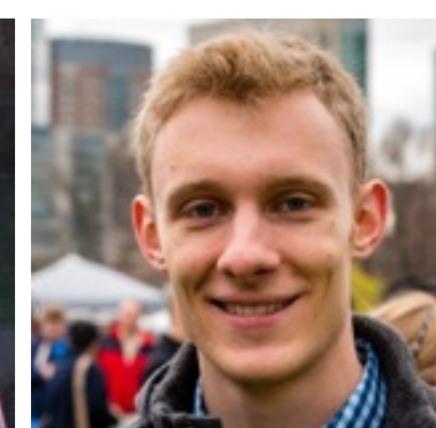
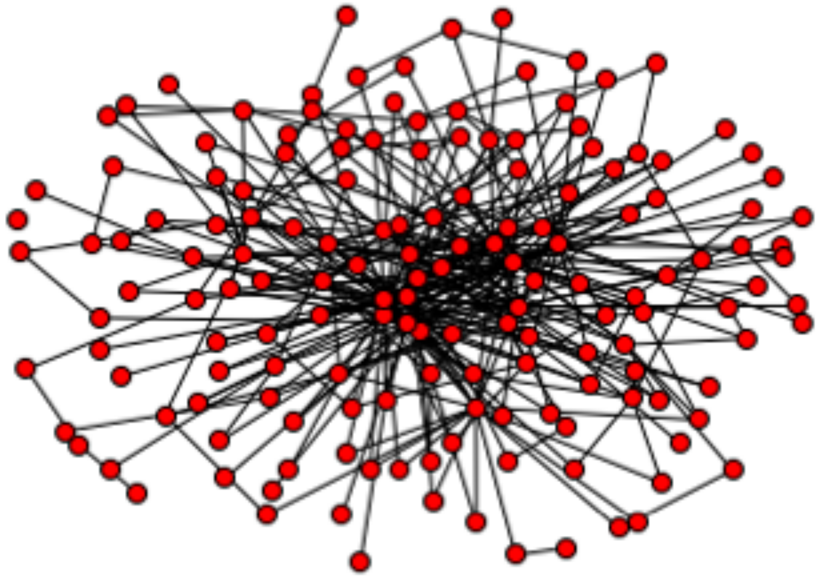
# Edge-exchangeable graphs and sparsity

Tamara Broderick
ITT Career Development
Assistant Professor,
MIT

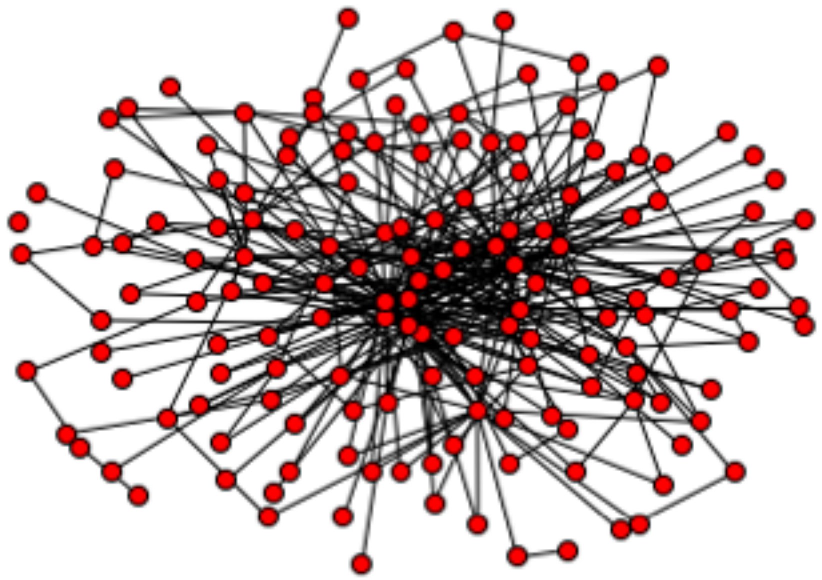
Joint work with: Diana Cai, Trevor Campbell







[Holland et al 1983; Kemp et al 2006; Xu et al 2007; Airoldi et al 2008; Lloyd et al 2012]

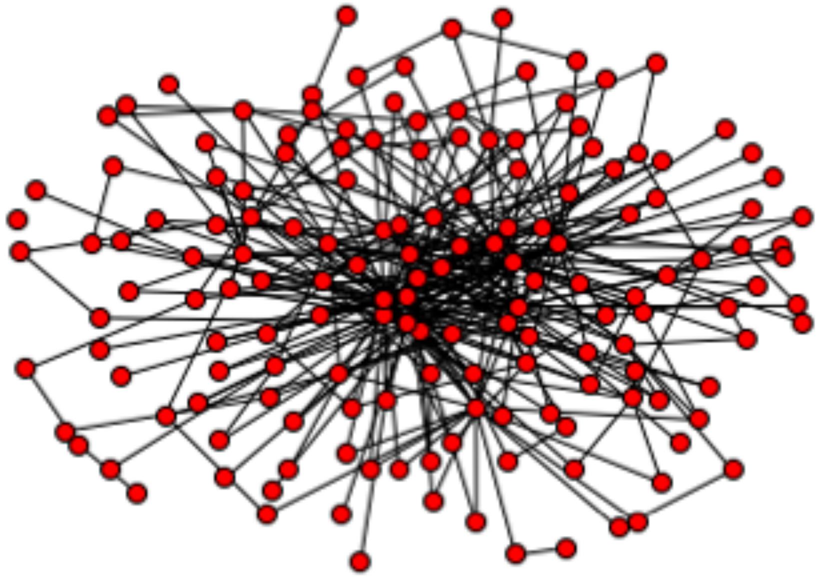


social: *Facebook, Twitter, email*
biological: *ecological, protein, gene*
transportation: *roads, railways*



[Holland et al 1983; Kemp et al 2006; Xu et al 2007; Airoldi et al 2008; Lloyd et al 2012]

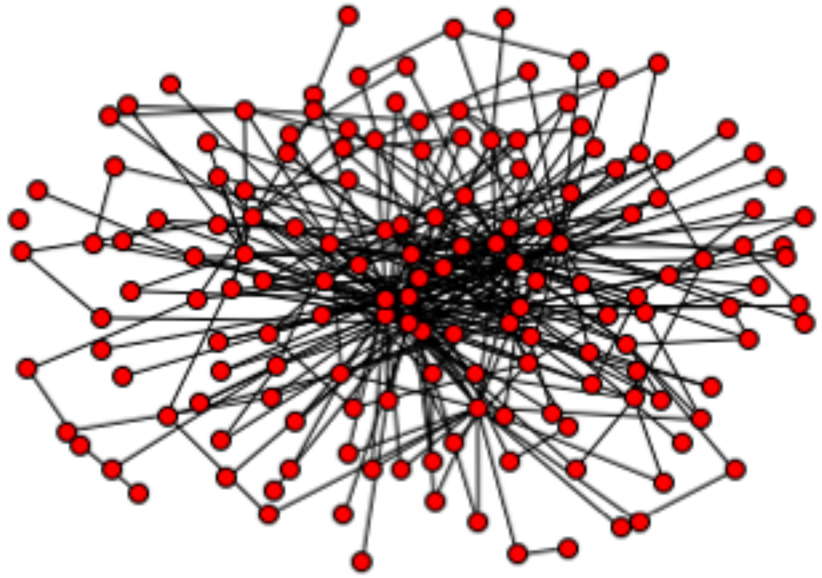
# Probabilistic models for graphs

$p($  $)$

social: *Facebook, Twitter, email*
biological: *ecological, protein, gene*
transportation: *roads, railways*

[Holland et al 1983; Kemp et al 2006; Xu et al 2007; Airoldi et al 2008; Lloyd et al 2012]

# Probabilistic models for graphs

$p($  $)$

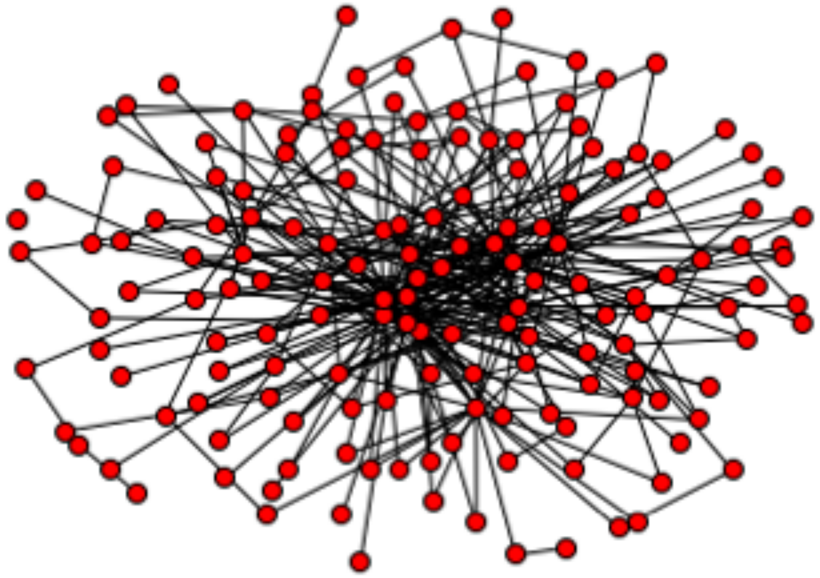
social: *Facebook, Twitter, email*
biological: *ecological, protein, gene*
transportation: *roads, railways*

- Rich relationships, coherent uncertainties, prior info



[Holland et al 1983; Kemp et al 2006; Xu et al 2007; Airoldi et al 2008; Lloyd et al 2012]

# Probabilistic models for graphs

$p($  $)$

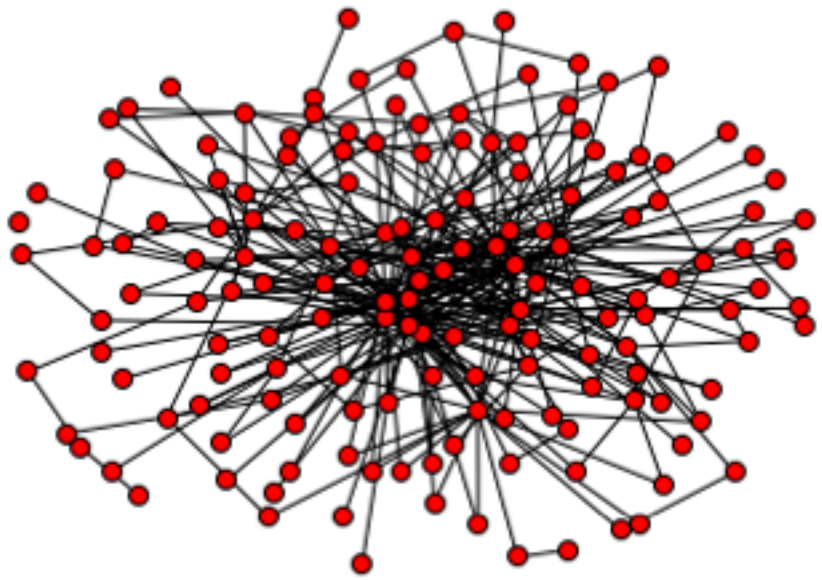
social: *Facebook, Twitter, email*
biological: *ecological, protein, gene*
transportation: *roads, railways*

- Rich relationships, coherent uncertainties, prior info
- Example models: Stochastic block, mixed membership stochastic block, infinite relational, and many more



[Holland et al 1983; Kemp et al 2006; Xu et al 2007; Airoldi et al 2008]

# Probabilistic models for graphs

$p($  $)$

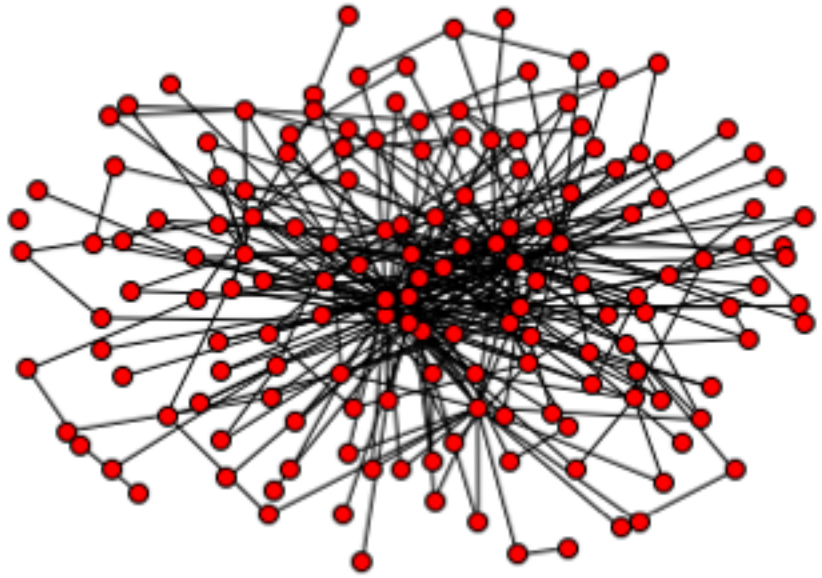
social: *Facebook, Twitter, email*
biological: *ecological, protein, gene*
transportation: *roads, railways*

- Rich relationships, coherent uncertainties, prior info
- Example models: Stochastic block, mixed membership stochastic block, infinite relational, and *many* more



[Holland et al 1983; Kemp et al 2006; Xu et al 2007; Airoldi et al 2008; Lloyd et al 2012]

# Probabilistic models for graphs

$p($  $)$

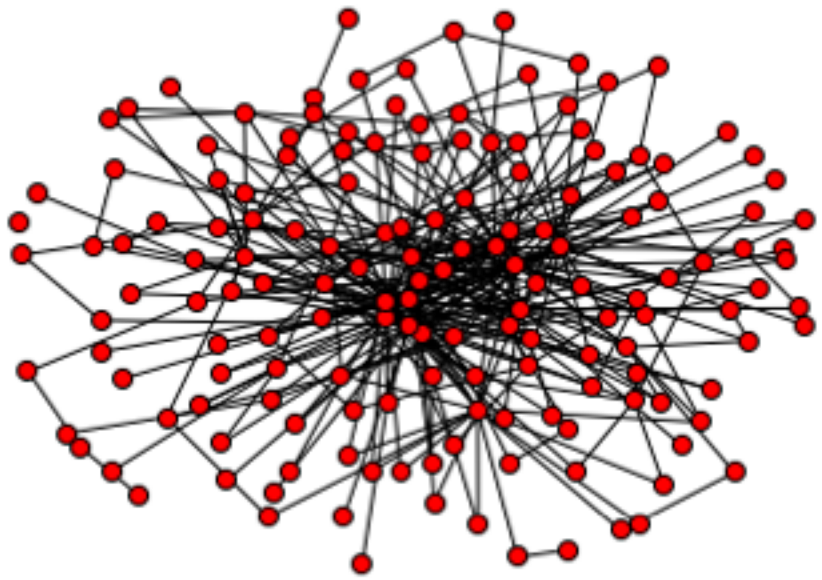
social: *Facebook, Twitter, email*
biological: *ecological, protein, gene*
transportation: *roads, railways*

- Rich relationships, coherent uncertainties, prior info
- Example models: Stochastic block, mixed membership stochastic block, infinite relational, and many more
- Assume: Adding more data doesn't change distribution of earlier data (*projectivity*)



[Holland et al 1983; Kemp et al 2006; Xu et al 2007; Airoldi et al 2008; Lloyd et al 2012]

# Probabilistic models for graphs

$p($  $)$

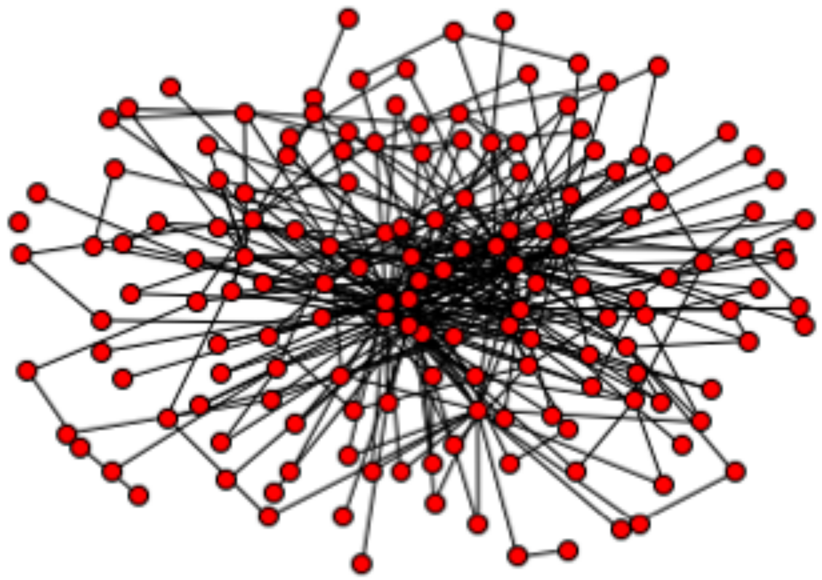
social: *Facebook, Twitter, email*
biological: *ecological, protein, gene*
transportation: *roads, railways*

- Rich relationships, coherent uncertainties, prior info
- Example models: Stochastic block, mixed membership stochastic block, infinite relational, and many more
- Assume: Adding more data doesn't change distribution of earlier data (*projectivity*)
- **Problem**: model misspecification, dense graphs



[Holland et al 1983; Kemp et al 2006; Xu et al 2007; Airoldi et al 2008; Lloyd et al 2012]

# Probabilistic models for graphs

$p($  $)$

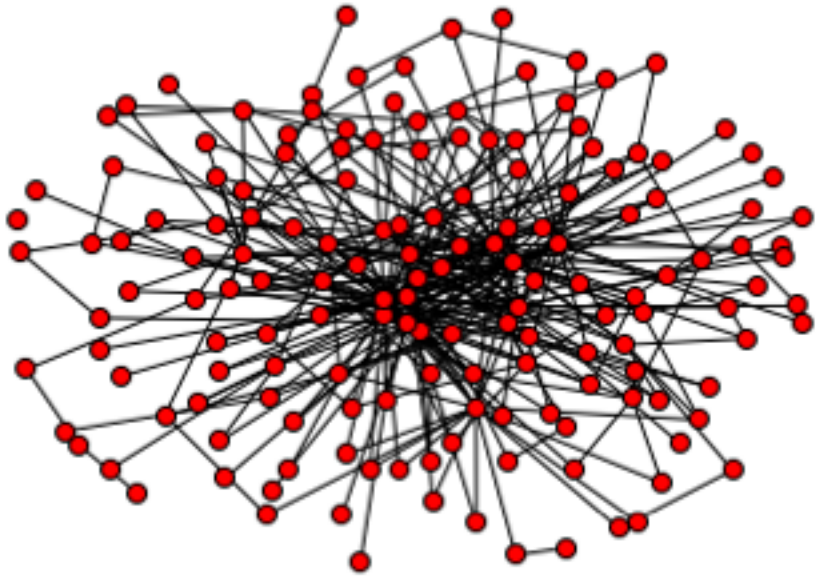
social: *Facebook, Twitter, email*
biological: *ecological, protein, gene*
transportation: *roads, railways*

- Rich relationships, coherent uncertainties, prior info
- Example models: Stochastic block, mixed membership stochastic block, infinite relational, and many more
- Assume: Adding more data doesn't change distribution of earlier data (*projectivity*)
- **Problem**: model misspecification, dense graphs
- **Solution**: a new framework for sparse graphs



[Holland et al 1983; Kemp et al 2006; Xu et al 2007; Airoldi et al 2008; Lloyd et al 2012]

# Probabilistic models for graphs

$p($  $)$

social: *Facebook, Twitter, email*
biological: *ecological, protein, gene*
transportation: *roads, railways*

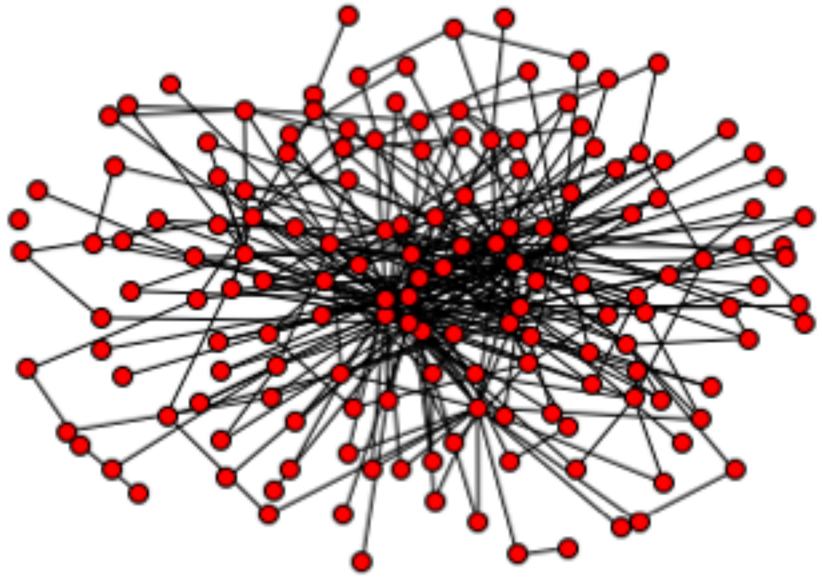
- Rich relationships, coherent uncertainties, prior info
- Example models: Stochastic block, mixed membership stochastic block, infinite relational, and many more
- Assume: Adding more data doesn't change distribution of earlier data (*projectivity*)
- **Problem**: **model misspecification, dense graphs**
- **Solution**: a new framework for sparse graphs



[Holland et al 1983; Kemp et al 2006; Xu et al 2007; Airoldi et al 2008; Lloyd et al 2012]

# Probabilistic models for graphs

$$p\Big(\quad\Big)$$



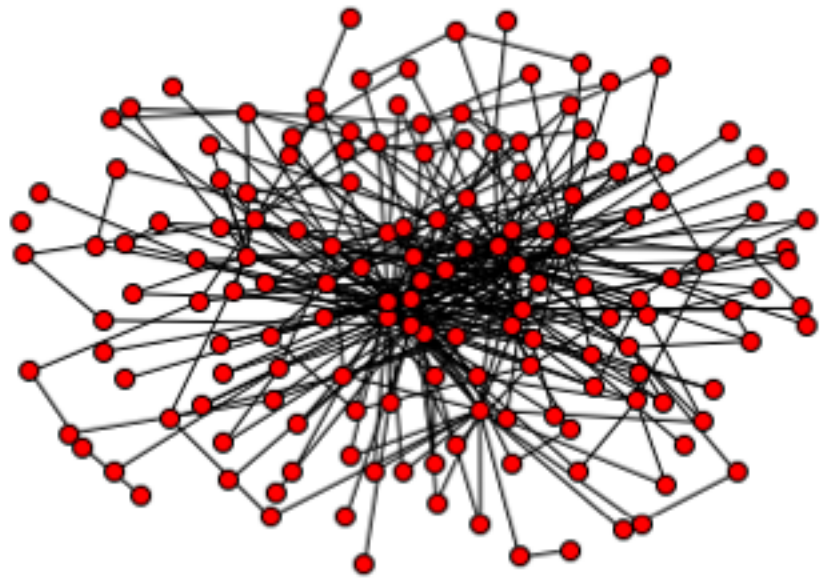
social: *Facebook, Twitter, email*
biological: *ecological, protein, gene*
transportation: *roads, railways*

- Rich relationships, coherent uncertainties, prior info
- Example models: Stochastic block, mixed membership stochastic block, infinite relational, and many more
- Assume: Adding more data doesn't change distribution of earlier data (*projectivity*)
- **Problem**: model misspecification, dense graphs
- **Solution**: a **new framework** for sparse graphs



[Holland et al 1983; Kemp et al 2006; Xu et al 2007; Airoldi et al 2008; Lloyd et al 2012]

# Probabilistic models for graphs

$p($  $)$

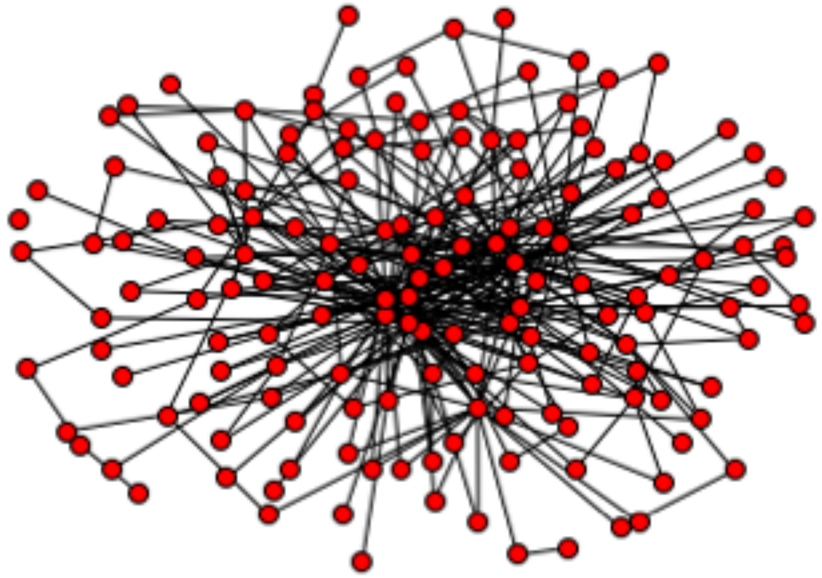
social: *Facebook, Twitter, email*
biological: *ecological, protein, gene*
transportation: *roads, railways*

- Rich relationships, coherent uncertainties, prior info
- Example models: Stochastic block, mixed membership stochastic block, infinite relational, and many more
- Assume: Adding more data doesn't change distribution of earlier data (*projectivity*)
- **Problem**: model misspecification, dense graphs
- **Solution**: a new framework for **sparse graphs**



[Holland et al 1983; Kemp et al 2006; Xu et al 2007; Airoldi et al 2008; Lloyd et al 2012]

# Probabilistic models for graphs

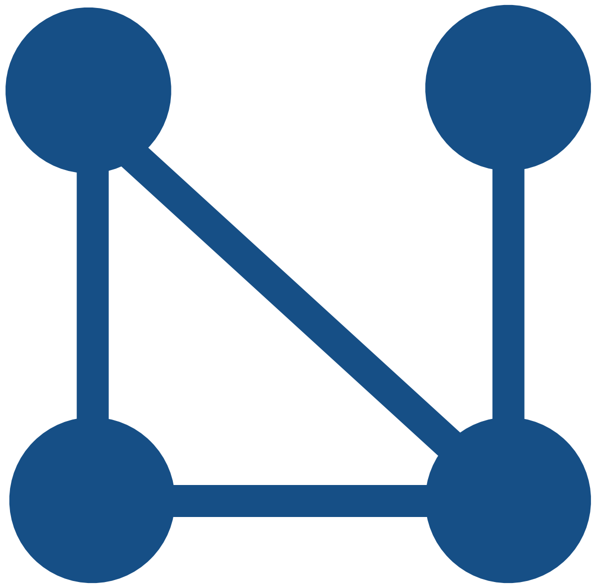
$p($  $)$

social: *Facebook, Twitter, email*
biological: *ecological, protein, gene*
transportation: *roads, railways*

- Rich relationships, coherent uncertainties, prior info
- Example models: Stochastic block, mixed membership stochastic block, infinite relational, and many more
- Assume: Adding more data doesn't change distribution of earlier data (*projectivity*)
- **Problem**: model misspecification, dense graphs
- **Solution**: a new framework for sparse graphs
  - Concurrent & independent graphs work by Crane & Dempsey

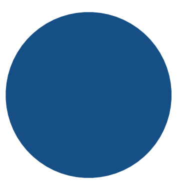
[B, Cai 2015; Cai, B 2015a,b; Crane, Dempsey 2015a,b,16a,b; Cai, Campbell, B 2016; Campbell, Cai, B 2016; Williamson 2016]

# Sequence of graphs

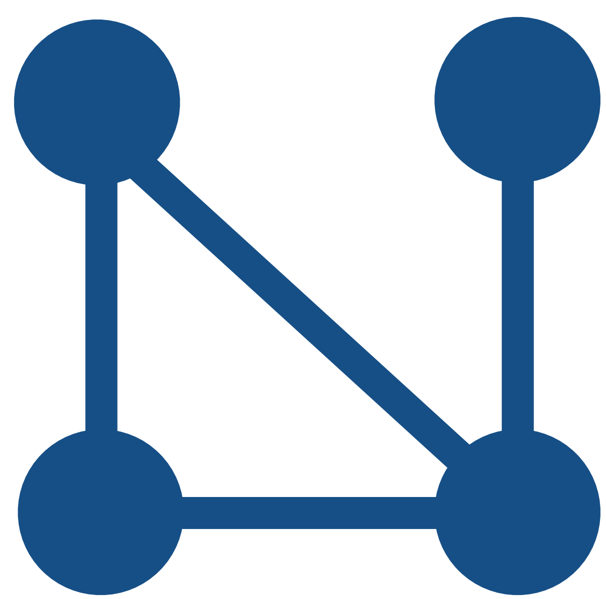


$G$

# Sequence of graphs

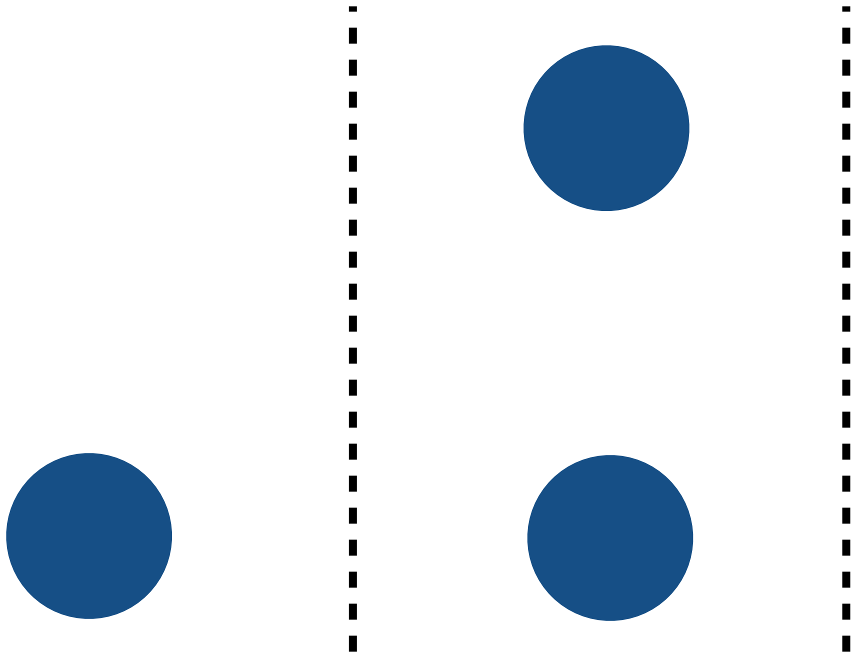


$G_1$

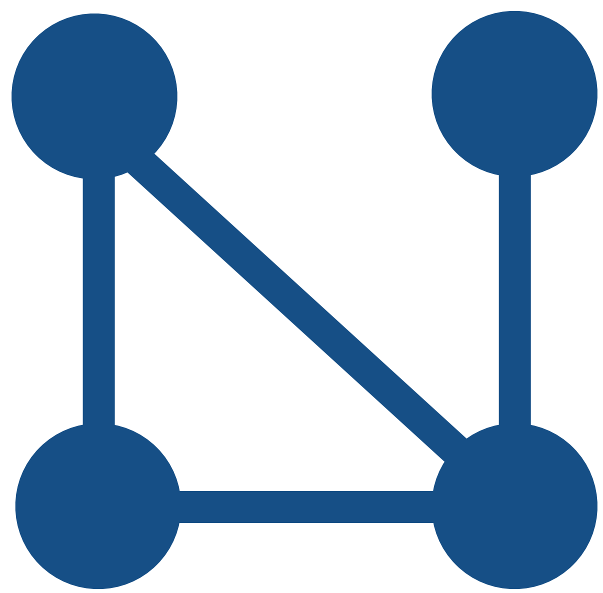


$G$

# Sequence of graphs

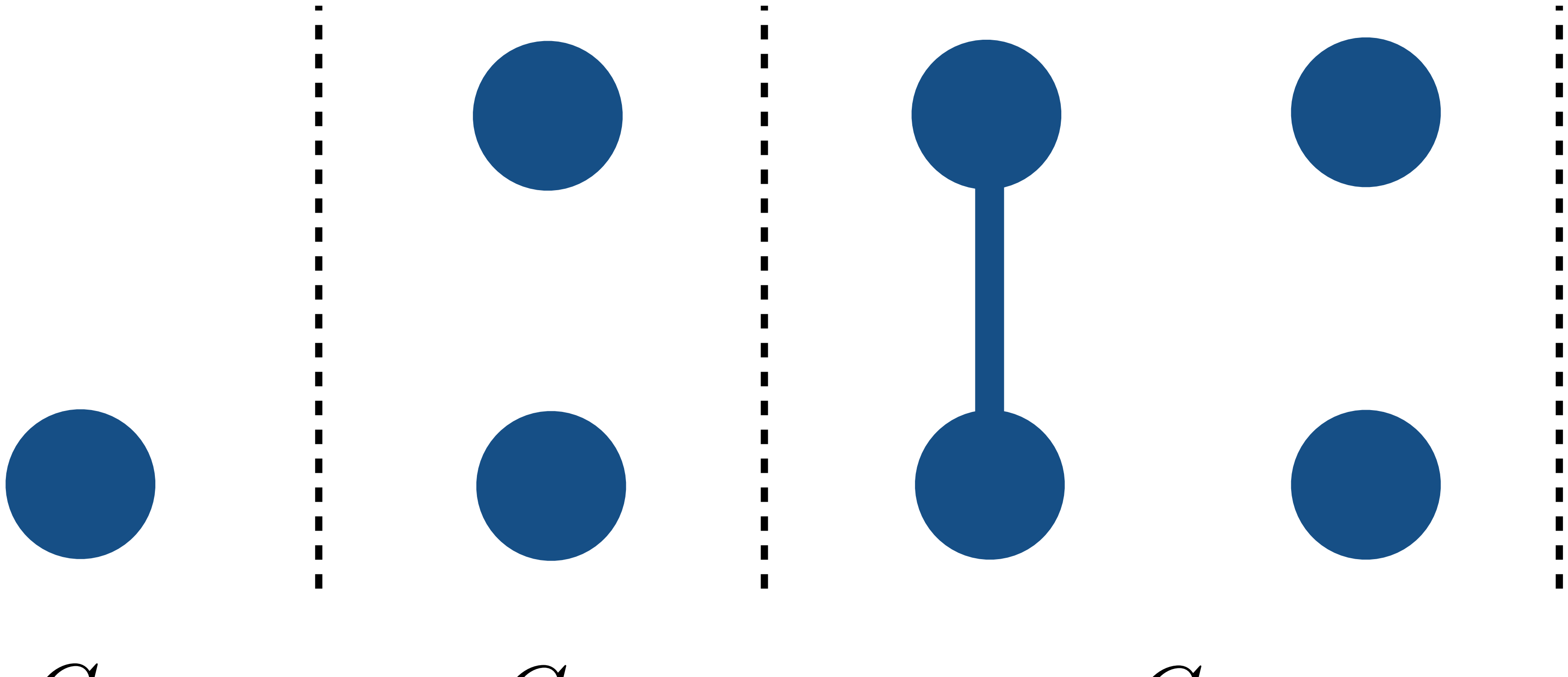


$G_1$ $G_2$

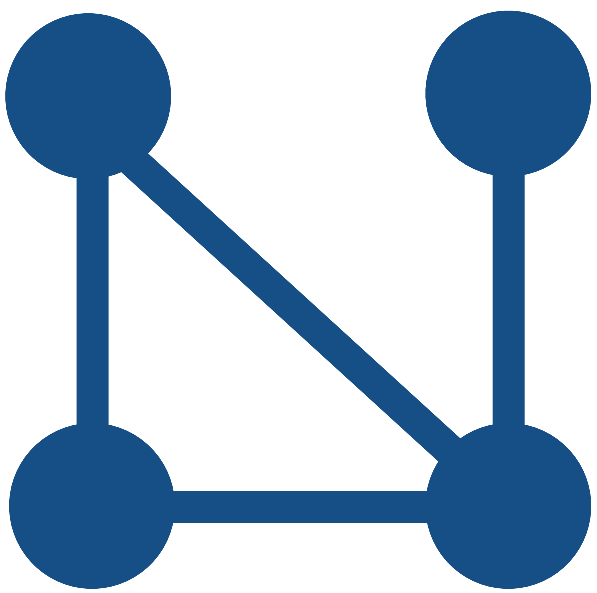


$G$

# Sequence of graphs

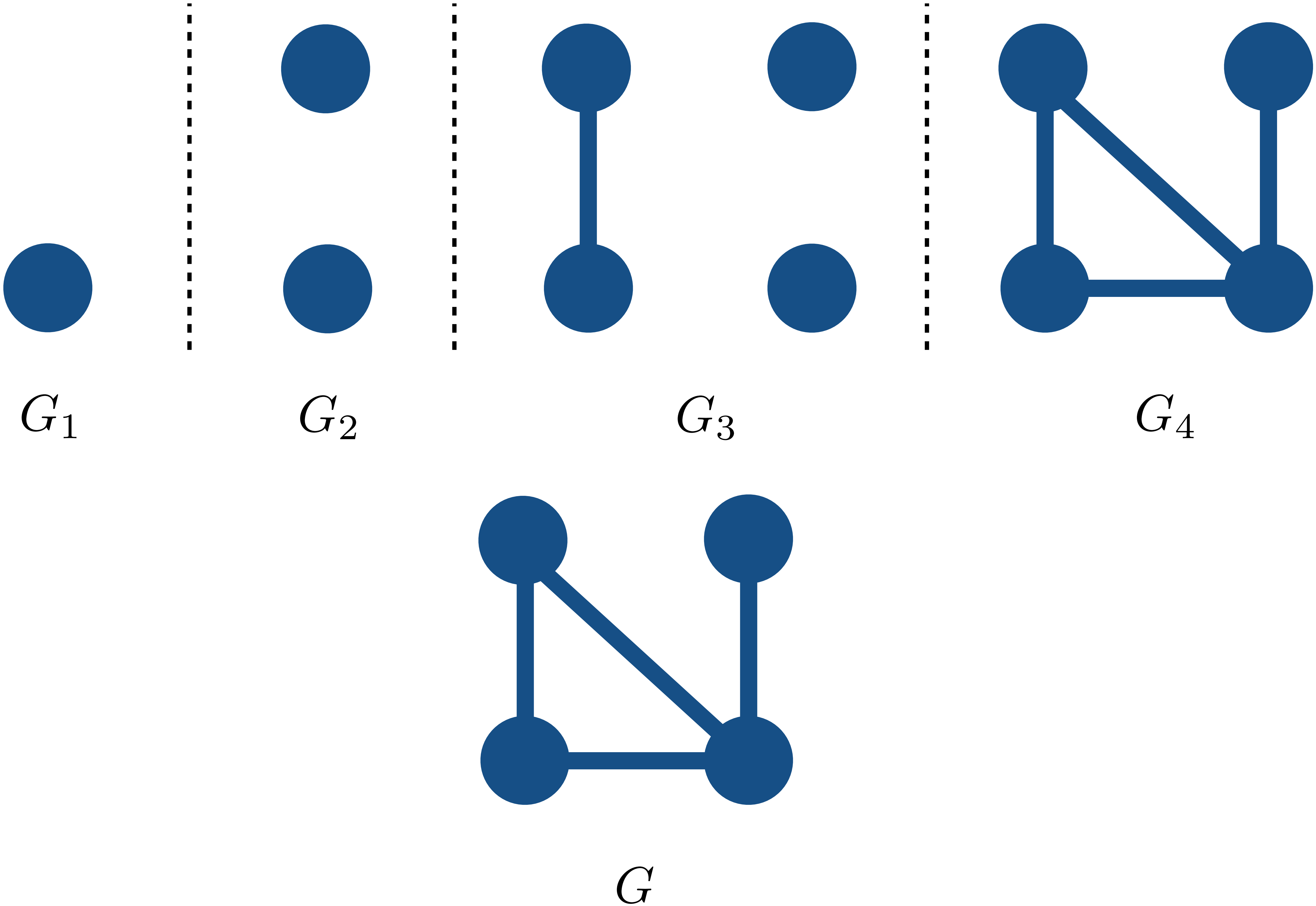


$G_1$ $G_2$ $G_3$

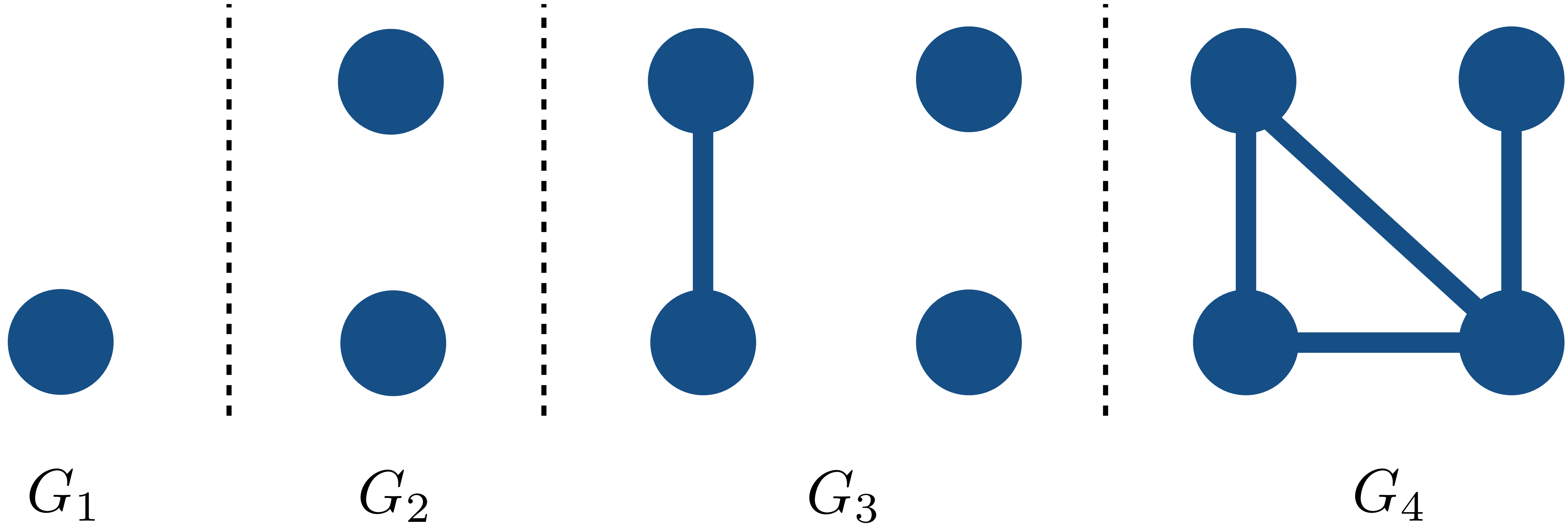


$G$

# Sequence of graphs

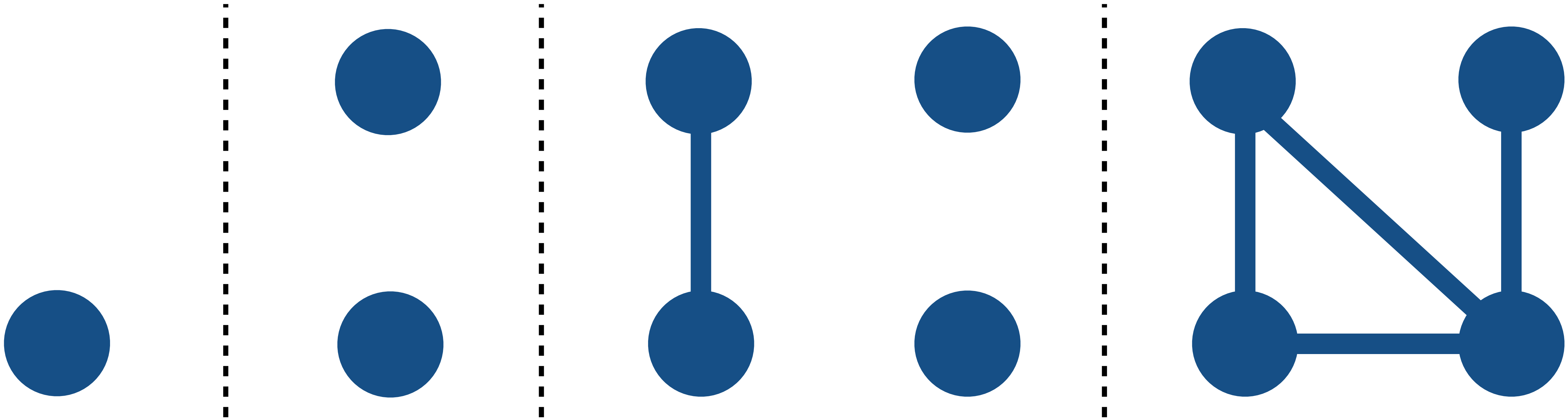


# Sequence of graphs



■ ■ ■

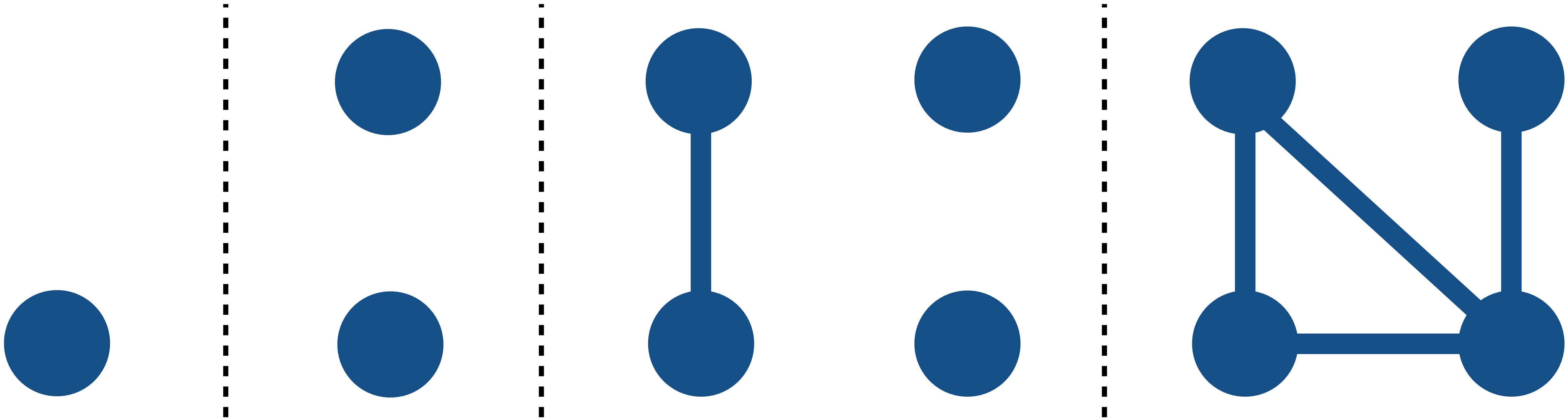
# Sequence of graphs



$G_1$ $G_2$ $G_3$ $G_4$

If $\#\text{nodes}(G_n) \to \infty$,

# Sequence of graphs

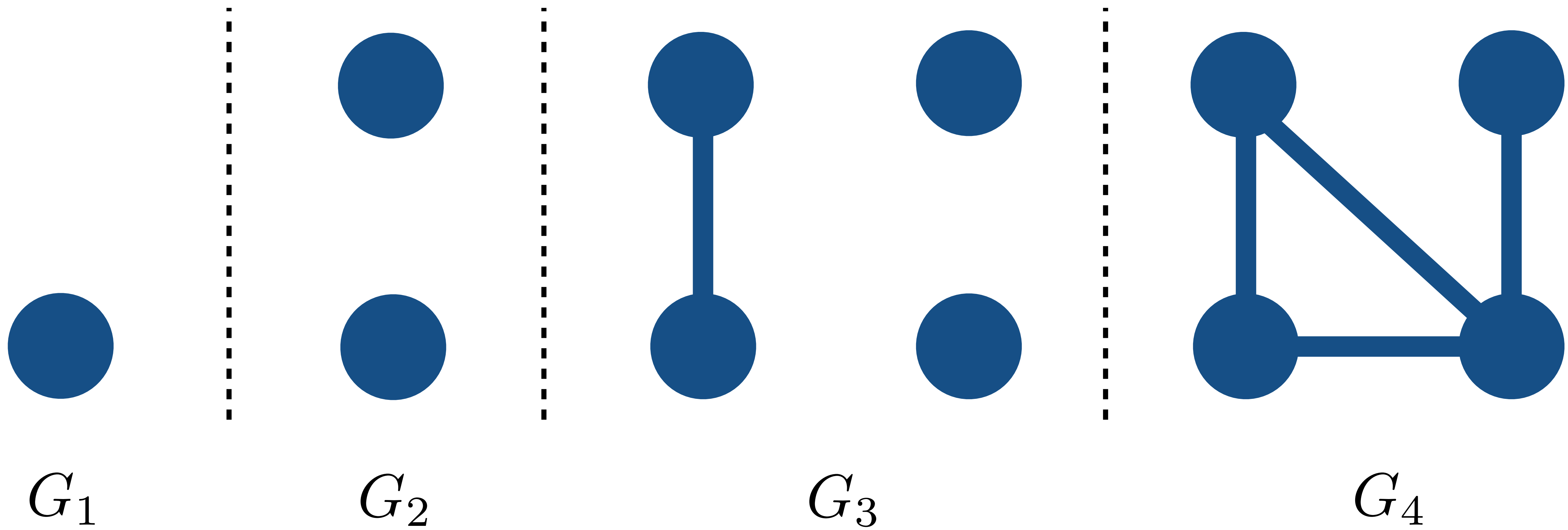


$G_1$ $G_2$ $G_3$ $G_4$

If $\#\text{nodes}(G_n) \to \infty$,

- *Dense* graph sequence $\#\text{edges}(G_n) \geq c \cdot [\#\text{nodes}(G_n)]^2$

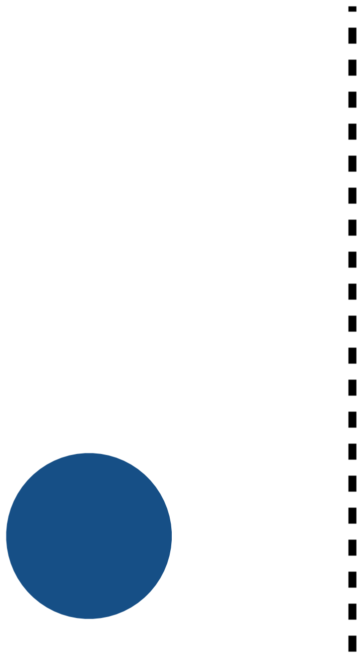
# Sequence of graphs



If $\#\text{nodes}(G_n) \to \infty$,

- *Dense* graph sequence $\#\text{edges}(G_n) \geq c \cdot [\#\text{nodes}(G_n)]^2$
- *Sparse* graph sequence $\#\text{edges}(G_n) \in o([\#\text{nodes}(G_n)]^2)$

# The Old Way: Nodes

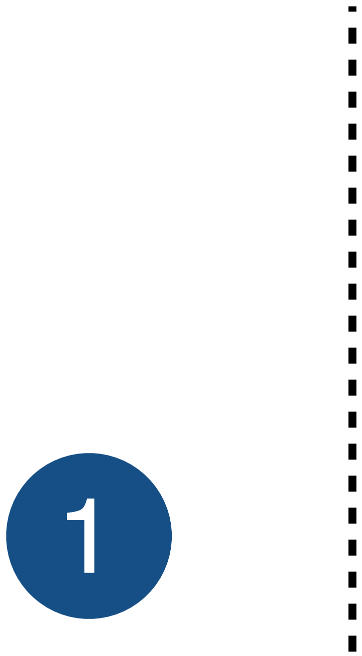


$G_1$



[Hoover 1979, Aldous 1981]

# The Old Way: Nodes

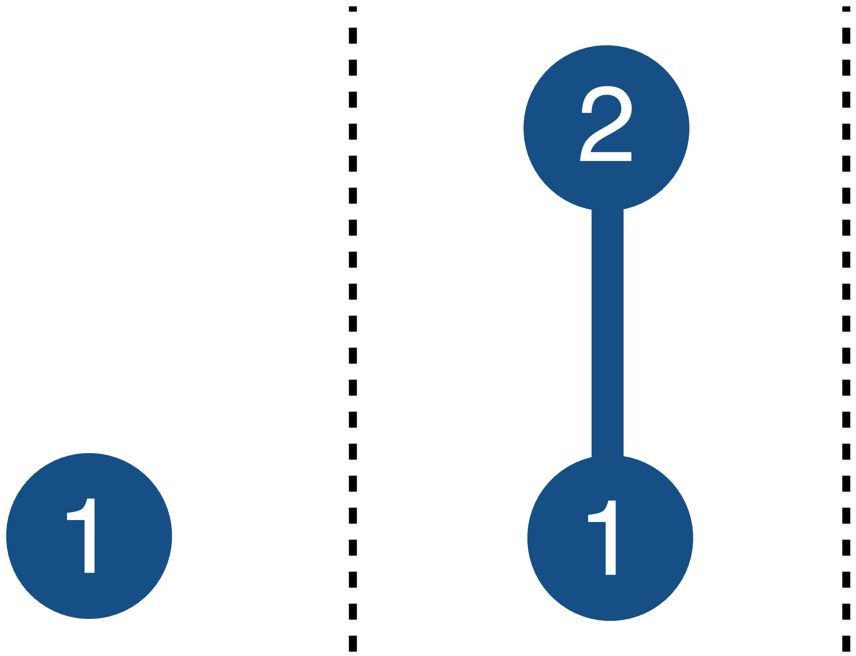



$G_1$



[Hoover 1979, Aldous 1981]

# The Old Way: Nodes

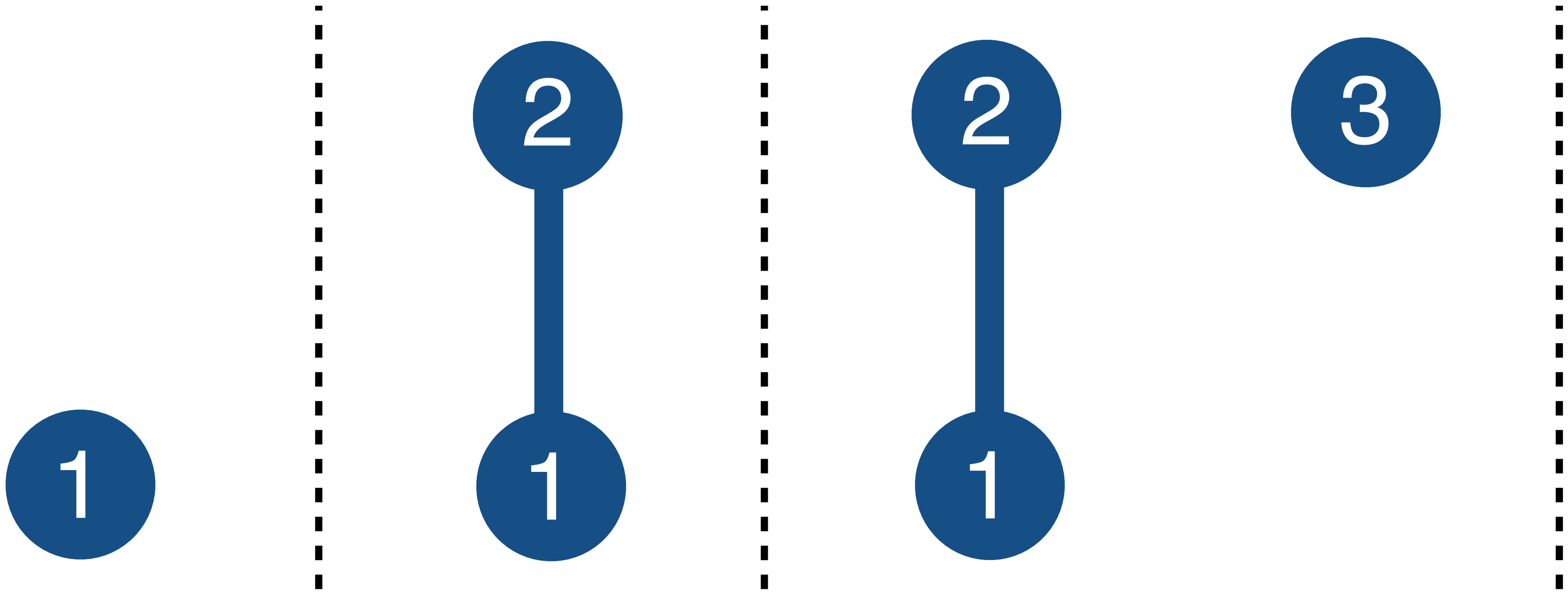



$G_1$ $G_2$



[Hoover 1979, Aldous 1981]

# The Old Way: Nodes

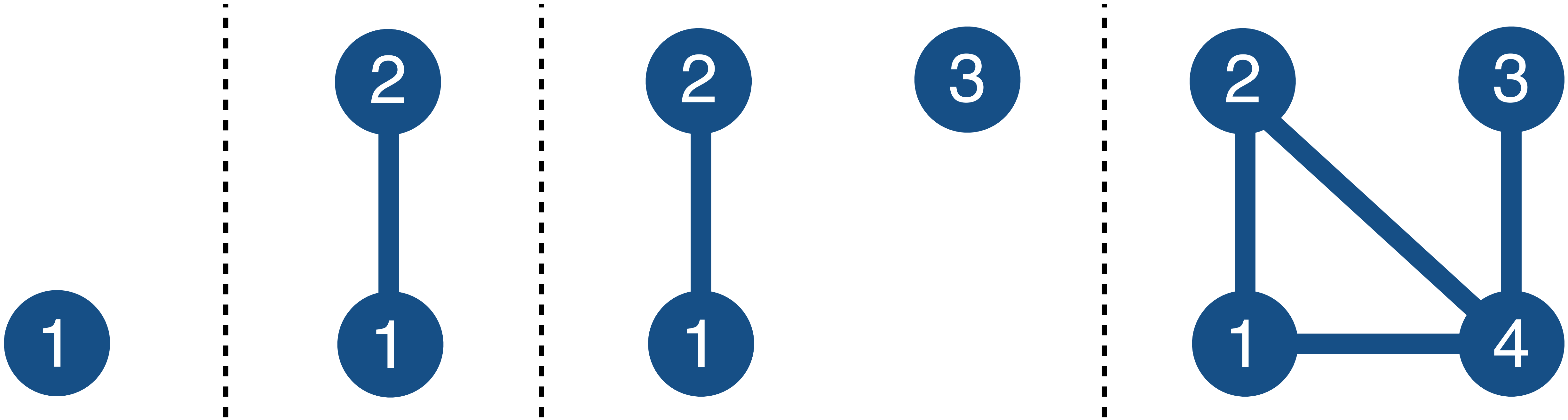



$G_1$ $G_2$ $G_3$



[Hoover 1979, Aldous 1981]

# The Old Way: Nodes

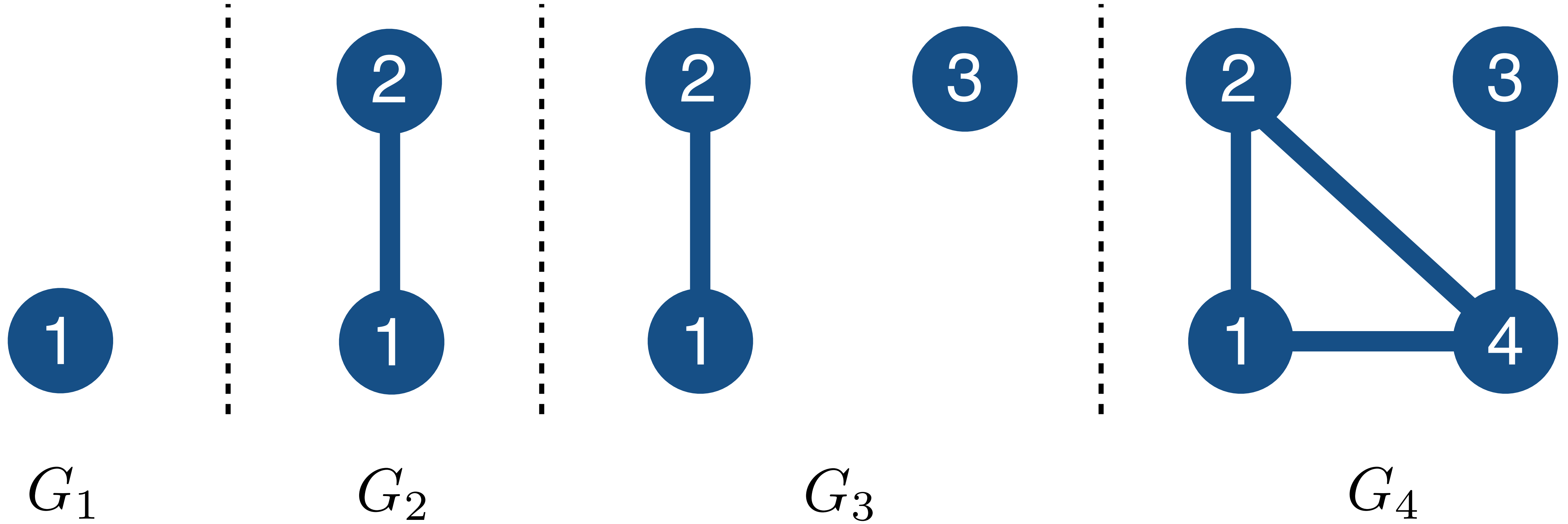



$G_1$ $G_2$ $G_3$ $G_4$



[Hoover 1979, Aldous 1981]

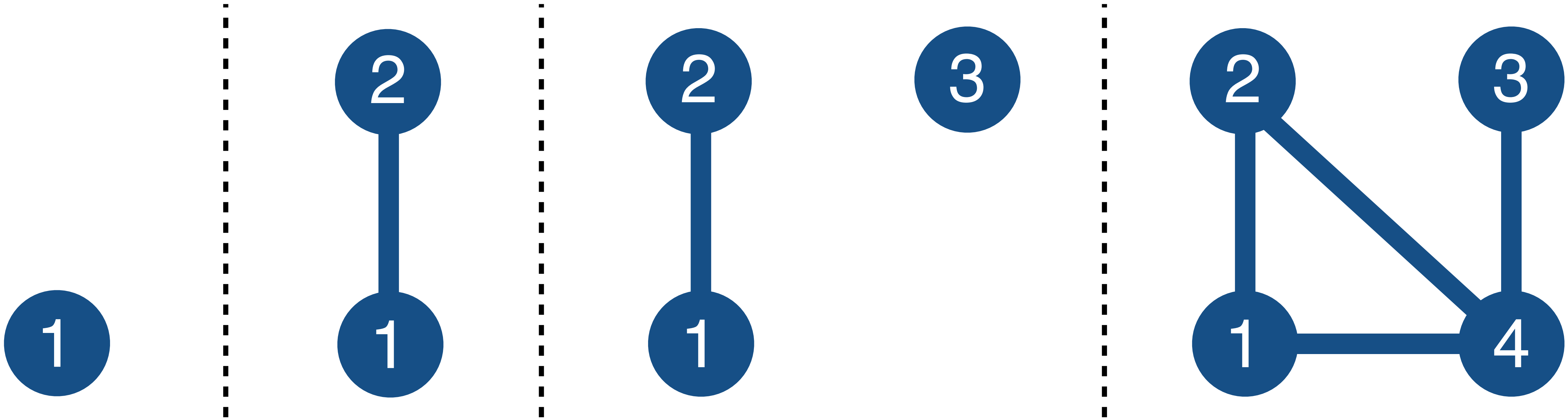
# The Old Way: Exchangeability



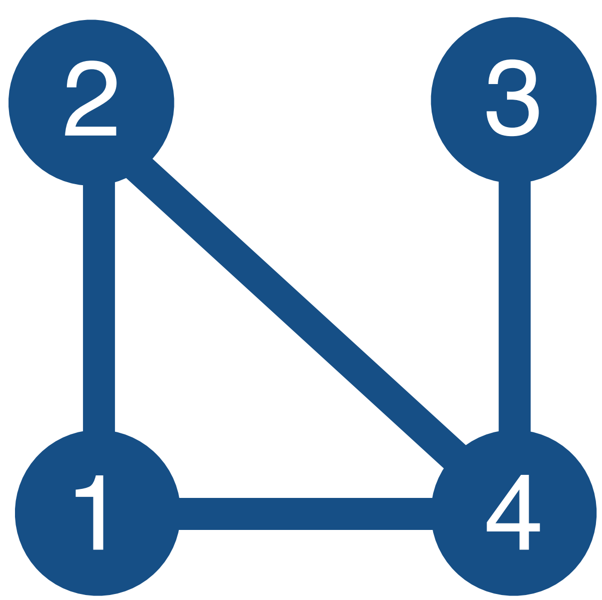



[Hoover 1979, Aldous 1981]

# The Old Way: Exchangeability



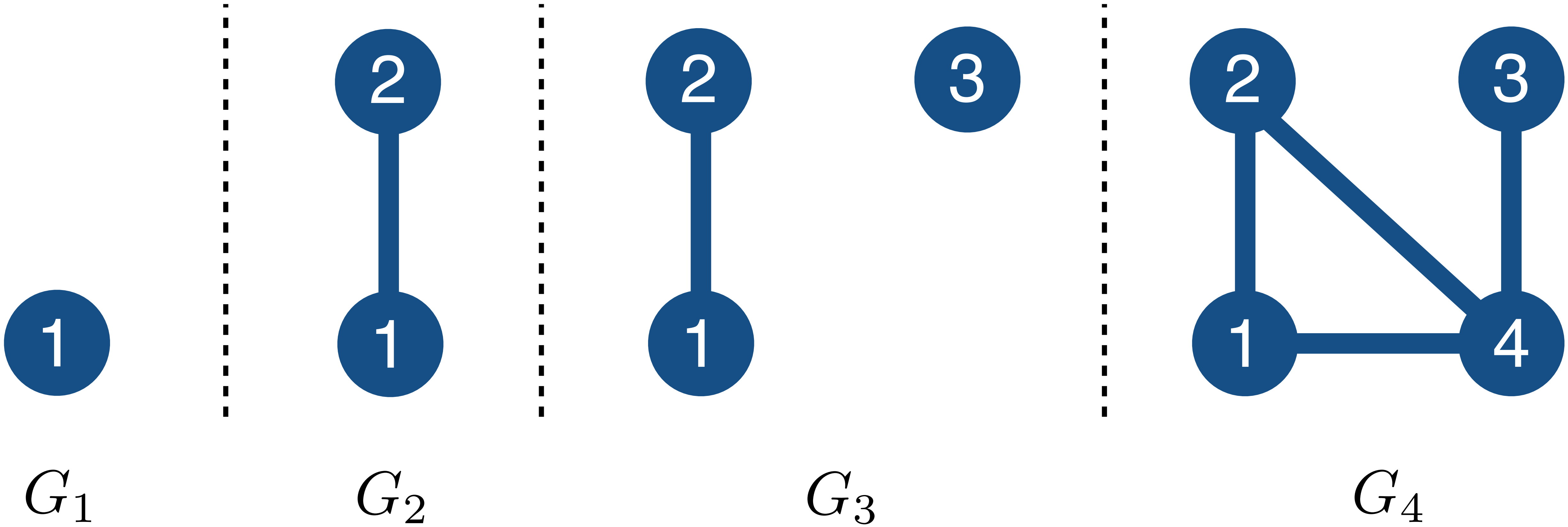

$G_1$ $G_2$ $G_3$ $G_4$

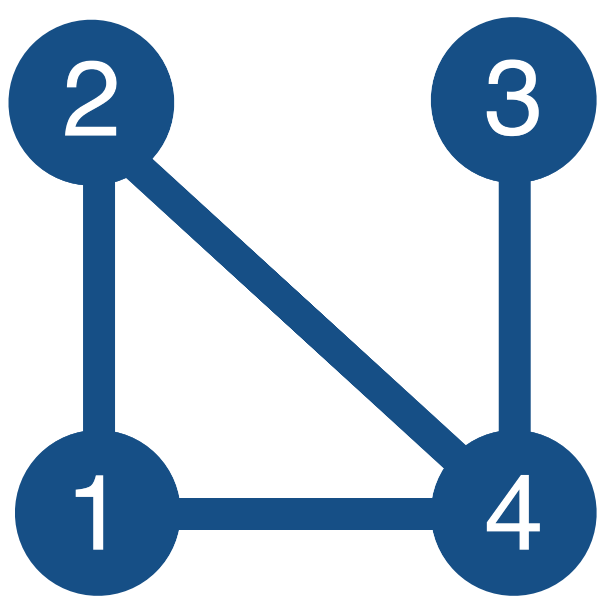





[Hoover 1979, Aldous 1981]

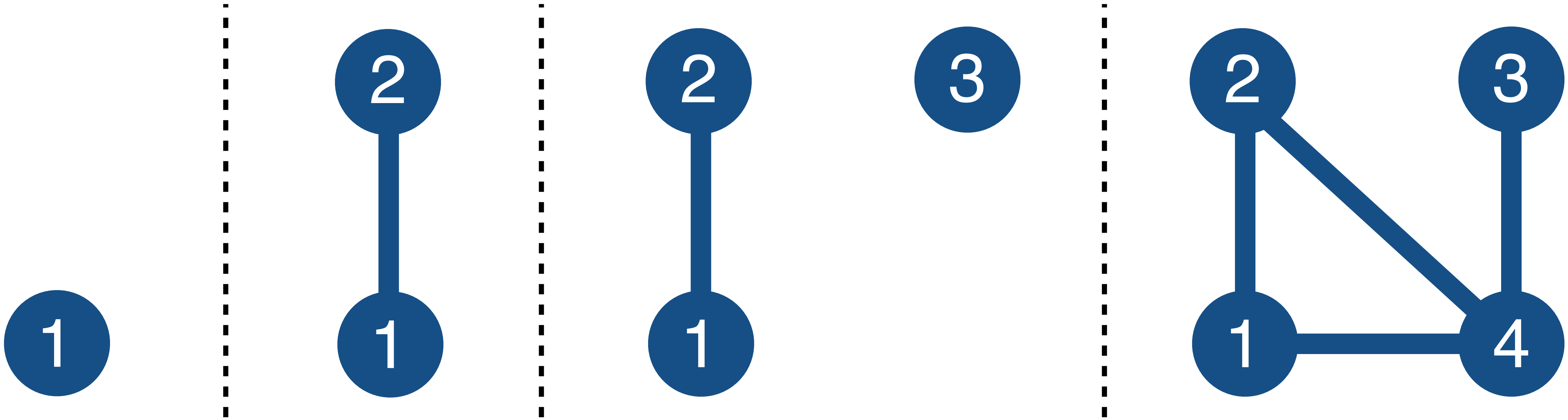
# The Old Way: Exchangeability



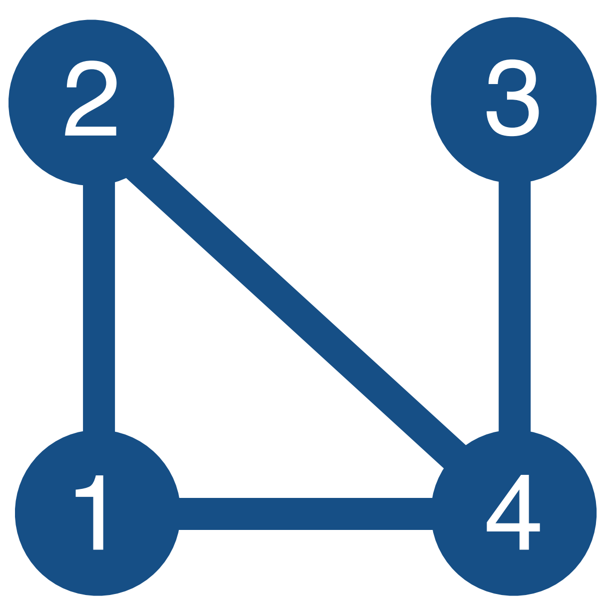

$p($  $)$

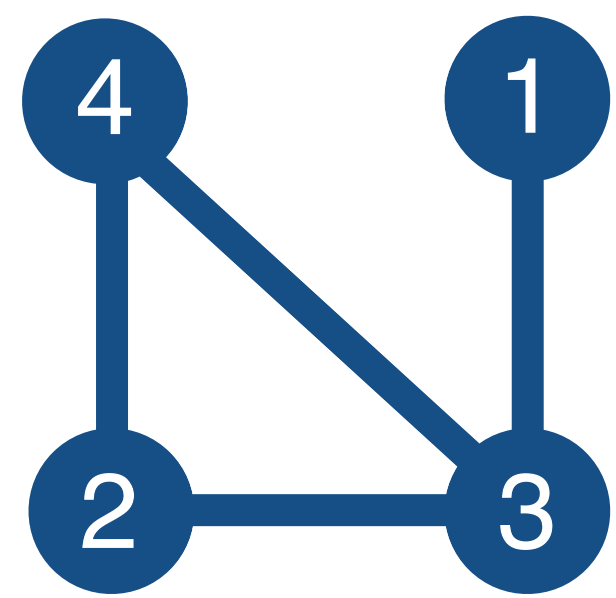

[Hoover 1979, Aldous 1981]

# The Old Way: Exchangeability




$G_1$ $G_2$ $G_3$ $G_4$

$p($ 

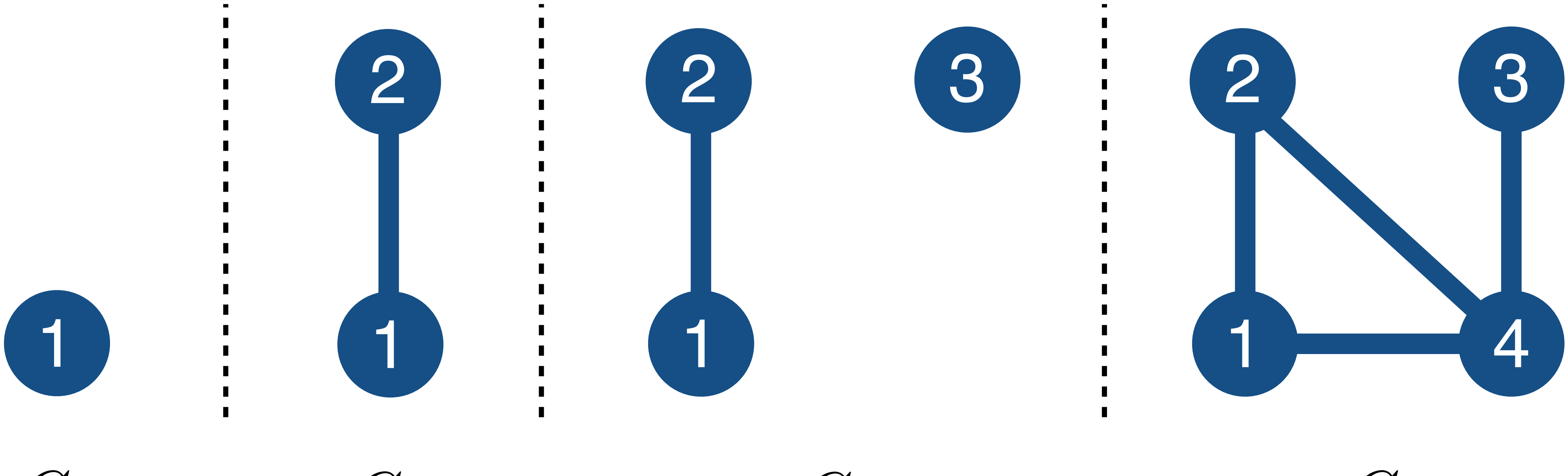
$) = p($ 

$)$

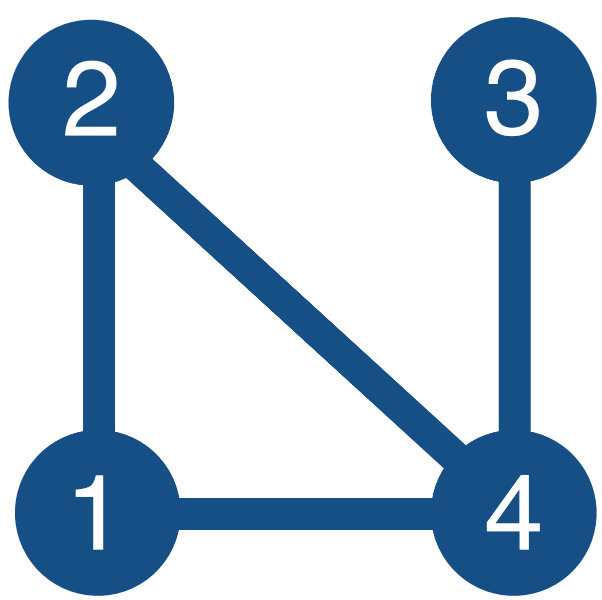


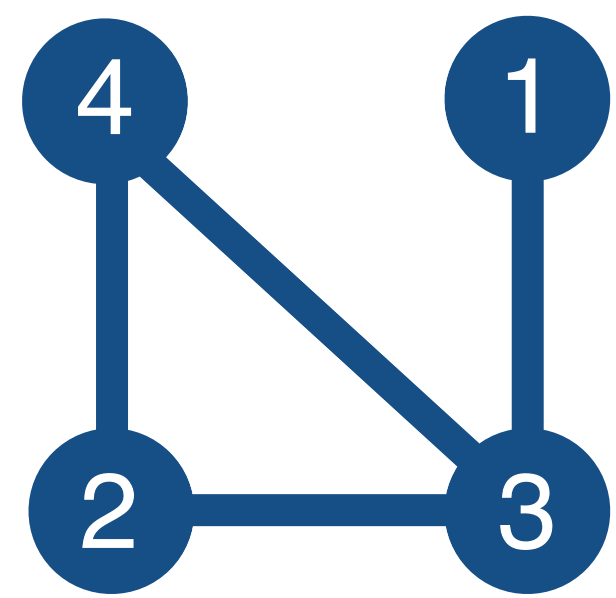
[Hoover 1979, Aldous 1981]

# The Old Way: Node exchangeability




$G_1$ $G_2$ $G_3$ $G_4$

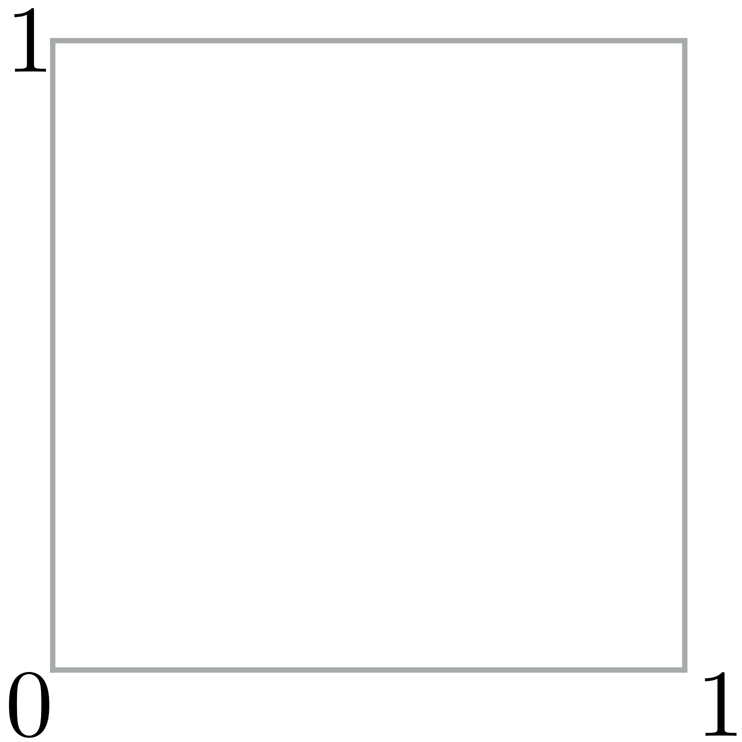
$p($ 

$) = p($ 

$)$

[Hoover 1979, Aldous 1981]

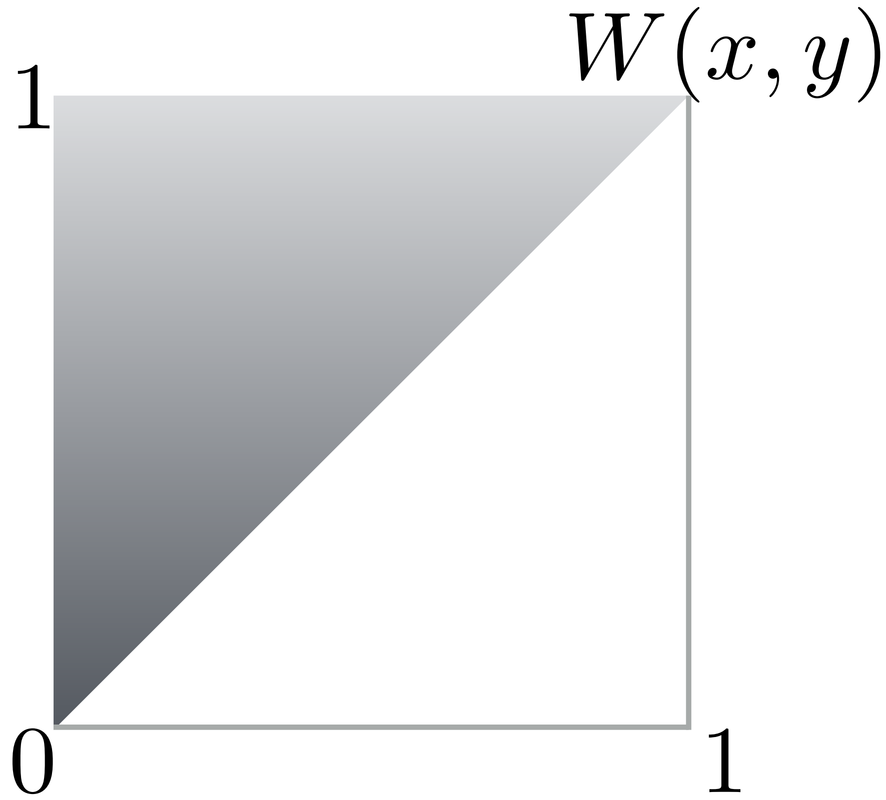
# Aldous-Hoover






[Hoover 1979, Aldous 1981]

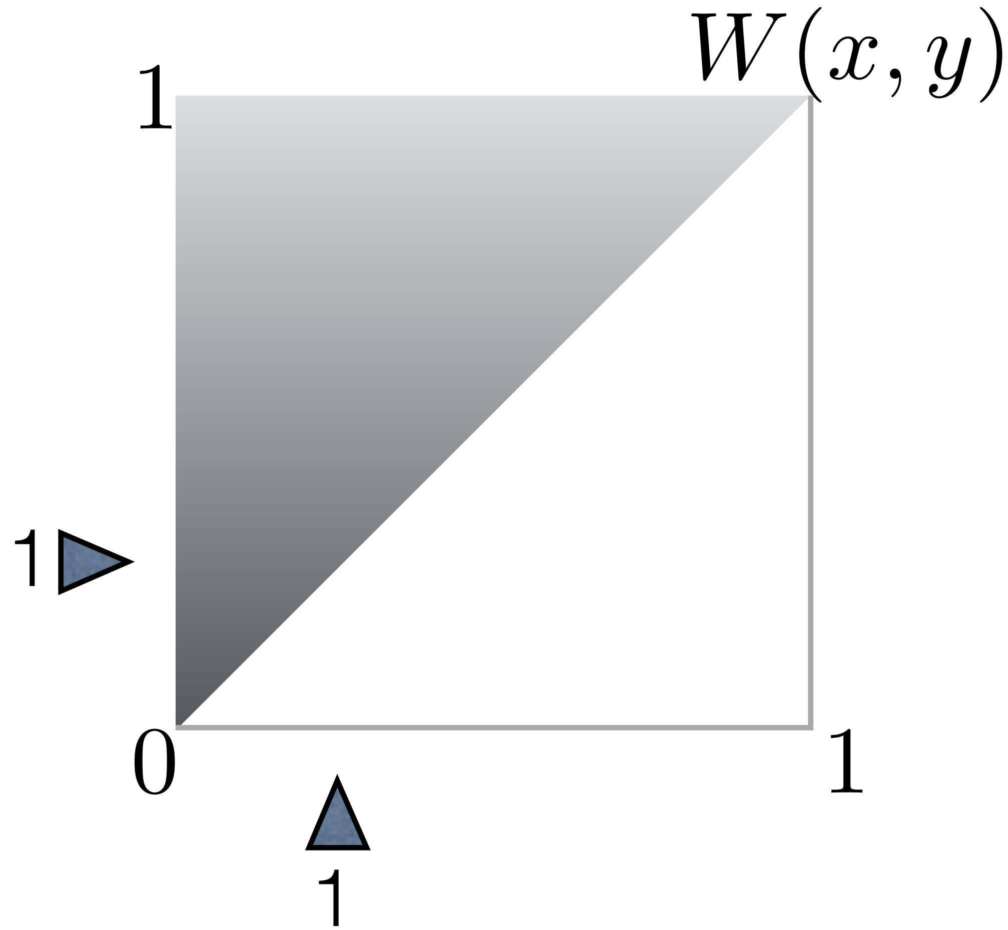
# Aldous-Hoover






[Hoover 1979, Aldous 1981]

# Aldous-Hoover

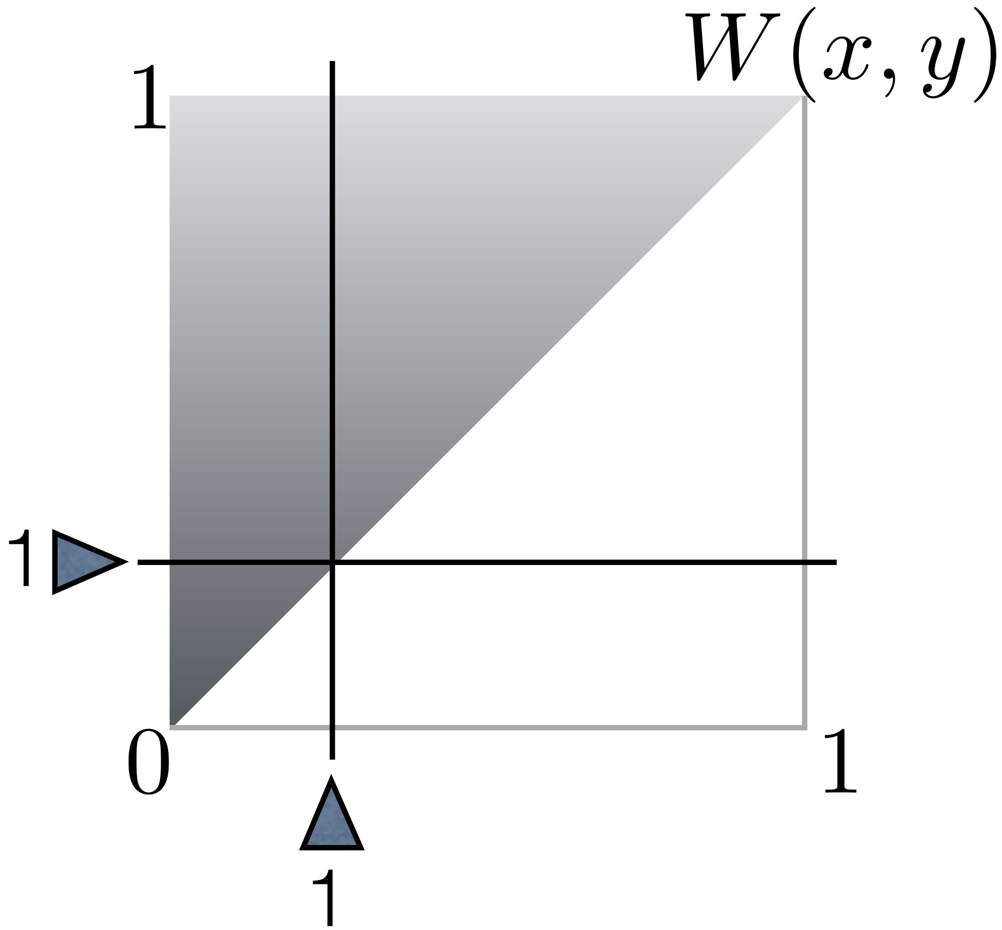






[Hoover 1979, Aldous 1981]

# Aldous-Hoover

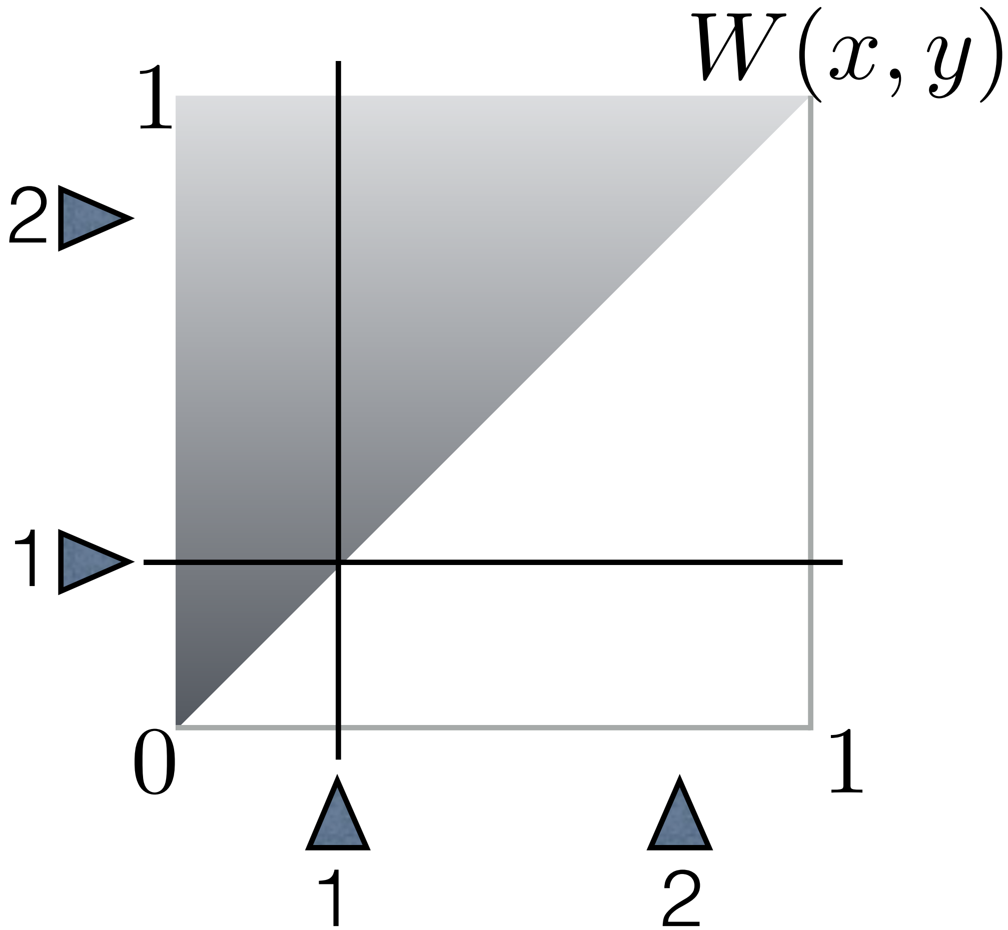






[Hoover 1979, Aldous 1981]

# Aldous-Hoover

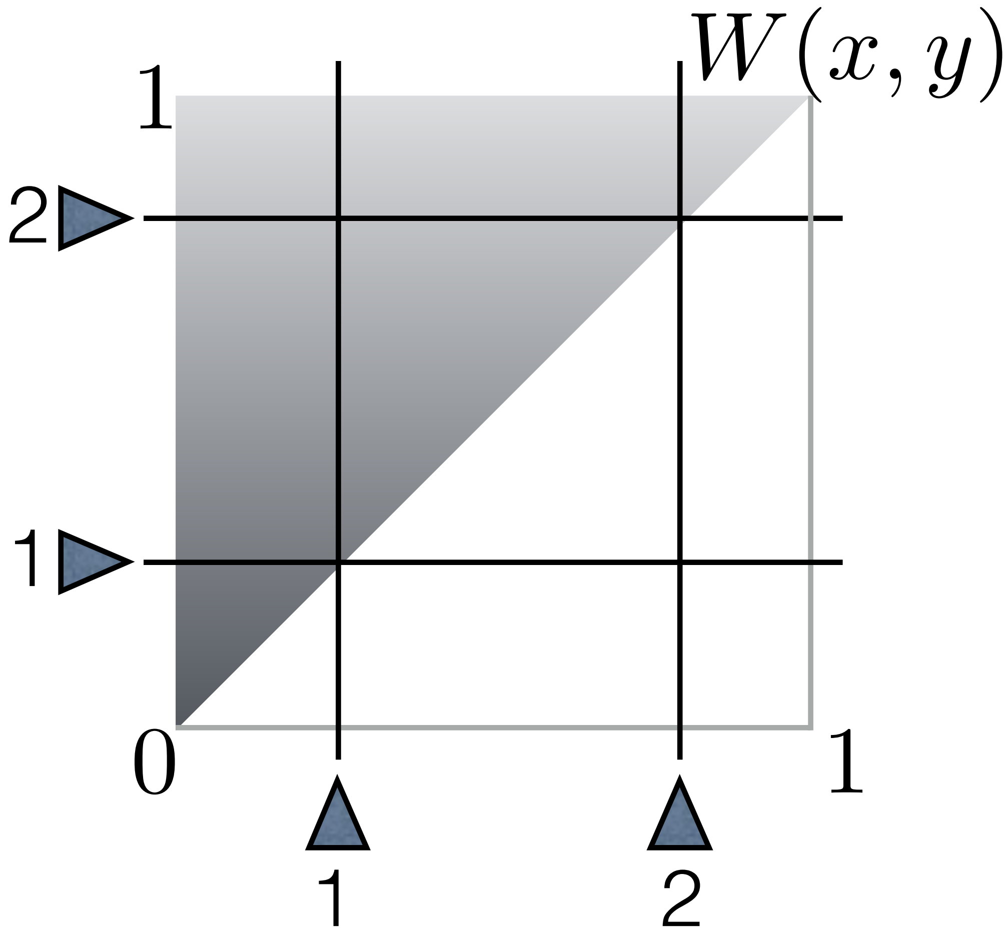






[Hoover 1979, Aldous 1981]

# Aldous-Hoover

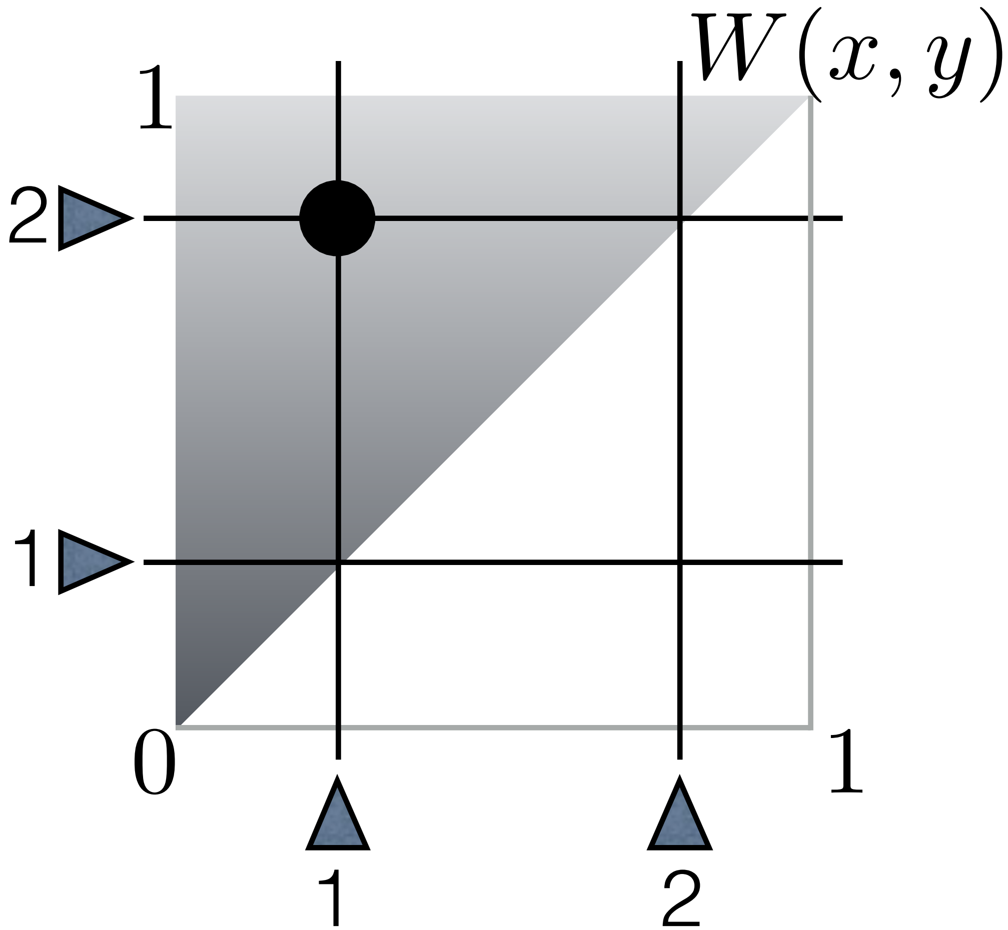


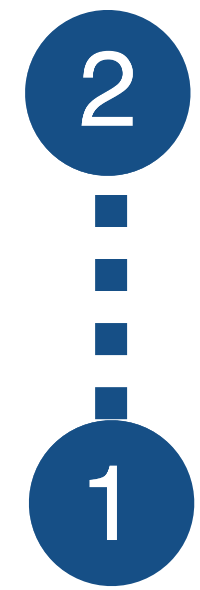




[Hoover 1979, Aldous 1981]

# Aldous-Hoover

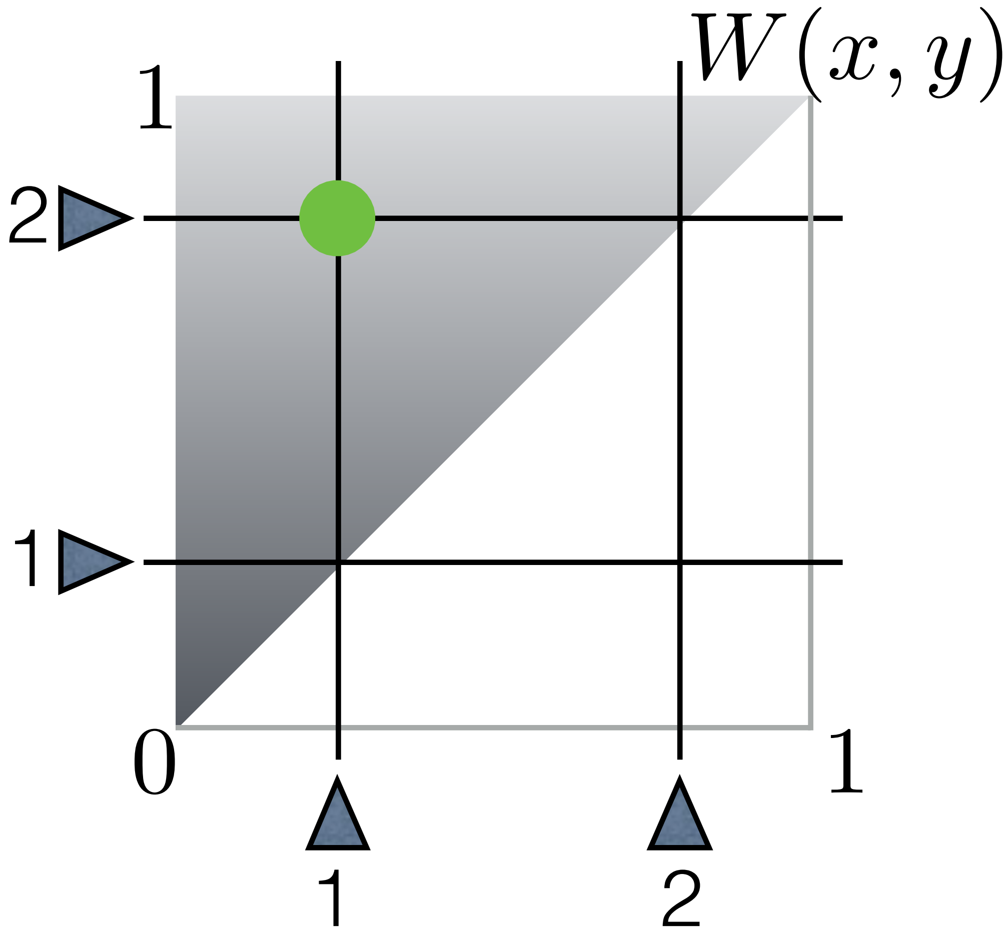


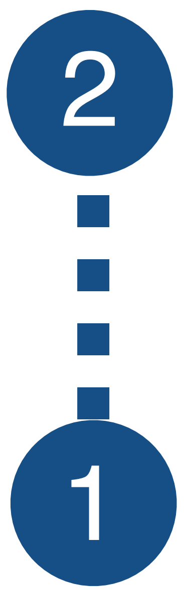






[Hoover 1979, Aldous 1981]

# Aldous-Hoover

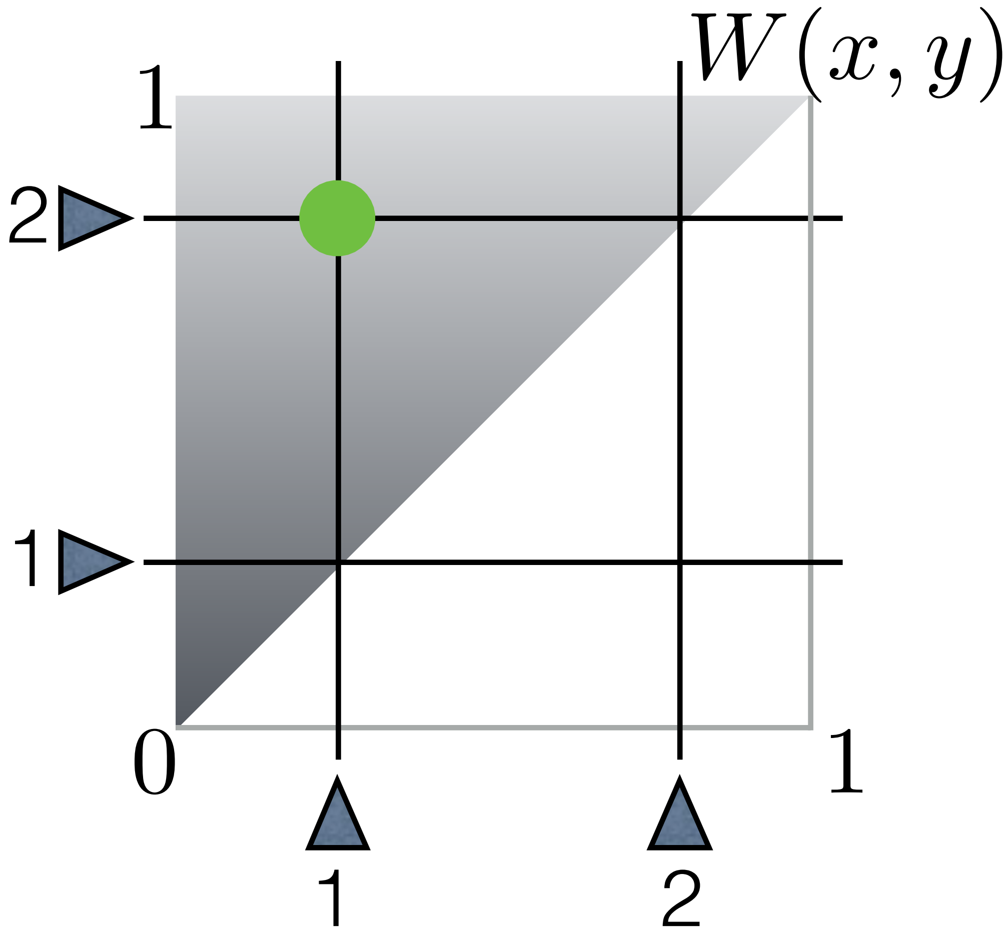





[Hoover 1979, Aldous 1981]

# Aldous-Hoover

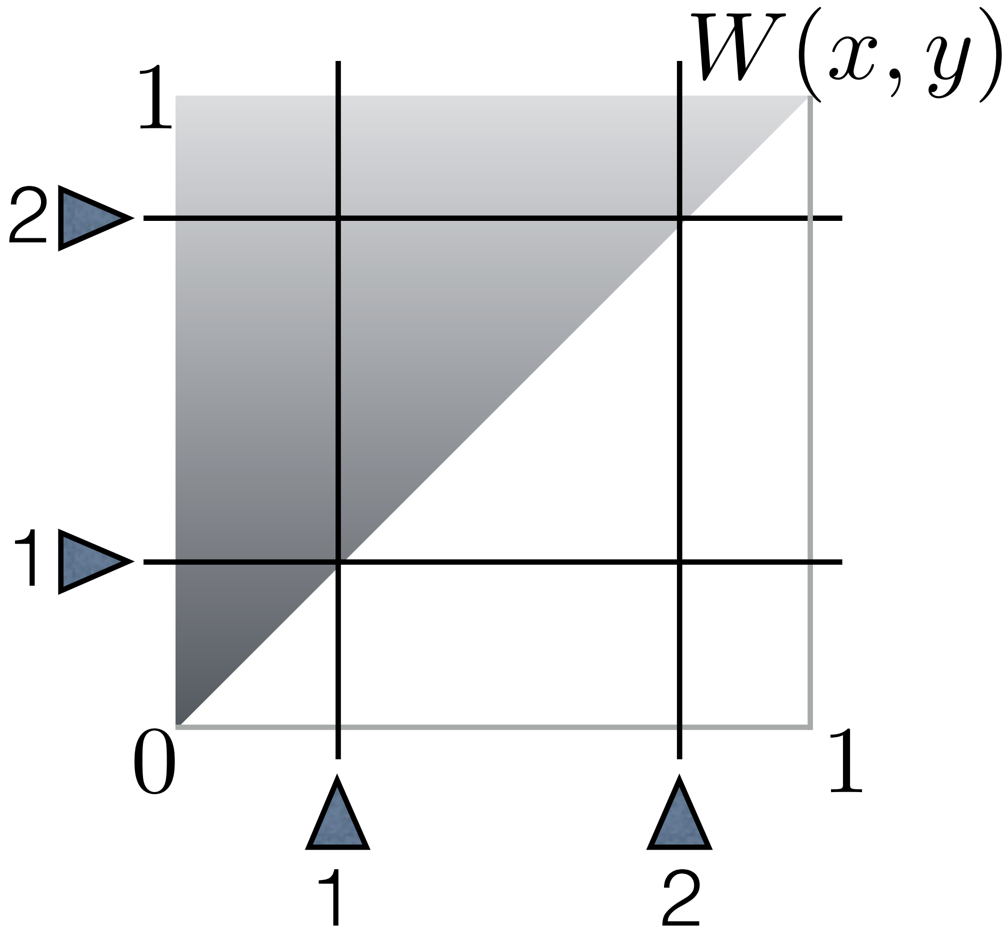


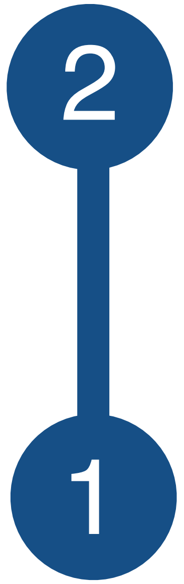




[Hoover 1979, Aldous 1981]

# Aldous-Hoover

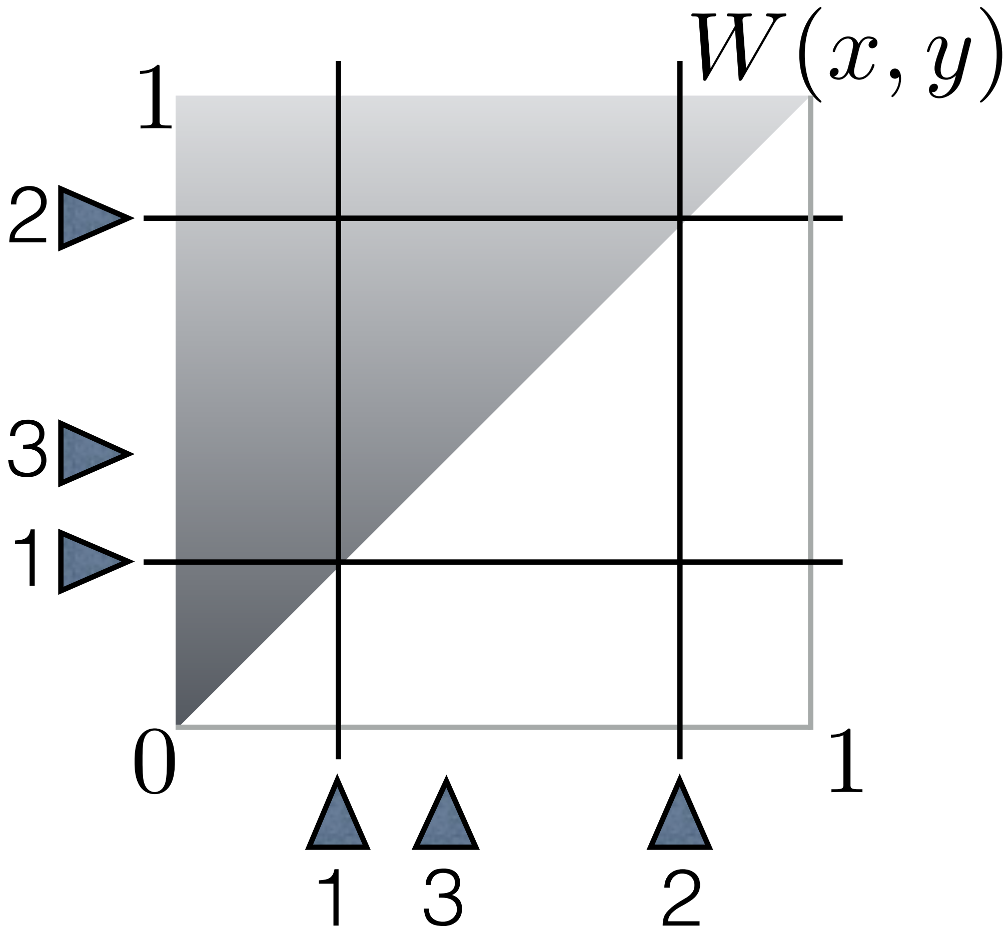


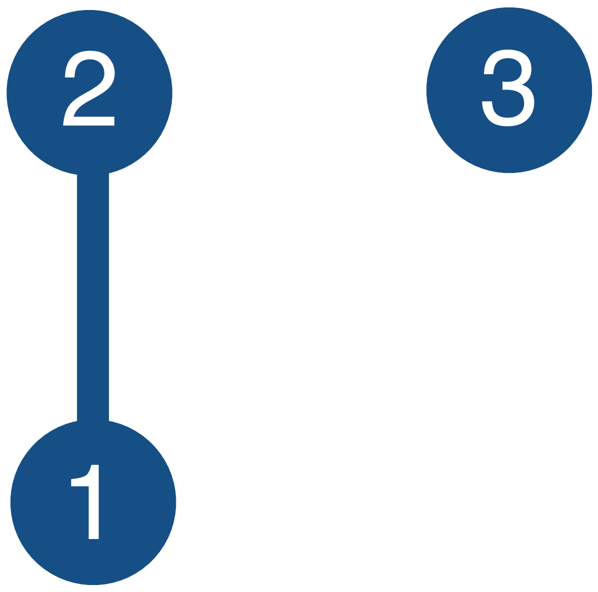






[Hoover 1979, Aldous 1981]

# Aldous-Hoover

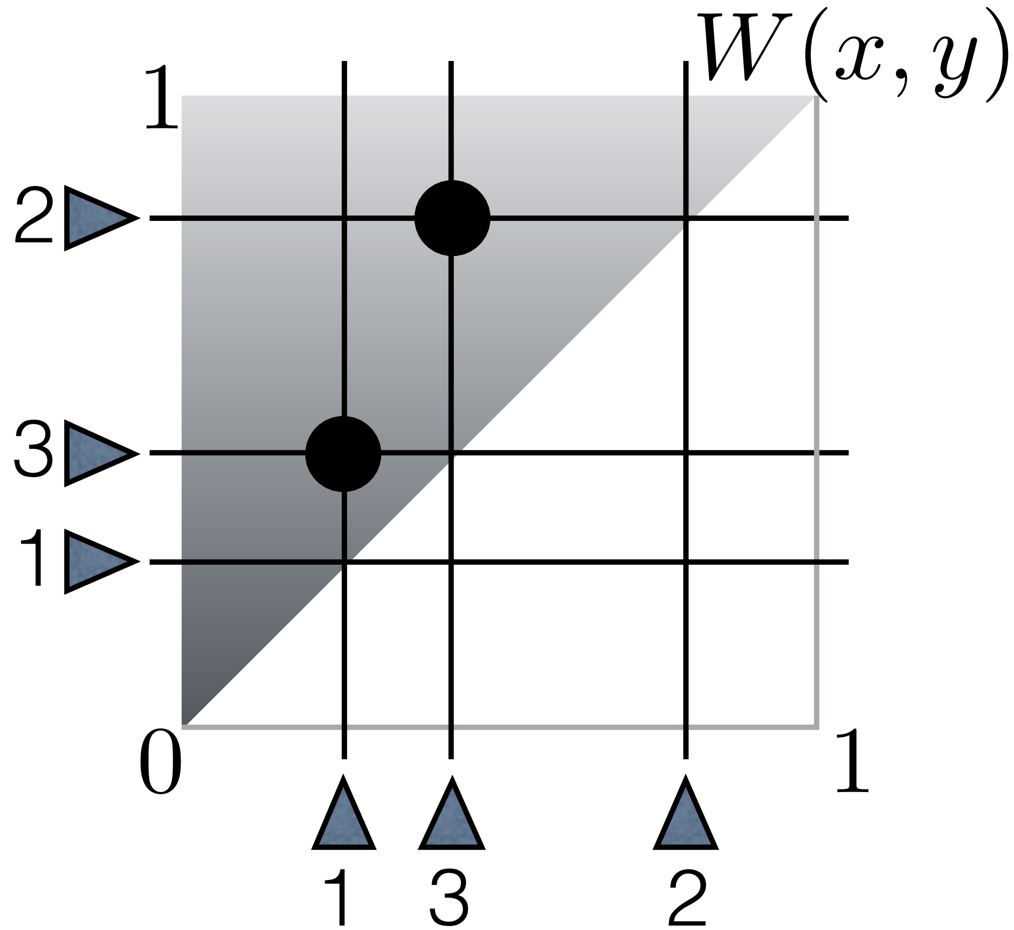


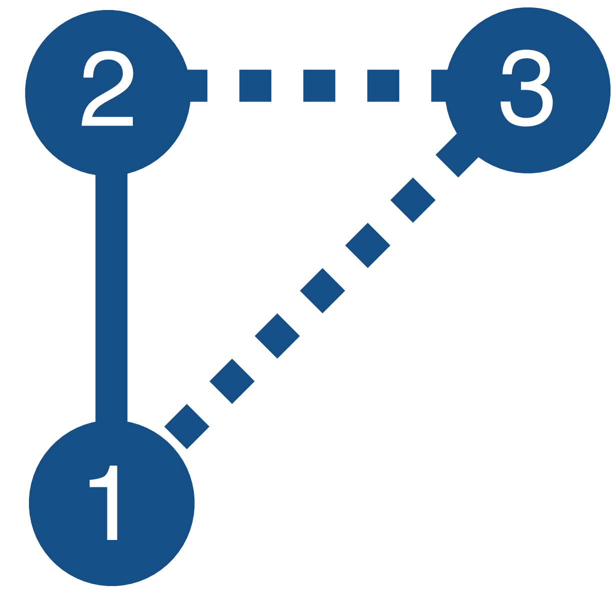






[Hoover 1979, Aldous 1981]

# Aldous-Hoover

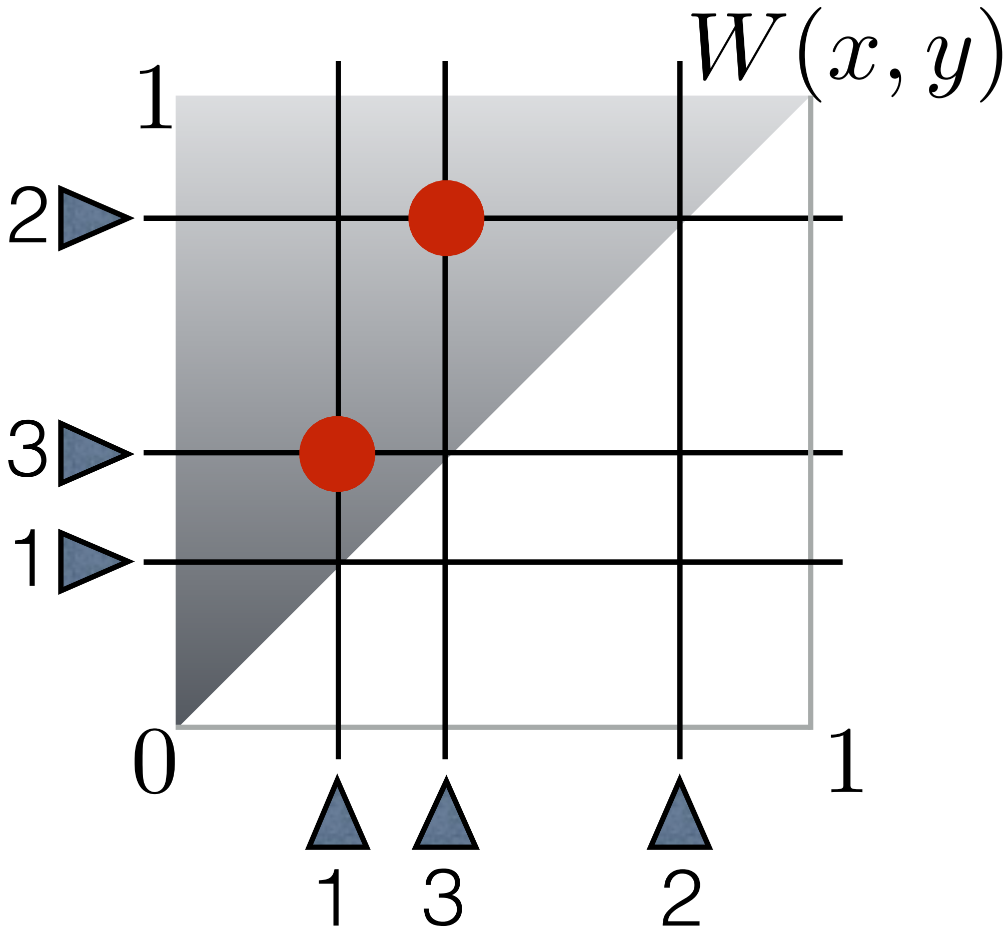


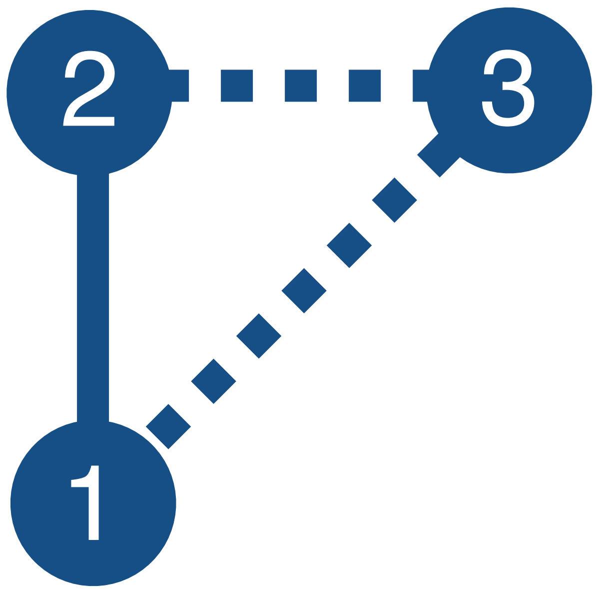






[Hoover 1979, Aldous 1981]

# Aldous-Hoover

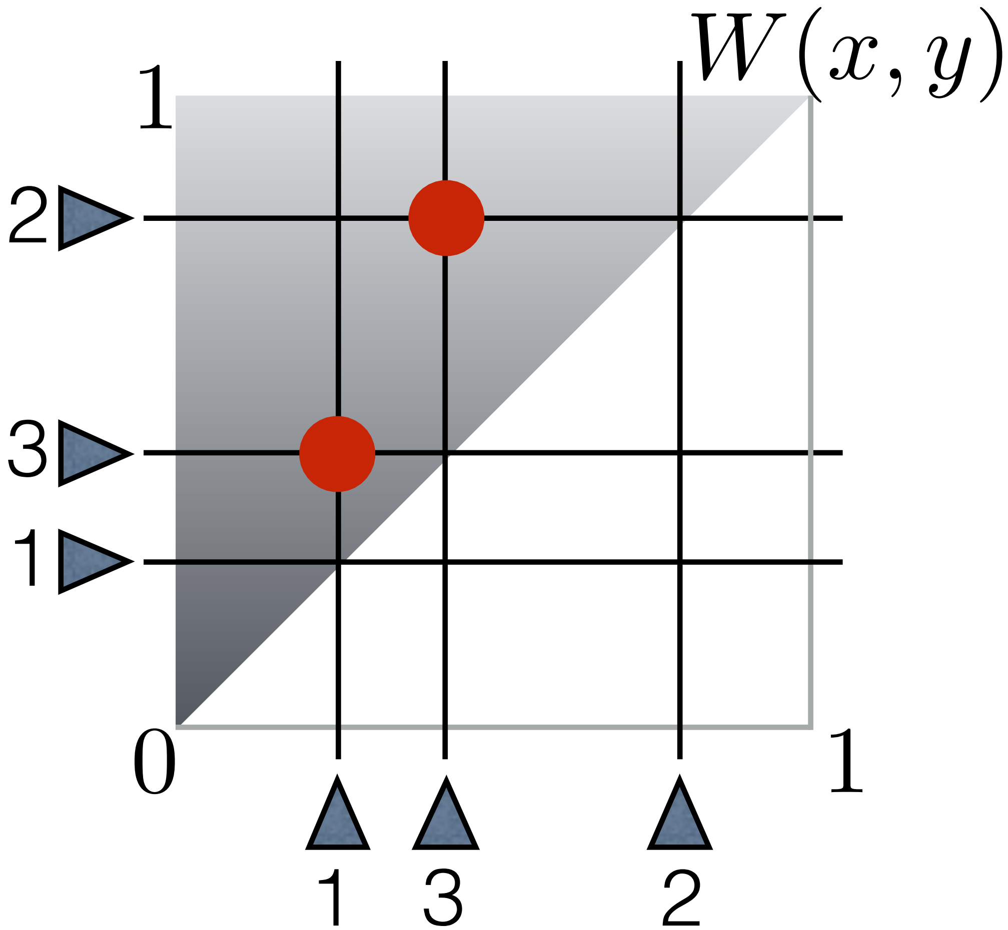


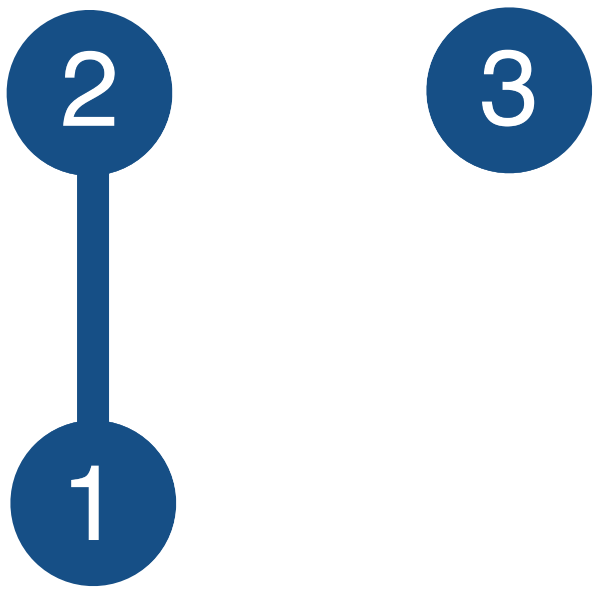






[Hoover 1979, Aldous 1981]

# Aldous-Hoover

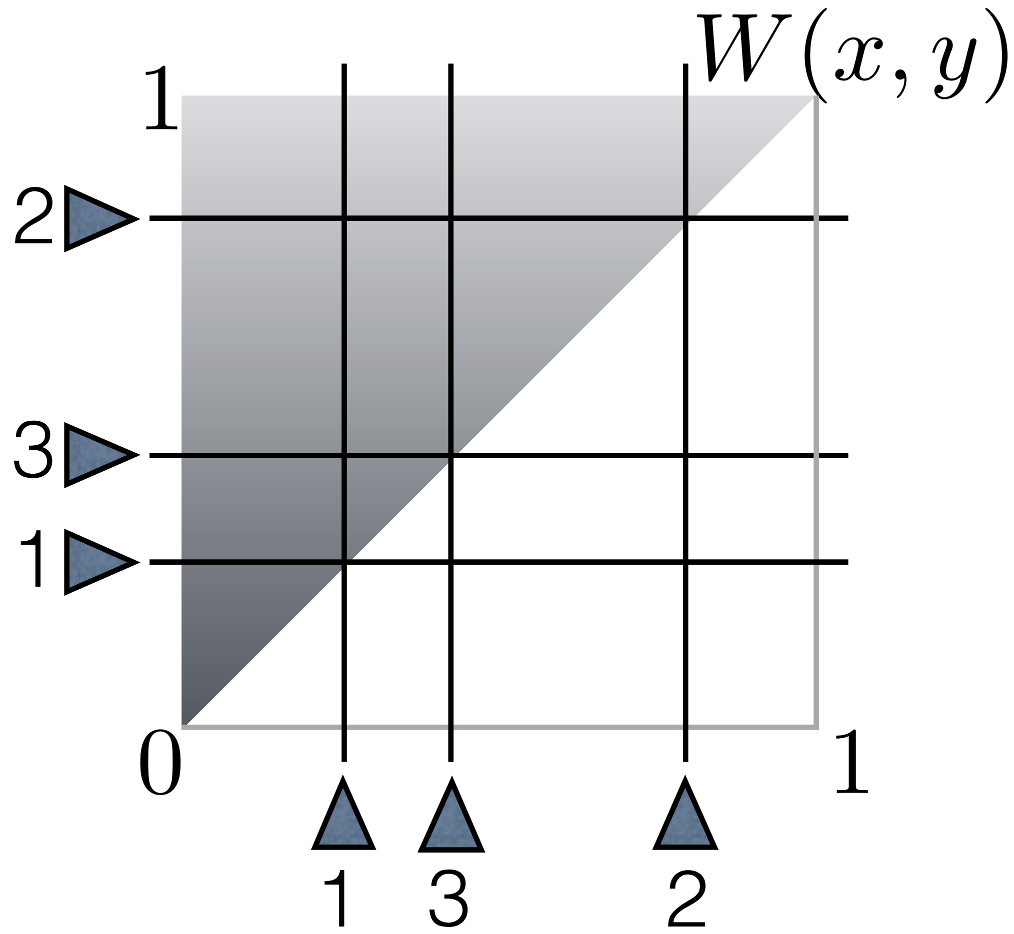


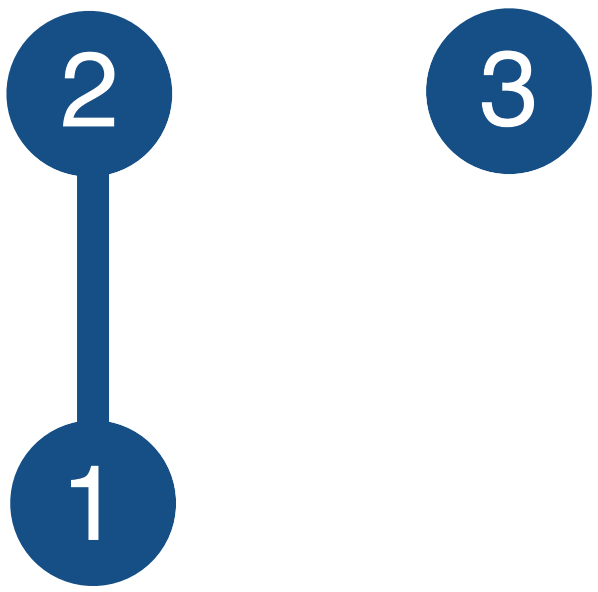






[Hoover 1979, Aldous 1981]

# Aldous-Hoover

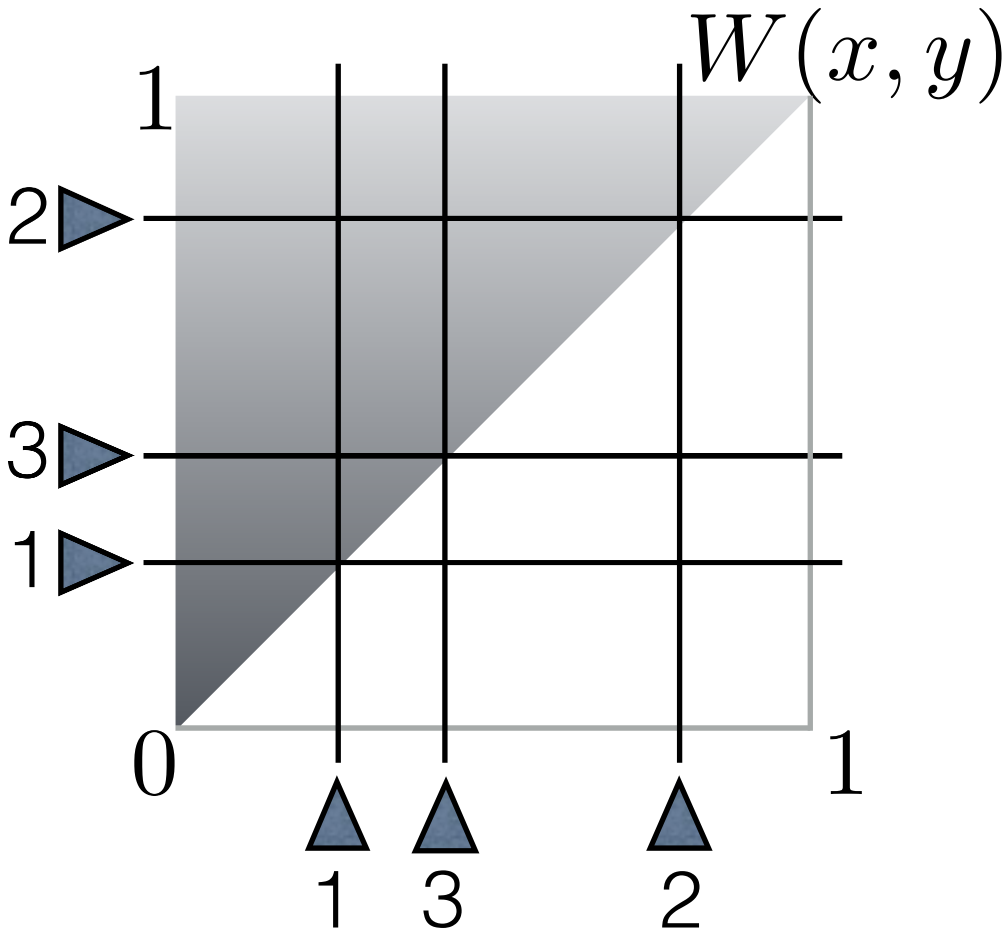


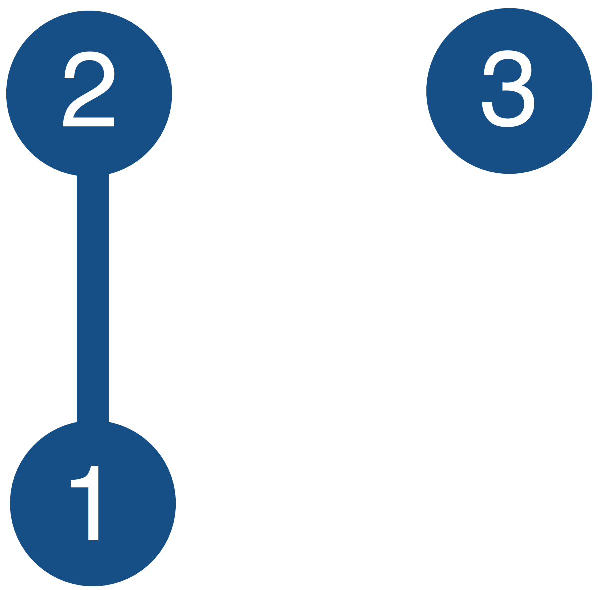






[Hoover 1979, Aldous 1981]

# Aldous-Hoover



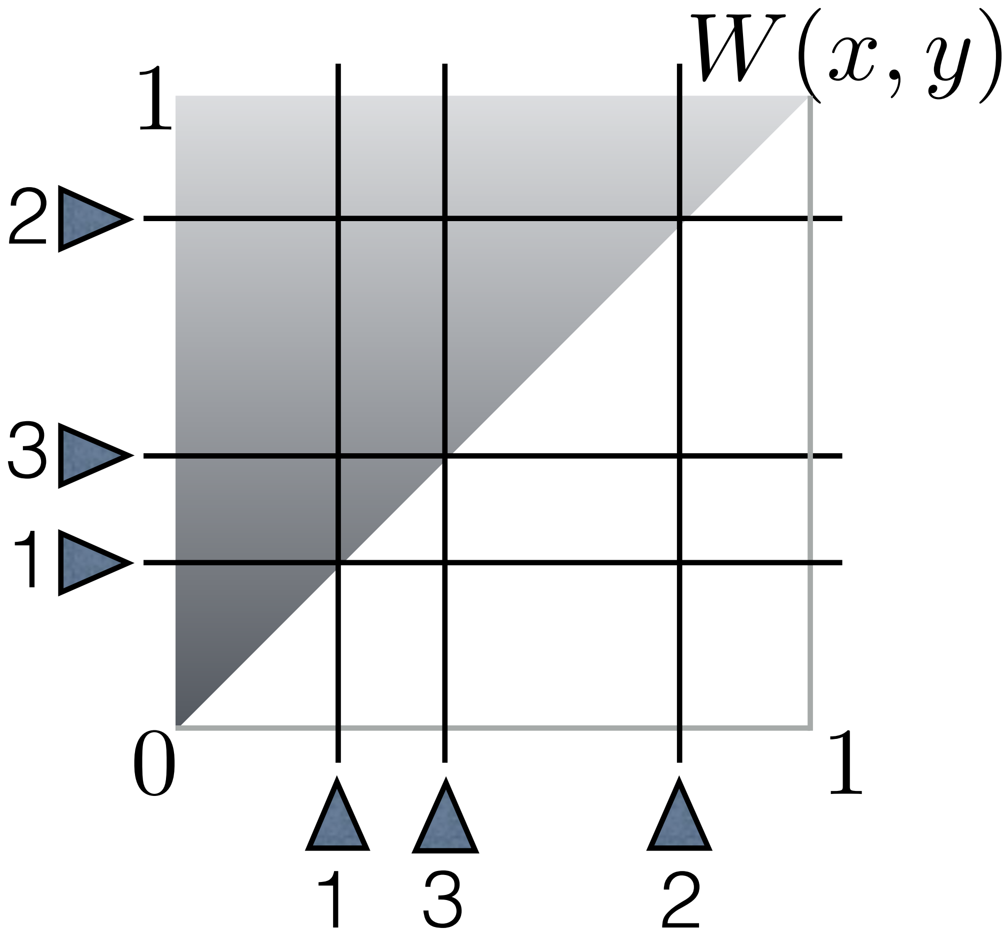



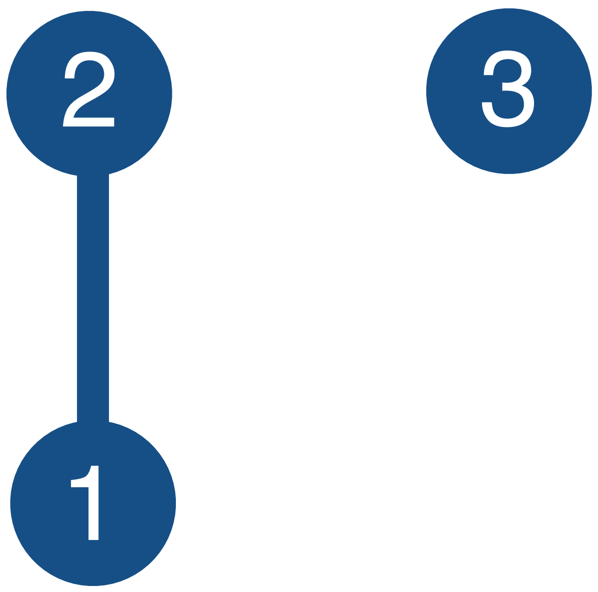

Thm (AH). Every node-exchangeable graph has a *graphon* rep



[Hoover 1979, Aldous 1981]

# Aldous-Hoover



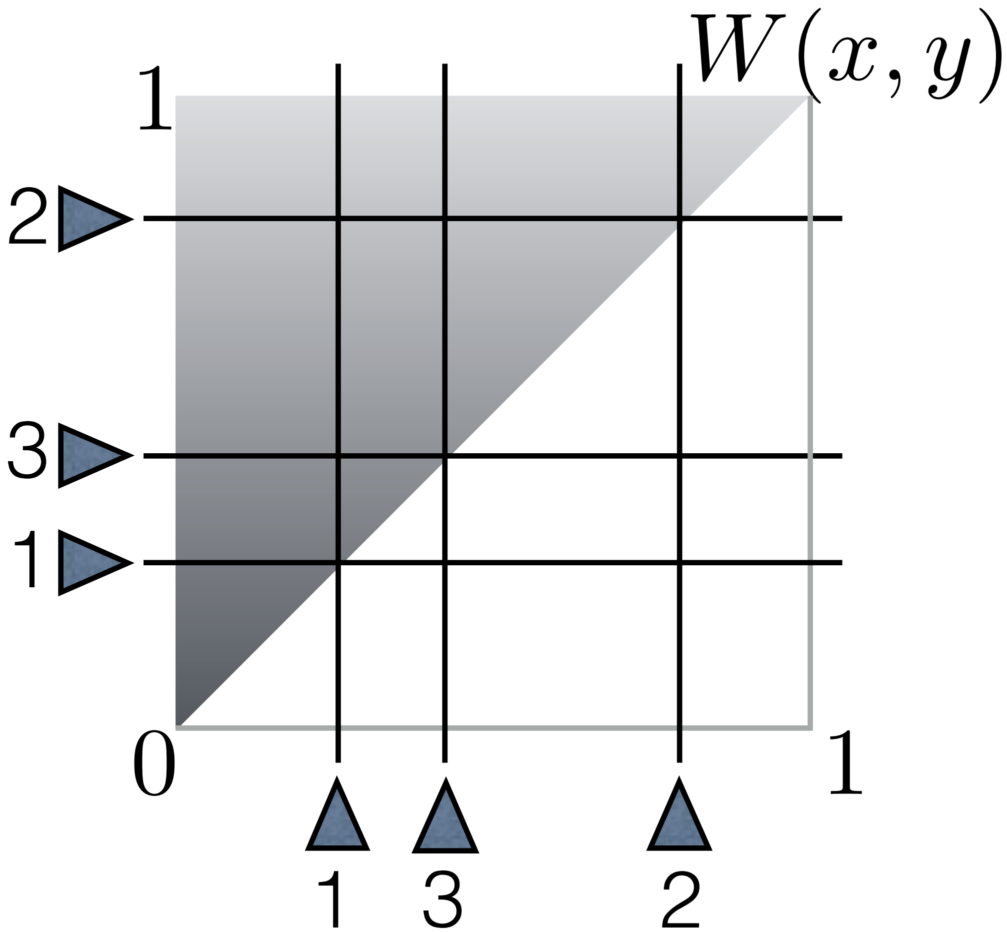

Thm (AH). Every node-exchangeable graph has a *graphon* rep

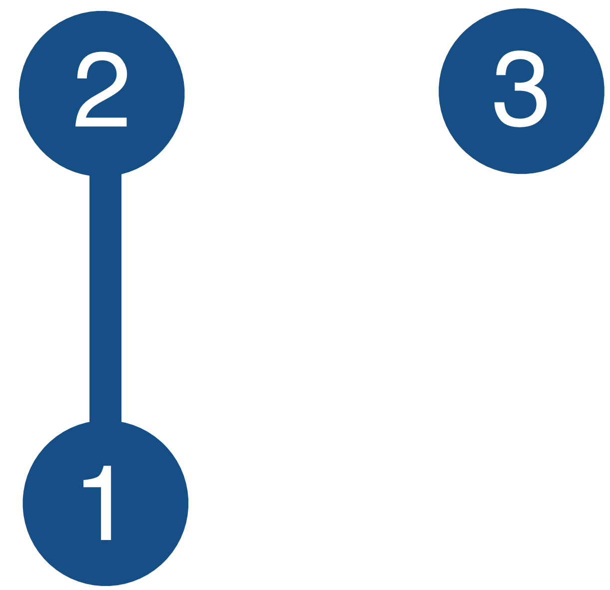
$$\mathbb{E}[\#\text{edges}(G_n)]$$



[Hoover 1979, Aldous 1981, Orbanz, Roy 2015]

# Aldous-Hoover



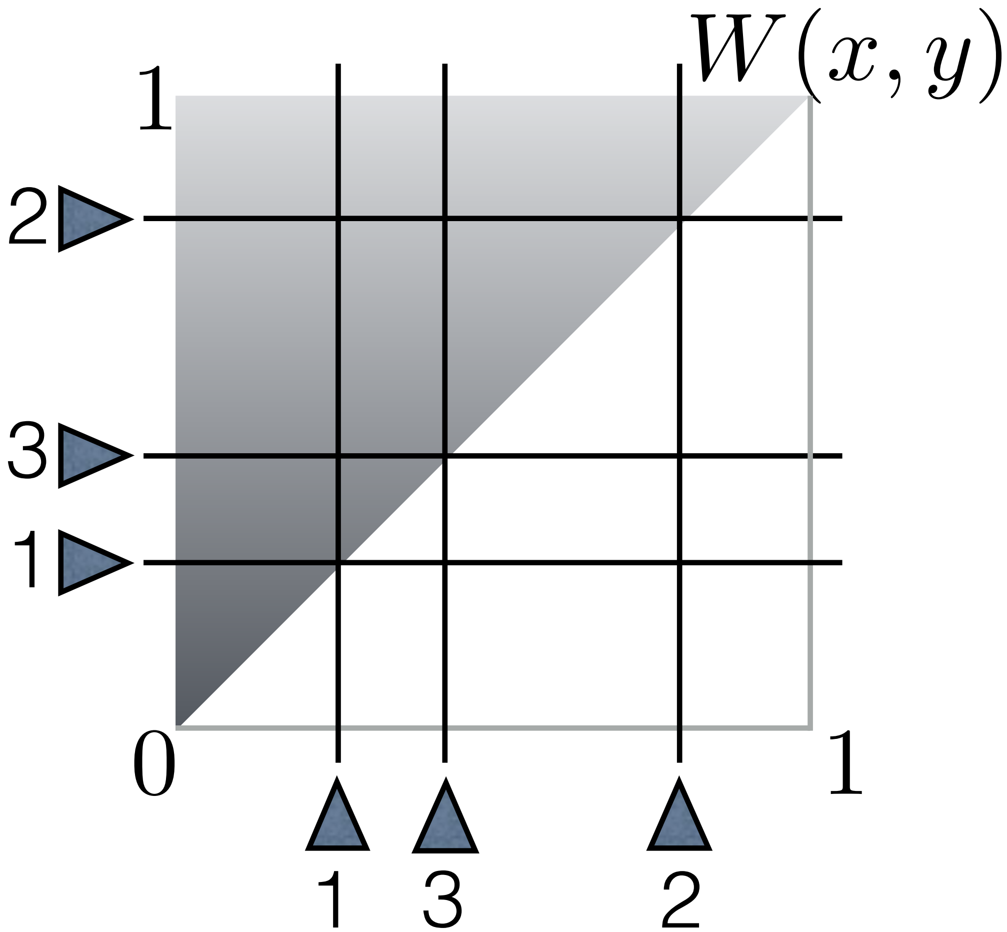

Thm (AH). Every node-exchangeable graph has a *graphon* rep

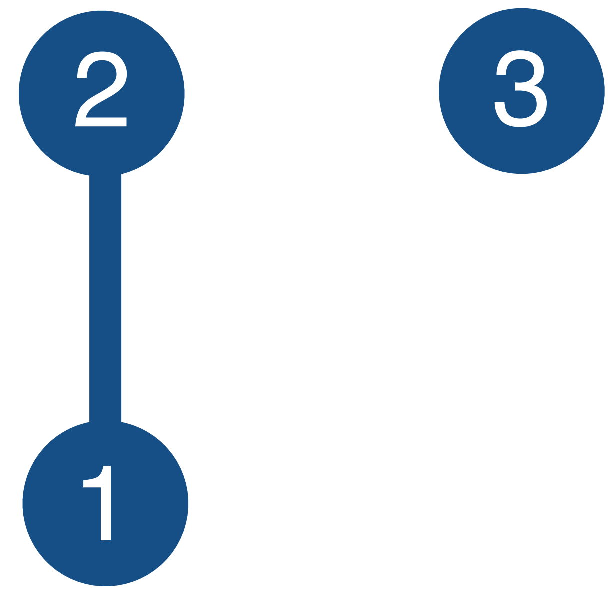
$$\mathbb{E}[\#\text{edges}(G_n)] = \mathbb{E}\left[\binom{n}{2}\frac{1}{2}\int_0^1\int_0^1 W(x,y)\,dx\,dy\right]$$



[Hoover 1979, Aldous 1981, Orbanz, Roy 2015]

# Aldous-Hoover




Thm (AH). Every node-exchangeable graph has a *graphon* rep

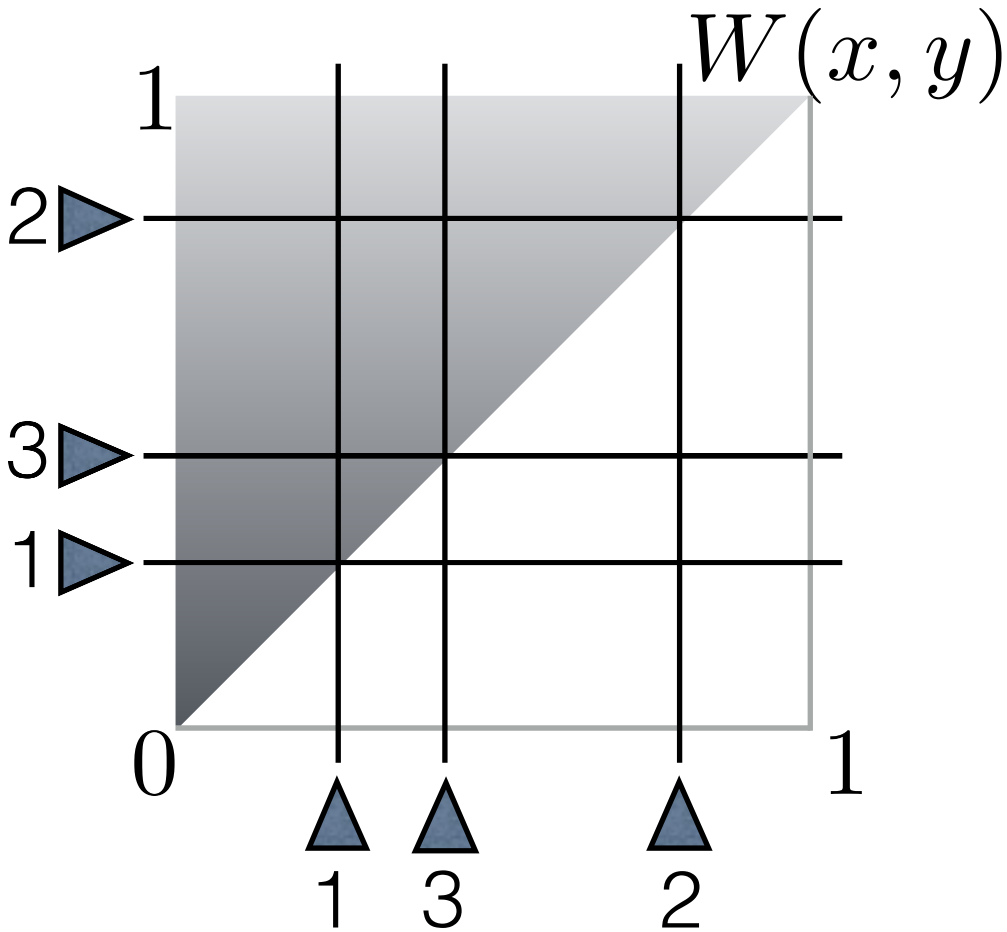
$$\mathbb{E}[\#\text{edges}(G_n)] = \mathbb{E}\left[\binom{n}{2}\frac{1}{2}\int_0^1\int_0^1 W(x,y)\,dx\,dy\right]$$

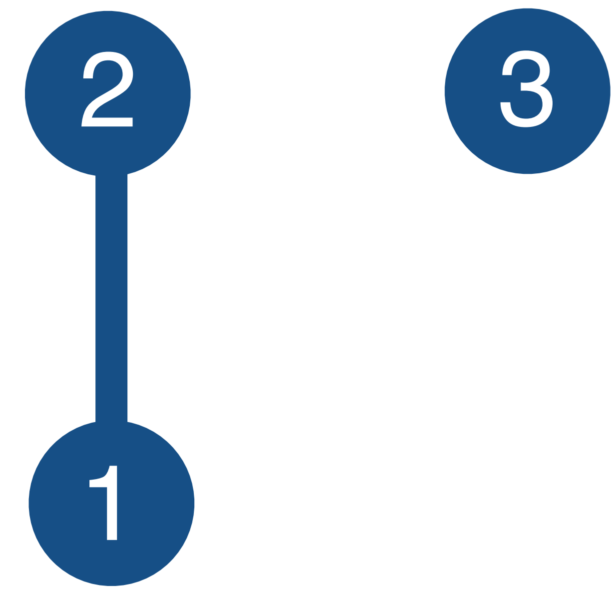
$$\sim cn^2$$



[Hoover 1979, Aldous 1981, Orbanz, Roy 2015]

# Aldous-Hoover




Thm (AH). Every node-exchangeable graph has a *graphon* rep

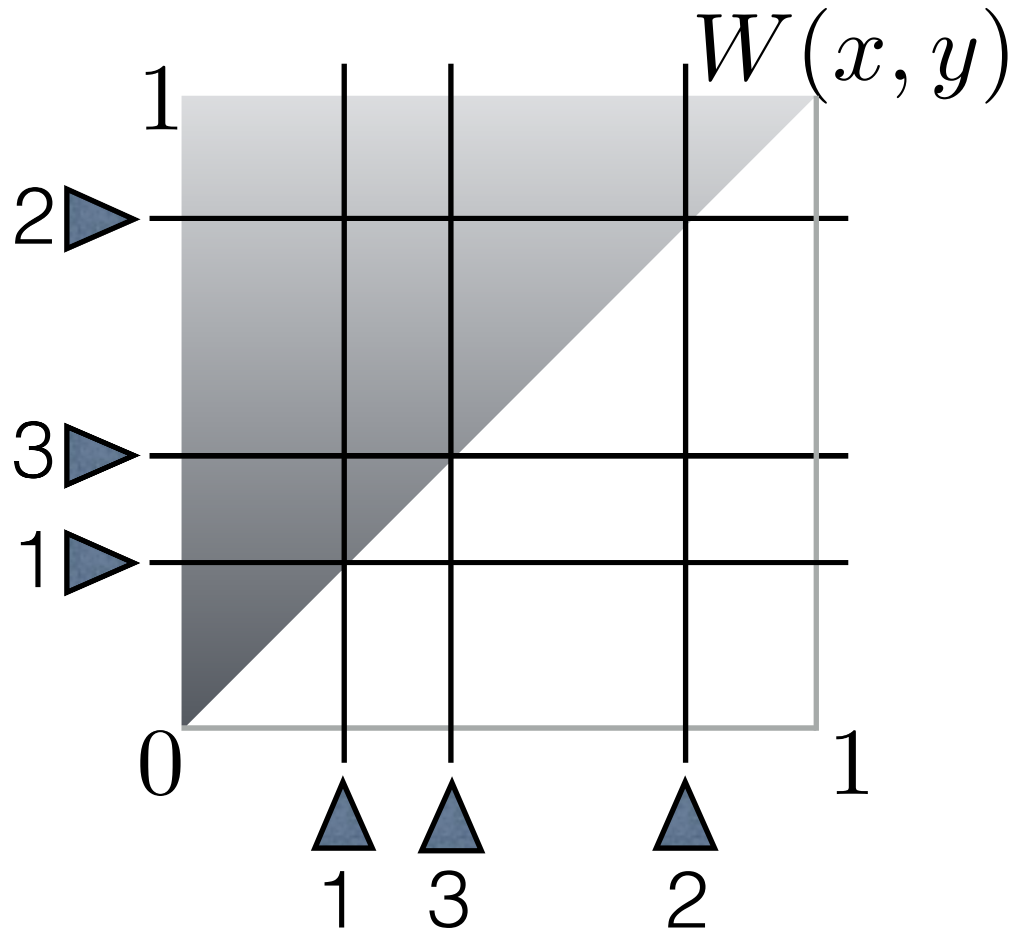
$$\mathbb{E}[\#\text{edges}(G_n)] = \mathbb{E}\left[\binom{n}{2}\frac{1}{2}\int_0^1\int_0^1 W(x,y)\,dx\,dy\right]$$

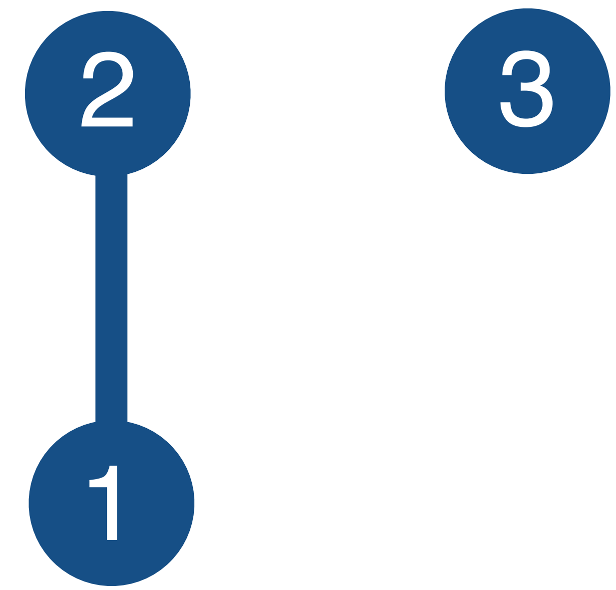
$$\sim cn^2 = c\cdot[\#\text{nodes}(G_n)]^2$$



[Hoover 1979, Aldous 1981, Orbanz, Roy 2015]

# Aldous-Hoover




Thm (AH). Every node-exchangeable graph has a *graphon* rep

$$\mathbb{E}[\#\text{edges}(G_n)] = \mathbb{E}\left[\binom{n}{2}\frac{1}{2}\int_0^1\int_0^1 W(x,y)\,dx\,dy\right]$$

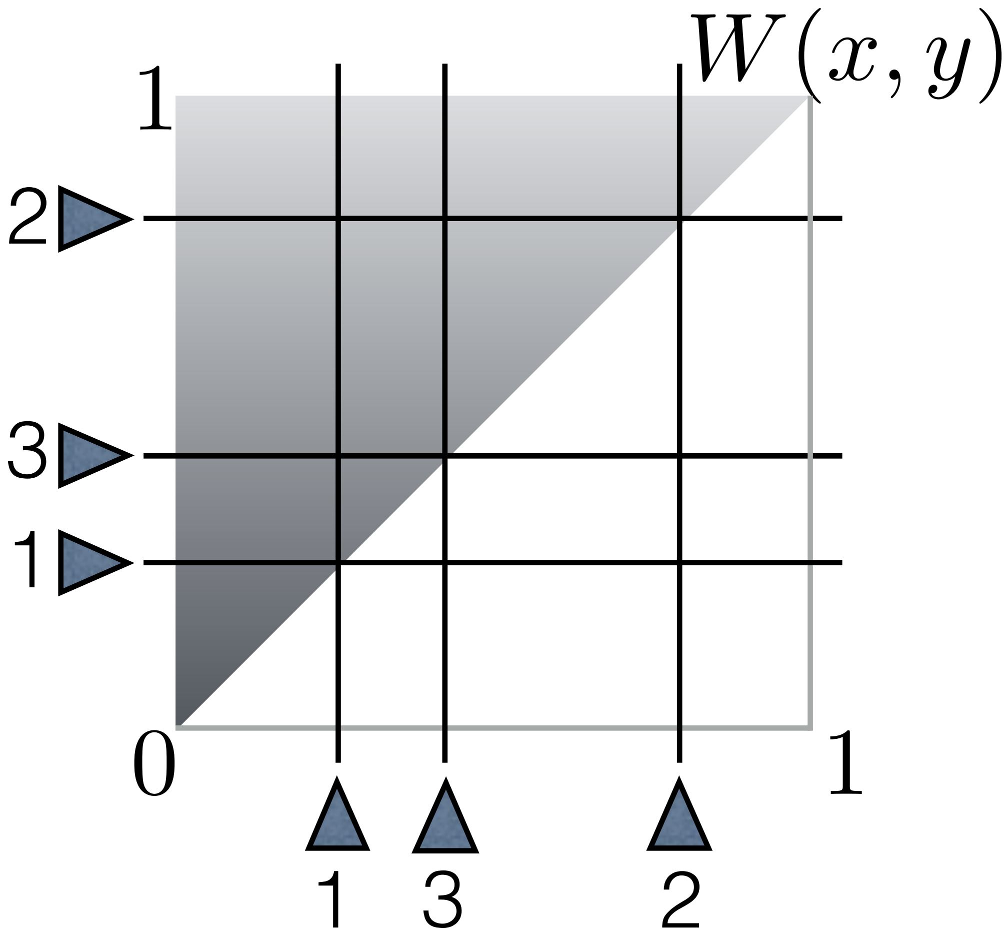
$$\sim cn^2 = c\cdot[\#\text{nodes}(G_n)]^2$$

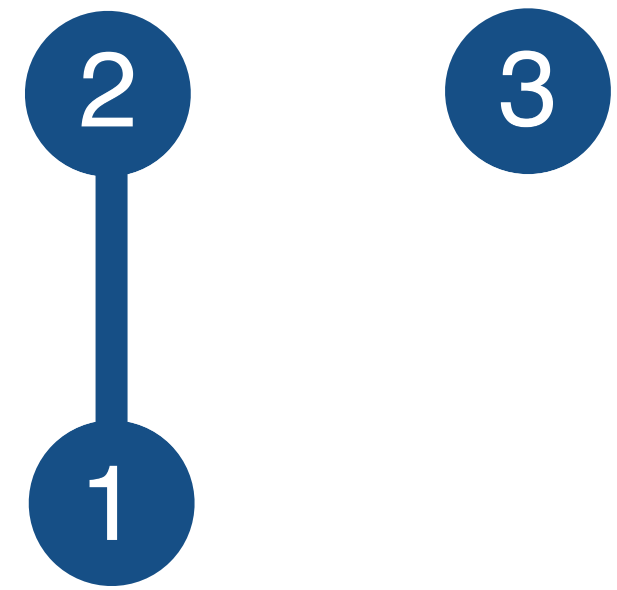
Cor. Every node-exch graph sequence is dense (or empty) a.s.



[Hoover 1979, Aldous 1981, Orbanz, Roy 2015]

# Aldous-Hoover




Thm (AH). Every node-exchangeable graph has a *graphon* rep

$$\mathbb{E}[\#\text{edges}(G_n)] = \mathbb{E}\left[\binom{n}{2}\frac{1}{2}\int_0^1\int_0^1 W(x,y)\,dx\,dy\right]$$

$$\sim cn^2 = c \cdot [\#\text{nodes}(G_n)]^2$$

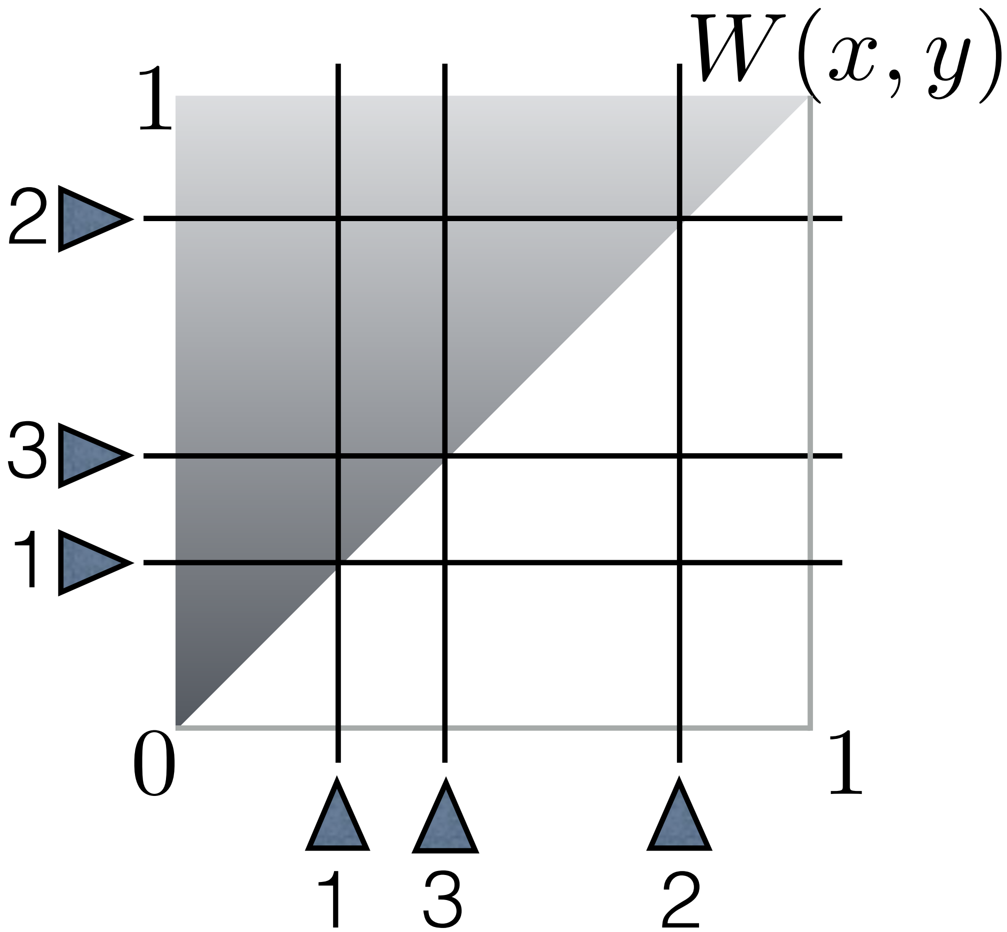
Cor. Every node-exch graph sequence is dense (or empty) a.s.

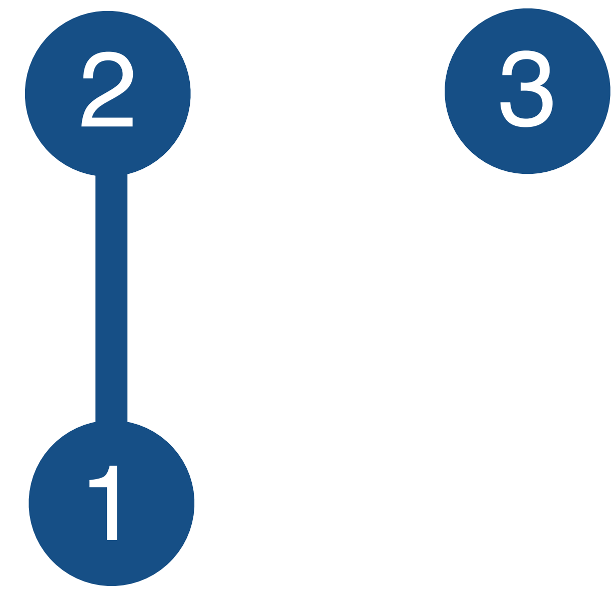
Intuition: To a given node, all other nodes look the same.



[Hoover 1979, Aldous 1981, Orbanz, Roy 2015]

# Aldous-Hoover




Thm (AH). Every node-exchangeable graph has a *graphon* rep

$$\mathbb{E}[\#\text{edges}(G_n)] = \mathbb{E}\left[\binom{n}{2}\frac{1}{2}\int_0^1\int_0^1 W(x,y)\,dx\,dy\right]$$

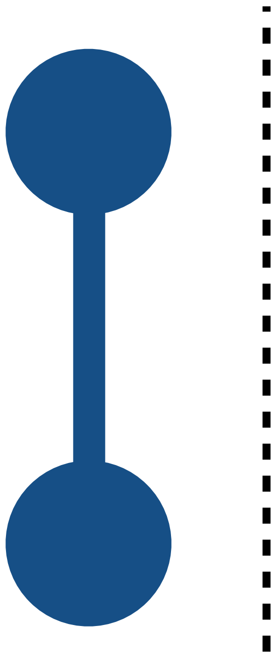
$$\sim cn^2 = c\cdot[\#\text{nodes}(G_n)]^2$$

Cor. Every node-exch graph sequence is dense (or empty) a.s.

Intuition: To a given node, all other nodes look the same.

[Caron, Fox 2014; Veitch, Roy 2015; Borgs, Chayes, Cohn, Holden 2016; Broderick, Cai 2015; Cai, Broderick 2015a,b; Crane, Dempsey 2015a,b, 2016a; Cai, Campbell, Broderick 2016]

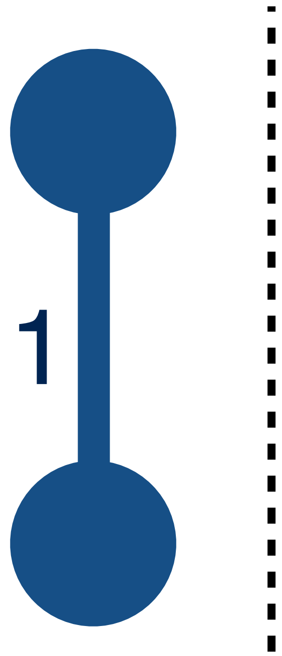
# A New Way: Edges



$G_1$

 [Broderick, Cai 2015; Cai, Broderick 2015a,b; Crane, Dempsey 2015a,b,16a; Cai, Campbell, Broderick 2016]

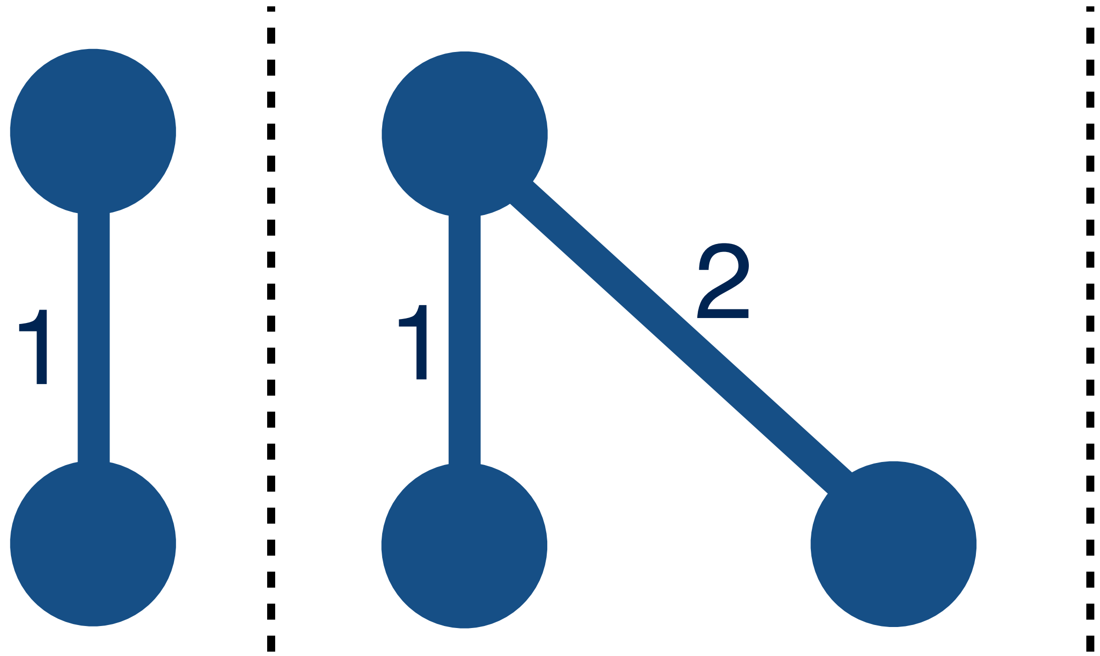
# A New Way: Edges




$G_1$

 [Broderick, Cai 2015; Cai, Broderick 2015a,b; Crane, Dempsey 2015a,b,16a; Cai, Campbell, Broderick 2016]

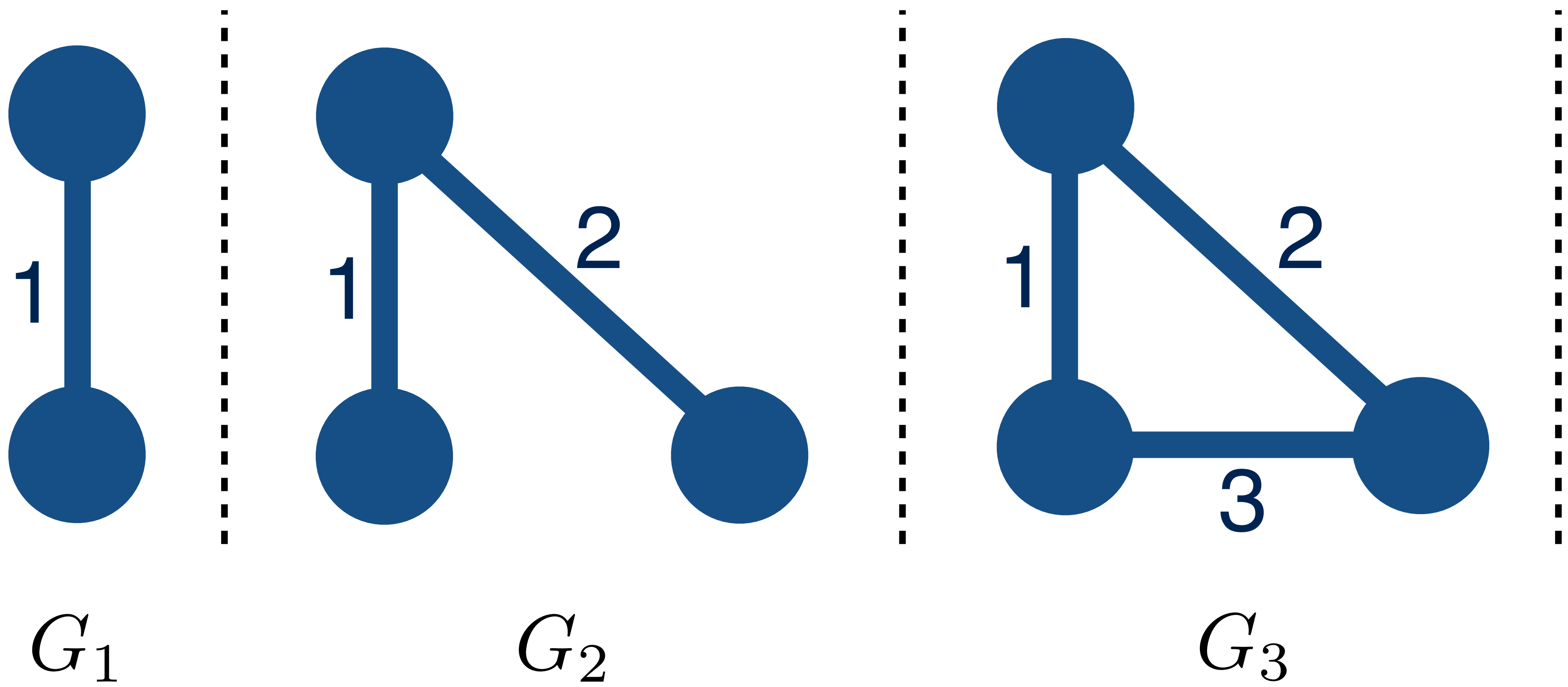
# A New Way: Edges




$G_1$ $G_2$

 [Broderick, Cai 2015; Cai, Broderick 2015a,b; Crane, Dempsey 2015a,b,16a; Cai, Campbell, Broderick 2016]

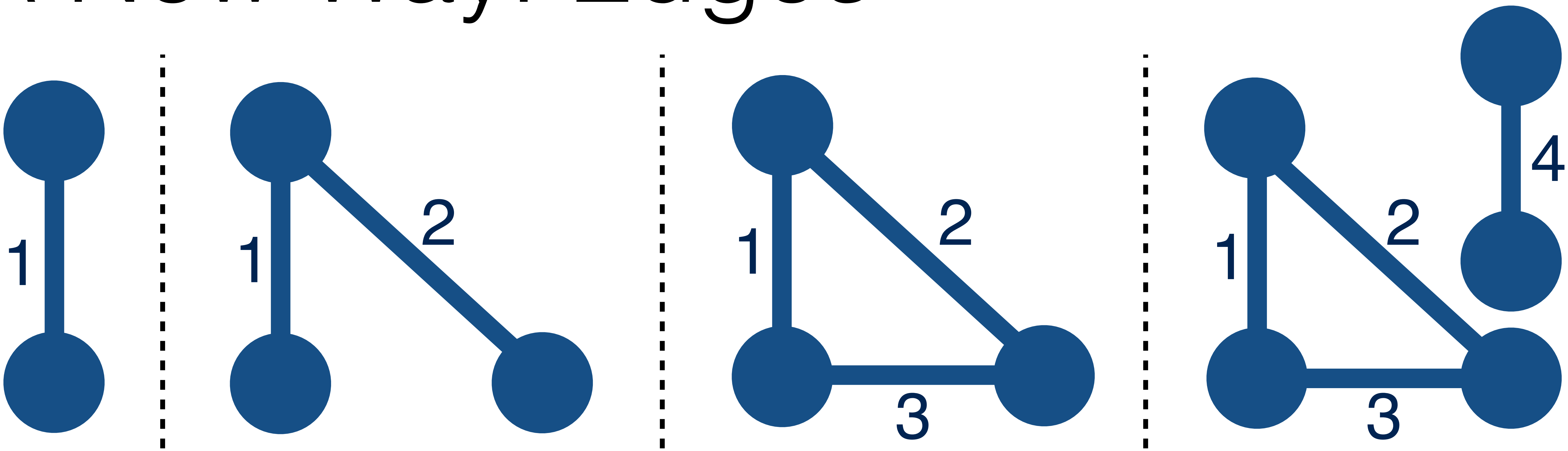
# A New Way: Edges






[Broderick, Cai 2015; Cai, Broderick 2015a,b; Crane, Dempsey 2015a,b,16a; Cai, Campbell, Broderick 2016]

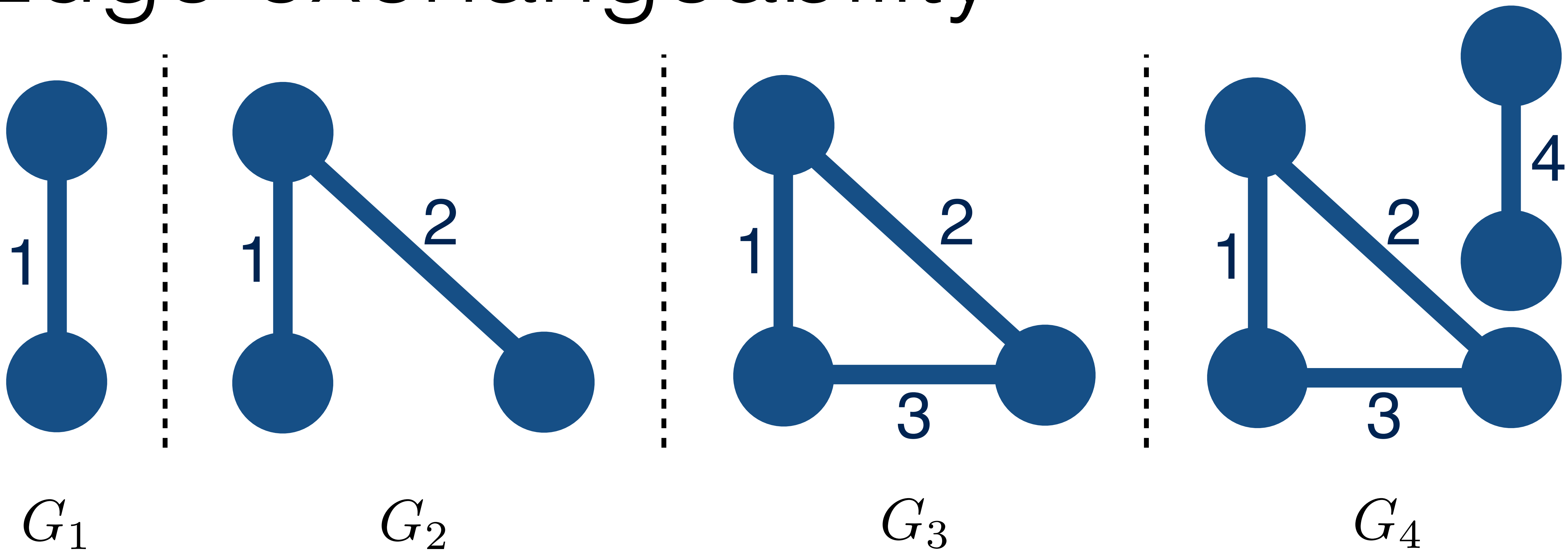
# A New Way: Edges




$G_1$ $G_2$ $G_3$ $G_4$

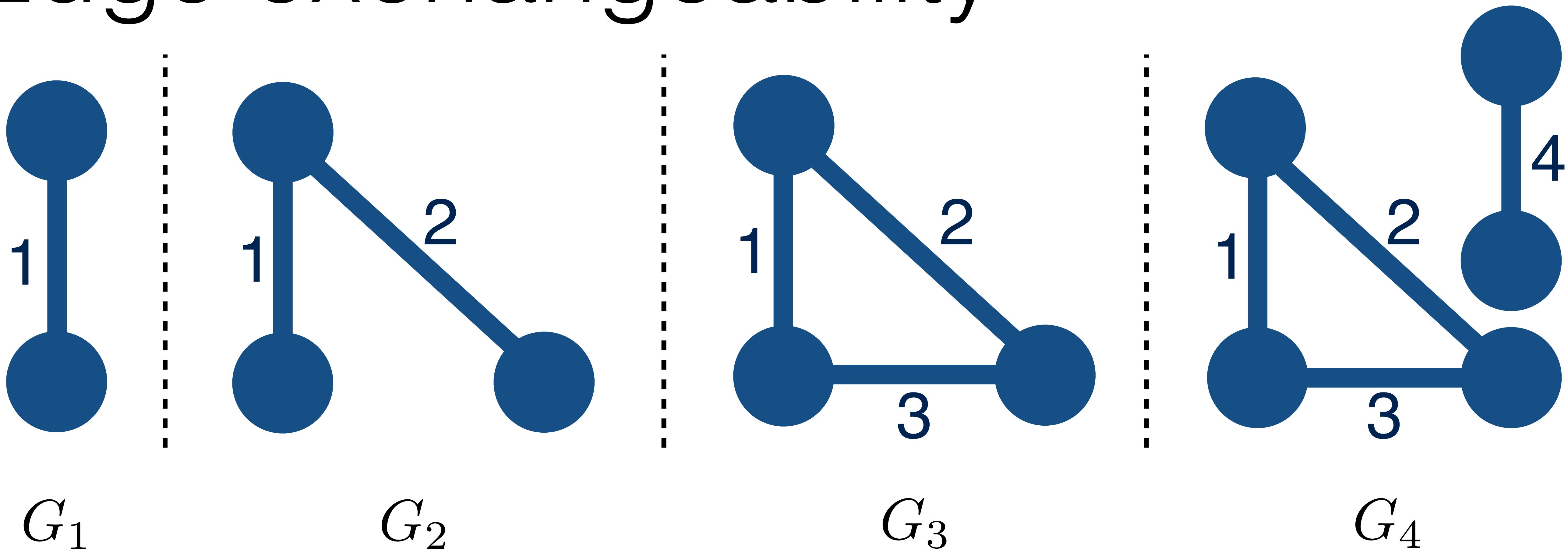

[Broderick, Cai 2015; Cai, Broderick 2015a,b; Crane, Dempsey 2015a,b,16a; Cai, Campbell, Broderick 2016]

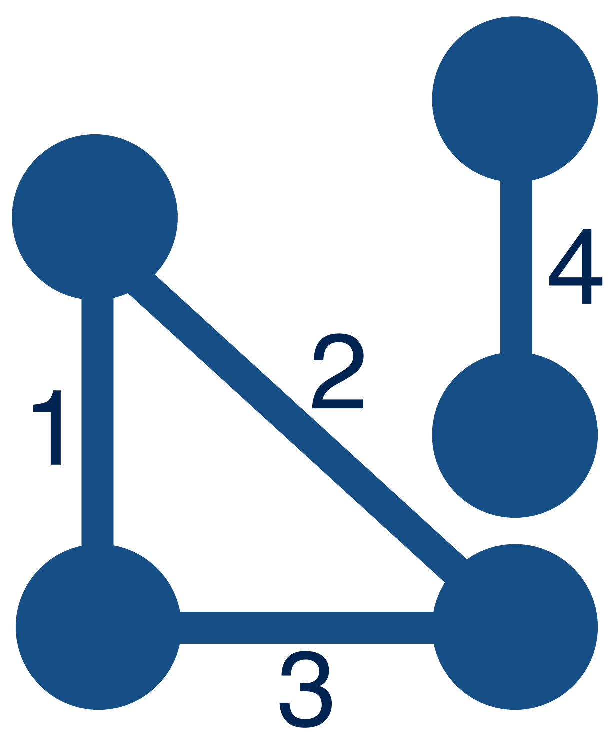
# Edge exchangeability




 [Broderick, Cai 2015; Cai, Broderick 2015a,b; Crane, Dempsey 2015a,b,16a; Cai, Campbell, Broderick 2016]

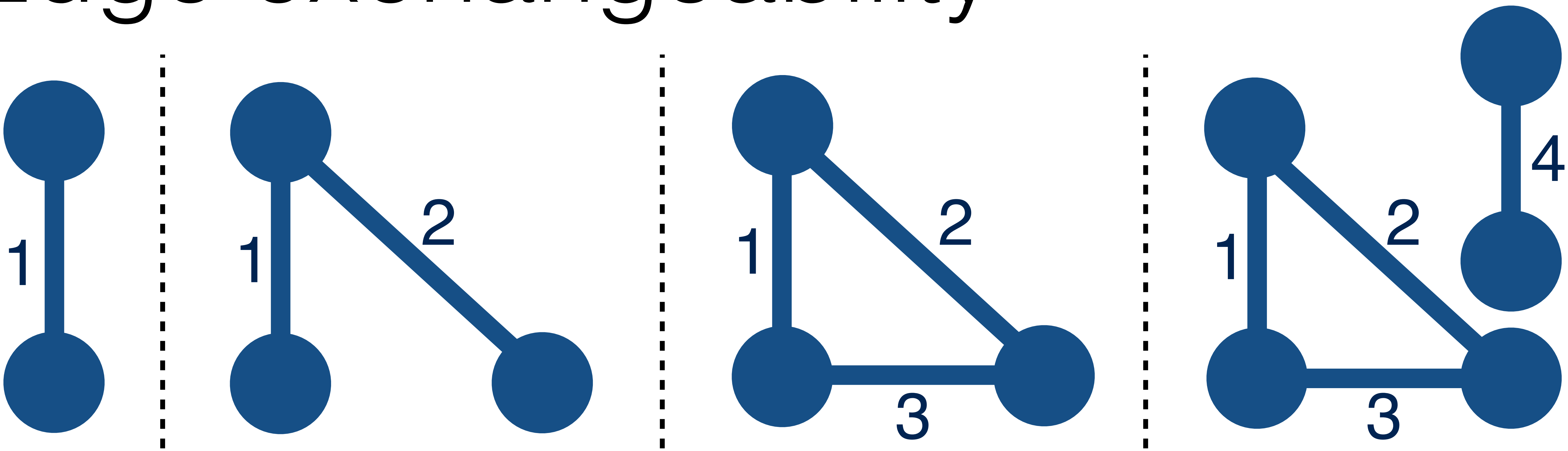
# Edge exchangeability



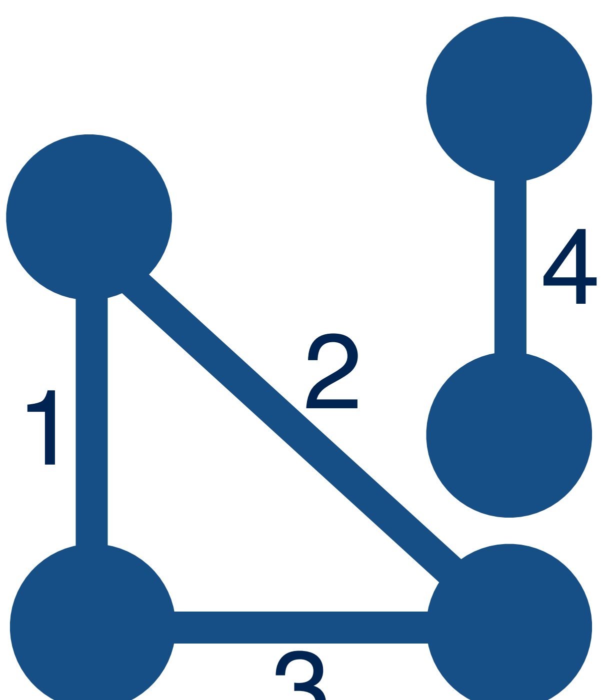




 [Broderick, Cai 2015; Cai, Broderick 2015a,b; Crane, Dempsey 2015a,b,16a; Cai, Campbell, Broderick 2016]

# Edge exchangeability

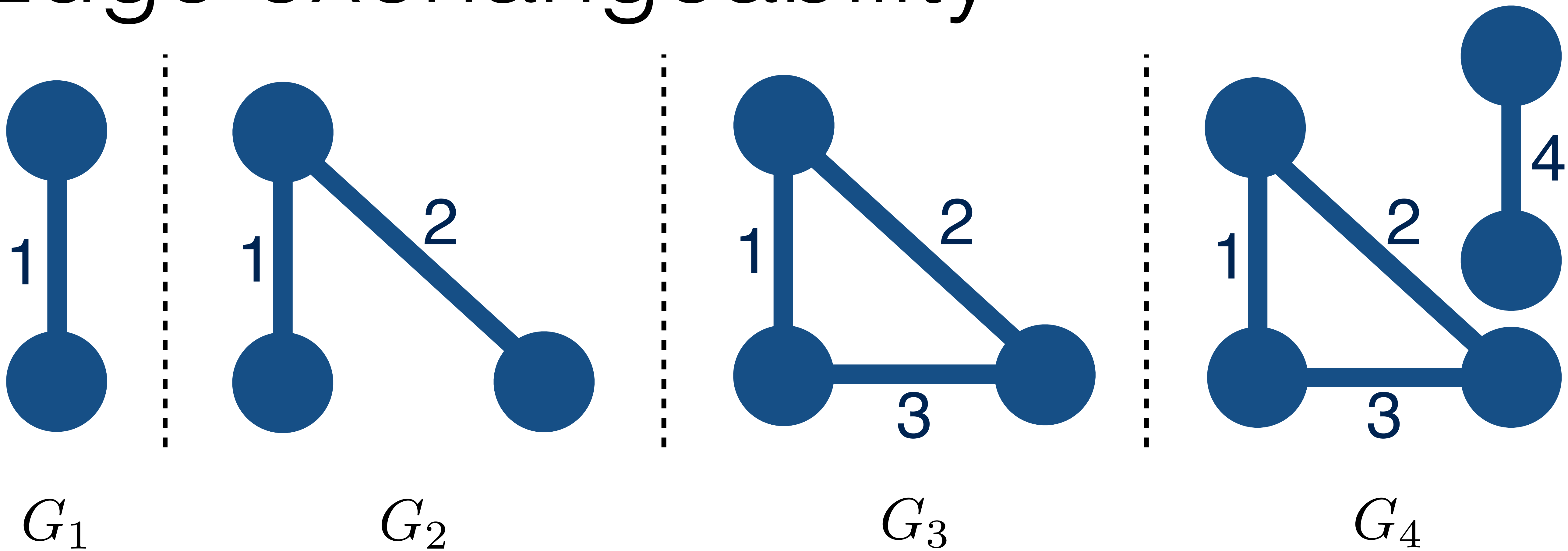


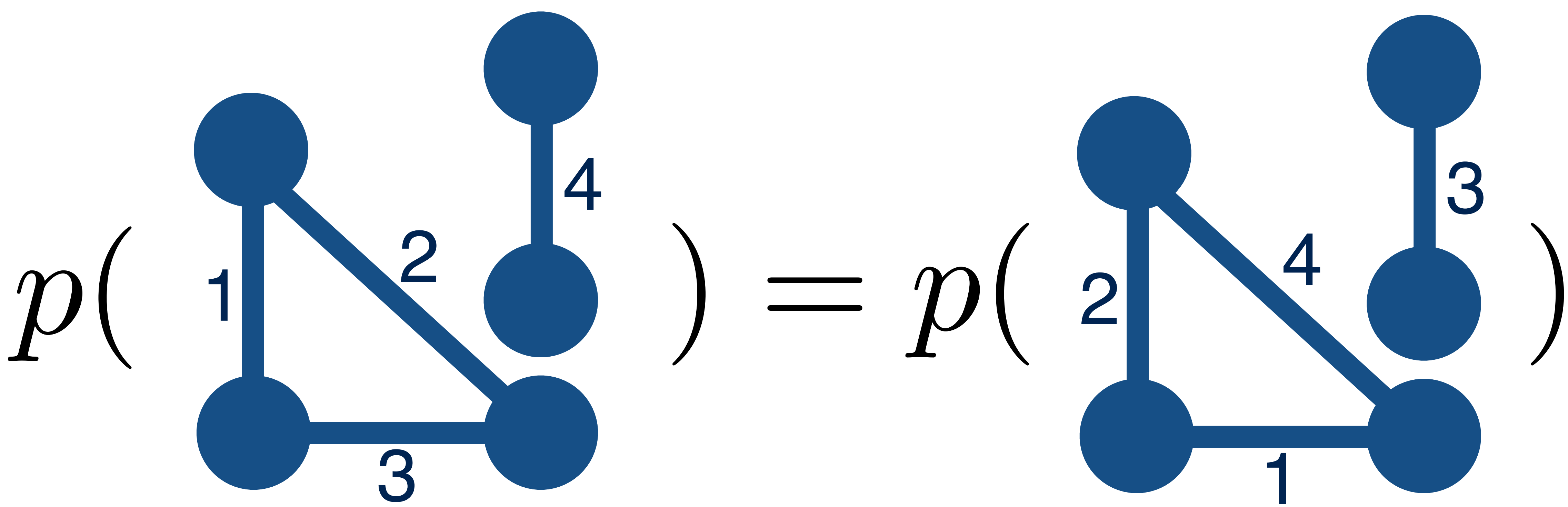

$G_1$ $G_2$ $G_3$ $G_4$

$p($  $)$


 [Broderick, Cai 2015; Cai, Broderick 2015a,b; Crane, Dempsey 2015a,b,16a; Cai, Campbell, Broderick 2016]

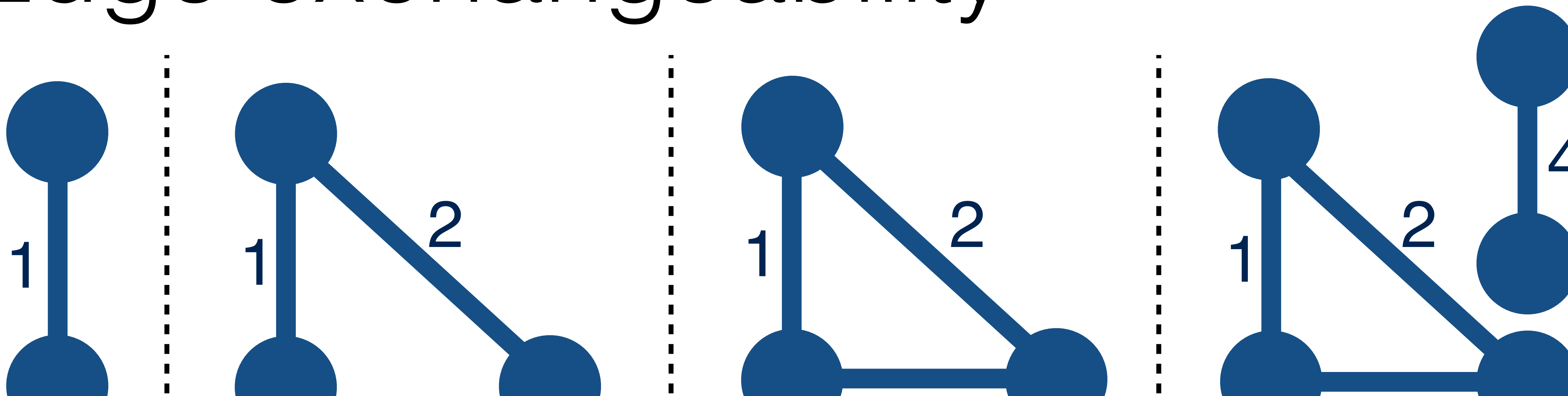
# Edge exchangeability



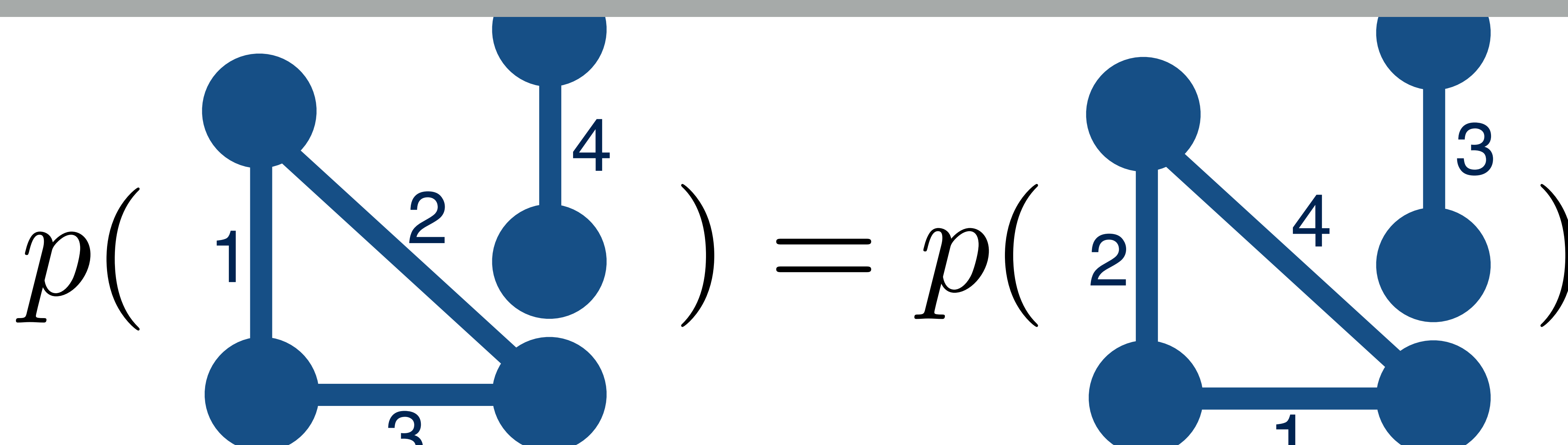




 [Broderick, Cai 2015; Cai, Broderick 2015a,b; Crane, Dempsey 2015a,b,16a; Cai, Campbell, Broderick 2016]

# Edge exchangeability

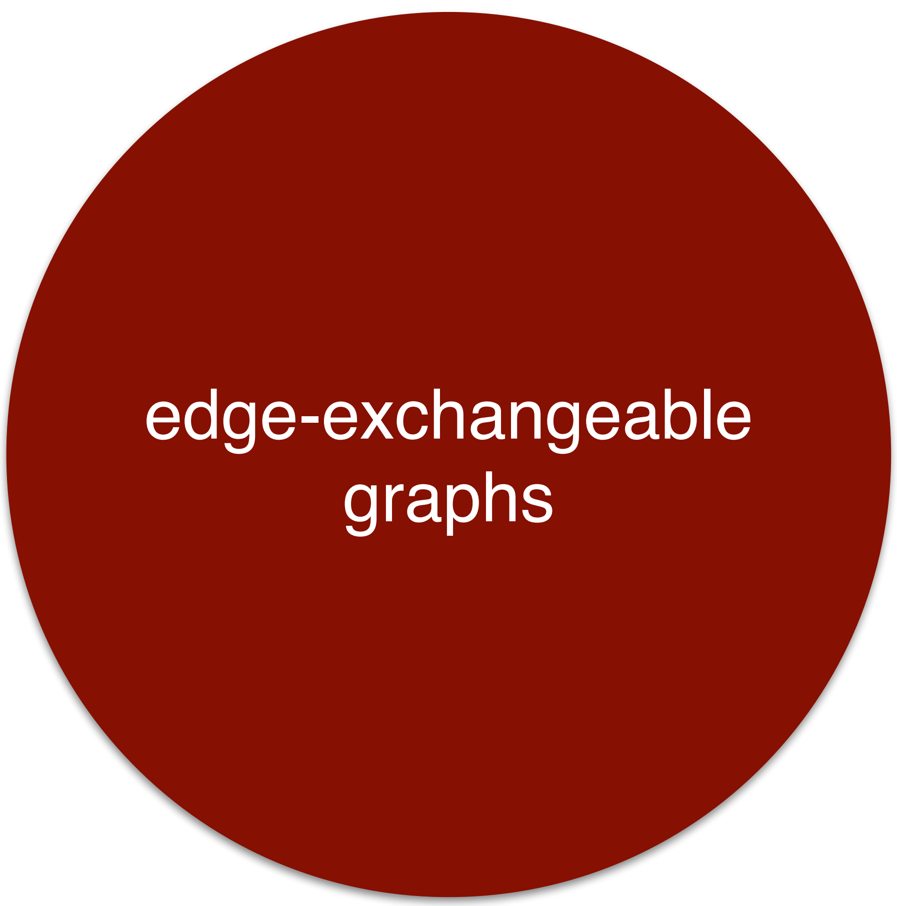



**Thm (CCB). A wide class of edge-exchangeable graph models yields sparse graph sequences.**



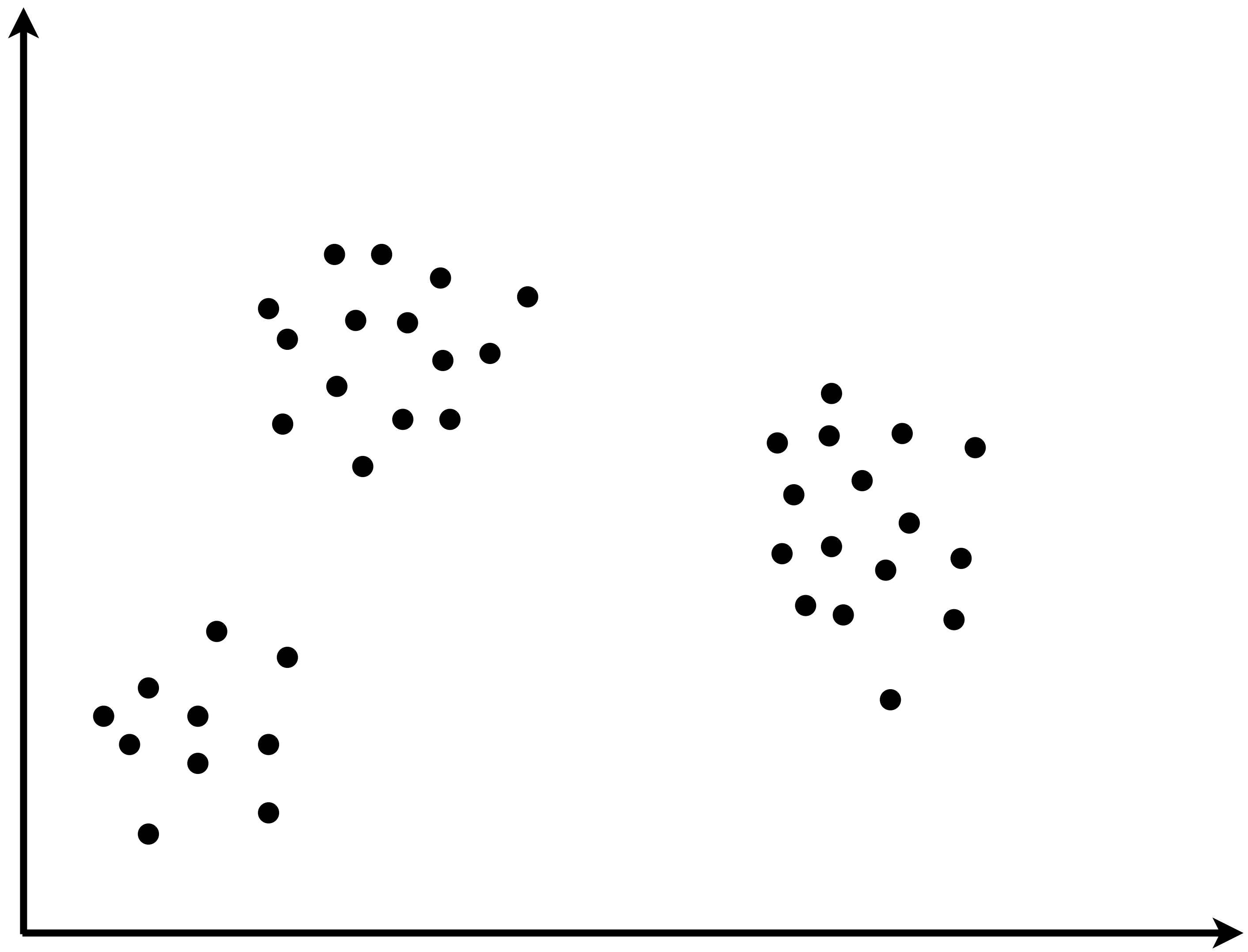

 [Broderick, Cai 2015; Cai, Broderick 2015a,b; Crane, Dempsey 2015a,b,16a; Cai, Campbell, Broderick 2016]

# What we know so far

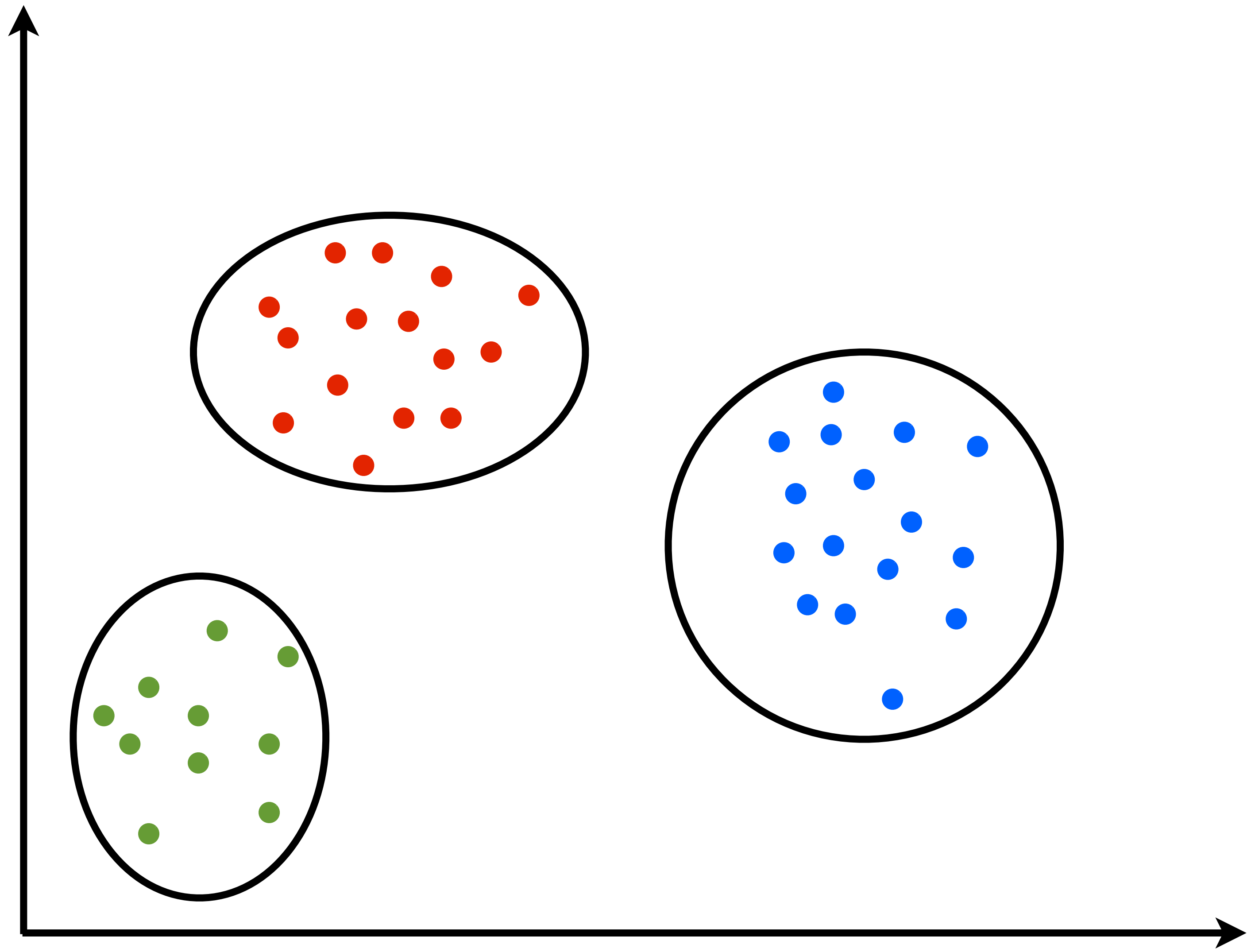



- Goal 1: characterization theorem for edge-exchangeable graphs
- Goal 2: sparsity theorem for edge-exchangeable graphs

# Clustering

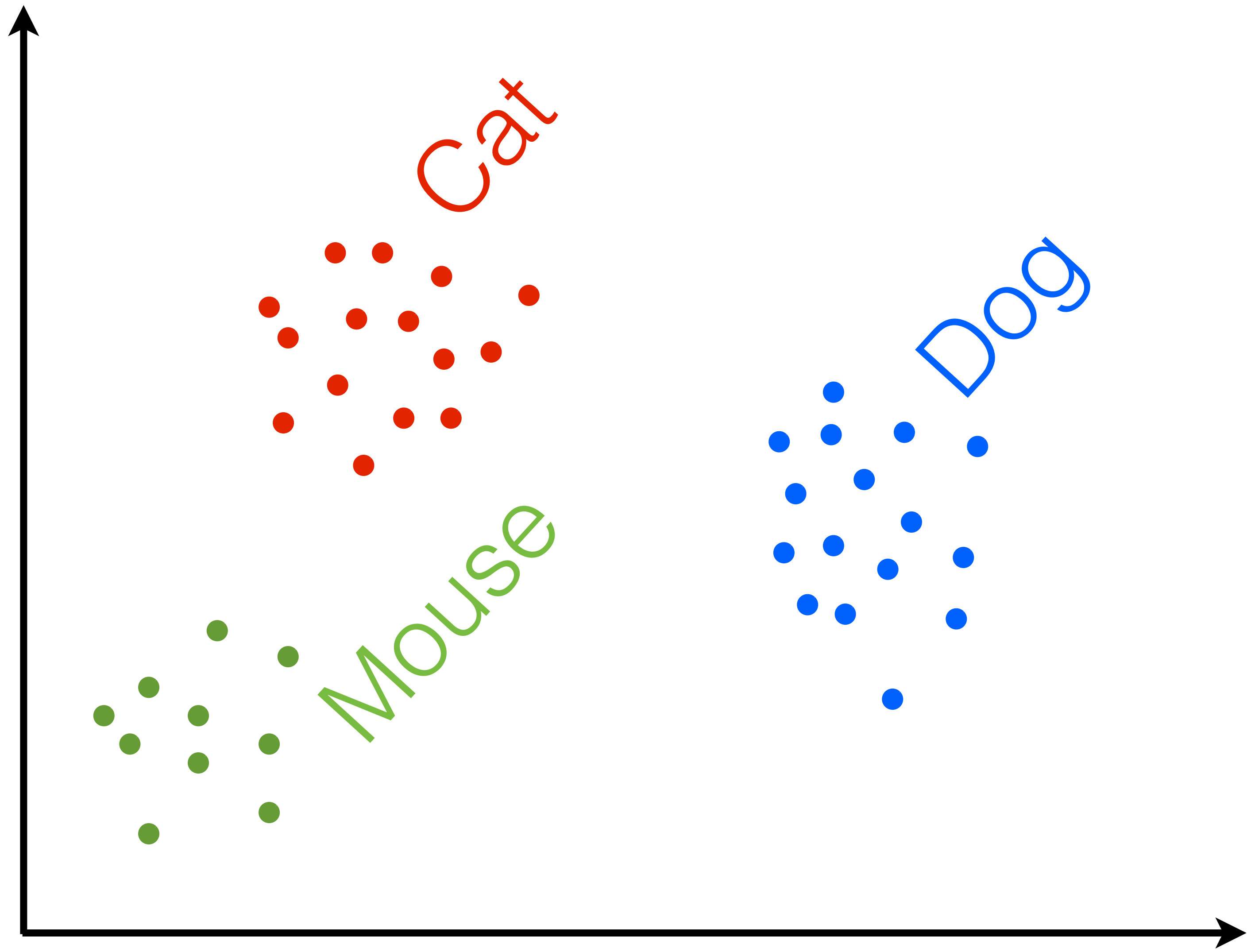


# Clustering



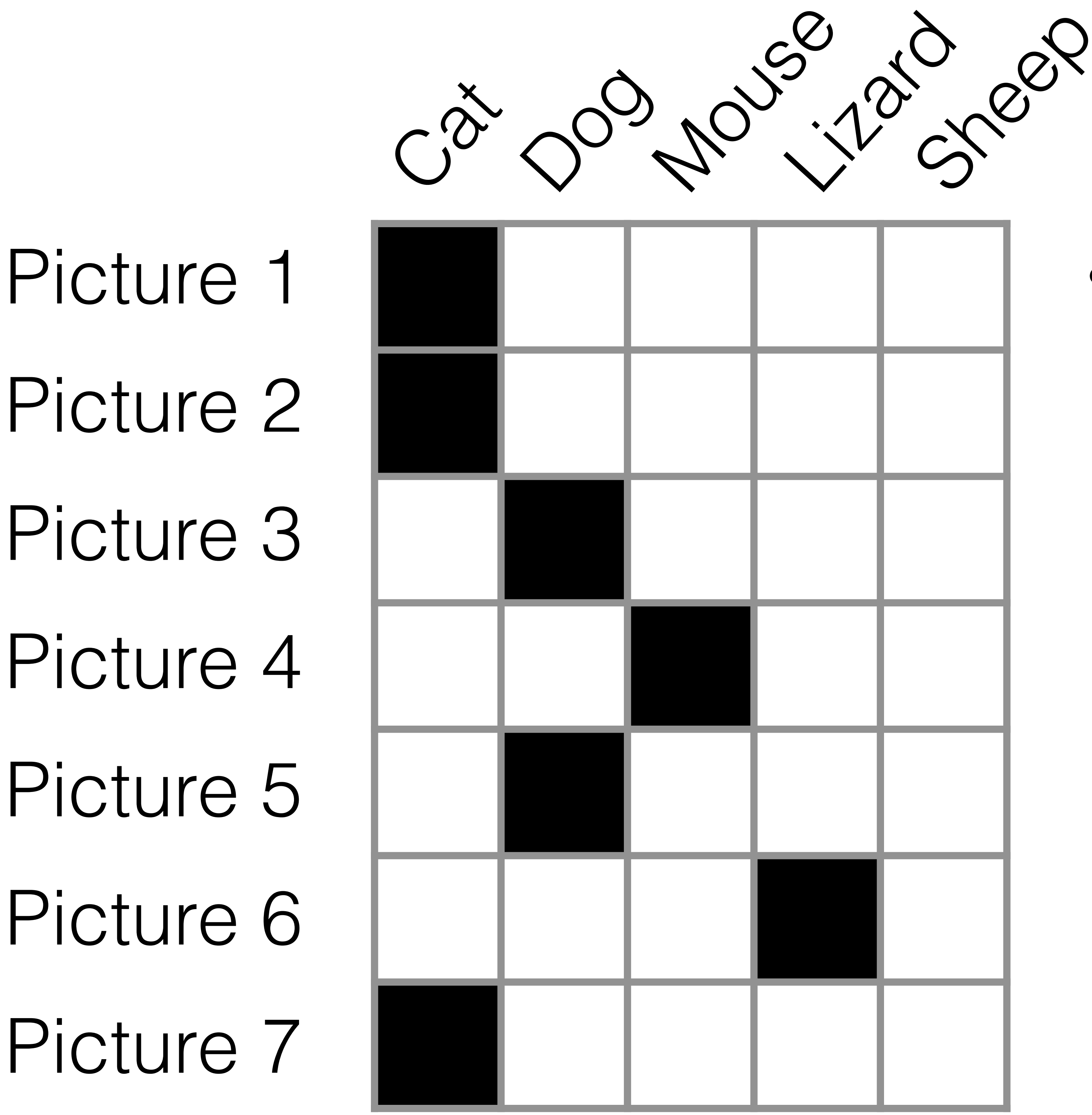
"Clusters"

# Clustering



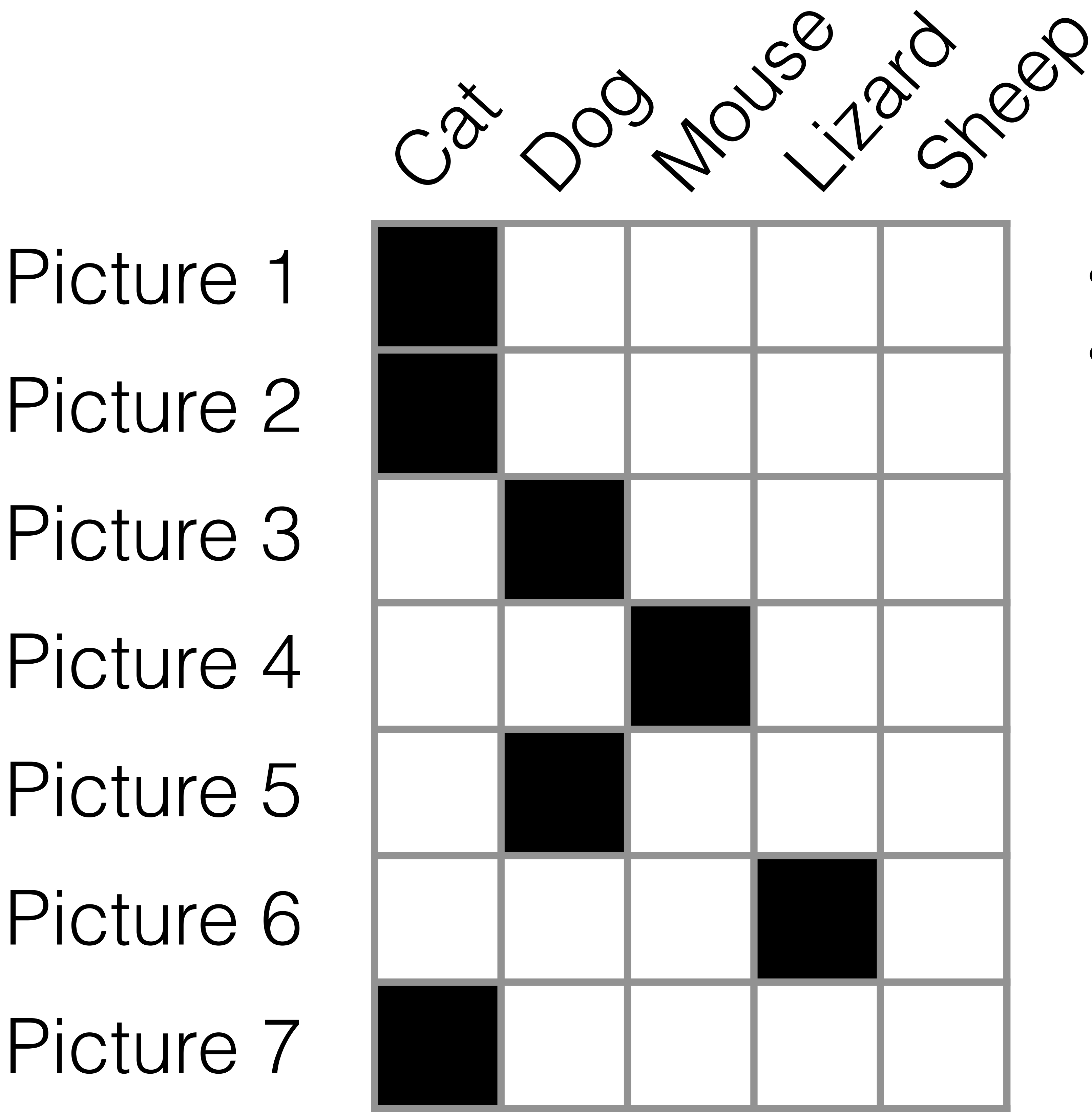

“Clusters”

# Clustering



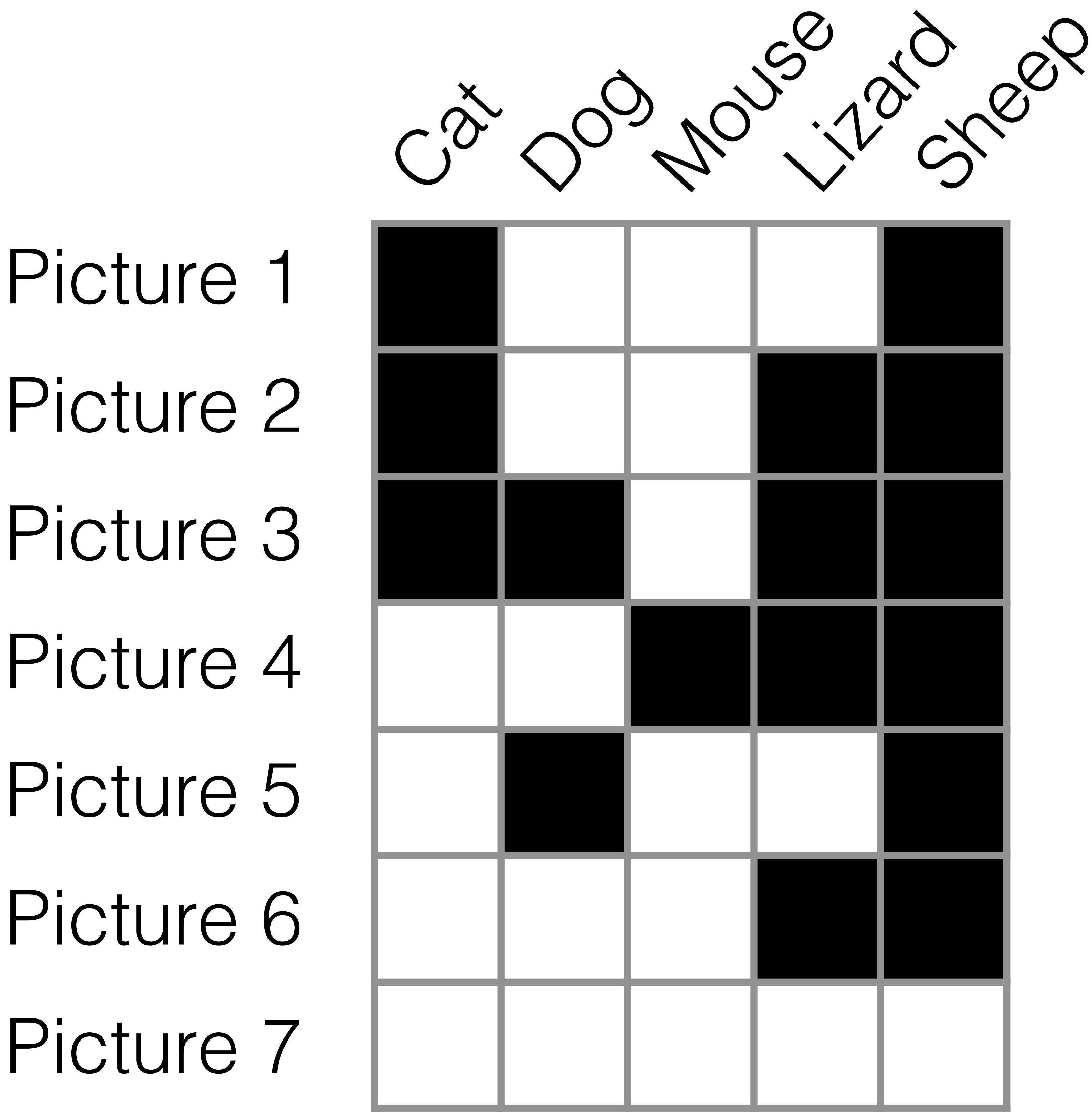

- Groups: clusters

# Clustering

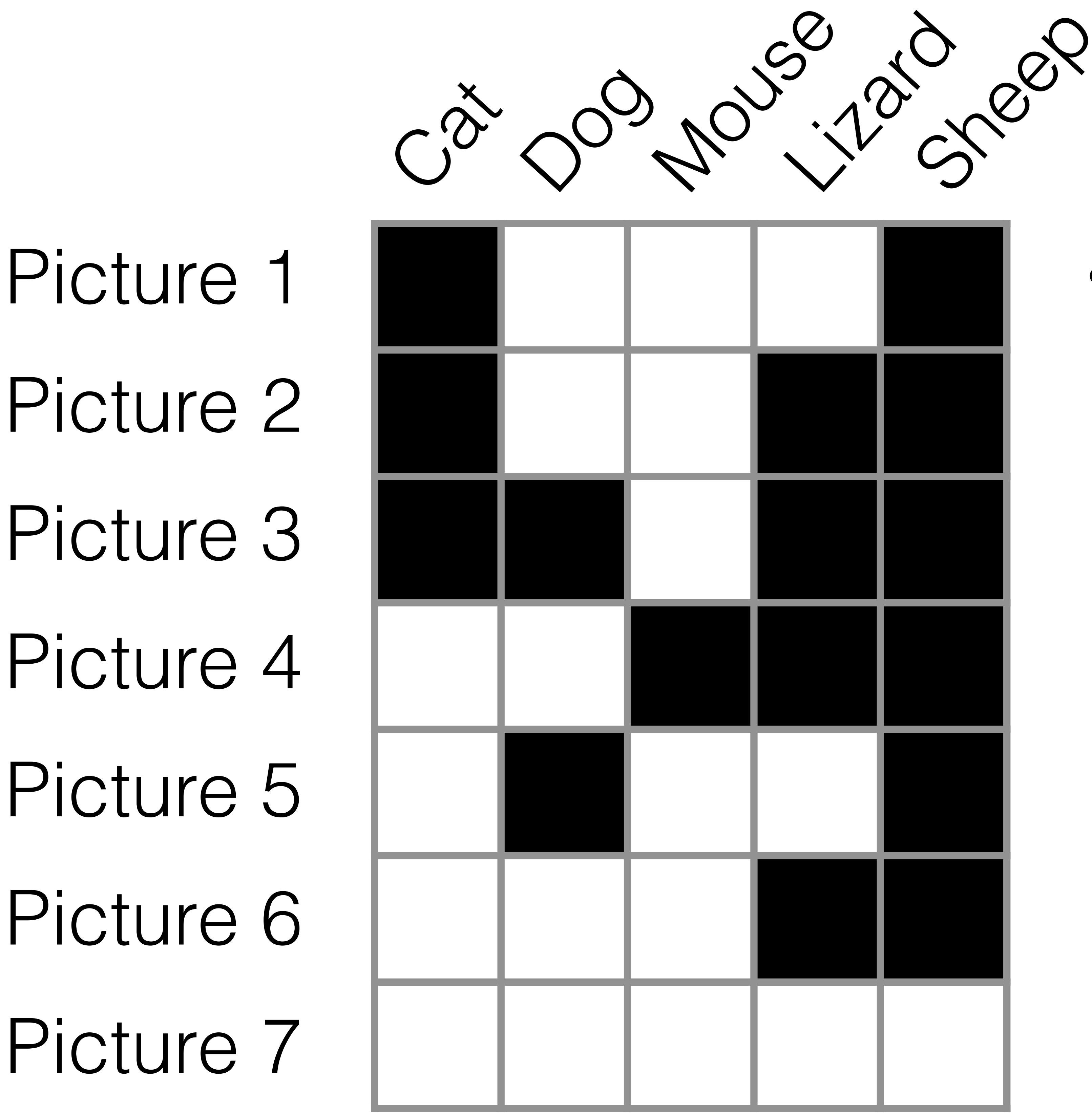



- Groups: clusters
- Exchangeable

# Feature allocation

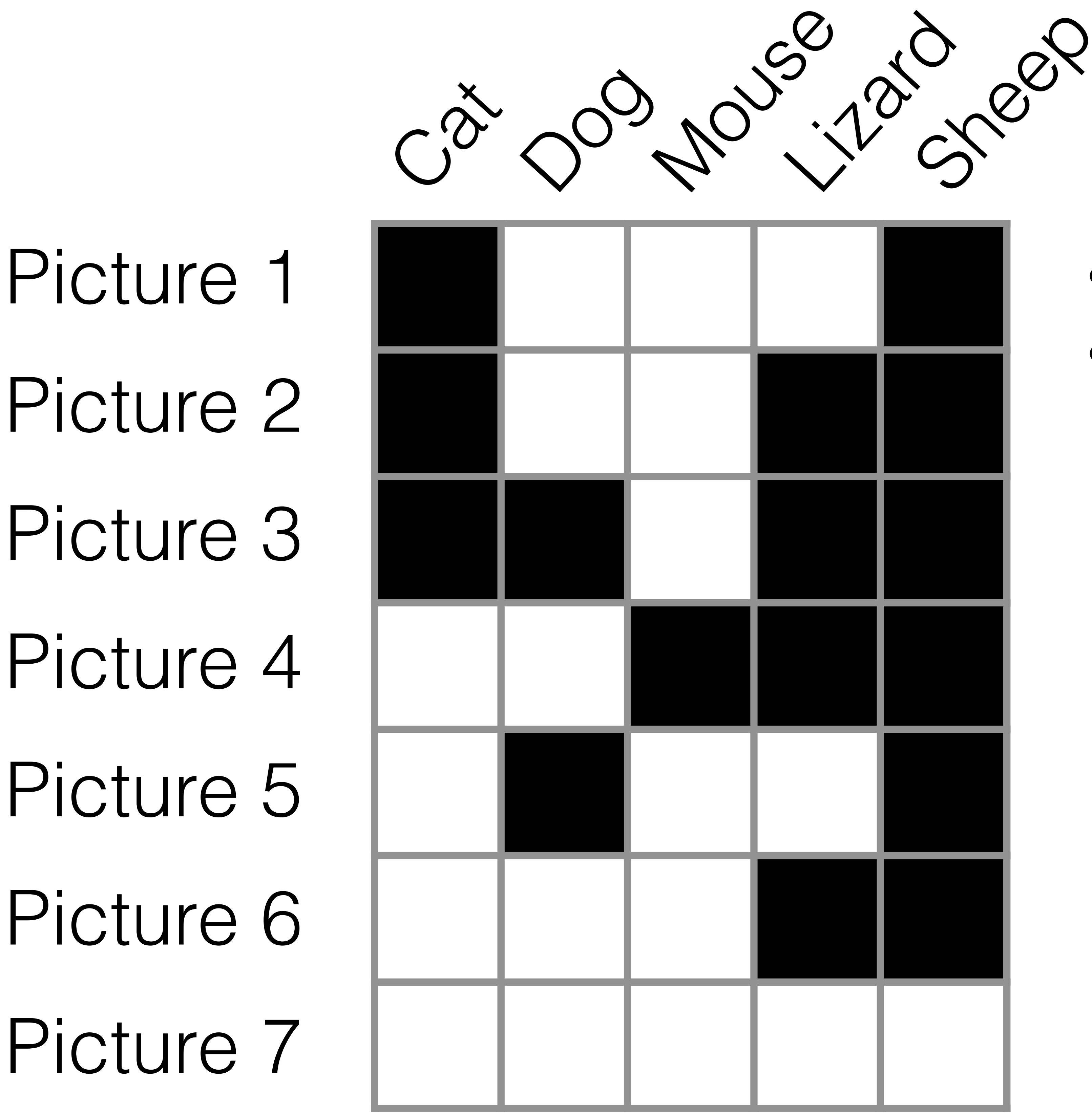


# Feature allocation




- Groups: features

# Feature allocation

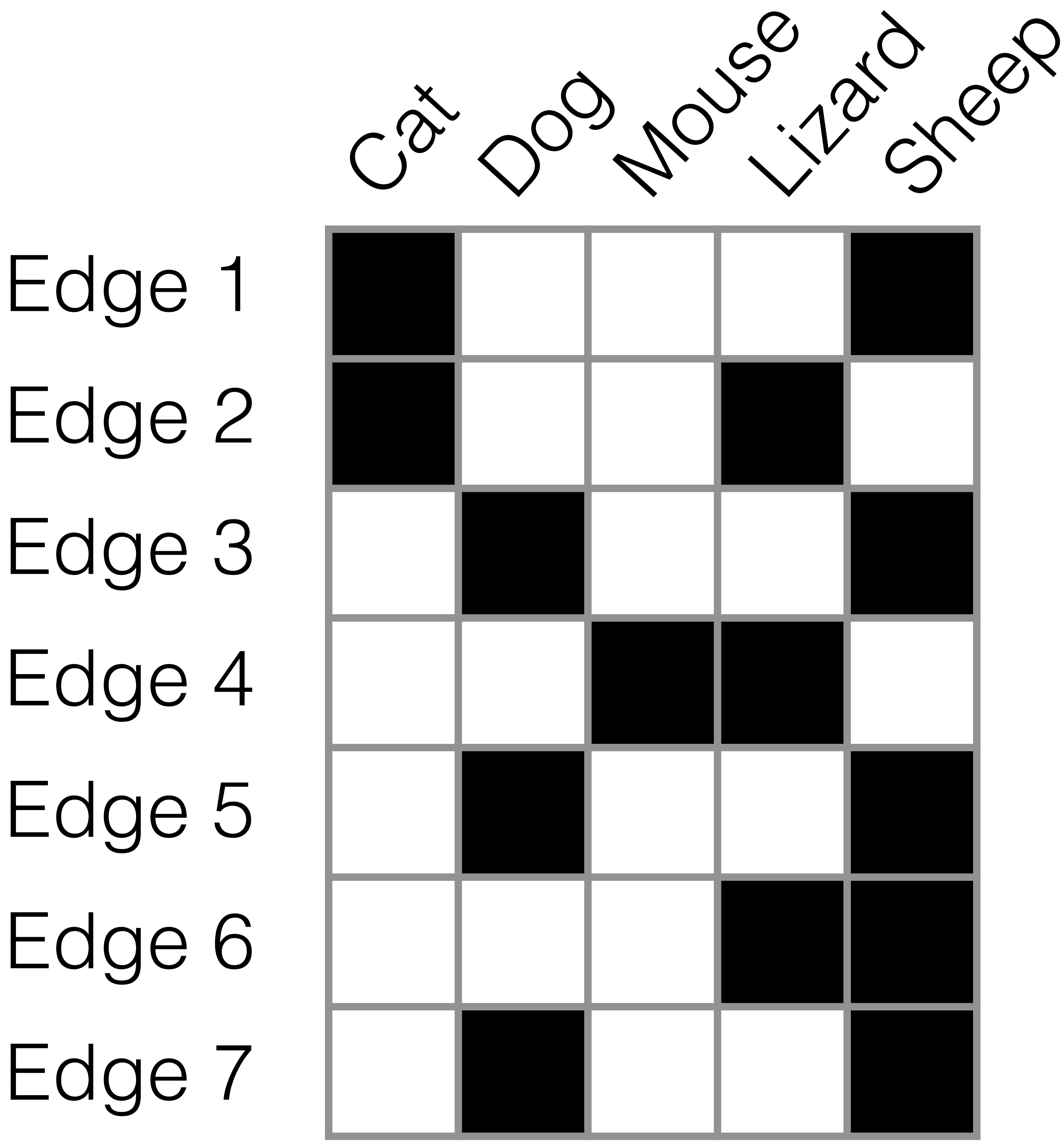



- Groups: features
- Exchangeable

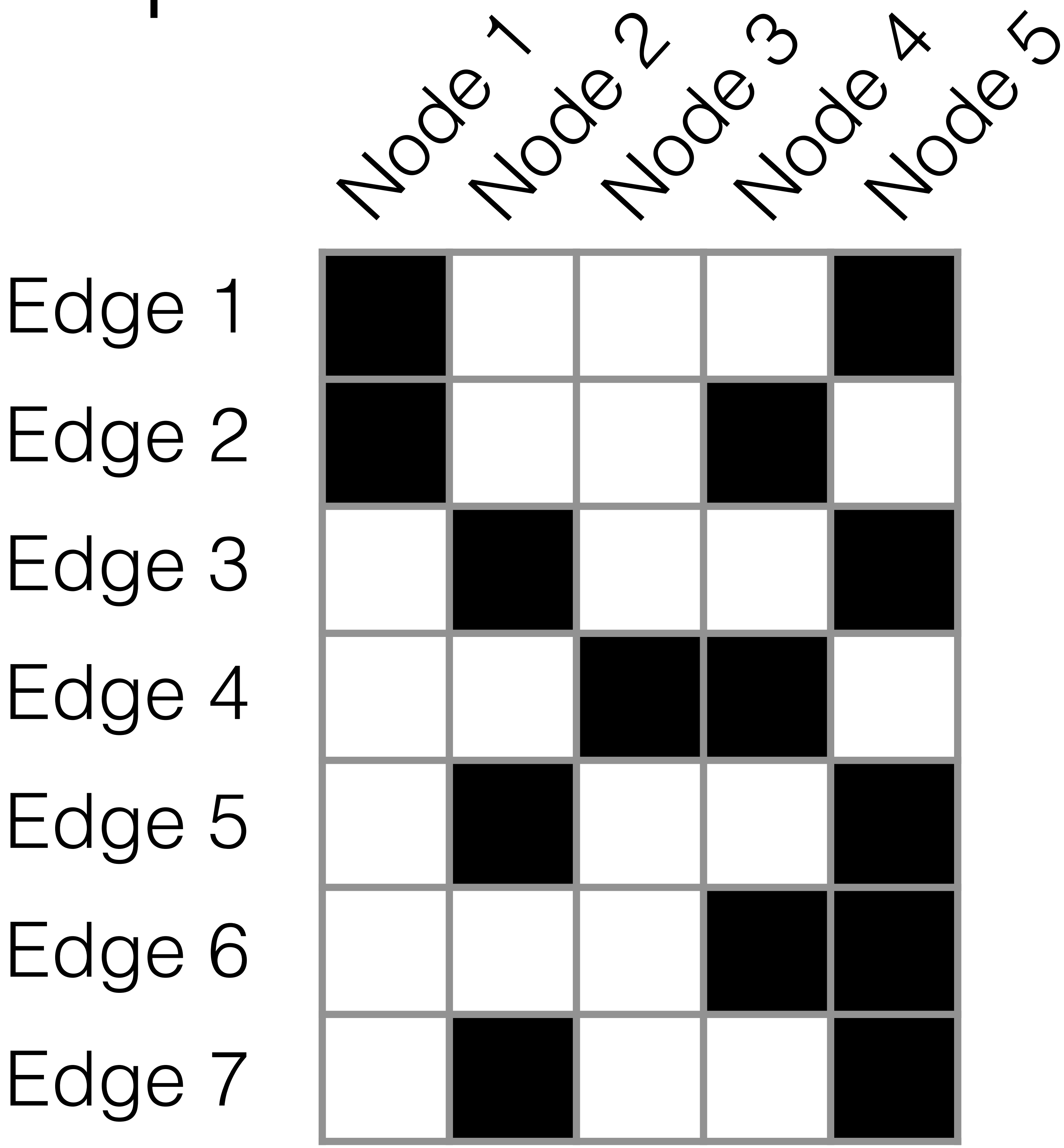


[Broderick, Jordan, Pitman 2013; Broderick, Pitman, Jordan 2013]

# Graph

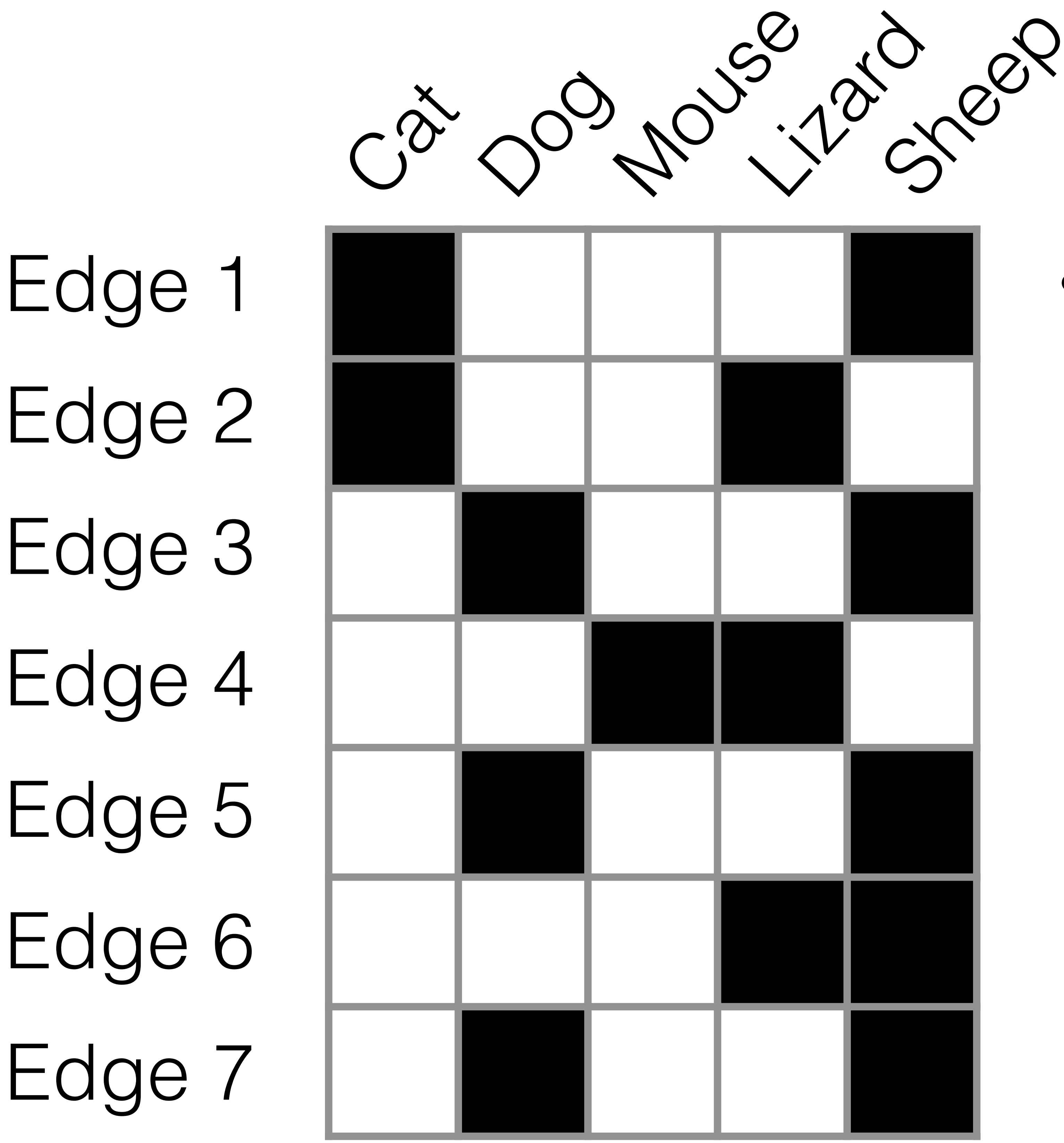

Cat
Dog
Mouse
Lizard
Sheep
Edge 1
Edge 2
Edge 3
Edge 4
Edge 5
Edge 6
Edge 7

# Graph



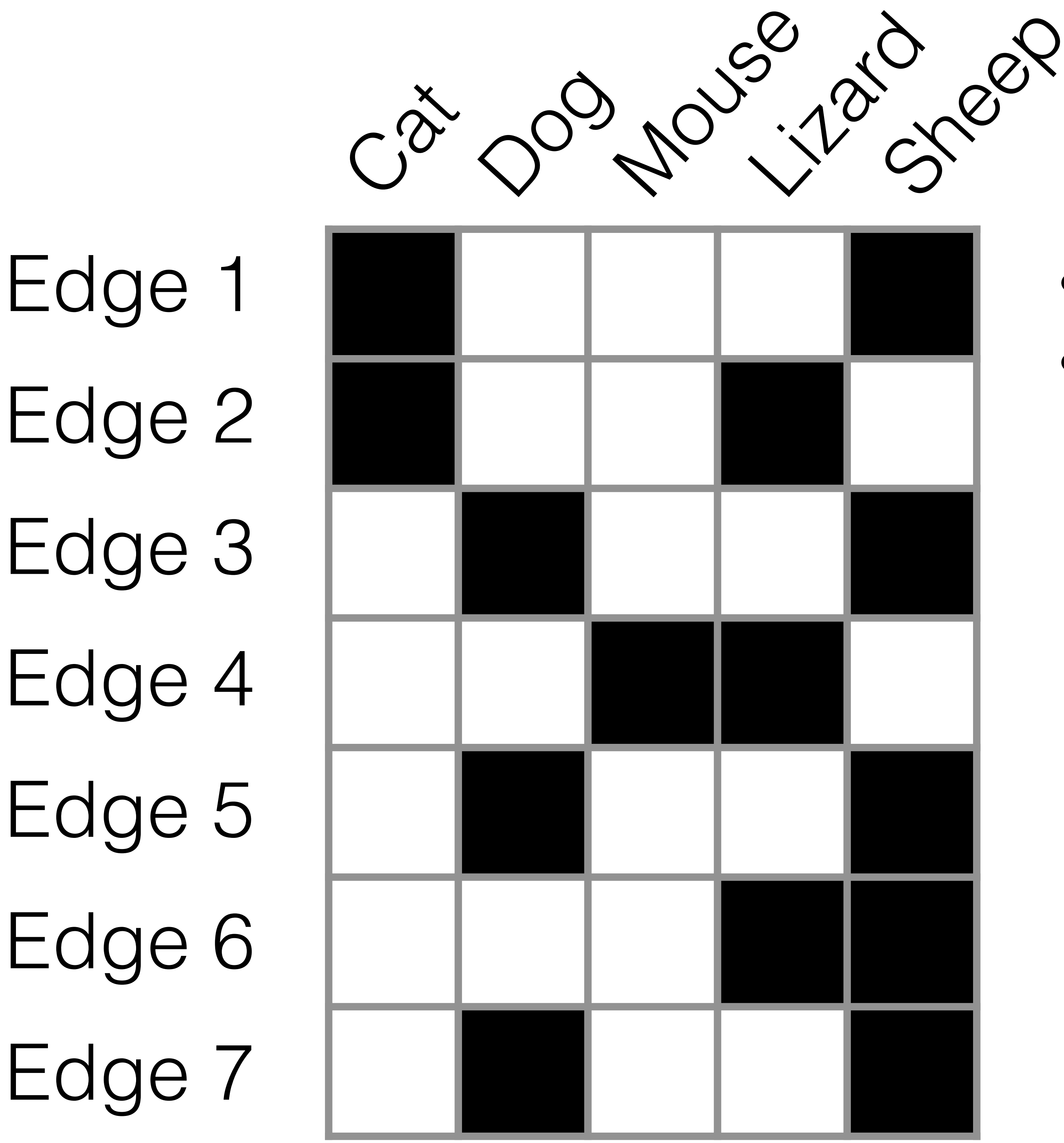
Node 1
Node 2
Node 3
Node 4
Node 5
Edge 1
Edge 2
Edge 3
Edge 4
Edge 5
Edge 6
Edge 7

# Graph



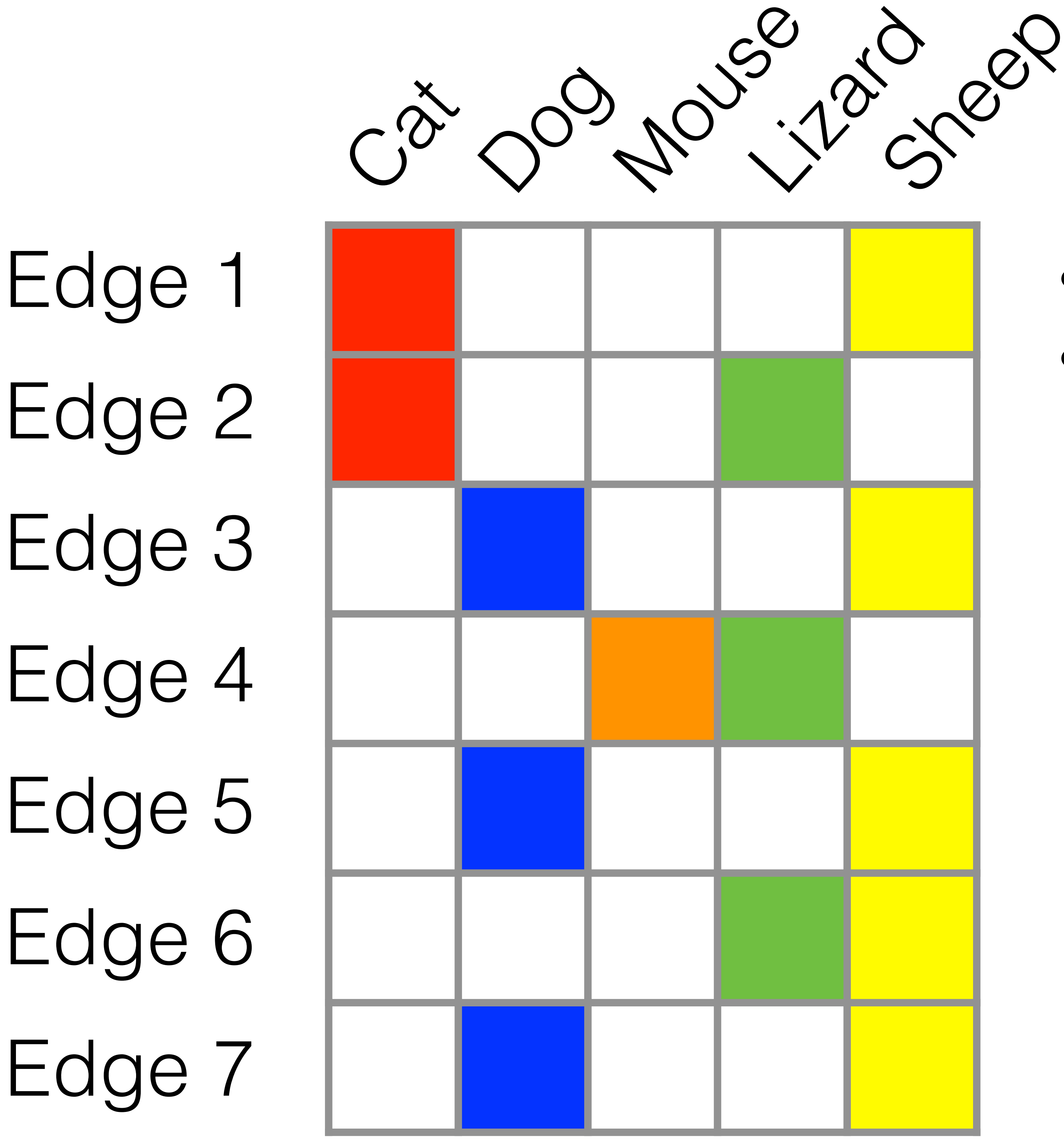

- Groups: vertices

# Graph



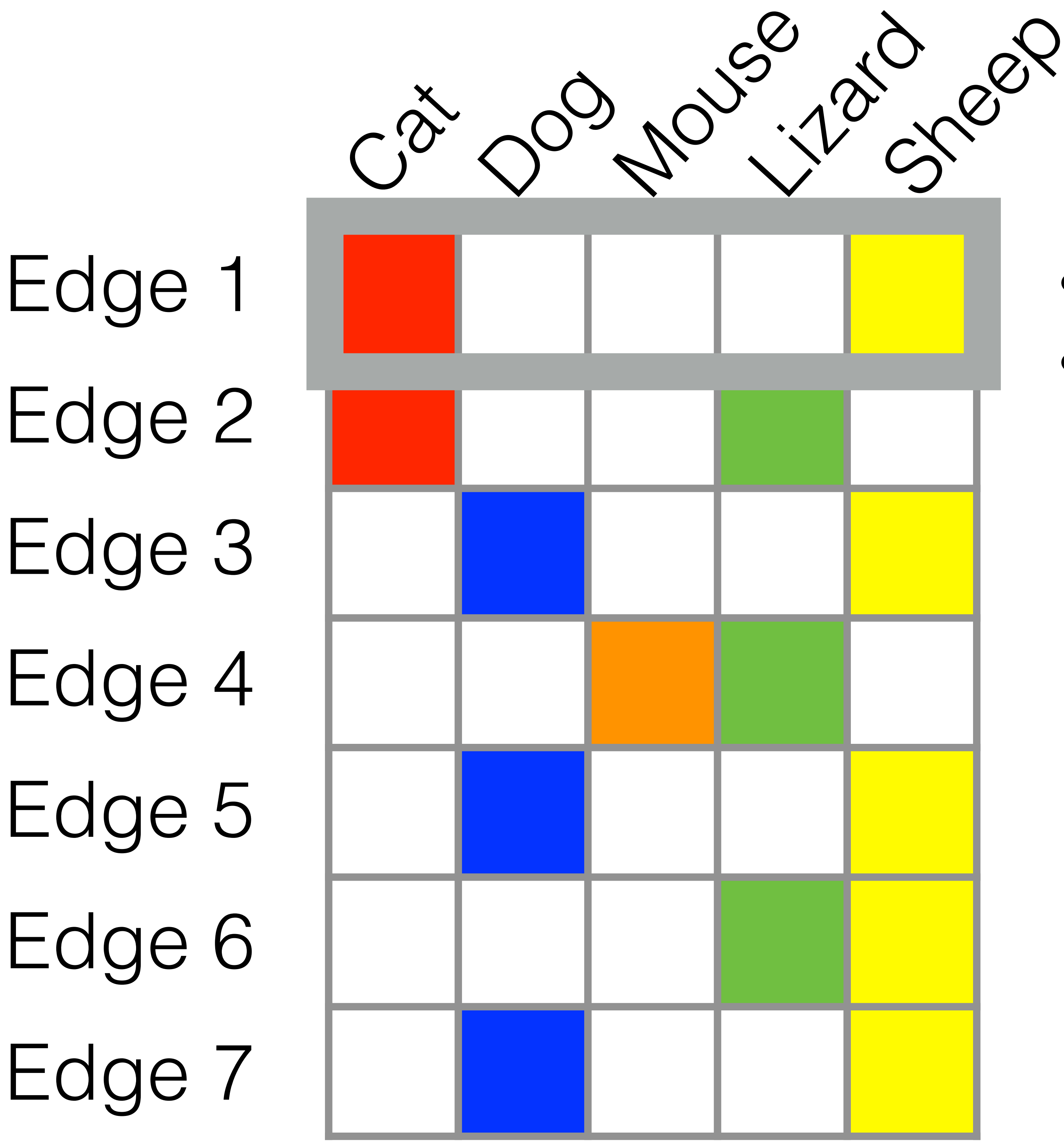

- Groups: vertices
- Edge-exchangeable

# Graph



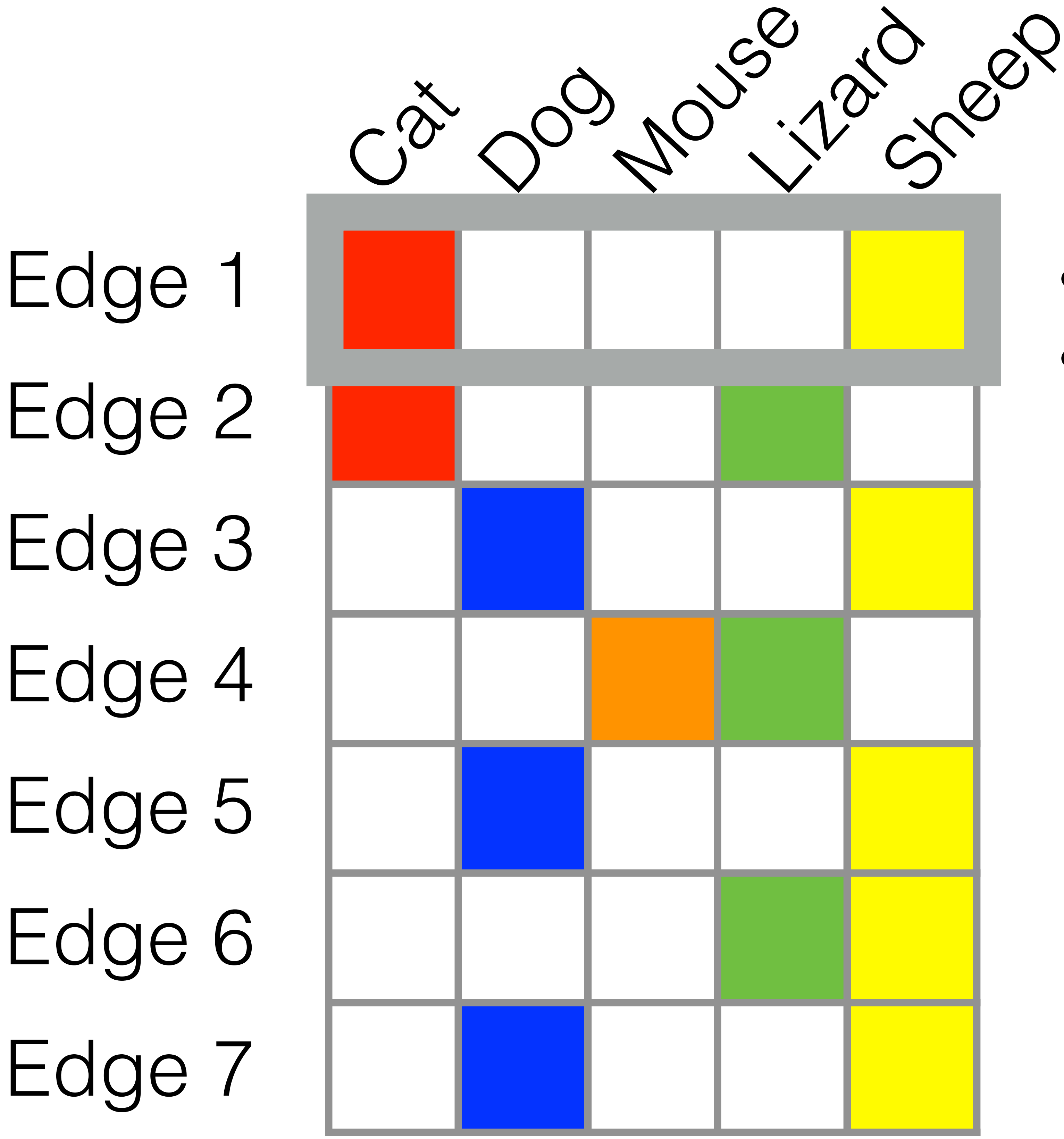

- Groups: vertices
- Edge-exchangeable

# Graph

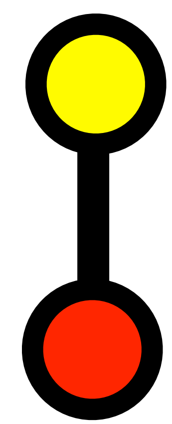



- Groups: vertices
- Edge-exchangeable

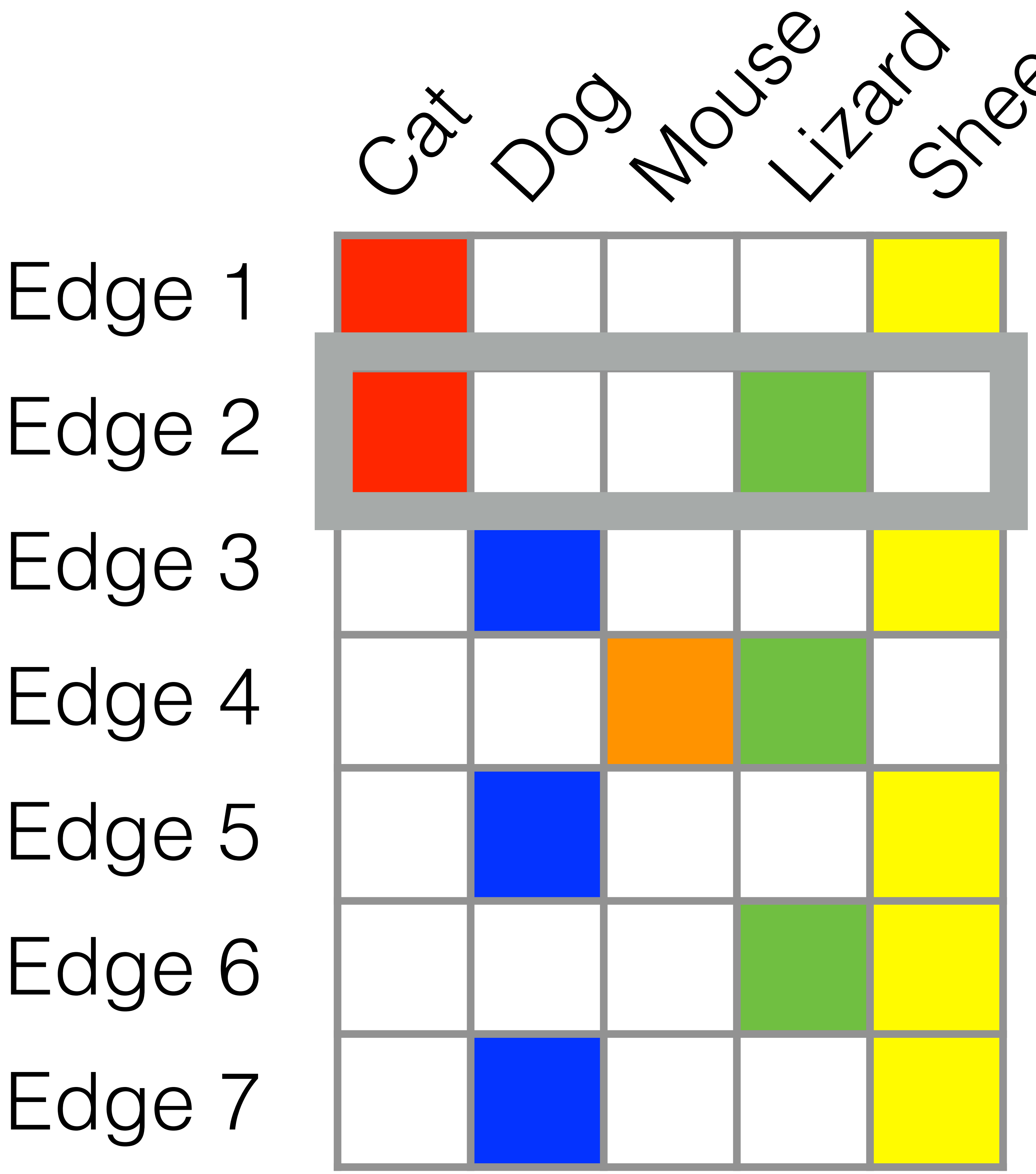
# Graph



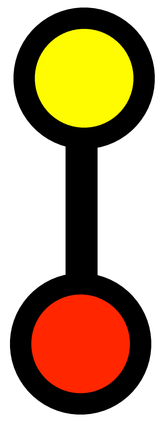

- Groups: vertices
- Edge-exchangeable



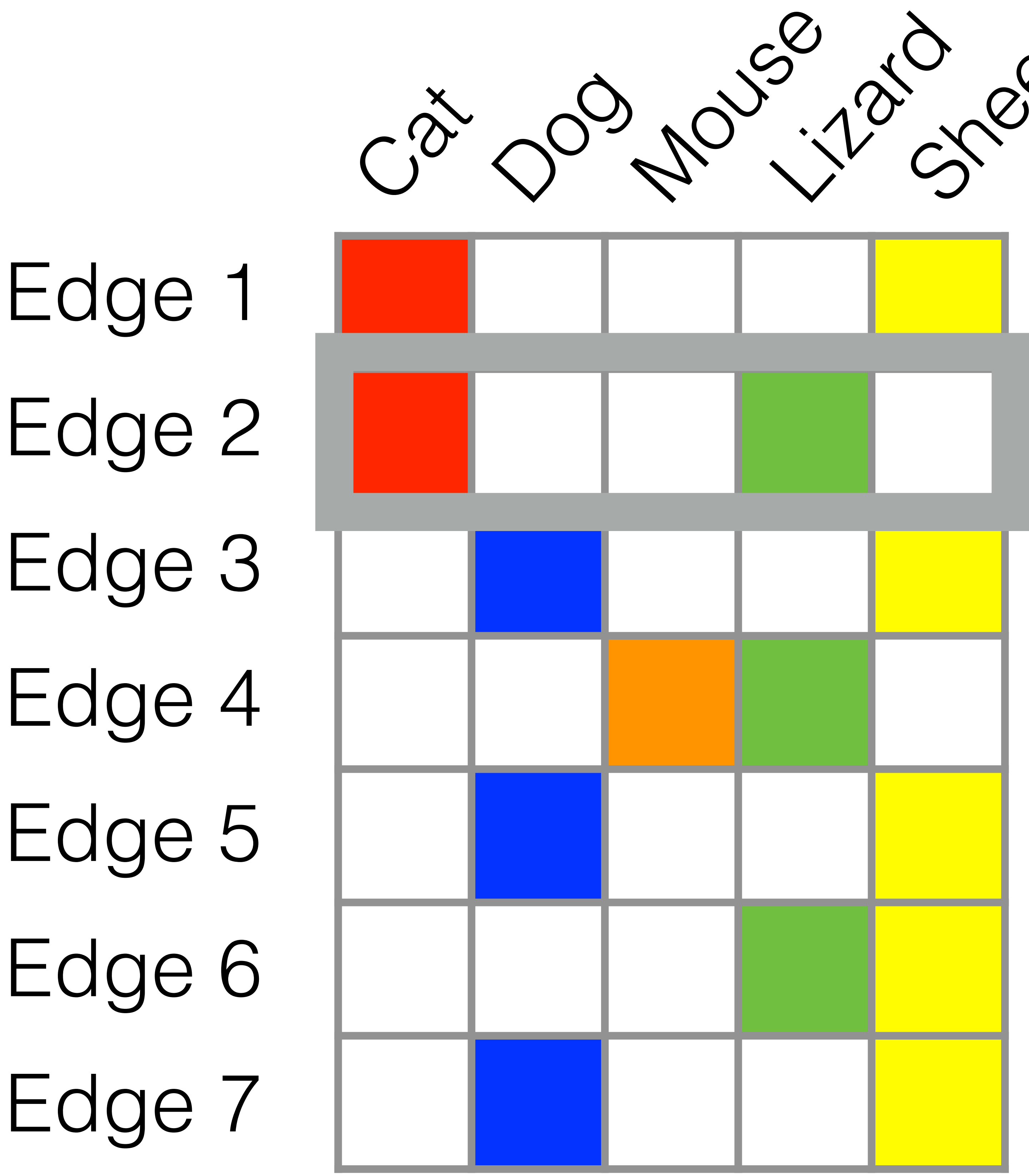
# Graph



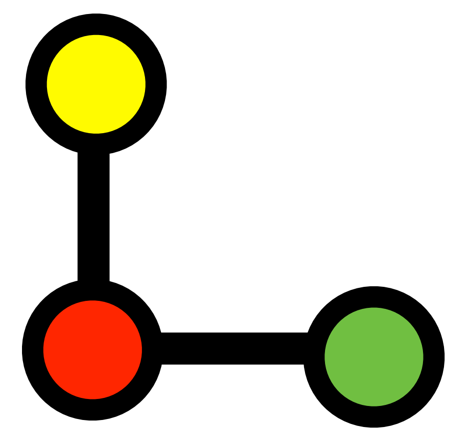

- Groups: vertices
- Edge-exchangeable



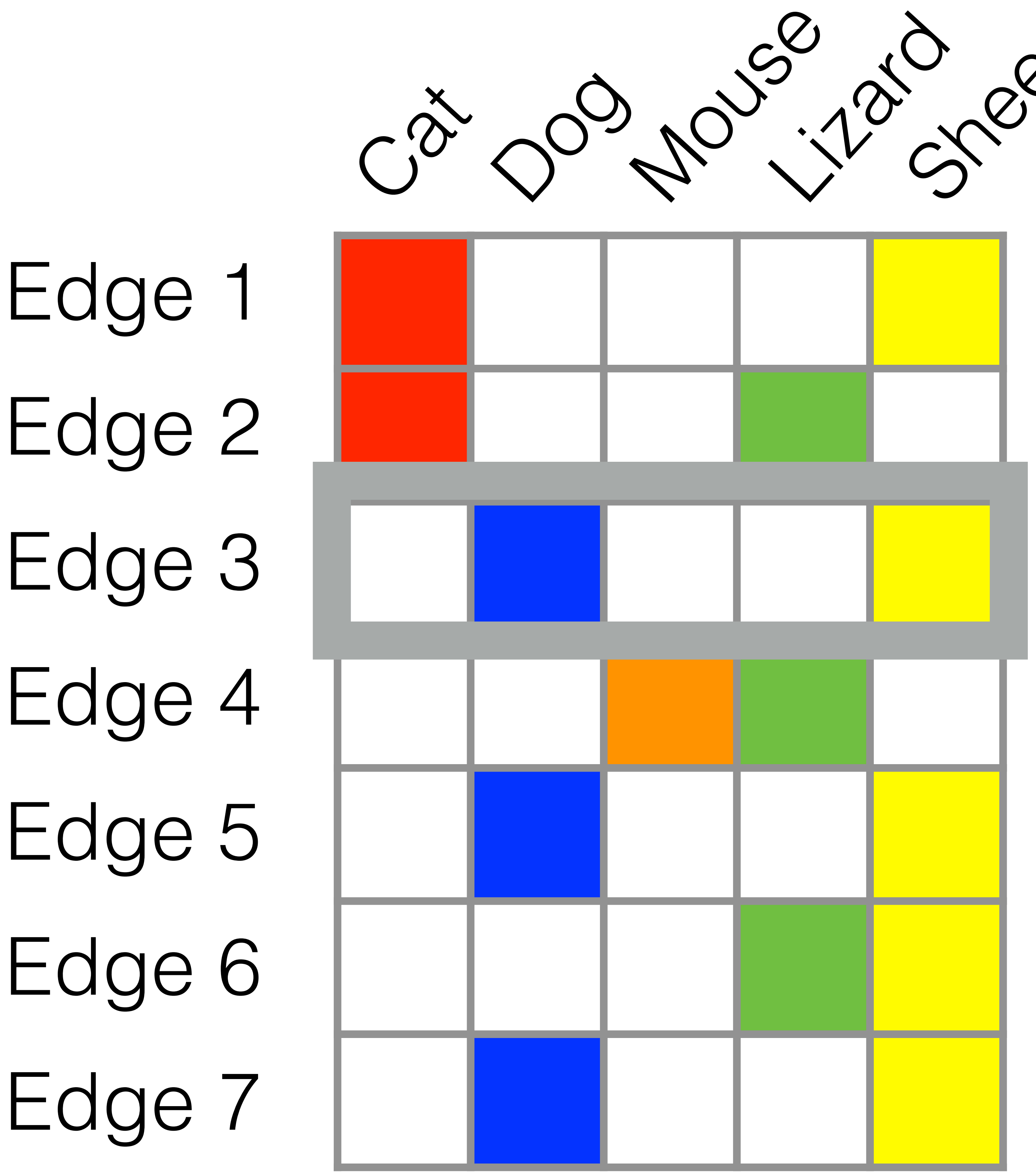
# Graph



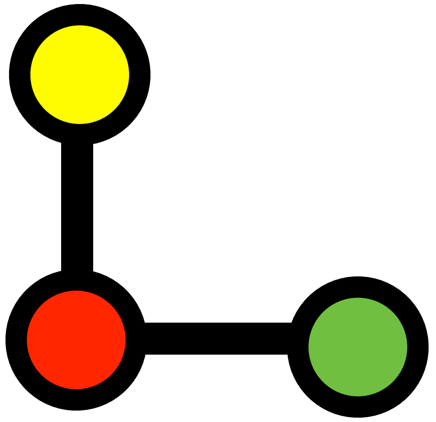

- Groups: vertices
- Edge-exchangeable



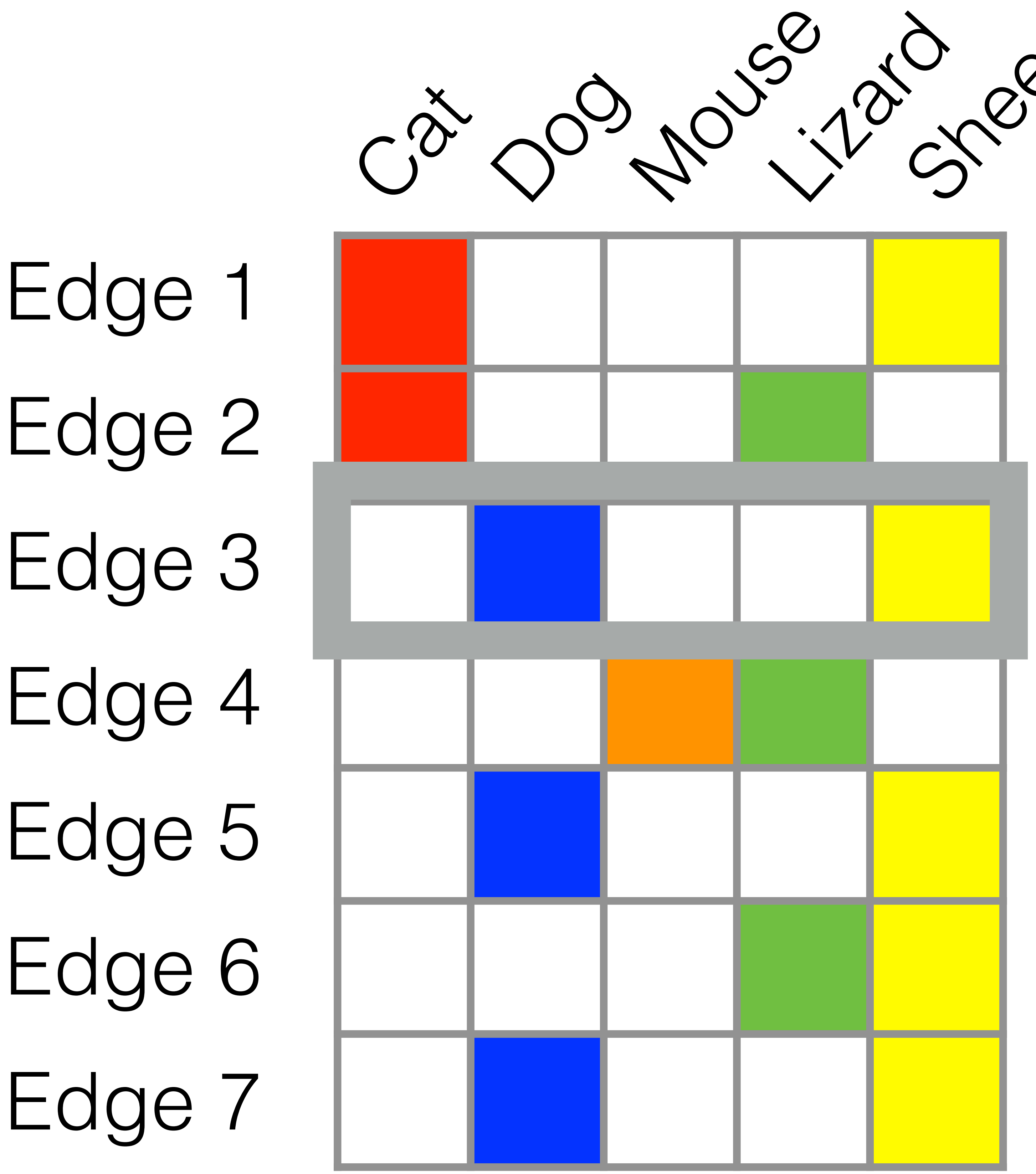
# Graph



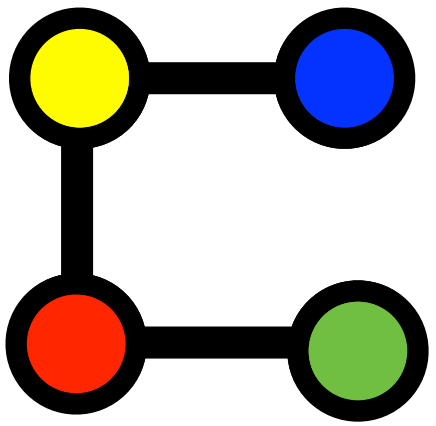

- Groups: vertices
- Edge-exchangeable



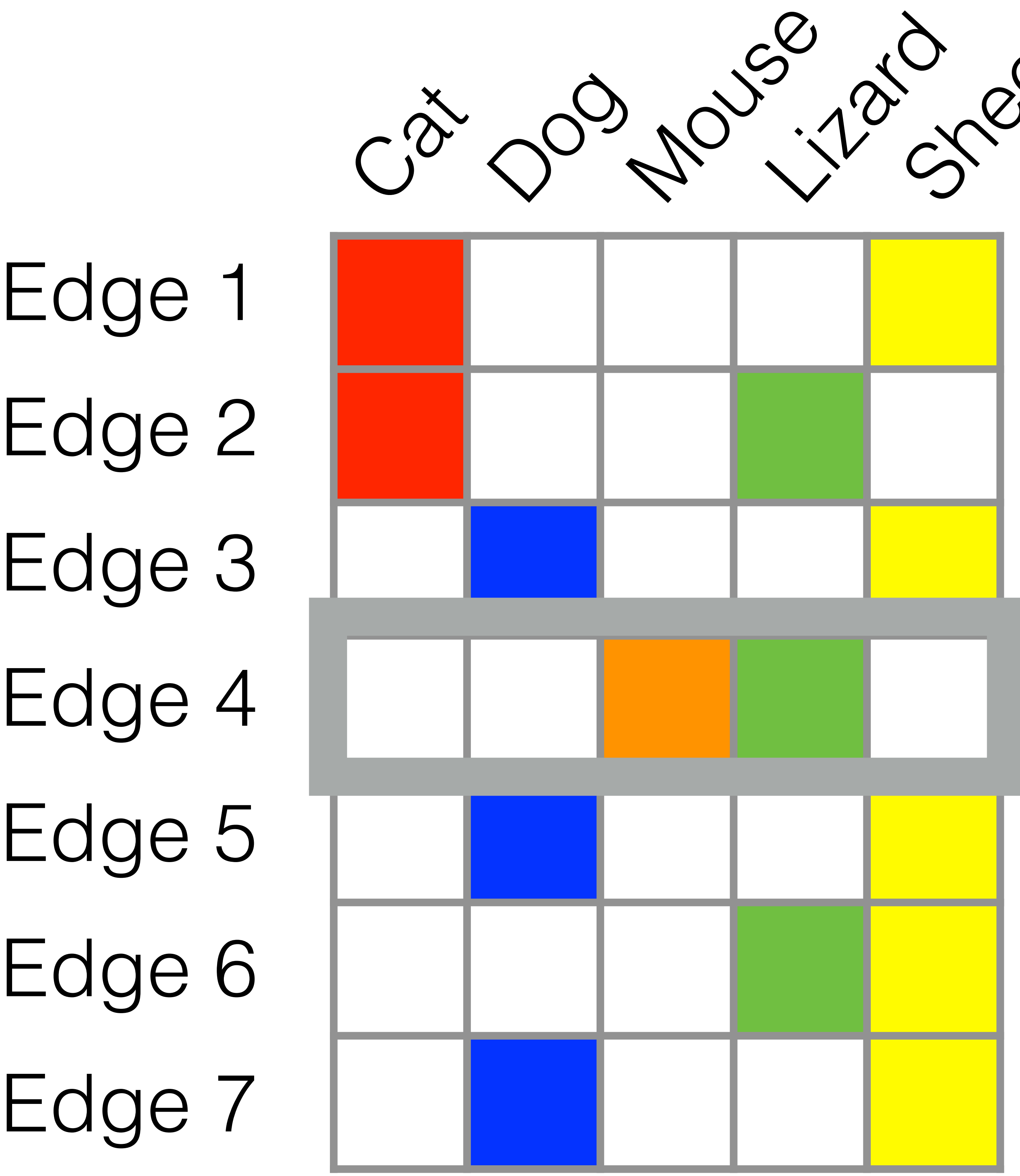
# Graph



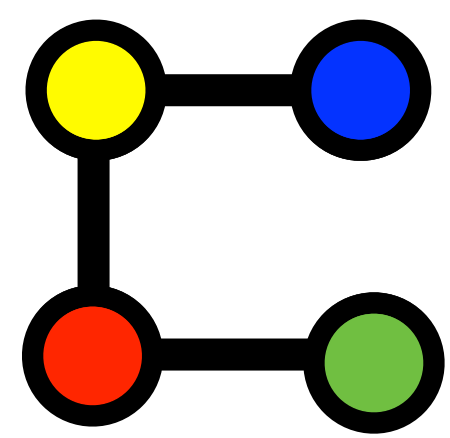

- Groups: vertices
- Edge-exchangeable



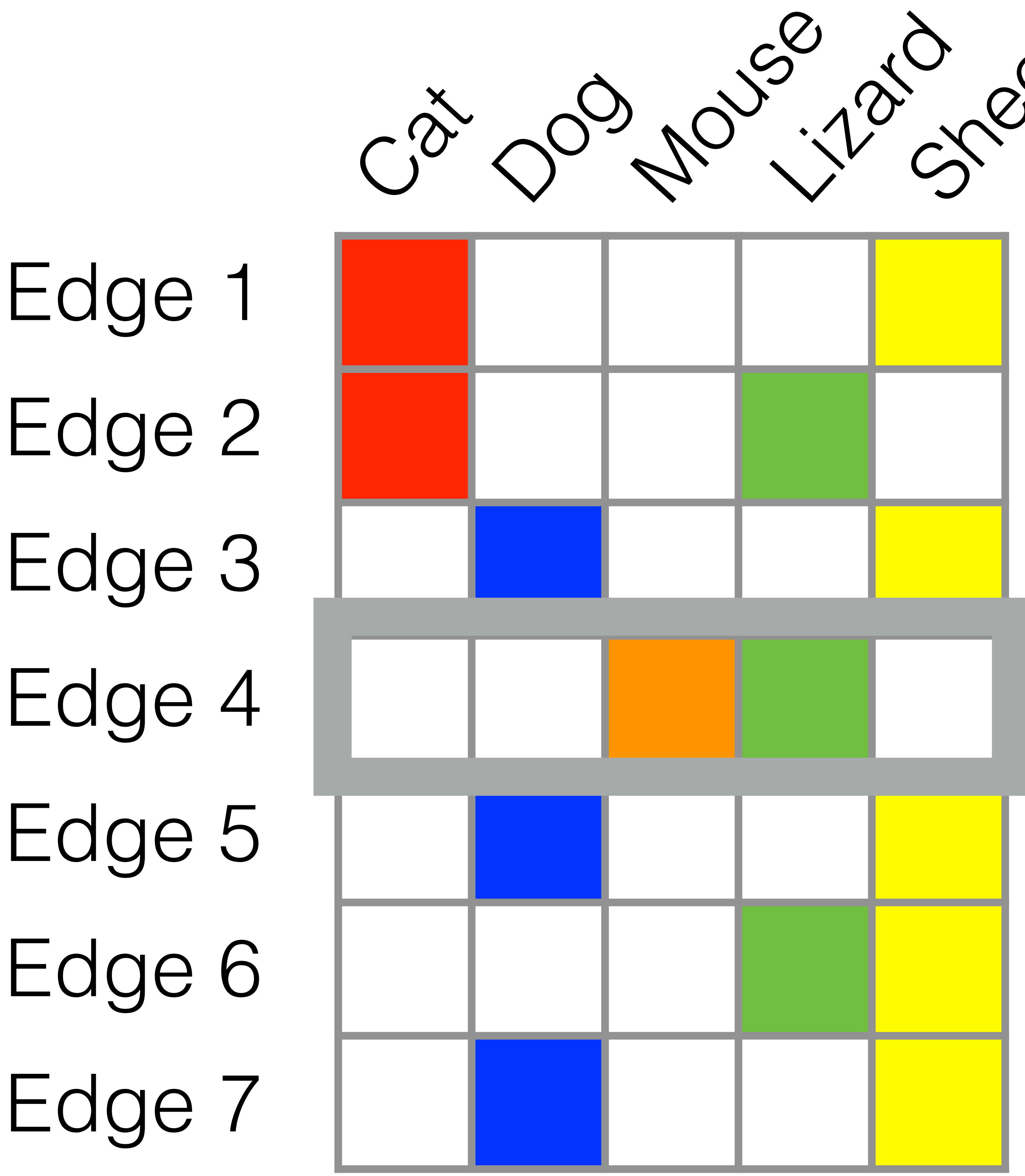
# Graph



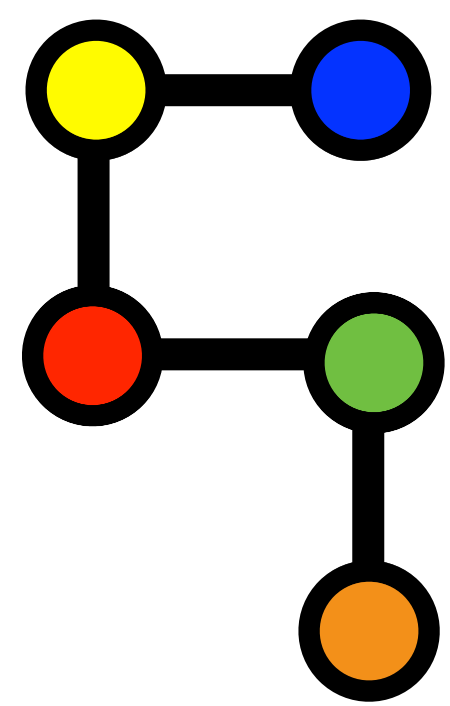

- Groups: vertices
- Edge-exchangeable



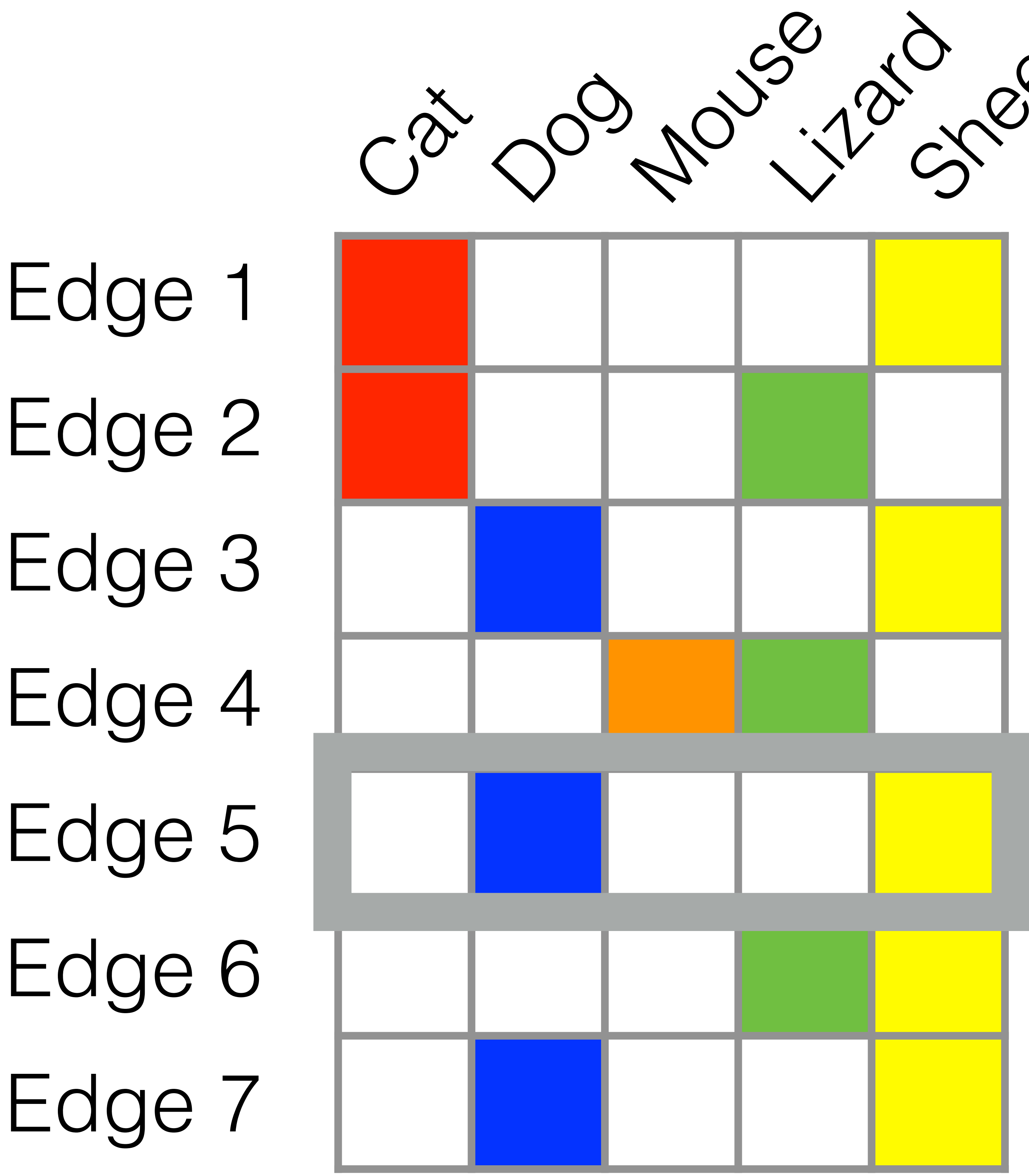
# Graph



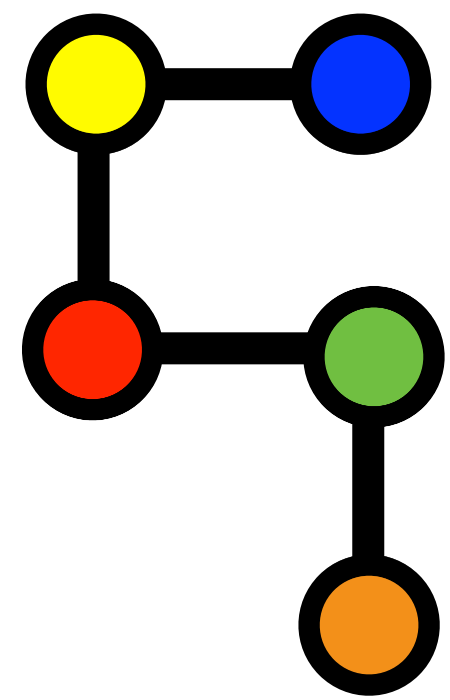

- Groups: vertices
- Edge-exchangeable



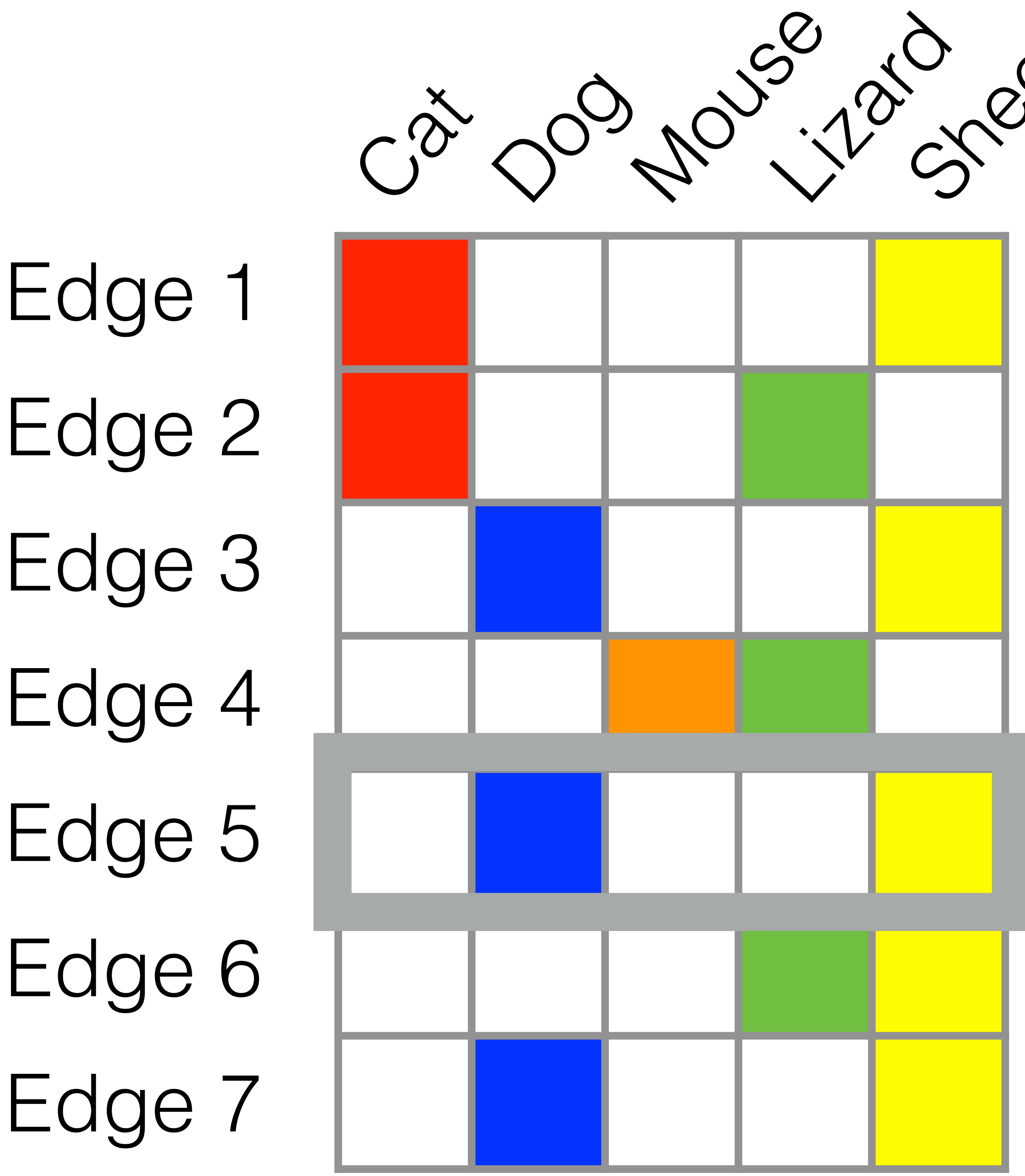
# Graph



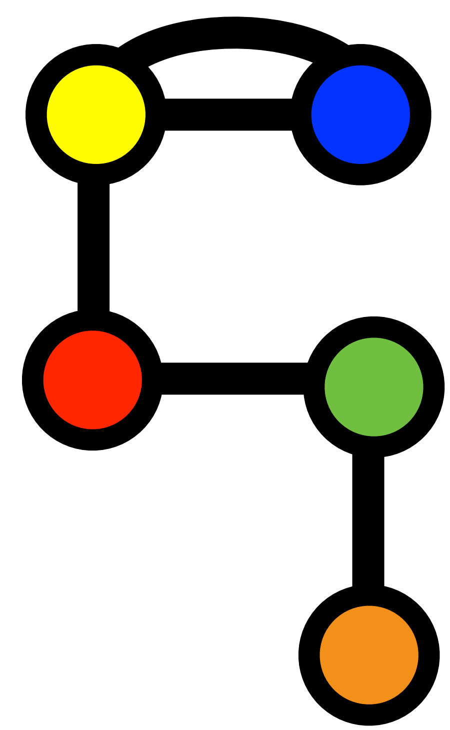

- Groups: vertices
- Edge-exchangeable



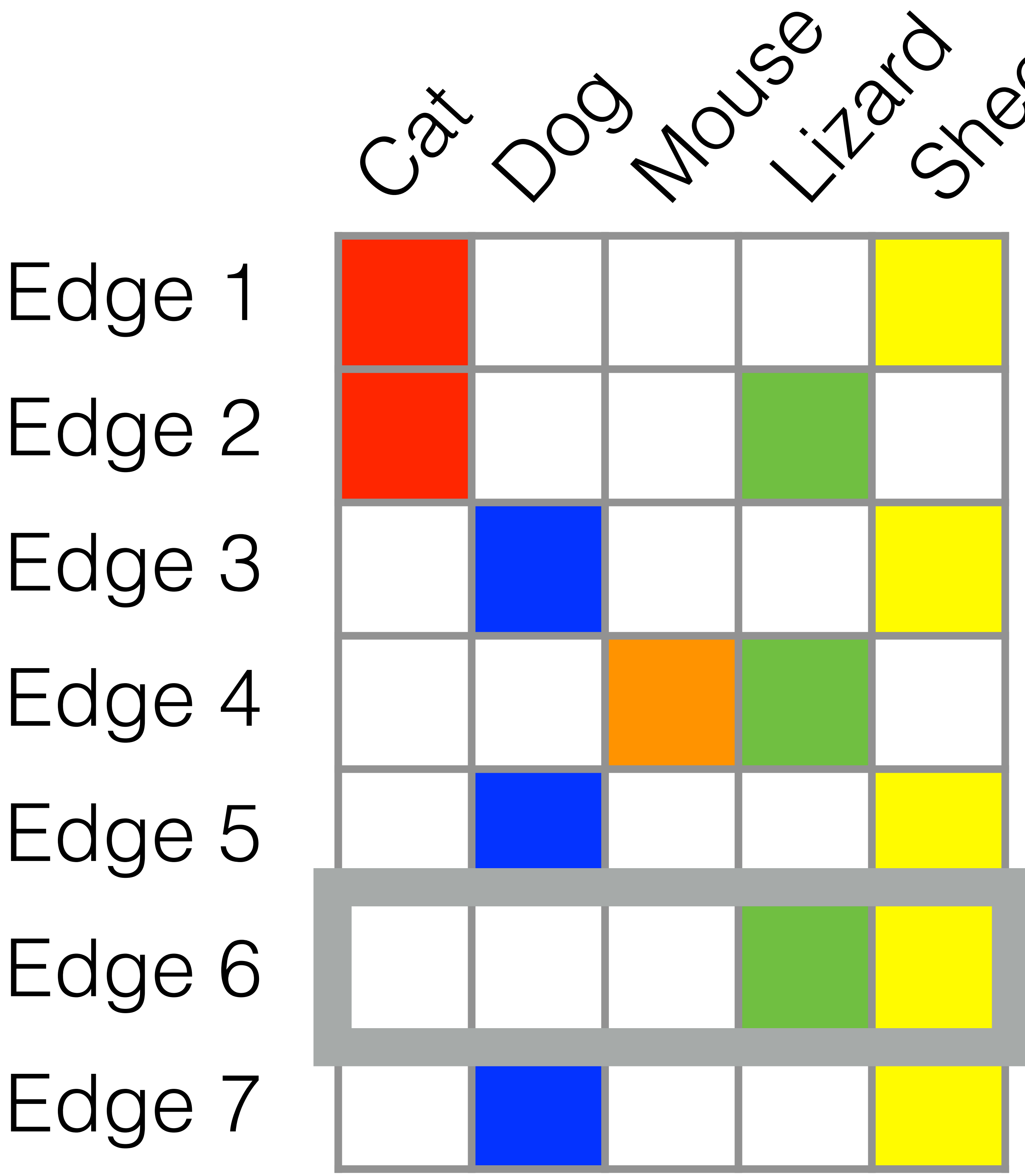
# Graph



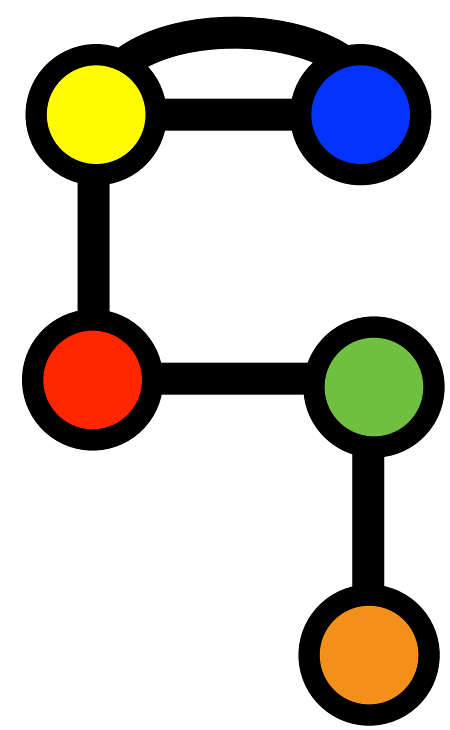

- Groups: vertices
- Edge-exchangeable



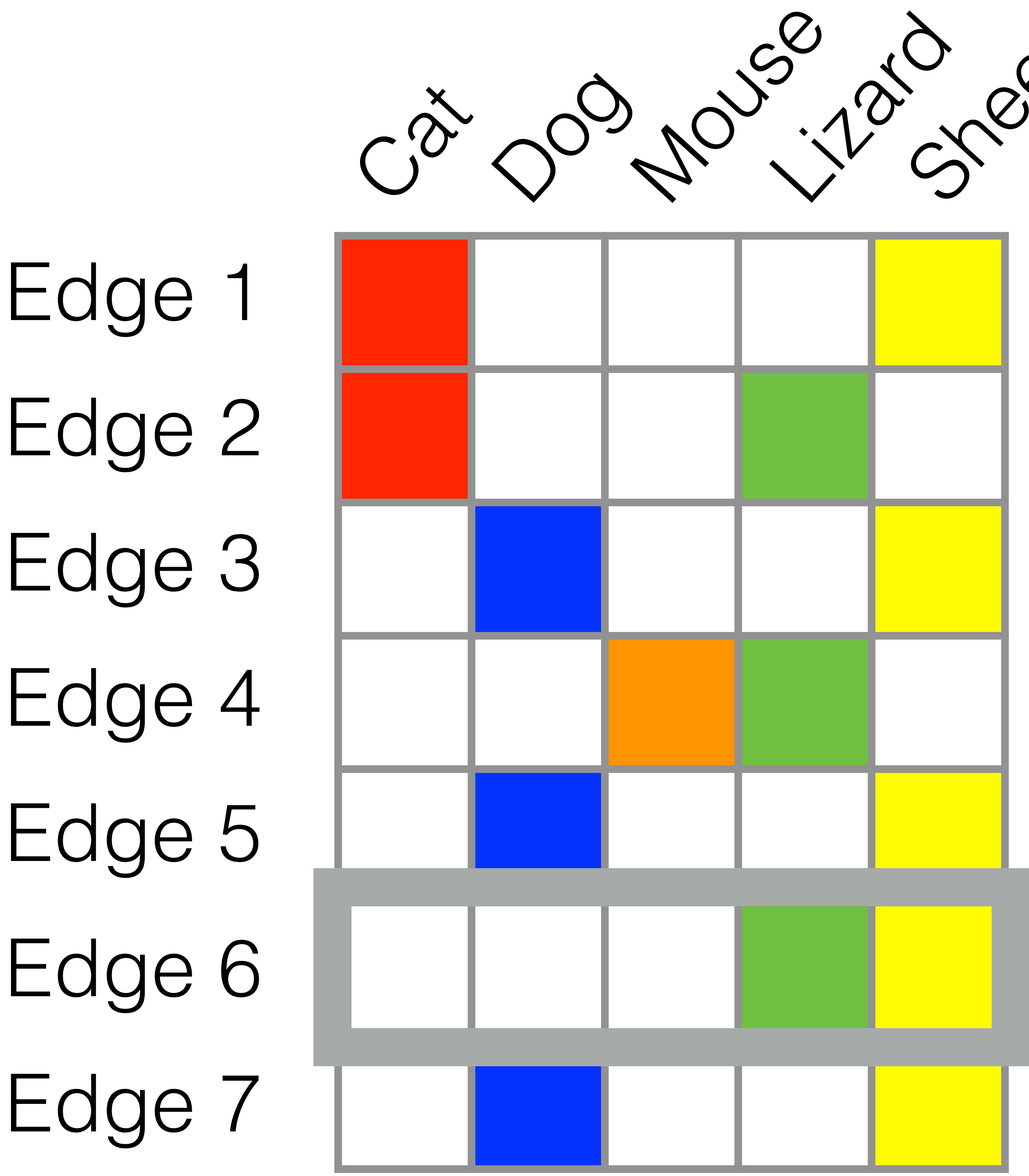
# Graph



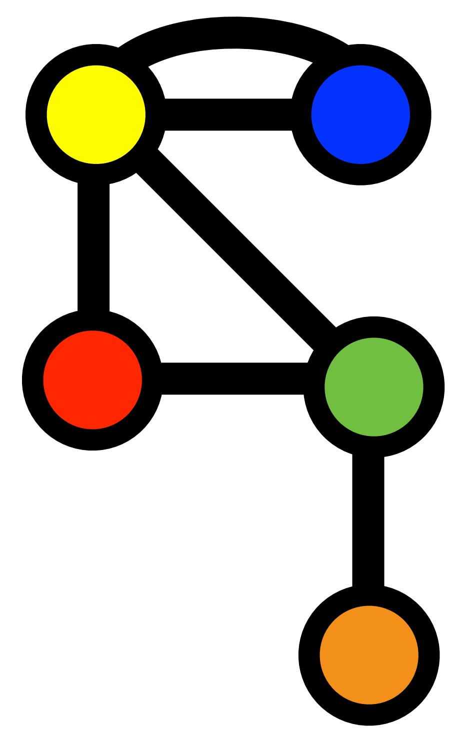

- Groups: vertices
- Edge-exchangeable



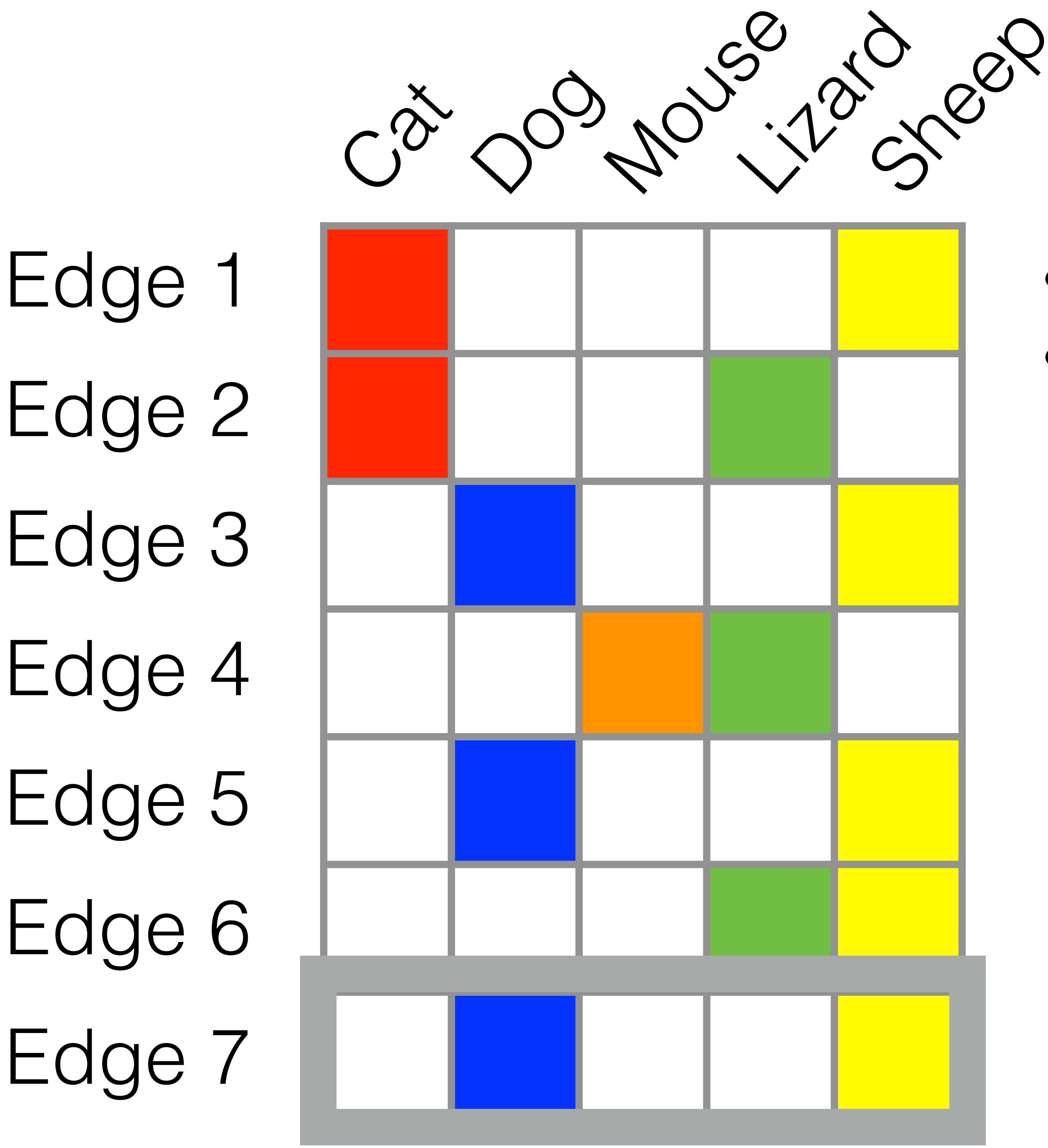
# Graph



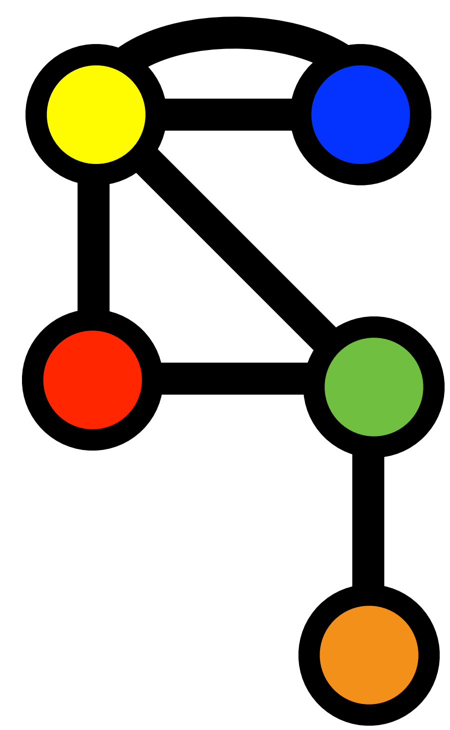

- Groups: vertices
- Edge-exchangeable



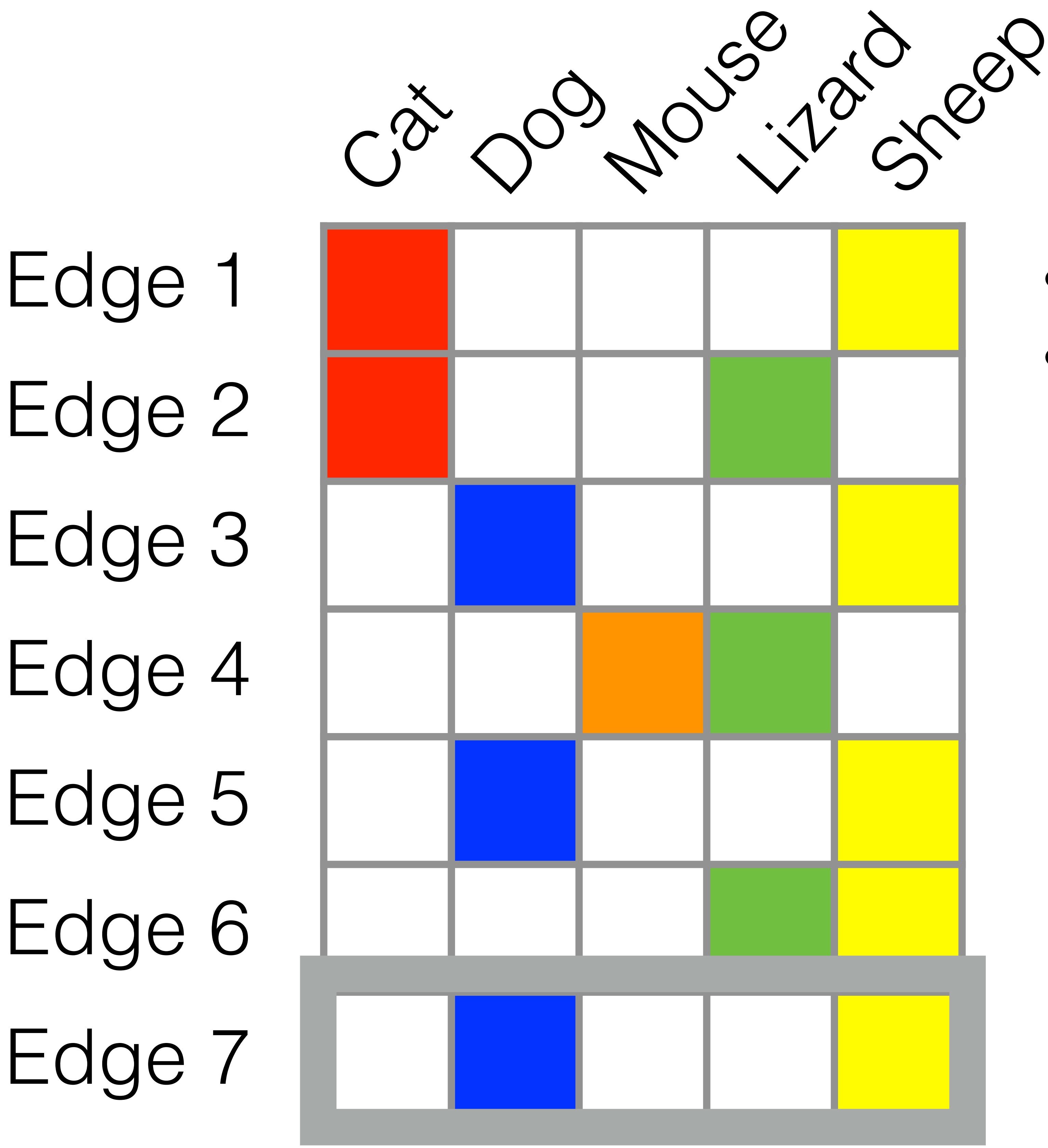
# Graph



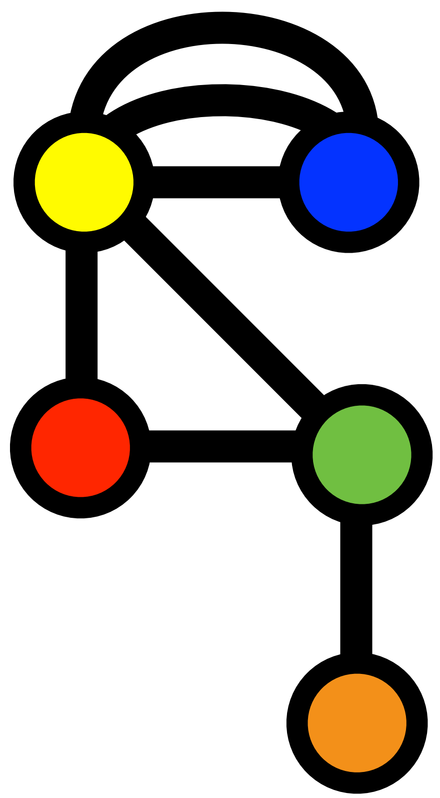

- Groups: vertices
- Edge-exchangeable



# Graph




- Groups: vertices
- Edge-exchangeable



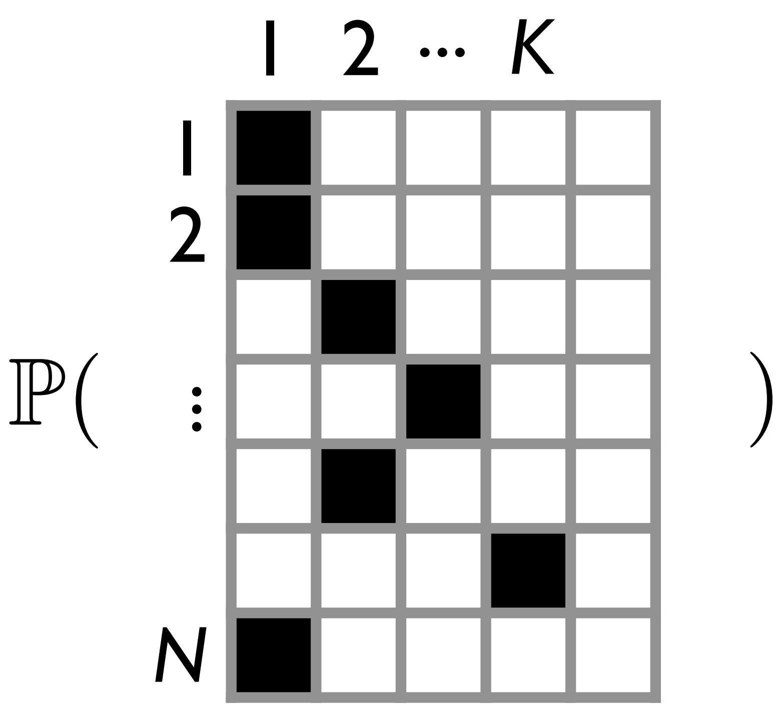
Exchangeable clustering distributions are characterized

What about:

Exchangeable feature allocations?

**Edge-exchangeable graphs?**

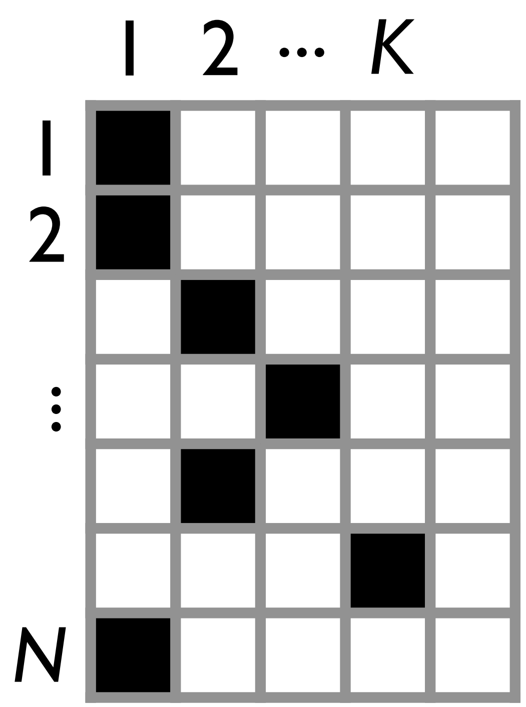
# Exchangeable probability functions






[Pitman 1995]

# Exchangeable probability functions

$\mathbb{P}($  $) = p\left(S_{N,1}, \ldots, S_{N,K}\right)$

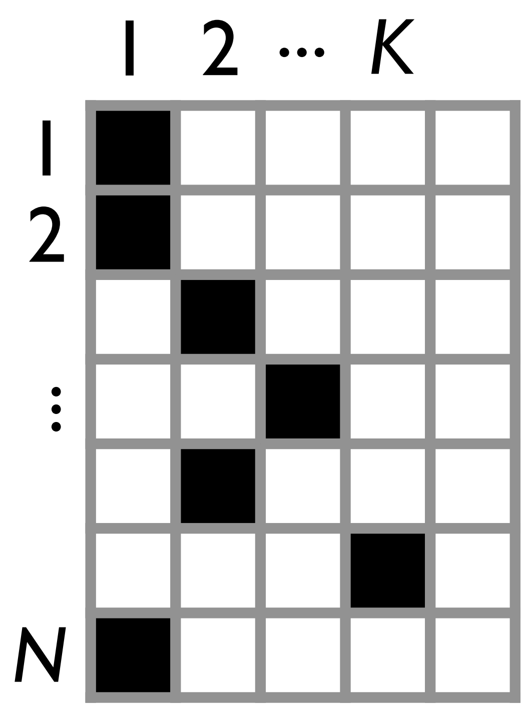




[Pitman 1995]

# Exchangeable probability functions

Size of *K*th cluster




$$\mathbb{P}(\ \cdots\ ) = p\left(S_{N,1}, \ldots, S_{N,K}\right)$$



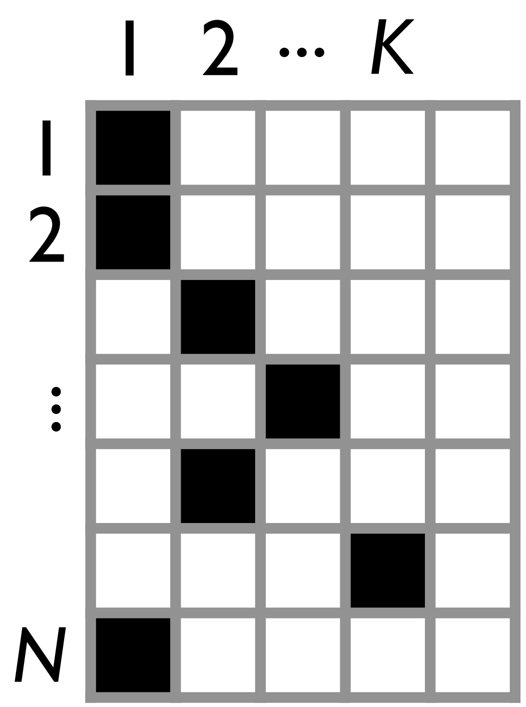
[Pitman 1995]

# Exchangeable probability functions

exchangeable **partition** probability function (E**P**PF)

Size of *K*th cluster

$\mathbb{P}($




$) = p\left(S_{N,1}, \ldots, S_{N,K}\right)$



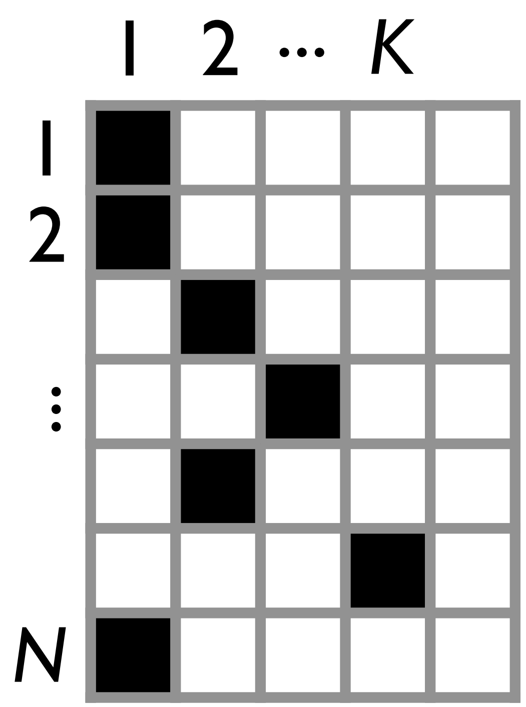
[Pitman 1995]

# Exchangeable probability functions

Thm (Pitman). Every exchangeable clustering has an exchangeable **partition** probability function (E**P**PF)

Size of *K*th cluster

$$\mathbb{P}\Big($$




$$\Big) = p\left(S_{N,1}, \ldots, S_{N,K}\right)$$

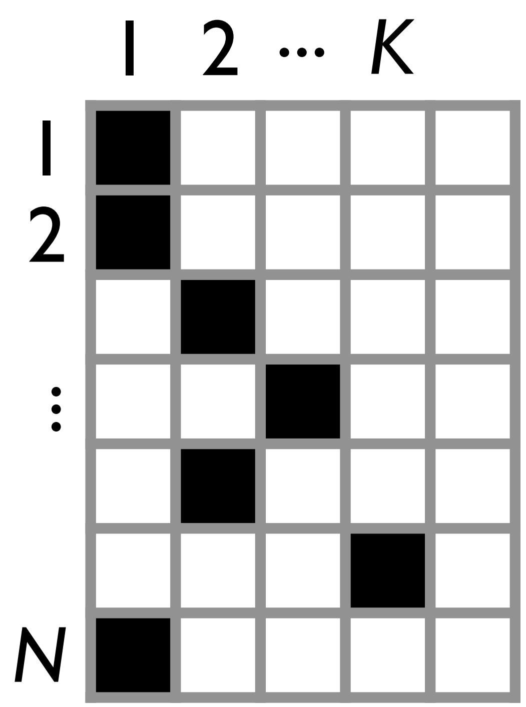


[Pitman 1995]

# Exchangeable probability functions

Thm (Pitman). Every exchangeable clustering has an exchangeable **partition** probability function (E**P**PF)

Size of *K*th cluster

$\mathbb{P}($ 

$) = p(N; S_{N,1}, \ldots, S_{N,K})$

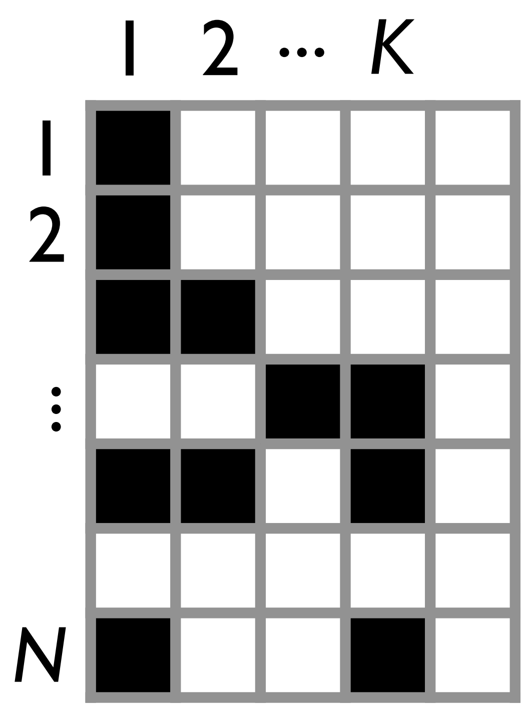


[Pitman 1995]

# Exchangeable probability functions

Thm (Pitman). Every exchangeable clustering has an exchangeable **partition** probability function (E**P**PF)

$\mathbb{P}($




$) = p(N; S_{N,1}, \ldots, S_{N,K})$

Size of *K*th cluster

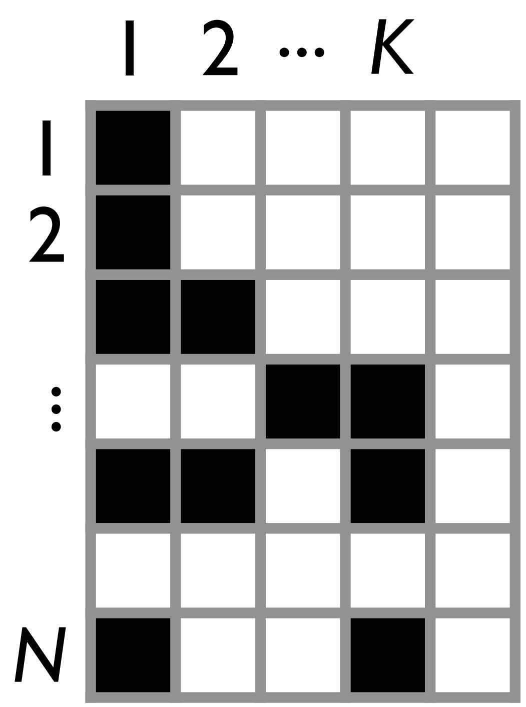


[Pitman 1995]

# Exchangeable probability functions

Thm (Pitman). Every exchangeable clustering has an exchangeable **partition** probability function (E**P**PF)

Size of *K*th feature

$\mathbb{P}($  $) = p(N; S_{N,1}, \ldots, S_{N,K})$





[Pitman 1995]

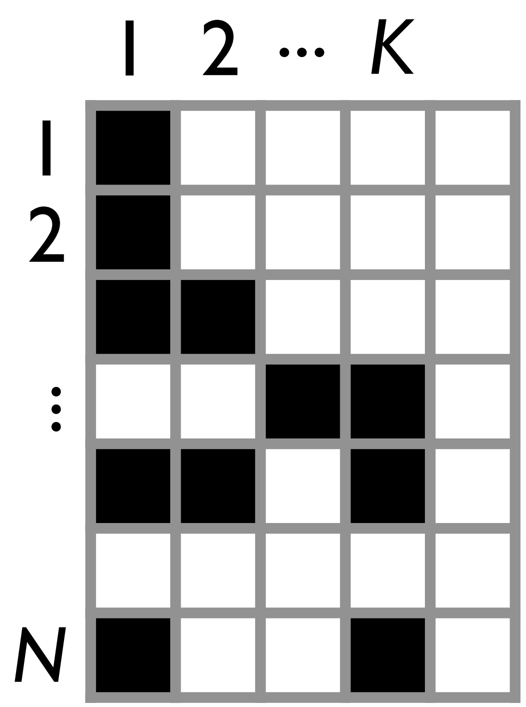
# Exchangeable probability functions

Thm (Pitman). Every exchangeable clustering has an exchangeable **partition** probability function (E**P**PF)

exchangeable **feature** probability function (E**F**PF)

Size of *K*th feature

$\mathbb{P}($




$) = p(N; S_{N,1}, \ldots, S_{N,K})$



[Pitman 1995; Broderick, Pitman, Jordan 2013]

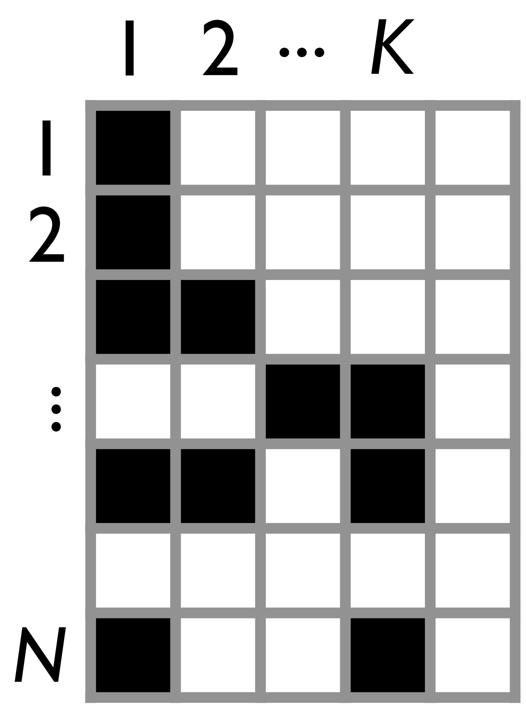
# Exchangeable probability functions

Thm (Pitman). Every exchangeable clustering has an exchangeable **partition** probability function (E**P**PF)

Prop (BPJ). Many, but not all, exchangeable feature allocations have an exchangeable **feature** probability function (E**F**PF)

Size of *K*th feature

$$\mathbb{P}\Bigg(\quad$$




$$\Bigg) = p(N; S_{N,1}, \ldots, S_{N,K})$$



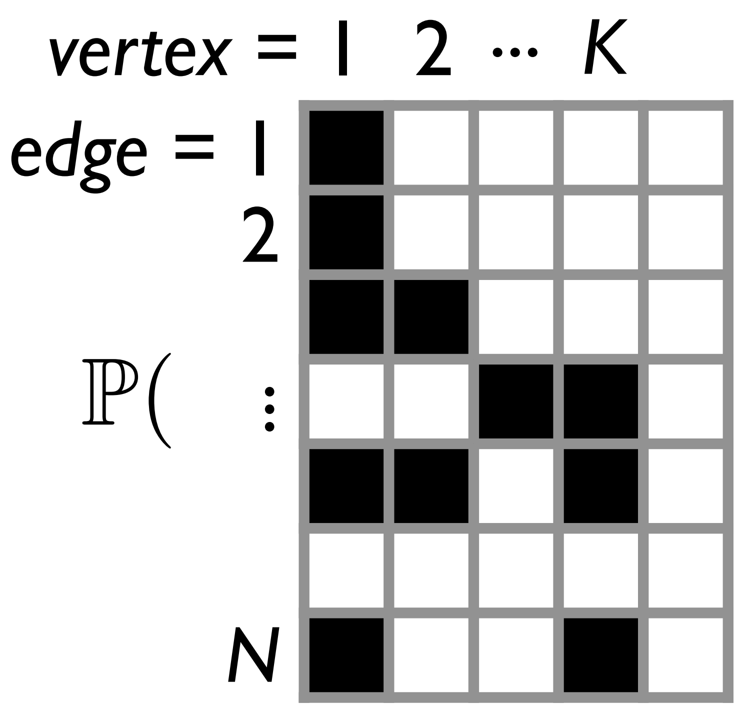
[Pitman 1995; Broderick, Pitman, Jordan 2013]

# Exchangeable probability functions

Thm (Pitman). Every exchangeable clustering has an exchangeable **partition** probability function (E**P**PF)

Prop (BPJ). Many, but not all, exchangeable feature allocations have an exchangeable **feature** probability function (E**F**PF)

Size of *K*th feature

$\mathbb{P}($ 

$) = p(N; S_{N,1}, \ldots, S_{N,K})$



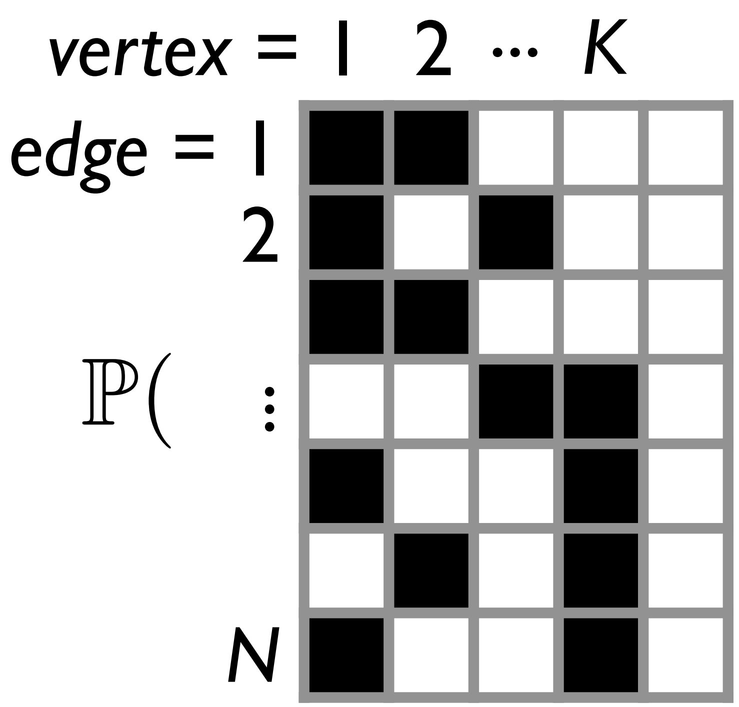
[Pitman 1995; Broderick, Pitman, Jordan 2013; Campbell, Cai, Broderick 2016]

# Exchangeable probability functions

Thm (Pitman). Every exchangeable clustering has an exchangeable **partition** probability function (E**P**PF)

Prop (BPJ). Many, but not all, exchangeable feature allocations have an exchangeable **feature** probability function (E**F**PF)

Size of *K*th feature




$$\mathbb{P}(\quad) = p(N; S_{N,1}, \ldots, S_{N,K})$$



[Pitman 1995; Broderick, Pitman, Jordan 2013; Campbell, Cai, Broderick 2016]

# Exchangeable probability functions

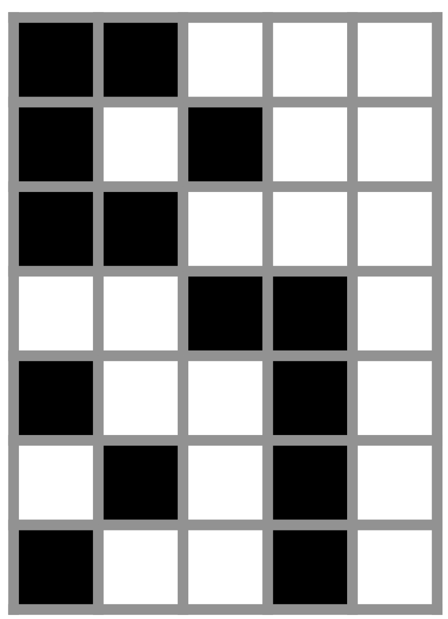
Thm (Pitman). Every exchangeable clustering has an exchangeable **partition** probability function (E**P**PF)

Prop (BPJ). Many, but not all, exchangeable feature allocations have an exchangeable **feature** probability function (E**F**PF)

Size of *K*th vertex

*vertex* = 1 2 ··· *K*

*edge* = 1, 2, ⋮, *N*

$\mathbb{P}($  $) = p(N; S_{N,1}, \ldots, S_{N,K})$

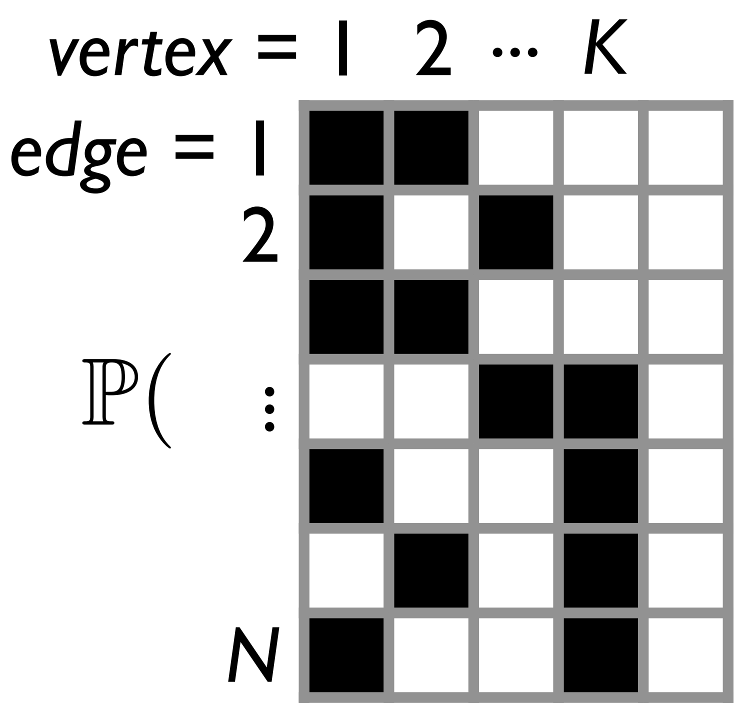


[Pitman 1995; Broderick, Pitman, Jordan 2013; Campbell, Cai, Broderick 2016]

# Exchangeable probability functions

Thm (Pitman). Every exchangeable clustering has an exchangeable **partition** probability function (E**P**PF)

Prop (BPJ). Many, but not all, exchangeable feature allocations have an exchangeable **feature** probability function (E**F**PF)




$$\mathbb{P}(\quad\cdot\quad) = p(N; S_{N,1}, \ldots, S_{N,K})$$

Degree of *K*th vertex



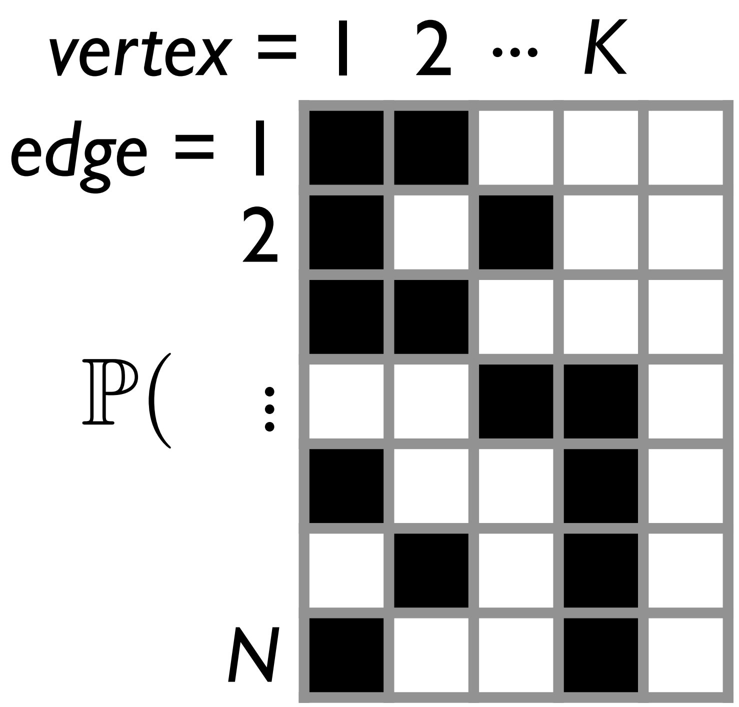
[Pitman 1995; Broderick, Pitman, Jordan 2013; Campbell, Cai, Broderick 2016]

# Exchangeable probability functions

Thm (Pitman). Every exchangeable clustering has an exchangeable **partition** probability function (E**P**PF)

Prop (BPJ). Many, but not all, exchangeable feature allocations have an exchangeable **feature** probability function (E**F**PF)

Degree of *K*th vertex




$$\mathbb{P}(\quad) = p\left(S_{N,1}, \ldots, S_{N,K}\right)$$



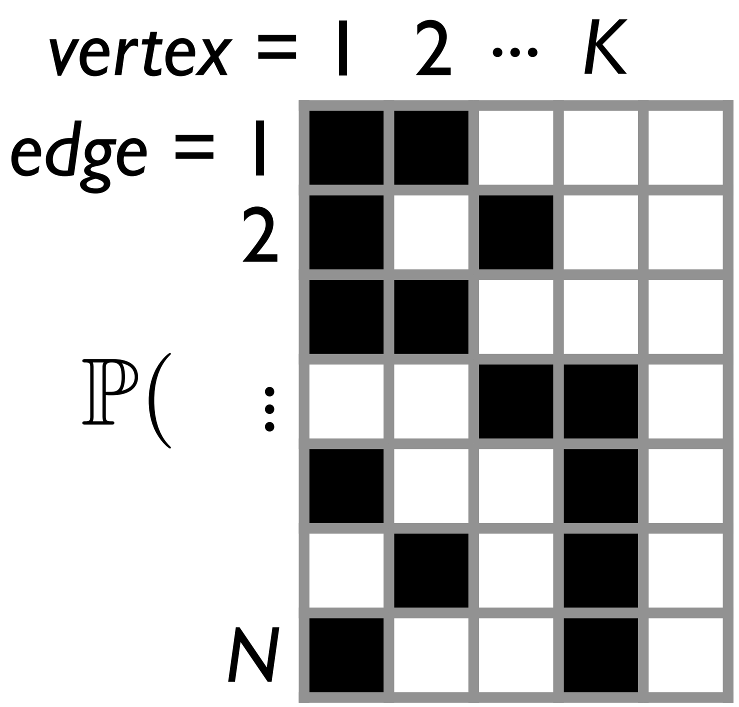
[Pitman 1995; Broderick, Pitman, Jordan 2013; Campbell, Cai, Broderick 2016]

# Exchangeable probability functions

Thm (Pitman). Every exchangeable clustering has an exchangeable **partition** probability function (E**P**PF)

Prop (BPJ). Many, but not all, exchangeable feature allocations have an exchangeable **feature** probability function (E**F**PF)

Degree of *K*th vertex




$$\mathbb{P}(\quad) = p\left(S_{N,1}, \ldots, S_{N,K}\right)$$

Definition (CCB). Exchangeable **vertex** probability function (E**V**PF)



[Pitman 1995; Broderick, Pitman, Jordan 2013; Campbell, Cai, Broderick 2016]

# Exchangeable probability functions

Q: Does every edge-exchangeable graph have an EVPF?



[Broderick, Jordan, Pitman 2013]

# Exchangeable probability functions

Q: Does every edge-exchangeable graph have an EVPF?

A: No. Counterexample:

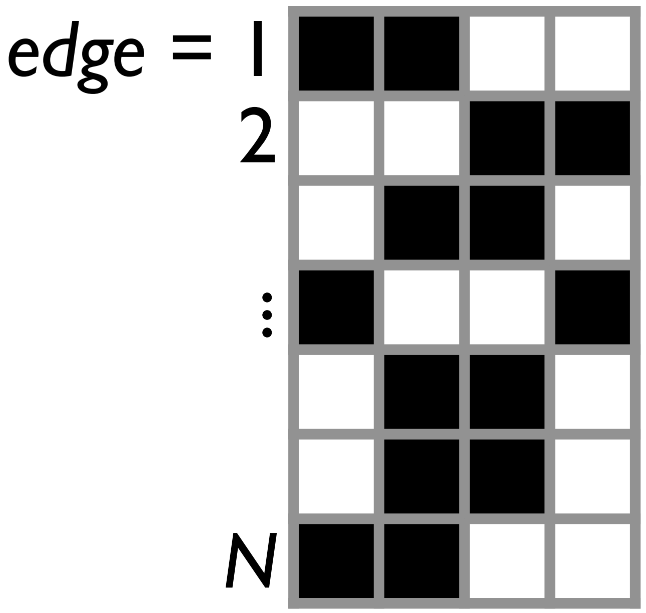


[Broderick, Jordan, Pitman 2013]

# Exchangeable probability functions

Q: Does every edge-exchangeable graph have an EVPF?

A: No. Counterexample:



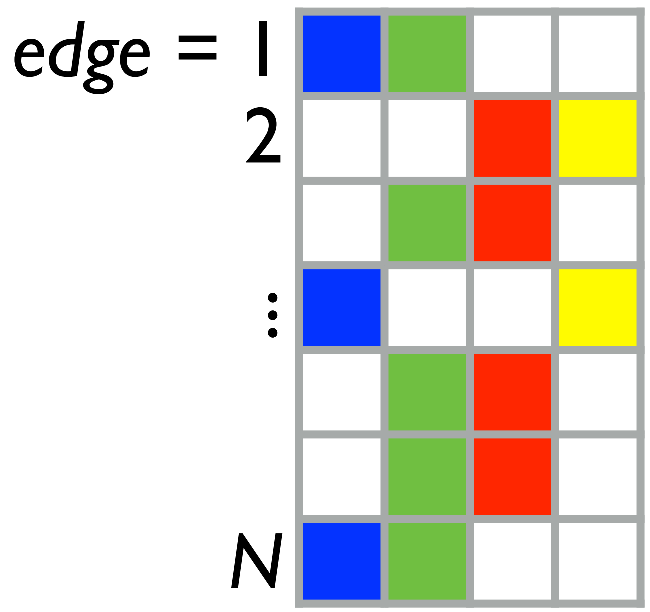



[Broderick, Jordan, Pitman 2013]

# Exchangeable probability functions

Q: Does every edge-exchangeable graph have an EVPF?

A: No. Counterexample:



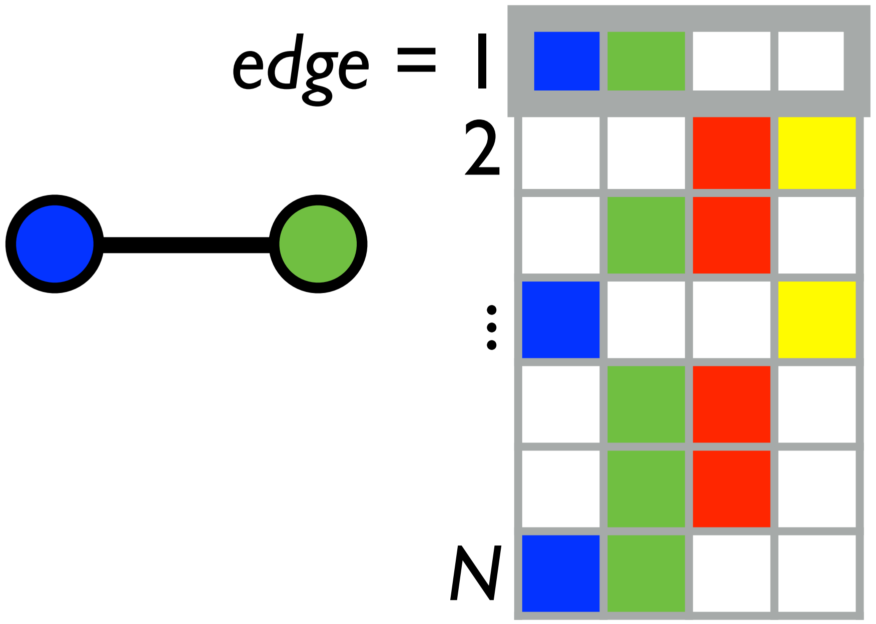



[Broderick, Jordan, Pitman 2013]

# Exchangeable probability functions

Q: Does every edge-exchangeable graph have an EVPF?

A: No. Counterexample:



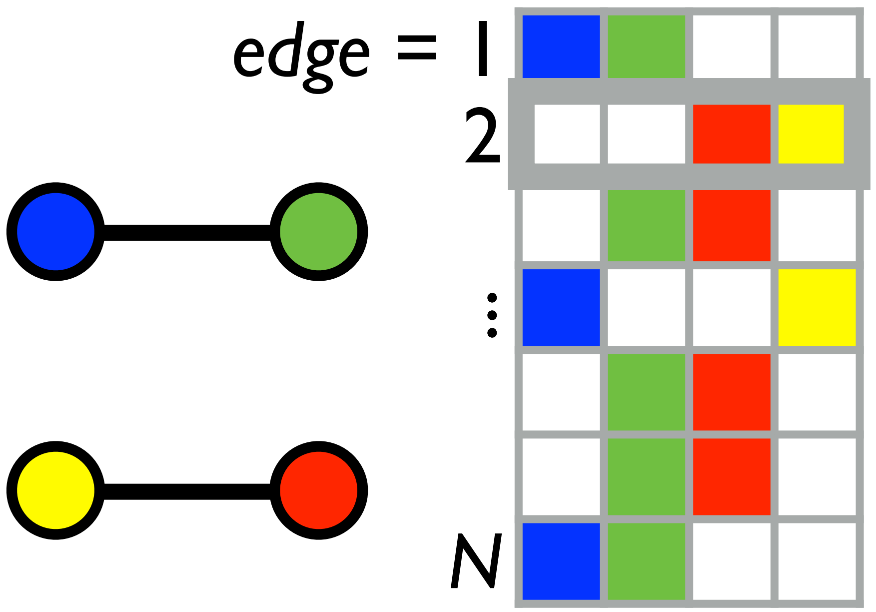



[Broderick, Jordan, Pitman 2013]

# Exchangeable probability functions

Q: Does every edge-exchangeable graph have an EVPF?

A: No. Counterexample:



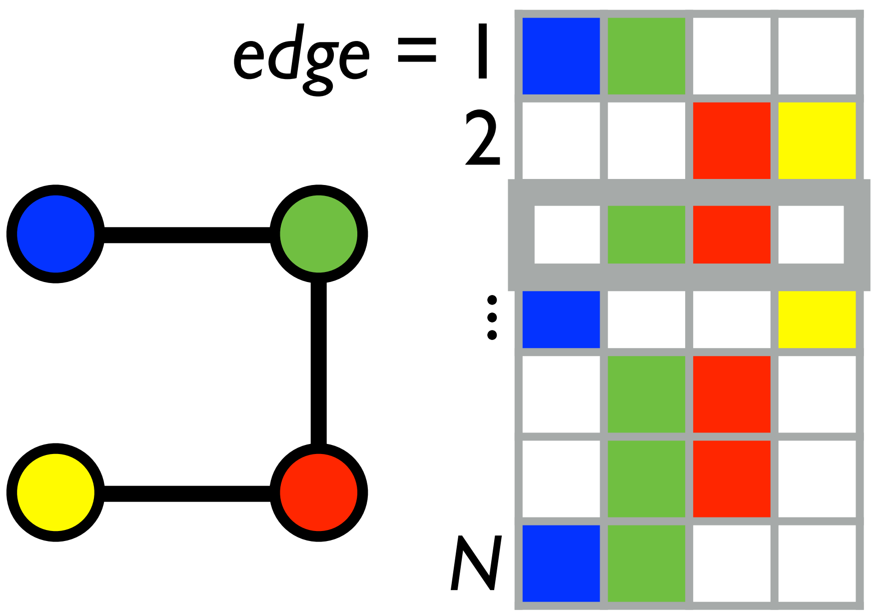



[Broderick, Jordan, Pitman 2013]

# Exchangeable probability functions

Q: Does every edge-exchangeable graph have an EVPF?

A: No. Counterexample:



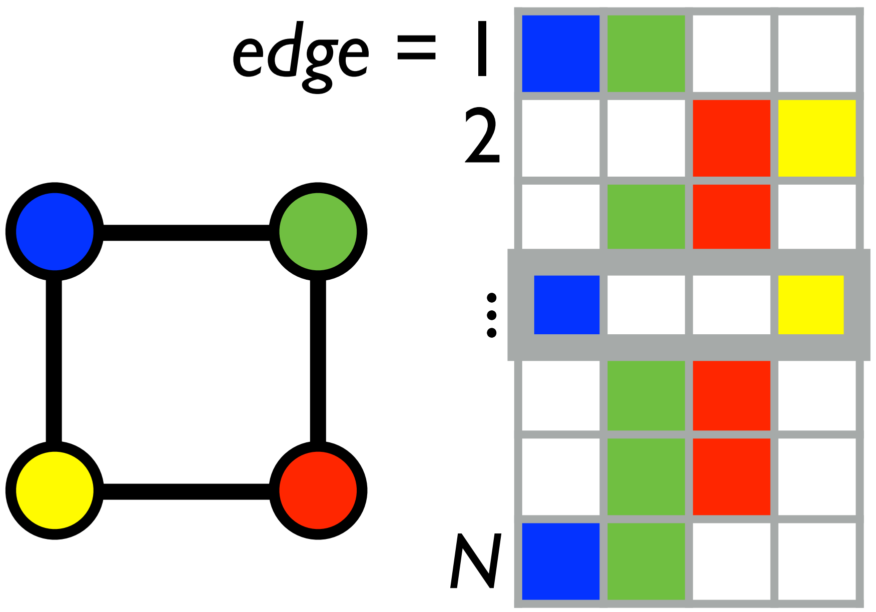



[Broderick, Jordan, Pitman 2013]

# Exchangeable probability functions

Q: Does every edge-exchangeable graph have an EVPF?

A: No. Counterexample:



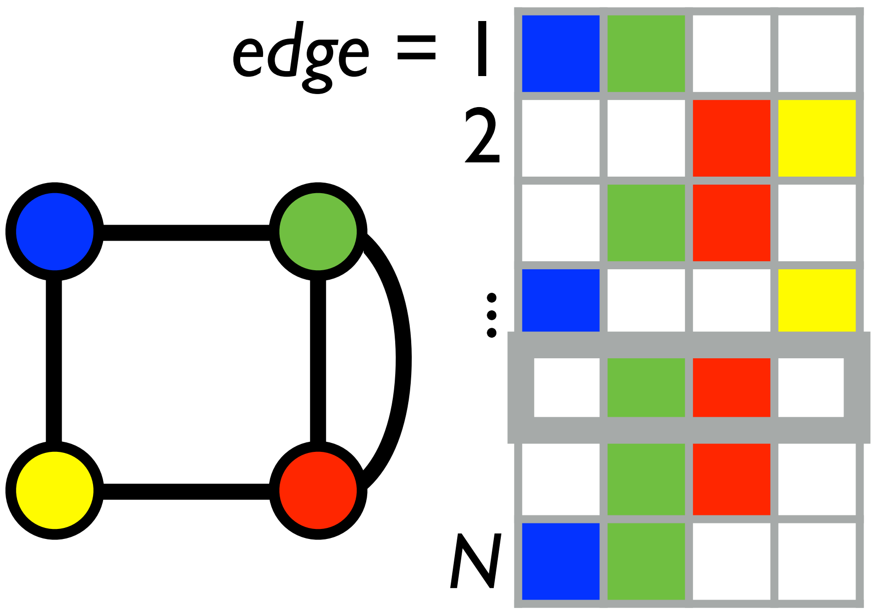



[Broderick, Jordan, Pitman 2013]

# Exchangeable probability functions

Q: Does every edge-exchangeable graph have an EVPF?

A: No. Counterexample:



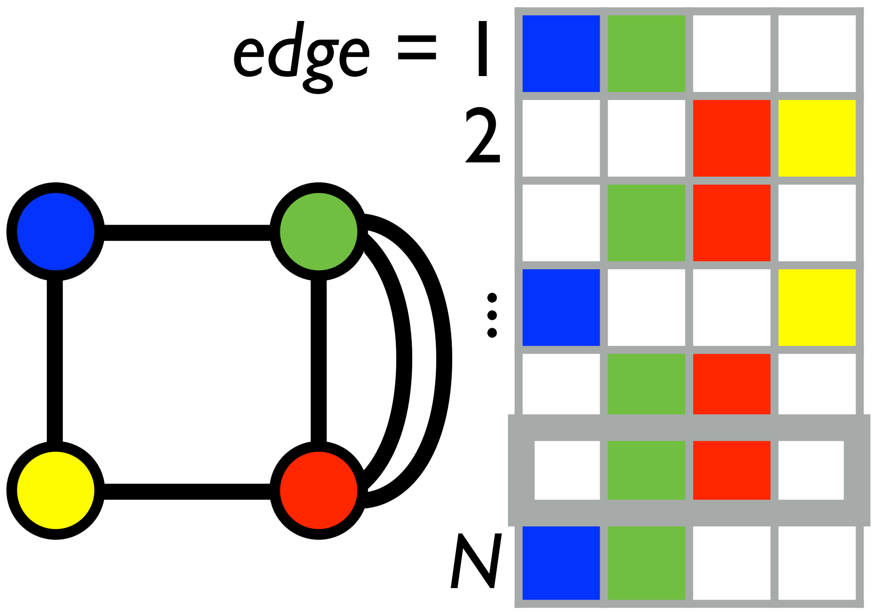



[Broderick, Jordan, Pitman 2013]

# Exchangeable probability functions

Q: Does every edge-exchangeable graph have an EVPF?

A: No. Counterexample:



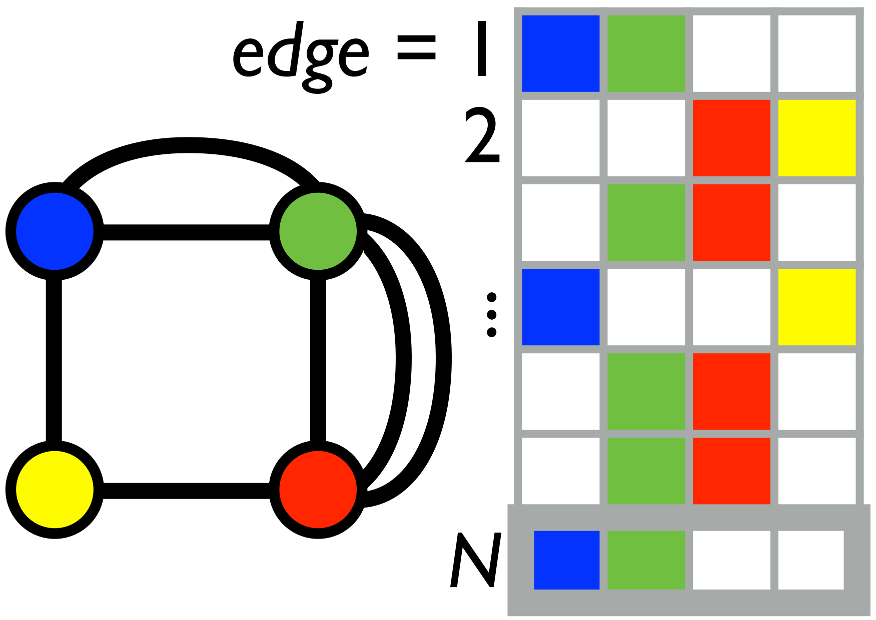



[Broderick, Jordan, Pitman 2013]

# Exchangeable probability functions

Q: Does every edge-exchangeable graph have an EVPF?

A: No. Counterexample:



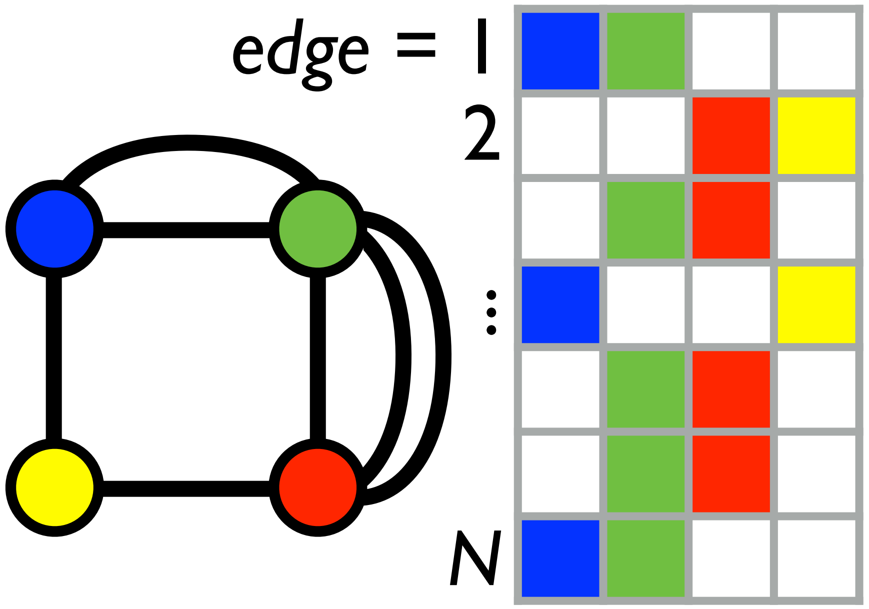



[Broderick, Jordan, Pitman 2013]

# Exchangeable probability functions

Q: Does every edge-exchangeable graph have an EVPF?

A: No. Counterexample:



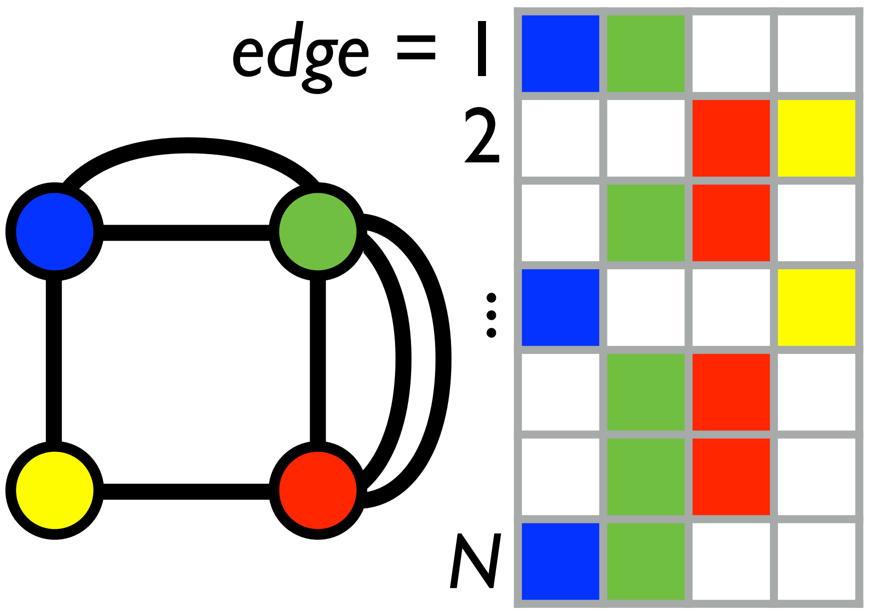



[Broderick, Jordan, Pitman 2013]

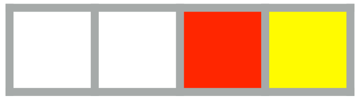
# Exchangeable probability functions

Q: Does every edge-exchangeable graph have an EVPF?

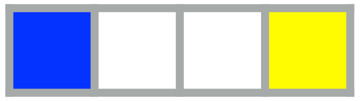
A: No. Counterexample:




$\mathbb{P}(\mathrm{row} =$  $) = p_1$

$\mathbb{P}(\mathrm{row} =$  $) = p_2$

$\mathbb{P}(\mathrm{row} =$  $) = p_3$

$\mathbb{P}(\mathrm{row} =$  $) = p_4$

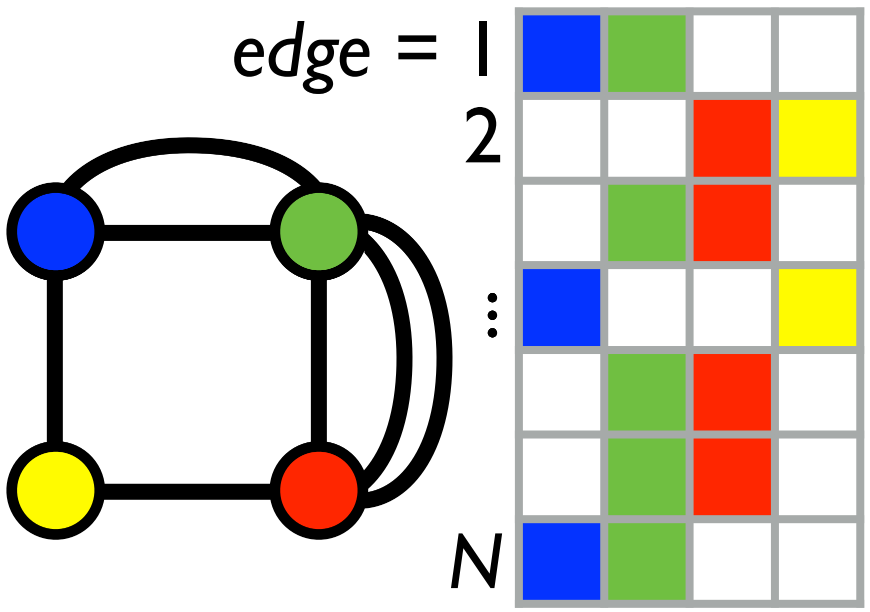


[Broderick, Jordan, Pitman 2013]

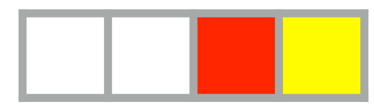
# Exchangeable probability functions

Q: Does every edge-exchangeable graph have an EVPF?

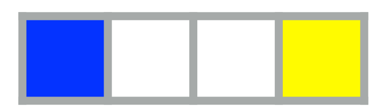
A: No. Counterexample:

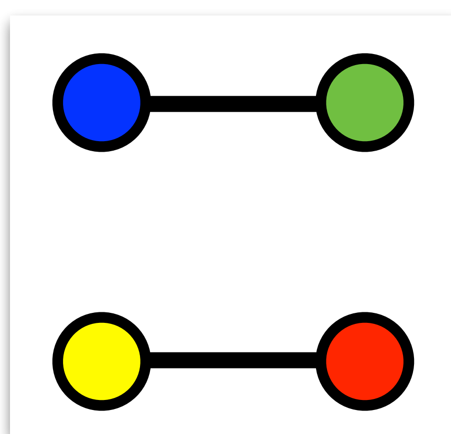


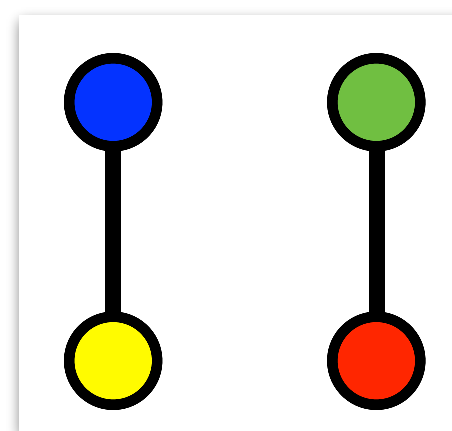

$\mathbb{P}(\text{row} =$  $) = p_1$

$\mathbb{P}(\text{row} =$  $) = p_2$

$\mathbb{P}(\text{row} =$  $) = p_3$

$\mathbb{P}(\text{row} =$  $) = p_4$





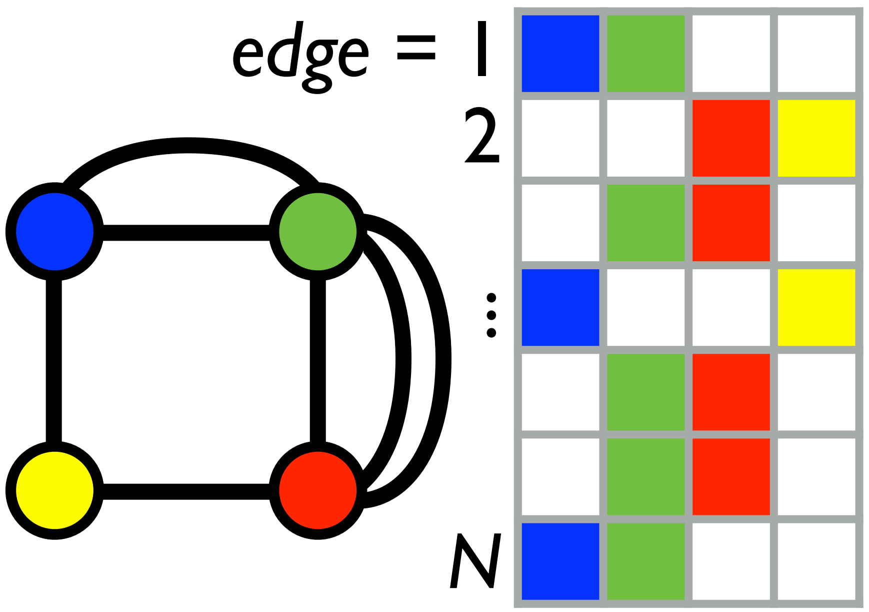


[Broderick, Jordan, Pitman 2013]

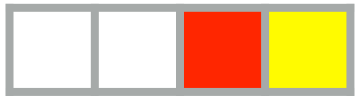
# Exchangeable probability functions

Q: Does every edge-exchangeable graph have an EVPF?

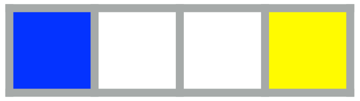
A: No. Counterexample:

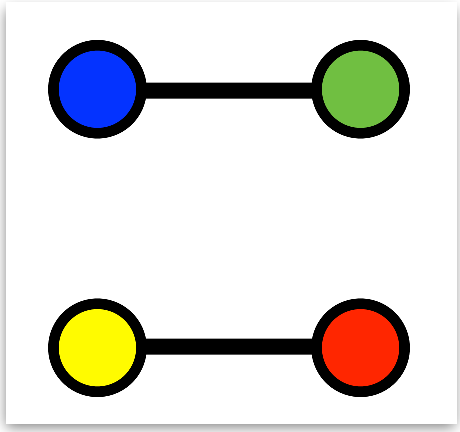
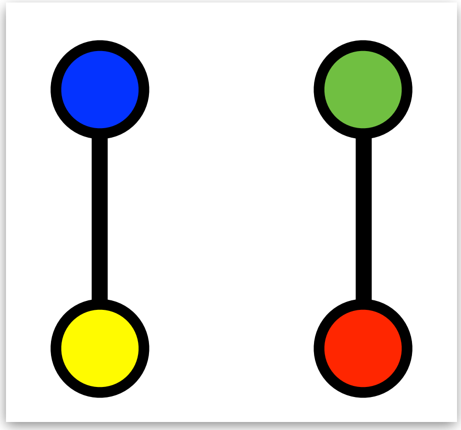



$\mathbb{P}($row $=$  $) = p_1$

$\mathbb{P}($row $=$  $) = p_2$

$\mathbb{P}($row $=$  $) = p_3$

$\mathbb{P}($row $=$  $) = p_4$

$\mathbb{P}($  $)$ $\mathbb{P}($  $)$

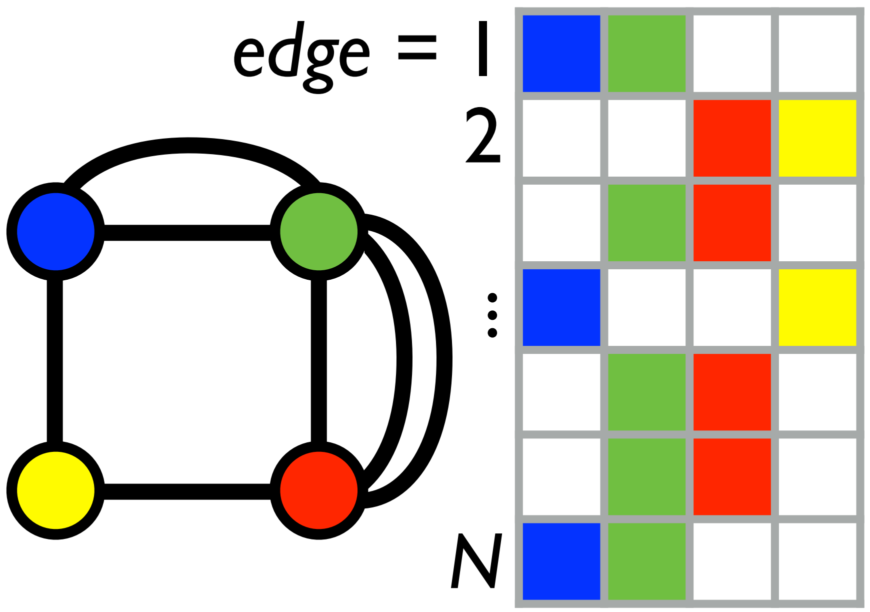


[Broderick, Jordan, Pitman 2013]

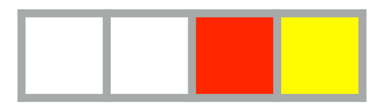
# Exchangeable probability functions

Q: Does every edge-exchangeable graph have an EVPF?

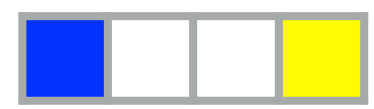
A: No. Counterexample:

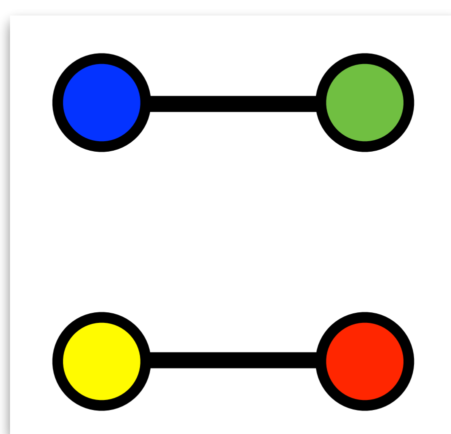
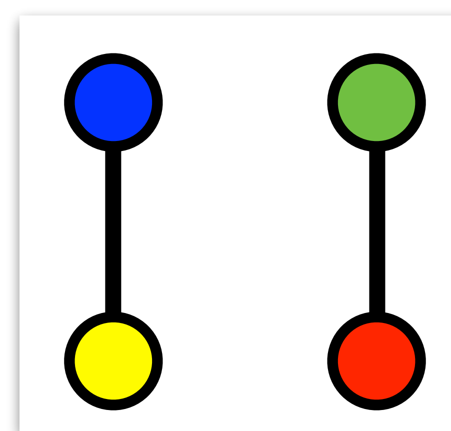



$\mathbb{P}(\text{row} =$  $) = p_1$

$\mathbb{P}(\text{row} =$  $) = p_2$

$\mathbb{P}(\text{row} =$  $) = p_3$

$\mathbb{P}(\text{row} =$  $) = p_4$

$\mathbb{P}($  $)$ $\quad$ $\mathbb{P}($  $)$

$p_1 p_2 \qquad p_3 p_4$

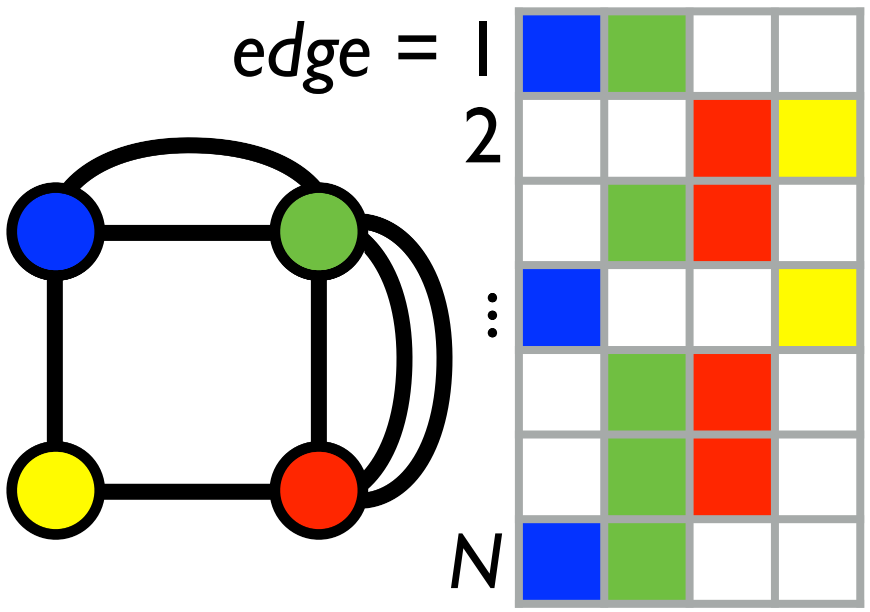


[Broderick, Jordan, Pitman 2013]

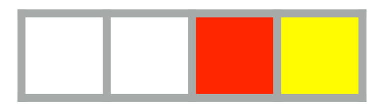
# Exchangeable probability functions

Q: Does every edge-exchangeable graph have an EVPF?

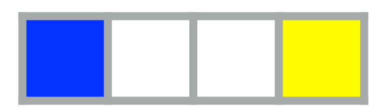
A: No. Counterexample:

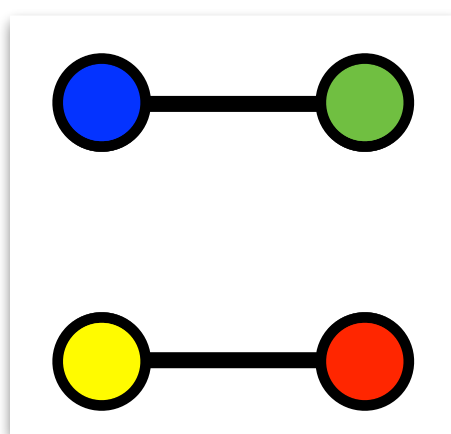
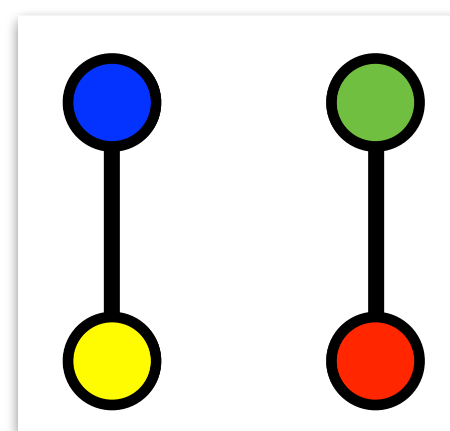



$\mathbb{P}(\text{row} =$  $) = p_1$

$\mathbb{P}(\text{row} =$  $) = p_2$

$\mathbb{P}(\text{row} =$  $) = p_3$

$\mathbb{P}(\text{row} =$  $) = p_4$

$\mathbb{P}($  $) \neq \mathbb{P}($  $)$

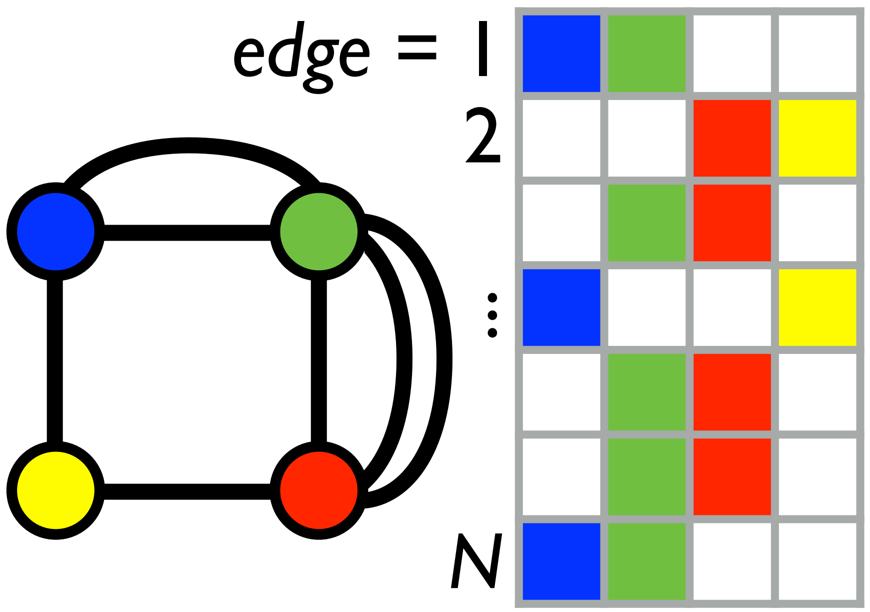
$$p_1 p_2 \neq p_3 p_4$$

[Broderick, Jordan, Pitman 2013]

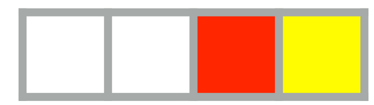
# Exchangeable probability functions

Q: Does every edge-exchangeable graph have an EVPF?

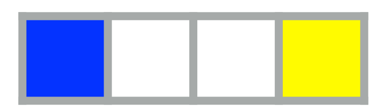
A: No. Counterexample:

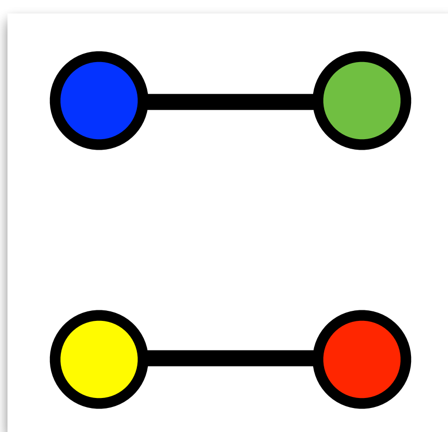
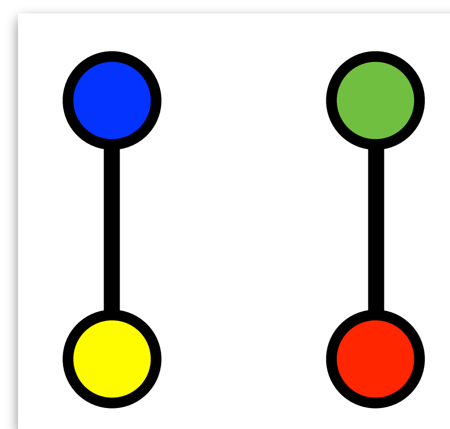



$\mathbb{P}(\text{row} =$  $) = p_1$

$\mathbb{P}(\text{row} =$  $) = p_2$

$\mathbb{P}(\text{row} =$  $) = p_3$

$\mathbb{P}(\text{row} =$  $) = p_4$

$\mathbb{P}($  $) \neq \mathbb{P}($  $)$

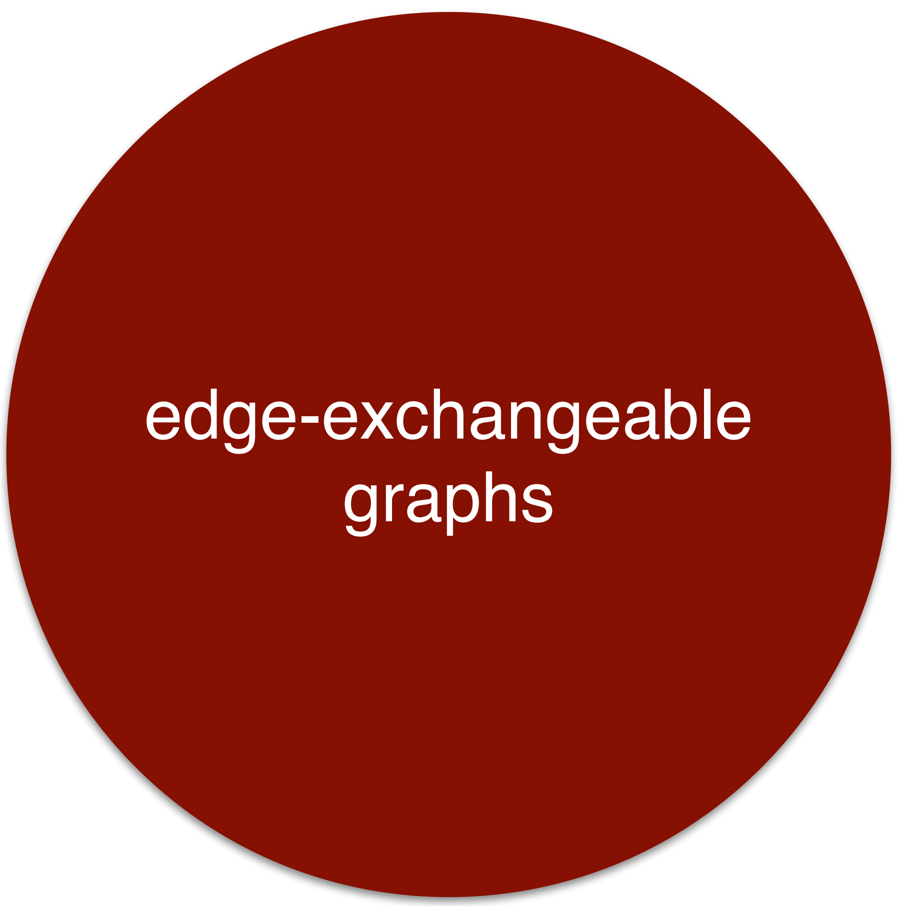
$$p_1 p_2 \neq p_3 p_4$$

Cor. Not every exchangeable feature allocation has an EFPF.



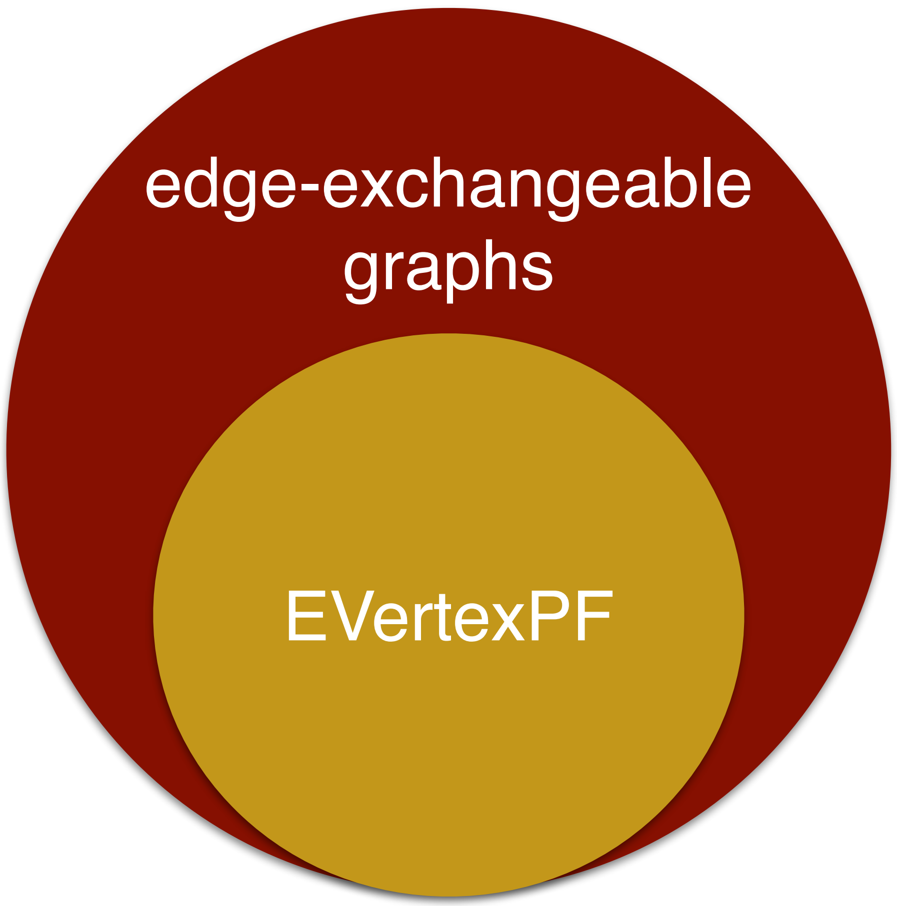
[Broderick, Jordan, Pitman 2013]

# What we know so far



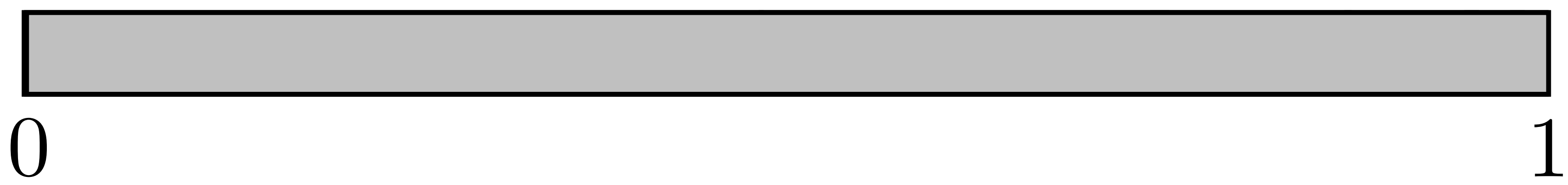

- Goal 1: characterization theorem for edge-exchangeable graphs
- Goal 2: sparsity theorem for edge-exchangeable graphs

# What we know so far



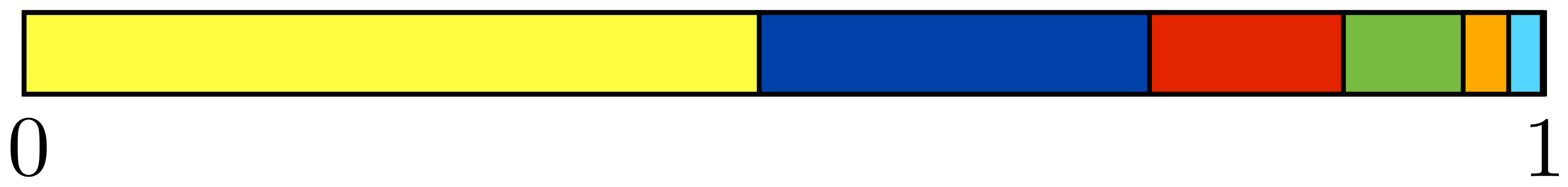

- Goal 1: characterization theorem for edge-exchangeable graphs
- Goal 2: sparsity theorem for edge-exchangeable graphs

# Clustering




[Kingman 1978]

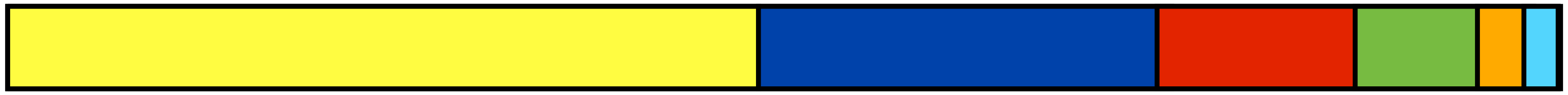
# Clustering

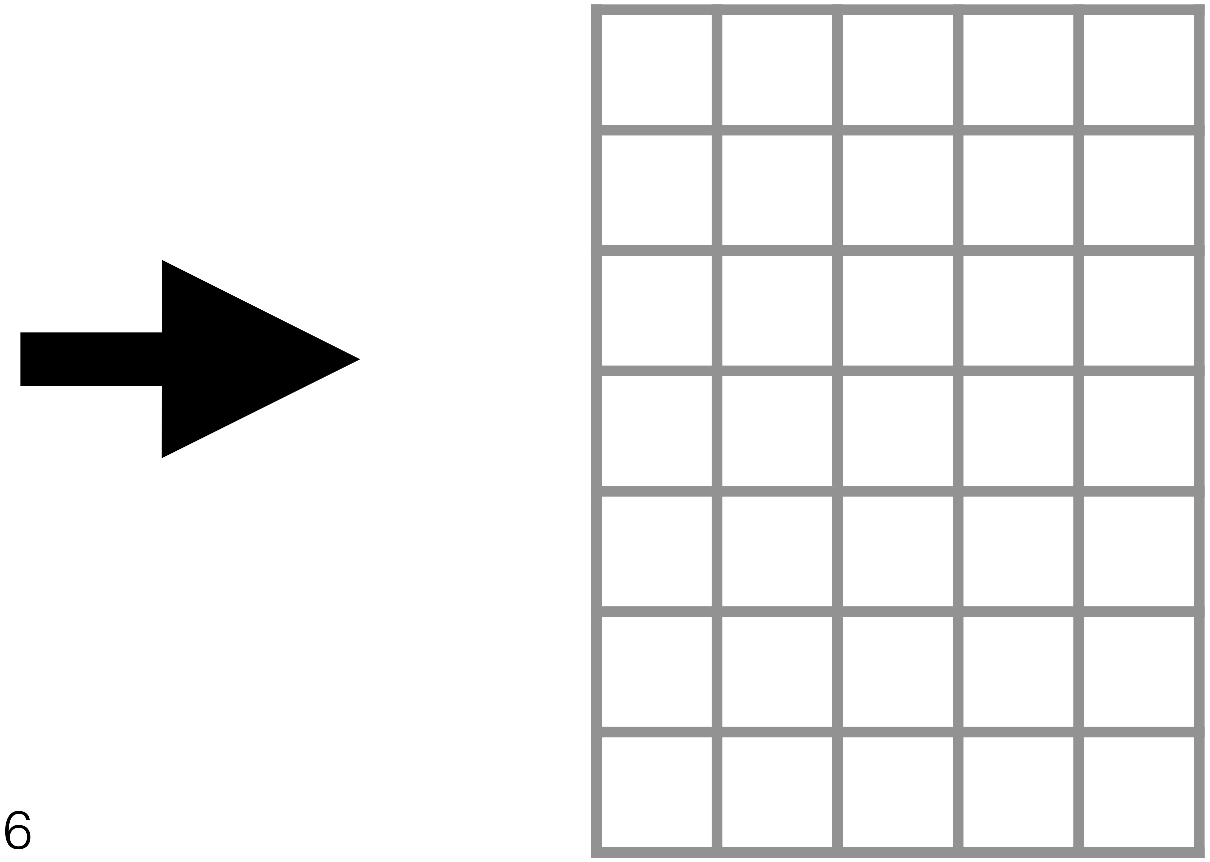





[Kingman 1978]

# Clustering

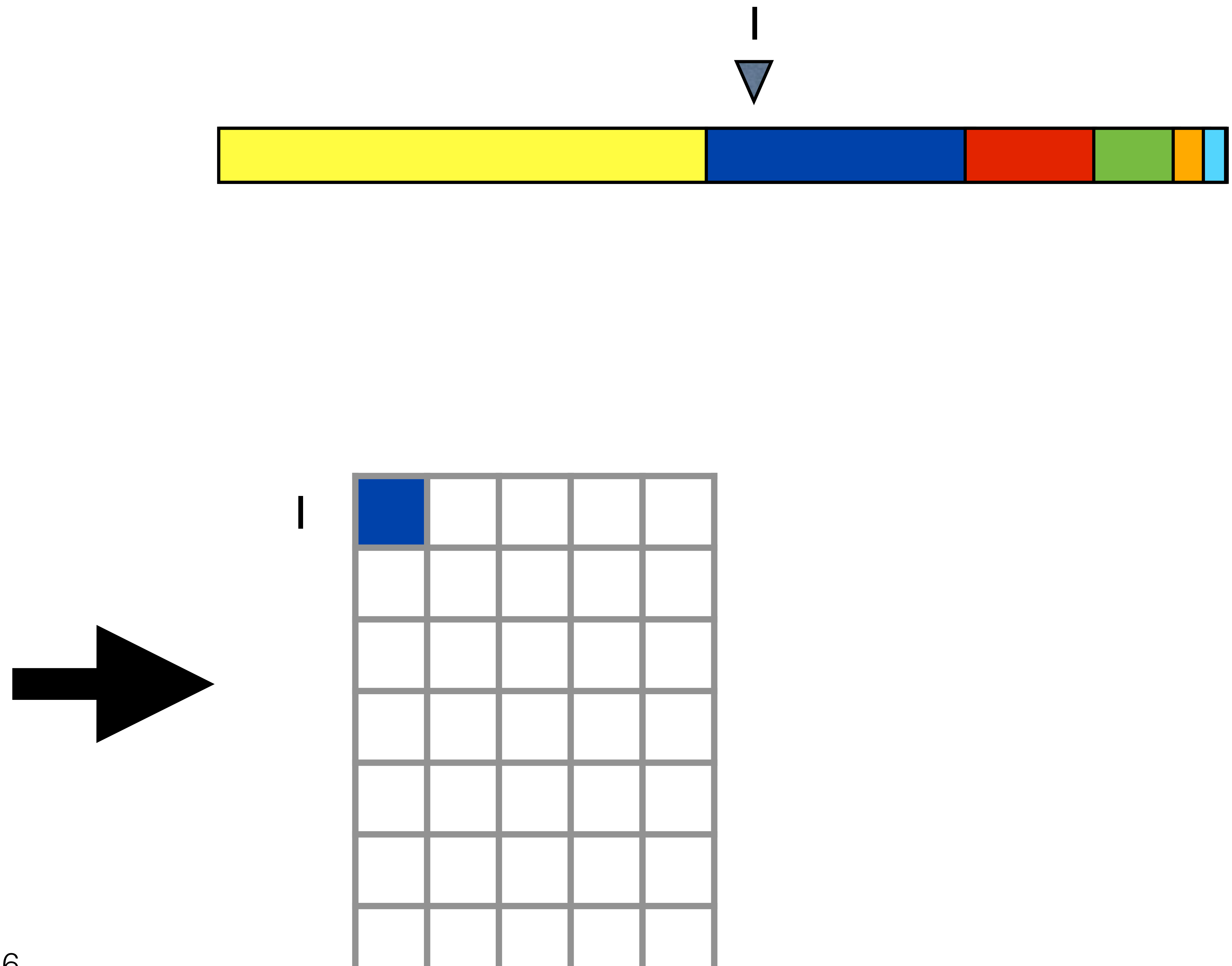






[Kingman 1978]

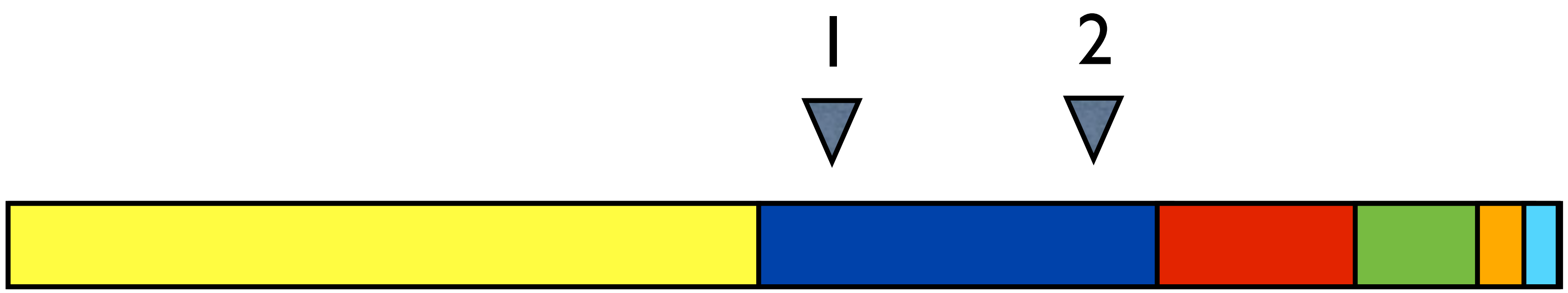
# Clustering

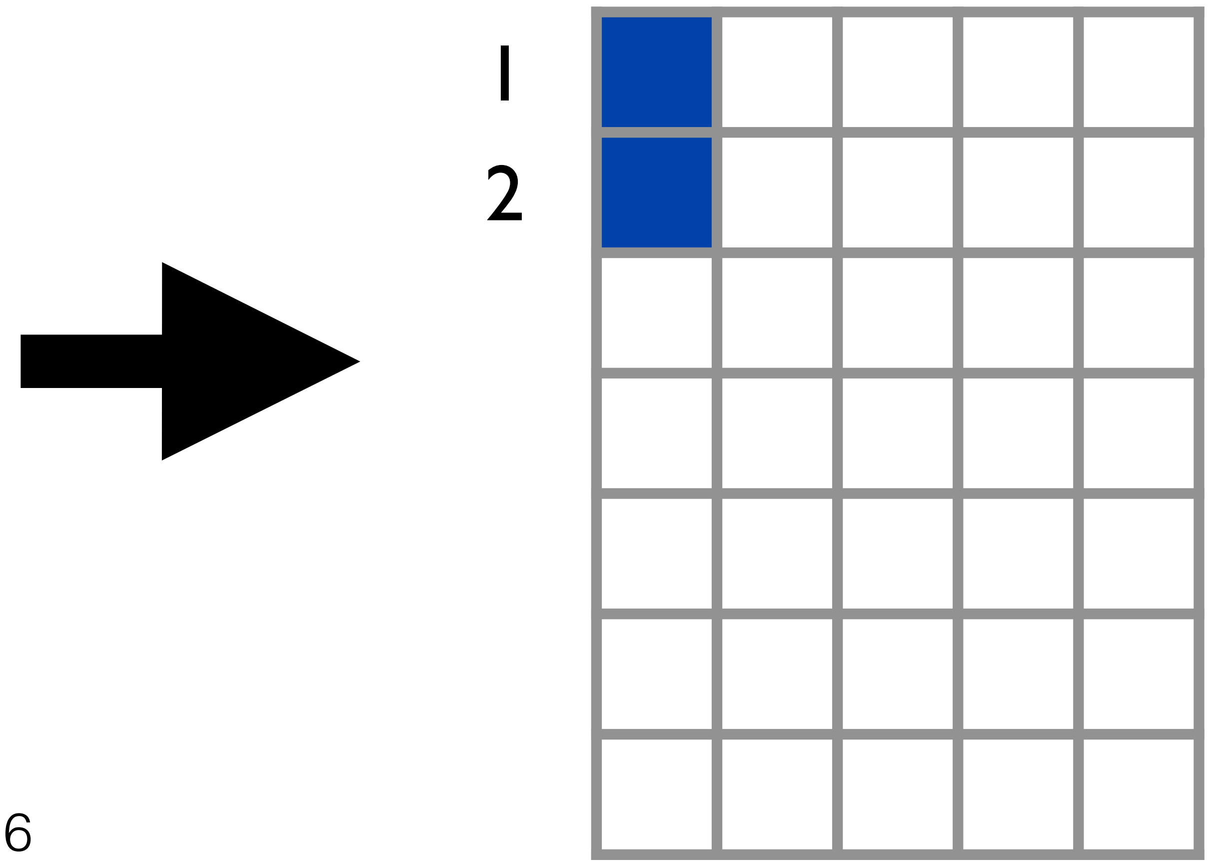





[Kingman 1978]

# Clustering

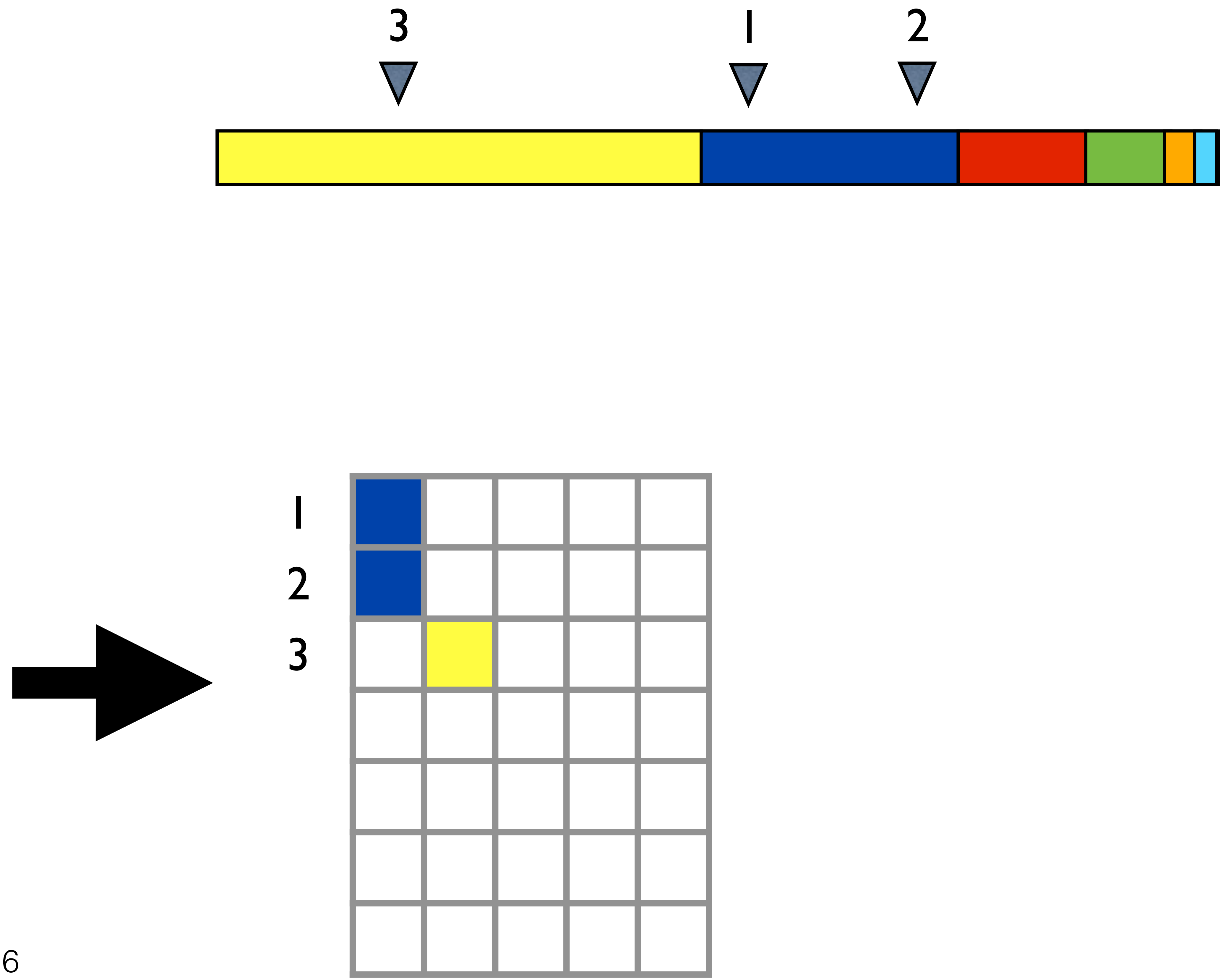





[Kingman 1978]

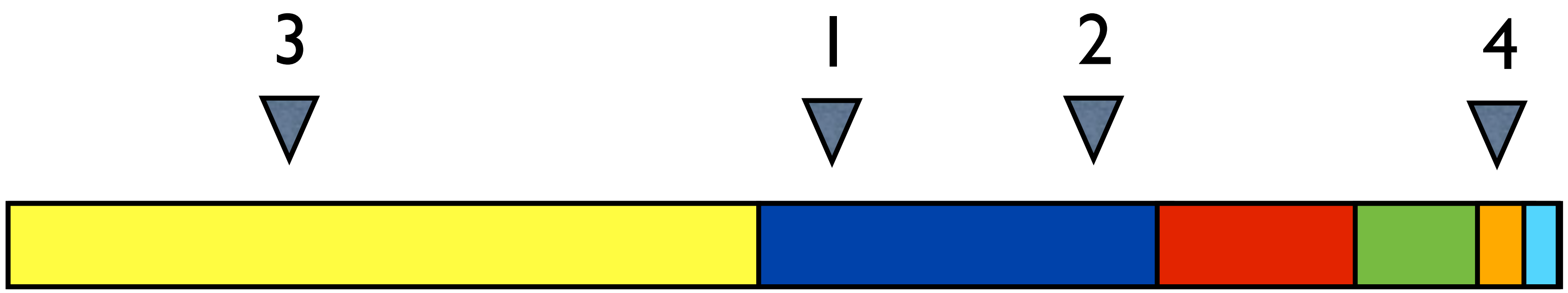
# Clustering

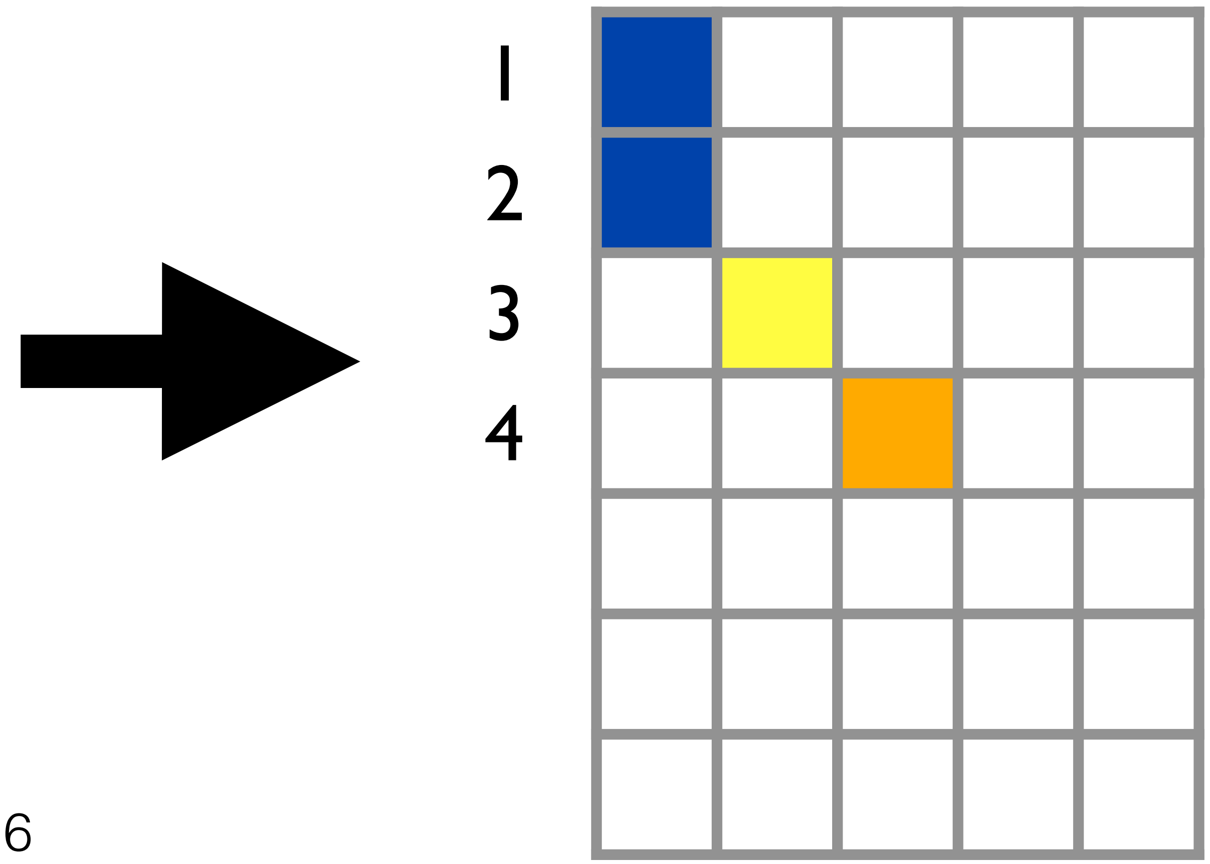





[Kingman 1978]

# Clustering

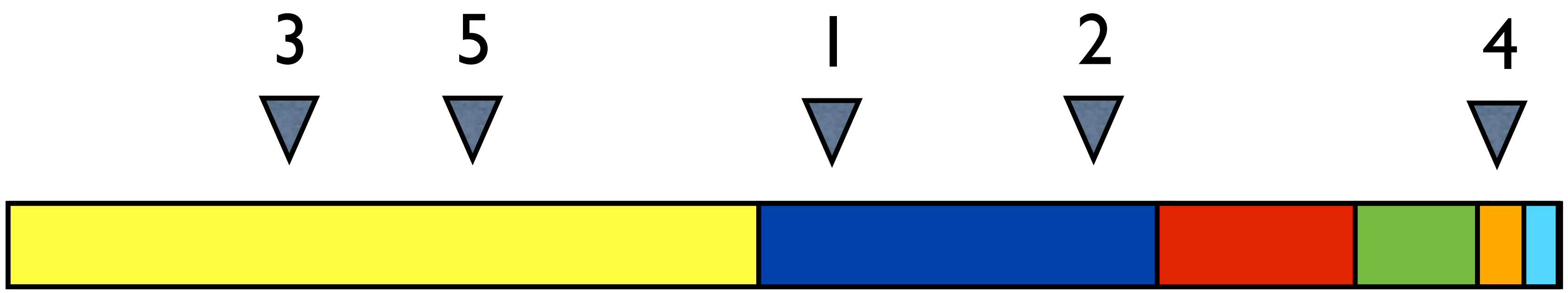


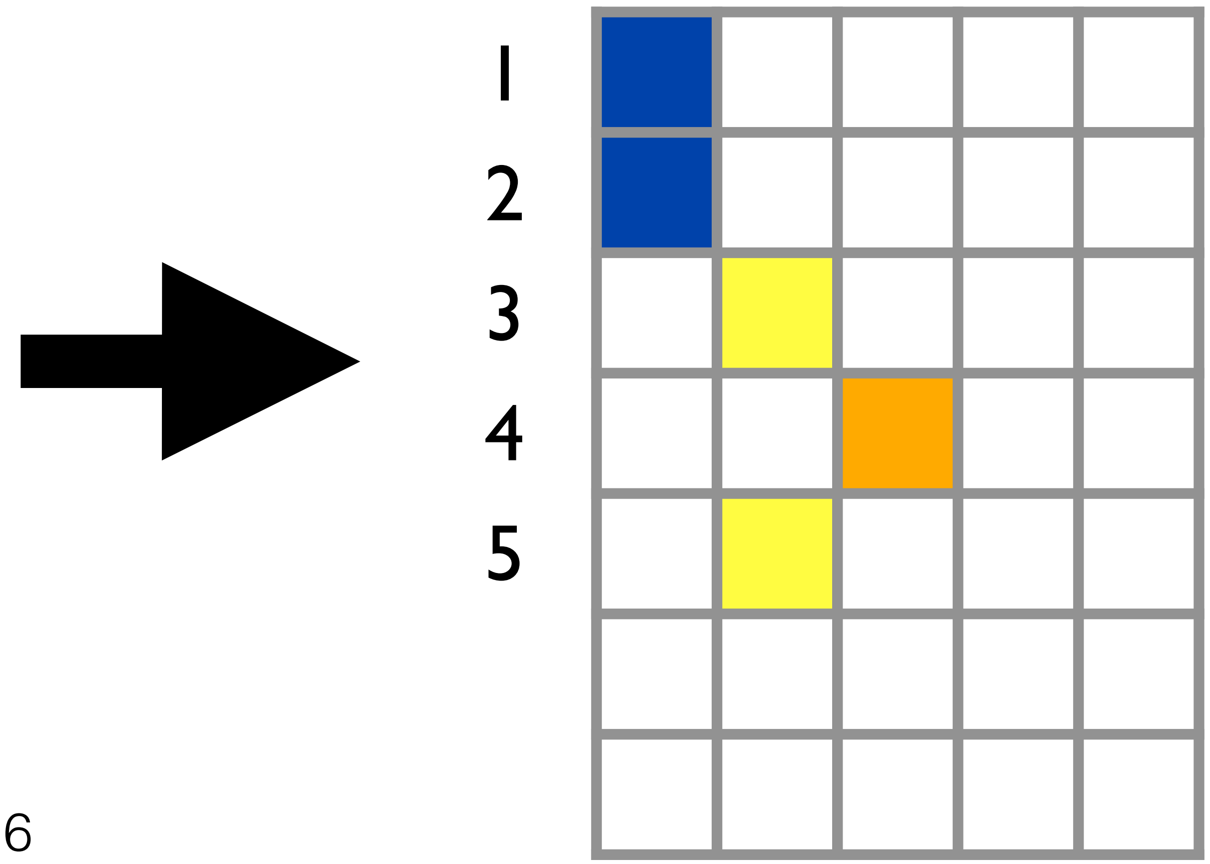






[Kingman 1978]

# Clustering

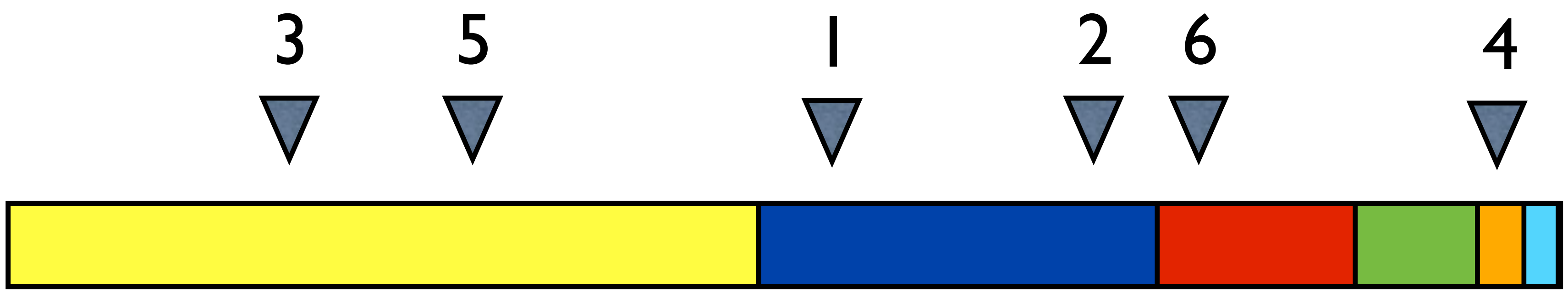


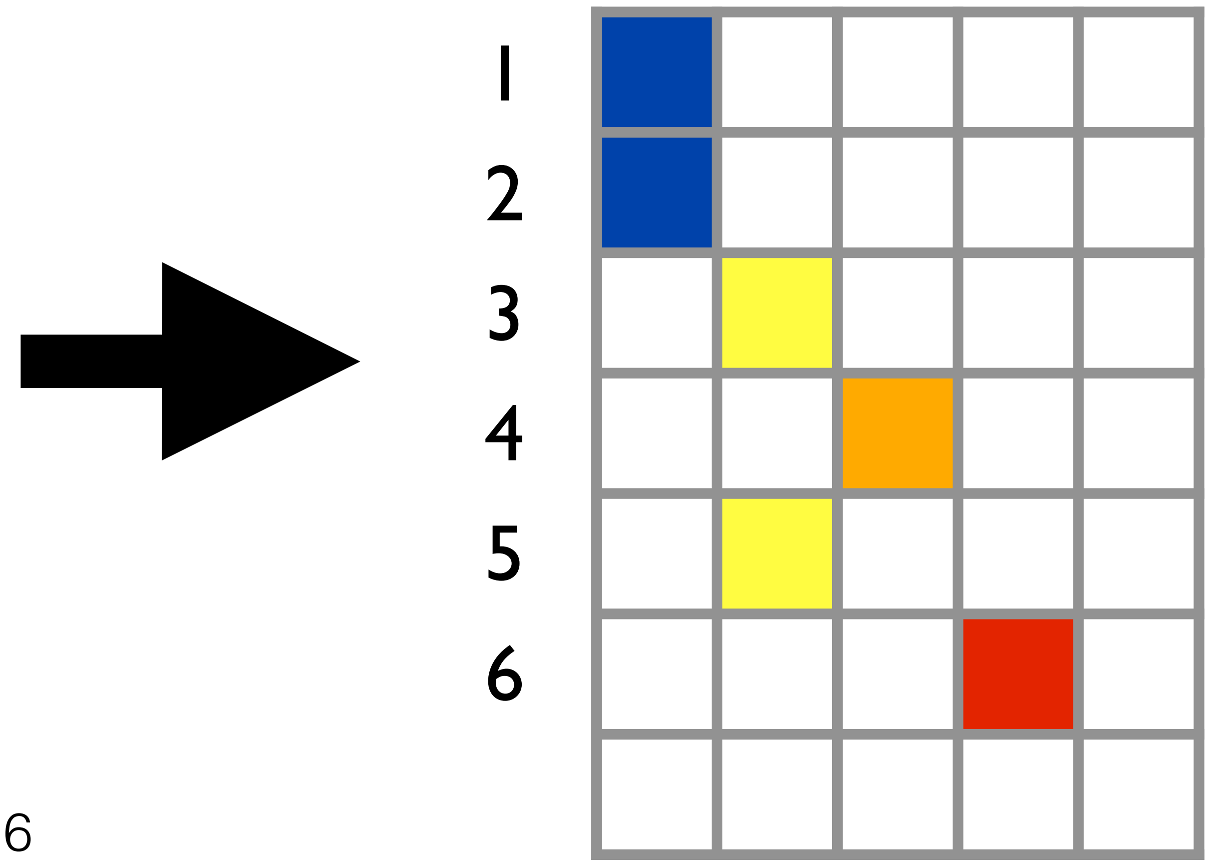






[Kingman 1978]

# Clustering

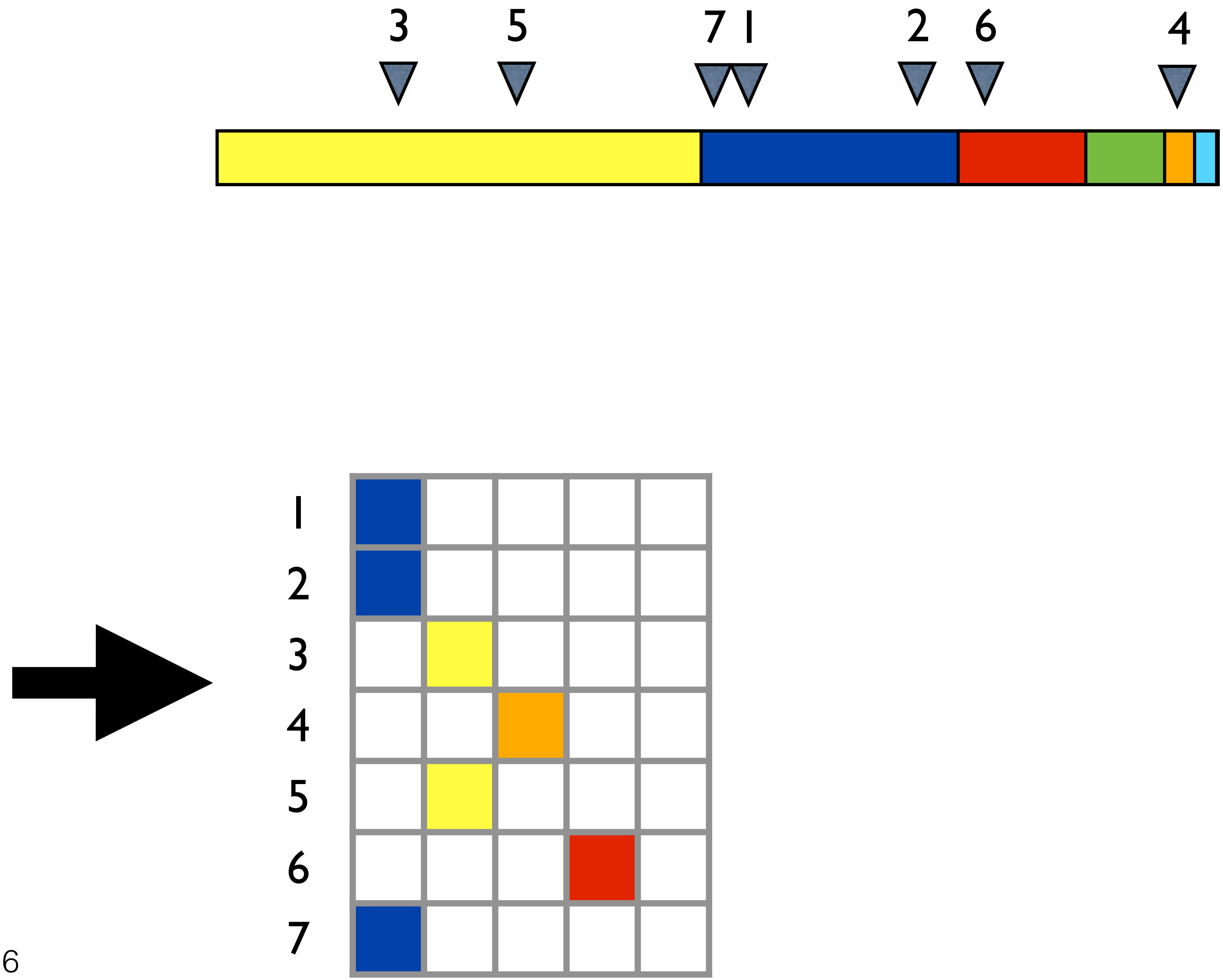





[Kingman 1978]

# Clustering

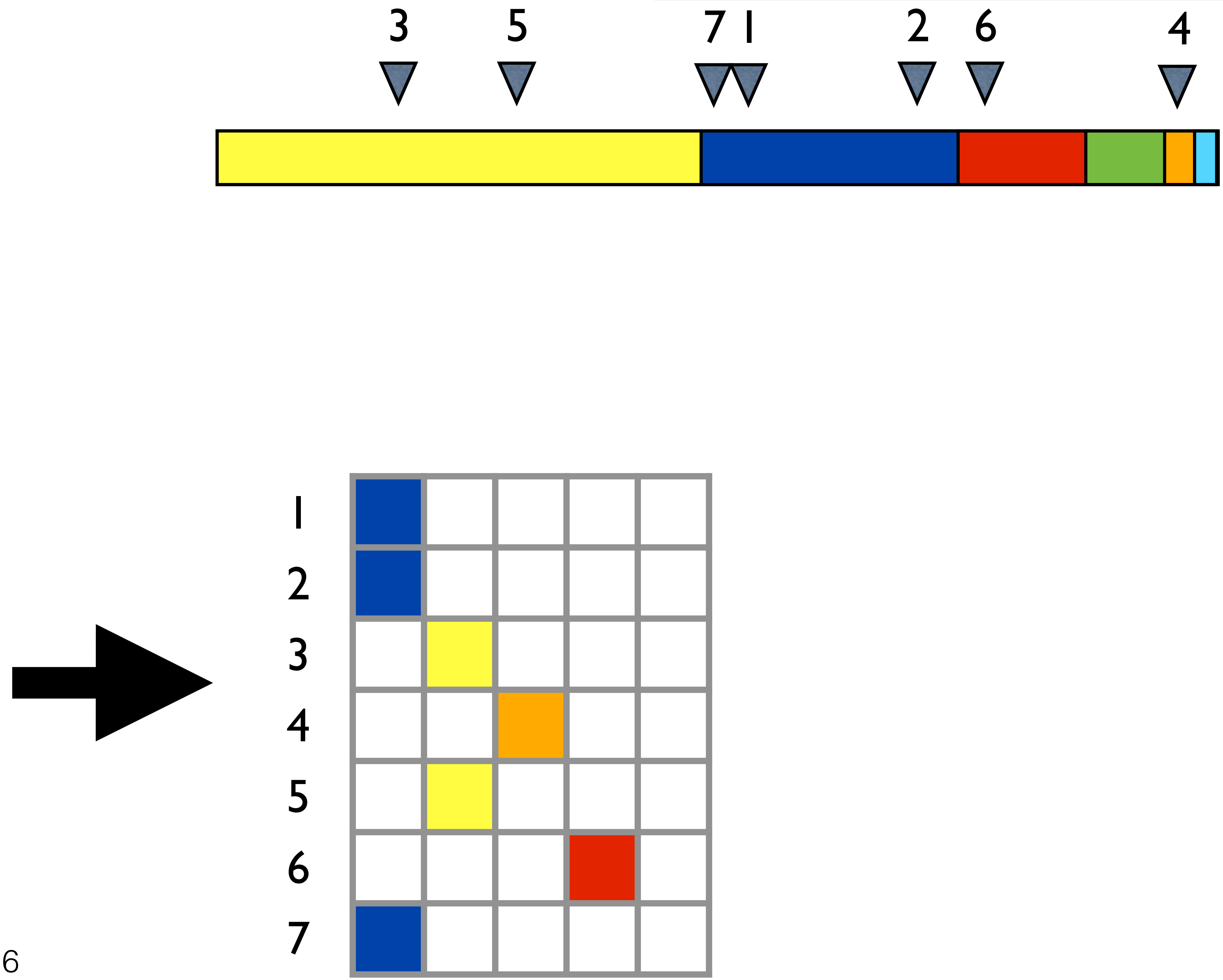





[Kingman 1978]

# Clustering

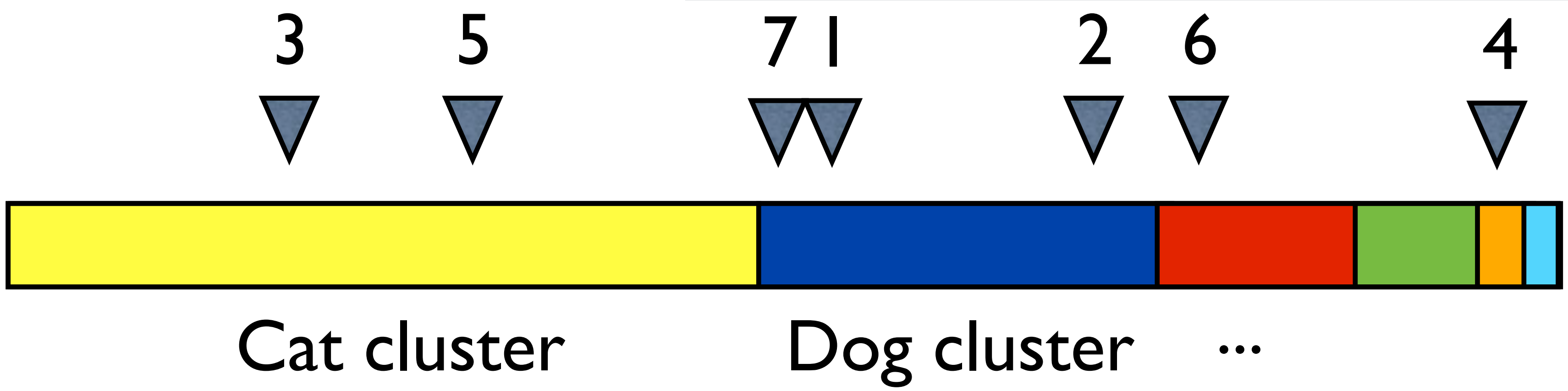
Thm (Kingman). A clustering is exchangeable iff it has a Kingman paintbox representation

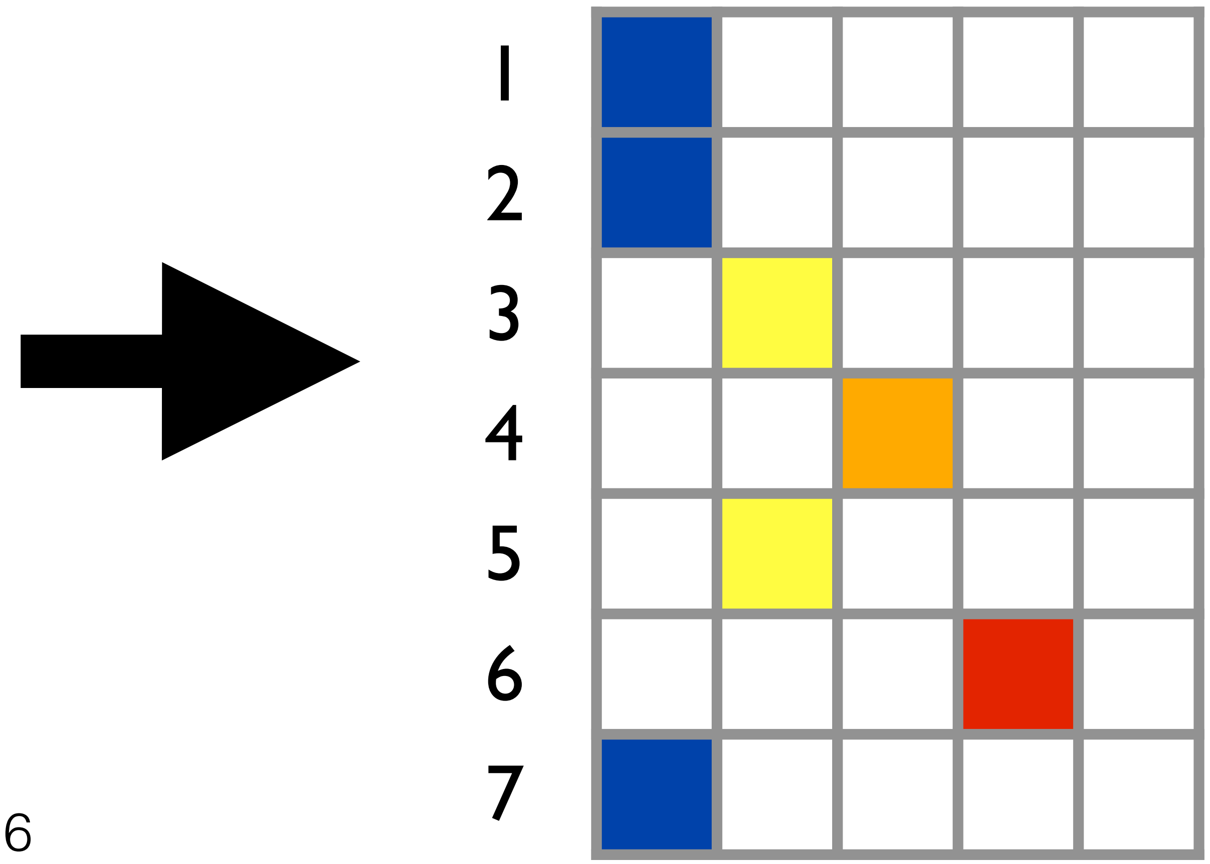



[Kingman 1978]

# Clustering

Thm (Kingman). A clustering is exchangeable iff it has a Kingman paintbox representation



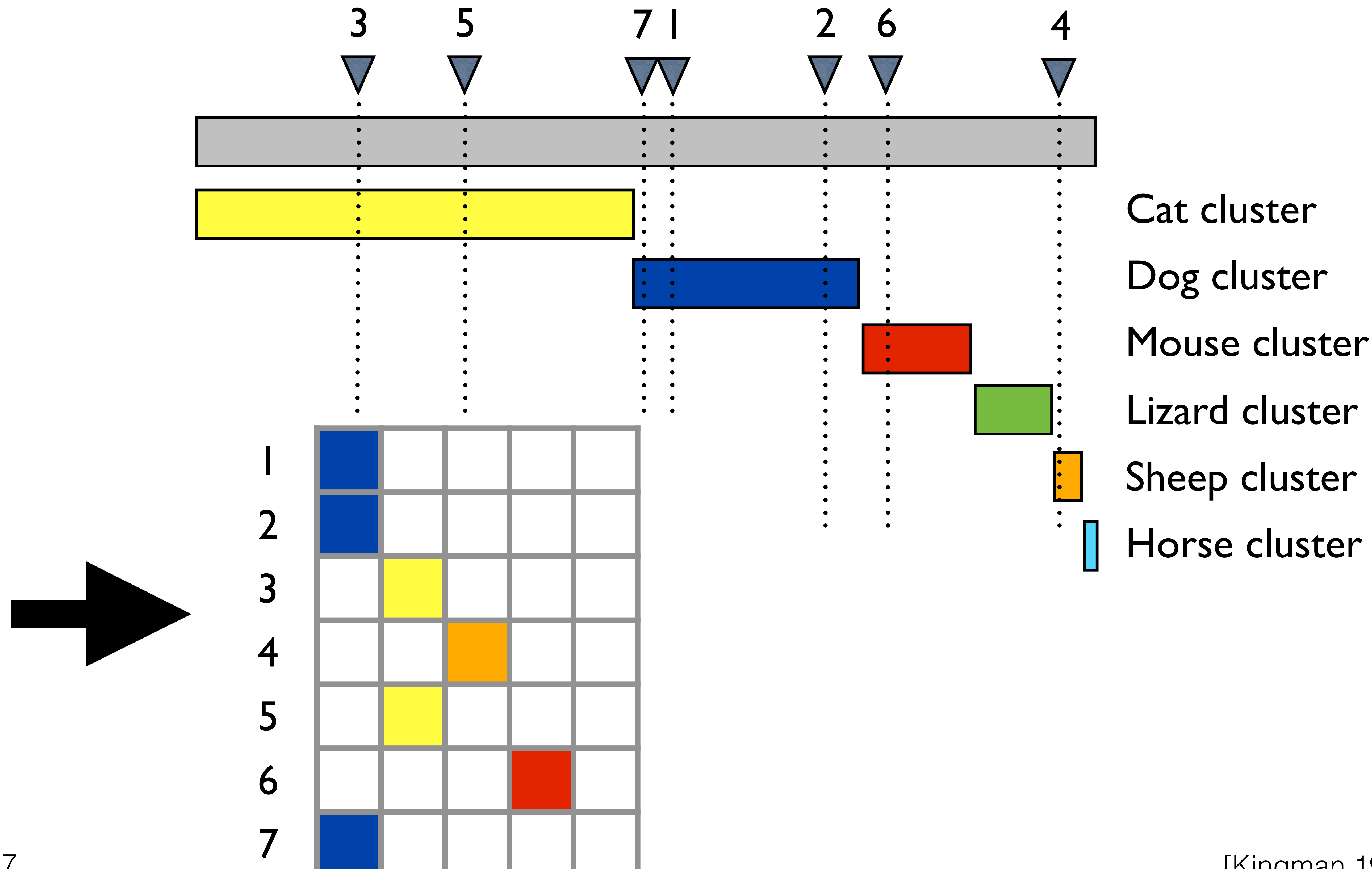






[Kingman 1978]

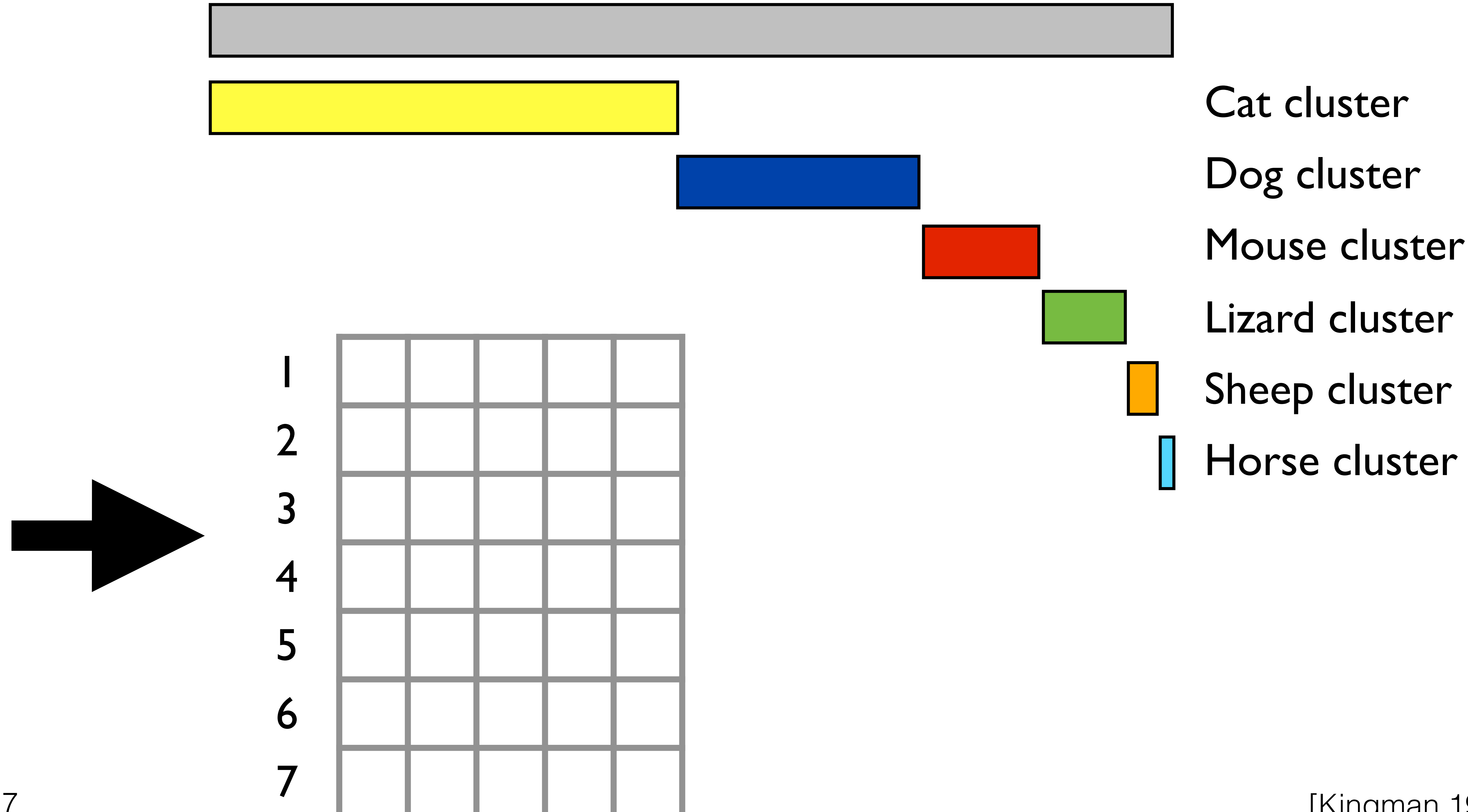
# Clustering

Thm (Kingman). A clustering is exchangeable iff it has a Kingman paintbox representation



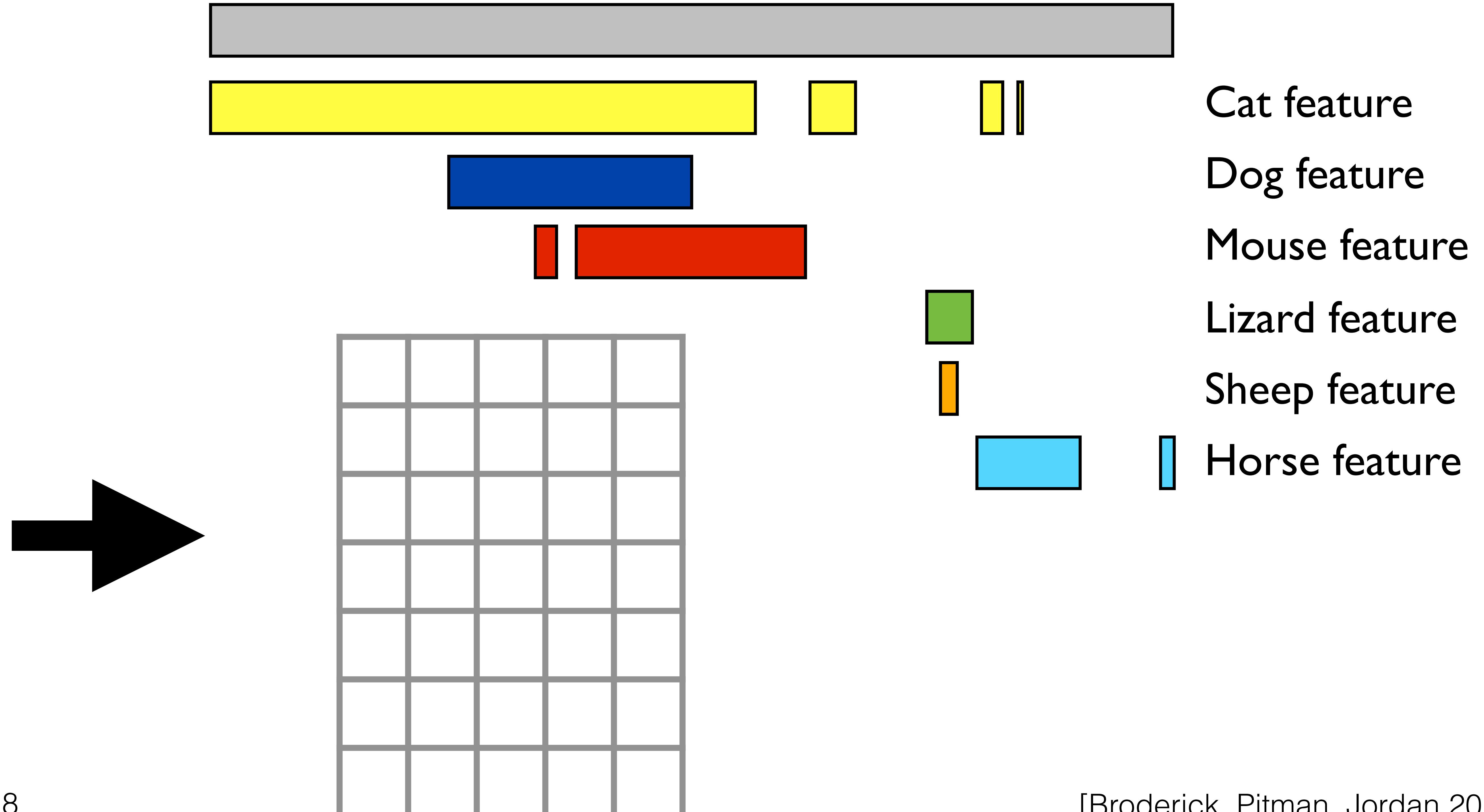



[Kingman 1978]





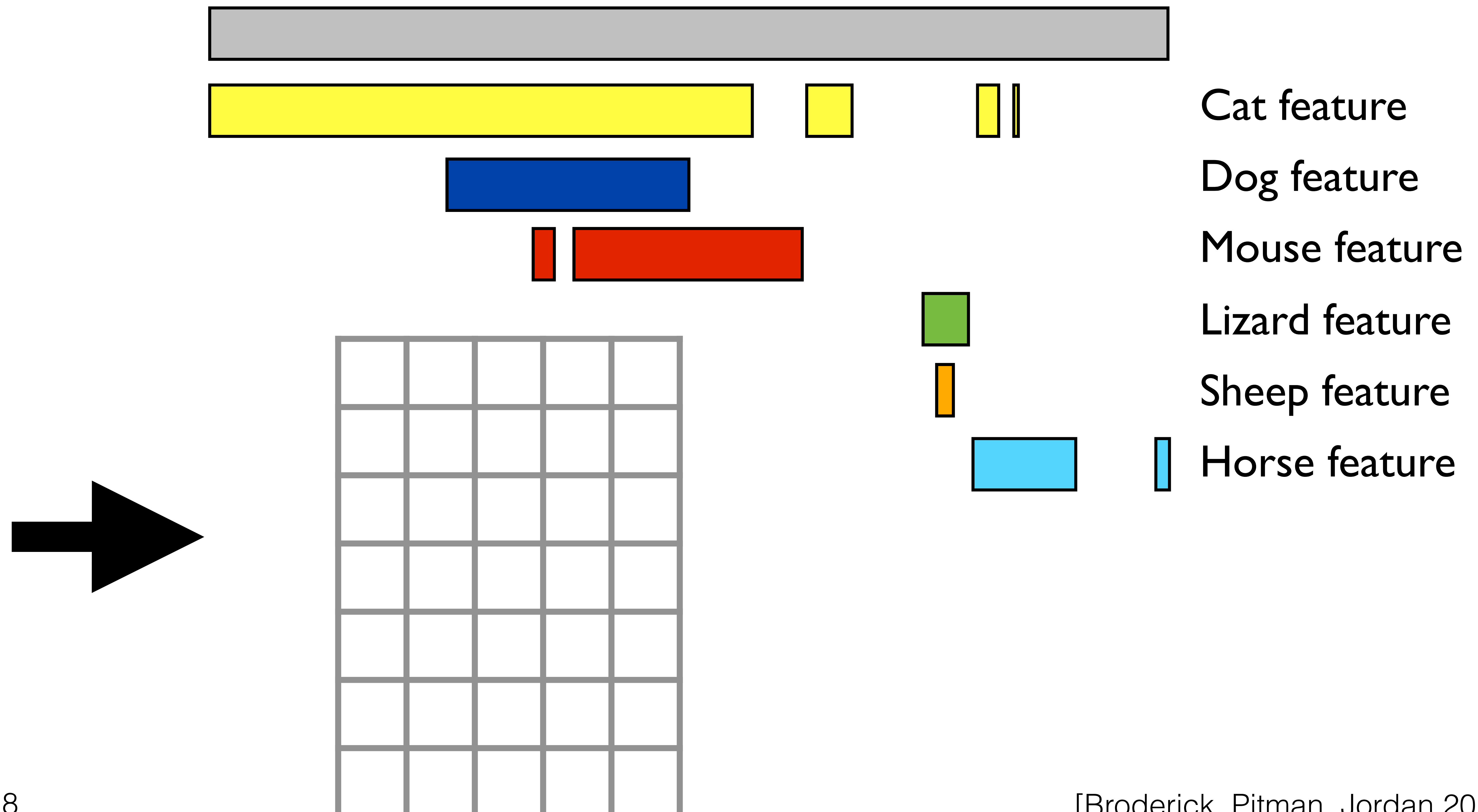

[Kingman 1978]





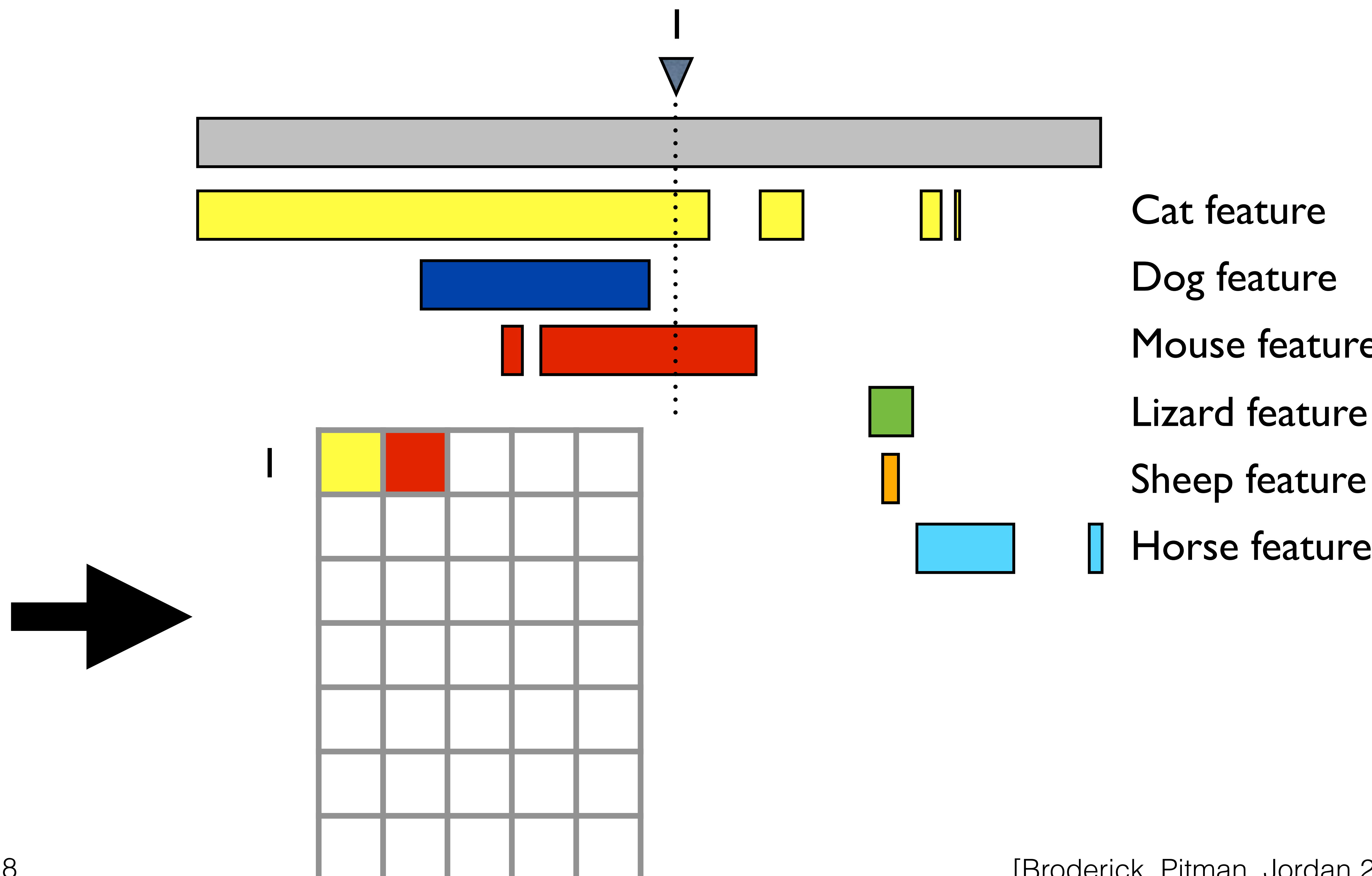

[Broderick, Pitman, Jordan 2013]

# Feature allocation




[Broderick, Pitman, Jordan 2013]

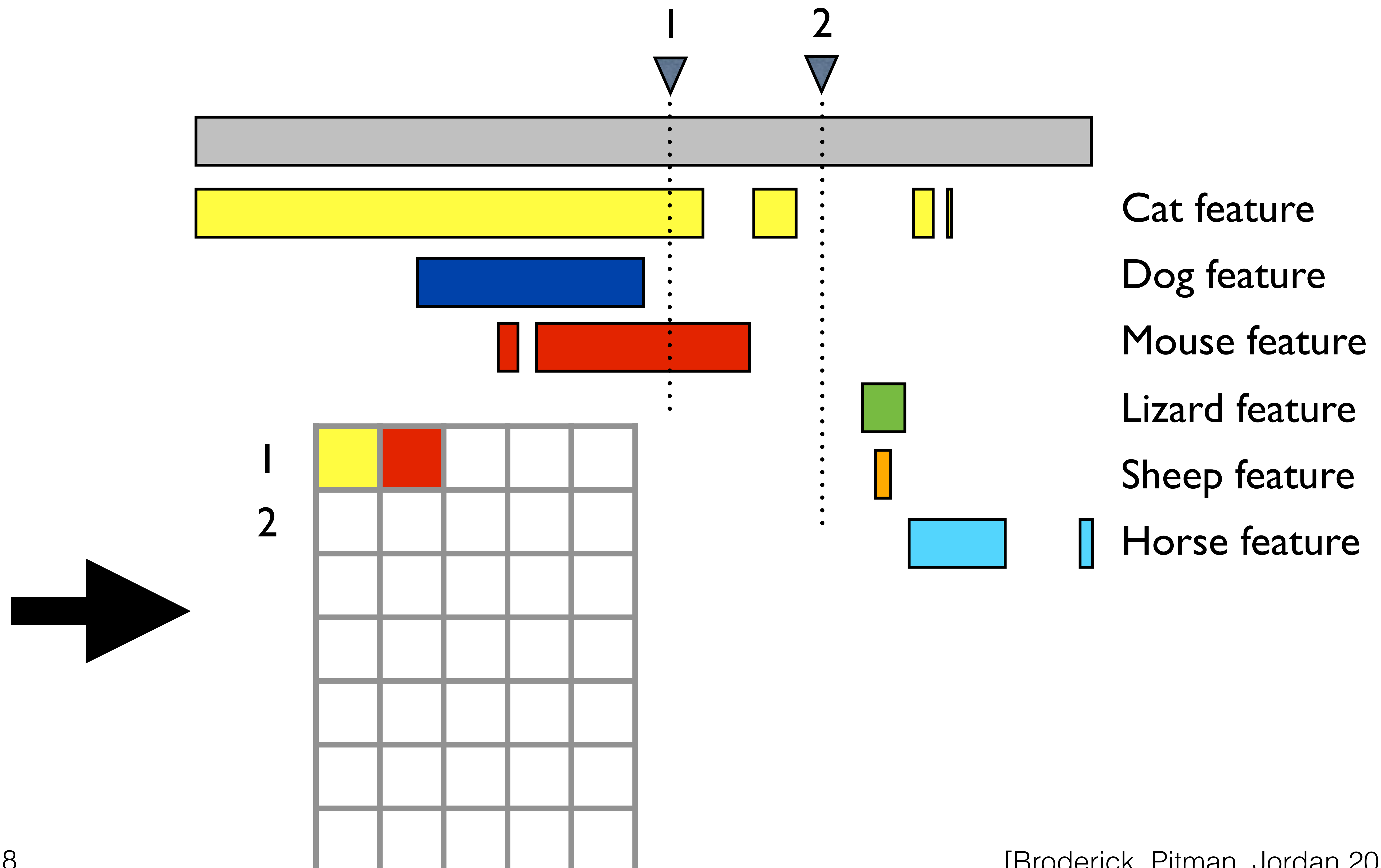
# Feature allocation






[Broderick, Pitman, Jordan 2013]

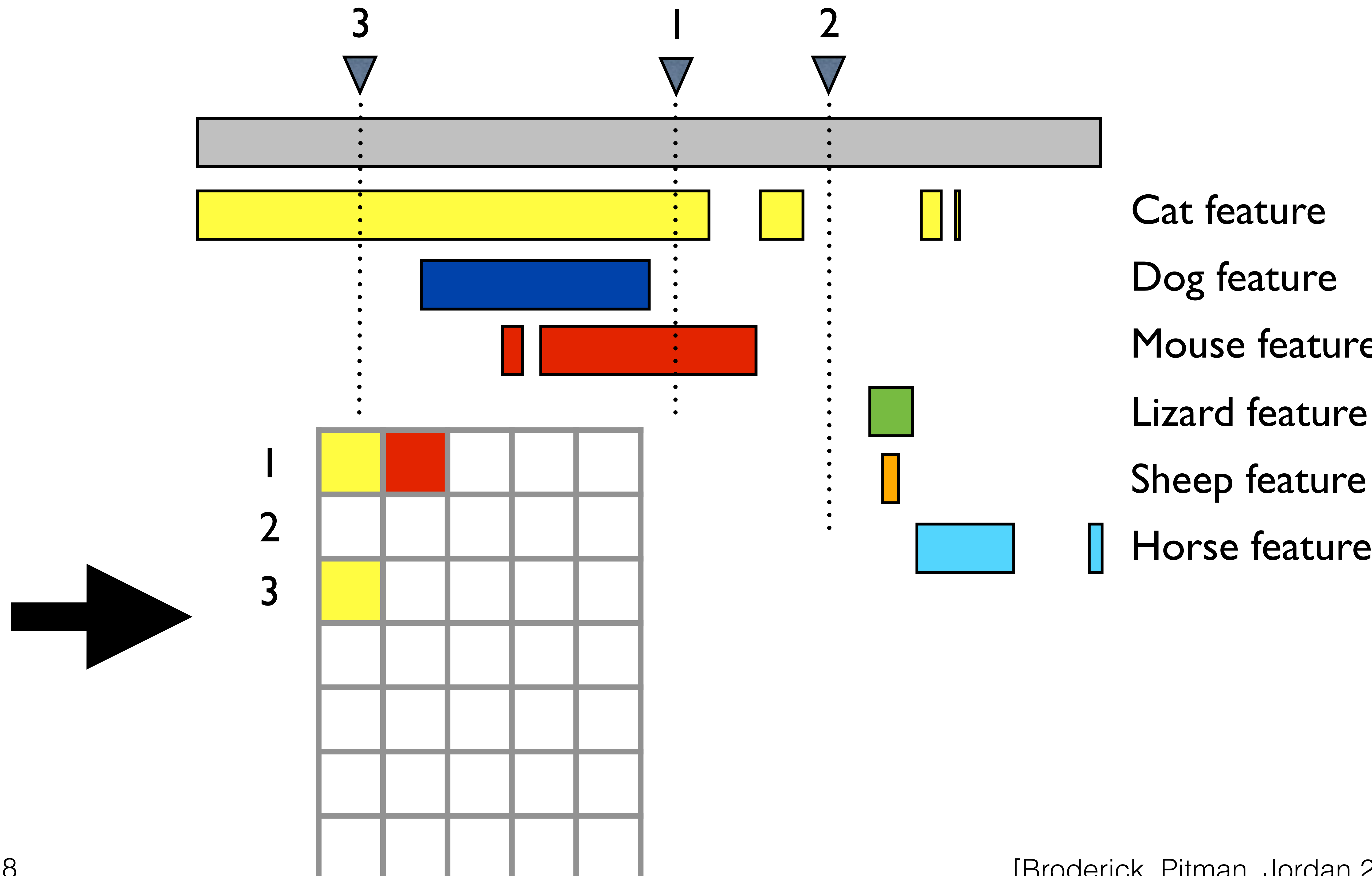
# Feature allocation





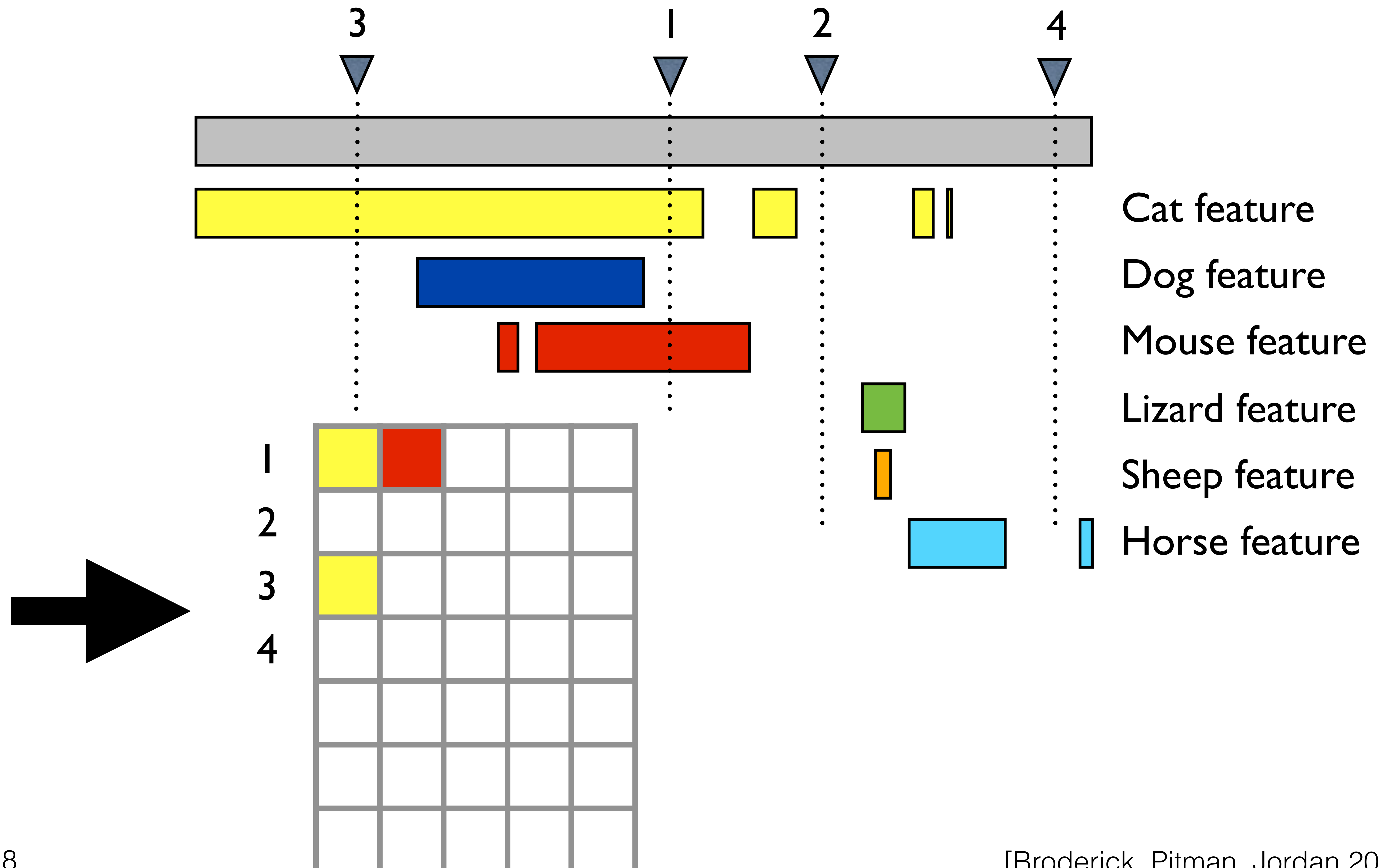

[Broderick, Pitman, Jordan 2013]

# Feature allocation




[Broderick, Pitman, Jordan 2013]

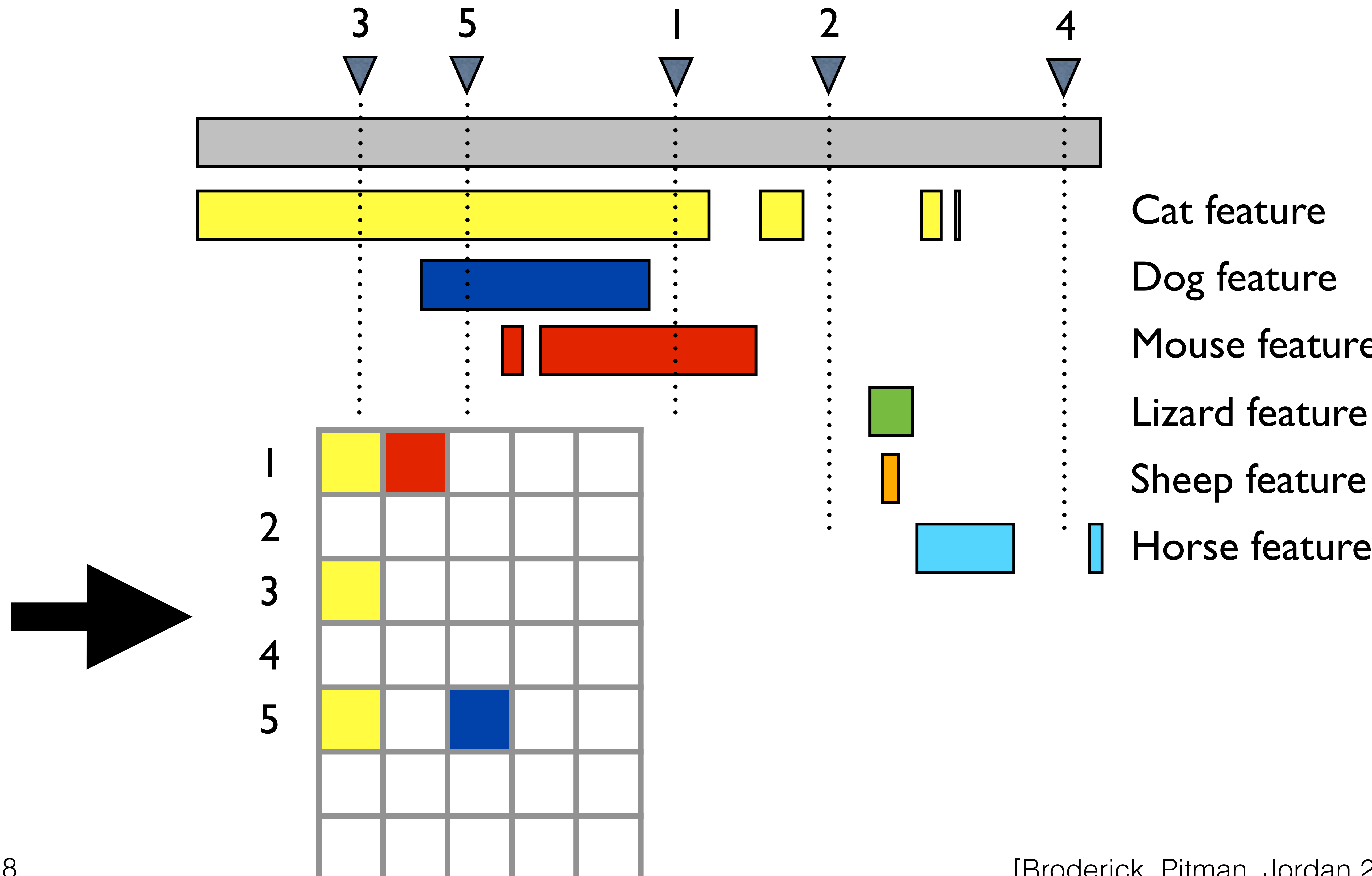
# Feature allocation





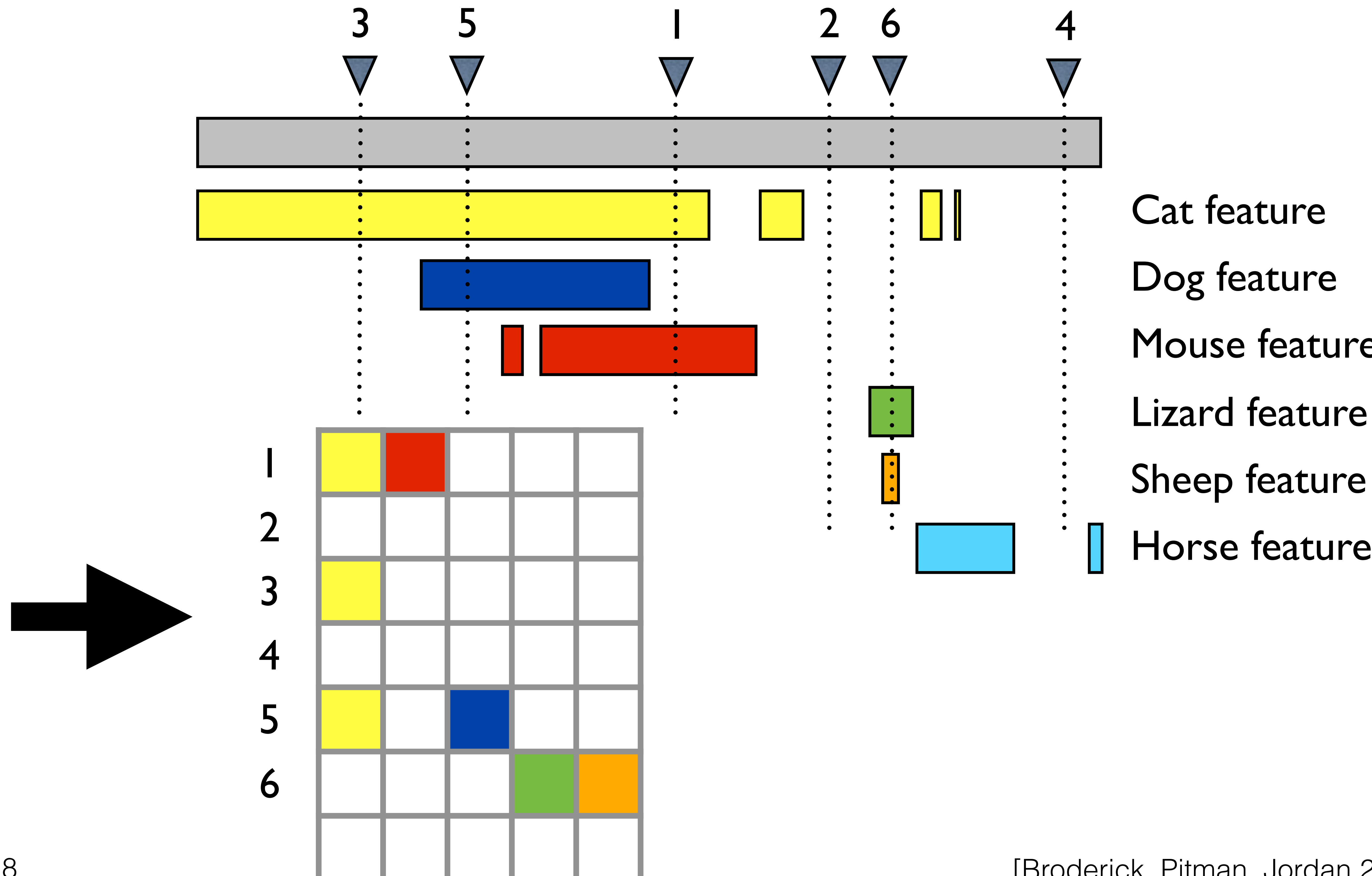

[Broderick, Pitman, Jordan 2013]

# Feature allocation




[Broderick, Pitman, Jordan 2013]

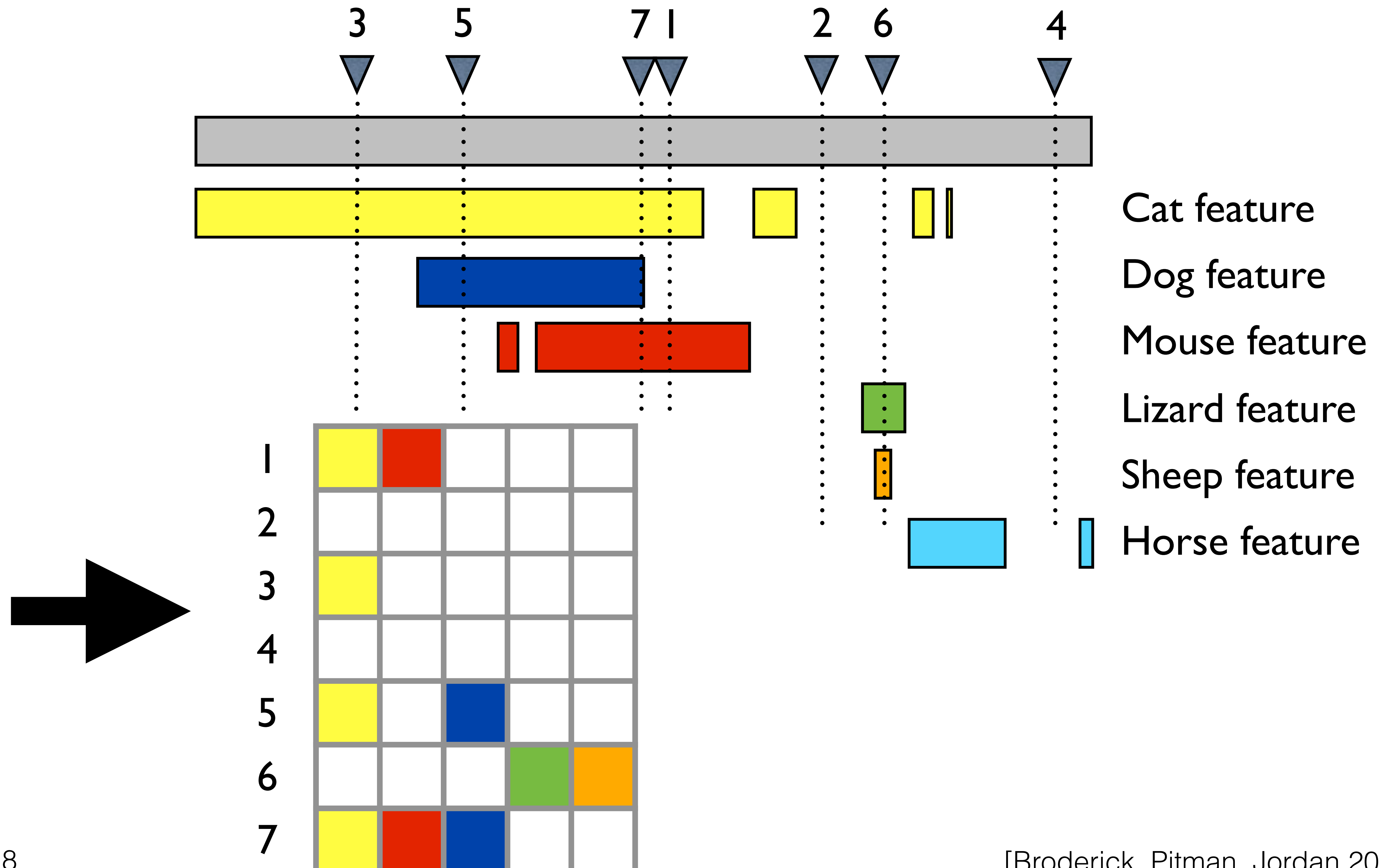
# Feature allocation






[Broderick, Pitman, Jordan 2013]

# Feature allocation

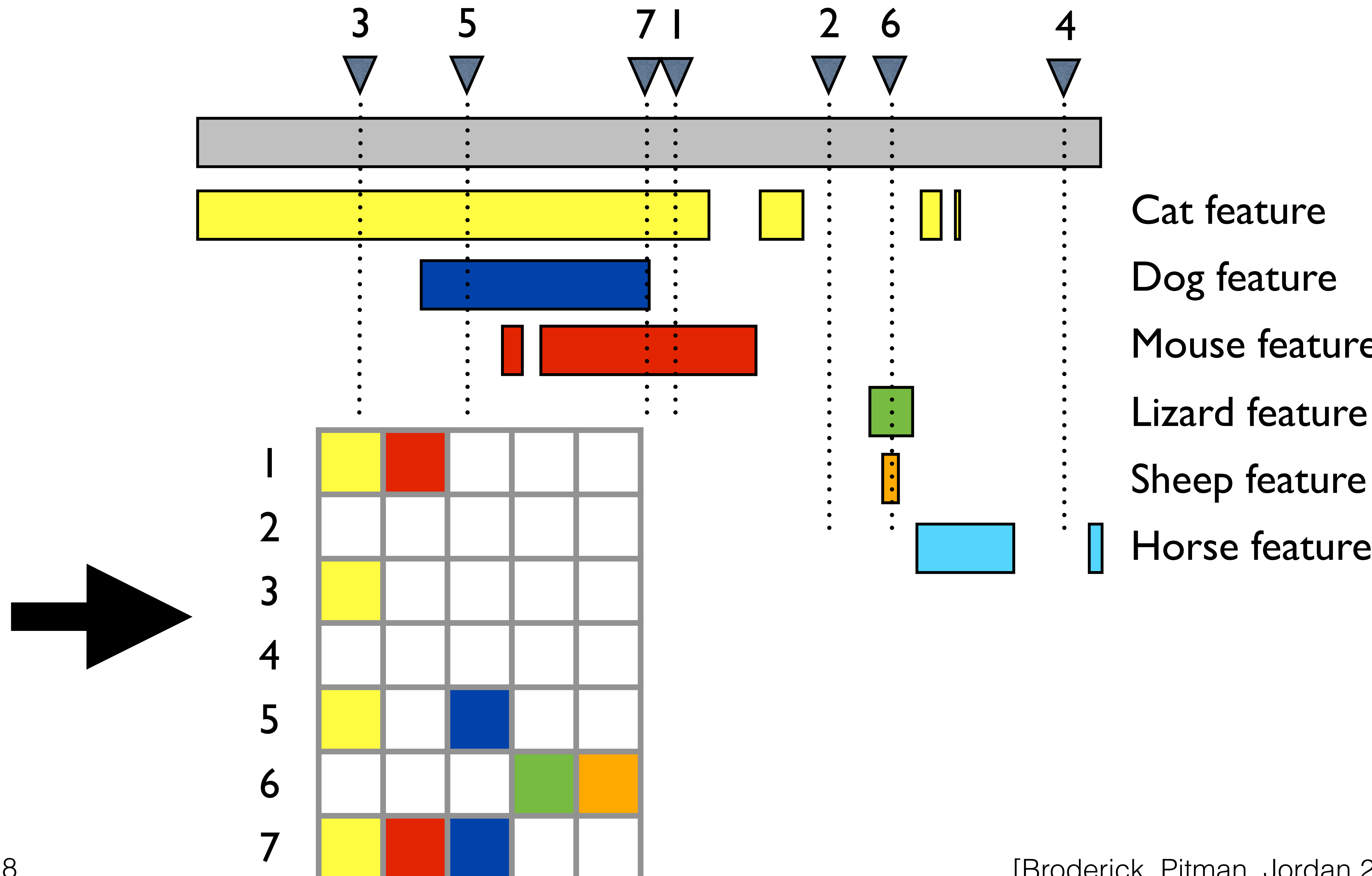





[Broderick, Pitman, Jordan 2013]

# Feature allocation

Thm (BPJ). A feature allocation is exchangeable iff it has a feature paintbox representation

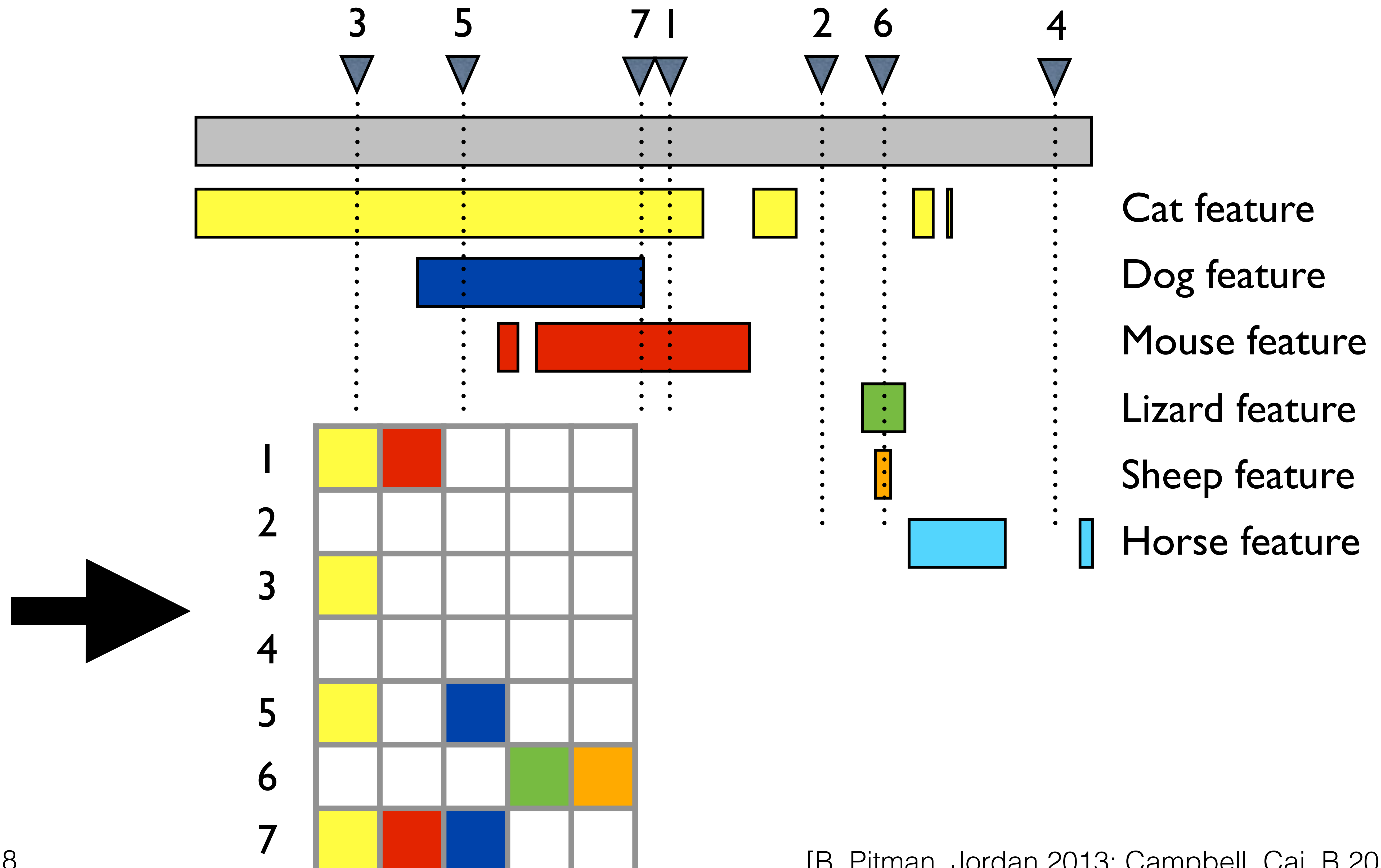





[Broderick, Pitman, Jordan 2013]

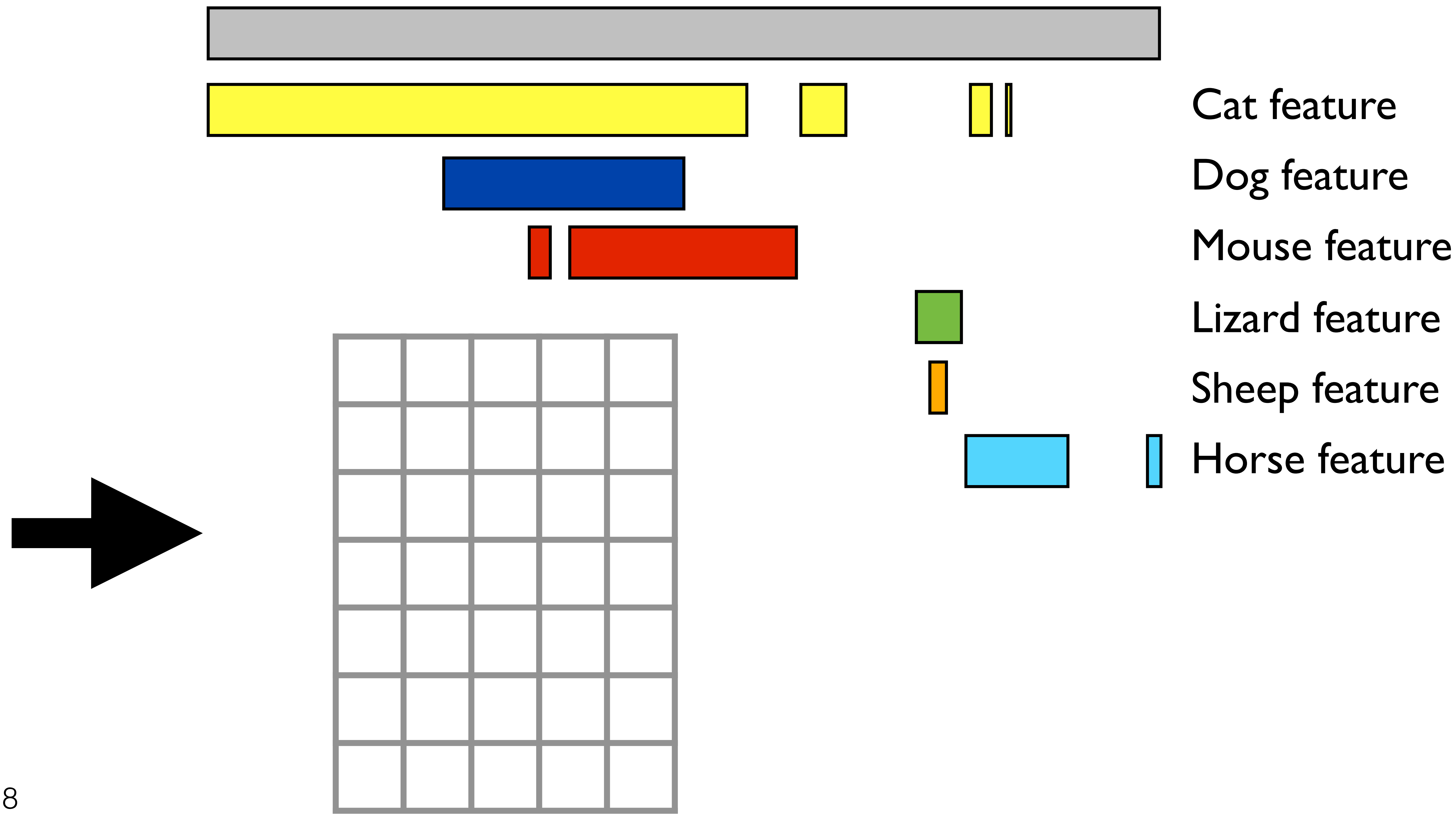
# Feature allocation

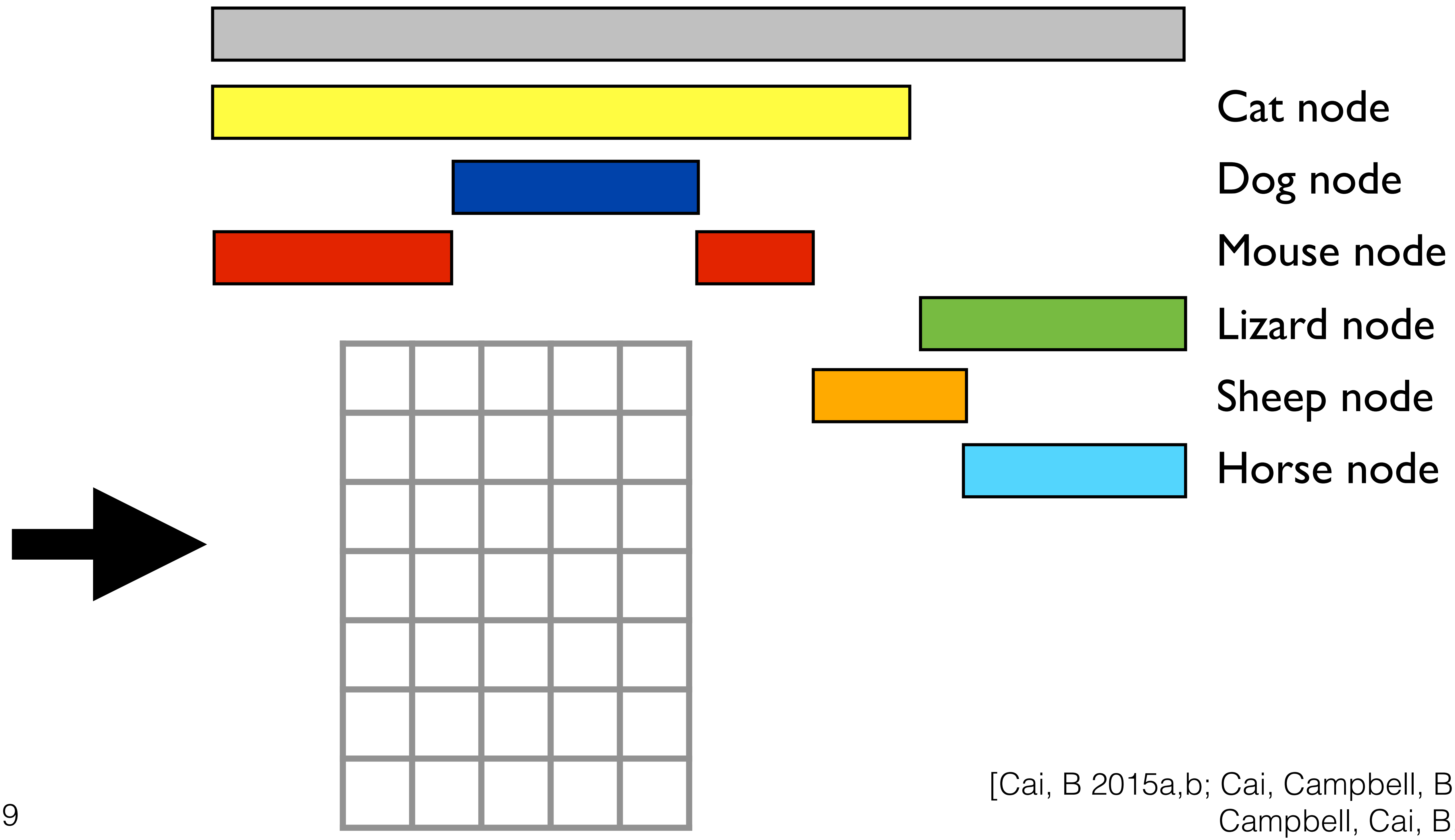
Thm (BPJ). A feature allocation is exchangeable iff it has a feature paintbox representation





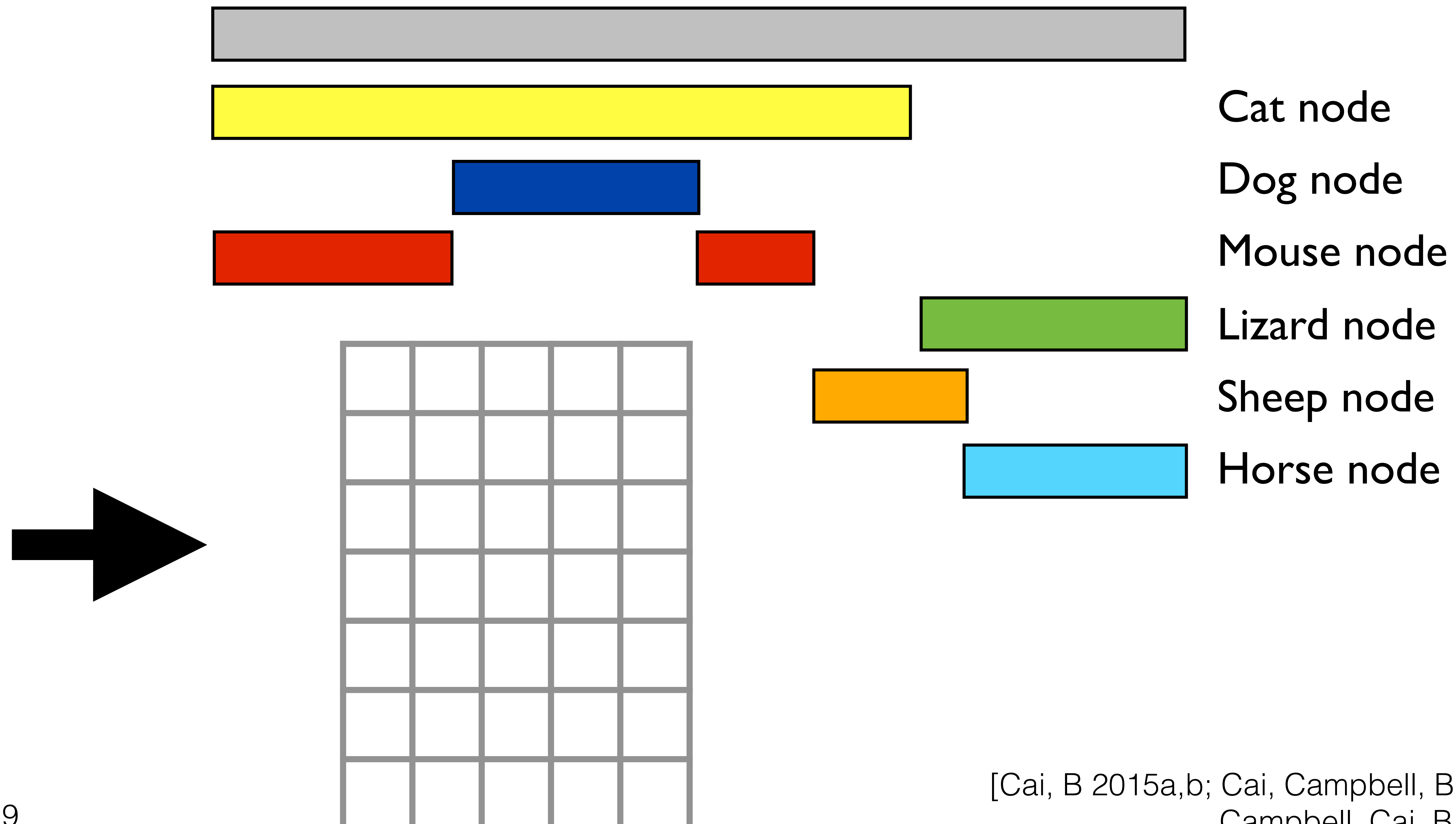

[B, Pitman, Jordan 2013; Campbell, Cai, B 2016]


Cat feature
Dog feature
Mouse feature
Lizard feature
Sheep feature
Horse feature



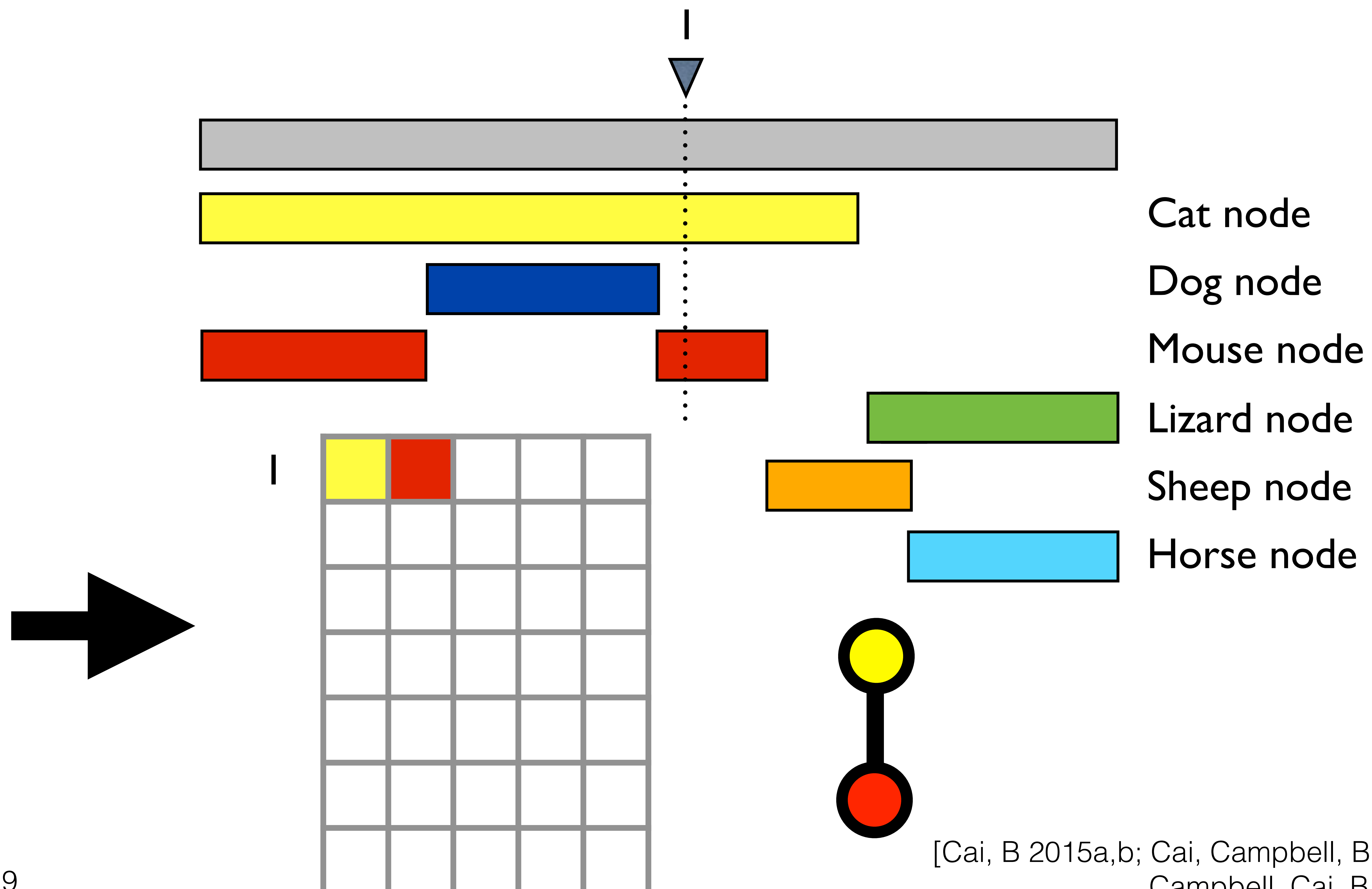

[Cai, B 2015a,b; Cai, Campbell, B, 2016;
Campbell, Cai, B, 2016]

# Edge-exchangeable graph



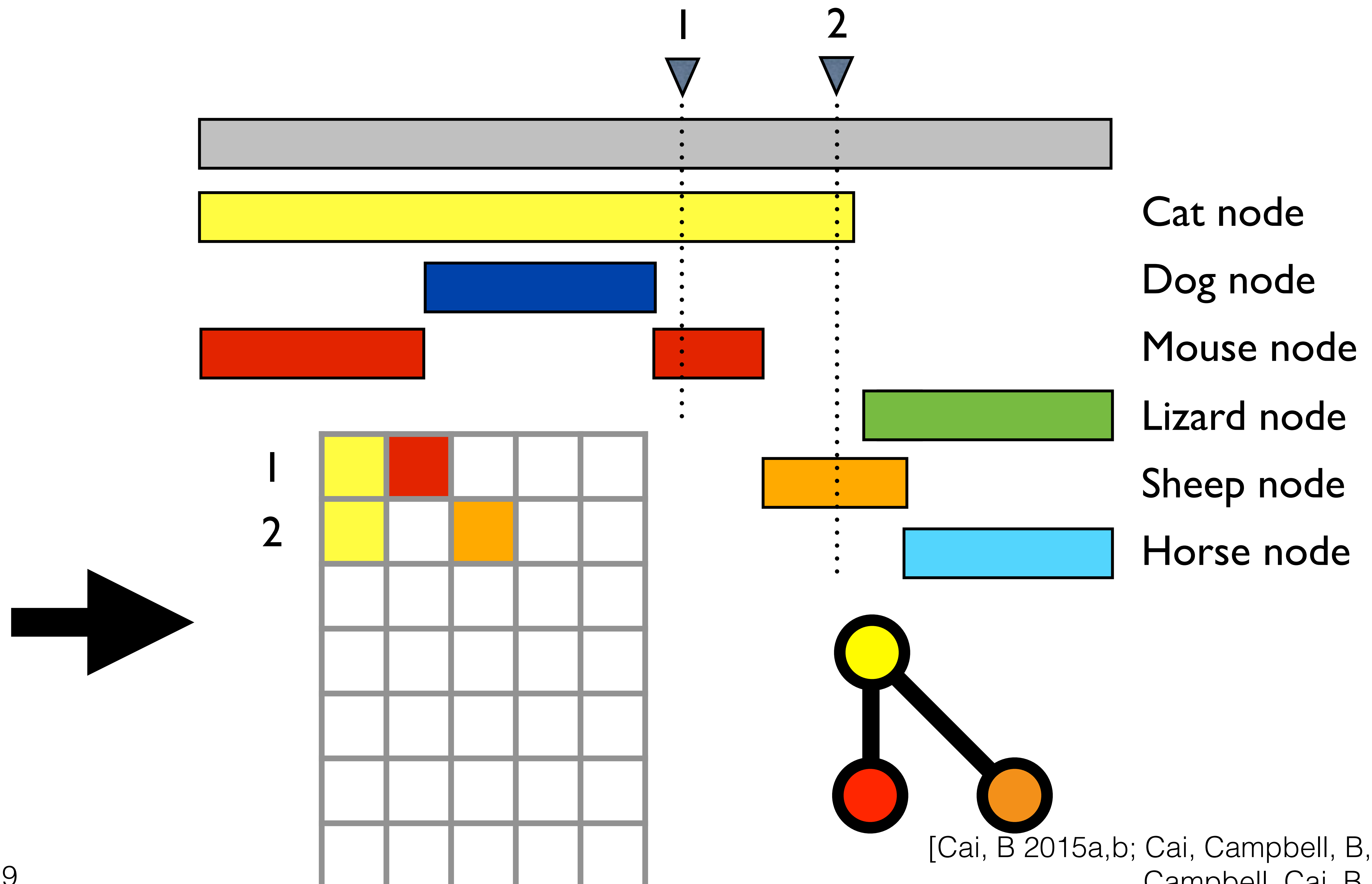

[Cai, B 2015a,b; Cai, Campbell, B, 2016;
Campbell, Cai, B, 2016]

# Edge-exchangeable graph



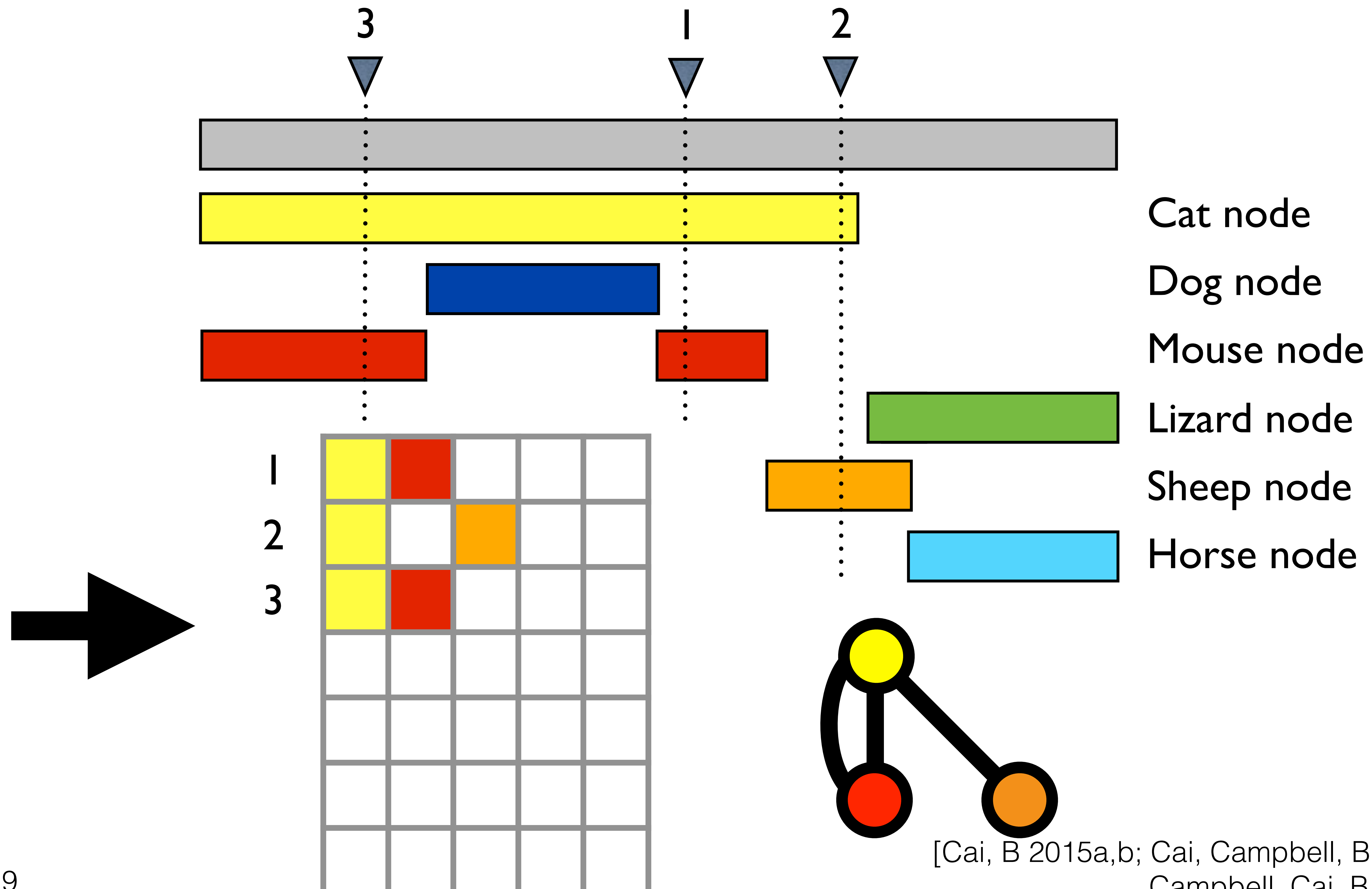

[Cai, B 2015a,b; Cai, Campbell, B, 2016; Campbell, Cai, B, 2016]

# Edge-exchangeable graph



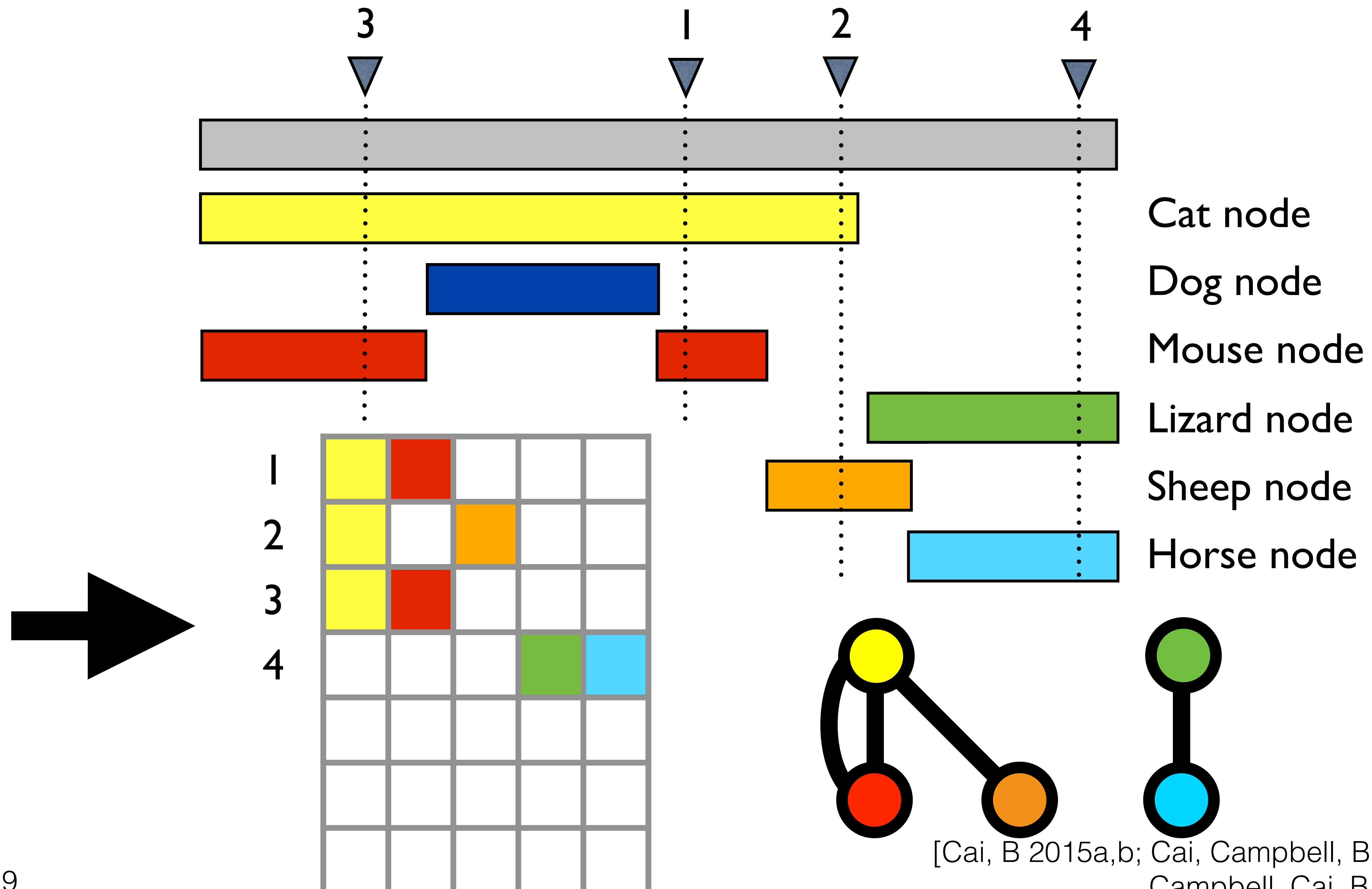

[Cai, B 2015a,b; Cai, Campbell, B, 2016; Campbell, Cai, B, 2016]

# Edge-exchangeable graph



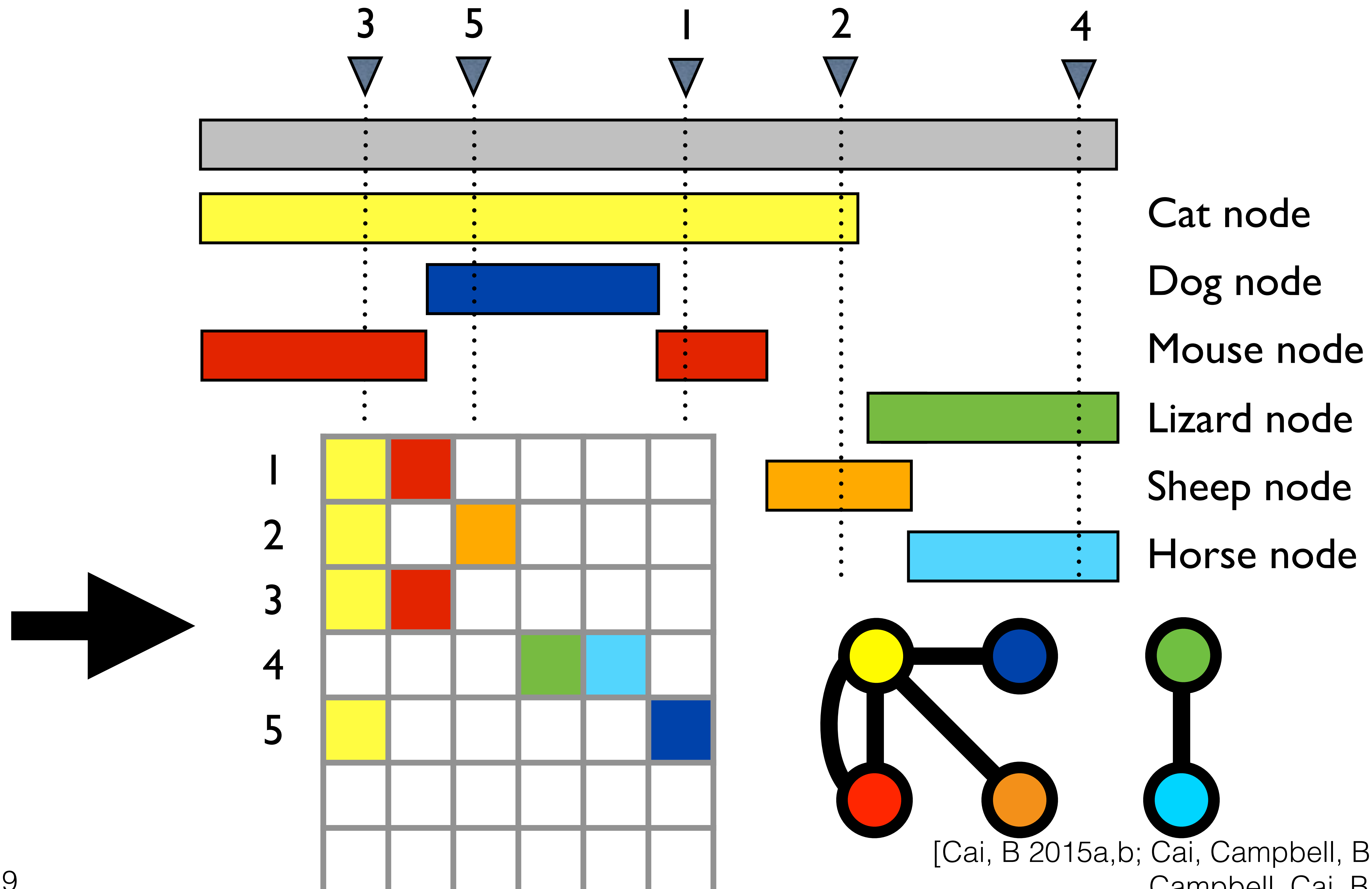

[Cai, B 2015a,b; Cai, Campbell, B, 2016; Campbell, Cai, B, 2016]

# Edge-exchangeable graph



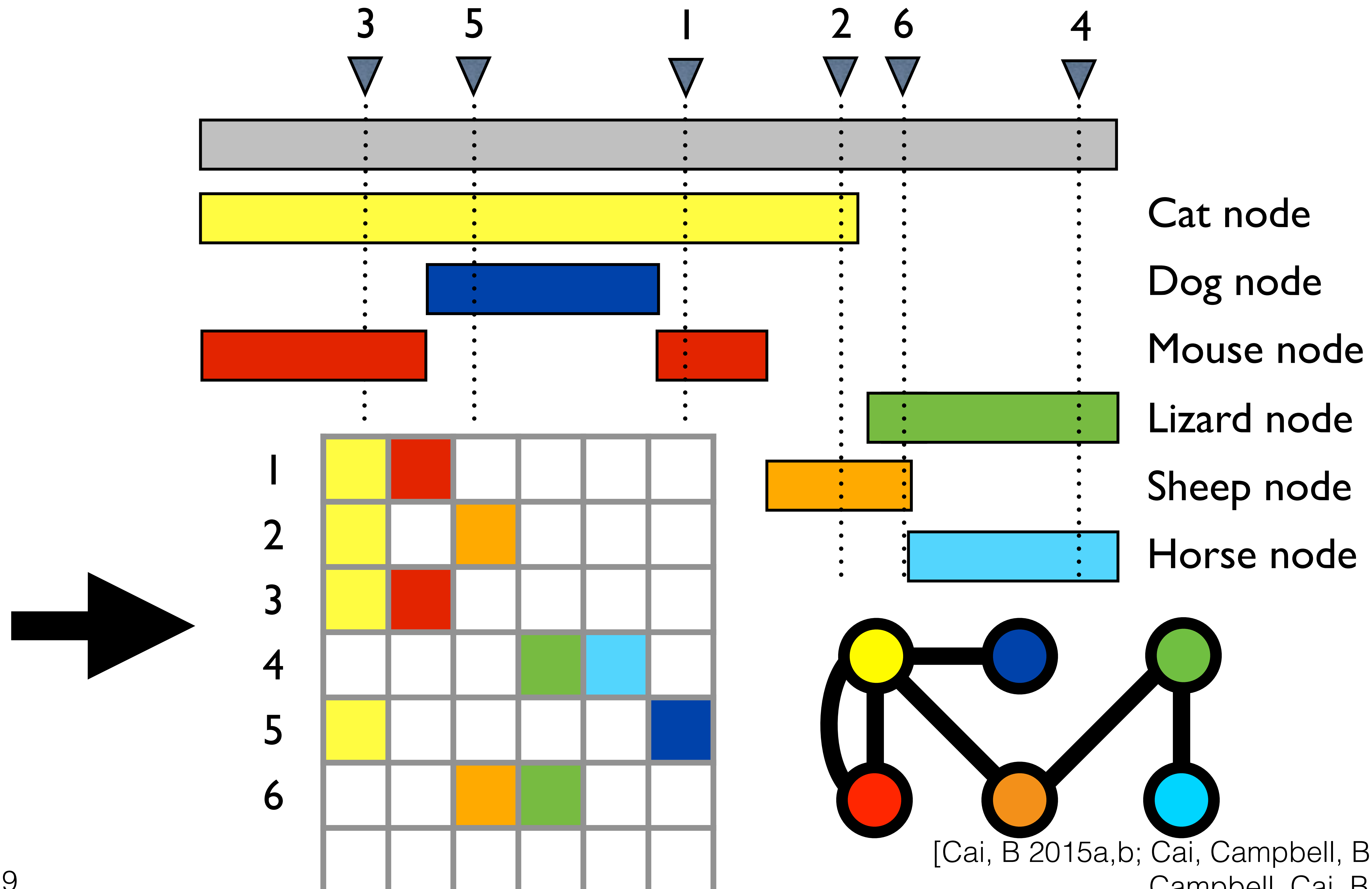

[Cai, B 2015a,b; Cai, Campbell, B, 2016; Campbell, Cai, B, 2016]

# Edge-exchangeable graph



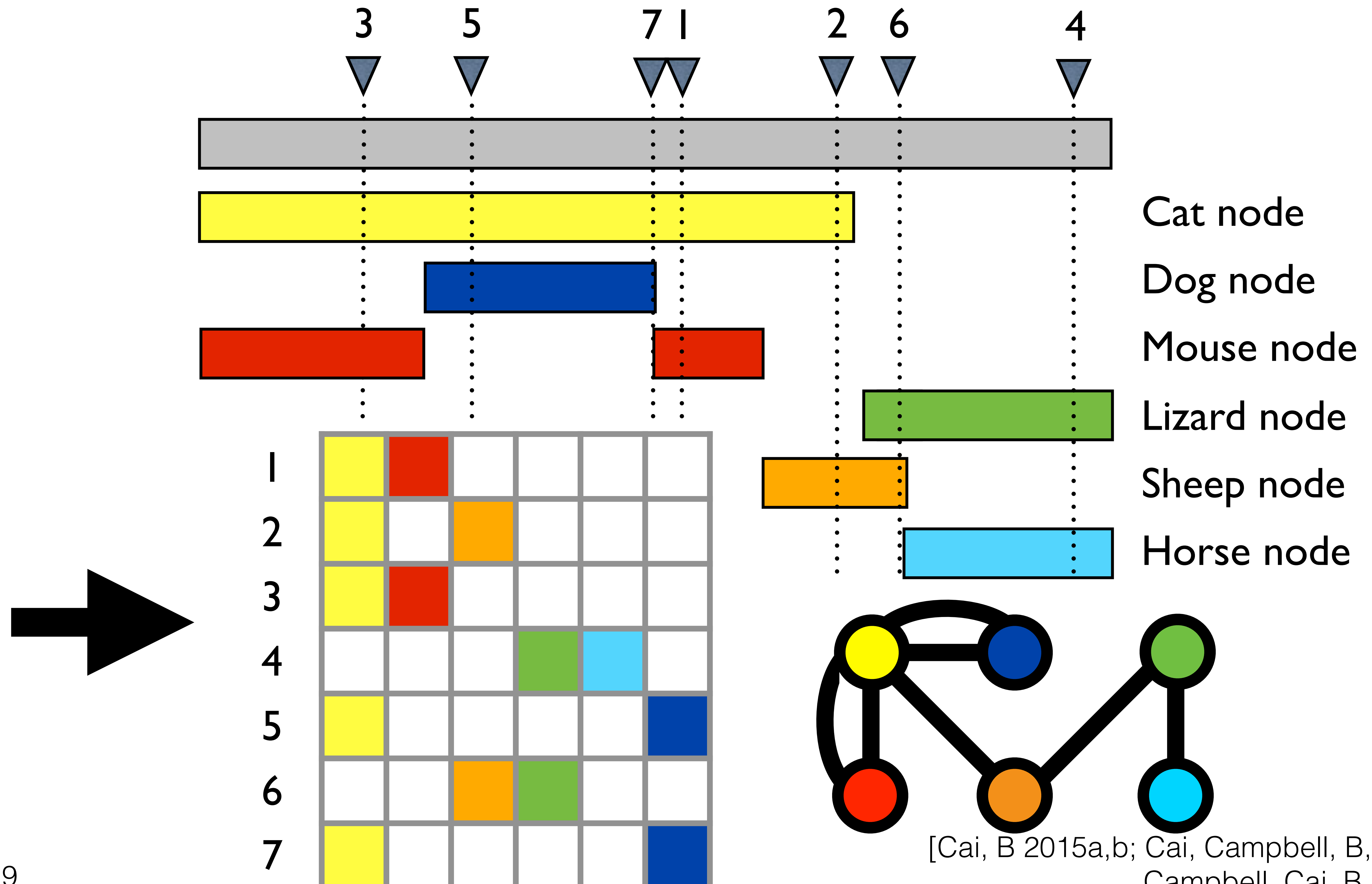

[Cai, B 2015a,b; Cai, Campbell, B, 2016;
Campbell, Cai, B, 2016]

# Edge-exchangeable graph



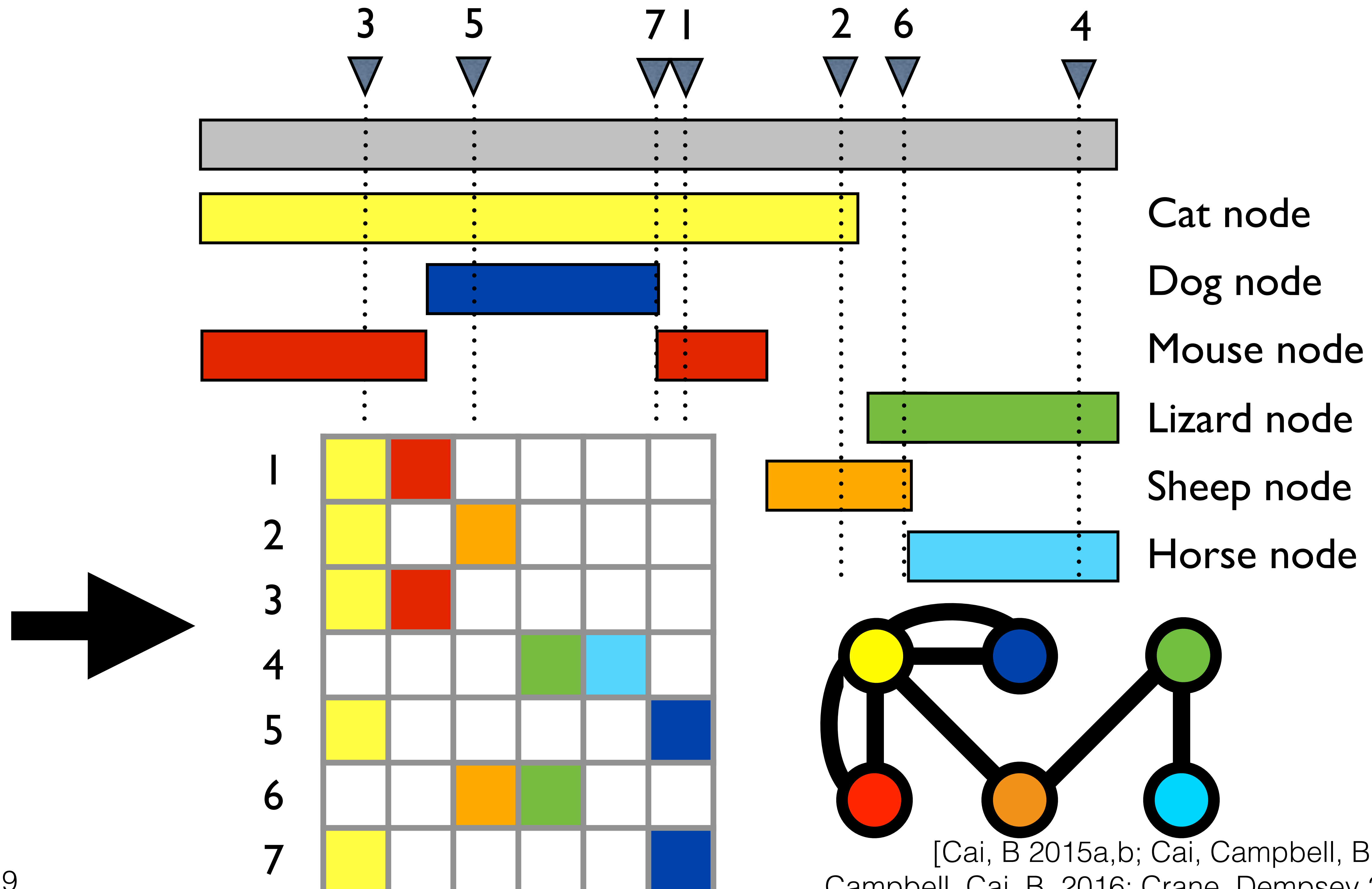

[Cai, B 2015a,b; Cai, Campbell, B, 2016; Campbell, Cai, B, 2016]

# Edge-exchangeable graph




[Cai, B 2015a,b; Cai, Campbell, B, 2016;
Campbell, Cai, B, 2016]

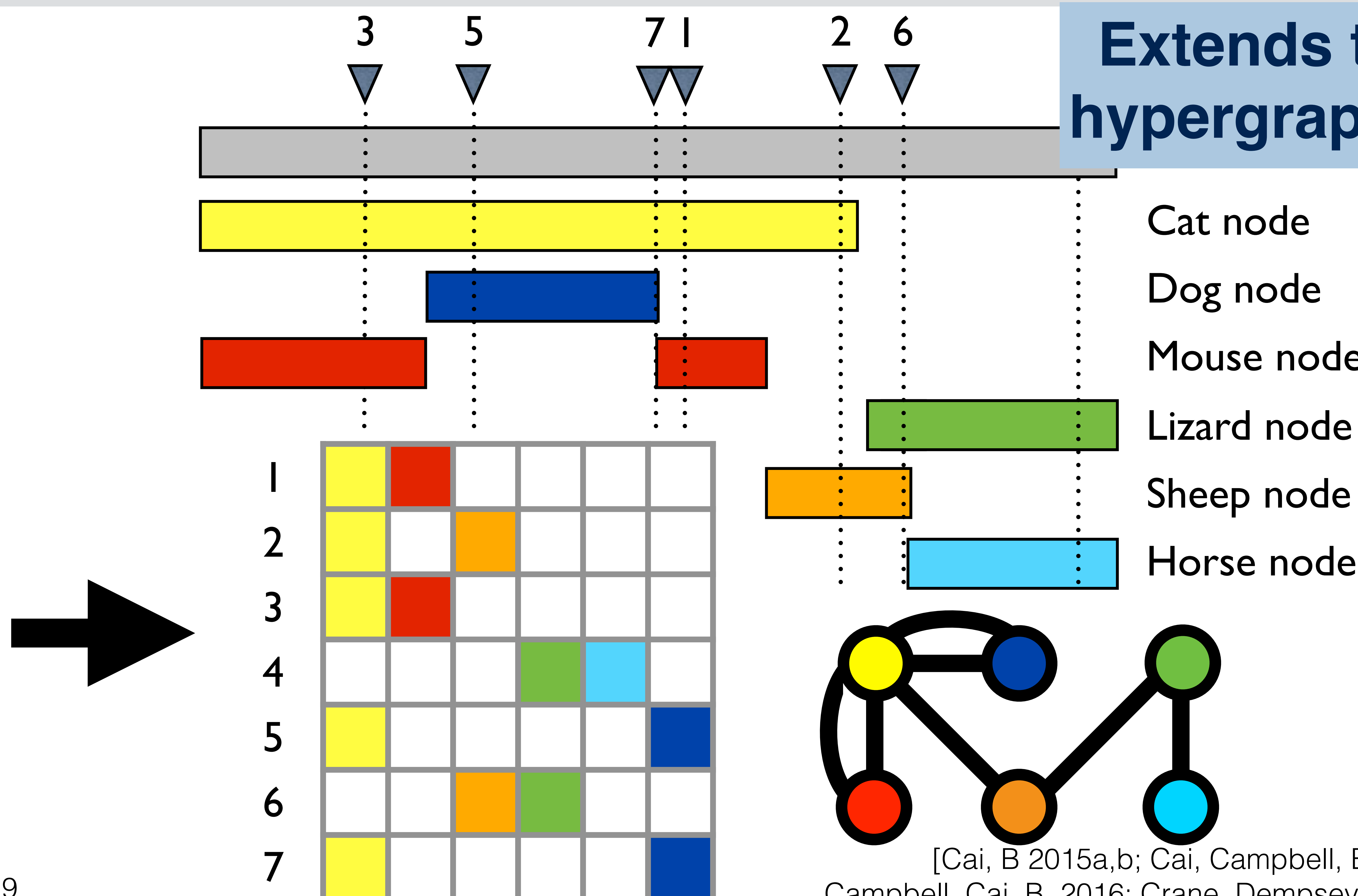
# Cor (CCB). A graph sequence is edge-exchangeable iff it has a graph paintbox




[Cai, B 2015a,b; Cai, Campbell, B, 2016; Campbell, Cai, B, 2016; Crane, Dempsey 2016b]

# Cor (CCB). A graph sequence is edge-exchangeable iff it has a graph paintbox

**Extends to hypergraphs**

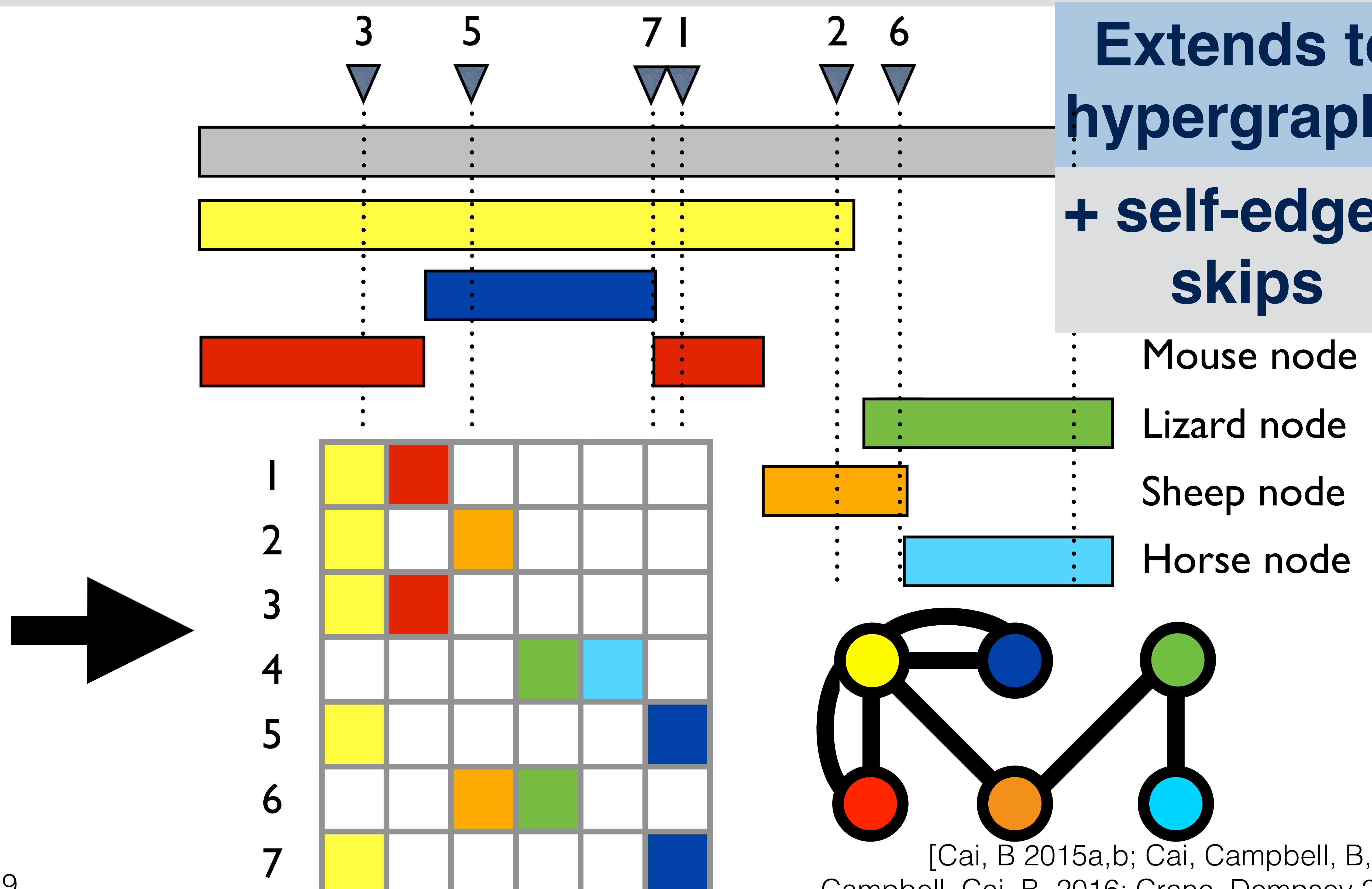



[Cai, B 2015a,b; Cai, Campbell, B, 2016; Campbell, Cai, B, 2016; Crane, Dempsey 2016b]

# Cor (CCB). A graph sequence is edge-exchangeable iff it has a graph paintbox

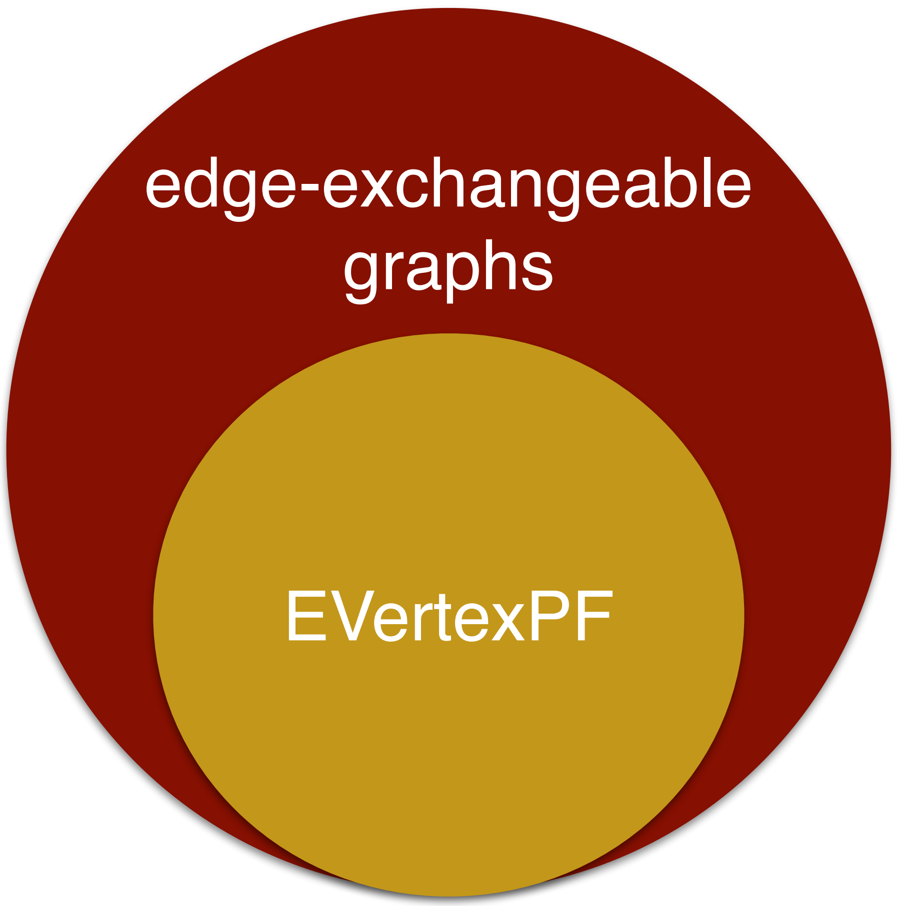
**Extends to hypergraphs**

**+ self-edges, skips**



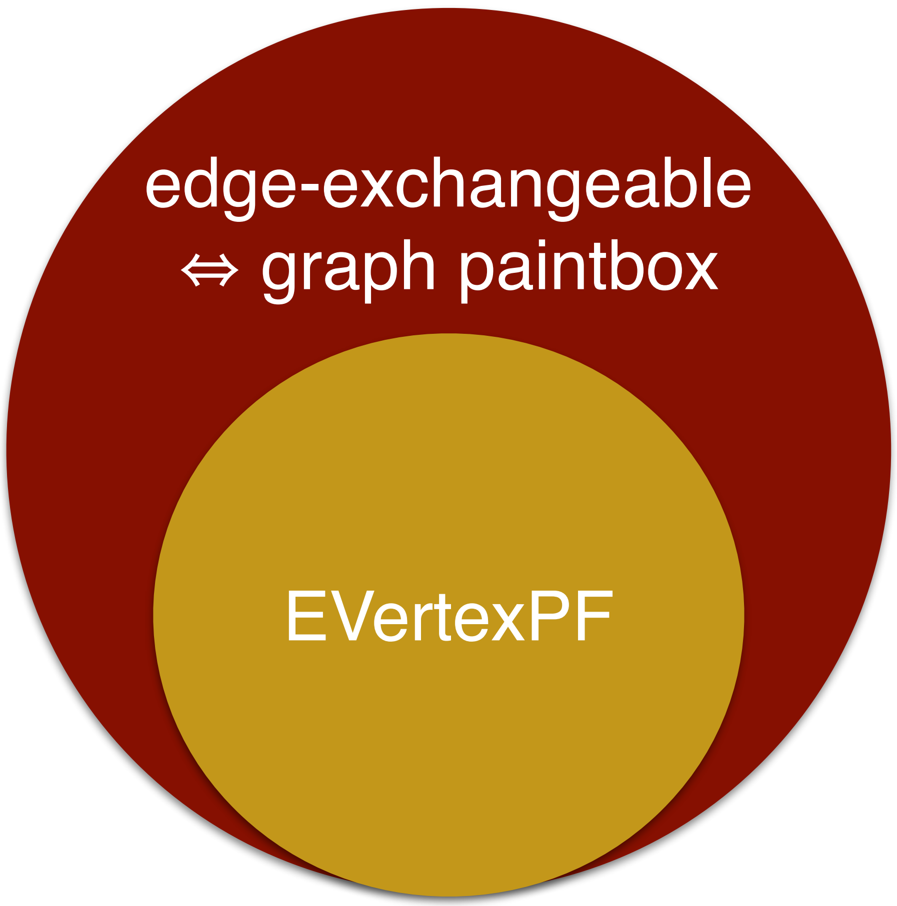

[Cai, B 2015a,b; Cai, Campbell, B, 2016; Campbell, Cai, B, 2016; Crane, Dempsey 2016b]

# What we know so far



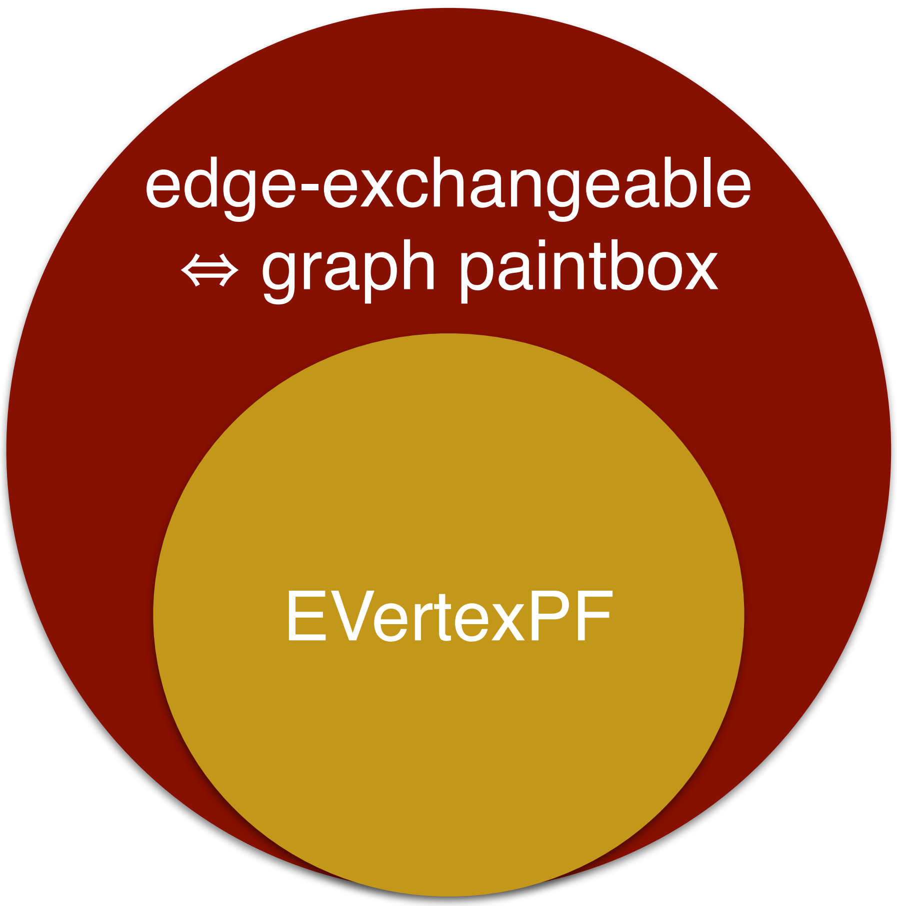

- Goal 1: characterization theorem for edge-exchangeable graphs
- Goal 2: sparsity theorem for edge-exchangeable graphs

# What we know so far




- Goal 1: characterization theorem for edge-exchangeable graphs
- Goal 2: sparsity theorem for edge-exchangeable graphs

# What we know so far




- Thm 1: characterization theorem for edge-exchangeable graphs
- Goal 2: sparsity theorem for edge-exchangeable graphs

# How to prove sparsity?

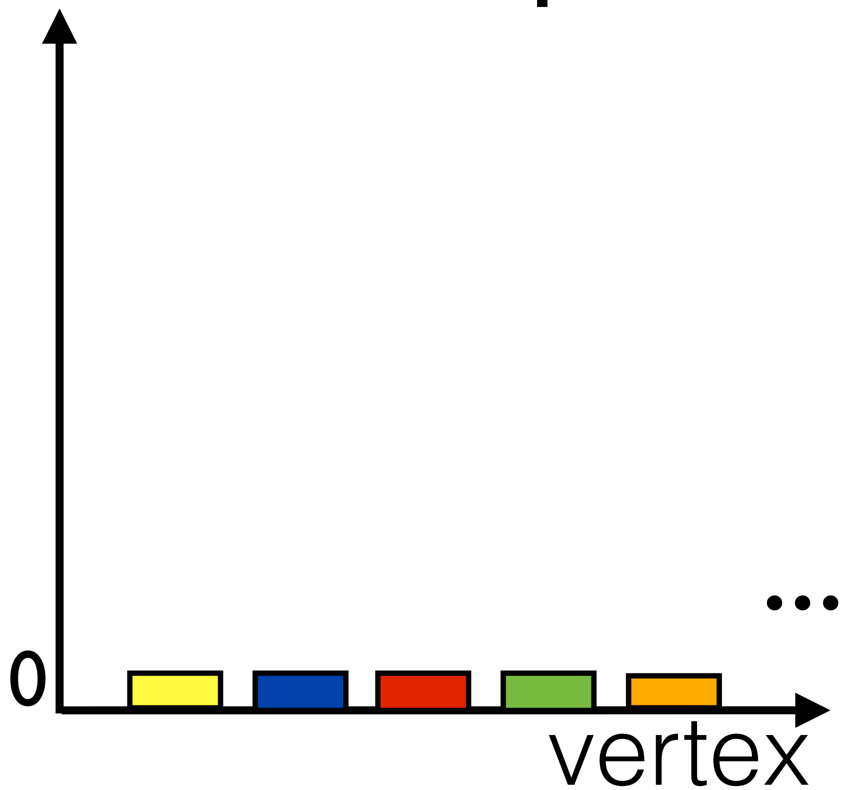
# How to prove sparsity?

- Need # nodes to go to infinity

# How to prove sparsity?

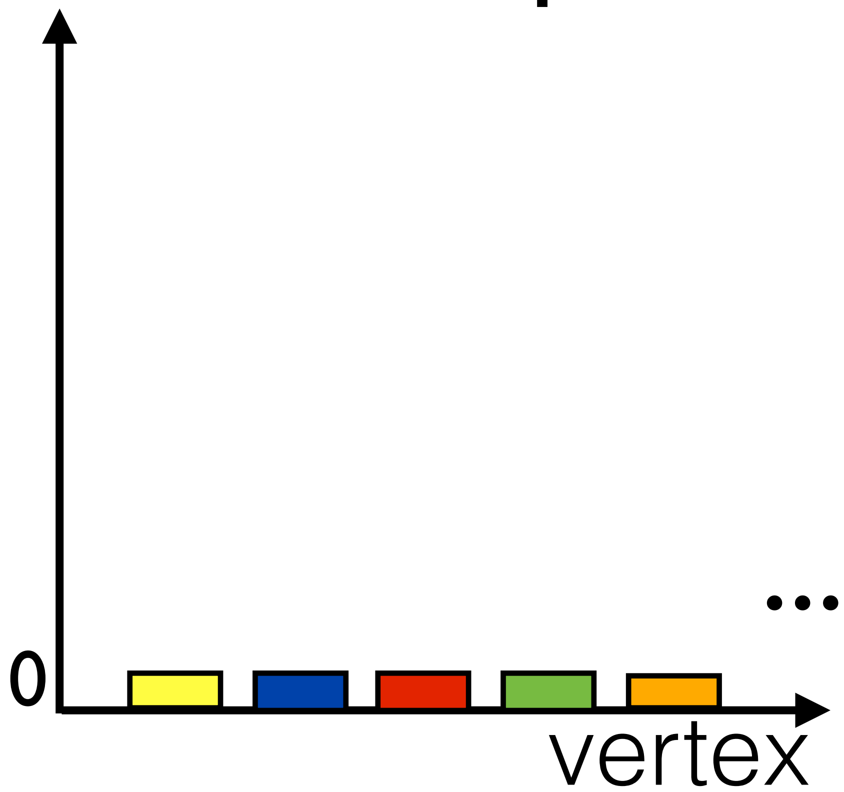
- Need # nodes to go to infinity
  - Need countable $\infty$ of latent nodes

# How to prove sparsity?




- Need # nodes to go to infinity
    - Need countable ∞ of latent nodes

# How to prove sparsity?

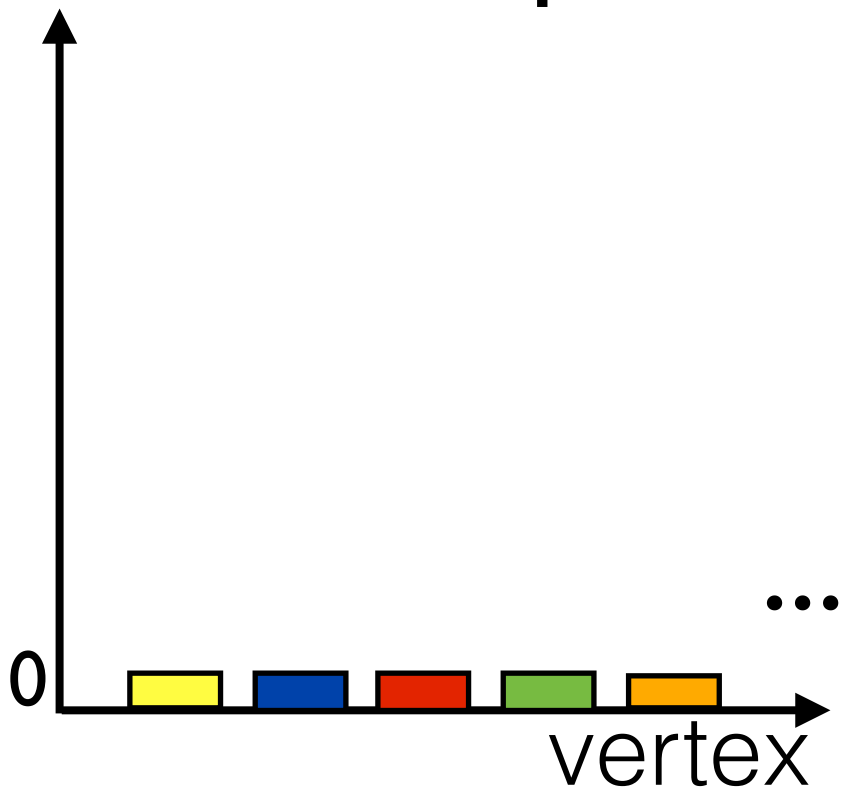



- Need # nodes to go to infinity
    - Need countable ∞ of latent nodes
- Graph frequency model/vertex popularity model



[Caron, Fox 2014; Cai, Campbell, B 2016; Crane, Dempsey 2016a; Palla, Caron, Teh 2016; Herlau, Schmidt 2016]

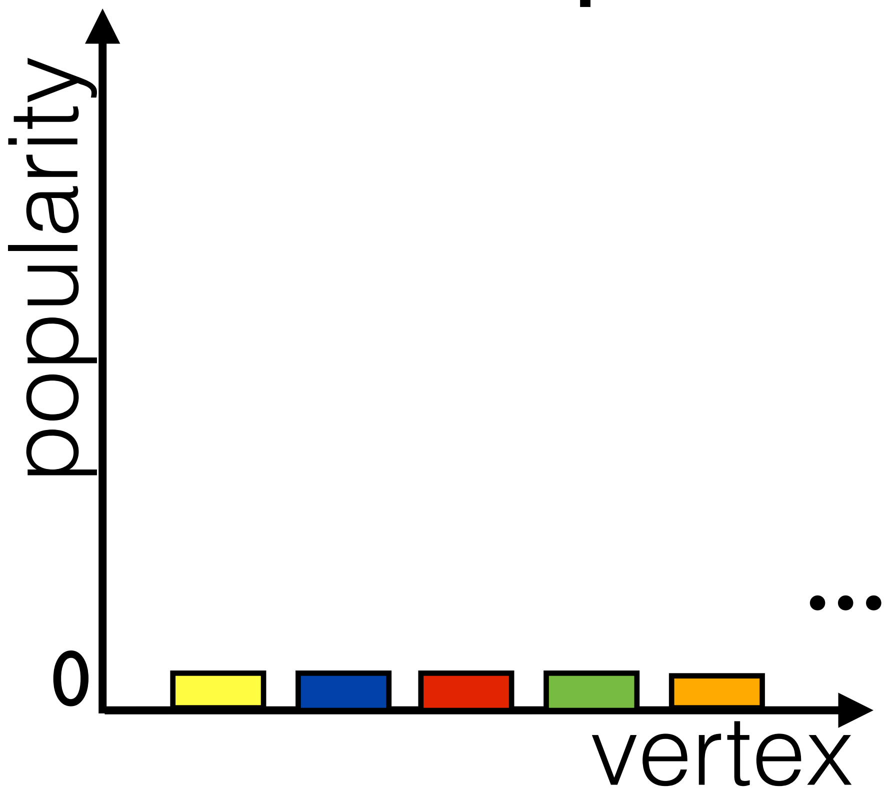
# How to prove sparsity?




- Need # nodes to go to infinity
  - Need countable ∞ of latent nodes
- Graph frequency model/vertex popularity model
  - Draw a frequency $w_i$ for each vertex $i$


[Caron, Fox 2014; Cai, Campbell, B 2016; Crane, Dempsey 2016a; Palla, Caron, Teh 2016; Herlau, Schmidt 2016]

# How to prove sparsity?

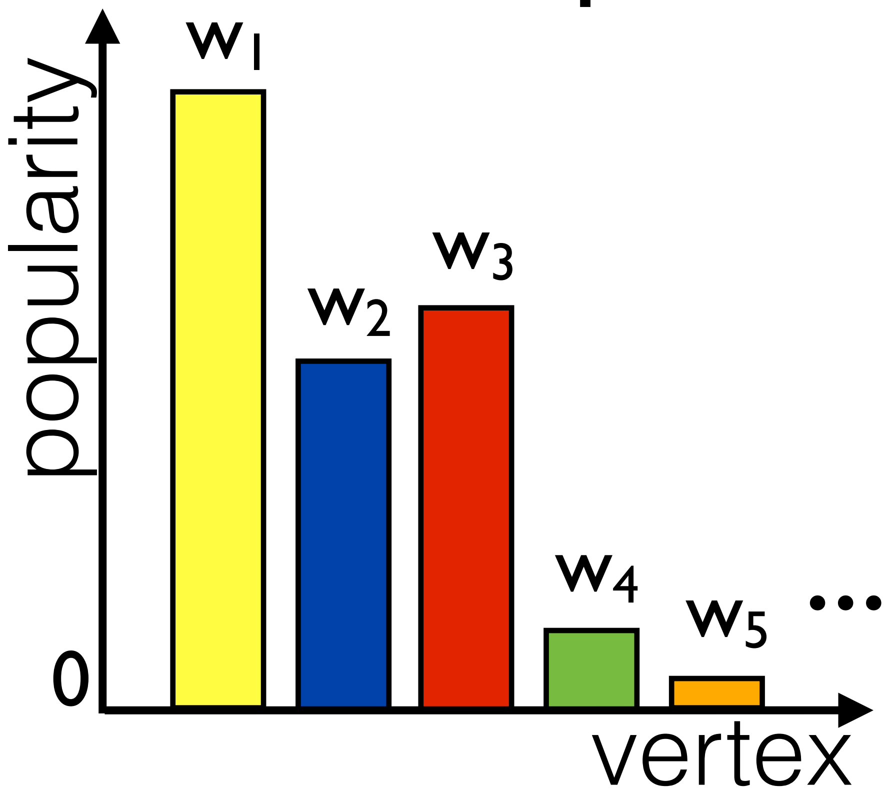



- Need # nodes to go to infinity
  - Need countable ∞ of latent nodes
- Graph frequency model/vertex popularity model
  - Draw a frequency $w_i$ for each vertex $i$



[Caron, Fox 2014; Cai, Campbell, B 2016; Crane, Dempsey 2016a; Palla, Caron, Teh 2016; Herlau, Schmidt 2016]

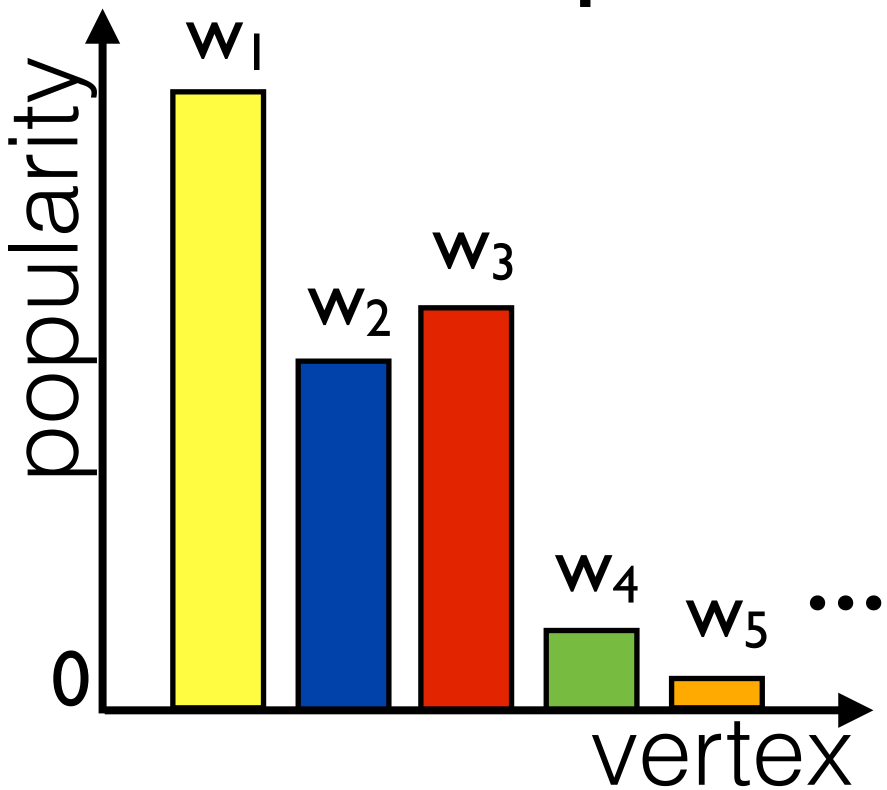
# How to prove sparsity?




- Need # nodes to go to infinity
  - Need countable ∞ of latent nodes
- Graph frequency model/vertex popularity model
  - Draw a frequency $w_i$ for each vertex $i$


[Caron, Fox 2014; Cai, Campbell, B 2016; Crane, Dempsey 2016a; Palla, Caron, Teh 2016; Herlau, Schmidt 2016]

# How to prove sparsity?

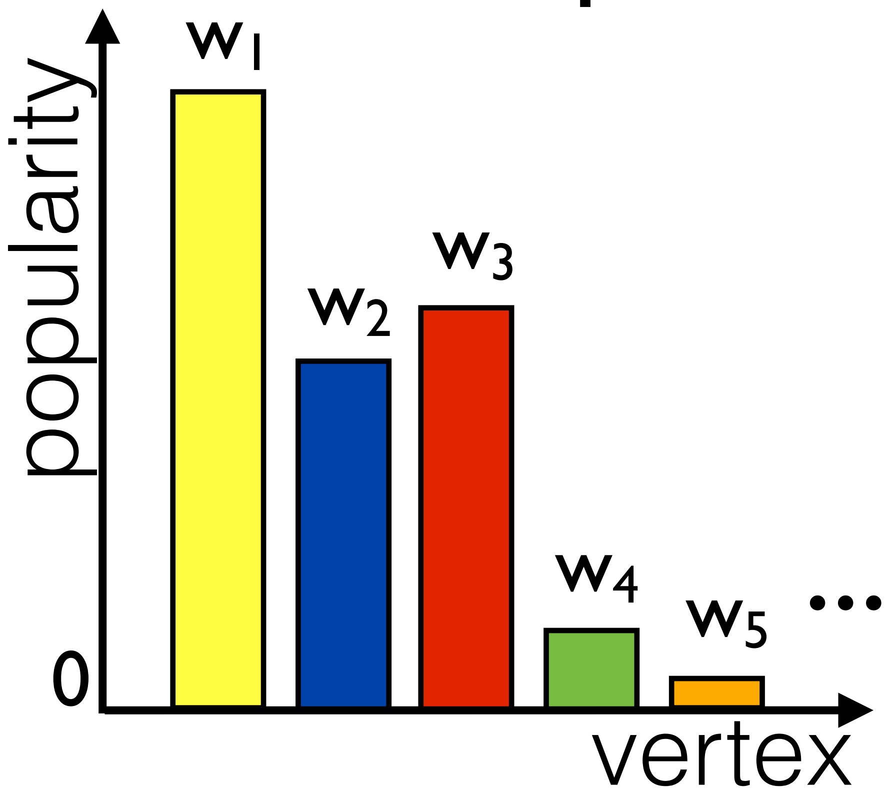



- Need # nodes to go to infinity
  - Need countable ∞ of latent nodes
- Graph frequency model/vertex popularity model
  - Draw a frequency $w_i$ for each vertex $i$
  - Draw edge {i,j} with probability $w_i w_j$

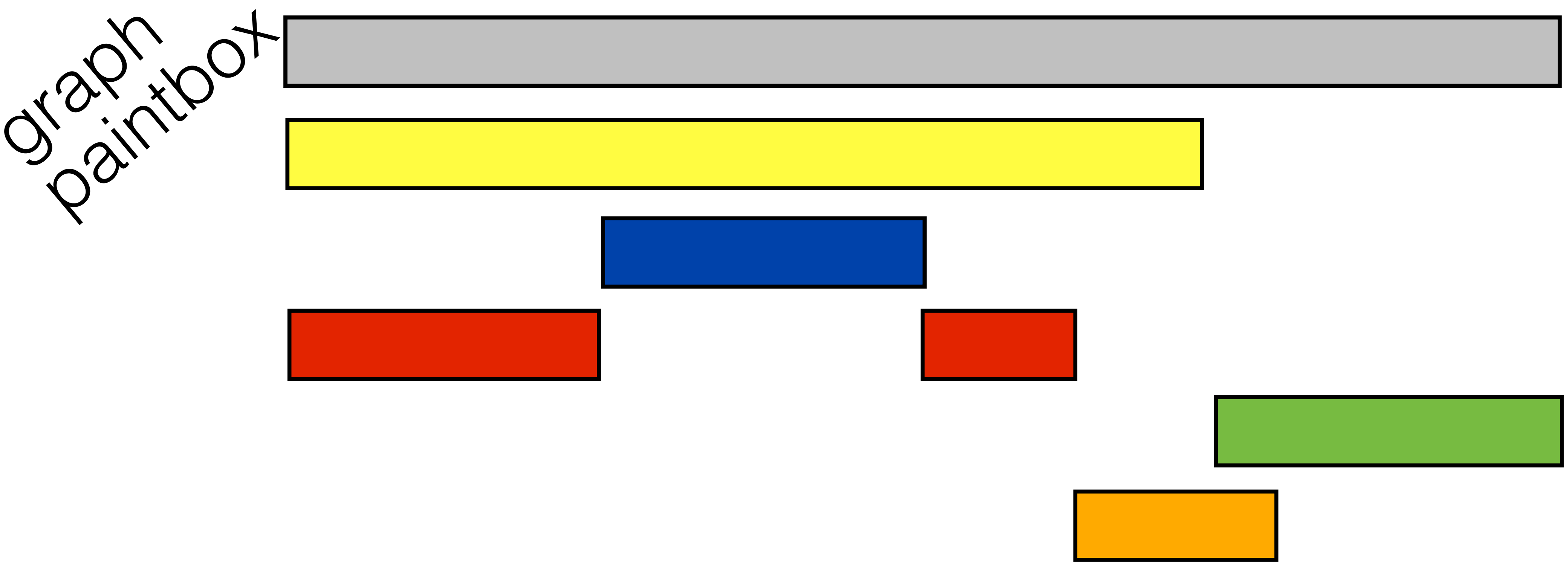


[Caron, Fox 2014; Cai, Campbell, B 2016; Crane, Dempsey 2016a; Palla, Caron, Teh 2016; Herlau, Schmidt 2016]

# How to prove sparsity?



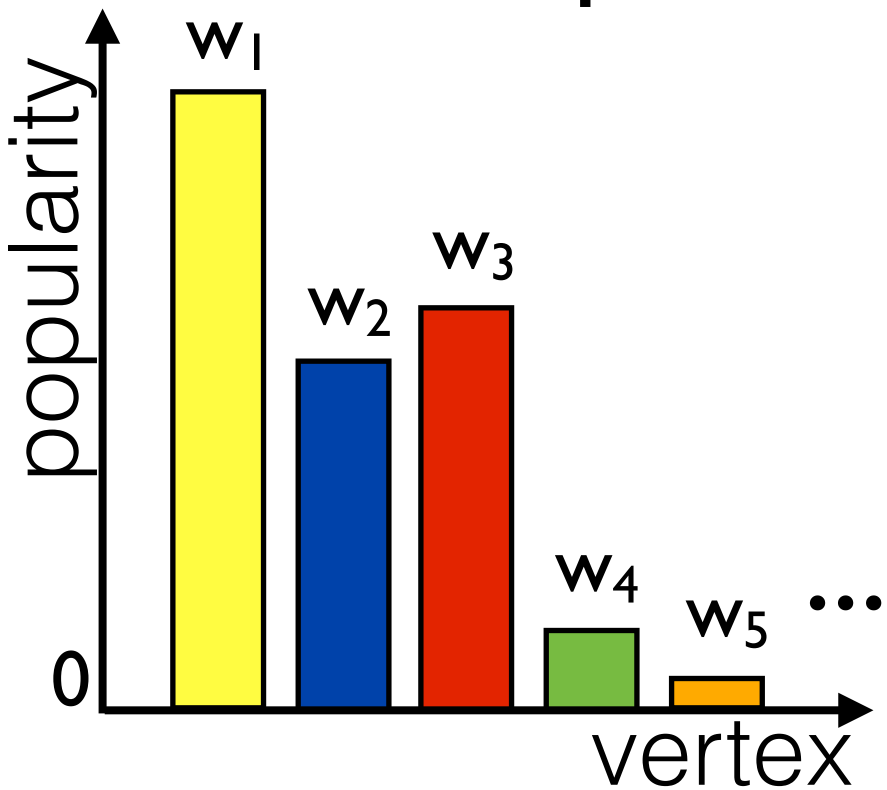

- Need # nodes to go to infinity
  - Need countable ∞ of latent nodes
- Graph frequency model/vertex popularity model
  - Draw a frequency $w_i$ for each vertex $i$
  - Draw edge {i,j} with probability $w_i w_j$



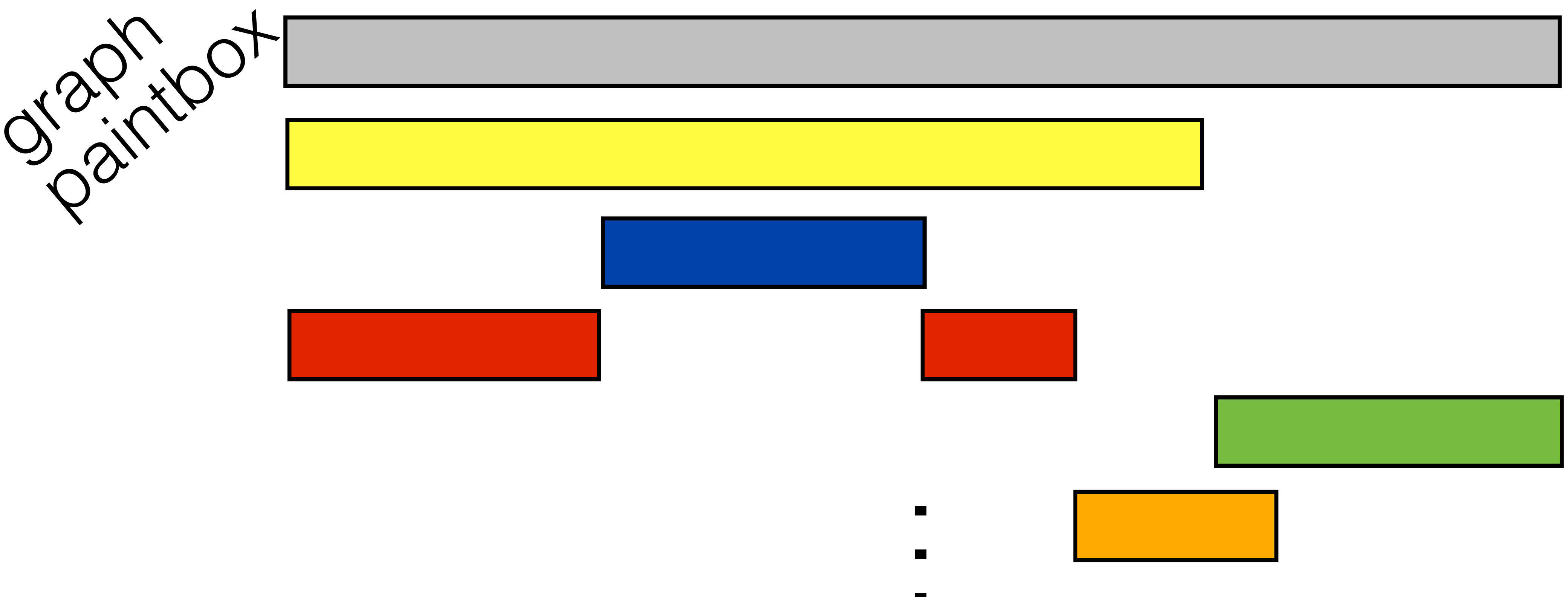



[Caron, Fox 2014; Cai, Campbell, B 2016; Crane, Dempsey 2016a; Palla, Caron, Teh 2016; Herlau, Schmidt 2016]

# How to prove sparsity?



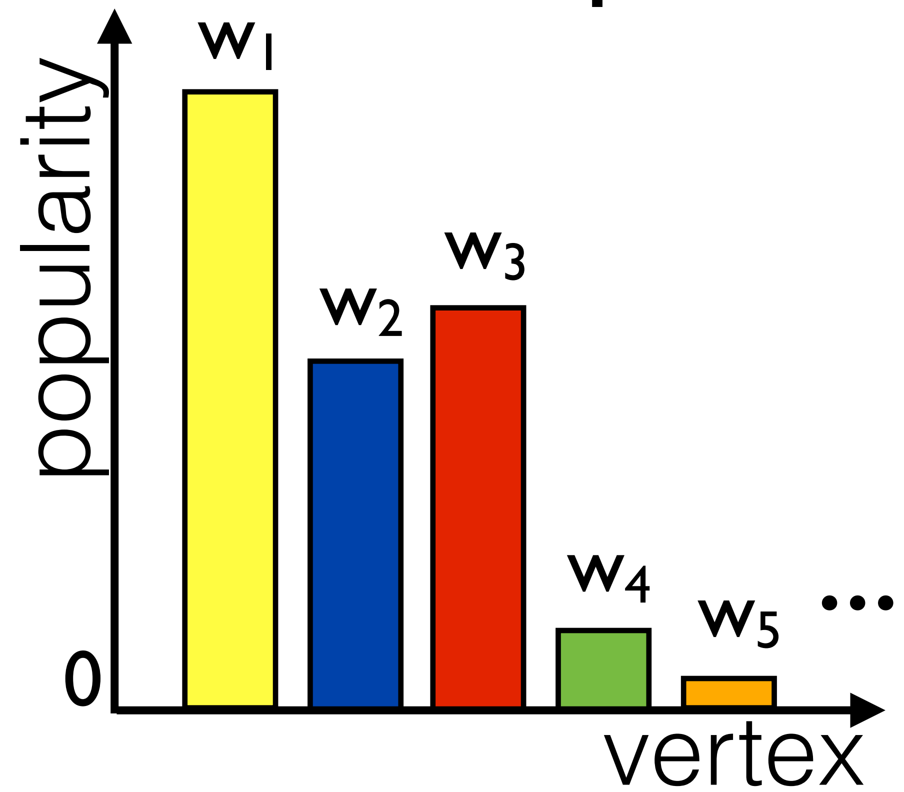

- Need # nodes to go to infinity
  - Need countable ∞ of latent nodes
- Graph frequency model/vertex popularity model
  - Draw a frequency $w_i$ for each vertex $i$
  - Draw edge {i,j} with probability $w_i w_j$



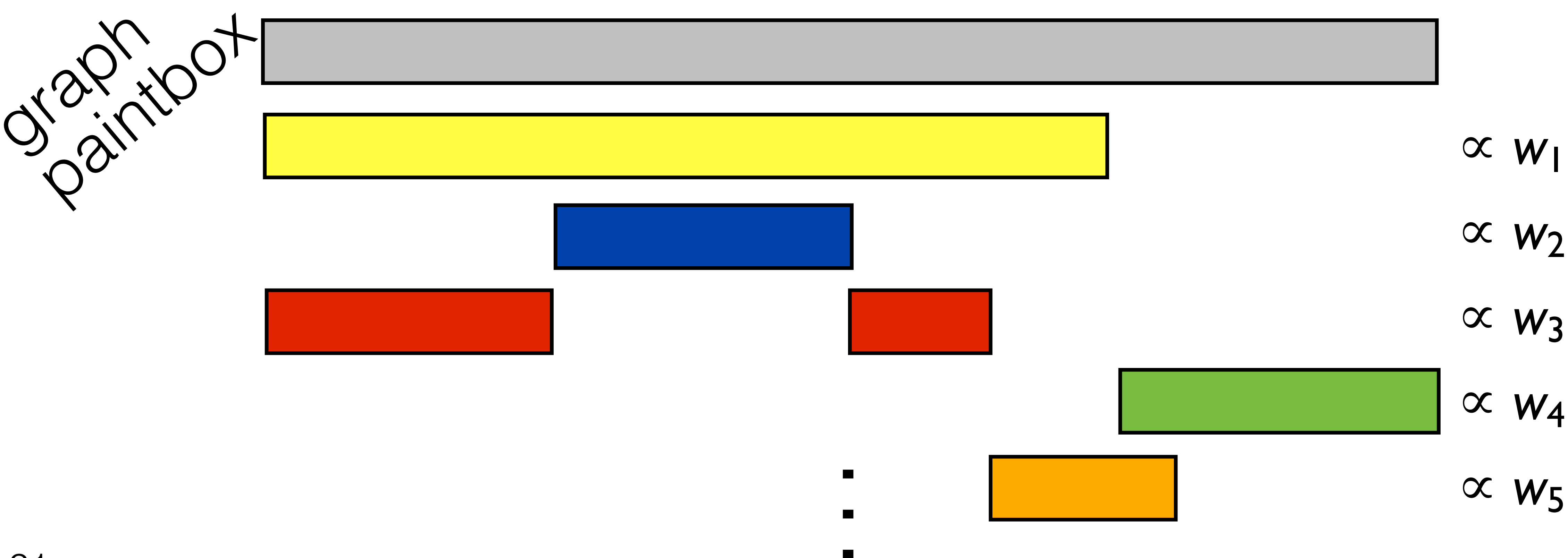



[Caron, Fox 2014; Cai, Campbell, B 2016; Crane, Dempsey 2016a; Palla, Caron, Teh 2016; Herlau, Schmidt 2016]

# How to prove sparsity?



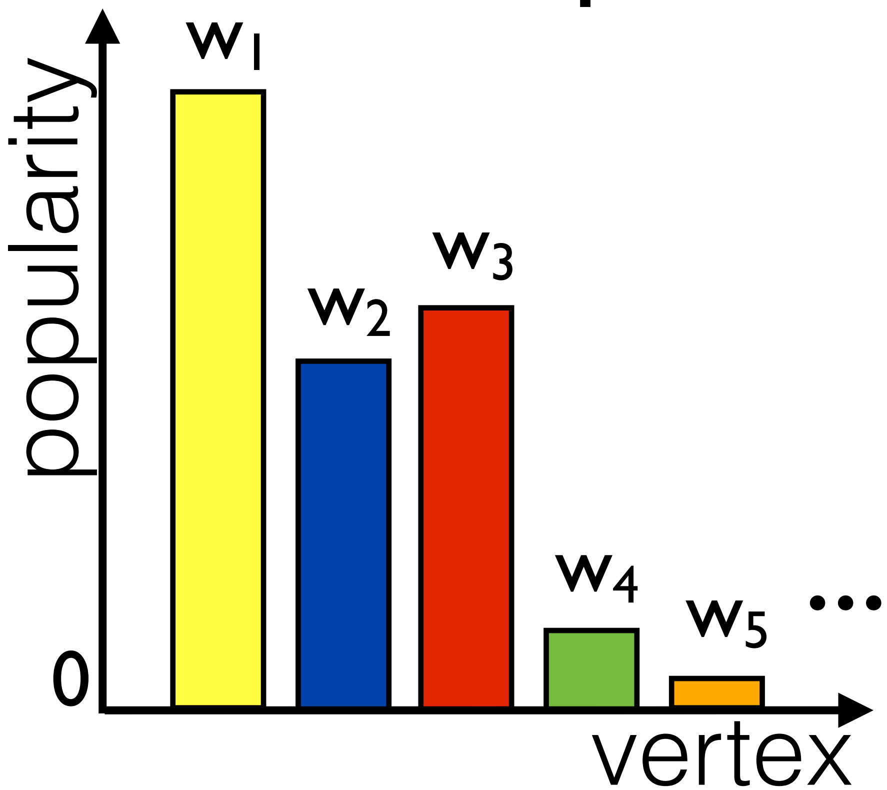

- Need # nodes to go to infinity
  - Need countable ∞ of latent nodes
- Graph frequency model/vertex popularity model
  - Draw a frequency $w_i$ for each vertex $i$
  - Draw edge {i,j} with probability $w_i w_j$



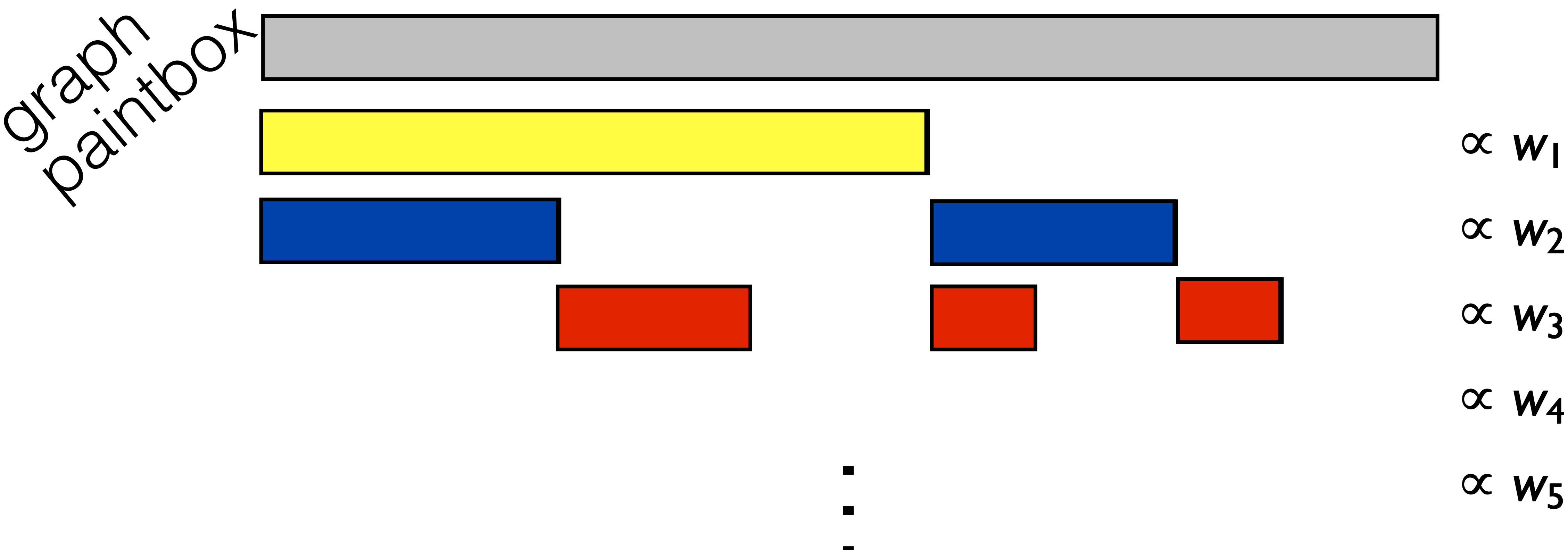



[Caron, Fox 2014; Cai, Campbell, B 2016; Crane, Dempsey 2016a; Palla, Caron, Teh 2016; Herlau, Schmidt 2016]

# How to prove sparsity?



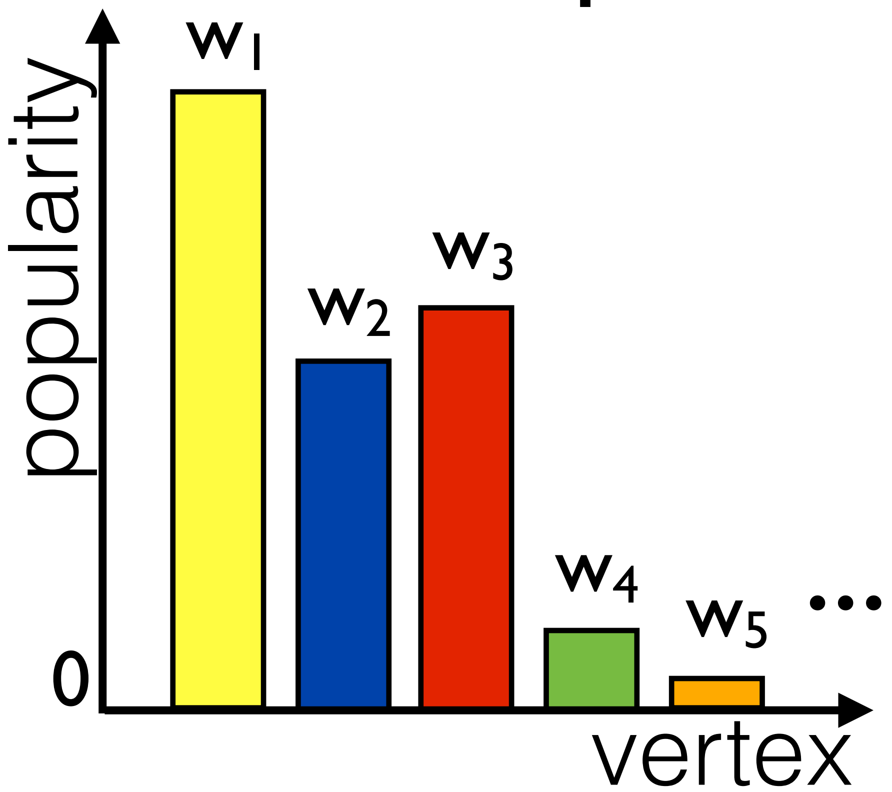

- Need # nodes to go to infinity
  - Need countable ∞ of latent nodes
- Graph frequency model/vertex popularity model
  - Draw a frequency $w_i$ for each vertex $i$
  - Draw edge {i,j} with probability $w_i w_j$



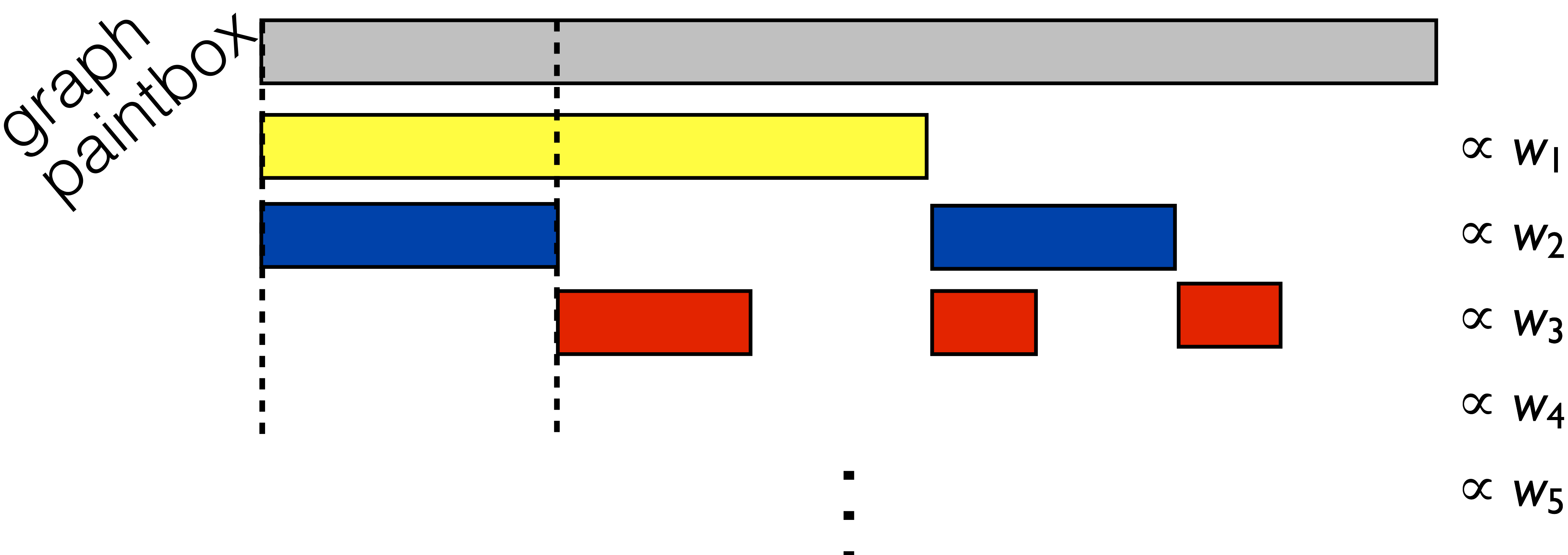



[Caron, Fox 2014; Cai, Campbell, B 2016; Crane, Dempsey 2016a; Palla, Caron, Teh 2016; Herlau, Schmidt 2016]

# How to prove sparsity?



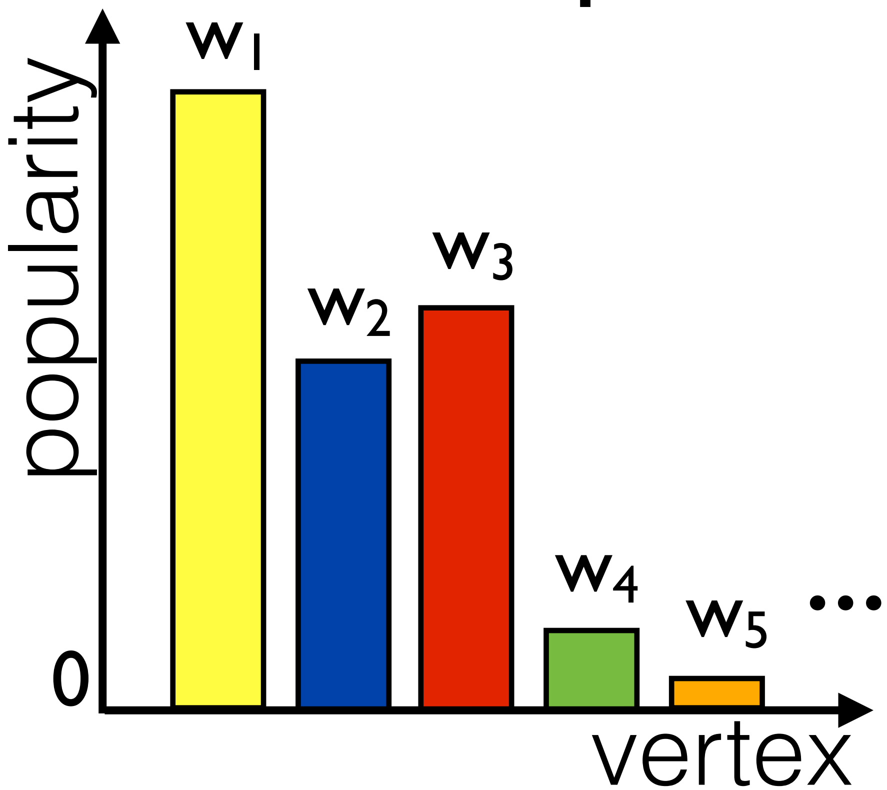

- Need # nodes to go to infinity
  - Need countable ∞ of latent nodes
- Graph frequency model/vertex popularity model
  - Draw a frequency $w_i$ for each vertex $i$
  - Draw edge {i,j} with probability $w_i w_j$



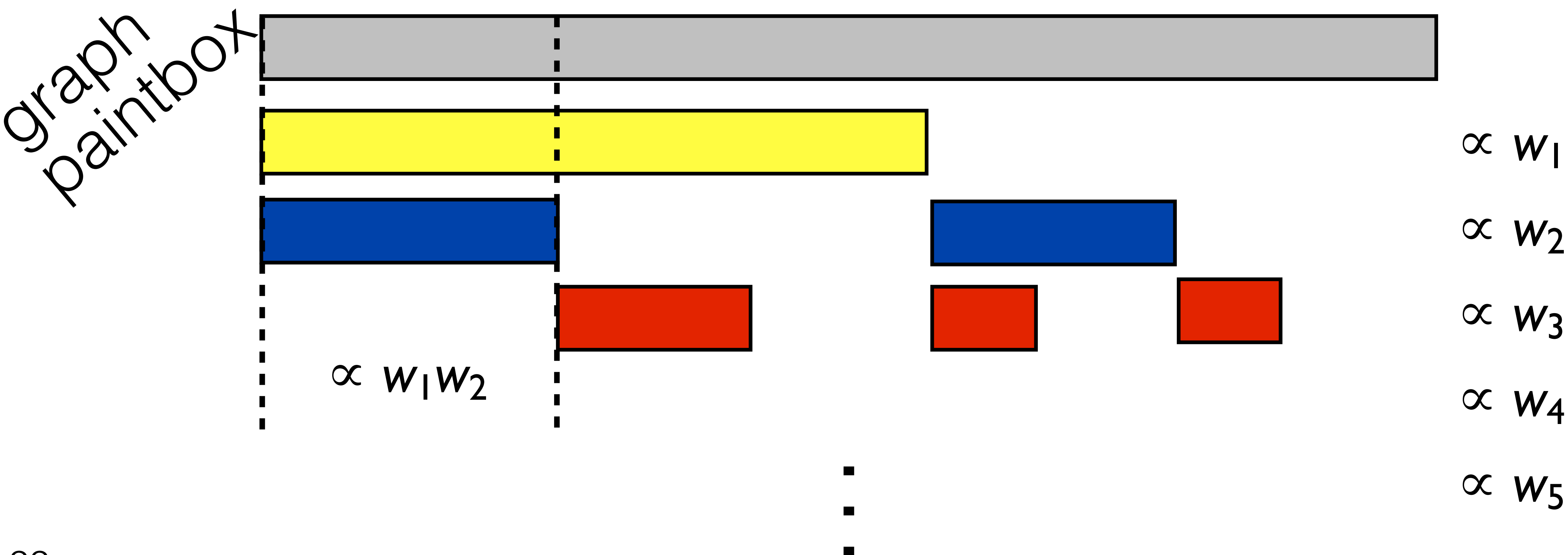



[Caron, Fox 2014; Cai, Campbell, B 2016; Crane, Dempsey 2016a; Palla, Caron, Teh 2016; Herlau, Schmidt 2016]

# How to prove sparsity?



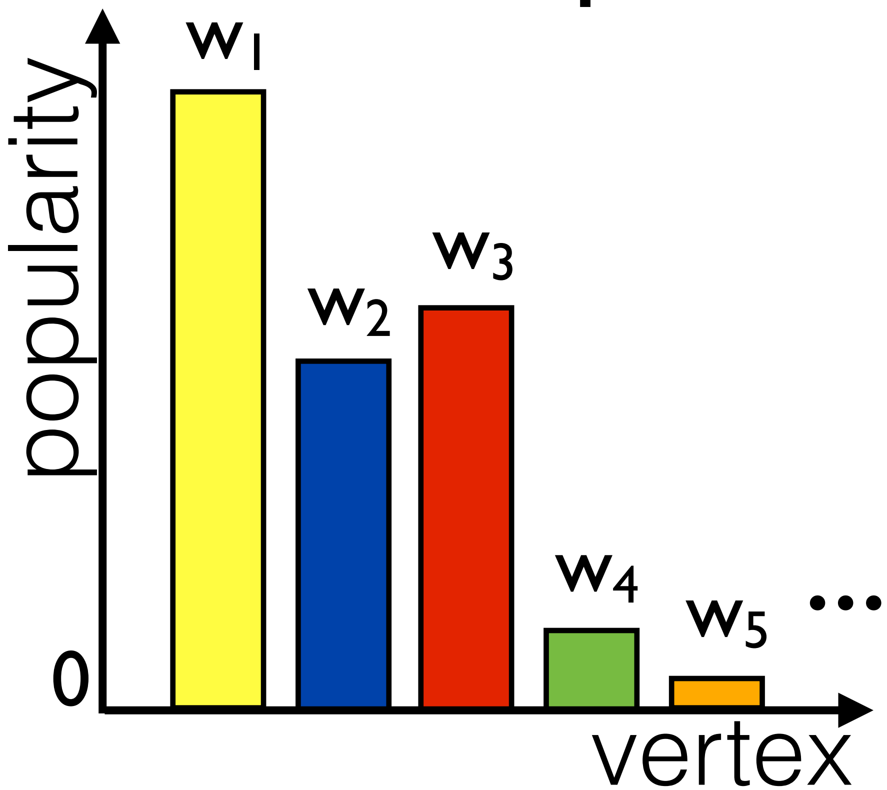

- Need # nodes to go to infinity
  - Need countable ∞ of latent nodes
- Graph frequency model/vertex popularity model
  - Draw a frequency $w_i$ for each vertex $i$
  - Draw edge {i,j} with probability $w_i w_j$





[Caron, Fox 2014; Cai, Campbell, B 2016; Crane, Dempsey 2016a; Palla, Caron, Teh 2016; Herlau, Schmidt 2016]

# How to prove sparsity?

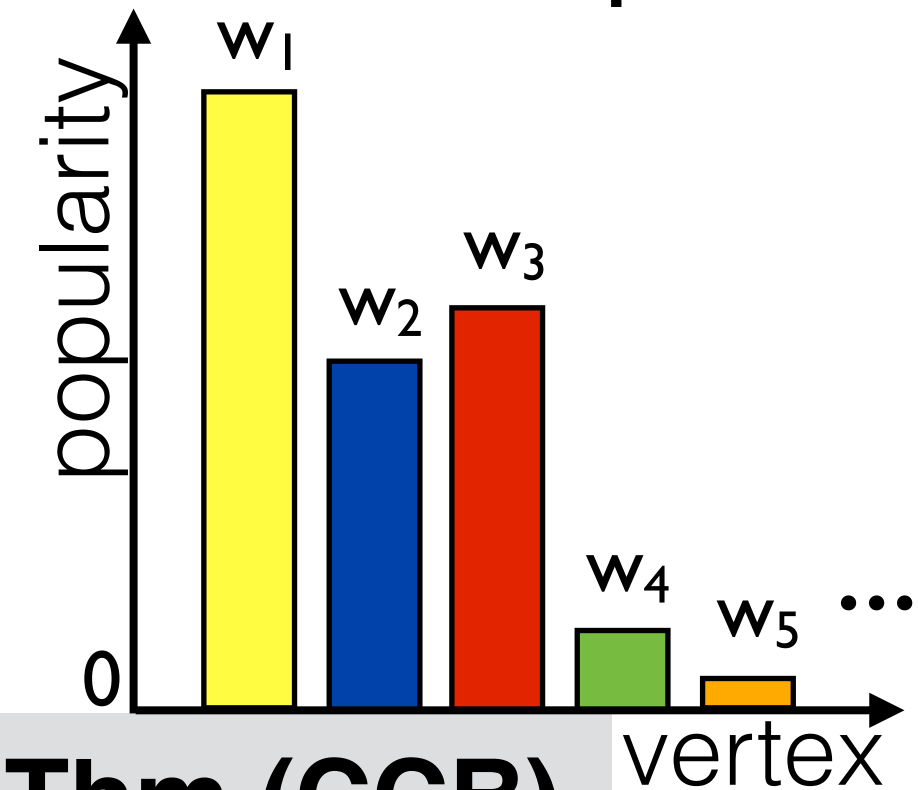



- Need # nodes to go to infinity
  - Need countable ∞ of latent nodes
- Graph frequency model/vertex popularity model
  - Draw a frequency $w_i$ for each vertex $i$
  - Draw edge {i,j} with probability $w_i w_j$



[Cai, Campbell, Broderick 2016]

# How to prove sparsity?



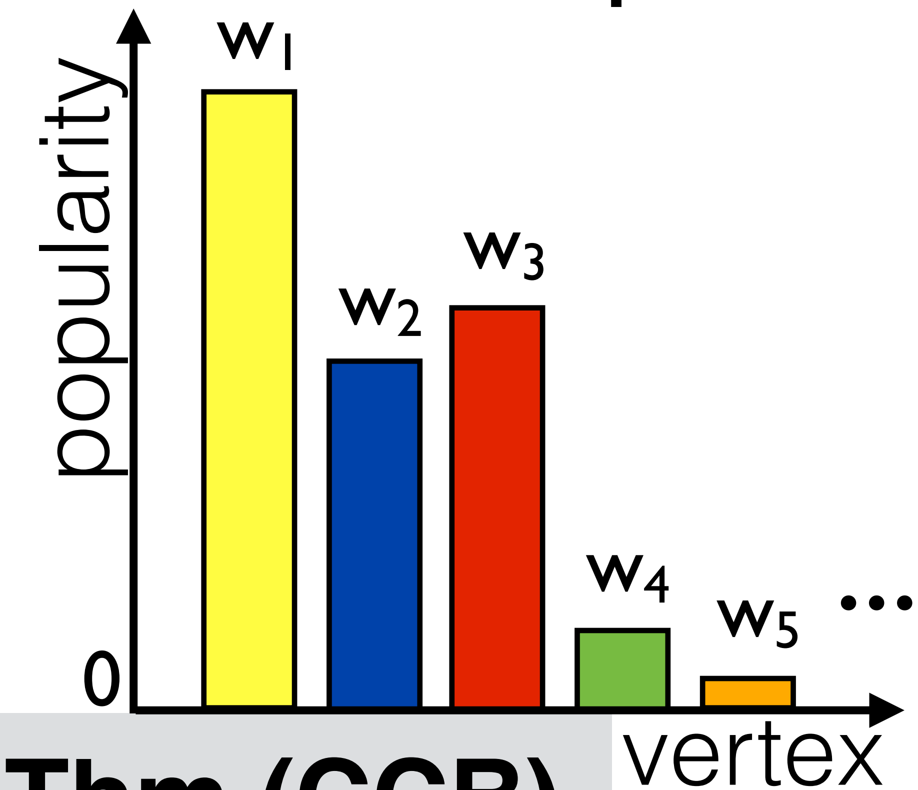

- Need # nodes to go to infinity
  - Need countable ∞ of latent nodes
- Graph frequency model/vertex popularity model
  - Draw a frequency $w_i$ for each vertex $i$
  - Draw edge {i,j} with probability $w_i w_j$

**Thm (CCB).**



[Cai, Campbell, Broderick 2016]

# How to prove sparsity?




- Need # nodes to go to infinity
  - Need countable ∞ of latent nodes
- Graph frequency model/vertex popularity model
  - Draw a frequency $w_i$ for each vertex $i$
  - Draw edge {i,j} with probability $w_i w_j$

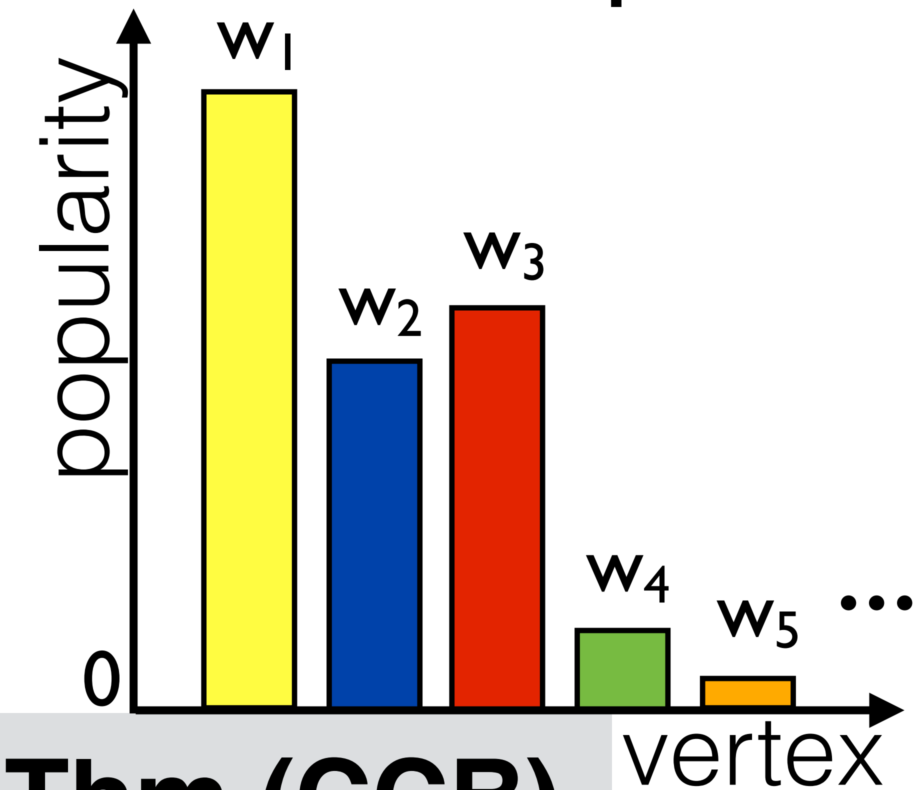
**Thm (CCB).**

- Suppose $w_i \sim \mathrm{PoissonPointProcess}(\nu)$, $\nu$ *regularly varying*:



[Cai, Campbell, Broderick 2016]

# How to prove sparsity?




- Need # nodes to go to infinity
  - Need countable ∞ of latent nodes
- Graph frequency model/vertex popularity model
  - Draw a frequency $w_i$ for each vertex $i$
  - Draw edge {i,j} with probability $w_i w_j$

**Thm (CCB).**

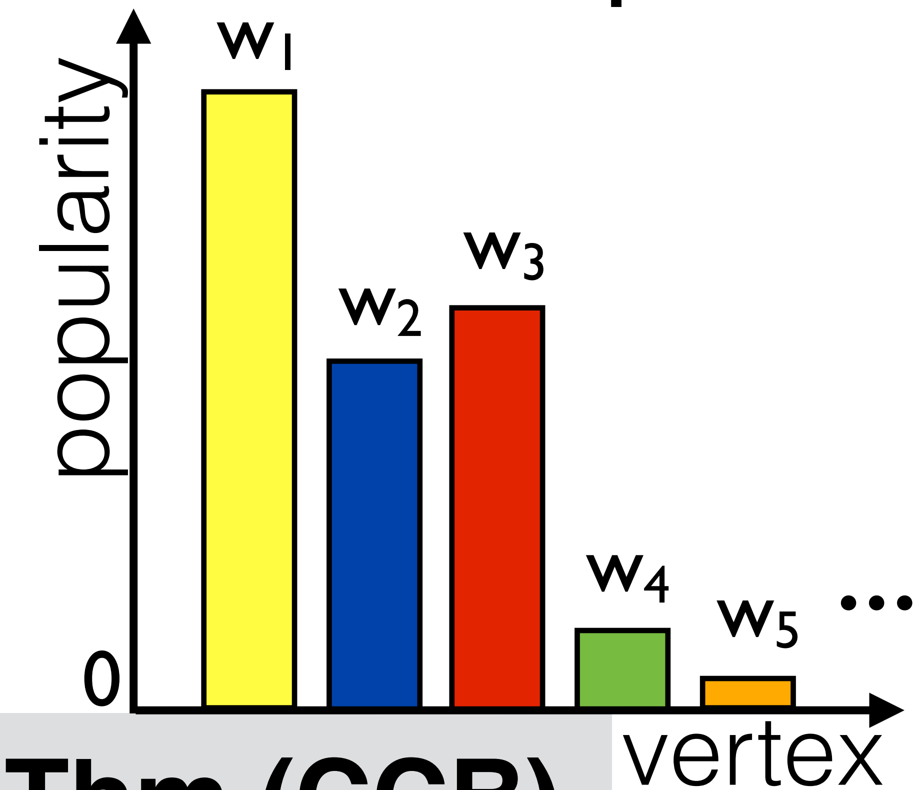
- Suppose $w_i \sim \mathrm{PoissonPointProcess}(\nu)$, $\nu$ *regularly varying*:

$$\int_x^1 \nu(dw) \sim x^{-\alpha} l(x^{-1}), x \to 0 \qquad \forall c > 0, \lim_{x\to\infty} \frac{l(cx)}{l(x)} = 1$$



[Cai, Campbell, Broderick 2016]

# How to prove sparsity?




- Need # nodes to go to infinity
  - Need countable ∞ of latent nodes
- Graph frequency model/vertex popularity model
  - Draw a frequency $w_i$ for each vertex $i$
  - Draw edge {i,j} with probability $w_i w_j$

**Thm (CCB).**

- Suppose $w_i \sim \mathrm{PoissonPointProcess}(\nu)$, $\nu$ *regularly varying*:

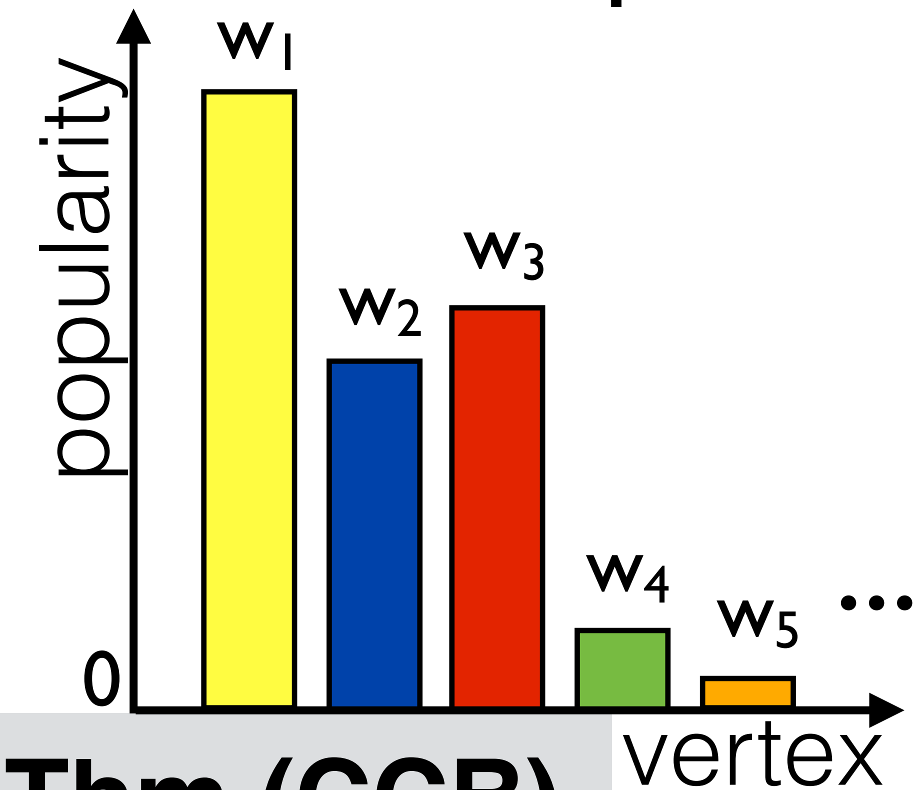
$$\int_x^1 \nu(dw) \sim x^{-\alpha} l(x^{-1}), x \to 0 \qquad \forall c > 0, \lim_{x\to\infty} \frac{l(cx)}{l(x)} = 1$$

- Then $|V_n| \overset{a.s.}{=} \Theta(n^\alpha l(n)), |E_n| \overset{a.s.}{=} \Theta(n)$



[Cai, Campbell, Broderick 2016]

# How to prove sparsity?




- Need # nodes to go to infinity
  - Need countable ∞ of latent nodes
- Graph frequency model/vertex popularity model
  - Draw a frequency $w_i$ for each vertex $i$
  - Draw edge {i,j} with probability $w_i w_j$

**Thm (CCB).**

- Suppose $w_i \sim \mathrm{PoissonPointProcess}(\nu)$, $\nu$ *regularly varying*:

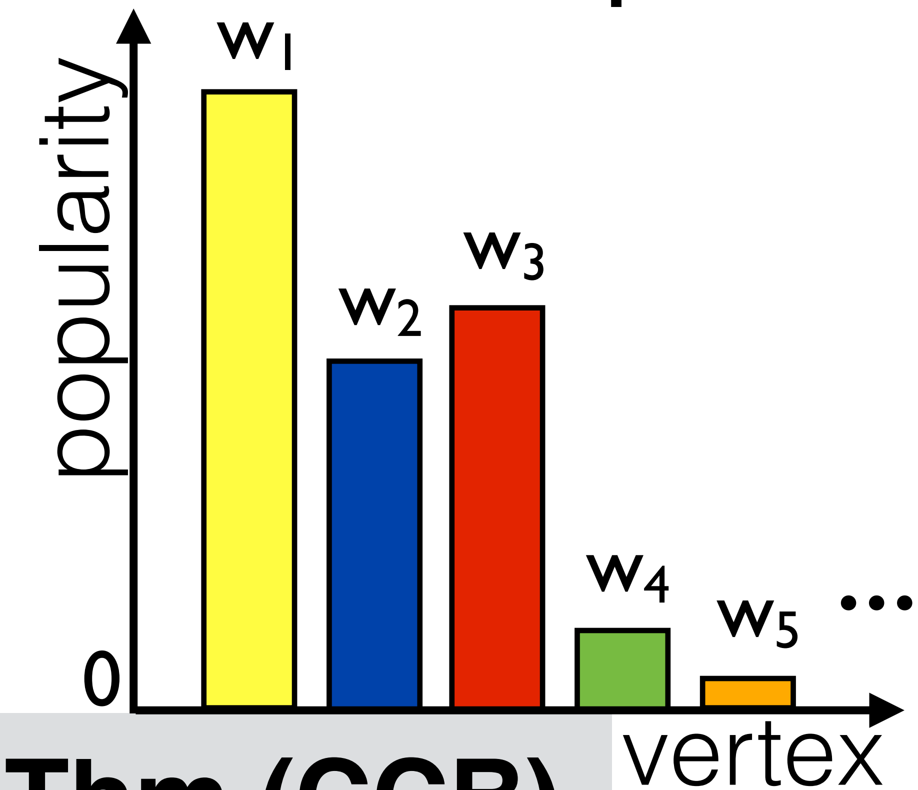
$$\int_x^1 \nu(dw) \sim x^{-\alpha} l(x^{-1}), x \to 0 \qquad \forall c > 0, \lim_{x \to \infty} \frac{l(cx)}{l(x)} = 1$$

- Then $|V_n| \stackrel{a.s.}{=} \Theta(n^\alpha l(n)), |E_n| \stackrel{a.s.}{=} \Theta(n)$
- & for binary edges: $|\bar{E}_n| \stackrel{a.s.}{=} O\left(l(n^{1/2}), \min\left(n^{\frac{1+\alpha}{2}}, l(n) n^{\frac{3\alpha}{2}}\right)\right)$



[Cai, Campbell, Broderick 2016]

# How to prove sparsity?




- Need # nodes to go to infinity
  - Need countable ∞ of latent nodes
- Graph frequency model/vertex popularity model
  - Draw a frequency $w_i$ for each vertex $i$
  - Draw edge {i,j} with probability $w_i w_j$

**Thm (CCB).**

- Suppose $w_i \sim \mathrm{PoissonPointProcess}(\nu)$, $\nu$ *regularly varying*:

$$\int_x^1 \nu(dw) \sim x^{-\alpha} l(x^{-1}), x \to 0 \qquad \forall c > 0, \lim_{x\to\infty} \frac{l(cx)}{l(x)} = 1$$

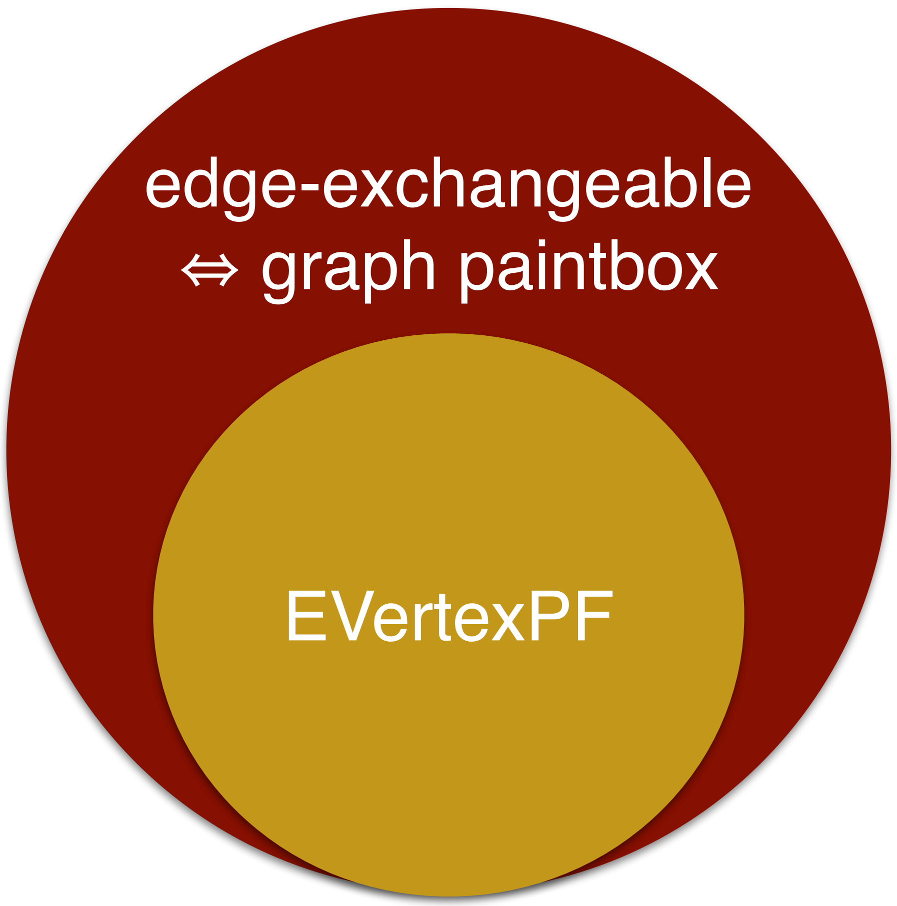
- Then $|V_n| \overset{a.s.}{=} \Theta(n^\alpha l(n)), |E_n| \overset{a.s.}{=} \Theta(n)$
- & for binary edges: $|\bar{E}_n| \overset{a.s.}{=} O\left(l(n^{1/2}), \min\left(n^{\frac{1+\alpha}{2}}, l(n) n^{\frac{3\alpha}{2}}\right)\right)$

**Cor (CCB). A wide class of edge-exchangeable graph models yields sparse graph sequences.**



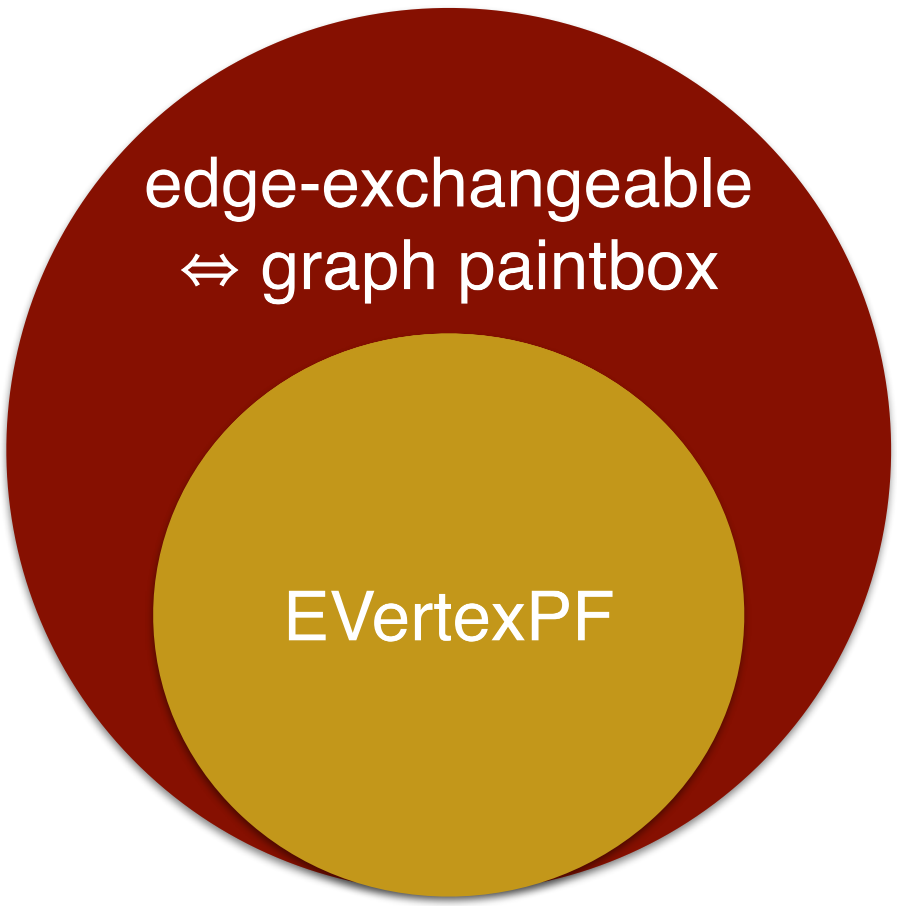
[Cai, Campbell, Broderick 2016]

# What we know so far



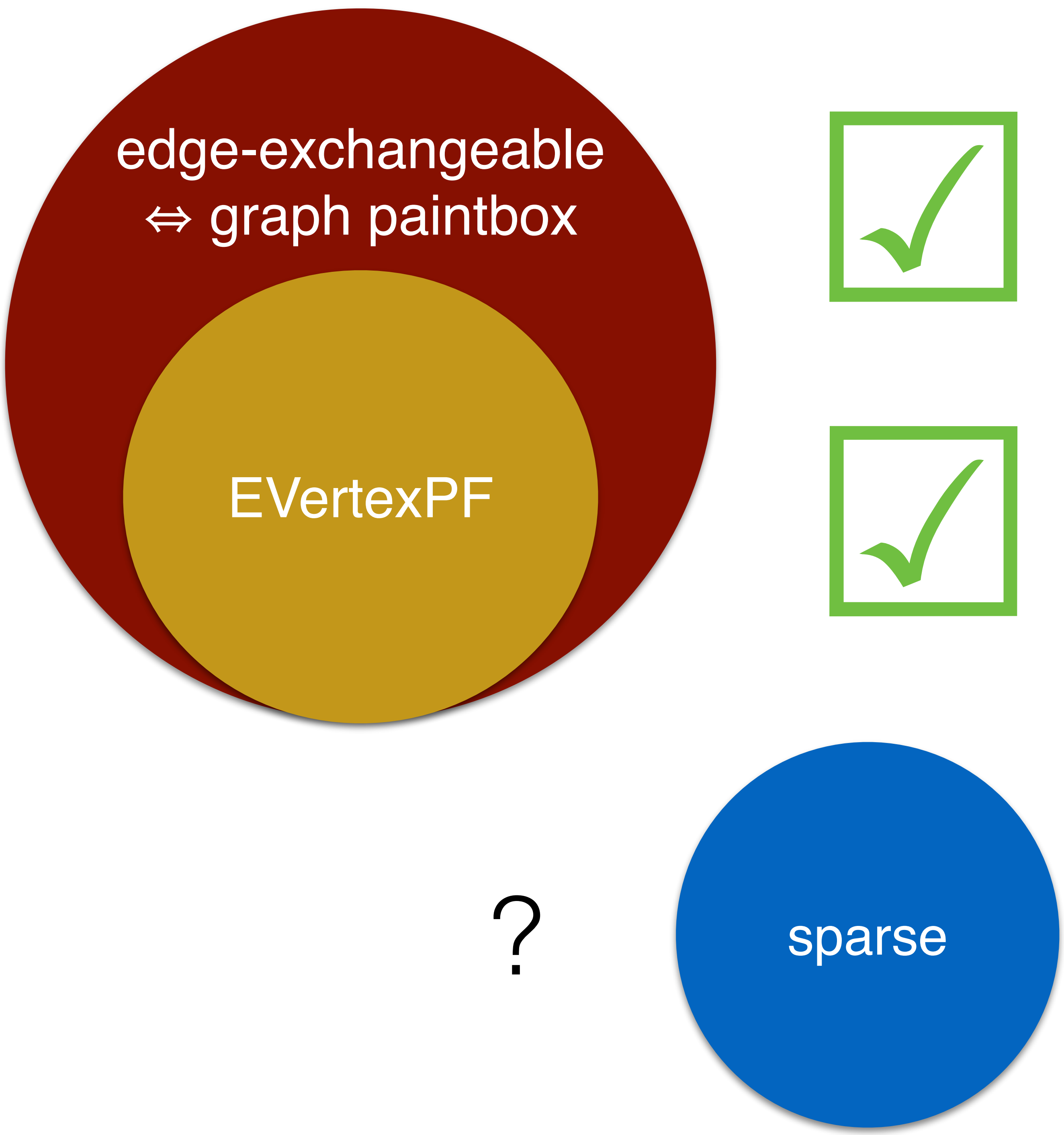

- Thm 1: characterization theorem for edge-exchangeable graphs
- Goal 2: sparsity theorem for edge-exchangeable graphs

# What we know so far



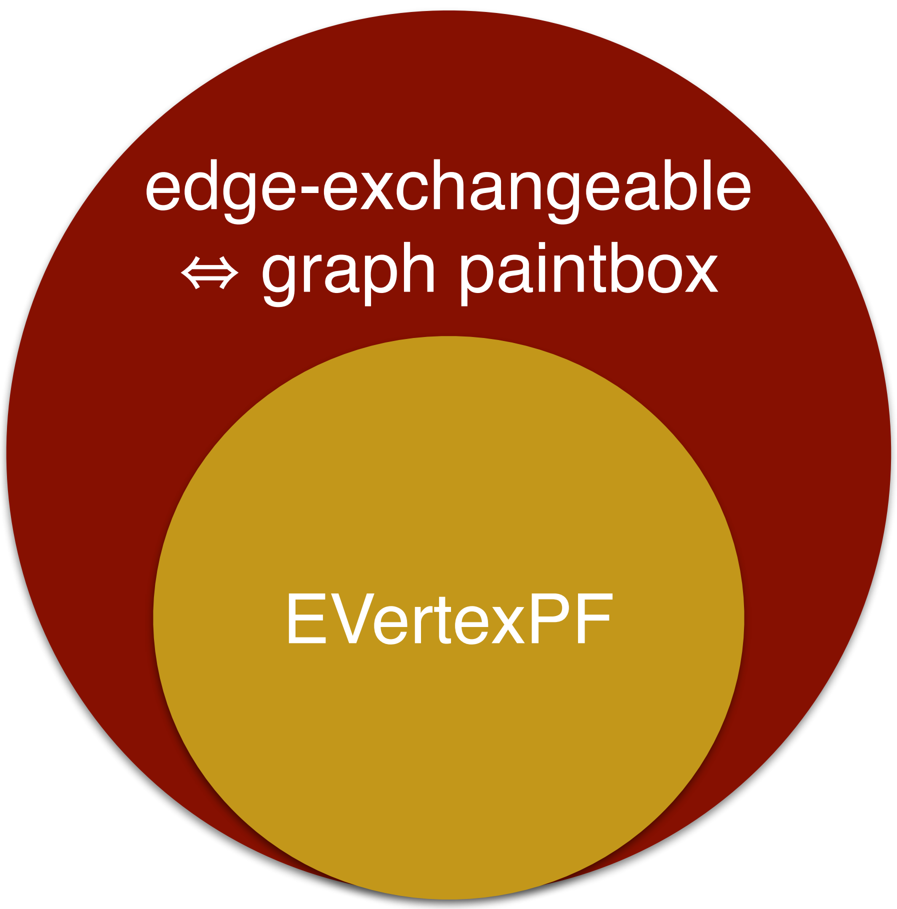

- Thm 1: characterization theorem for edge-exchangeable graphs
- Thm 2: sparsity exists in edge-exchangeable graphs

# What we know so far

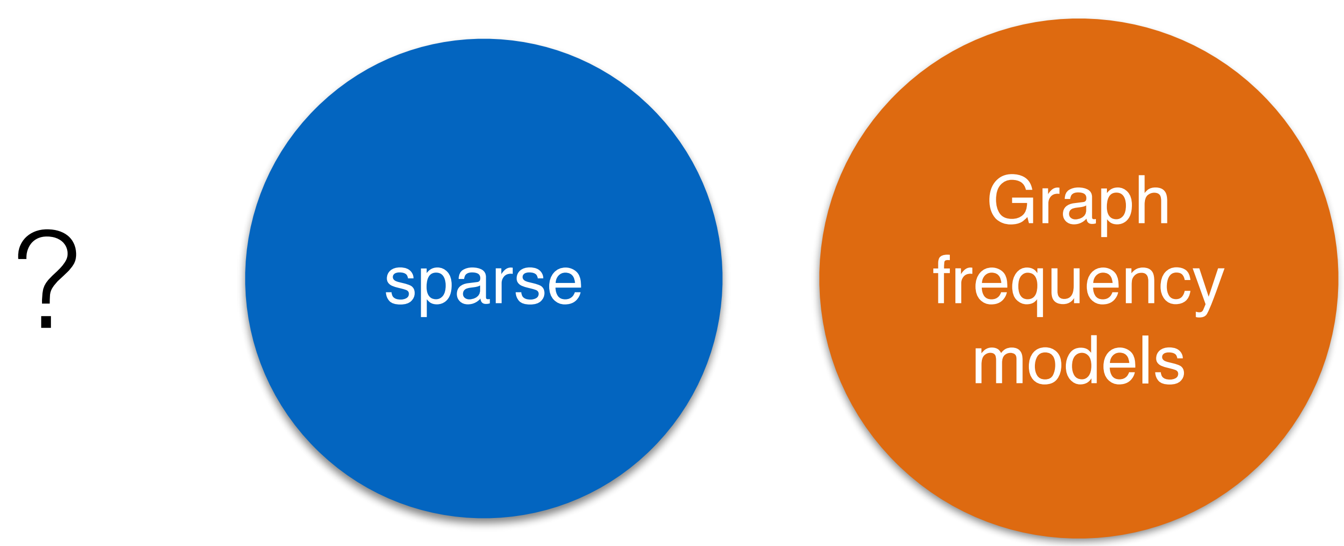



- Thm 1: characterization theorem for edge-exchangeable graphs
- Thm 2: sparsity exists in edge-exchangeable graphs

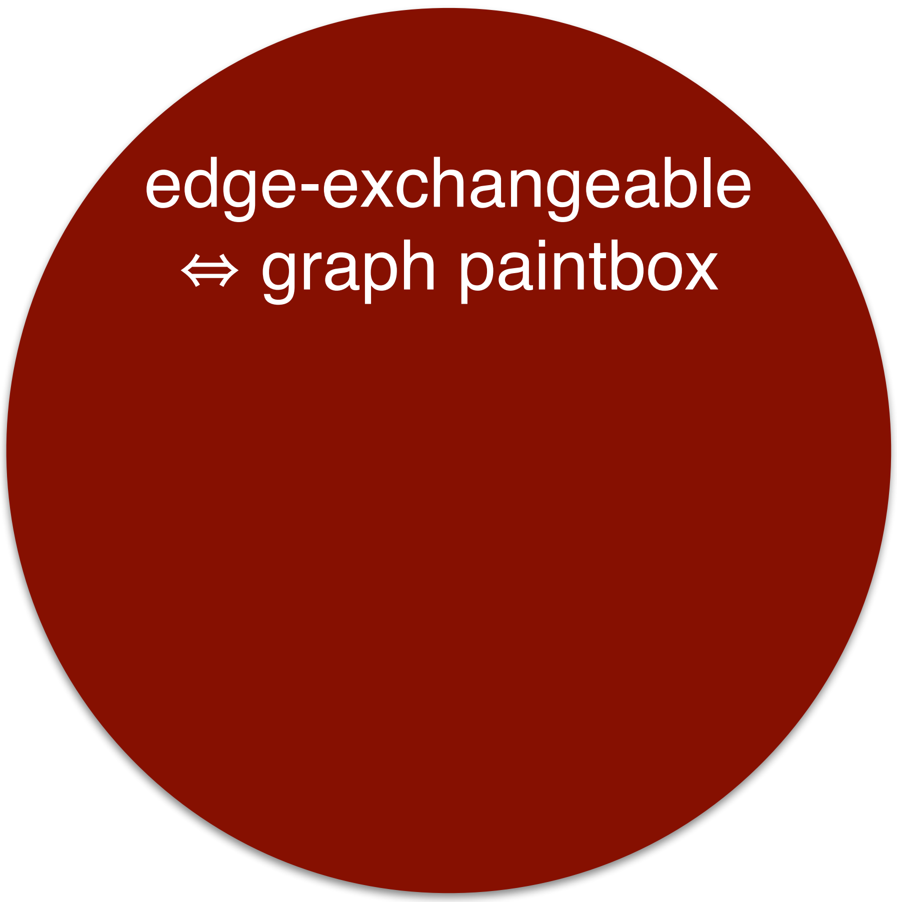
# What we know so far



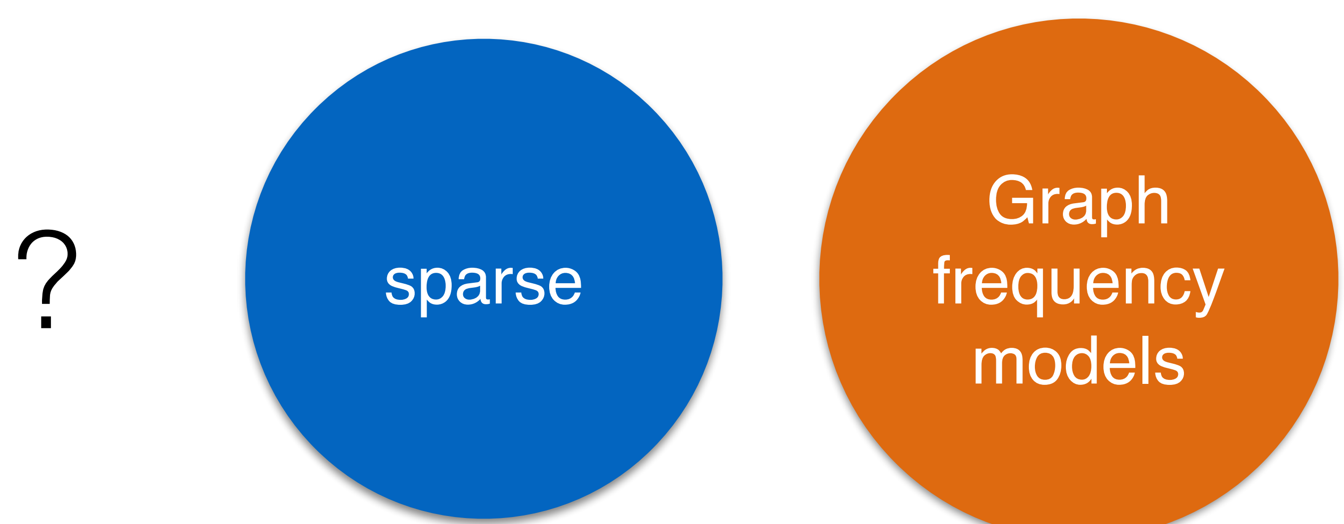

- Thm 1: characterization theorem for edge-exchangeable graphs
- Thm 2: sparsity exists in edge-exchangeable graphs



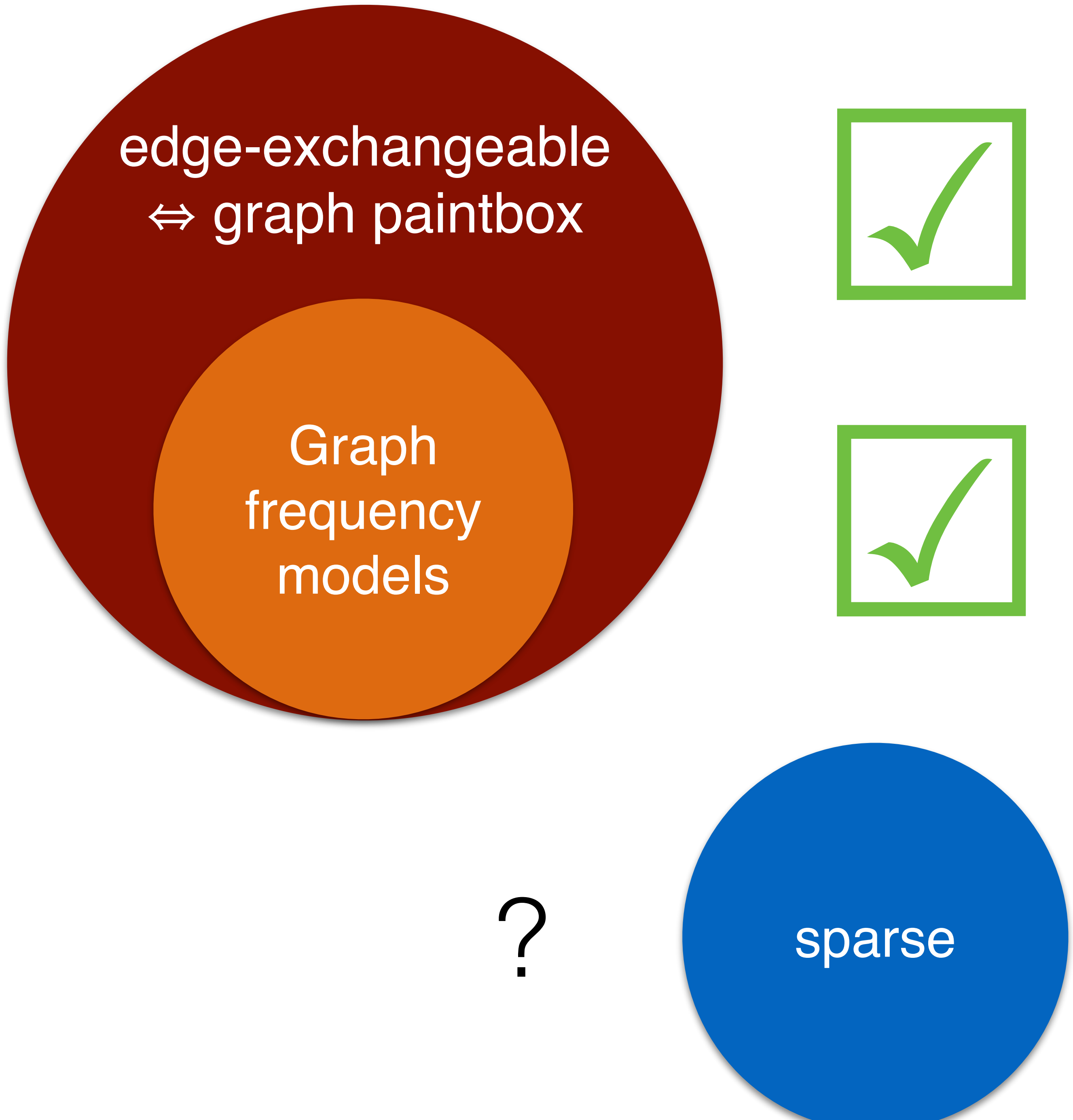
# What we know so far




- Thm 1: characterization theorem for edge-exchangeable graphs
- Thm 2: sparsity exists in edge-exchangeable graphs

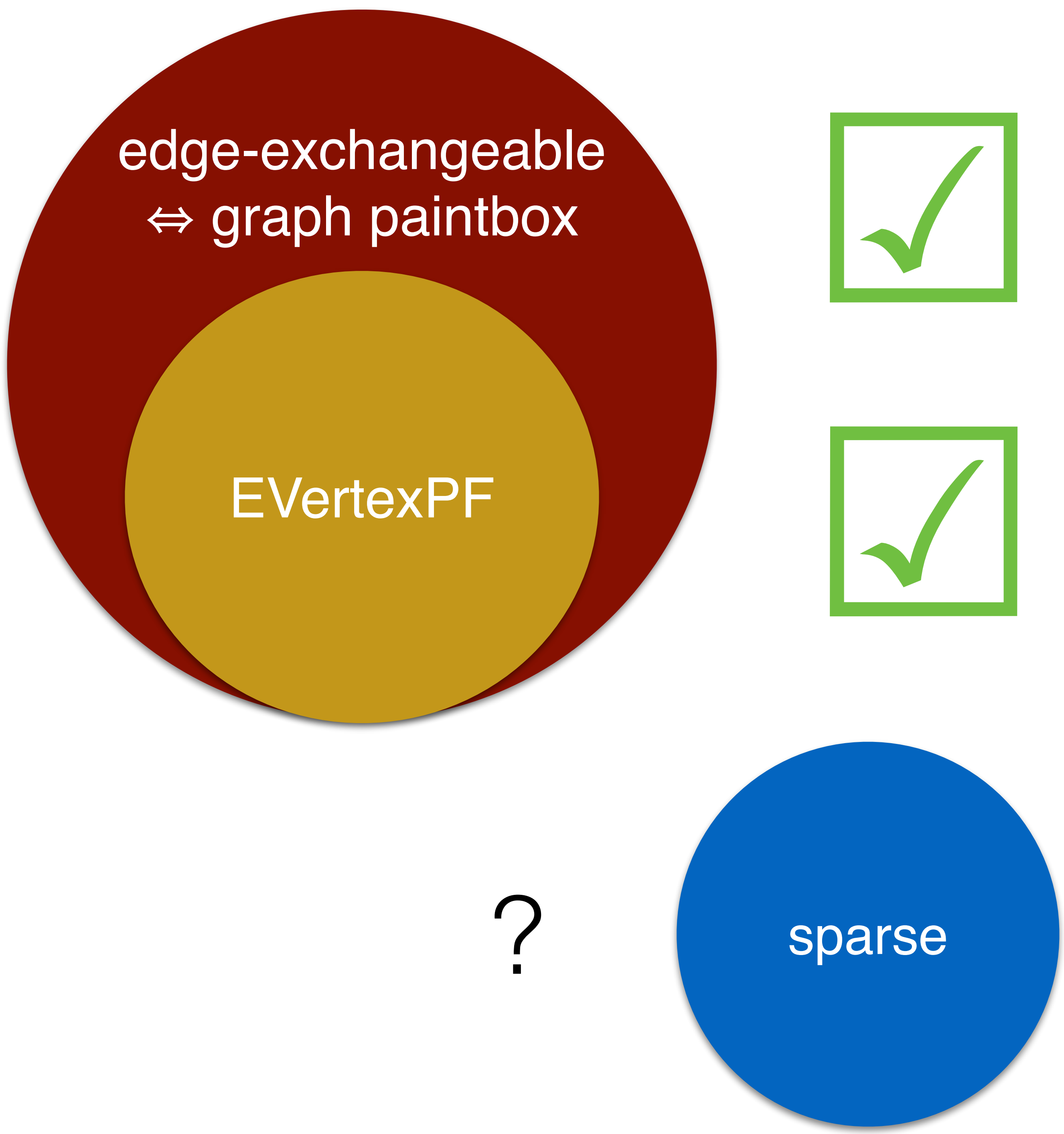


# What we know so far



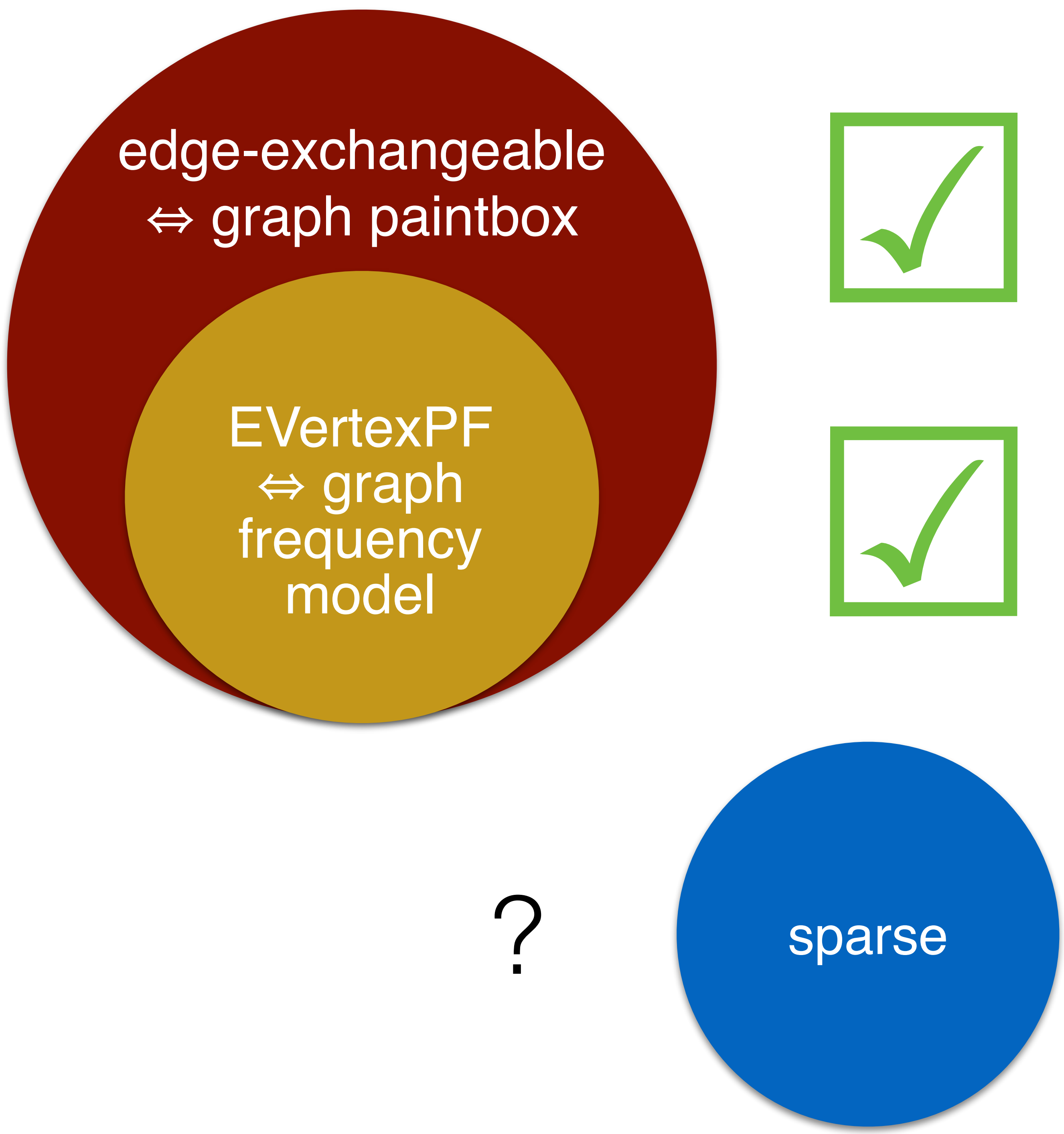

- Thm 1: characterization theorem for edge-exchangeable graphs
- Thm 2: sparsity exists in edge-exchangeable graphs

# What we know so far



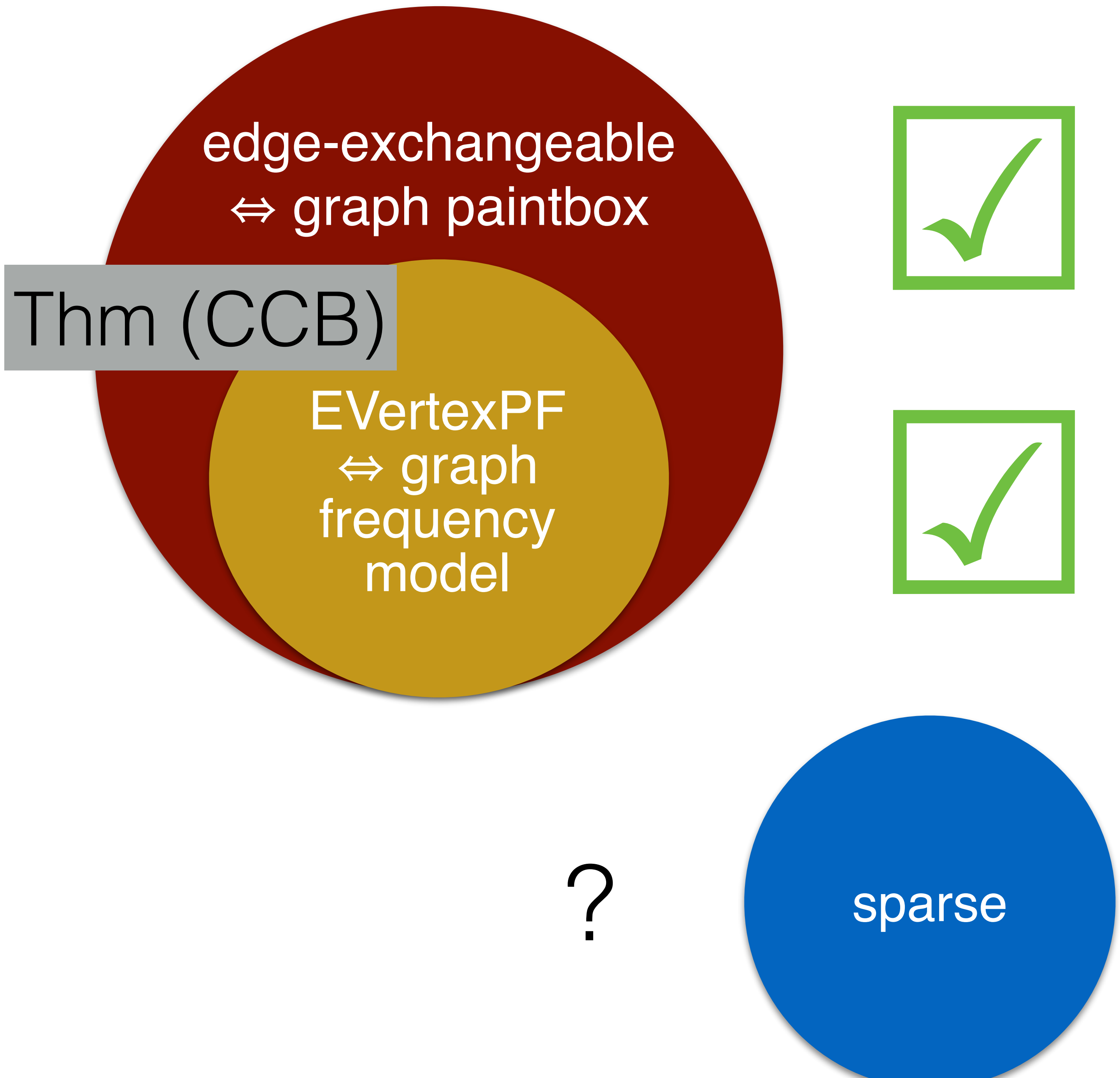

- Thm 1: characterization theorem for edge-exchangeable graphs
- Thm 2: sparsity exists in edge-exchangeable graphs

# What we know so far




- Thm 1: characterization theorem for edge-exchangeable graphs
- Thm 2: sparsity exists in edge-exchangeable graphs

# What we know so far

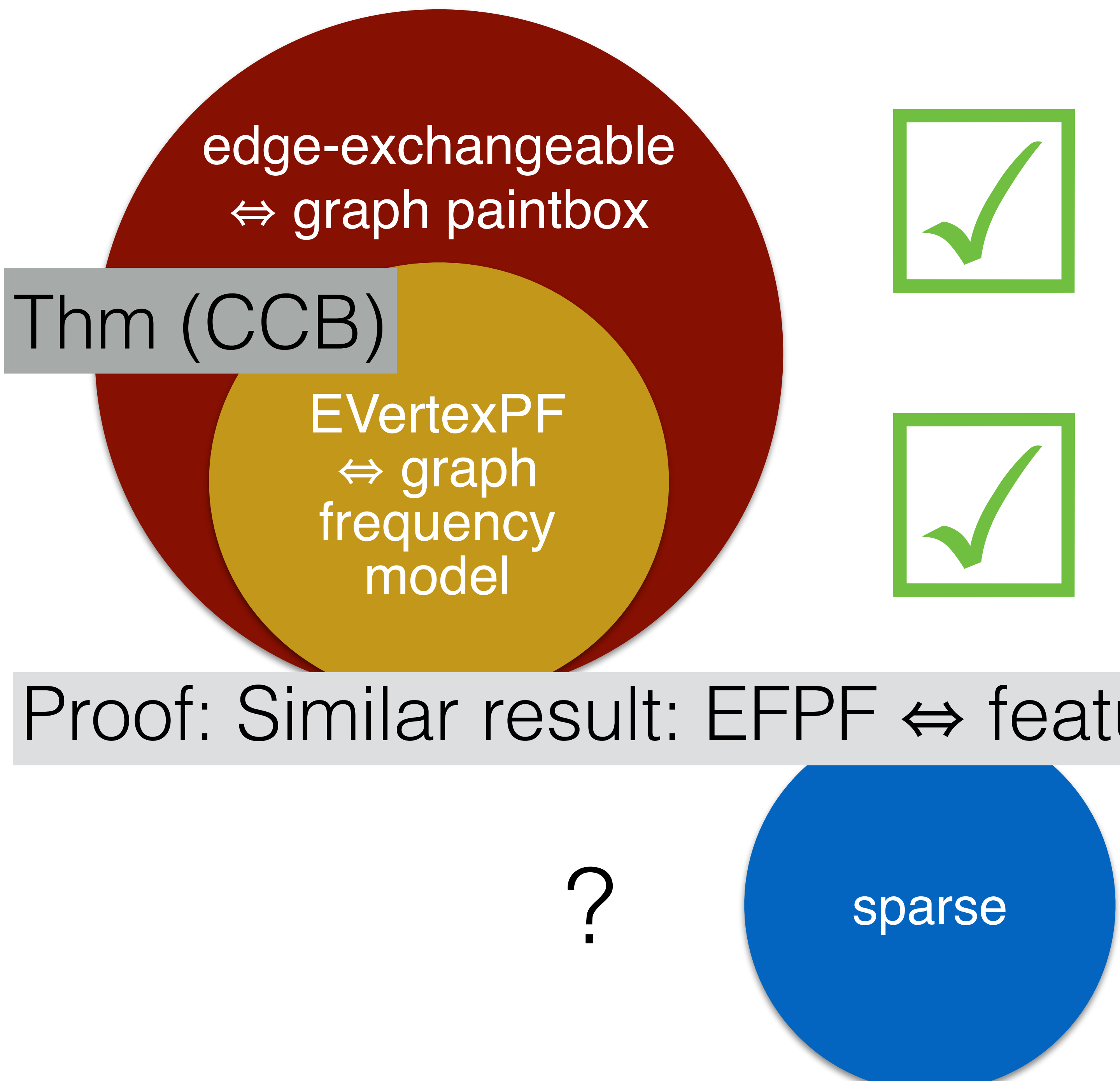



- Thm 1: characterization theorem for edge-exchangeable graphs
- Thm 2: sparsity exists in edge-exchangeable graphs



[Campbell, Cai, Broderick 2016]

# What we know so far

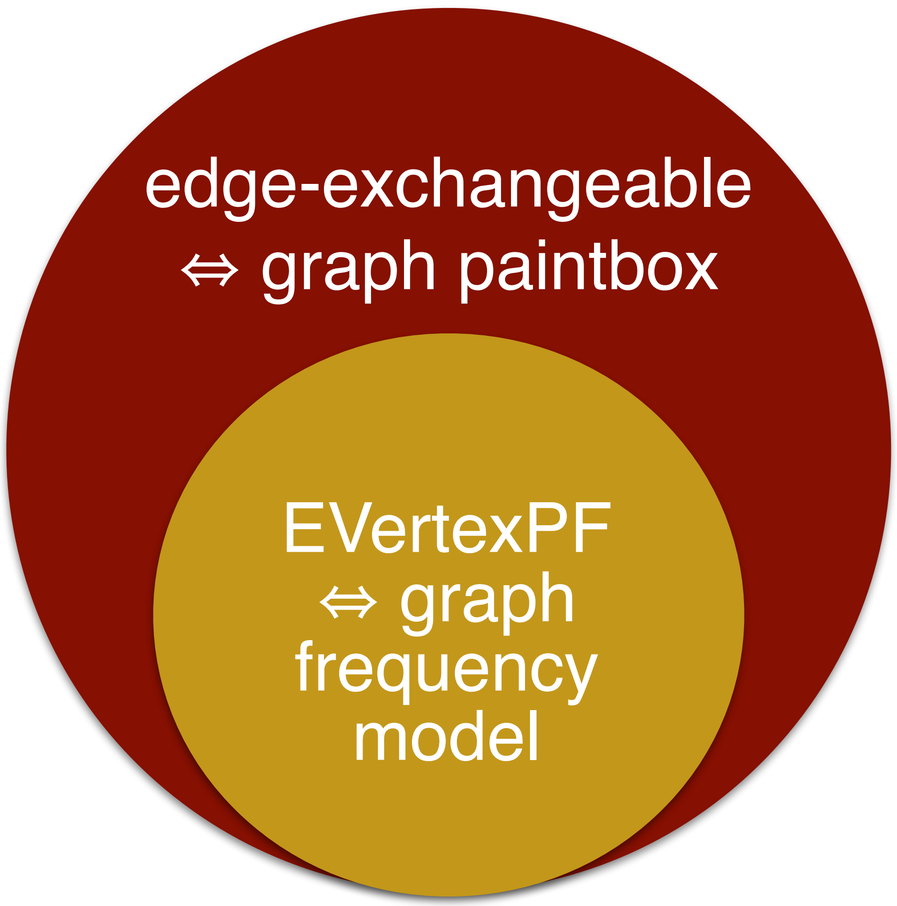



- Thm 1: characterization theorem for edge-exchangeable graphs
- Thm 2: sparsity exists in edge-exchangeable graphs

Proof: Similar result: EFPF ⇔ feature frequency model

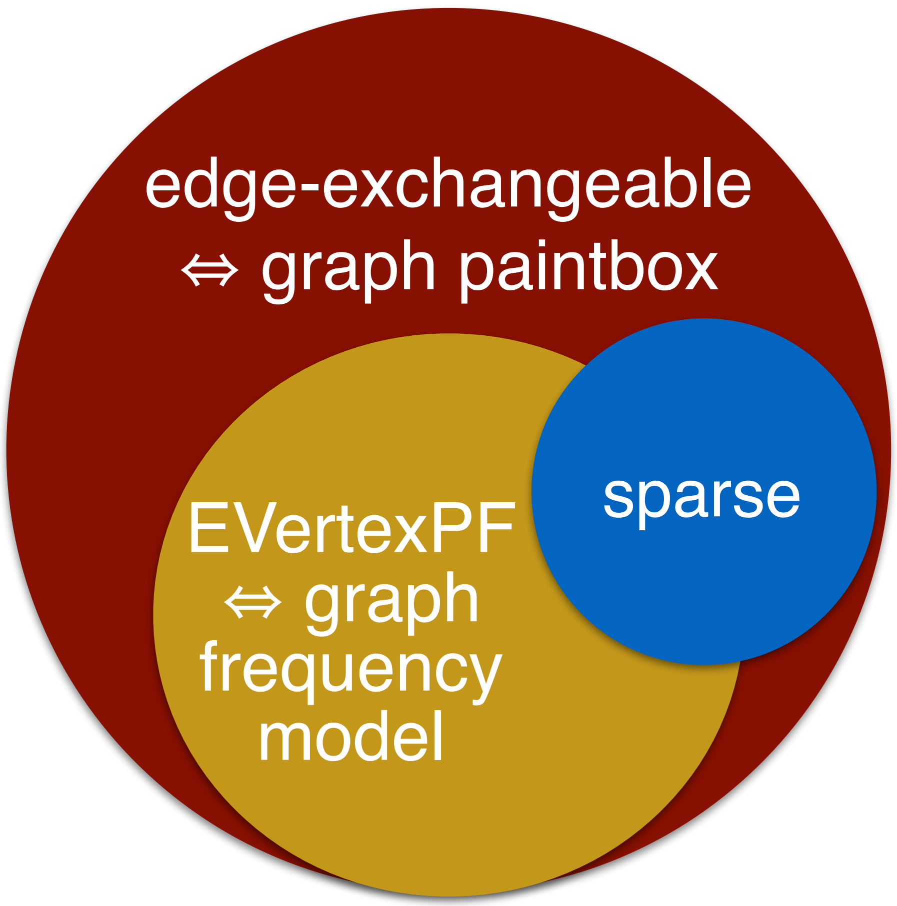
[Campbell, Cai, Broderick 2016; Broderick, Jordan, Pitman 2013]

# What we know so far



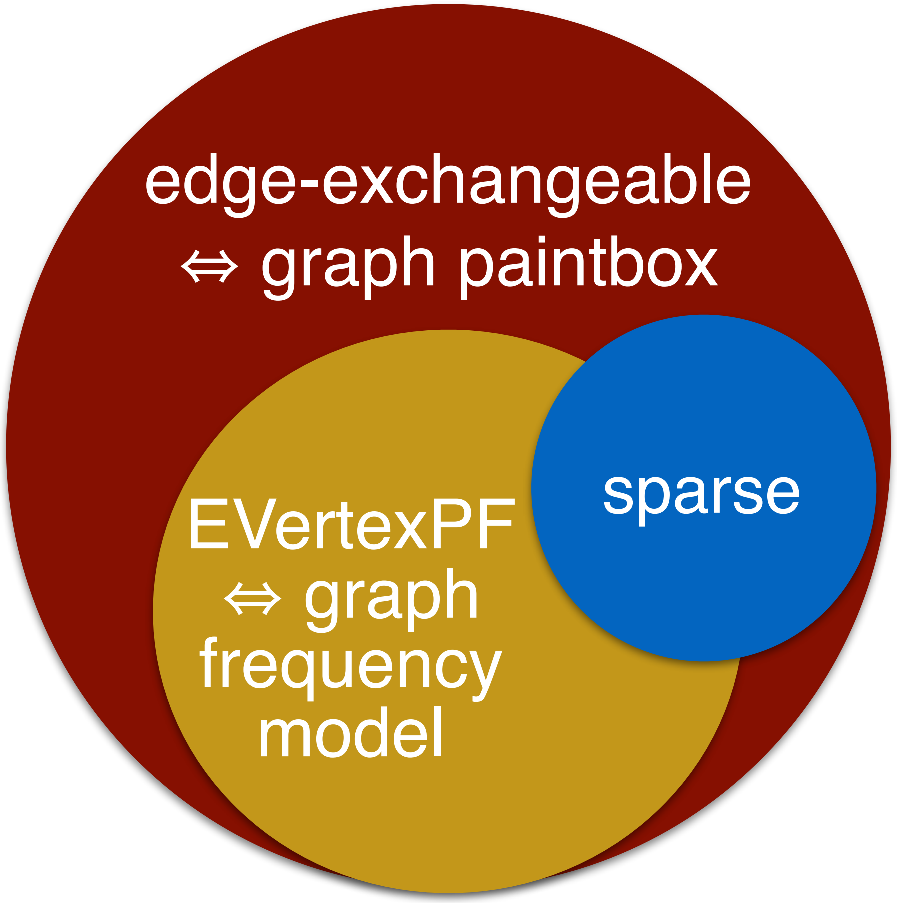

- Thm 1: characterization theorem for edge-exchangeable graphs
- Thm 2: sparsity exists in edge-exchangeable graphs

# What we know so far



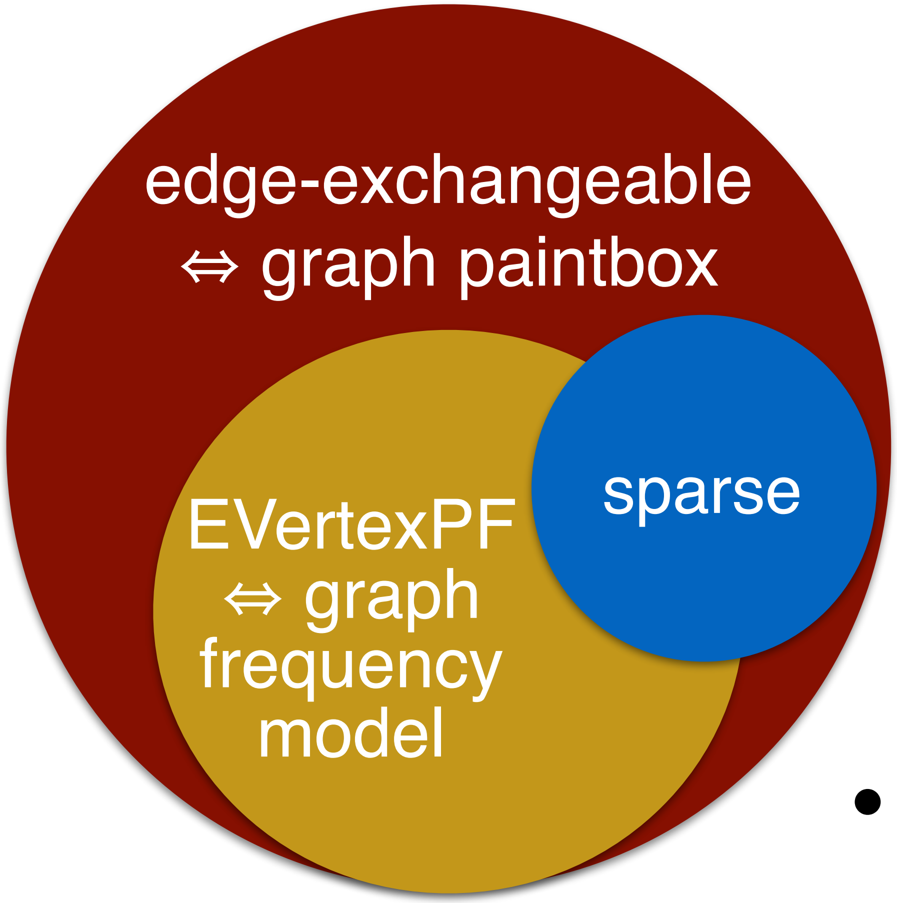

- Thm 1: characterization theorem for edge-exchangeable graphs
- Thm 2: sparsity exists in edge-exchangeable graphs

# What we know so far




- Thm 1: characterization theorem for edge-exchangeable graphs
- Thm 2: sparsity exists in edge-exchangeable graphs

# What we know so far

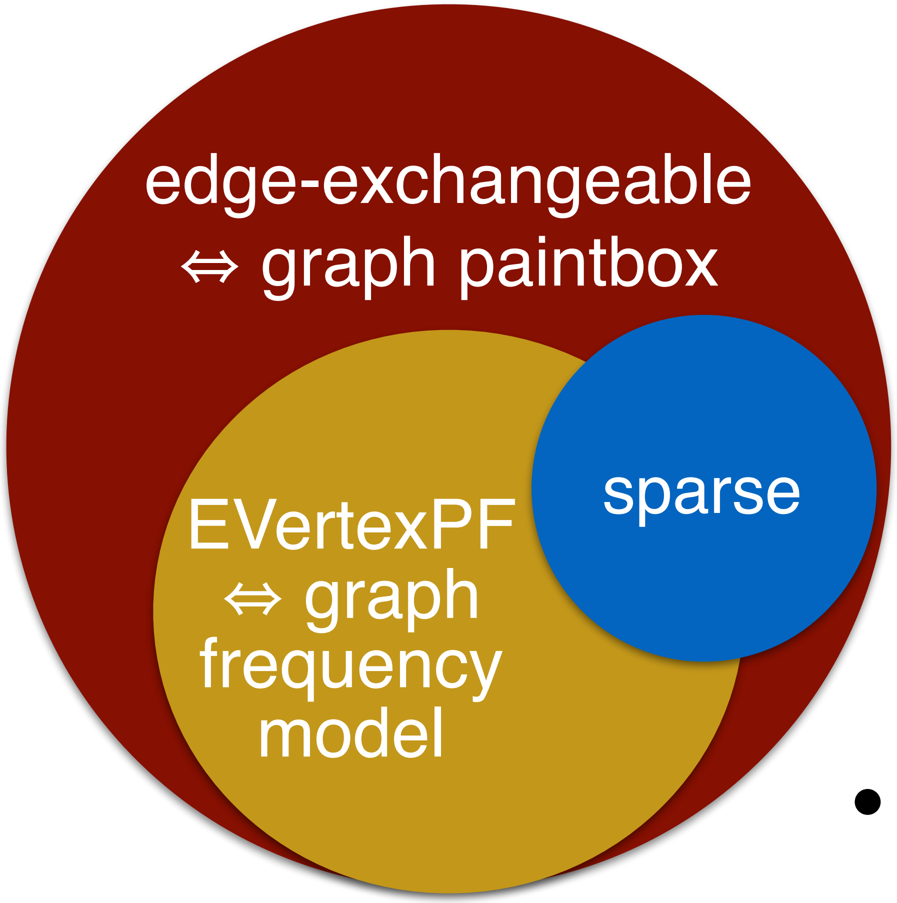



- Thm 1: characterization theorem for edge-exchangeable graphs
- Thm 2: sparsity exists in edge-exchangeable graphs
- Also: Characterization of “dust” (new for features, traits, and graphs)



[Campbell, Cai, Broderick 2016]

# What we know so far




- Thm 1: characterization theorem for edge-exchangeable graphs

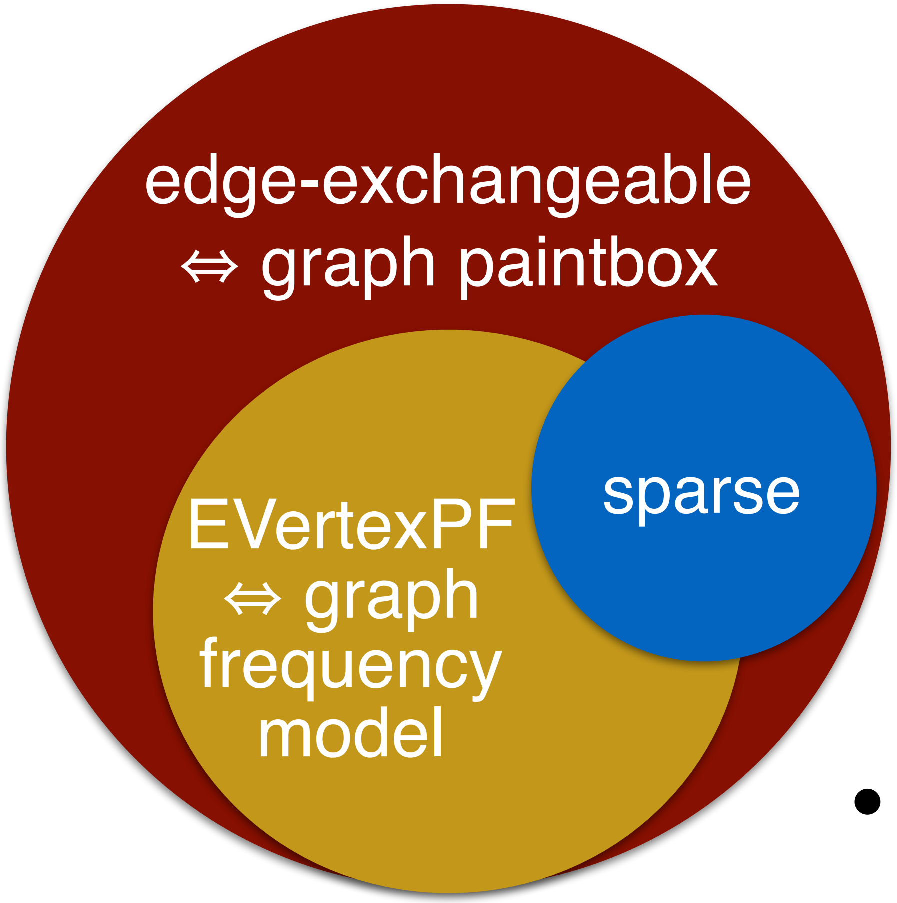
- Thm 2: sparsity exists in edge-exchangeable graphs
- Also: Characterization of “dust” (new for features, traits, and graphs)

# Frontier



[Campbell, Cai, Broderick 2016]

# What we know so far




- Thm 1: characterization theorem for edge-exchangeable graphs
- Thm 2: sparsity exists in edge-exchangeable graphs
- Also: Characterization of “dust” (new for features, traits, and graphs)

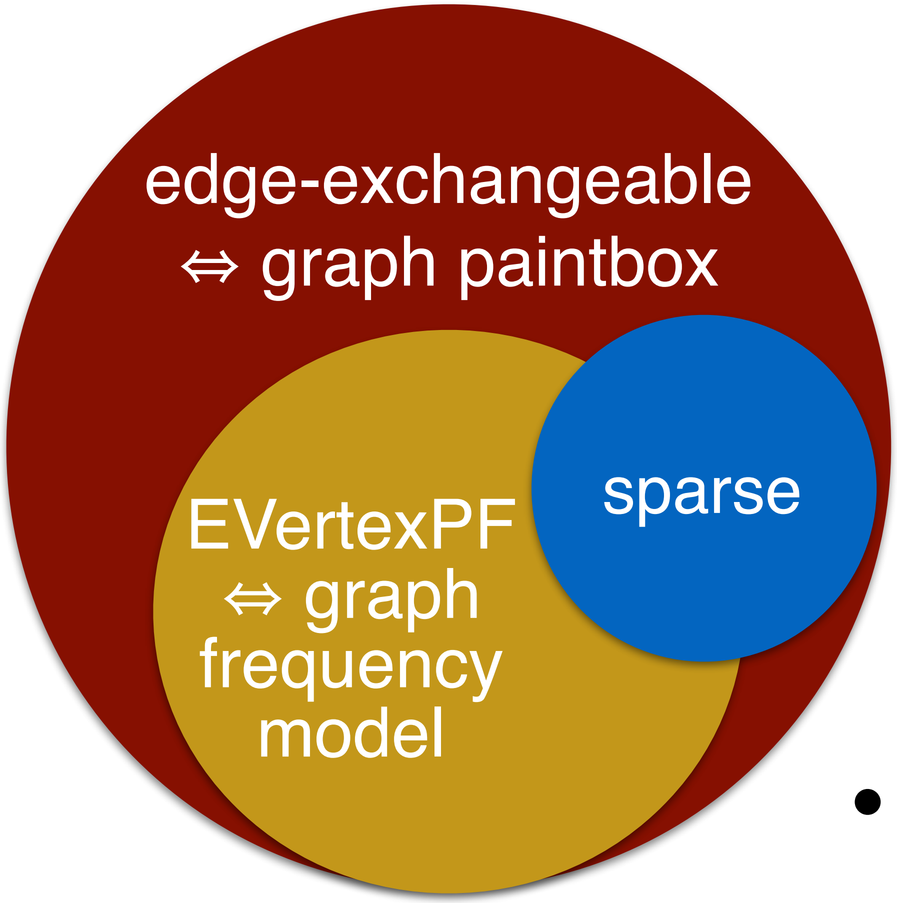
# Frontier

- Characterize all sparse, edge-exchangeable graphs



[Campbell, Cai, Broderick 2016]

# What we know so far




- Thm 1: characterization theorem for edge-exchangeable graphs
- Thm 2: sparsity exists in edge-exchangeable graphs
- Also: Characterization of “dust” (new for features, traits, and graphs)

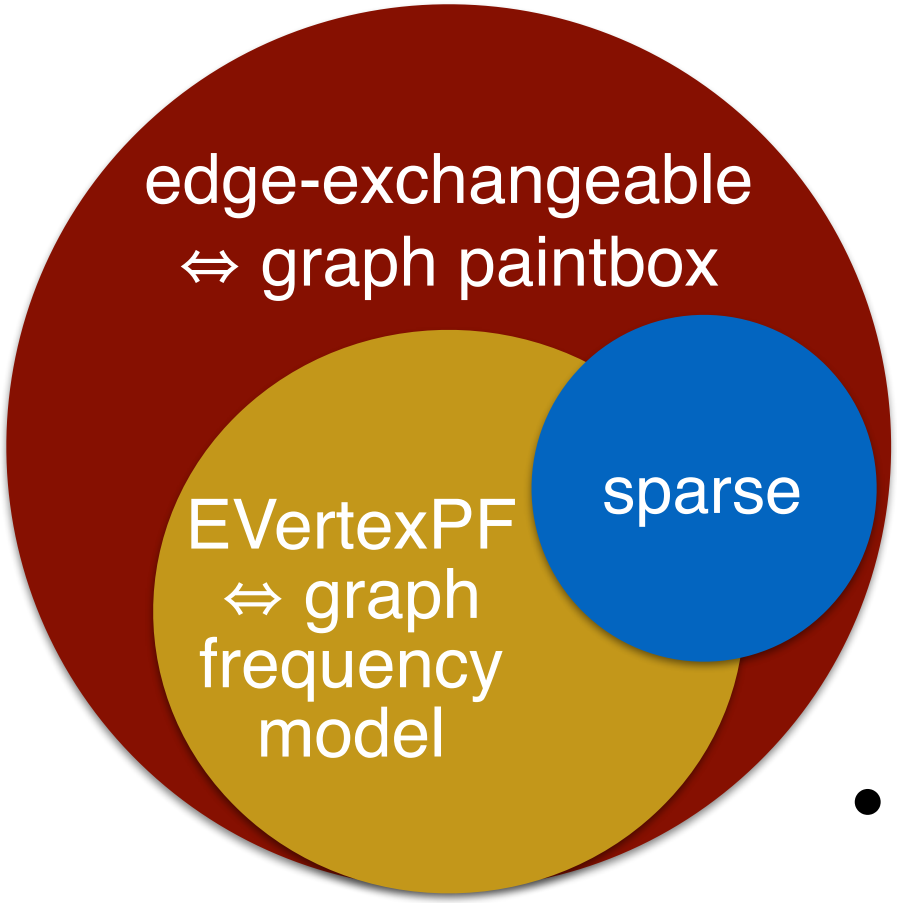
# Frontier

- Characterize all sparse, edge-exchangeable graphs
- Models and inference



[Campbell, Cai, Broderick 2016; Broderick, Jordan, Pitman 2012]

# What we know so far




- ✓ Thm 1: characterization theorem for edge-exchangeable graphs
- ✓ Thm 2: sparsity exists in edge-exchangeable graphs
- Also: Characterization of "dust" (new for features, traits, and graphs)

# Frontier

- Characterize all sparse, edge-exchangeable graphs
- Models and inference
- Truncation approximations

 [Campbell, Cai, Broderick 2016; Broderick, Jordan, Pitman 2012; Campbell*, Huggins*, How, Broderick 2016]

# References (page 1 of 2)


T Broderick, MI Jordan, and J Pitman. Beta processes, stick-breaking, and power laws. *Bayesian Analysis*, 2012.

T Broderick, J Pitman, and MI Jordan. Feature allocations, probability functions, and paintboxes. *Bayesian Analysis*, 2013.

T Broderick, MI Jordan, and J Pitman. Cluster and feature modeling from combinatorial stochastic processes. *Statistical Science*, 2013.

**D Cai, T Campbell, and T Broderick. Edge-exchangeable graphs and sparsity. *NIPS*, 2016.**

- ***NIPS 2015 Workshop on Networks in the Social & Information Sciences.***
- ***NIPS 2015 Workshop, Bayesian Nonparametrics: The Next Generation.***

**T Campbell, D Cai, and Broderick T. Exchangeable trait allocations. Submitted. ArXiv:1609.09147.**

- ***NIPS 2016 Workshop on Adaptive & Scalable Nonparametric Methods in ML.***
- ***NIPS 2016 Workshop on Practical Bayesian Nonparametrics.***

T Campbell*, JH Huggins*, J How, and T Broderick. Truncated random measures. Submitted. arXiv 1603.00861. Poster at ISBA 2016.

H Crane and W Dempsey. Atypical scaling behavior persists in real world interaction networks. arXiv 1509.08184, 2015.

H Crane and W Dempsey. A framework for statistical network modeling. arXiv 1509.08185, 2015.

H Crane and W Dempsey. Edge exchangeable models for network data. arXiv 1603.04571, 2016.

H Crane and W Dempsey. Relational exchangeability. arXiv 1607.06762, 2016.

S Williamson. Nonparametric Network Models for Link Prediction. *JMLR*, 2016.


* Shared first authorship

# References (page 2 of 2)

EM Airoldi, DM Blei, SE Fienberg, & EP Xing. Mixed membership stochastic blockmodels. *Journal of Machine Learning Research*, 2008.

DJ Aldous. Representations for partially exchangeable arrays of random variables. *Journal of Multivariate Analysis*, 1981.

DJ Aldous. Exchangeability and related topics. *École d'été de probabilités de Saint-Flour, XIII—1983, Lecture Notes in Mathematics*, 1985.

C Borgs, J Chayes, H Cohn, and N Holden. Sparse exchangeable graphs and their limits via graphon processes. arXiv 1601.07134, 2016.

F Caron and E Fox. Sparse graphs using exchangeable random measures. arXiv 1401.1137, 2014.

T Herlau and M Schmidt. Completely random measures for modelling block-structured sparse networks. NIPS, 2016.

PW Holland, KB Laskey, and S Leinhardt. Stochastic blockmodels: first steps. *Social Networks*, 1983.

DN Hoover. Relations on probability spaces and arrays of random variables. Preprint, Institute for Advanced Study, Princeton, NJ, 1979.

C Kemp, JB Tenenbaum, TL Griffiths, T Yamada, and N Ueda. Learning systems of concepts with an infinite relational model. *AAAI*, 2006.

JFC Kingman. The representation of partition structures. *Journal of the London Mathematical Society*, 1978.

J Lloyd, P Orbanz, Z Ghahramani, & DM Roy. Random function priors for exchangeable arrays with applications to graphs and relational data. *NIPS*, 2012.

P Orbanz and DM Roy. Bayesian models of graphs, arrays and other exchangeable random structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015.

K Palla, F Caron, and YW Teh. Bayesian nonparametrics for sparse dynamic networks. arXiv 1607.01624, 2016.

J Pitman. Exchangeable and partially exchangeable random partitions. *Probability Theory and Related Fields*, 1995.

V Veitch and DM Roy. The class of random graphs arising from exchangeable random measures. arXiv 1512.03099, 2015.

Z Xu, V Tresp, S Yu, K Yu, and H Kriegel. Fast inference in infinite hidden relational models. *Proceedings of Mining and Learning with Graphs*, 2007.