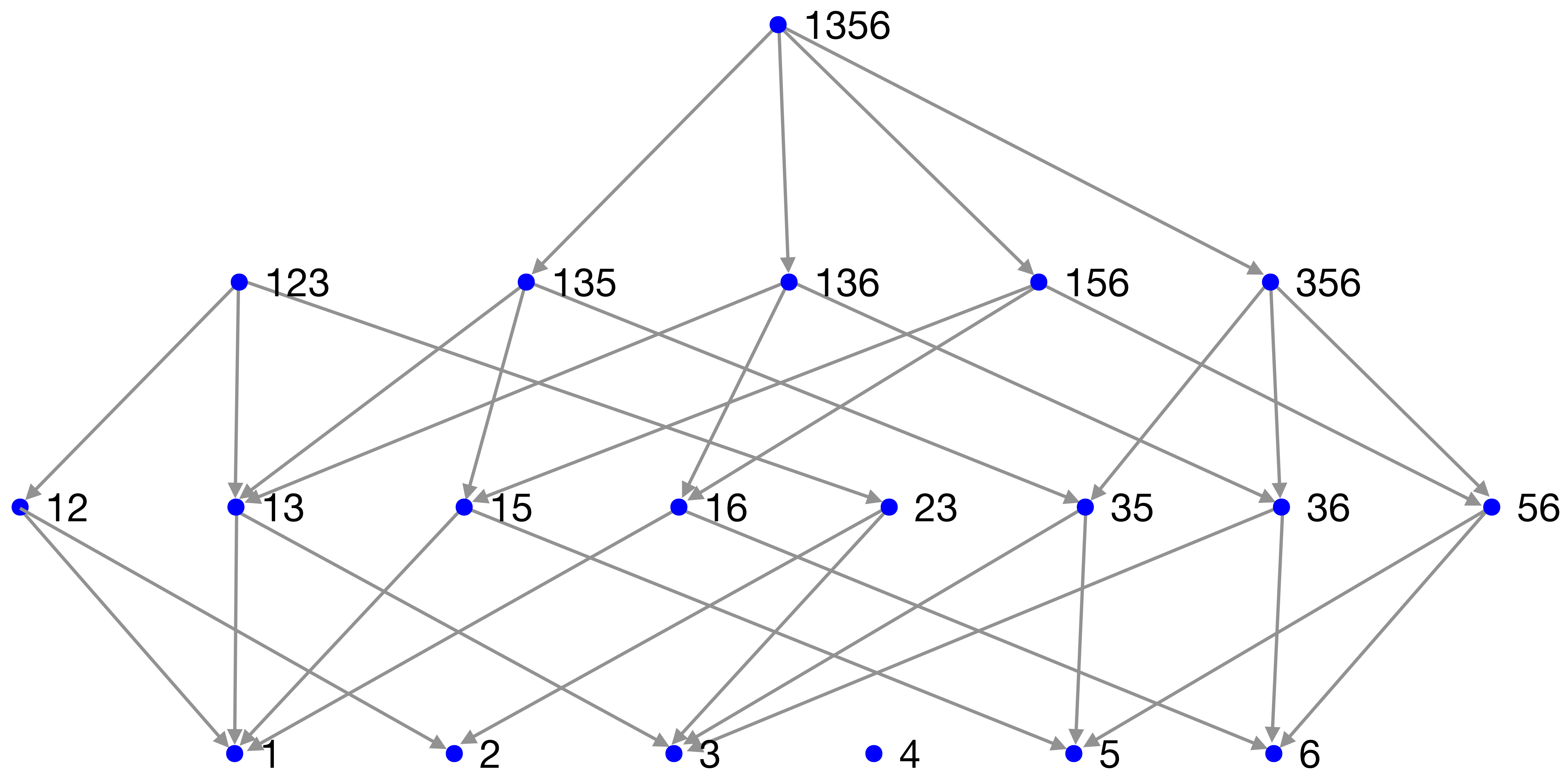
Hasse diagram

Hasse diagram
Simplicial complex



Simplicial complex

Closed under inclusion
Edges of different dimensions



Hasse diagram

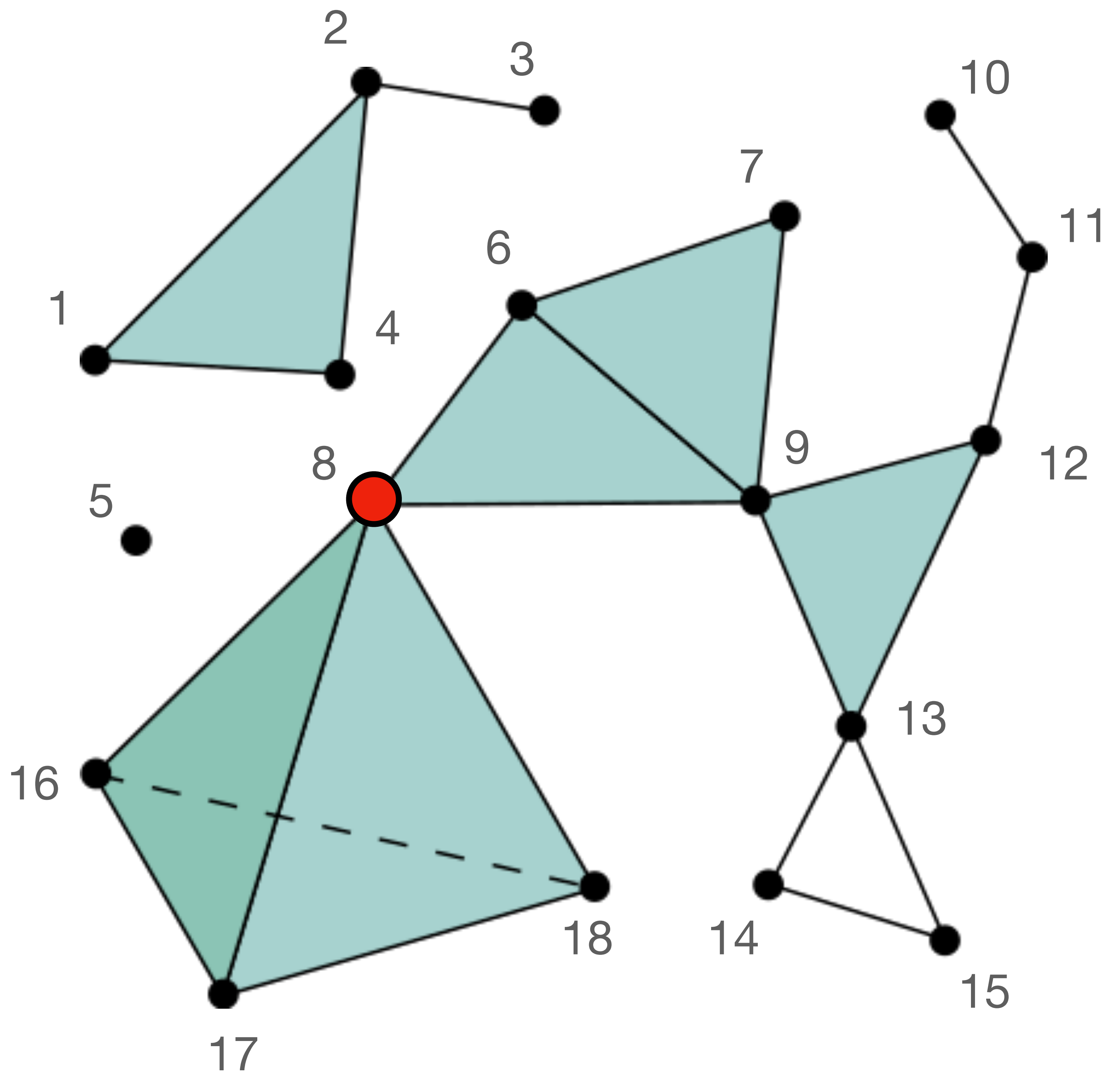
Incidence, adjacency, degree

Incidence

For a pair of simplices $\tau \subseteq \sigma$

- τ is a *face* of σ,
- σ is a *coface* of τ.

A p-simplex $\sigma_1^{(p)}$ is *q-incident* to a q-simplex $\sigma_2^{(q)}$, denoted $\sigma_1^{(p)} \rightarrow_q \sigma_2^{(q)}$ if $p \neq q$ and $\sigma_1^{(p)}$ is a face or coface of $\sigma_2^{(q)}$.



Simplicial complex *K*

Incidence, adjacency, degree

Incidence

For a pair of simplices $\tau \subseteq \sigma$

- τ is a *face* of σ,
- σ is a *coface* of τ.

A p-simplex $\sigma_1^{(p)}$ is *q-incident* to a q-simplex $\sigma_2^{(q)}$, denoted $\sigma_1^{(p)} \rightarrow_q \sigma_2^{(q)}$ if $p \neq q$ and $\sigma_1^{(p)}$ is a face or coface of $\sigma_2^{(p)}$.

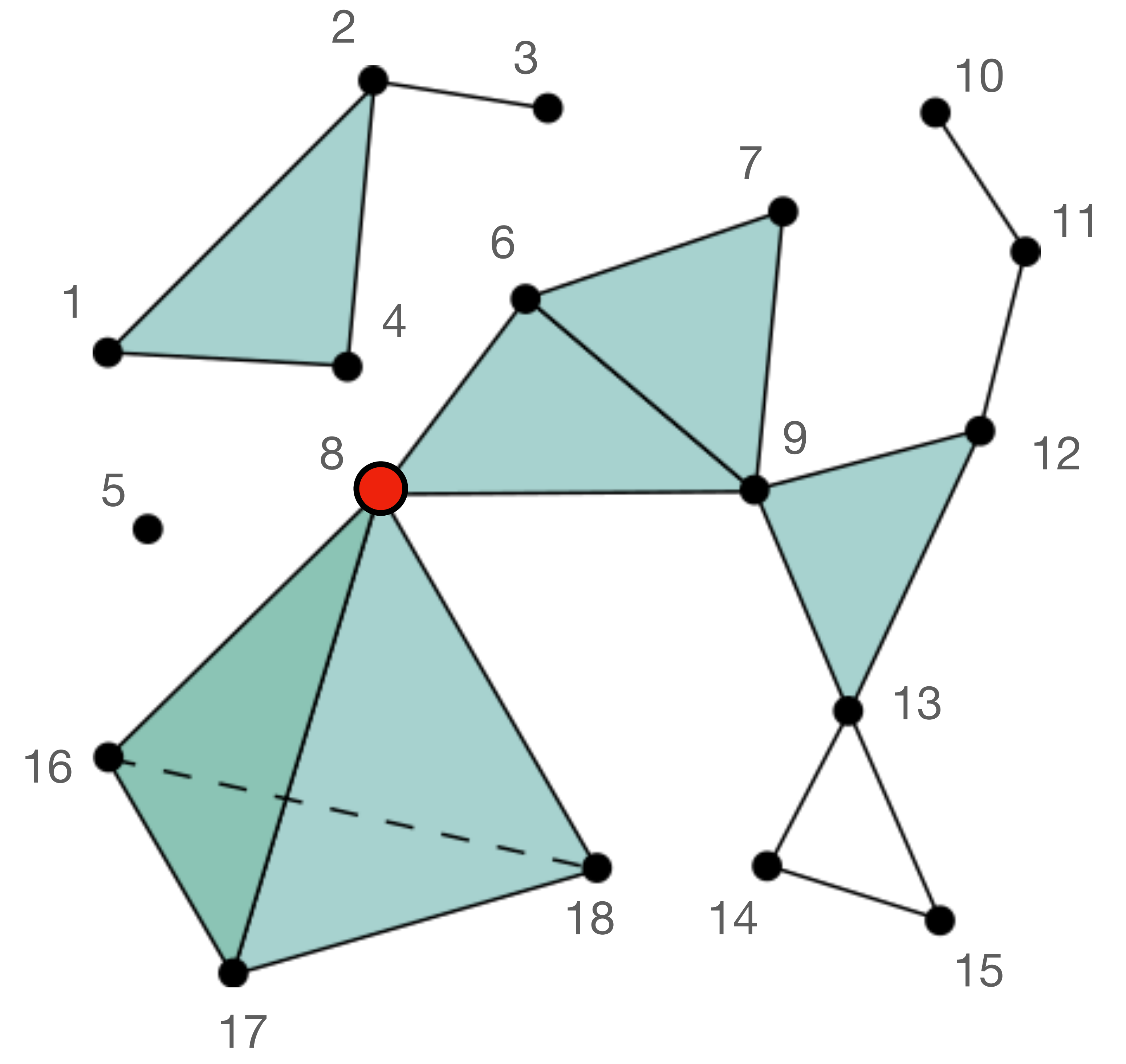
Upper incidence $p < q$

$\{8\} \rightarrow_1 \{6{,}8\}$ $\quad$ $\{8\} \rightarrow_2 \{8{,}17{,}18\}$ $\quad$ $\{8\} \rightarrow_3 \{8{,}16{,}17{,}18\}$

$\{8{,}16\} \rightarrow_3 \{8{,}16{,}17{,}18\}$

Lower incidence $p < q$

$\{6{,}7{,}9\} \rightarrow_1 \{6{,}9\}$ $\quad$ $\{8{,}16{,}17{,}18\} \rightarrow_1 \{16{,}18\}$



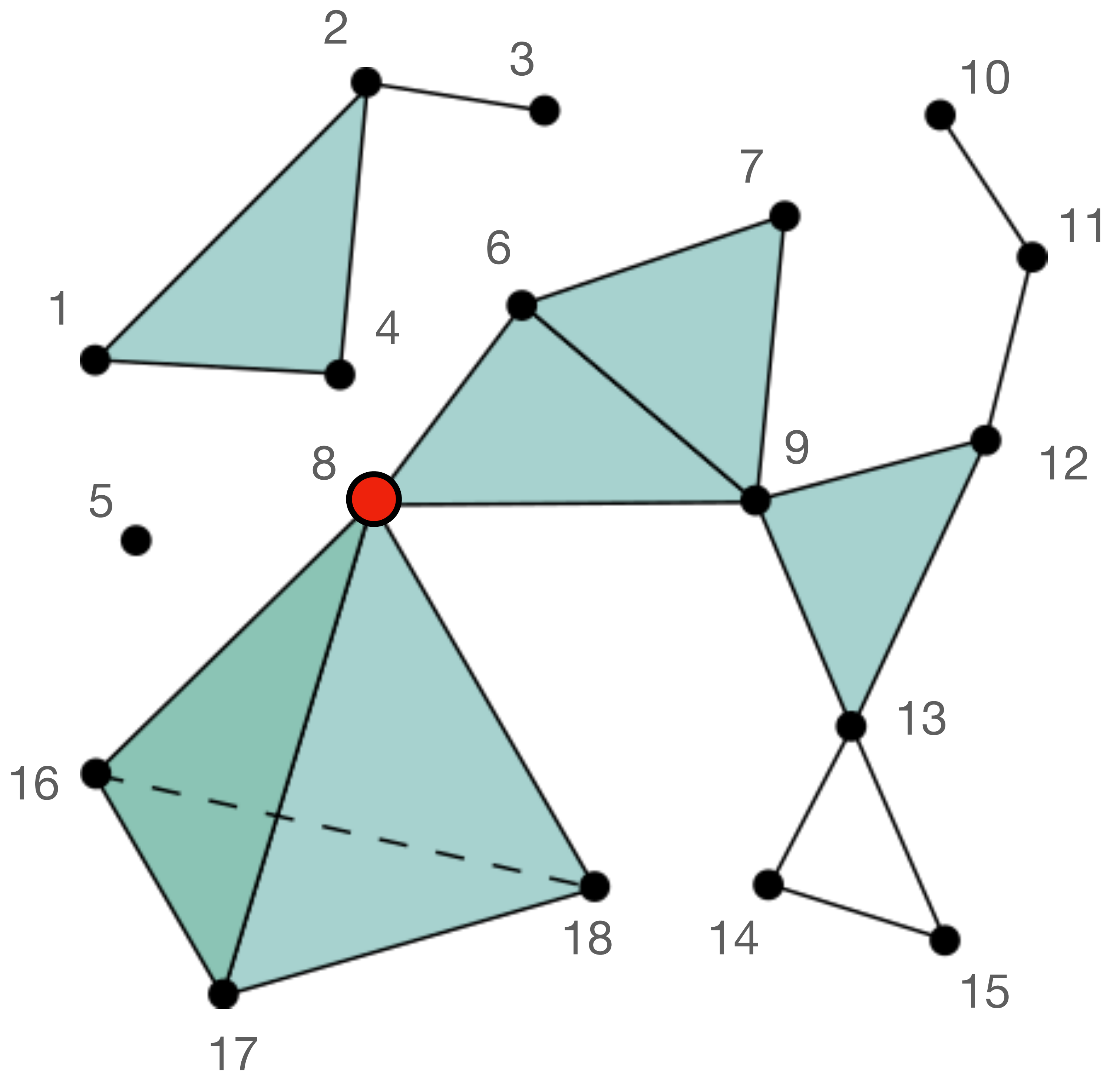
Simplicial complex *K*

Incidence, adjacency, degree

Incidence

(p, q)-incidence of a p-simplex is the # q-simplices q-incident to it.

$i_1(\{8\}) = 5$ $\quad$ $i_2(\{8\}) = 4$ $\quad$ $i_3(\{8\}) = 1$



Simplicial complex *K*

Incidence, adjacency, degree

Adjacency

A set of p-simplices $\{\sigma_1^{(p)}, \ldots, \sigma_n^{(p)}\}$ is *q-adjacent* via p-simplex τ^q denoted $\sigma_1^{(p)} \sim_q \sigma_2^{(p)}$ pairwise, or $\sim_q \{\sigma_1^{(p)}, \ldots, \sigma_n^{(p)}\}$ for all p-simplices if

- all p-simplices $\sigma_i^{(p)}$ are p-faces of q-simplex $\tau^{(q)}$, conversely
- q-simplex $\tau^{(q)}$ is a q-coface for all k-simplices $\sigma_i^{(p)}$.

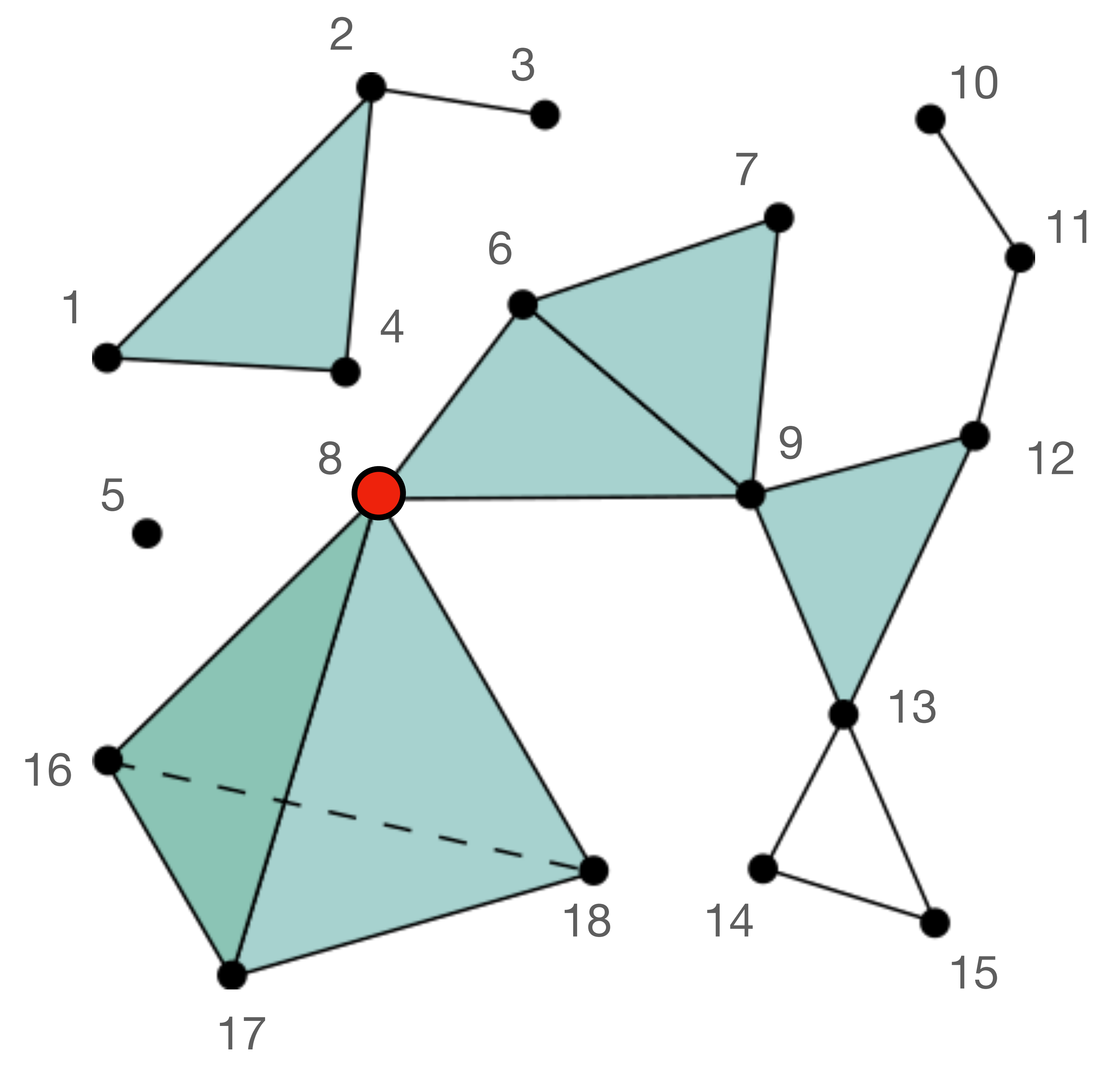
Upper adjacency $p < q$

$$\{6{,}9\} \sim_2 \{6{,}7\} \qquad \{8\} \sim_3 \{18\}$$

Lower adjacency $p > q$

$$\{8{,}16{,}17\} \sim_0 \{8{,}16{,}17{,}18\} \qquad \{6{,}7{,}9\} \sim_1 \{6{,}8{,}9\}$$

(p, q)-degree of a p-simplex is the # of q-adjacent to it p-simplices.



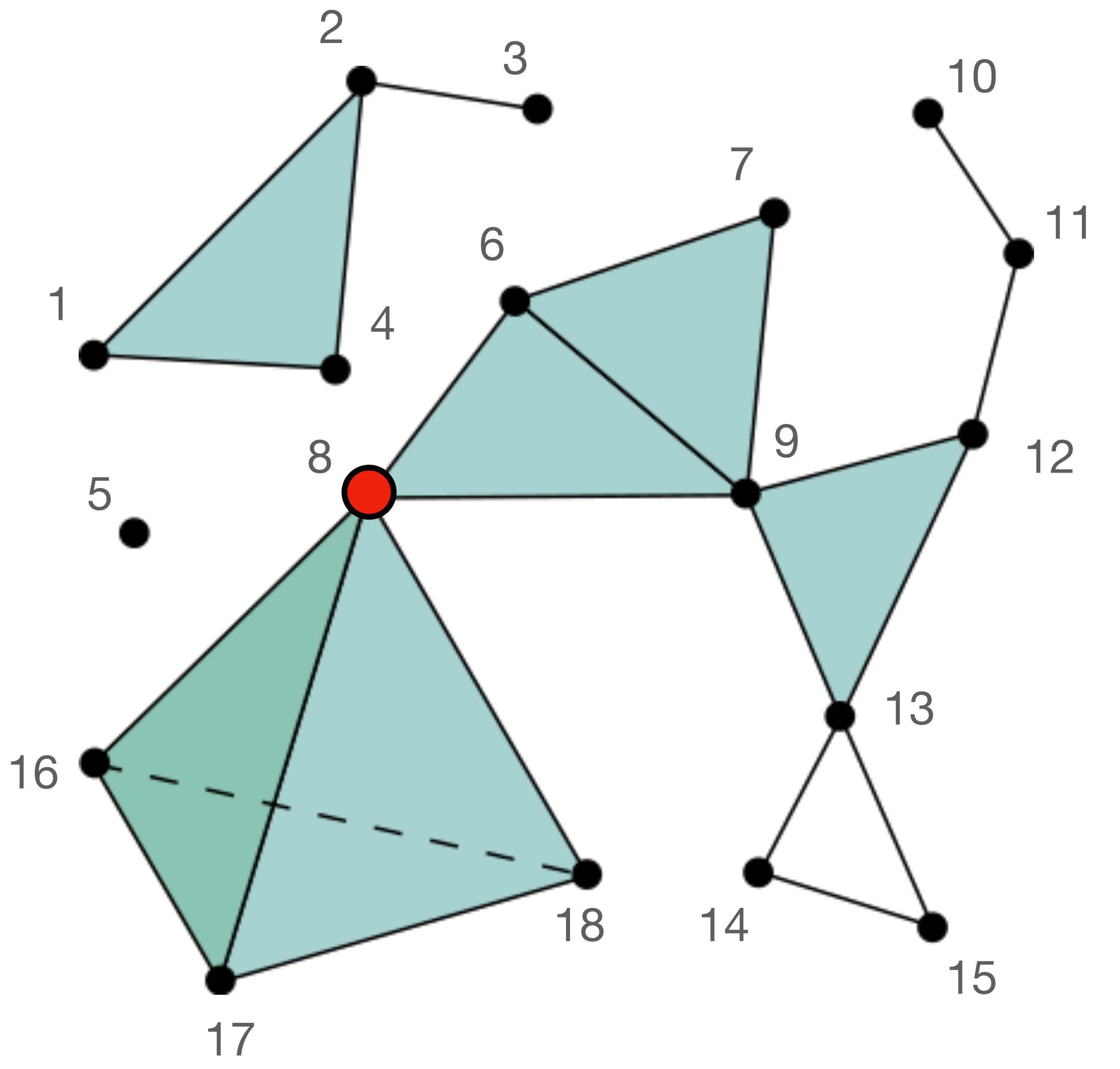
Simplicial complex *K*

Incidence, adjacency, degree

Degree

(p, q)-degree of a p-simplex is the # of q-adjacent to it p-simplices.

$$d_1(\{8\}) = 5 \qquad d_2(\{8\}) = 5 \qquad d_3(\{8\}) = 3$$



Simplicial complex *K*

Incidence, adjacency, degree

Degree

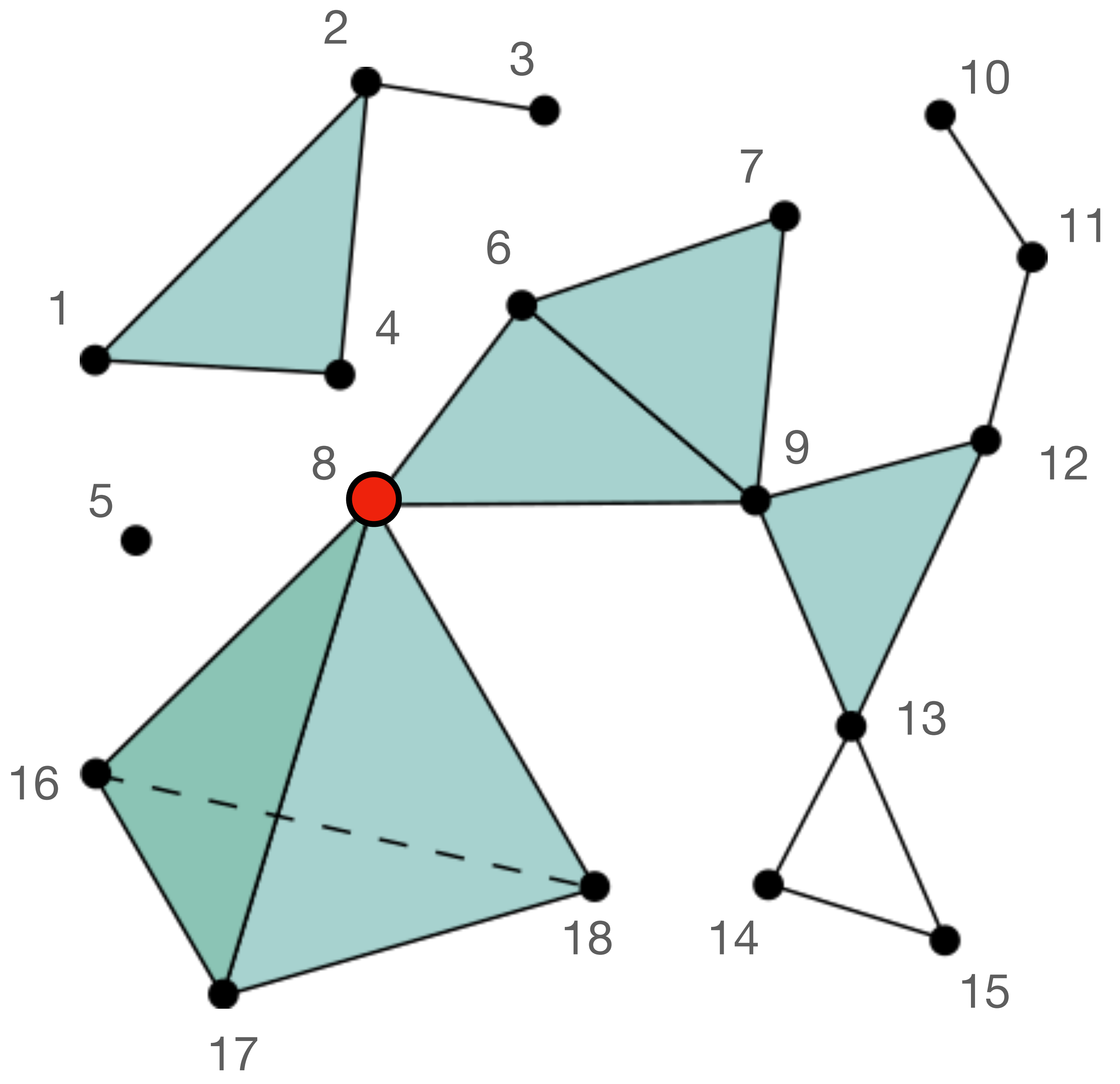
(p, q)-degree of a p-simplex is the # of q-adjacent to it p-simplices.

$d_1(\{8\}) = 5$ $d_2(\{8\}) = 5$ $d_3(\{8\}) = 3$

same for q = 1

different for q > 1

$i_1(\{8\}) = 5$ $i_2(\{8\}) = 4$ $i_3(\{8\}) = 1$



Simplicial complex *K*

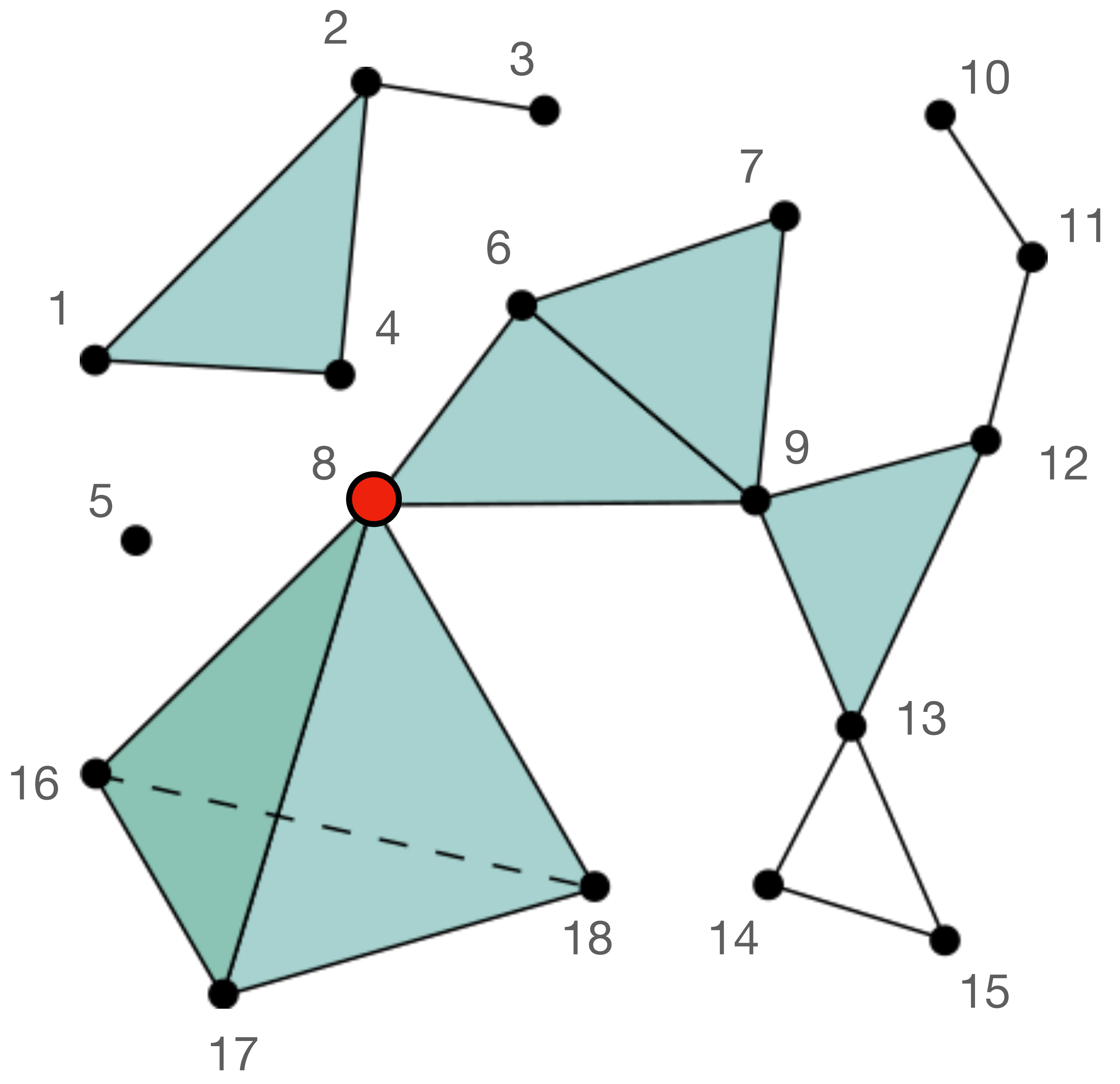
Connected components

A *(p, q)-path* is a sequence of simplices $\{\sigma_1^{(p)}, \sigma_2^{(q)}, \sigma_3^{(p)}, \sigma_4^{(q)}, \ldots, \sigma_{n-1}^{(q)}, \sigma_n^{(p)}\}$ such that $\sigma_i^{(p)} \to_q \sigma_{i+1}^{(p)} \to_q \sigma_{i+2}^{(p)} \quad \forall i$.

Two simplices $\sigma_1^{(p)}, \sigma_2^{(p)}$ are *(p, q)-connected* if there exists (p, q)-path between them.

A *(p, q)-connected component* is an equivalence relation on *K*.

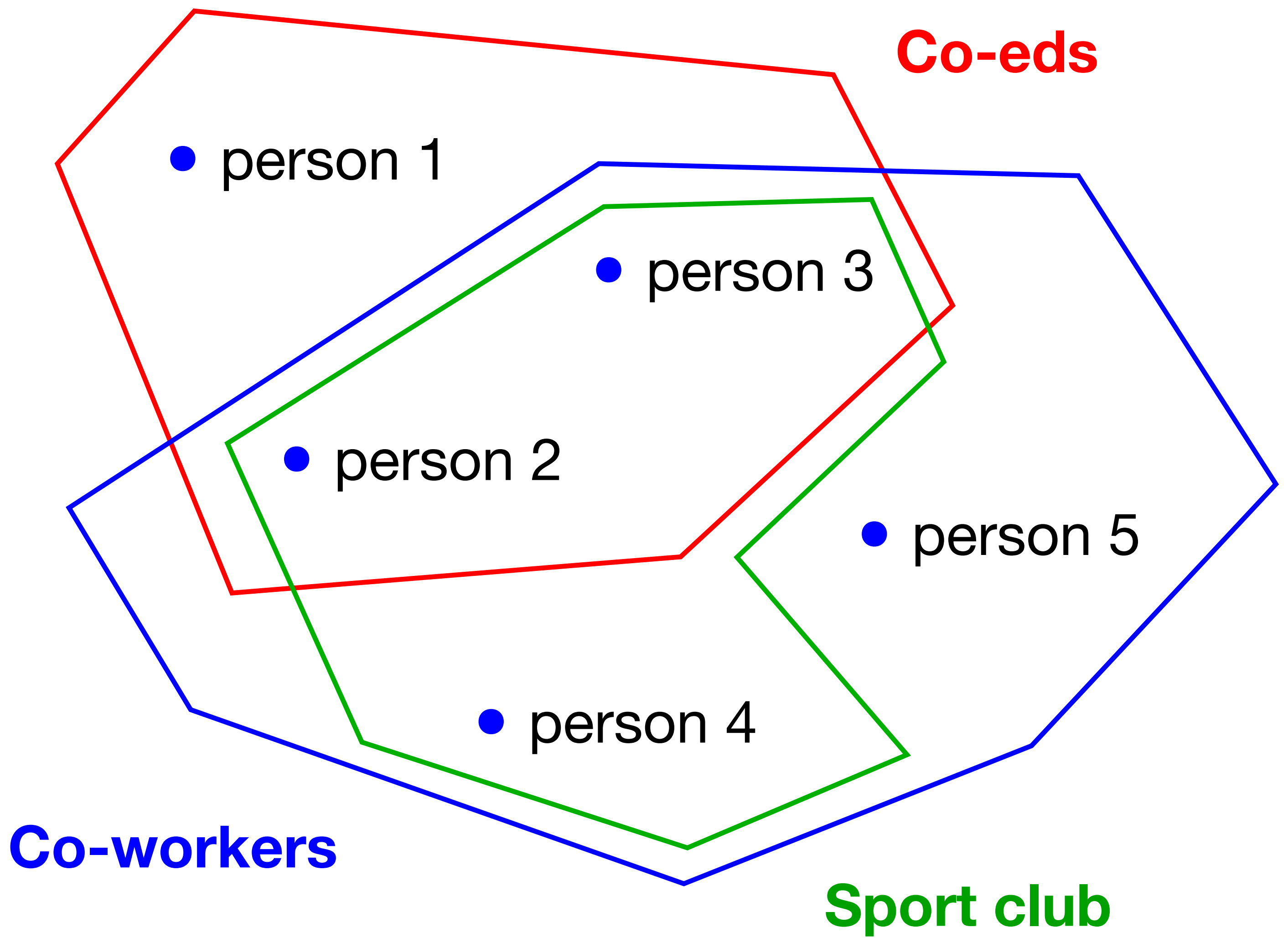
$$|C_{0,1}(K)| = 3 \qquad |C_{0,2}(K)| = 8$$



Simplicial complex *K*

Q-analysis

Social networks



From Cohomology in Physics to Q-connectivity in Social Science
Atkin R. *International Journal of Man-Machine Studies (1972)*

Chess networks

Multi-dimensional Structure in the Game of Chess
Atkin R. *International Journal of Man-Machine Studies (1972)*

Dowker complex

Binary relation

Let *X* and *Y* be sets or cardinalities *m* and *n*, respectively

$$X = \{x_1, \dots, x_m\} \qquad Y = \{y_1, \dots, y_n\}$$

Binary relation *R* is a subset of Cartesian product *X x Y*

$$R \subseteq X \times Y$$

$$xRy \iff (x, y) \in R$$

Matrix representation

R on sets *X* and *Y* is represented by m x n matrix ***A***

$$a_{ij} = \begin{cases} 1, & x_i R y_j, \\ 0, & \text{otherwise}. \end{cases}$$

	y_1	y_2	y_3
x_1	1		
x_2	1	1	
x_3	1		1
x_4		1	1

Dowker complex

Let R is a binary relation on sets *X* and *Y* of cardinalities *m* and *n*, respectively.

A *Dowker complex K* of a binary relation R on sets X and Y is defined

$$K = \{\sigma^{(m)} = \{x_{\sigma_0}, \ldots, x_{\sigma_m}\} \mid \exists y \text{ s.t. } x_i R y \quad \forall x_i \in \sigma^{(m)}\}$$

Analogously, a *Dowker complex L* is defined

$$L = \{\sigma^{(n)} = \{y_{\sigma_0}, \ldots, y_{\sigma_n}\} \mid \exists x \text{ s.t. } x_i R y \quad \forall y_j \in \sigma^{(n)}\}$$

Homology Groups of Relations
Dowker C. *The Annals of Mathematics (1952)*

Dowker complex

Let R is a binary relation on sets *X* and *Y* of cardinalities *m* and *n*, respectively.

A *Dowker complex K* of a binary relation R on sets X and Y is defined

$$K = \{\sigma^{(m)} = \{x_{\sigma_0}, \ldots, x_{\sigma_m}\} \mid \exists y \text{ s.t. } x_i R y \quad \forall x_i \in \sigma^{(m)}\}$$

Analogously, a *Dowker complex L* is defined

$$L = \{\sigma^{(n)} = \{y_{\sigma_0}, \ldots, y_{\sigma_n}\} \mid \exists x \text{ s.t. } x_i R y \quad \forall y_j \in \sigma^{(n)}\}$$

	y_1	y_2	y_3
x_1	1		
x_2	1	1	
x_3	1		1
x_4		1	1

$$\sigma^{(2)} = \{x_1, x_2, x_3\} \in K$$

Homology Groups of Relations
Dowker C. *The Annals of Mathematics (1952)*

Dowker complex

Let R is a binary relation on sets *X* and *Y* of cardinalities *m* and *n*, respectively.

A *Dowker complex K* of a binary relation R on sets X and Y is defined

$$K = \{\sigma^{(m)} = \{x_{\sigma_0}, \ldots, x_{\sigma_m}\} \mid \exists y \text{ s.t. } x_i R y \quad \forall x_i \in \sigma^{(m)}\}$$

Analogously, a *Dowker complex L* is defined

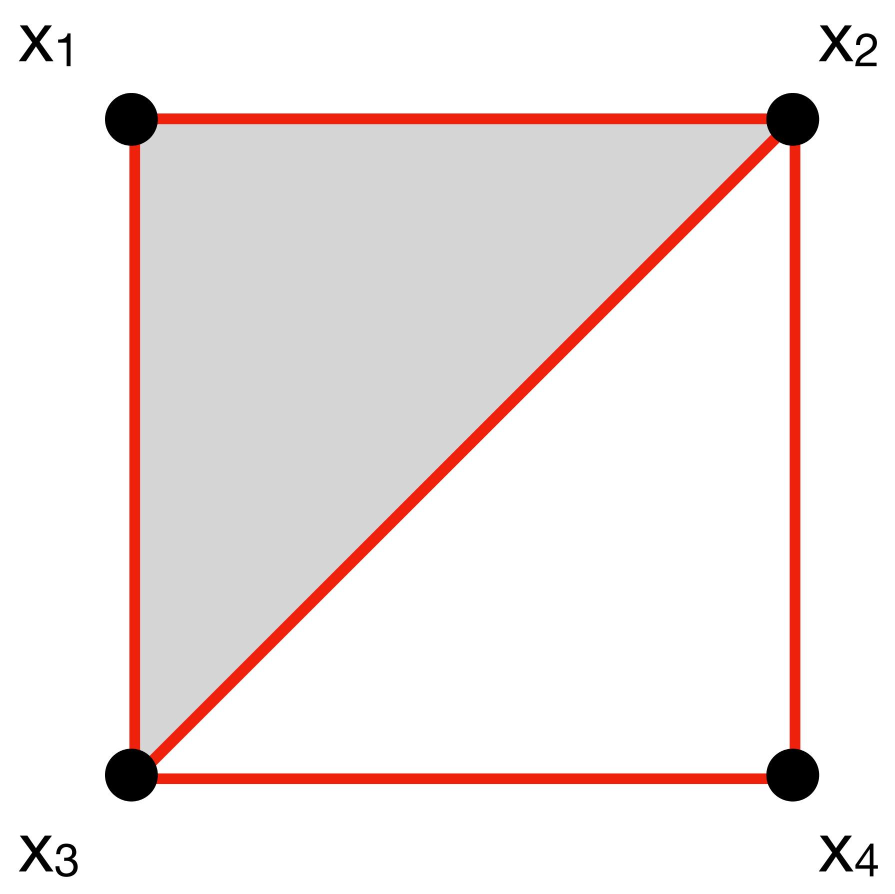
$$L = \{\sigma^{(n)} = \{y_{\sigma_0}, \ldots, y_{\sigma_n}\} \mid \exists x \text{ s.t. } x_i R y \quad \forall y_j \in \sigma^{(n)}\}$$

	y_1	y_2	y_3
x_1	1		
x_2	1	1	
x_3	1		1
x_4		1	1

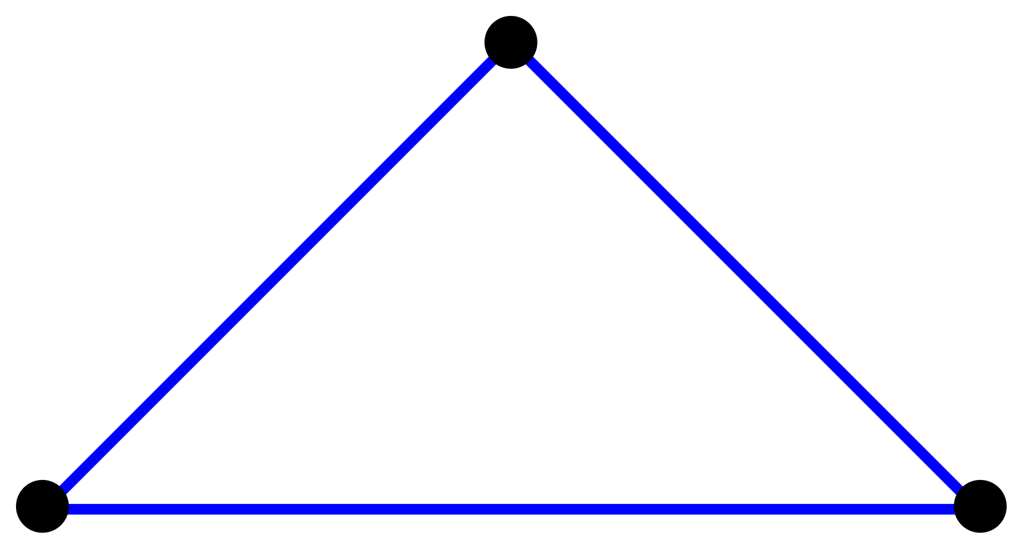
$$\sigma^{(1)} = \{y_1, y_3\} \in L$$

Homology Groups of Relations
Dowker C. *The Annals of Mathematics (1952)*

Dowker complex



K



L

	y_1	y_2	y_3
x_1	1		
x_2	1	1	
x_3	1		1
x_4		1	1

$\mathbf{A}$

Homology groups of Dowker complexes *K* and *L* are isomorphic, i.e. $H_\bullet(K) \simeq H_\bullet(L)$ [Dowker1952, Thm. 1].

Homology Groups of Relations
Dowker C. *The Annals of Mathematics (1952)*

Dowker complex

Chess network

Chess board

	f_1	…	f_{16}
s_1	1		
s_2	1	1	
…	1		1
s_{64}		1	1

Relation "figure attacks square"

K

vertices — figures
simplices — squares attacked by a given figure

L

vertices — squares
simplices — figures attacking a given square

Dowker complexes *K* and *L*

Multi-dimensional Structure in the Game of Chess
Atkin R. *International Journal of Man-Machine Studies (1972)*

Q-analysis

Q-connected components

Given a Dowker complex *K (L)* Atkin's Q-analysis is to associate
a Q-vector defined as a vector of #(max, q)-connected components of *K (L)*
for all *q.*

Q-vector of *K*

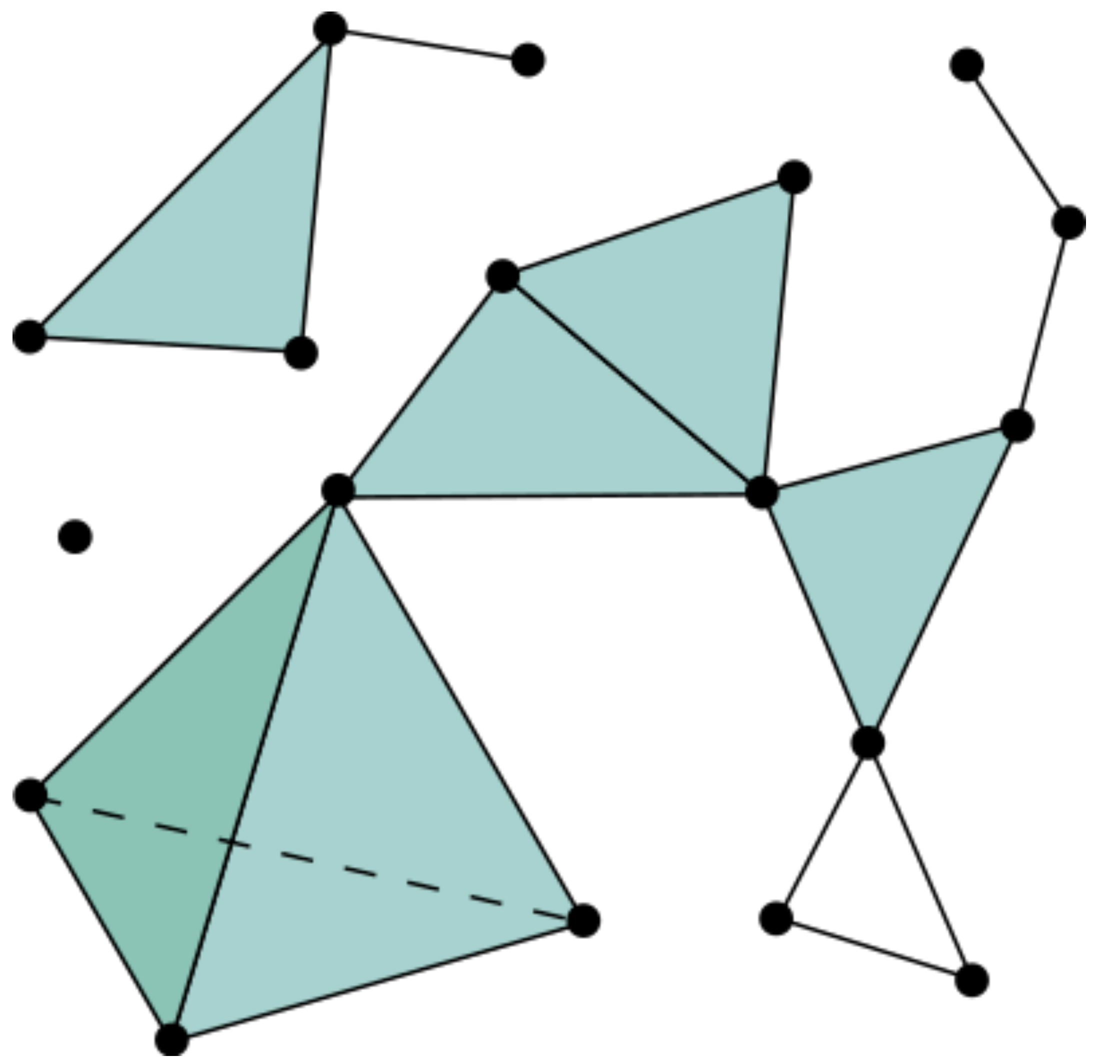
$$(3,\ 4+m)$$

0 1

(0, q)-vector of *K*

$$(3,\ 6,\ 15)$$

1 2 3



K

F-vector $(18,\ 19,\ 8,\ 1)$

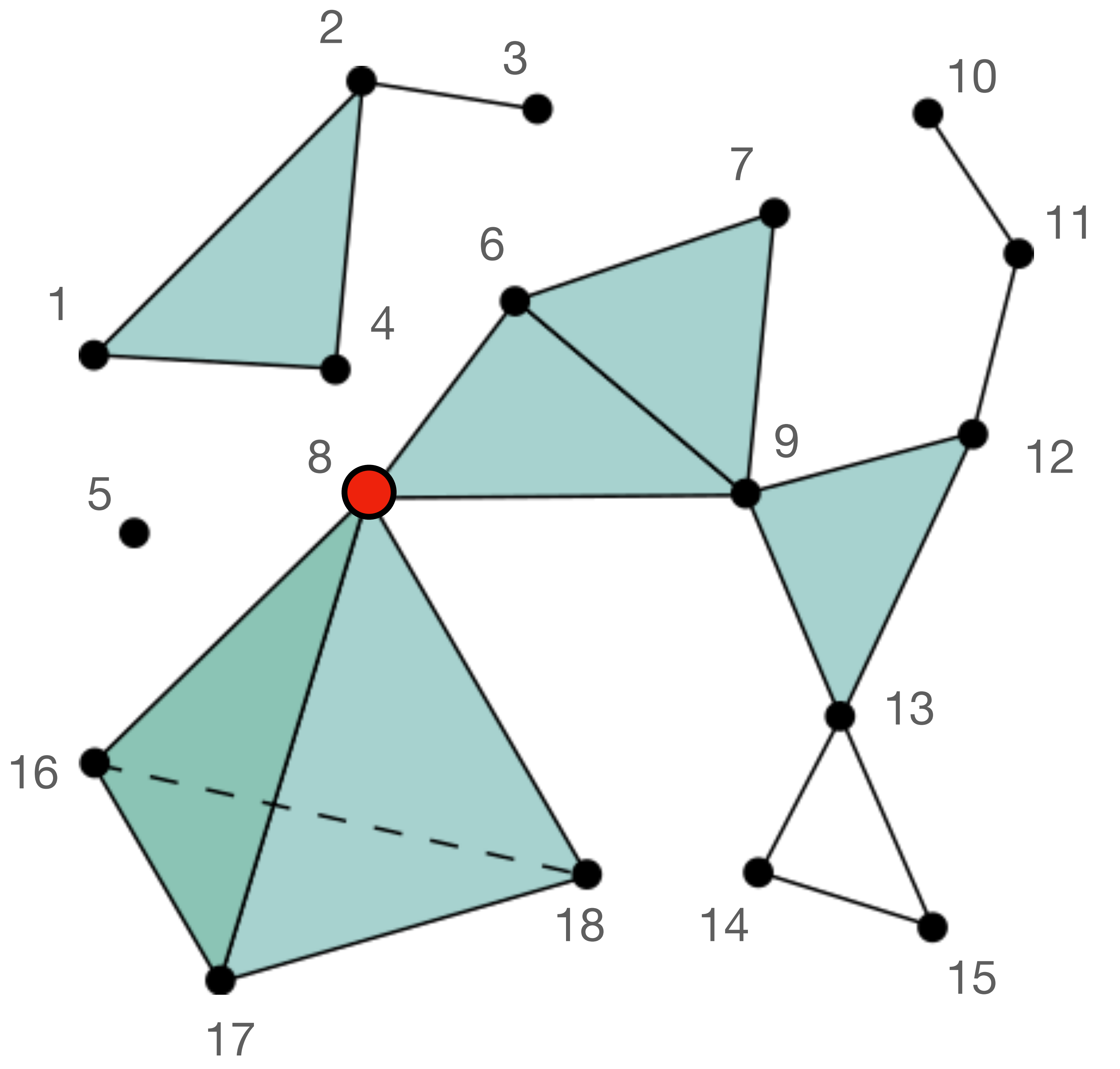
From Cohomology in Physics to Q-connectivity in Social Science
Atkin R. *International Journal of Man-Machine Studies (1972)*

Multi-dimensional Structure in the Game of Chess
Atkin R. *International Journal of Man-Machine Studies (1972)*

Line graph

A *(p, q)-line graph* of a simplicial complex *K* is a graph *G(V, E)* where

- *V* consists of p-simplices of *K,*
- $(\sigma_1^{(p)}, \sigma_2^{(p)}) \in E_G$ if $\sigma_1^{(p)} \sim_q \sigma_2^{(p)}$.



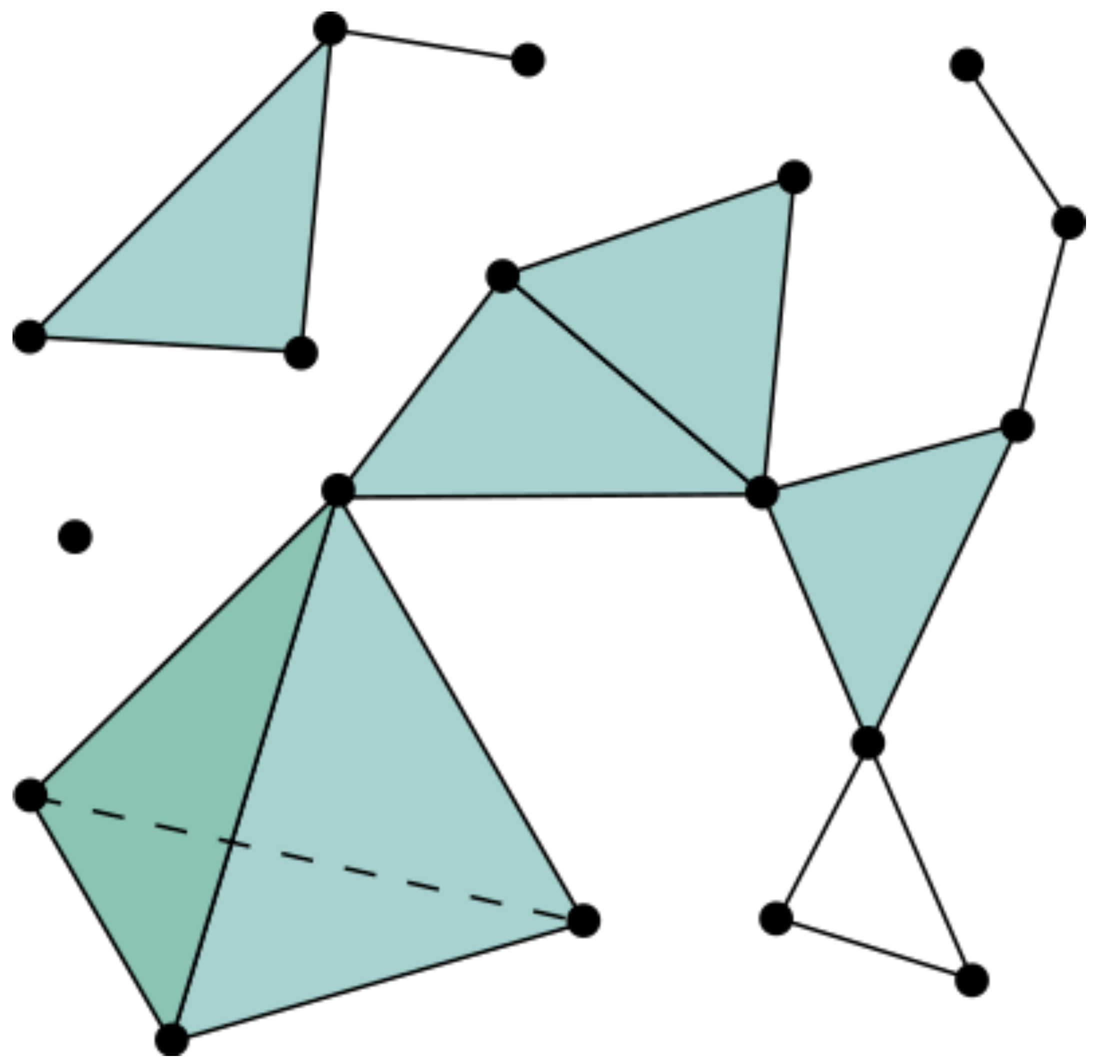
Simplicial complex *K*

Centralities

Betweenness, closeness centrality

***(p, q)*-betweenness centrality**
shortest (p, q)-paths passing through given p-simplex

***(p, q)*-closeness centrality**
Inverse of sum of the length of (p, q)-paths between given p-simplex and other p-simplices



K

Can be computed using (p, q)-line graph