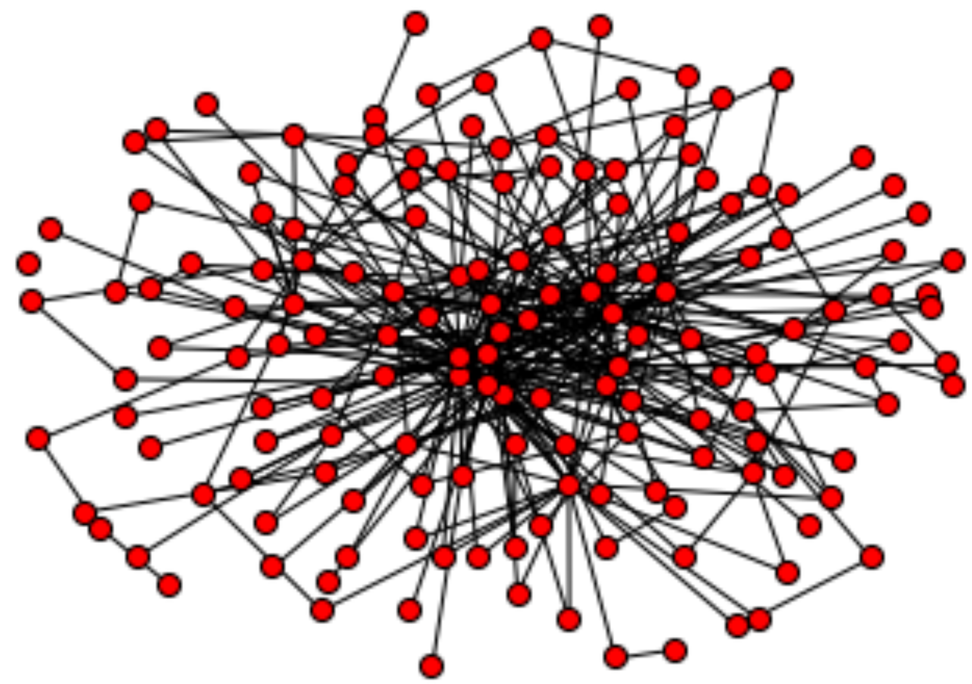
Edge-exchangeable graphs, sparsity, and power laws

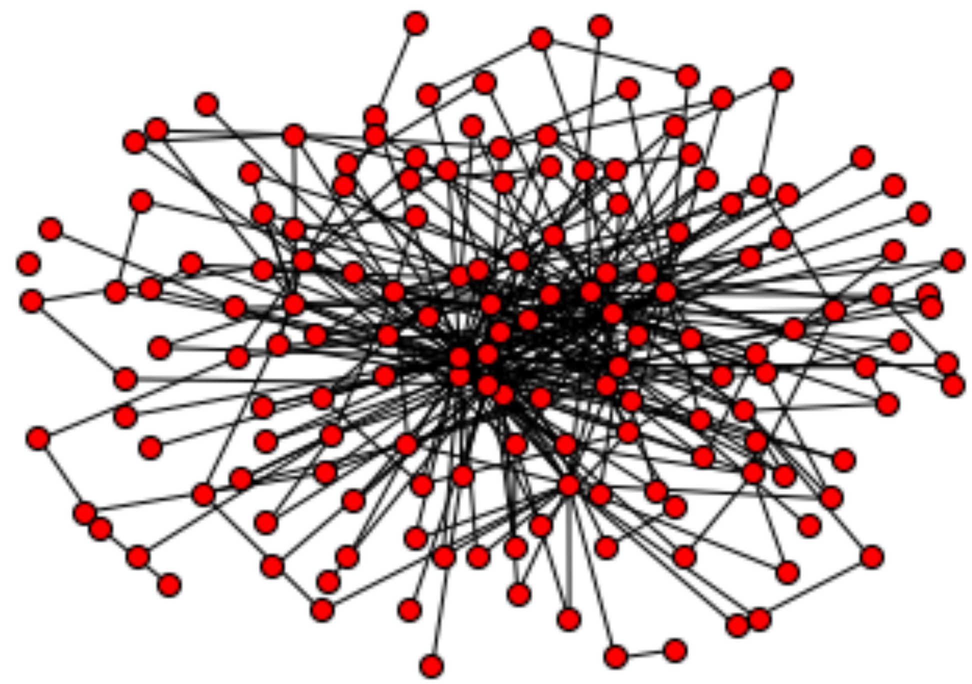
Tamara Broderick
ITT Career Development
Assistant Professor,
MIT

Joint work with: Diana Cai, Trevor Campbell

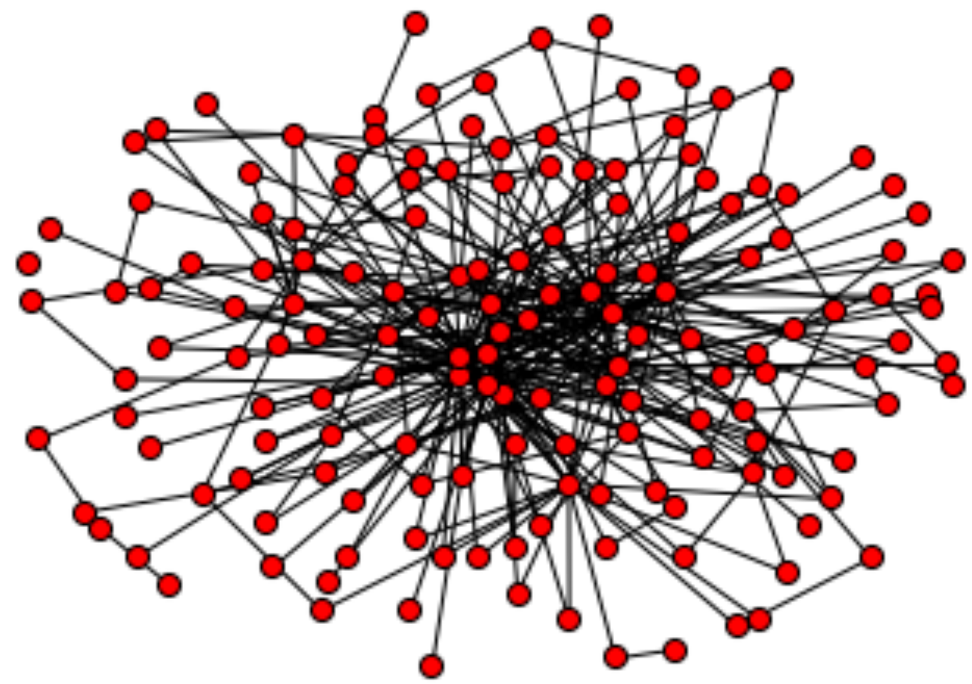
Probabilistic models for graphs



Probabilistic models for graphs

$p($  $)$

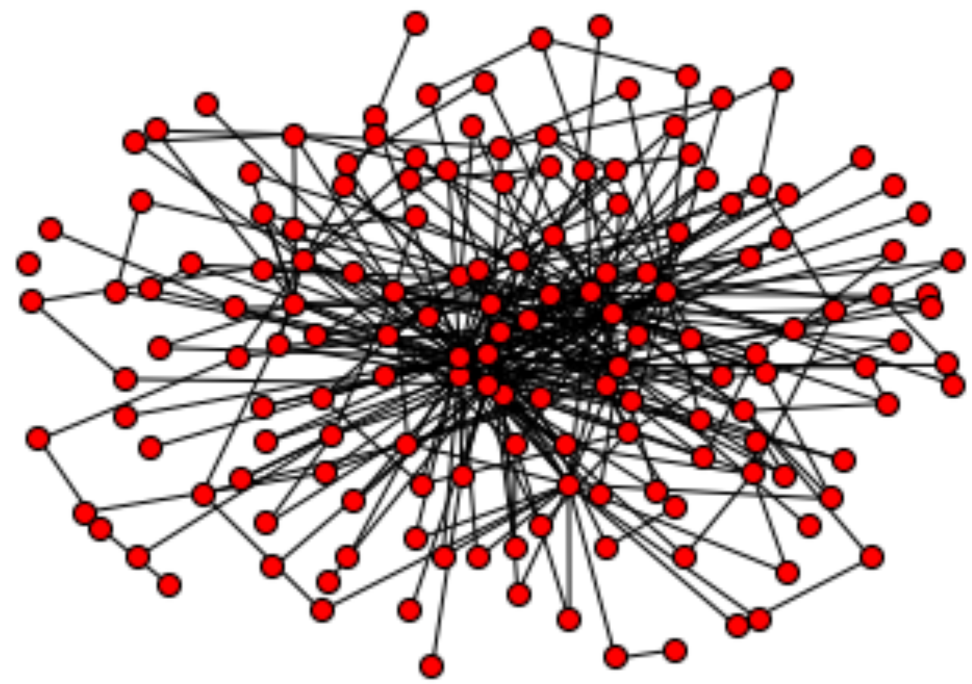
Probabilistic models for graphs

$p($  $)$

- Stochastic block model, mixed membership stochastic block model, infinite relational model, etc

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

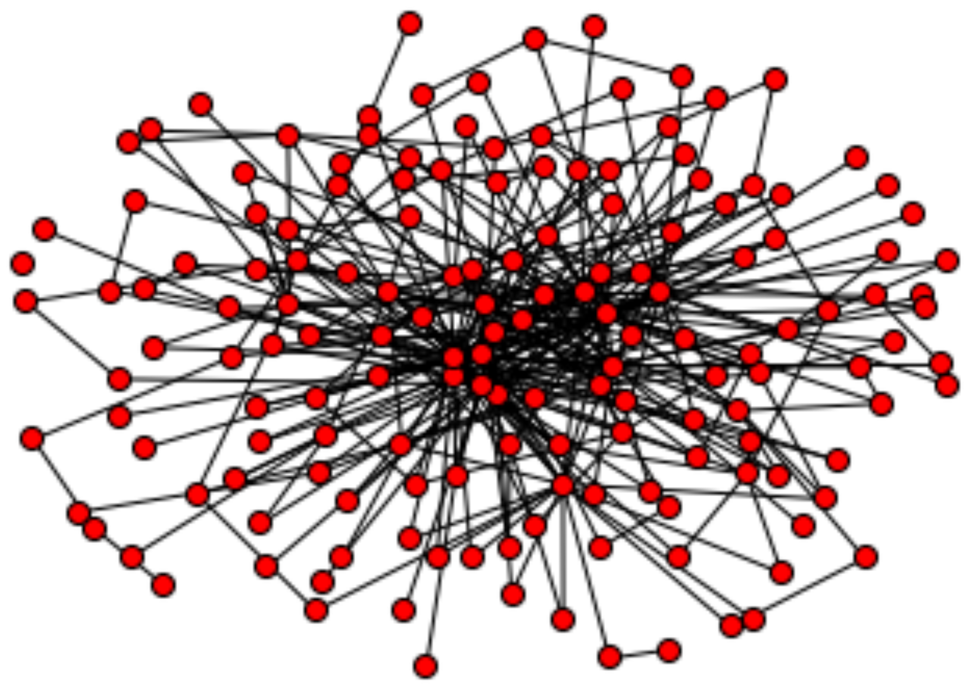
Probabilistic models for graphs

$p($  $)$

- Stochastic block model, mixed membership stochastic block model, infinite relational model, etc
- Learn hidden structure, uncertainties, modularity

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

Probabilistic models for graphs

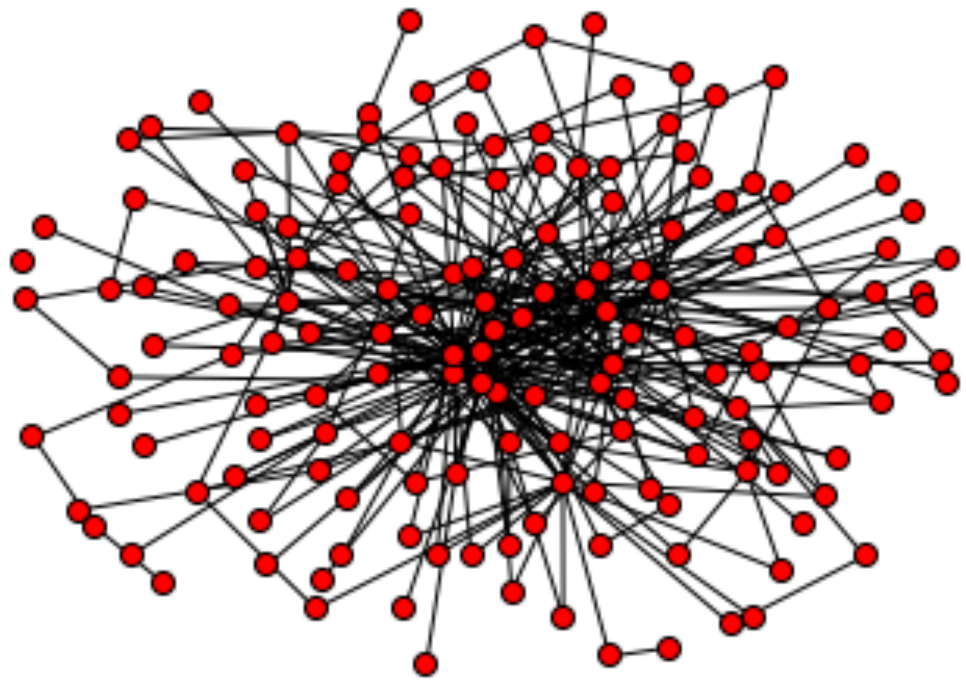
$p($  $)$

- Stochastic block model, mixed membership stochastic block model, infinite relational model, etc
- Learn hidden structure, uncertainties, modularity
- Problem: model misspecification, dense graphs

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

Probabilistic models for graphs

$p($



$)$

- Stochastic block model, mixed membership stochastic block model, infinite relational model, etc
- Learn hidden structure, uncertainties, modularity
- Problem: model misspecification, dense graphs
- Solution: a new set of models for sparse graphs

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

De Finetti

- A data sequence $X_1, X_2, \ldots$

[Hewitt, Savage 1955; Aldous 1983]

De Finetti

- A data sequence is *infinitely exchangeable* if the distribution of any *N* data points doesn't change when permuted: $p(X_1, \ldots, X_N) = p(X_{\sigma(1)}, \ldots, X_{\sigma(N)})$

[Hewitt, Savage 1955; Aldous 1983]

De Finetti

- A data sequence is *infinitely exchangeable* if the distribution of any *N* data points doesn't change when permuted: $p(X_1, \ldots, X_N) = p(X_{\sigma(1)}, \ldots, X_{\sigma(N)})$
- *De Finetti's Theorem* (roughly): A sequence $X_1, X_2, \ldots$ is infinitely exchangeable if and only if, for all *N* and some distribution *P*:

$$p(X_1, \ldots, X_N) = \int_\theta \prod_{n=1}^{N} p(X_n|\theta) P(d\theta)$$

[Hewitt, Savage 1955; Aldous 1983]

De Finetti

- A data sequence is *infinitely exchangeable* if the distribution of any *N* data points doesn't change when permuted: $p(X_1, \ldots, X_N) = p(X_{\sigma(1)}, \ldots, X_{\sigma(N)})$
- *De Finetti's Theorem* (roughly): A sequence $X_1, X_2, \ldots$ is infinitely exchangeable if and only if, for all *N* and some distribution *P*:

$$p(X_1, \ldots, X_N) = \int_\theta \prod_{n=1}^{N} p(X_n|\theta) P(d\theta)$$

- Motivates:

[Hewitt, Savage 1955; Aldous 1983]

De Finetti

- A data sequence is *infinitely exchangeable* if the distribution of any *N* data points doesn't change when permuted: $p(X_1, \ldots, X_N) = p(X_{\sigma(1)}, \ldots, X_{\sigma(N)})$
- *De Finetti's Theorem* (roughly): A sequence $X_1, X_2, \ldots$ is infinitely exchangeable if and only if, for all *N* and some distribution *P*:
$$p(X_1, \ldots, X_N) = \int_\theta \prod_{n=1}^{N} p(X_n|\theta) P(d\theta)$$
- Motivates:
  - Parameters and likelihoods

[Hewitt, Savage 1955; Aldous 1983]

De Finetti

- A data sequence is *infinitely exchangeable* if the distribution of any *N* data points doesn't change when permuted: $p(X_1, \ldots, X_N) = p(X_{\sigma(1)}, \ldots, X_{\sigma(N)})$
- *De Finetti's Theorem* (roughly): A sequence $X_1, X_2, \ldots$ is infinitely exchangeable if and only if, for all *N* and some distribution *P*:

$$p(X_1, \ldots, X_N) = \int_\theta \prod_{n=1}^{N} p(X_n|\theta) P(d\theta)$$

- Motivates:
  - Parameters and likelihoods
  - Priors

[Hewitt, Savage 1955; Aldous 1983]

De Finetti

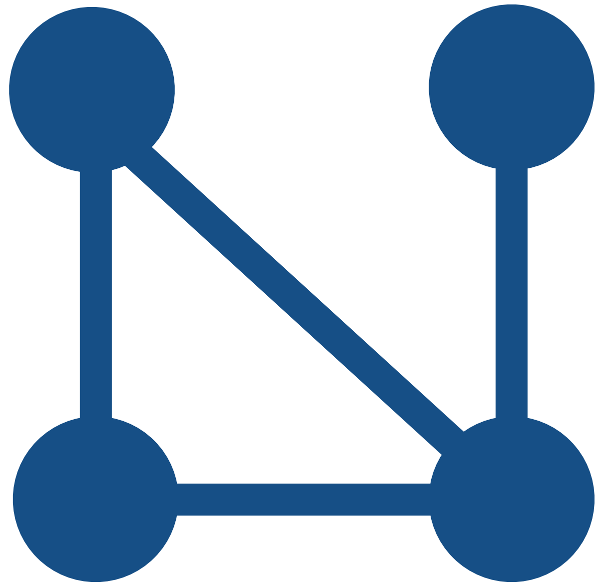
- A data sequence is *infinitely exchangeable* if the distribution of any *N* data points doesn't change when permuted: $p(X_1, \ldots, X_N) = p(X_{\sigma(1)}, \ldots, X_{\sigma(N)})$
- *De Finetti's Theorem* (roughly): A sequence $X_1, X_2, \ldots$ is infinitely exchangeable if and only if, for all *N* and some distribution *P*:

$$p(X_1, \ldots, X_N) = \int_\theta \prod_{n=1}^{N} p(X_n|\theta) P(d\theta)$$

- Motivates:
    - Parameters and likelihoods
    - Priors
    - *Nonparametric Bayesian* priors

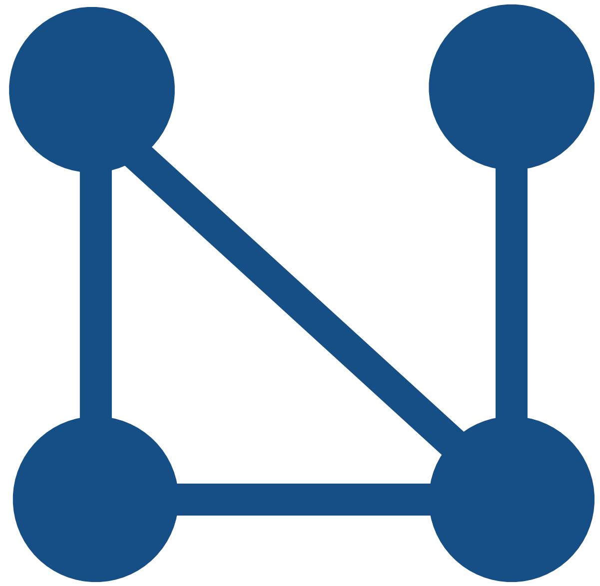
[Hewitt, Savage 1955; Aldous 1983]

Sequence of graphs



G

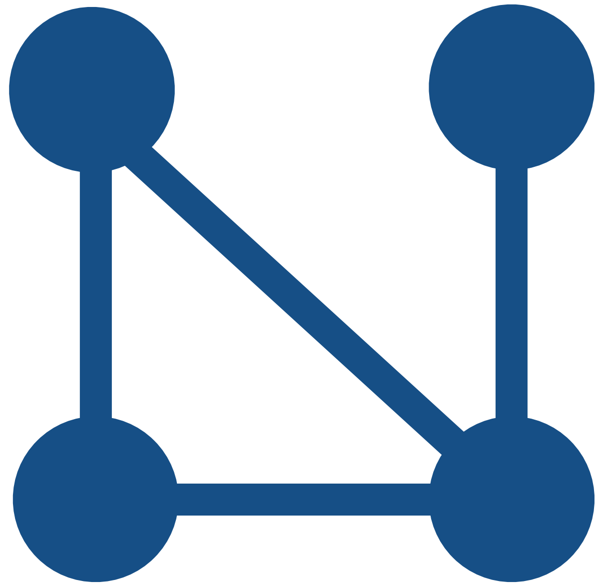
Sequence of graphs



G

- Nodes/vertices

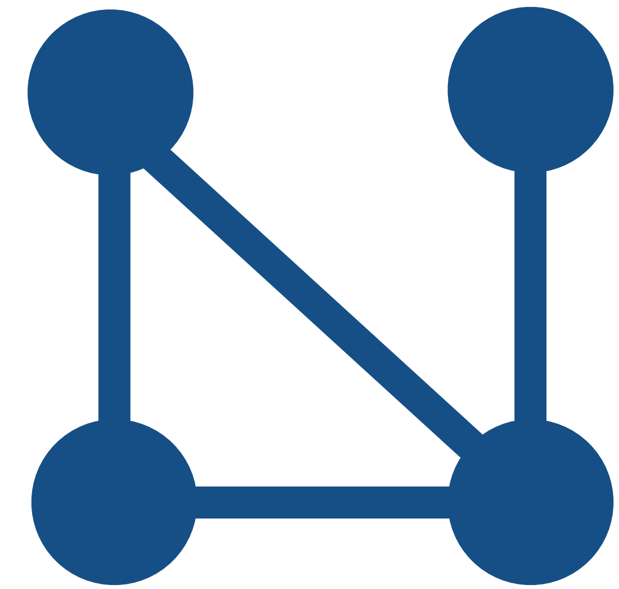
Sequence of graphs



G

- Nodes/vertices
- Edges

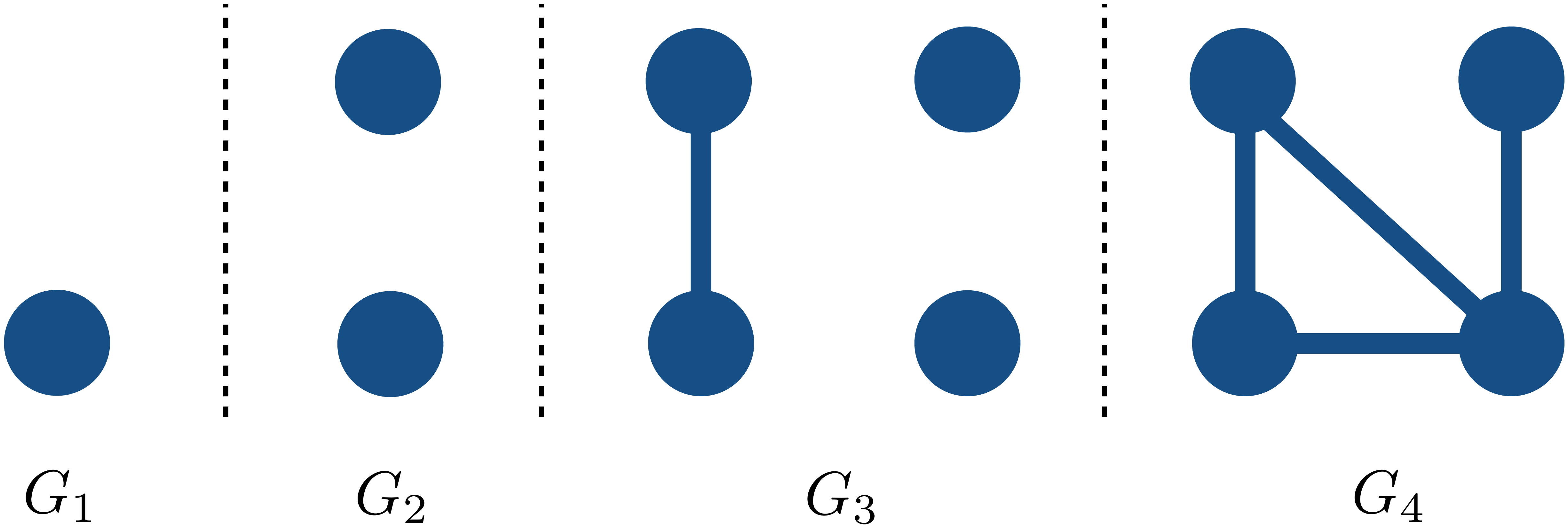
Sequence of graphs



G

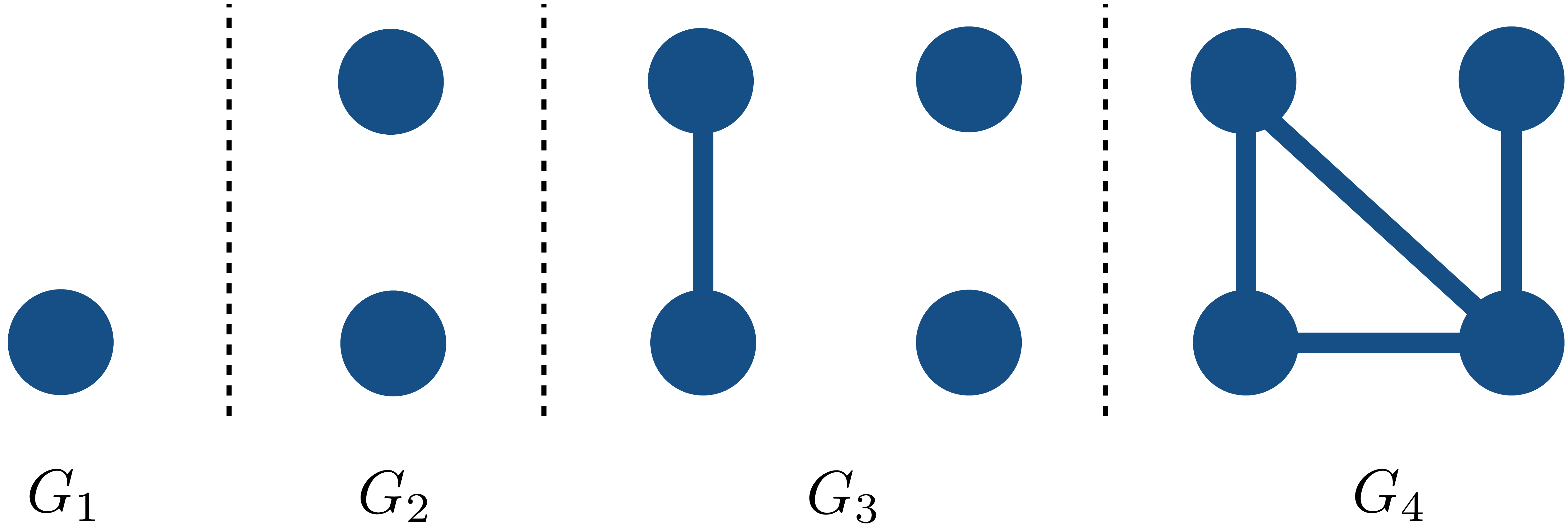
- Nodes/vertices
- Edges
- Active nodes

Sequence of graphs



- Nodes/vertices
- Edges
- Active nodes

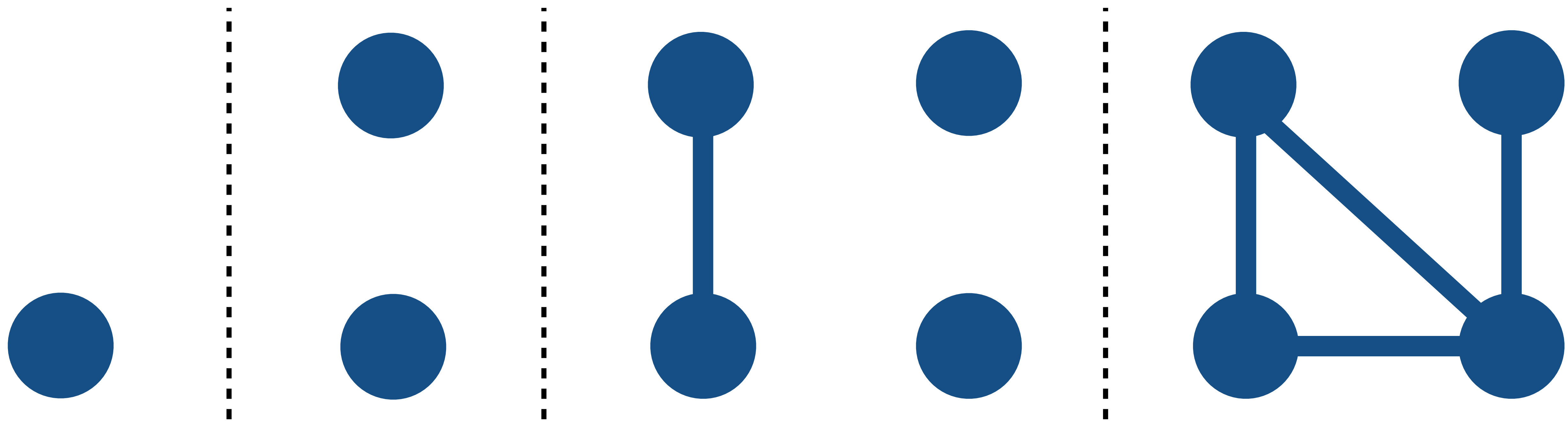
Sequence of graphs



- Nodes/vertices
- Edges
- Active nodes

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

Sequence of graphs

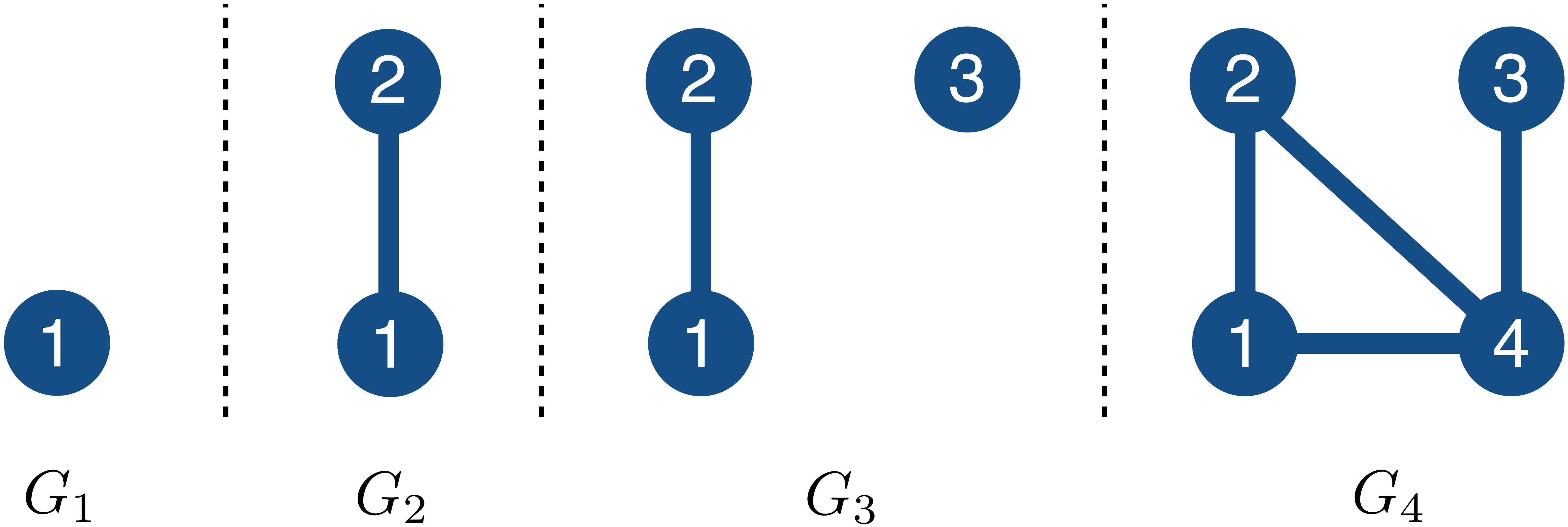


G_1 G_2 G_3 G_4

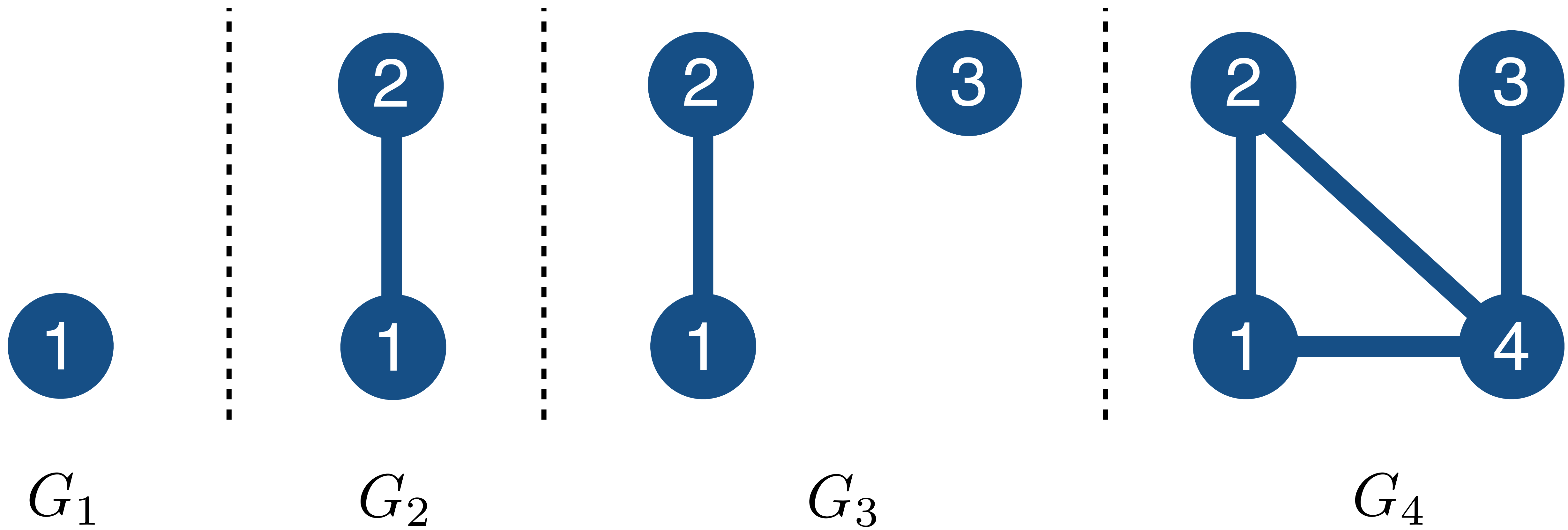
- Nodes/vertices
- Edges
- Active nodes

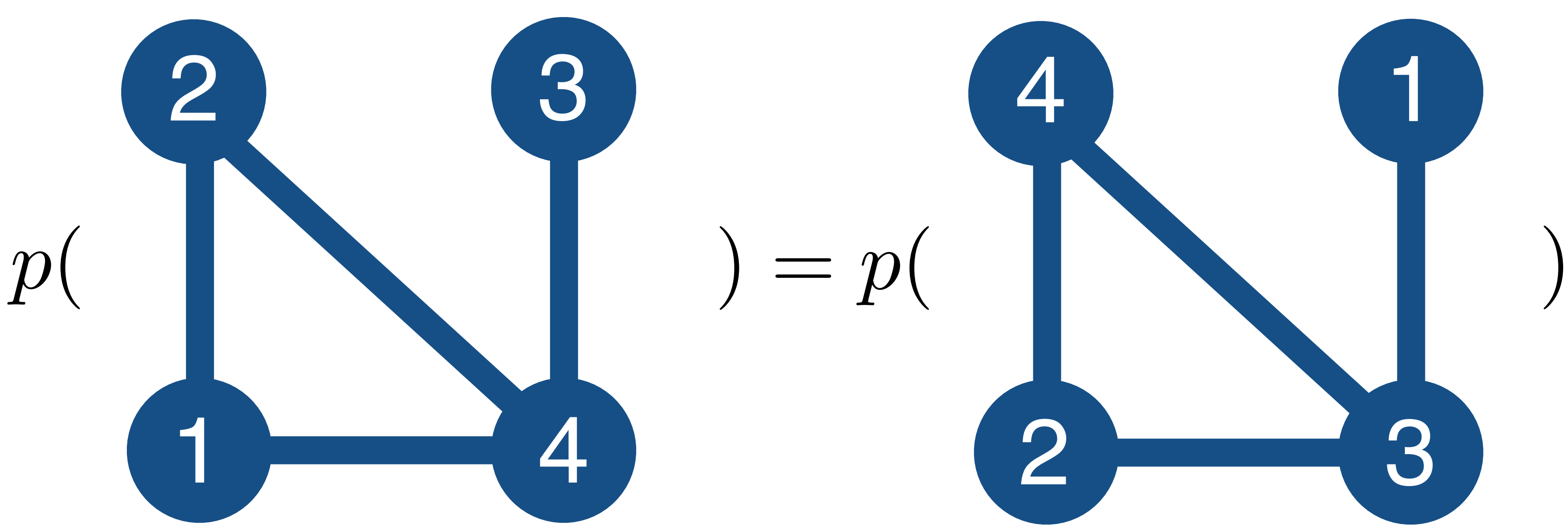
- *Dense* graph sequence $\#\text{edges}(G_n) \geq c \cdot [\#\text{nodes}(G_n)]^2$
- *Sparse* graph sequence $\#\text{edges}(G_n) \in o([\#\text{nodes}(G_n)]^2)$

(Node) Exchangeability

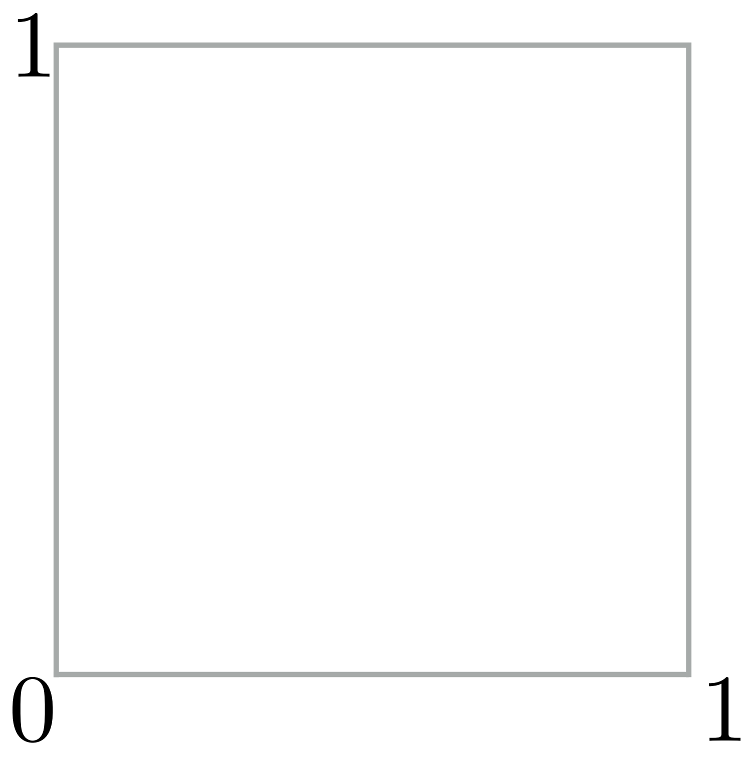


(Node) Exchangeability



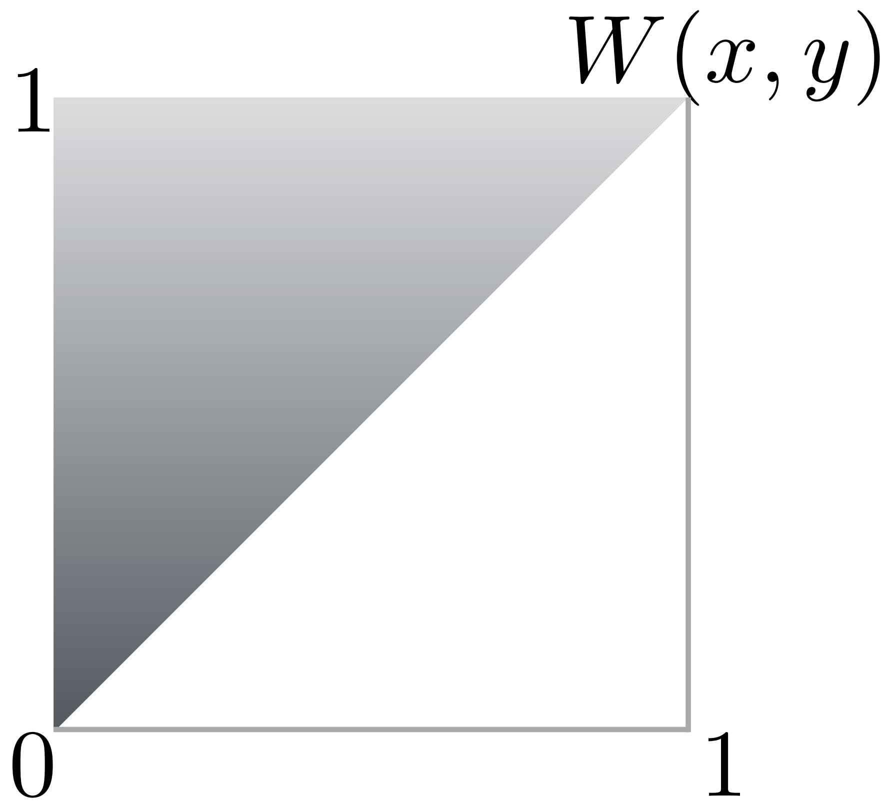


Aldous-Hoover



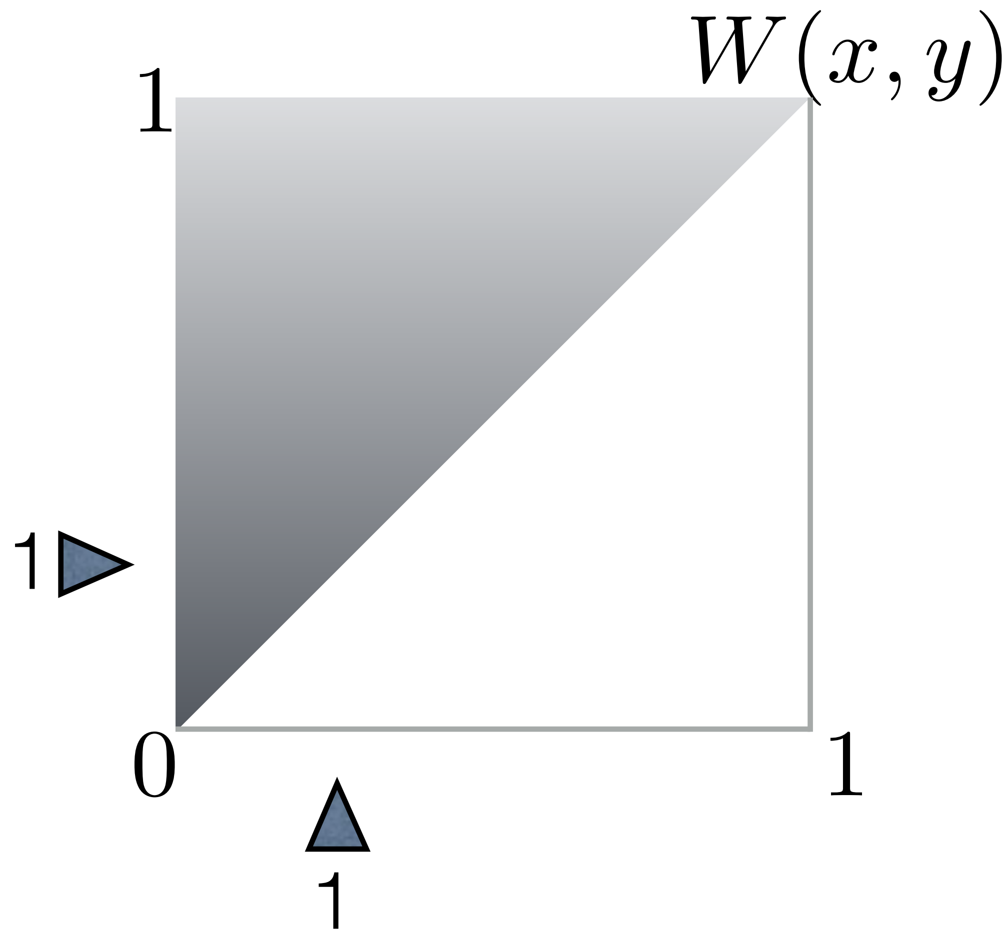
[Hoover 1979, Aldous 1981, Orbanz, Roy 2015]

Aldous-Hoover



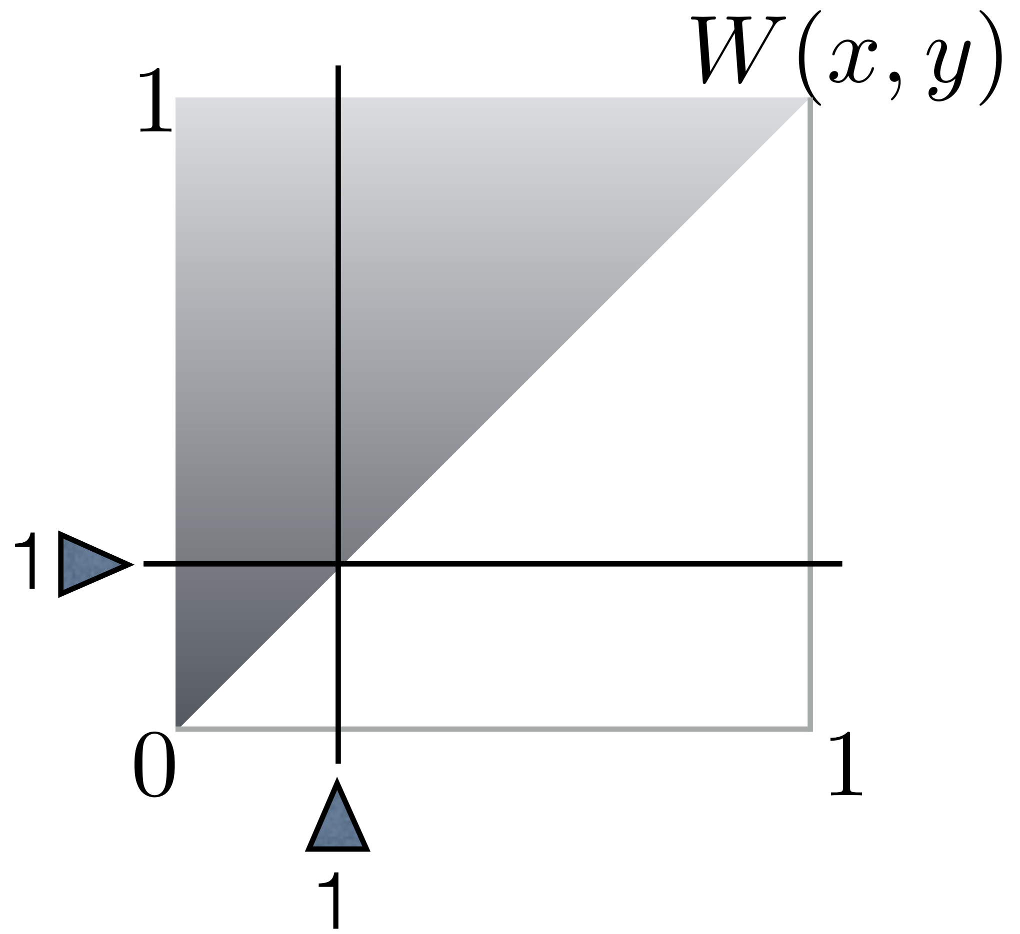
[Hoover 1979, Aldous 1981, Orbanz, Roy 2015]

Aldous-Hoover



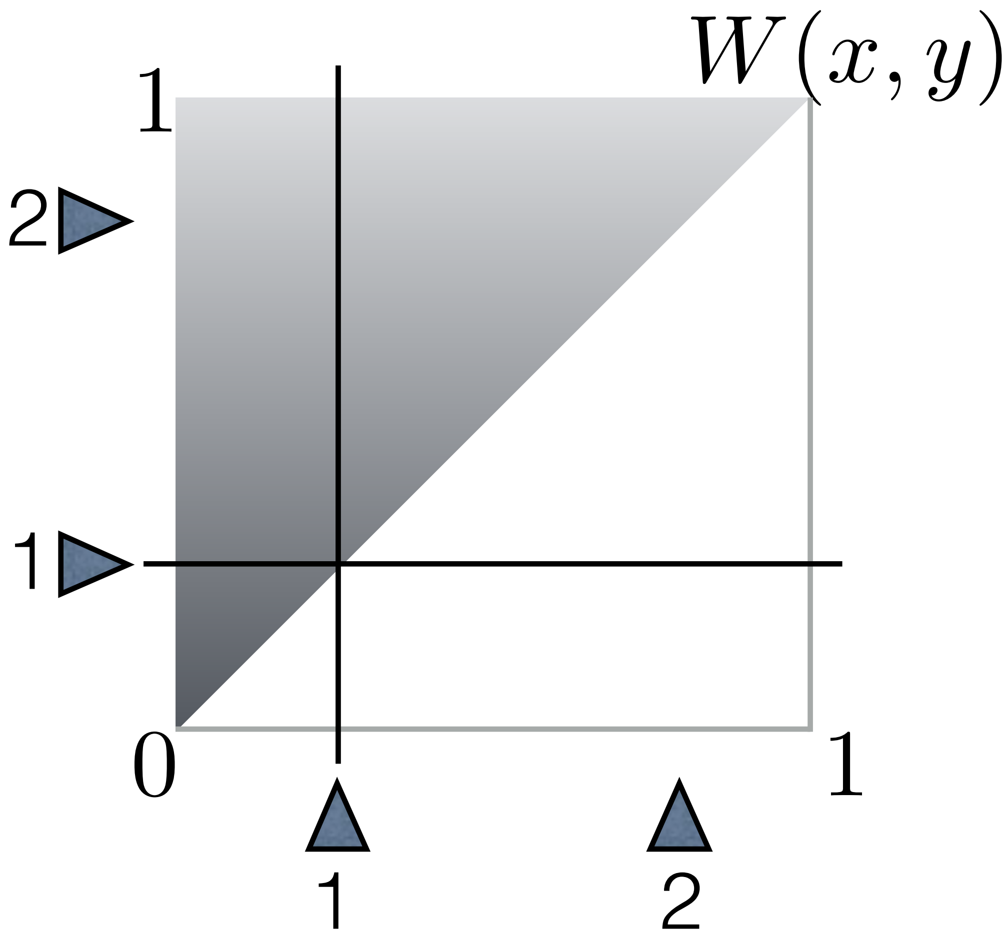
[Hoover 1979, Aldous 1981, Orbanz, Roy 2015]

Aldous-Hoover



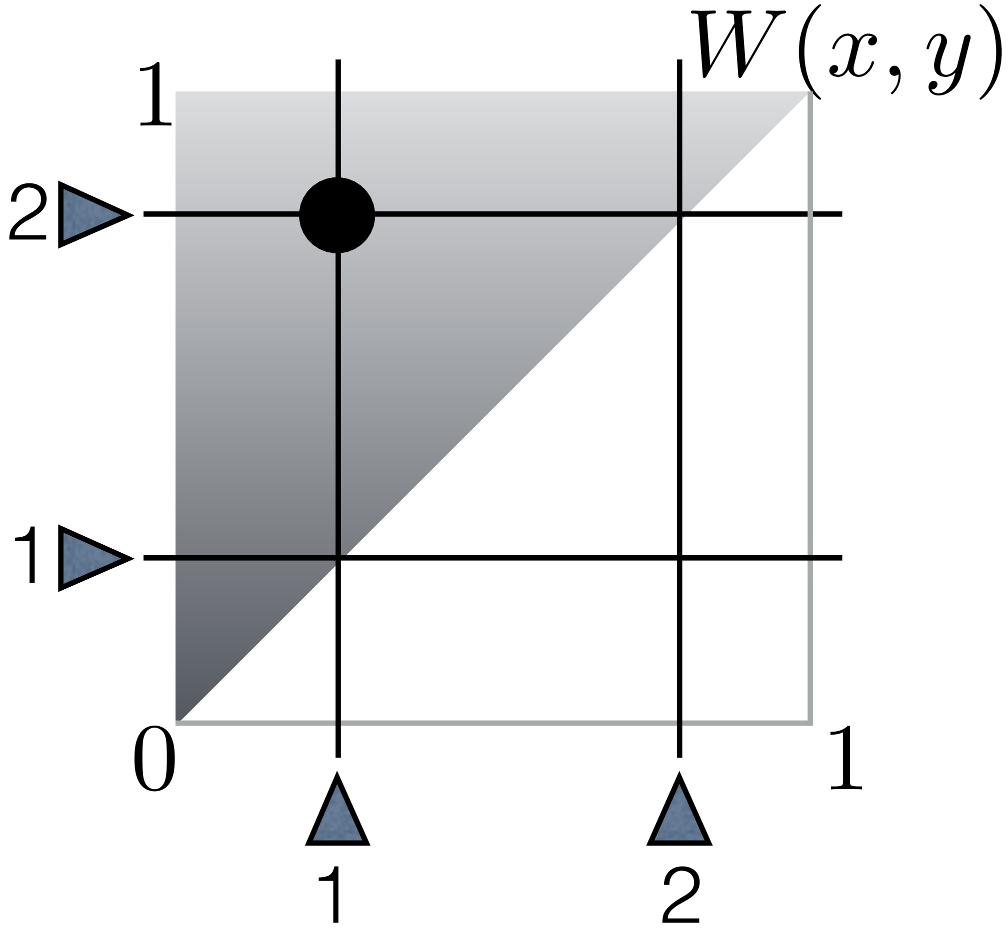
[Hoover 1979, Aldous 1981, Orbanz, Roy 2015]

Aldous-Hoover



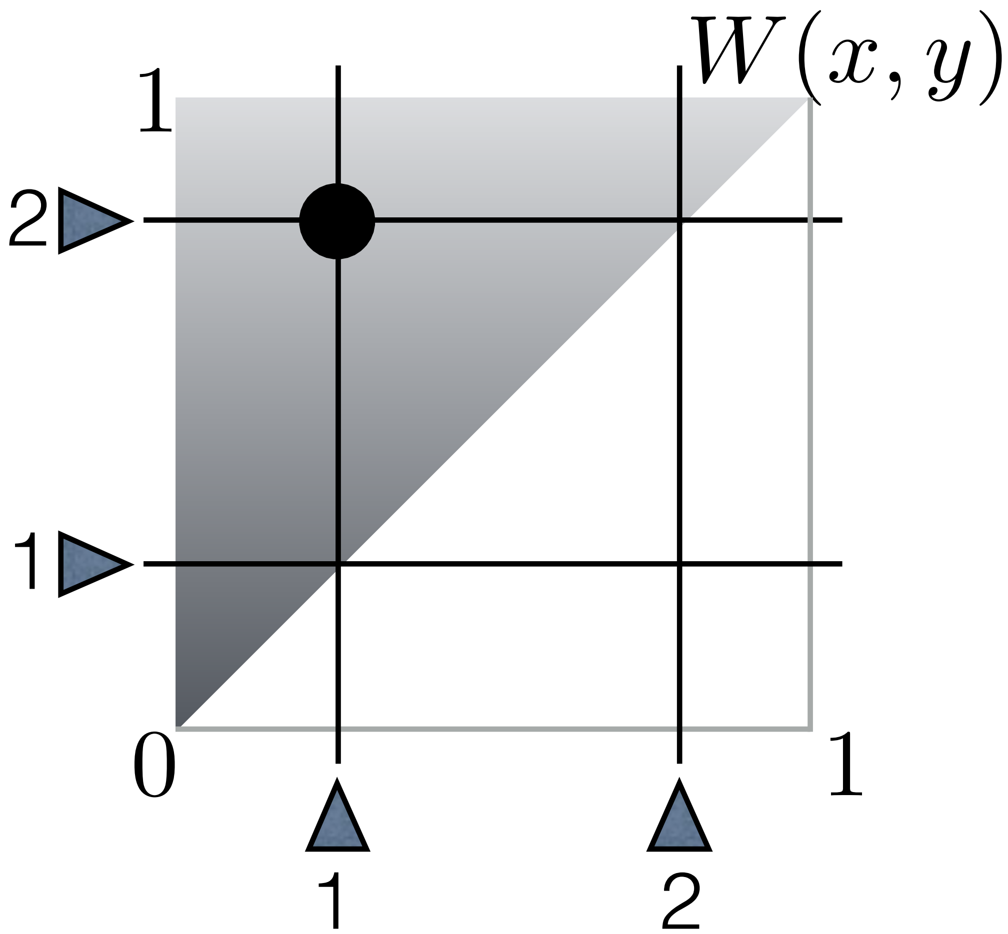
[Hoover 1979, Aldous 1981, Orbanz, Roy 2015]

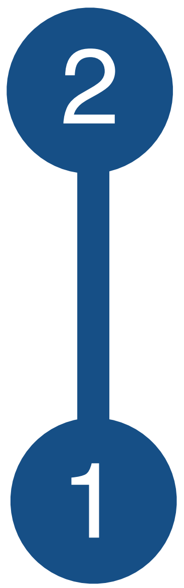
Aldous-Hoover



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

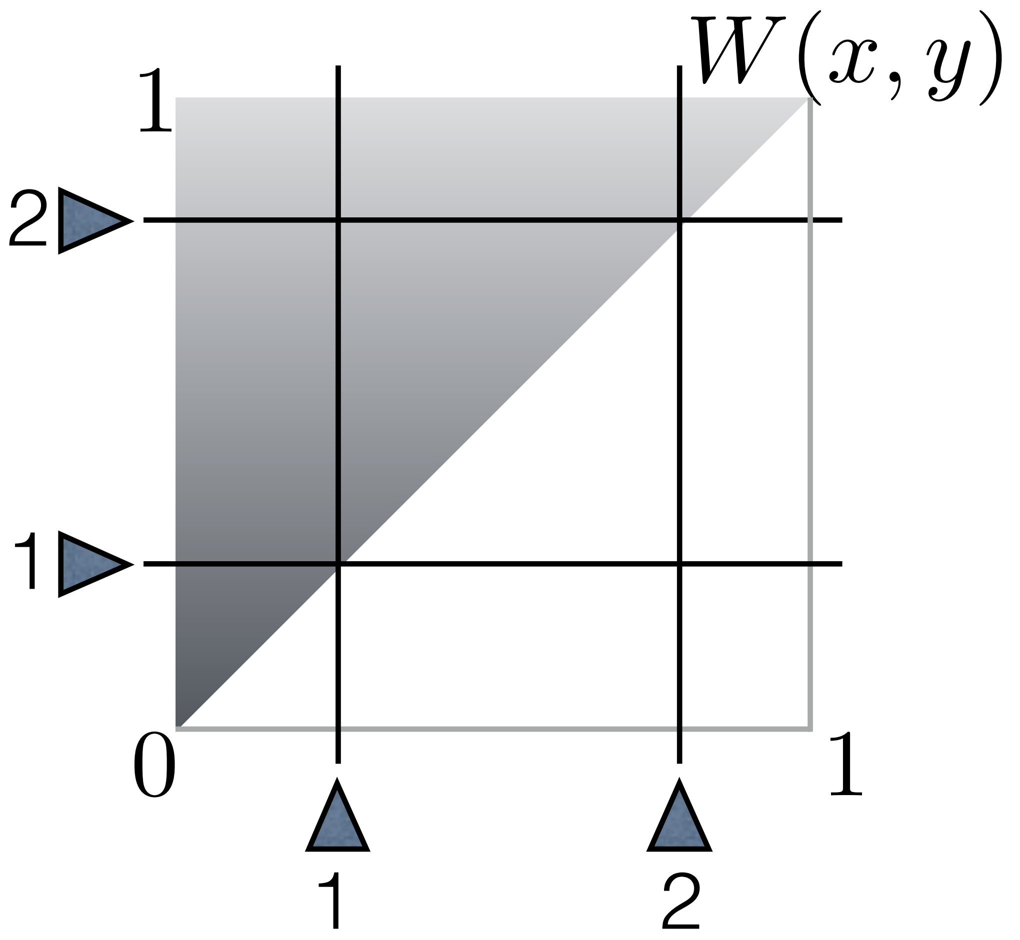
Aldous-Hoover

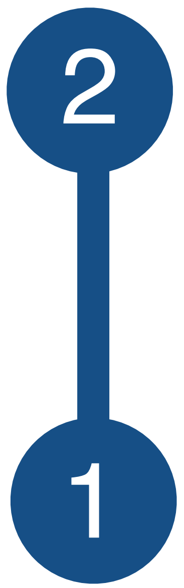




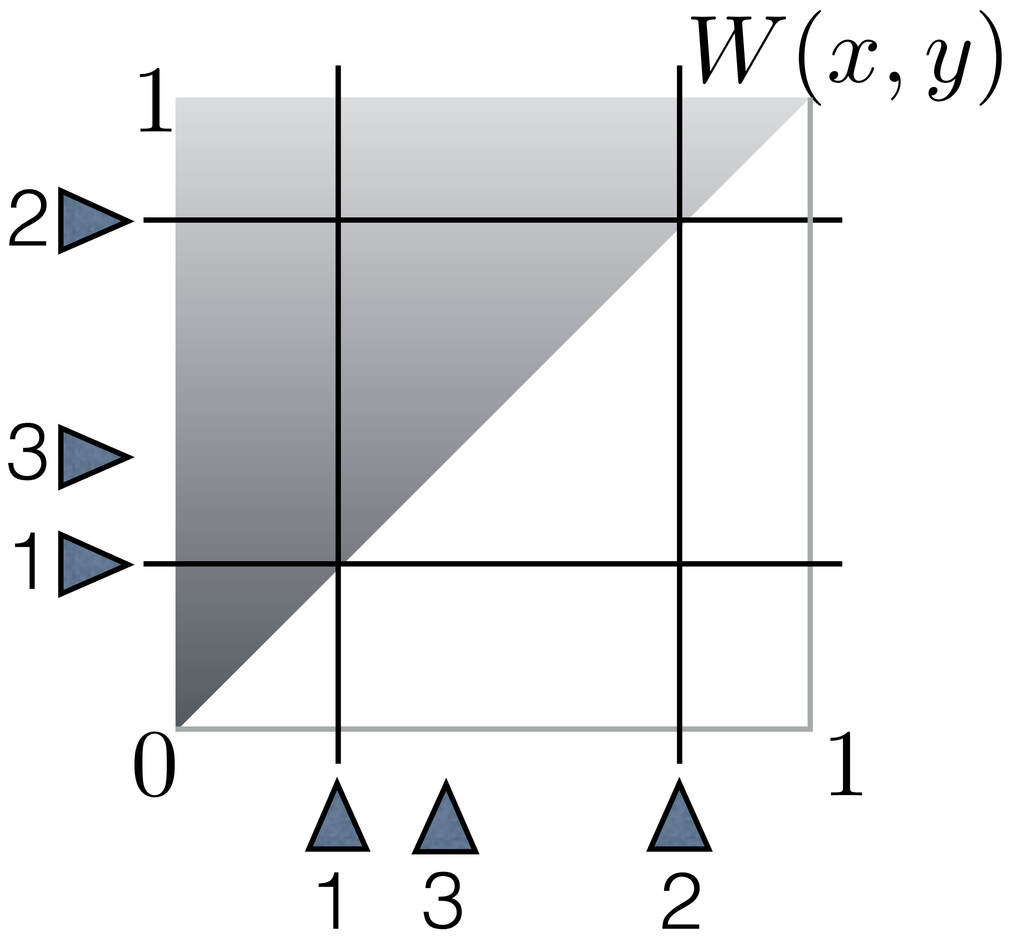
[Hoover 1979, Aldous 1981, Orbanz, Roy 2015]

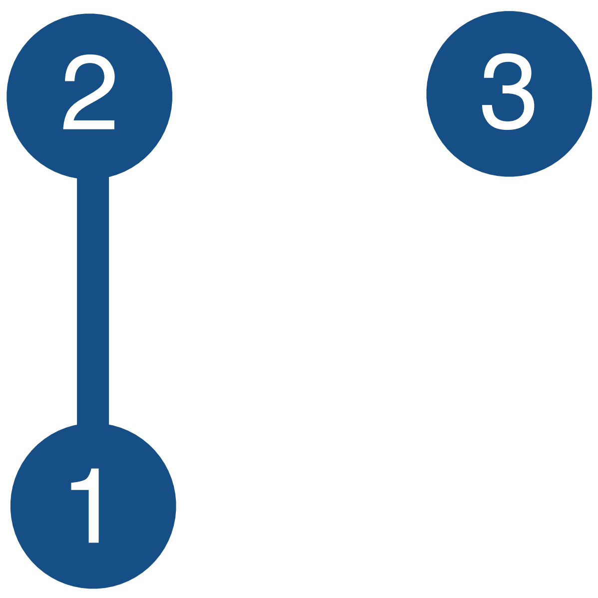
Aldous-Hoover





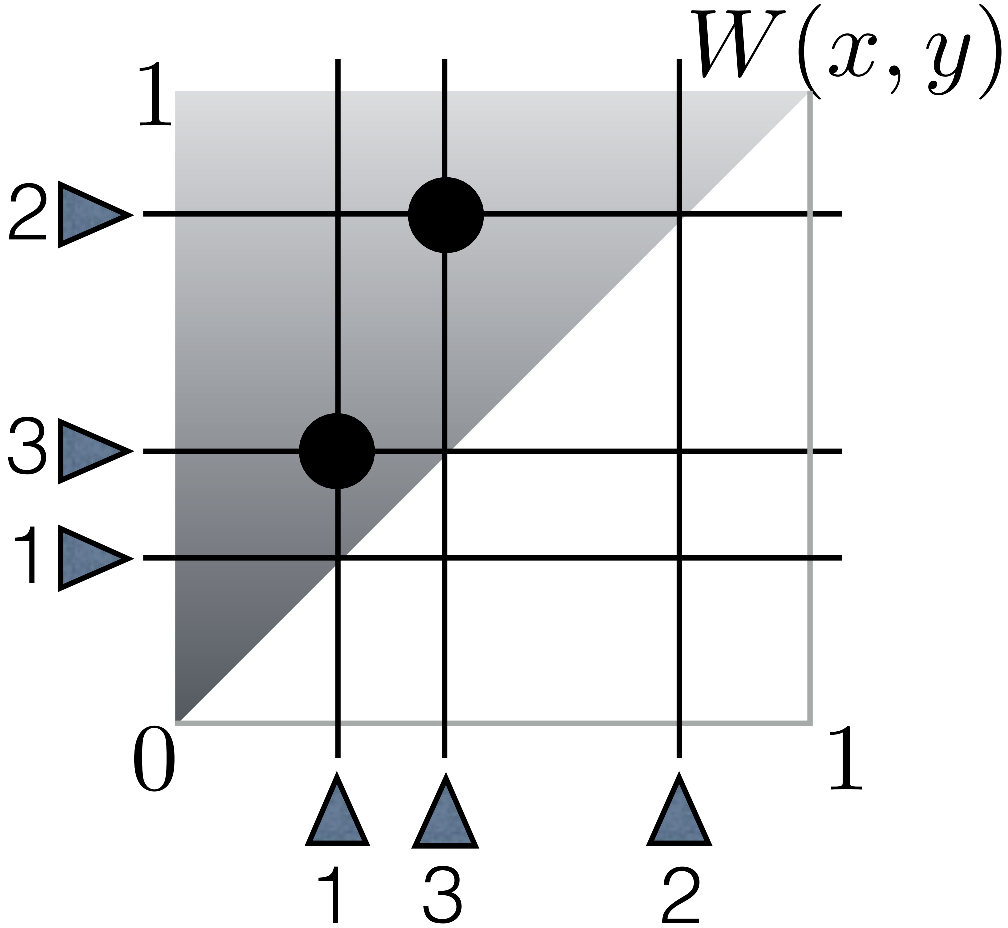
Aldous-Hoover

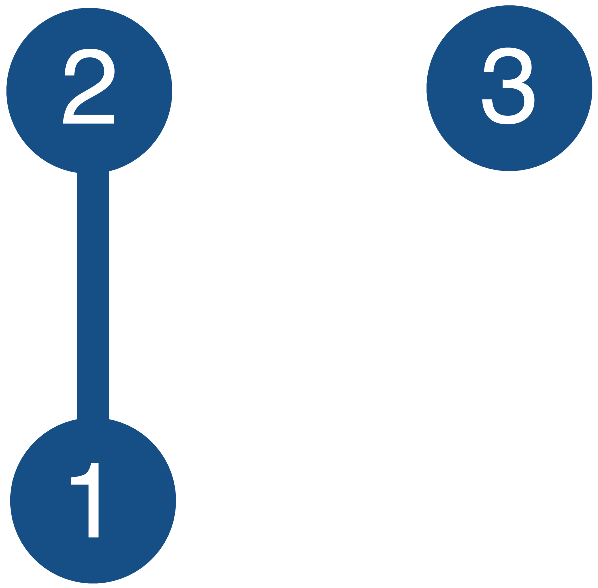




[Hoover 1979, Aldous 1981, Orbanz, Roy 2015]

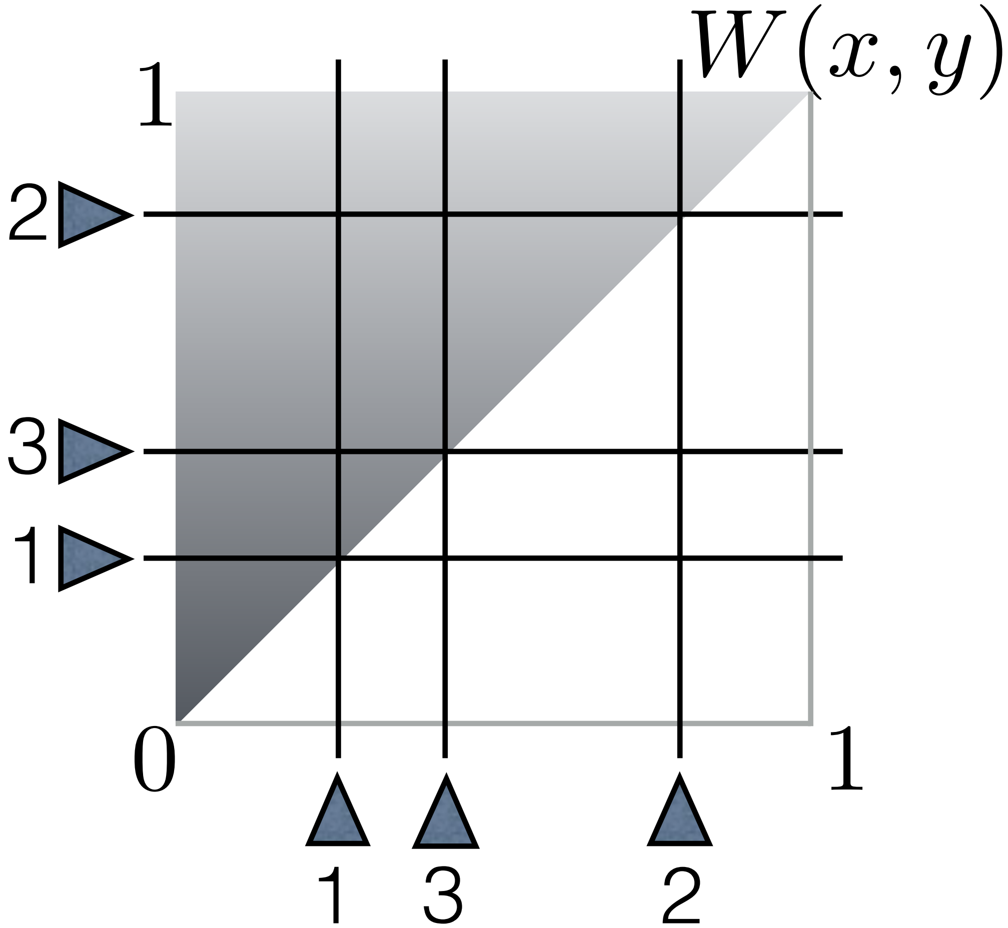
Aldous-Hoover

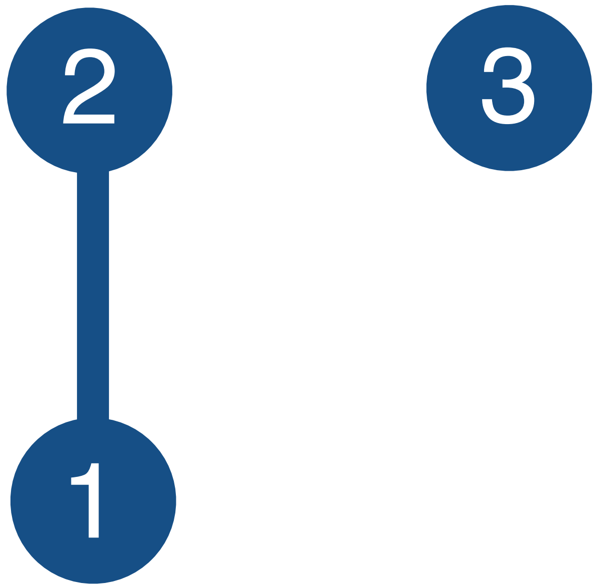




[Hoover 1979, Aldous 1981, Orbanz, Roy 2015]

Aldous-Hoover

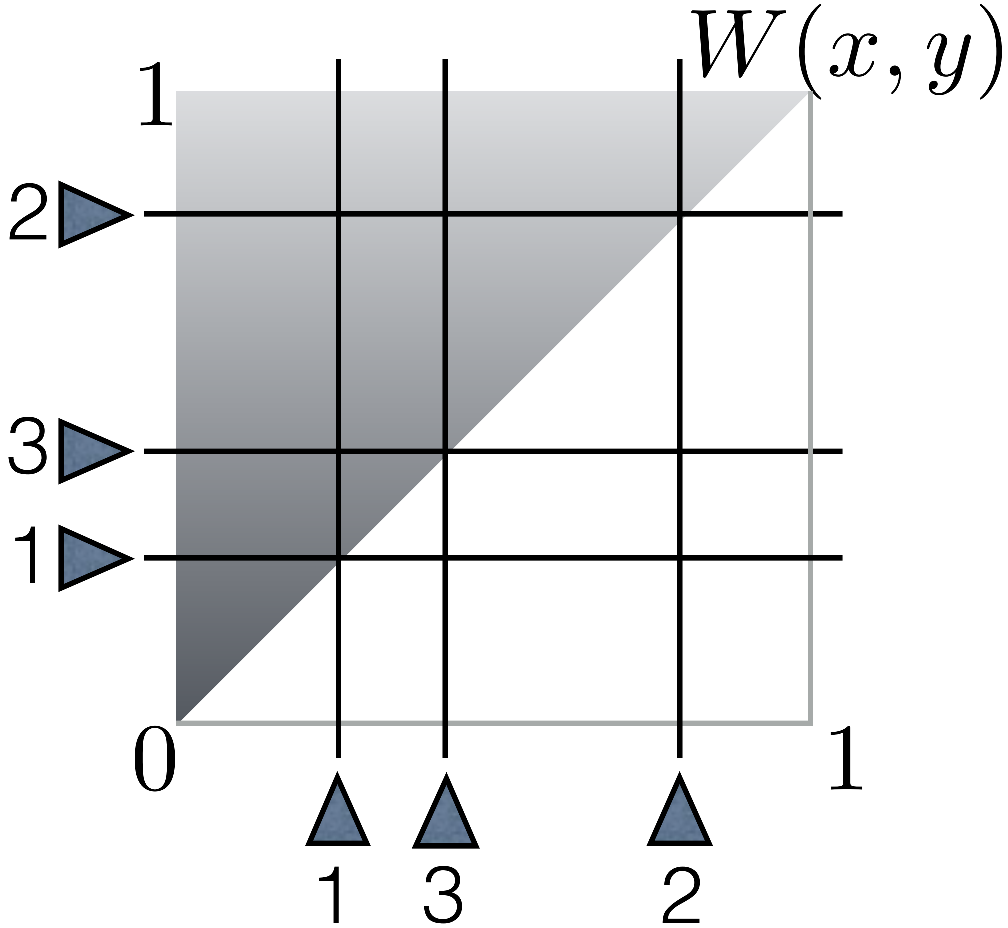


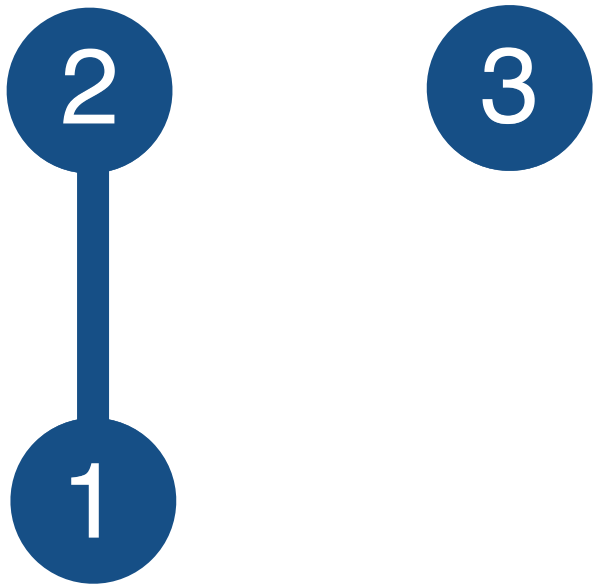


$\mathbb{E}[\#\text{edges}(G_n)]$

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

Aldous-Hoover

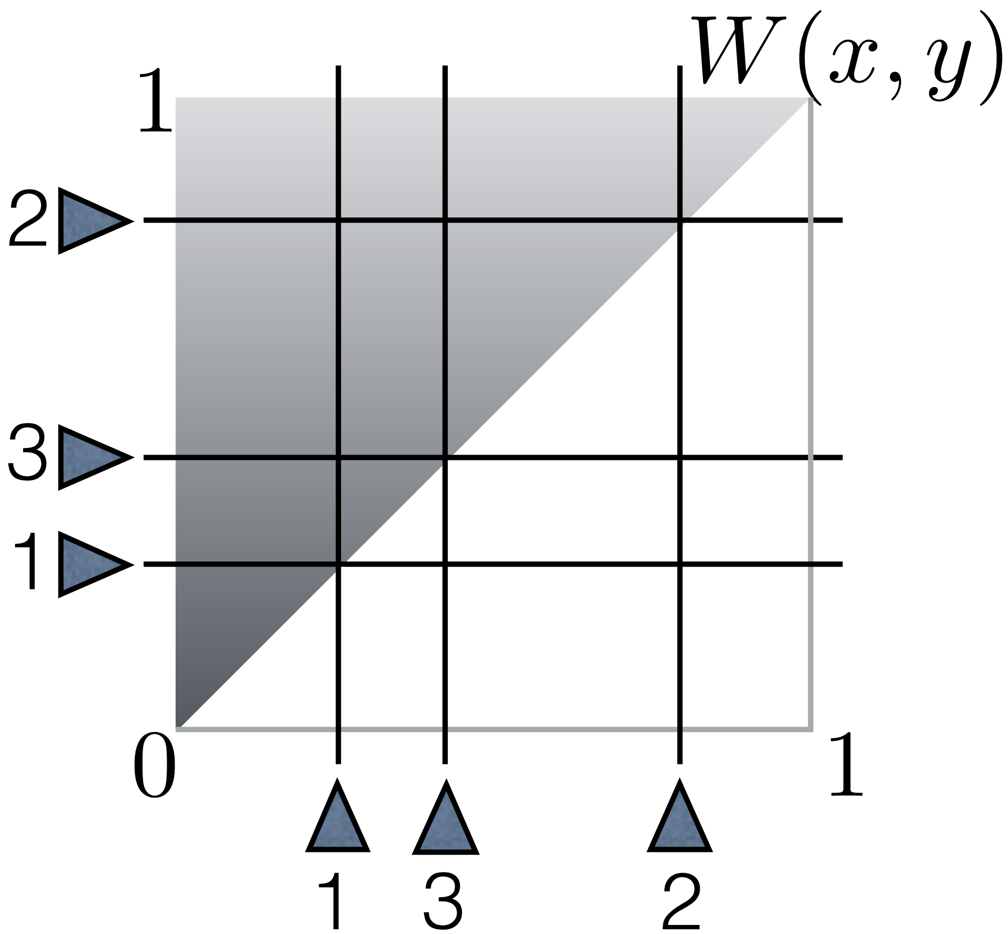


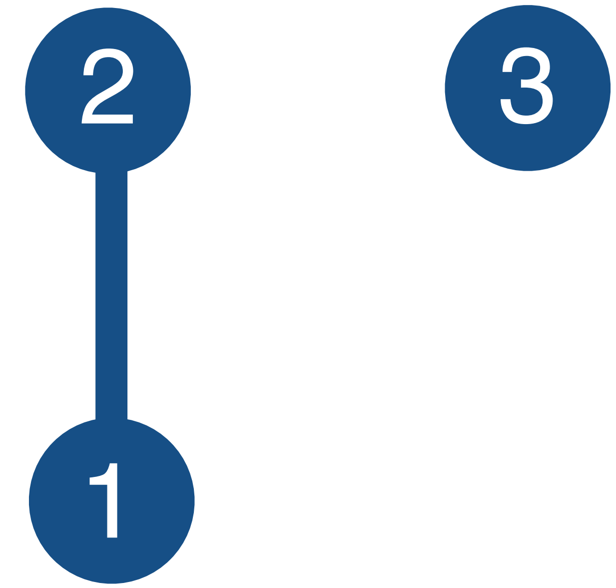


$$\mathbb{E}[\#\text{edges}(G_n)] = \binom{n}{2} \frac{1}{2} \int_0^1 \int_0^1 W(x,y)dx\, dy$$

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

Aldous-Hoover



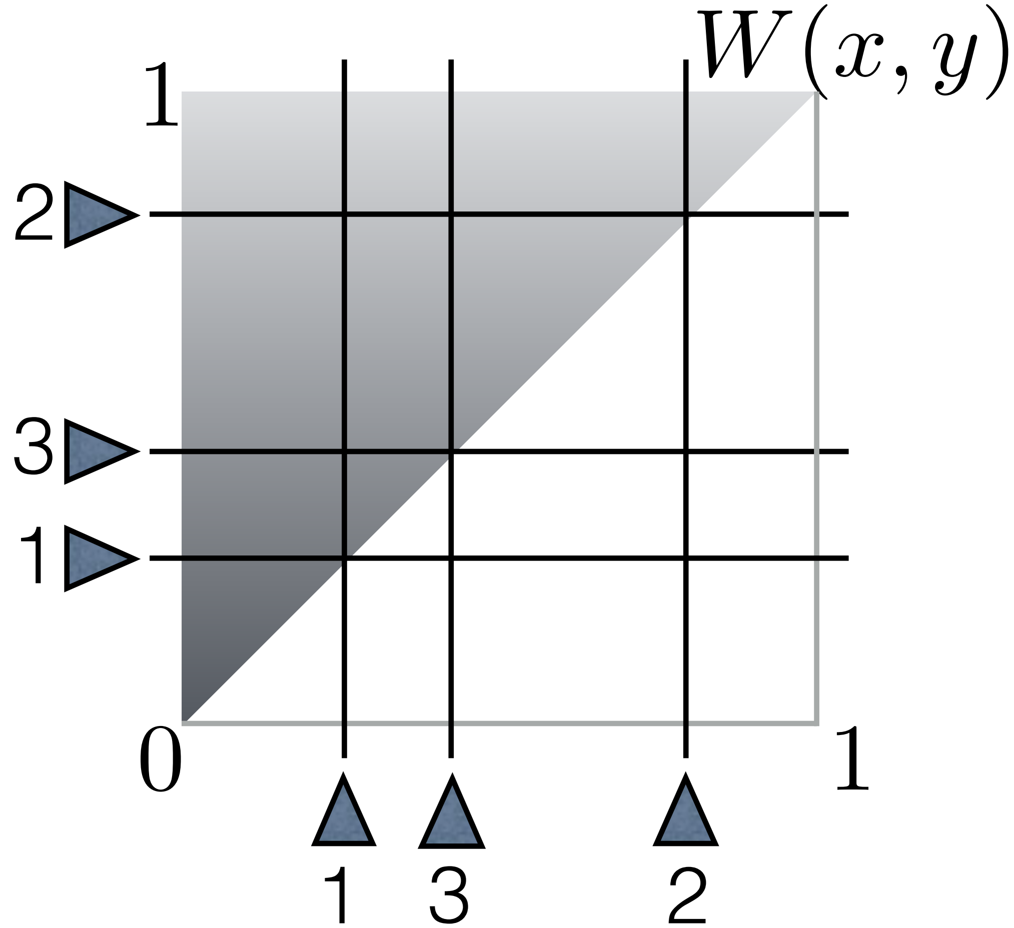


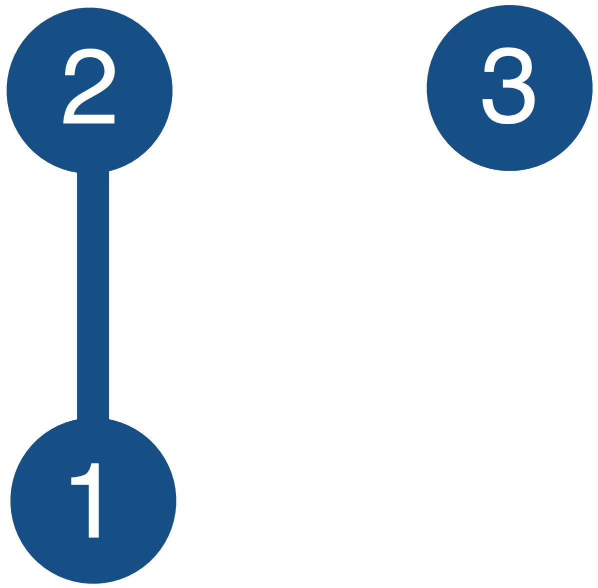
$$\mathbb{E}[\#\text{edges}(G_n)] = \binom{n}{2}\frac{1}{2}\int_0^1\int_0^1 W(x,y)dx\,dy$$

$$= cn^2 = c \cdot \#\text{vertices}(G_n)^2$$

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

Aldous-Hoover





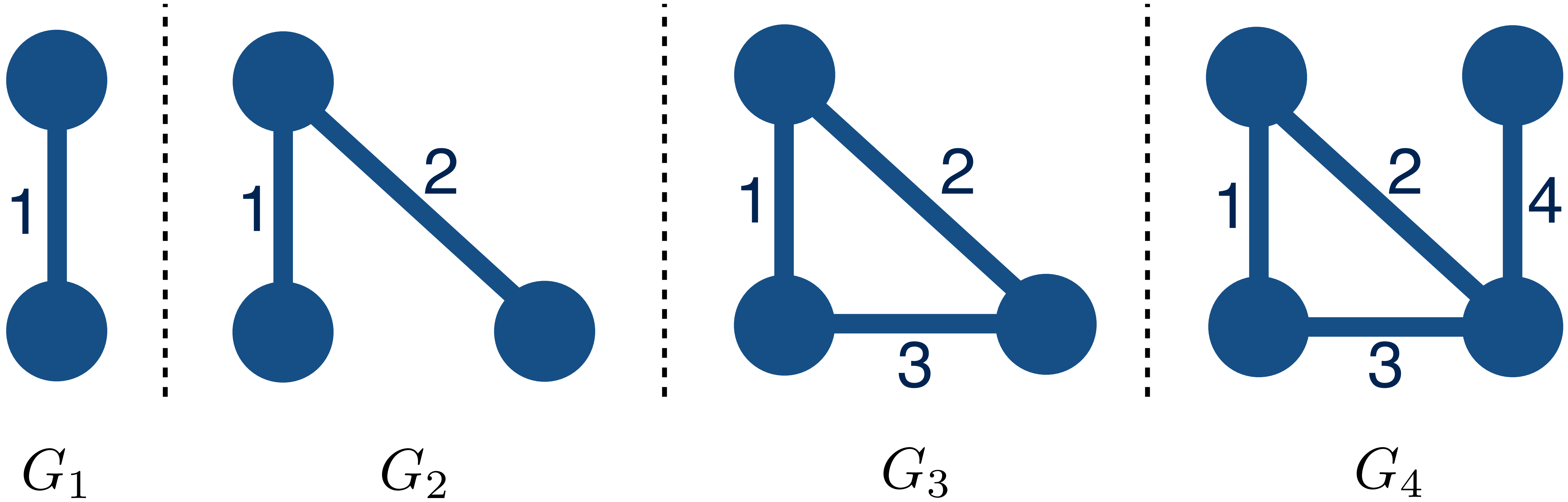
$$\mathbb{E}[\#\text{edges}(G_n)] = \binom{n}{2} \frac{1}{2} \int_0^1 \int_0^1 W(x,y) dx\, dy$$

$$= cn^2 = c \cdot \#\text{vertices}(G_n)^2$$

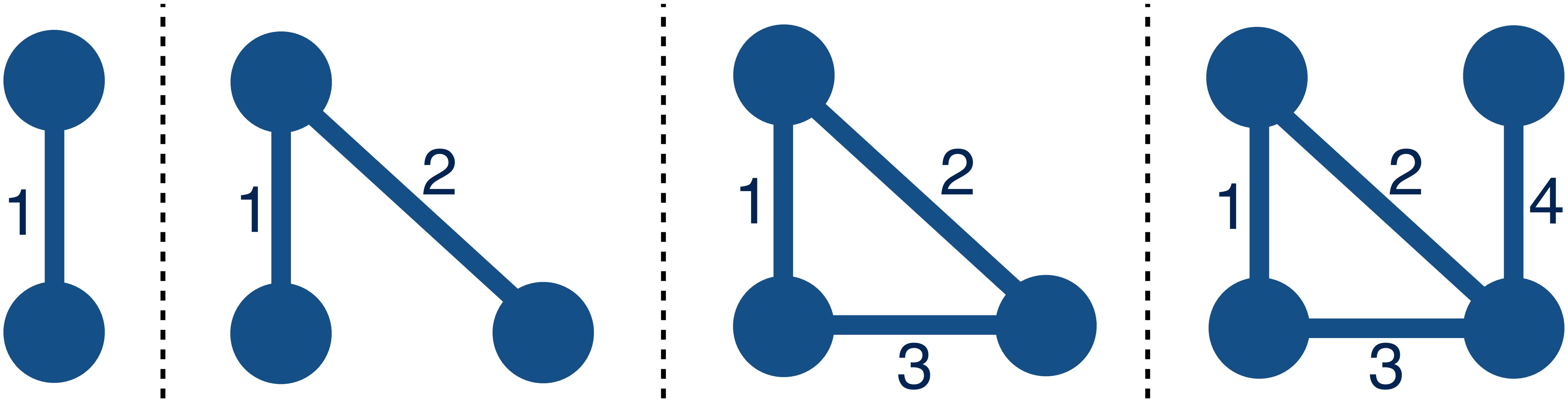
(Might be empty)

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

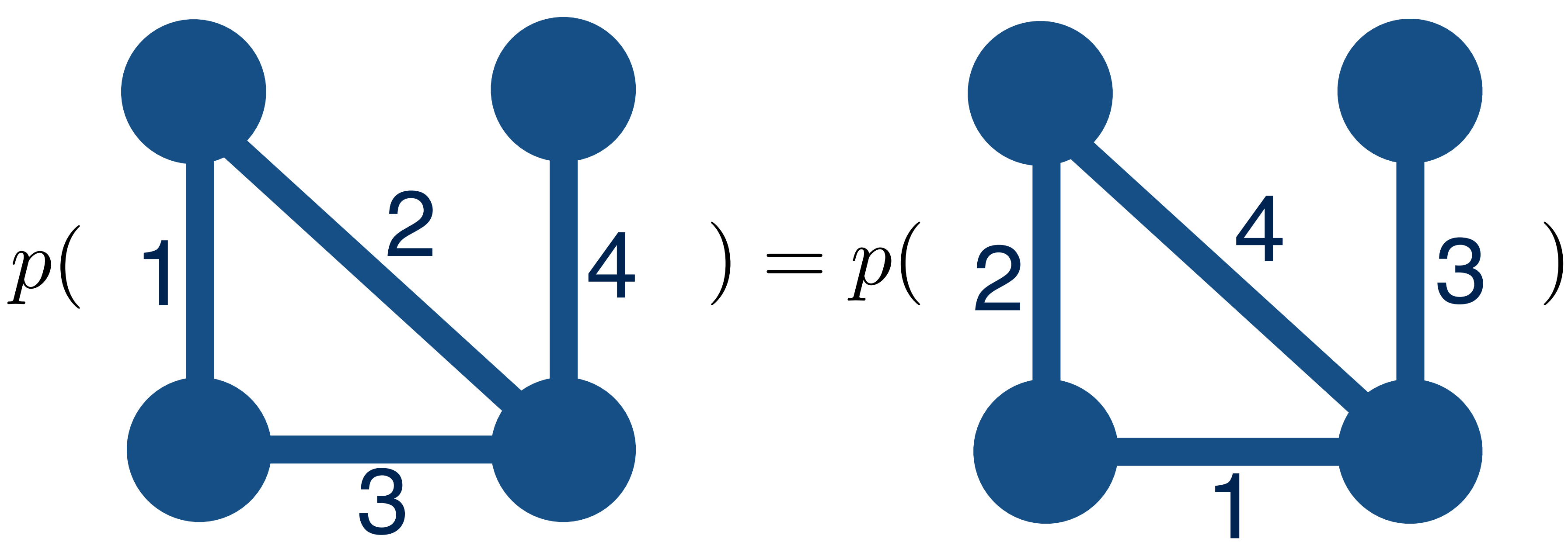
Sequence of graphs



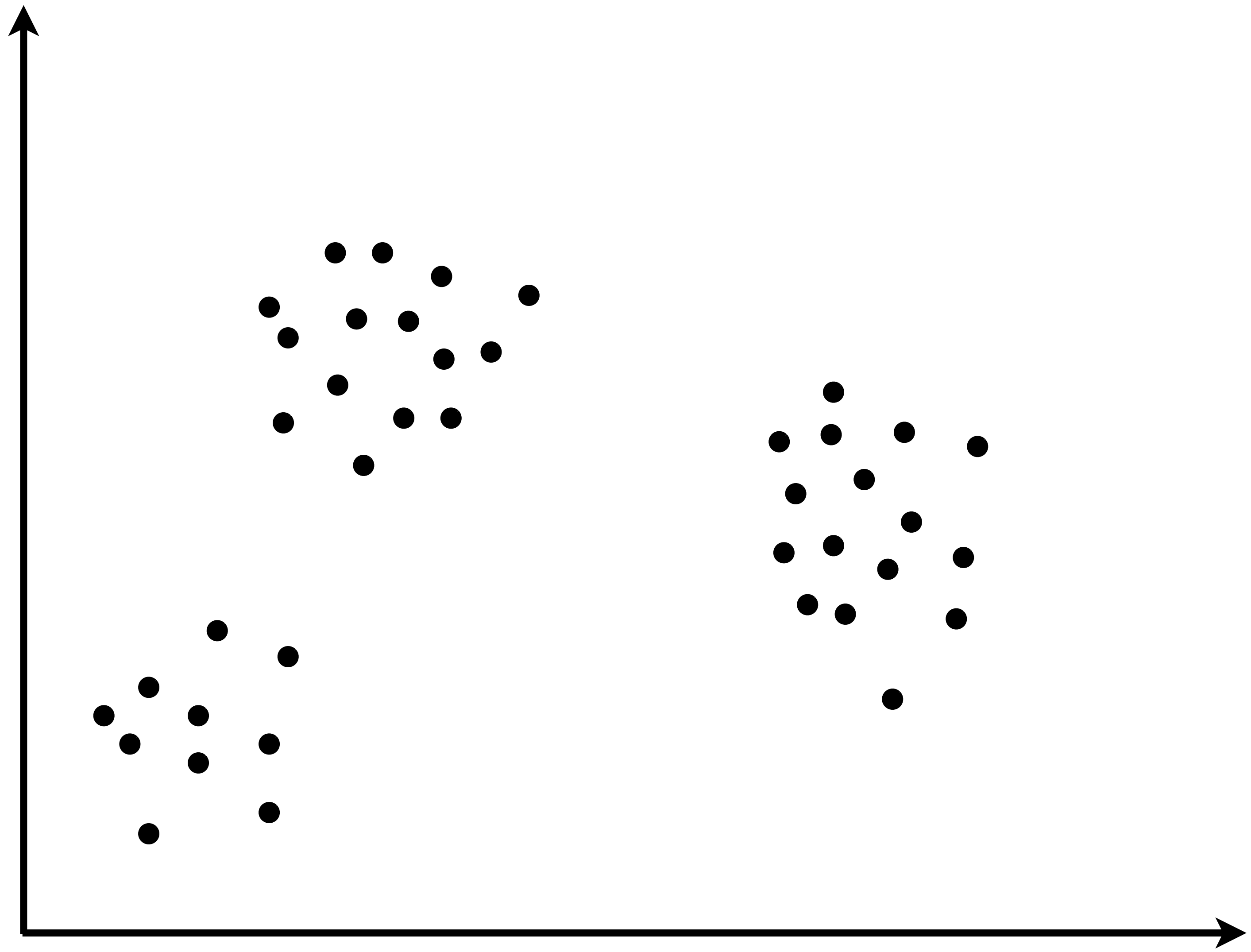
Sequence of graphs



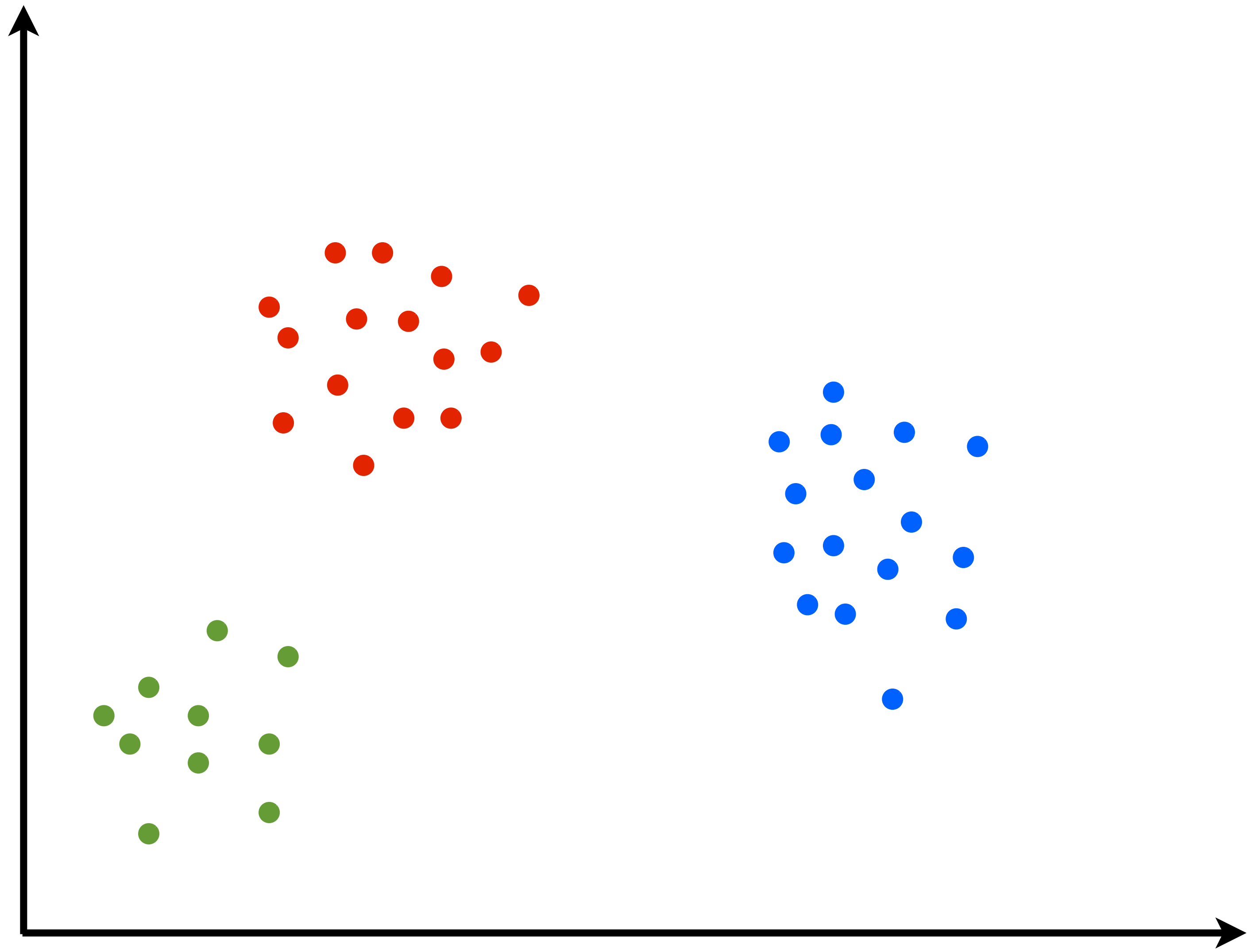
G_1 G_2 G_3 G_4



Clustering/partition

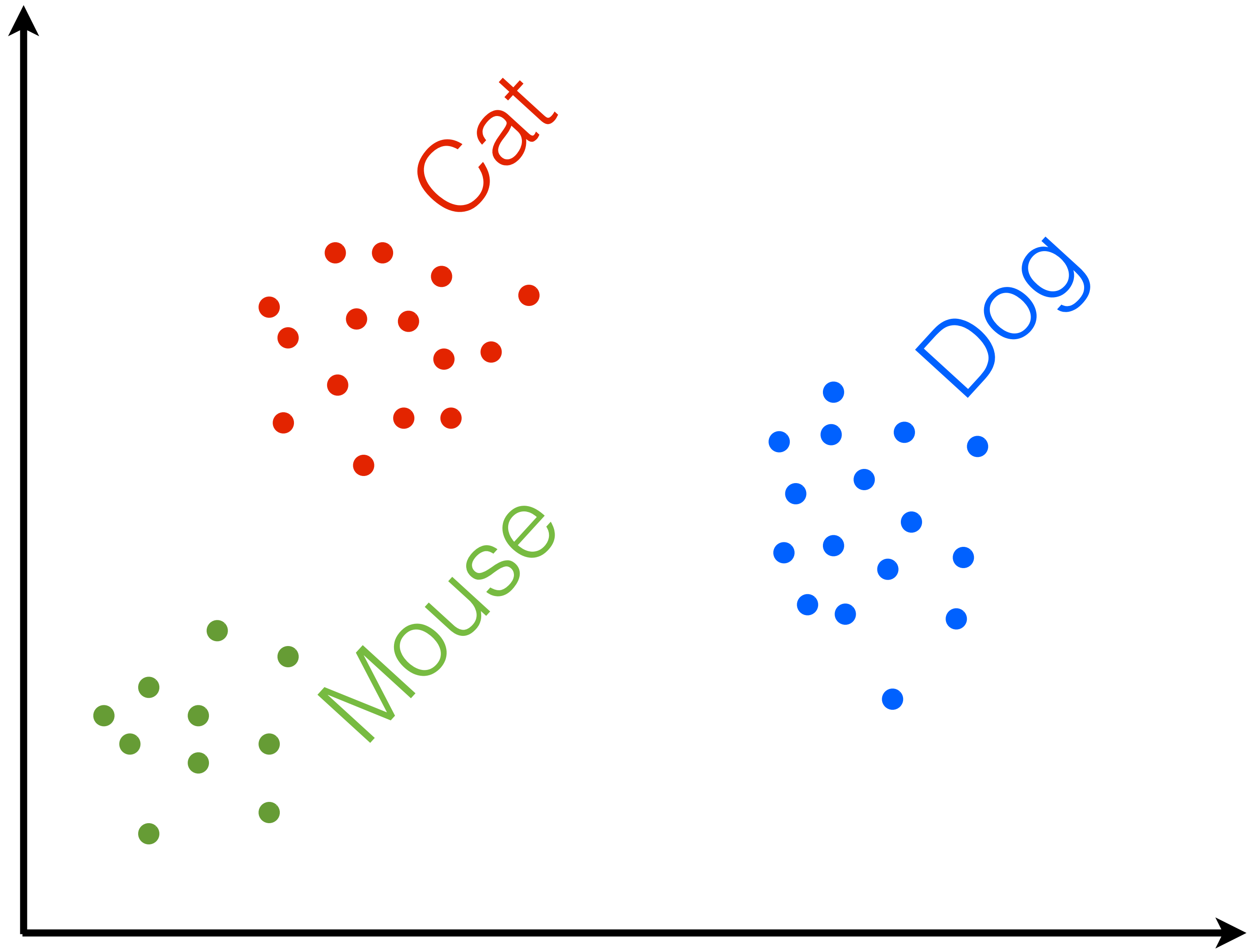


Clustering/partition



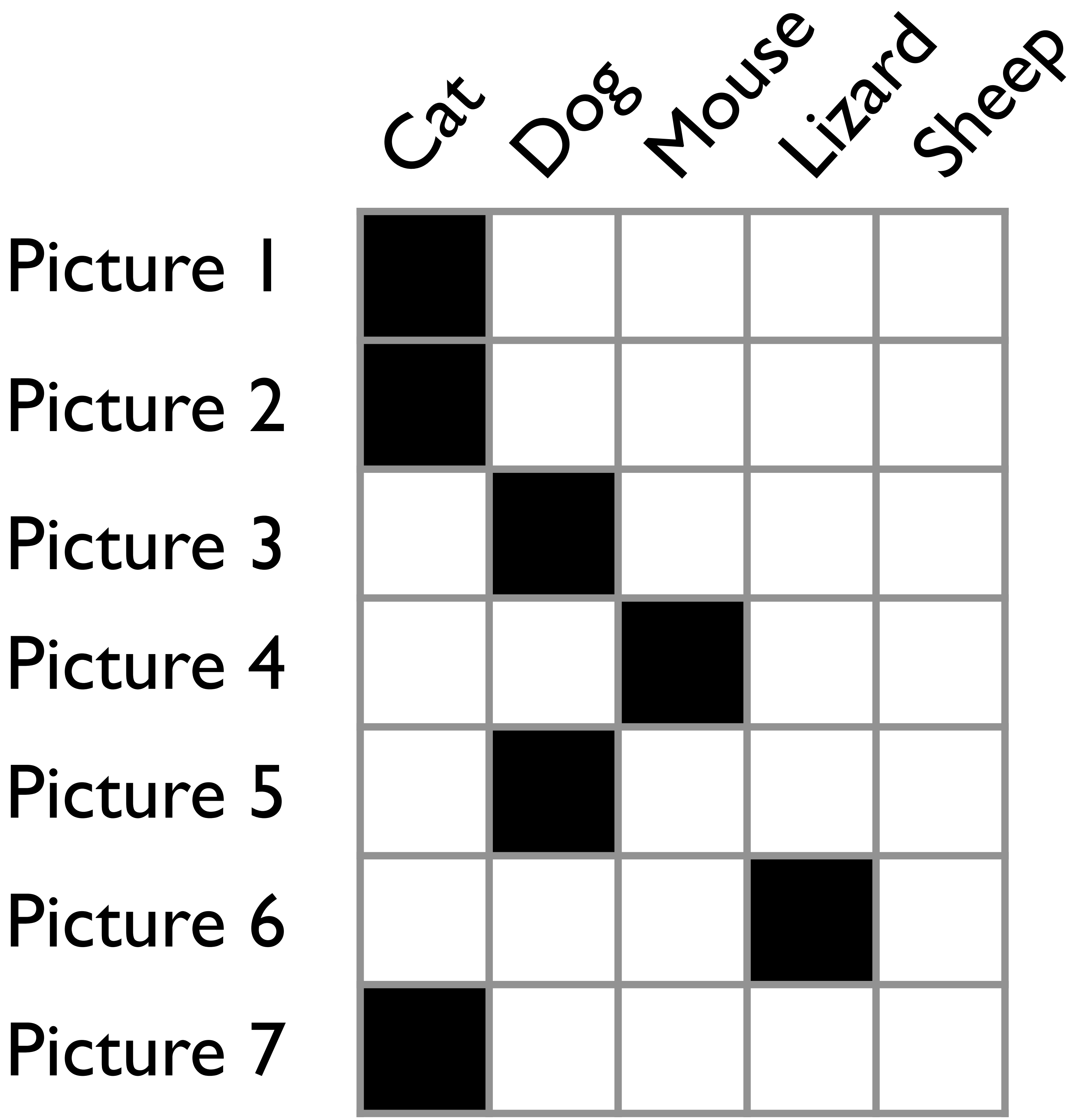
“Clusters”,
“classes”,
“blocks of a partition”

Clustering/partition

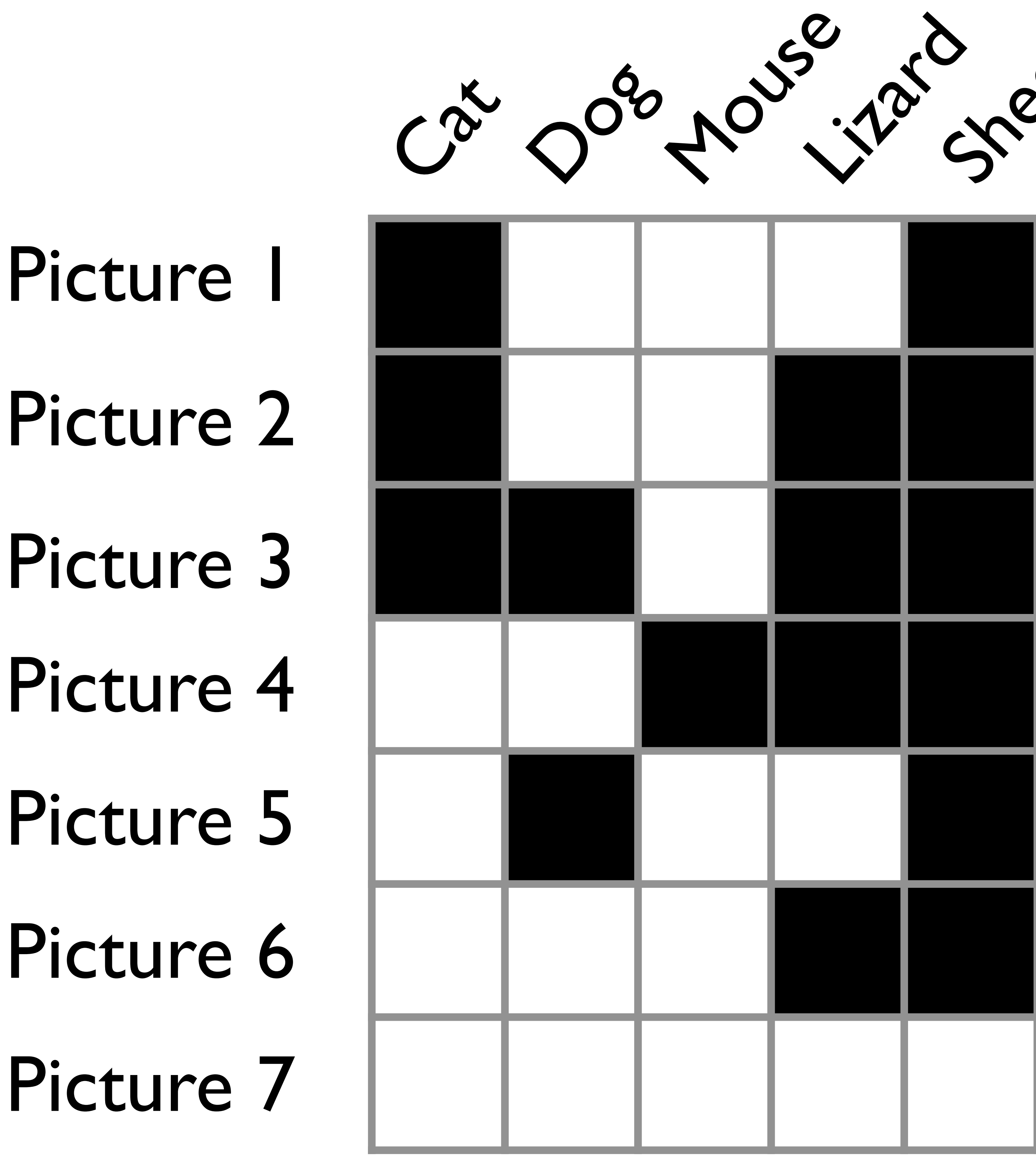


“Clusters”,
“classes”,
“blocks of a partition”

Clustering/partition

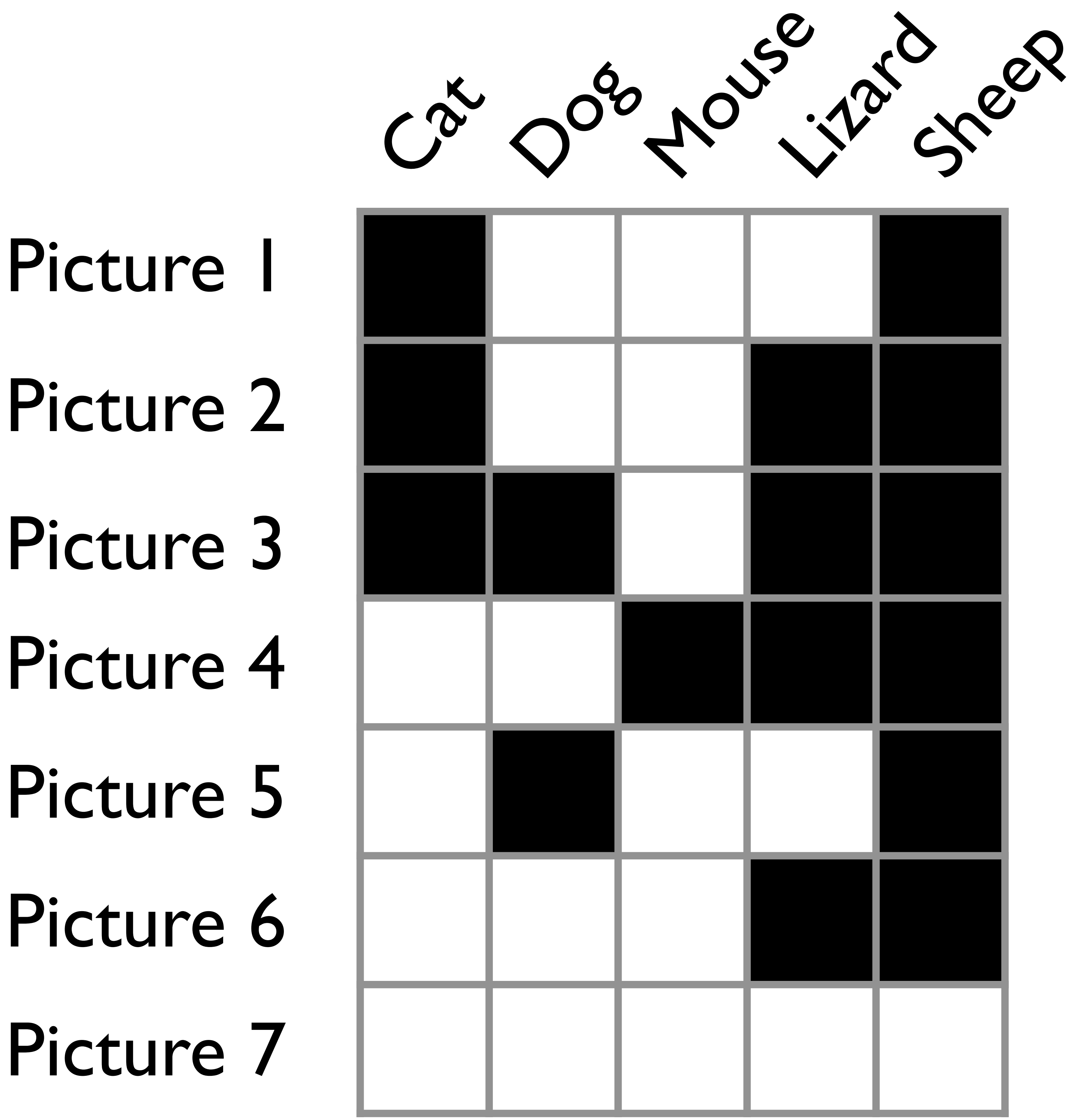


Feature allocation



“Features”,
“topics”

Feature allocation



“Features”,
“topics”

- Exchangeable
- Finite # of features per data point

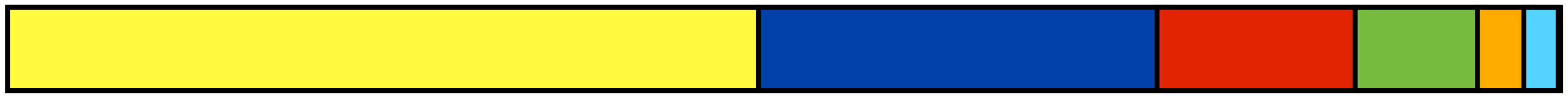
Paintboxes

Exchangeable partition: Kingman paintbox

[Kingman 1978]

Paintboxes

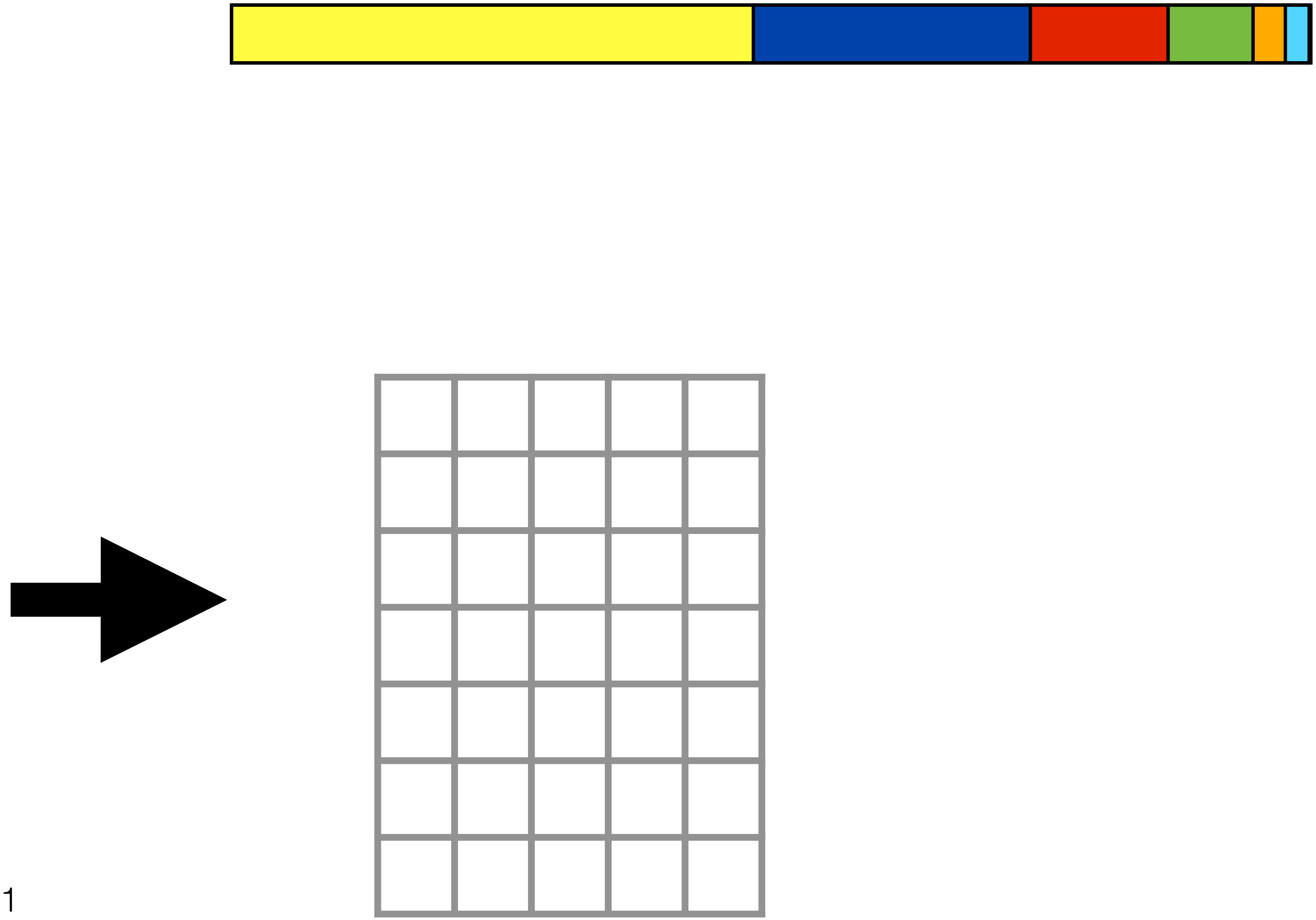
Exchangeable partition: Kingman paintbox



[Kingman 1978]

Paintboxes

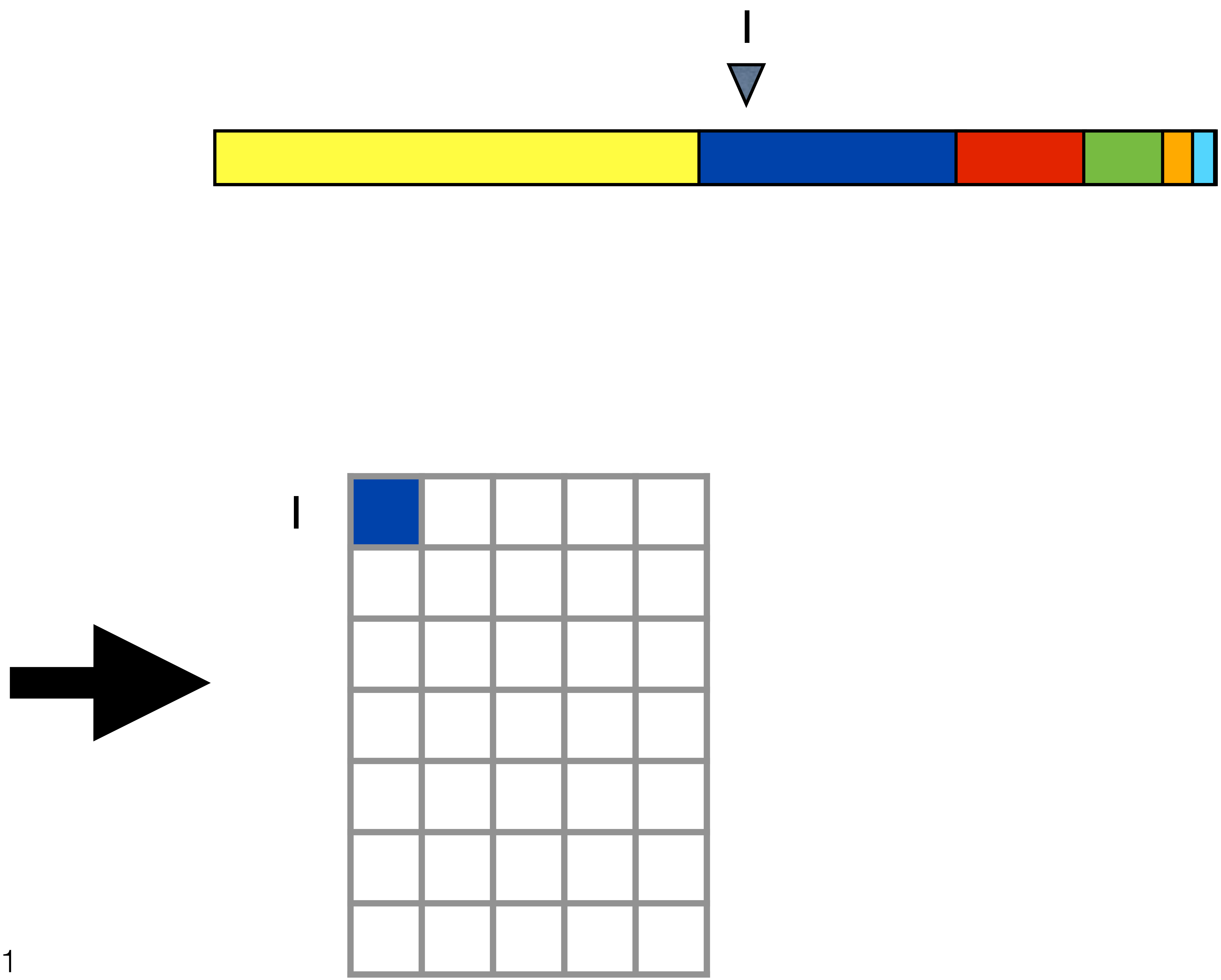
Exchangeable partition: Kingman paintbox



[Kingman 1978]

Paintboxes

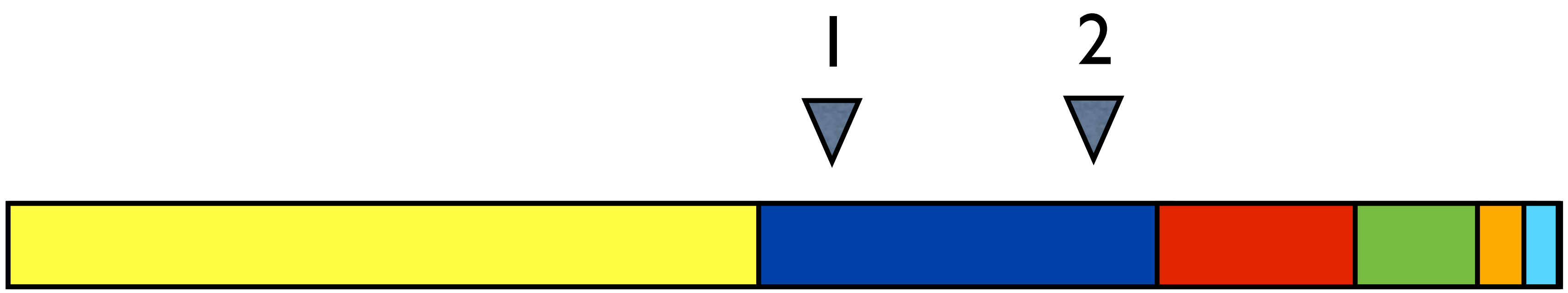
Exchangeable partition: Kingman paintbox

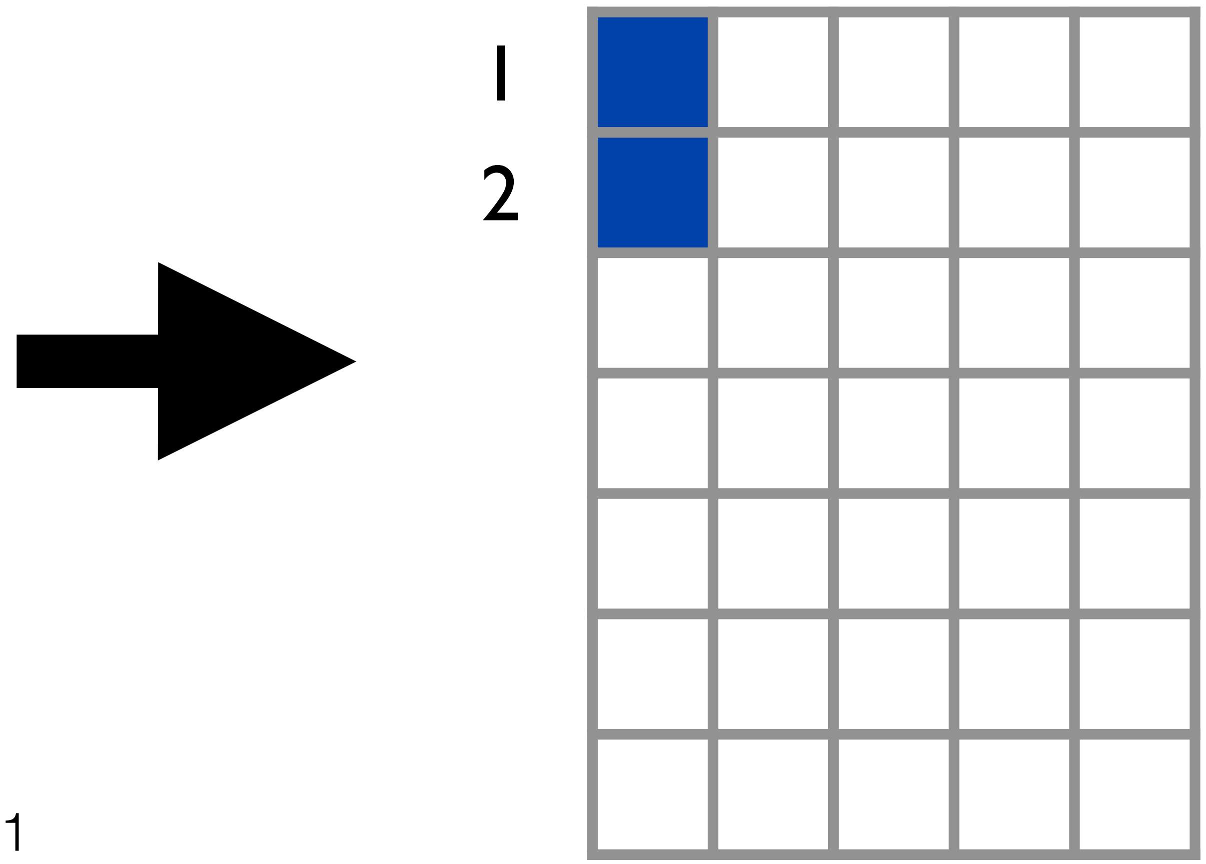


[Kingman 1978]

Paintboxes

Exchangeable partition: Kingman paintbox

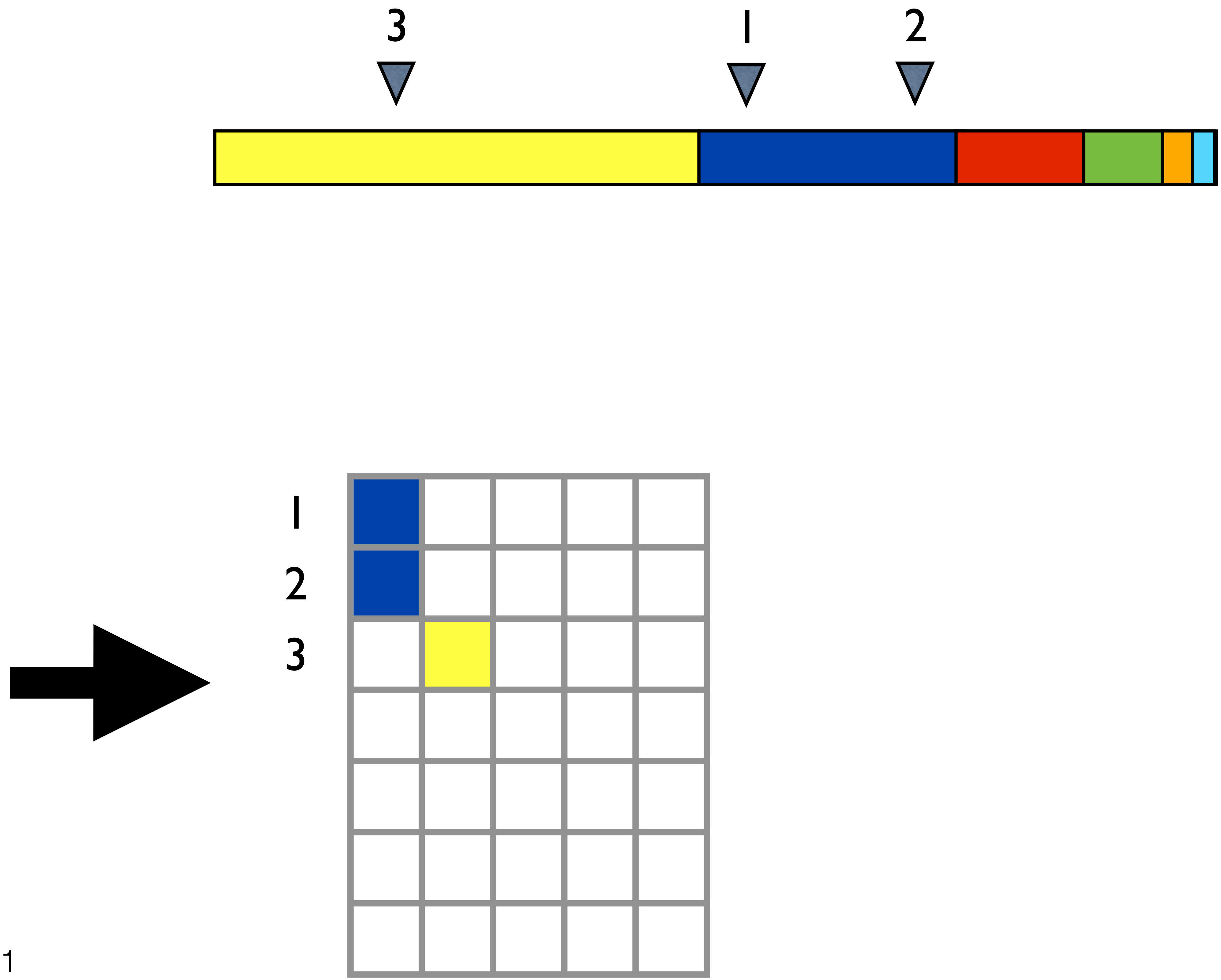




[Kingman 1978]

Paintboxes

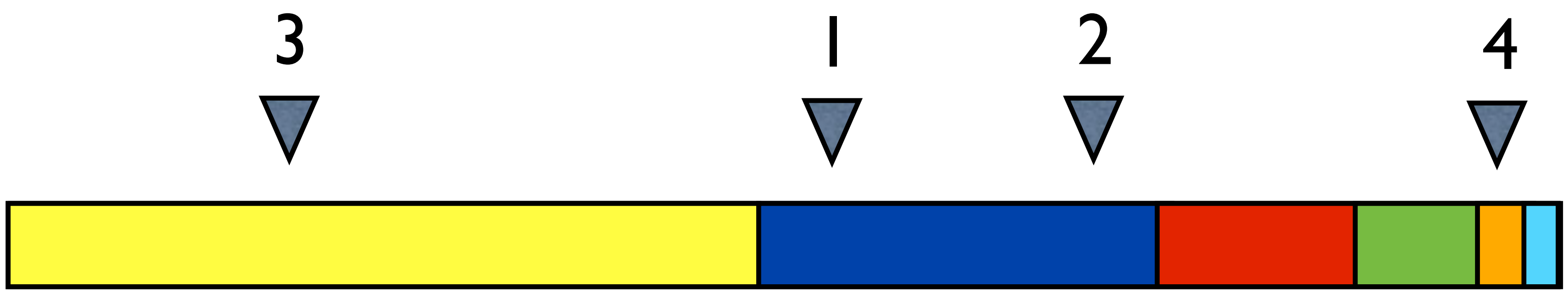
Exchangeable partition: Kingman paintbox

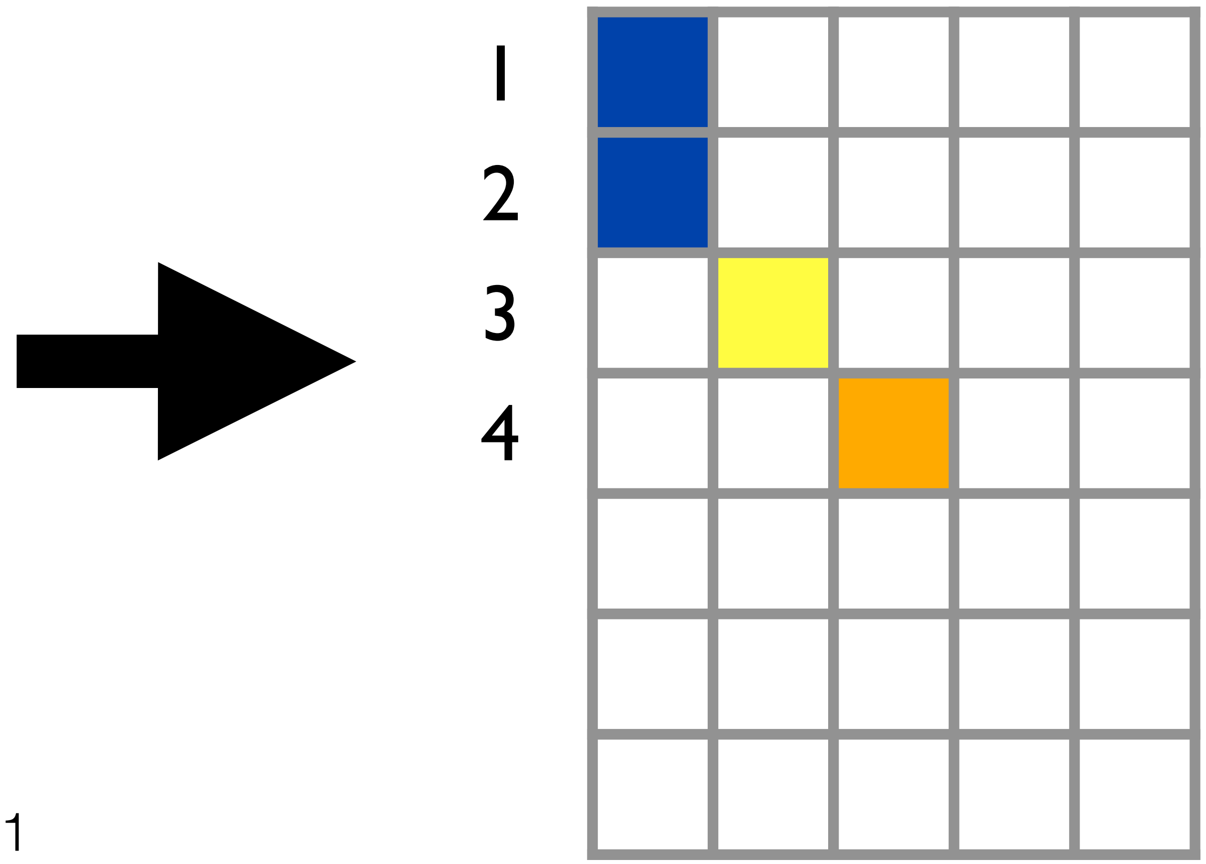


[Kingman 1978]

Paintboxes

Exchangeable partition: Kingman paintbox

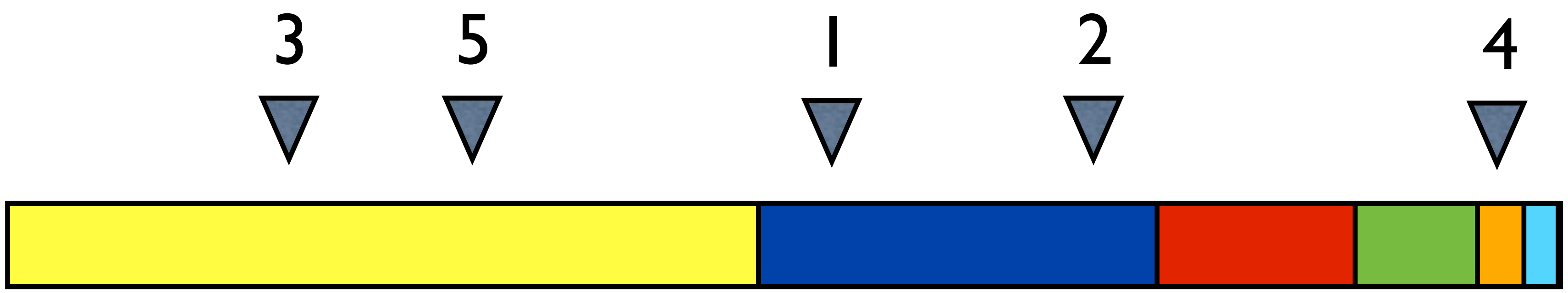


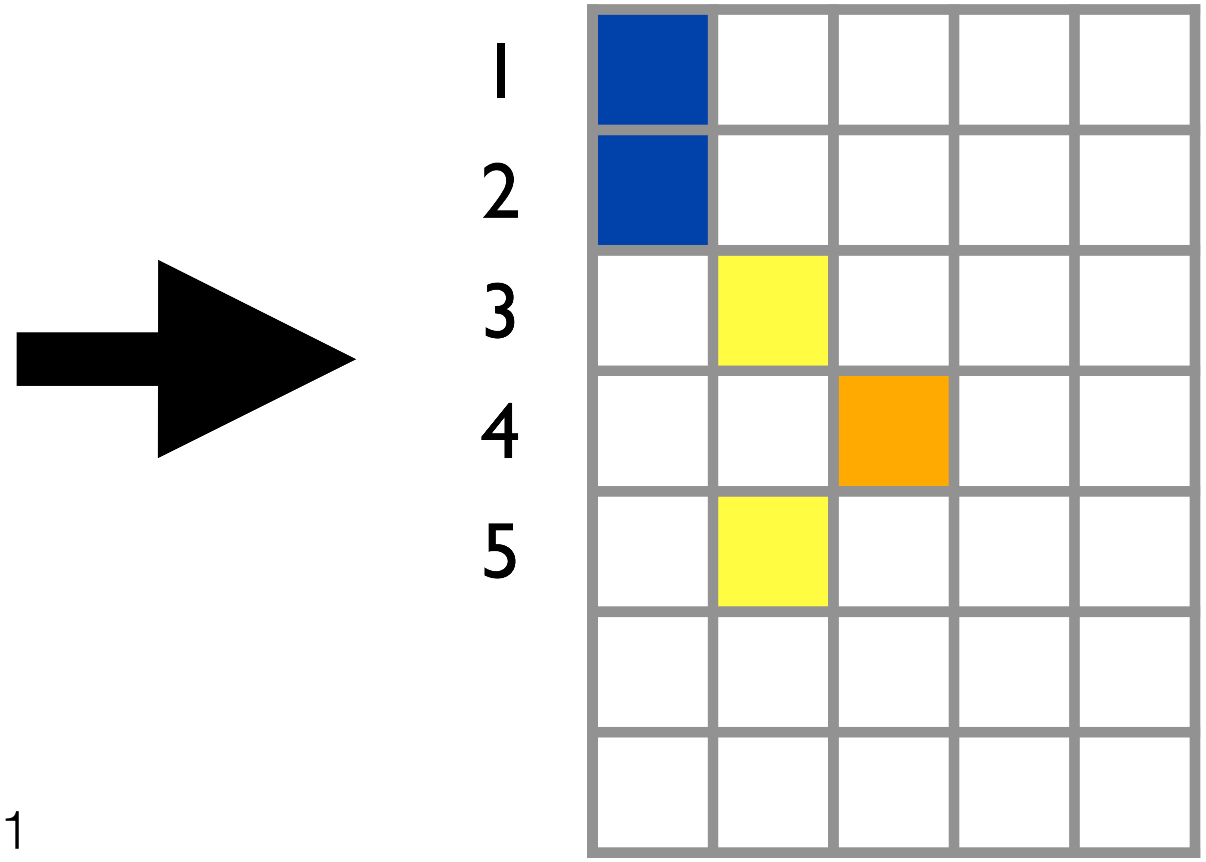


[Kingman 1978]

Paintboxes

Exchangeable partition: Kingman paintbox

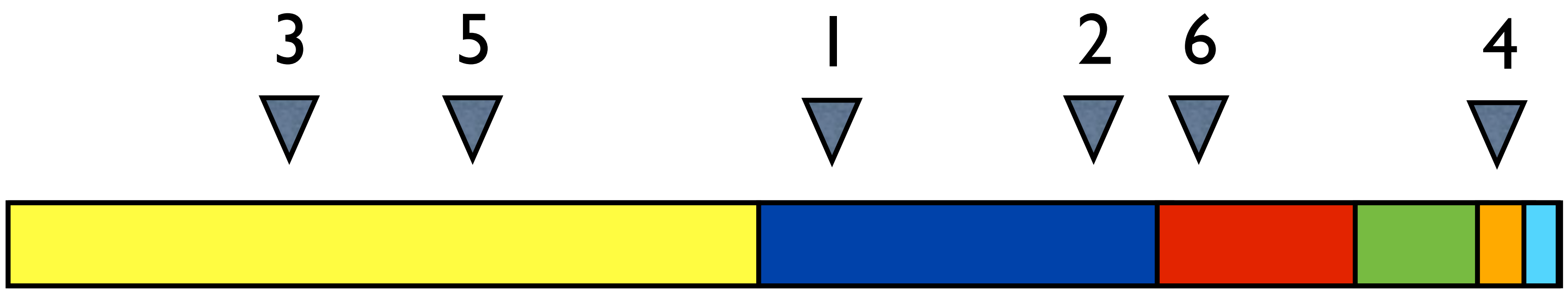


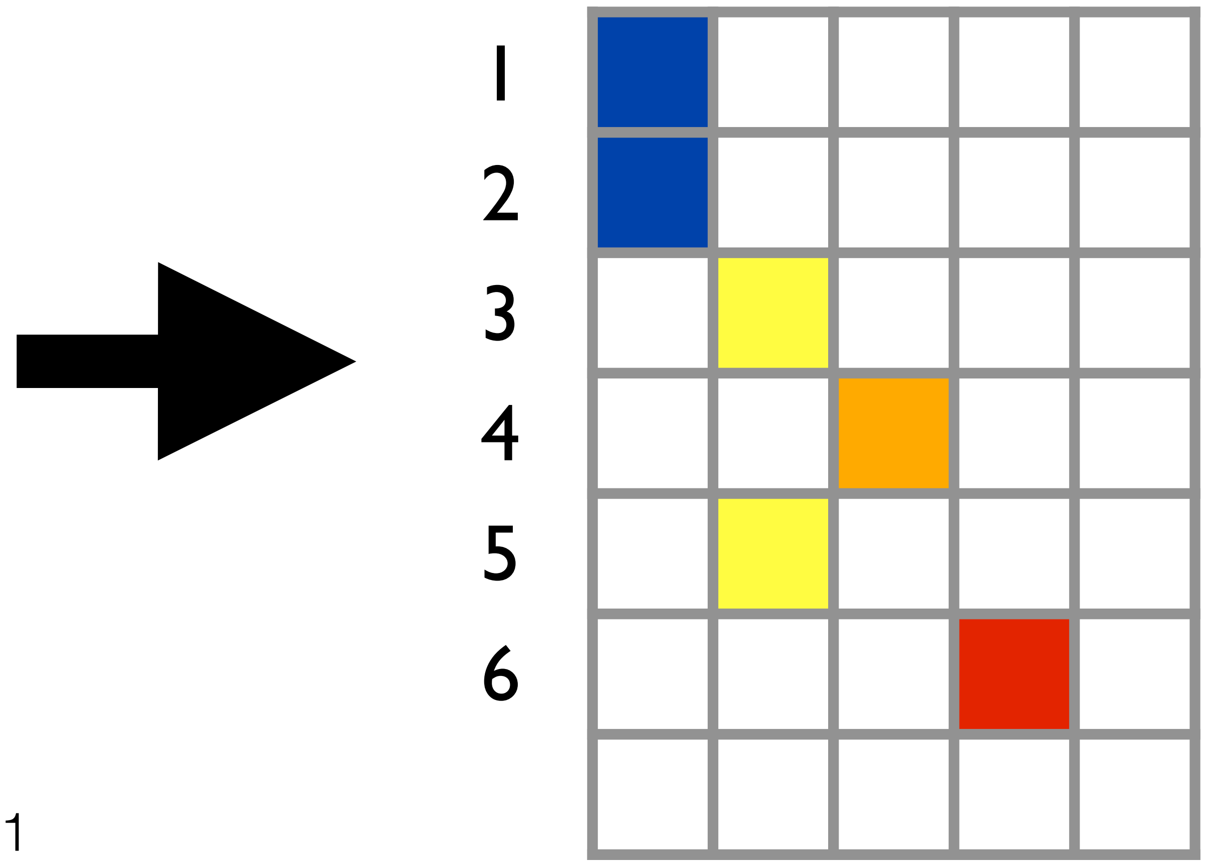


[Kingman 1978]

Paintboxes

Exchangeable partition: Kingman paintbox

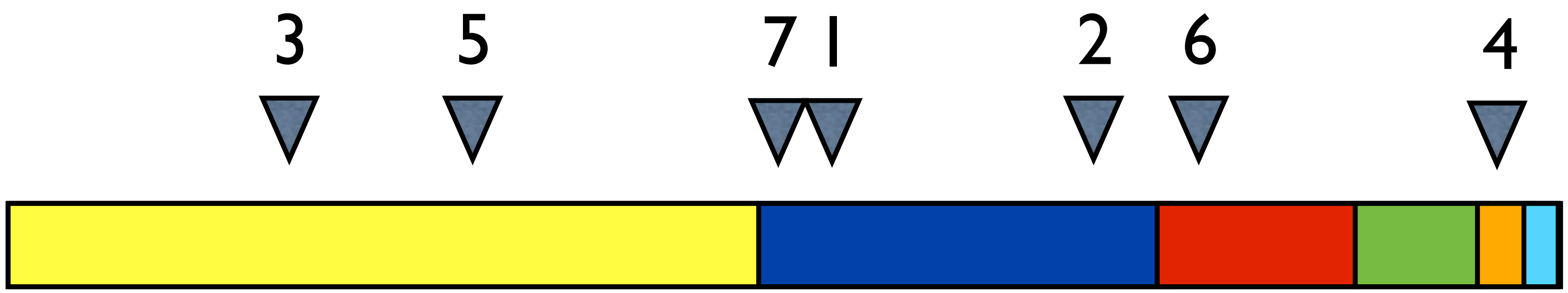


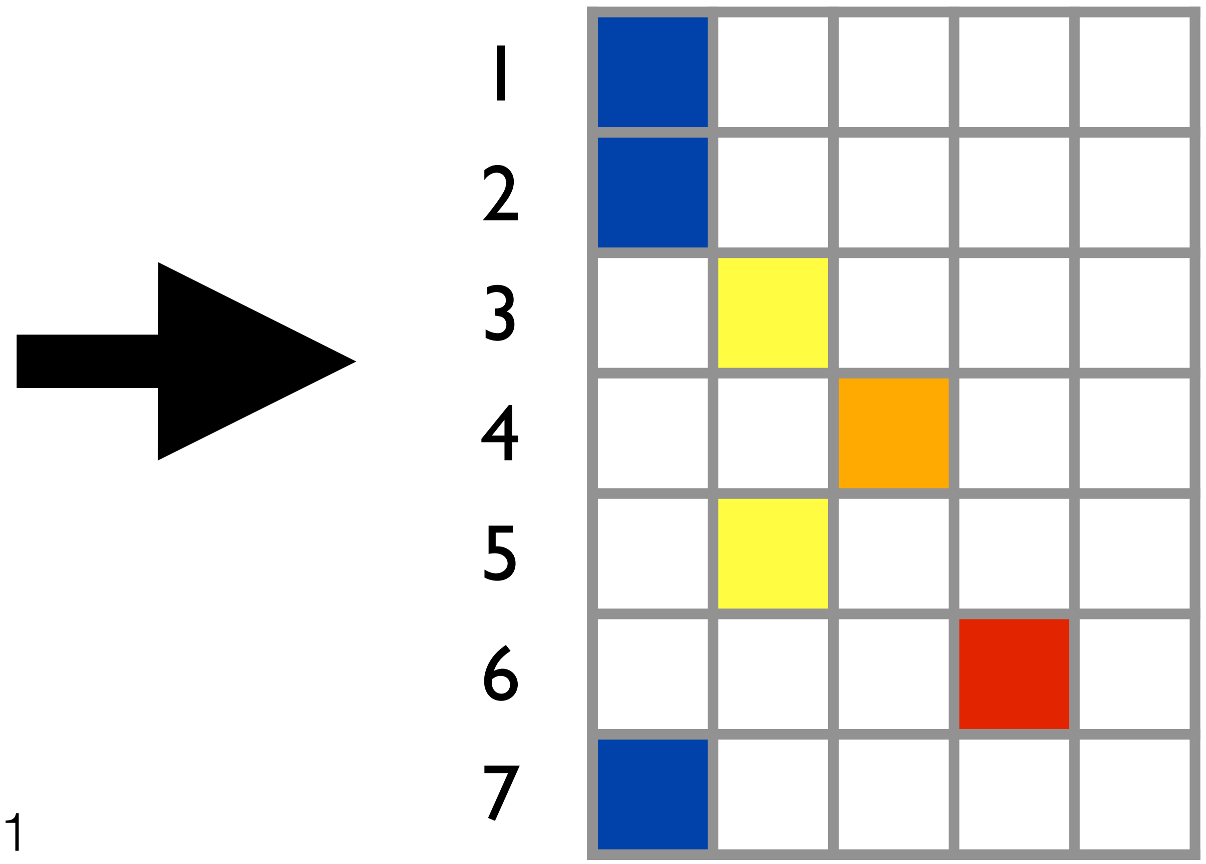


[Kingman 1978]

Paintboxes

Exchangeable partition: Kingman paintbox

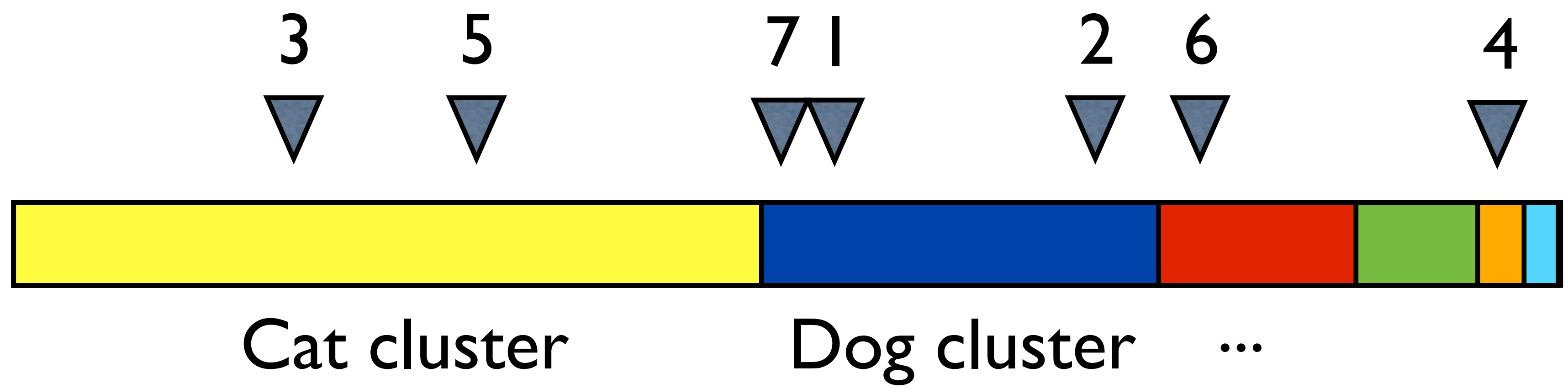


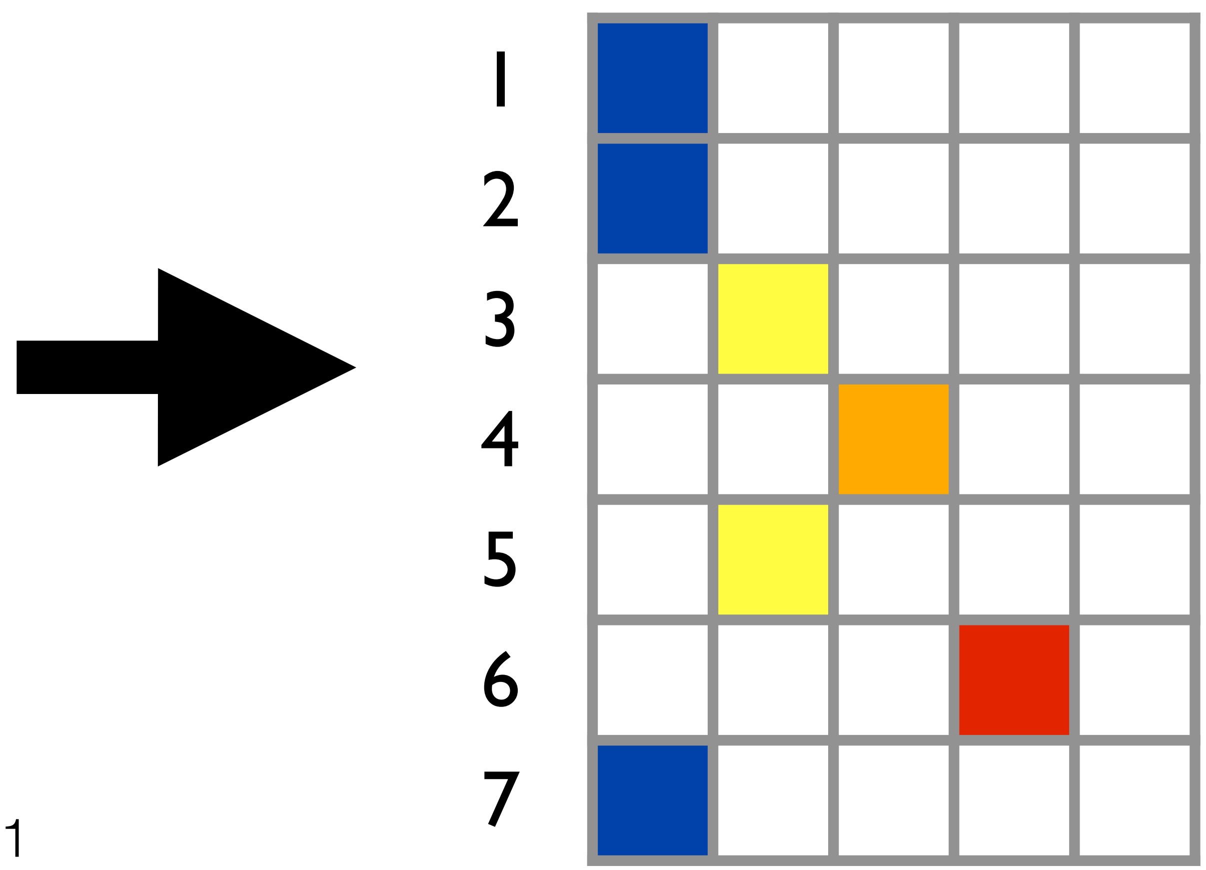


[Kingman 1978]

Paintboxes

Exchangeable partition: Kingman paintbox

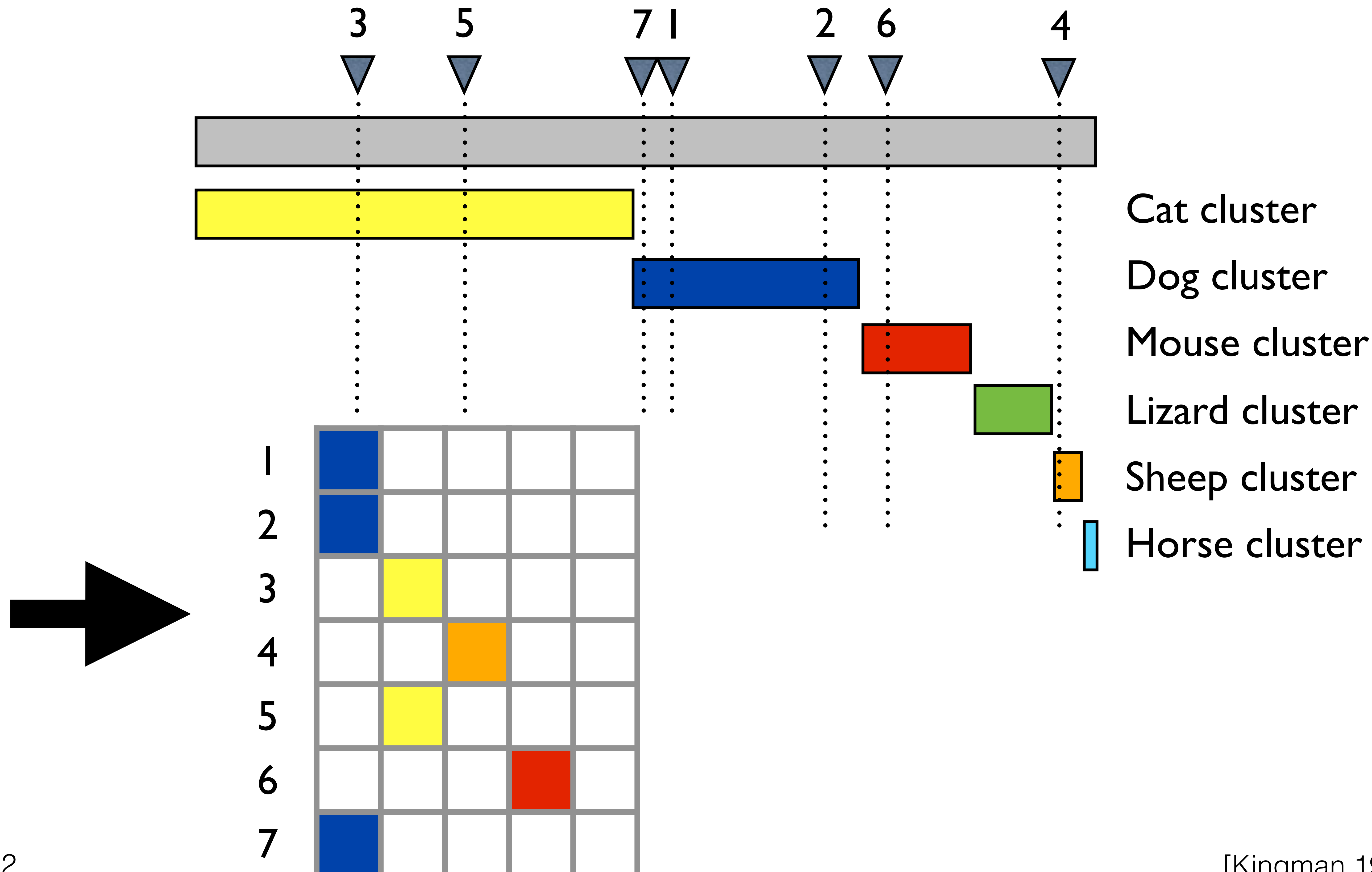




[Kingman 1978]

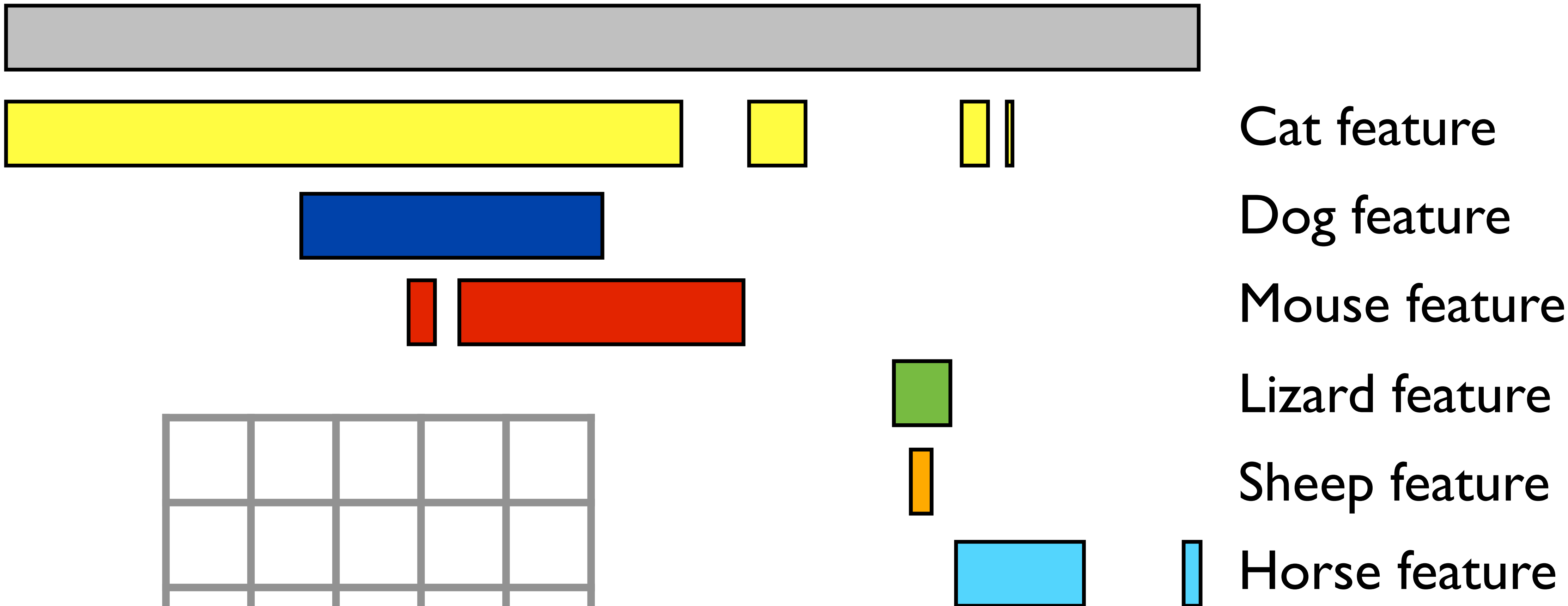
Paintboxes

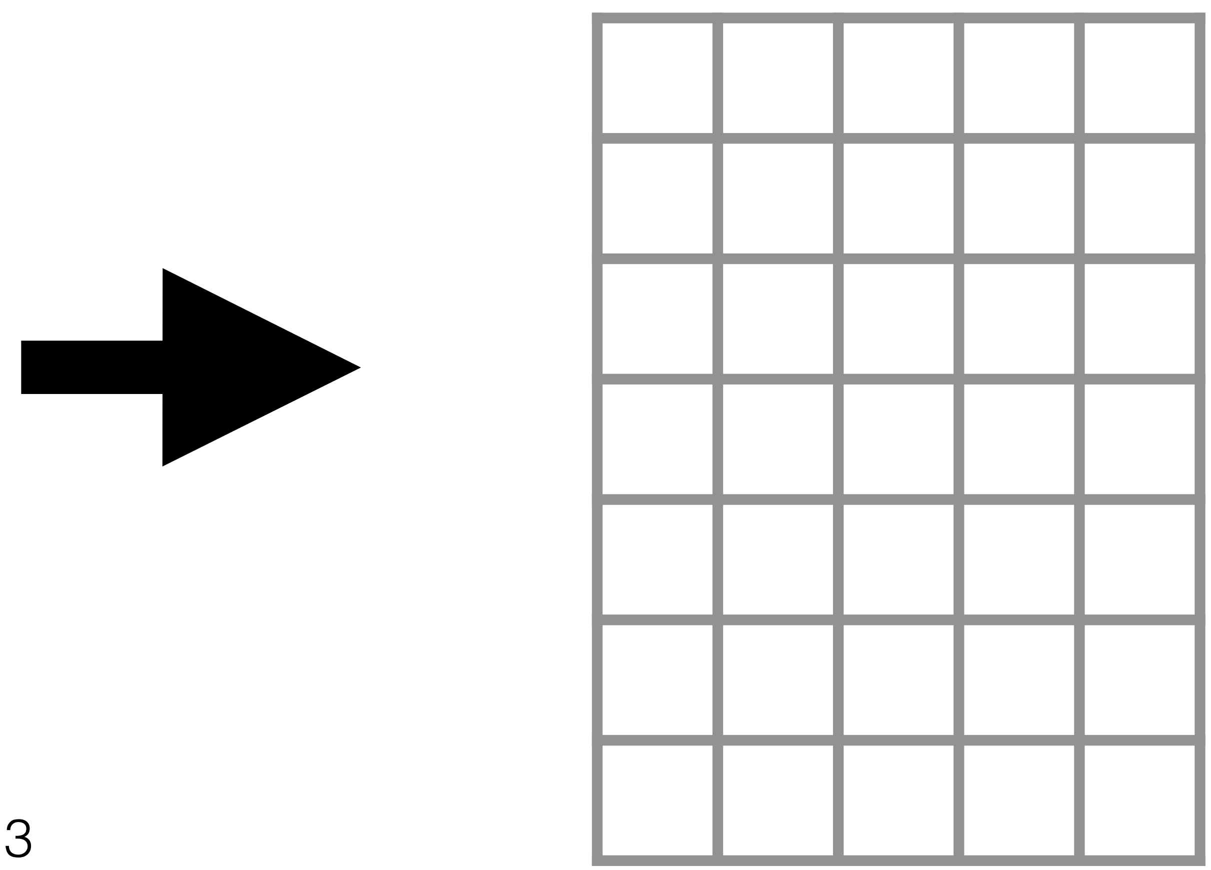
Exchangeable partition: Kingman paintbox



[Kingman 1978]

Paintboxes

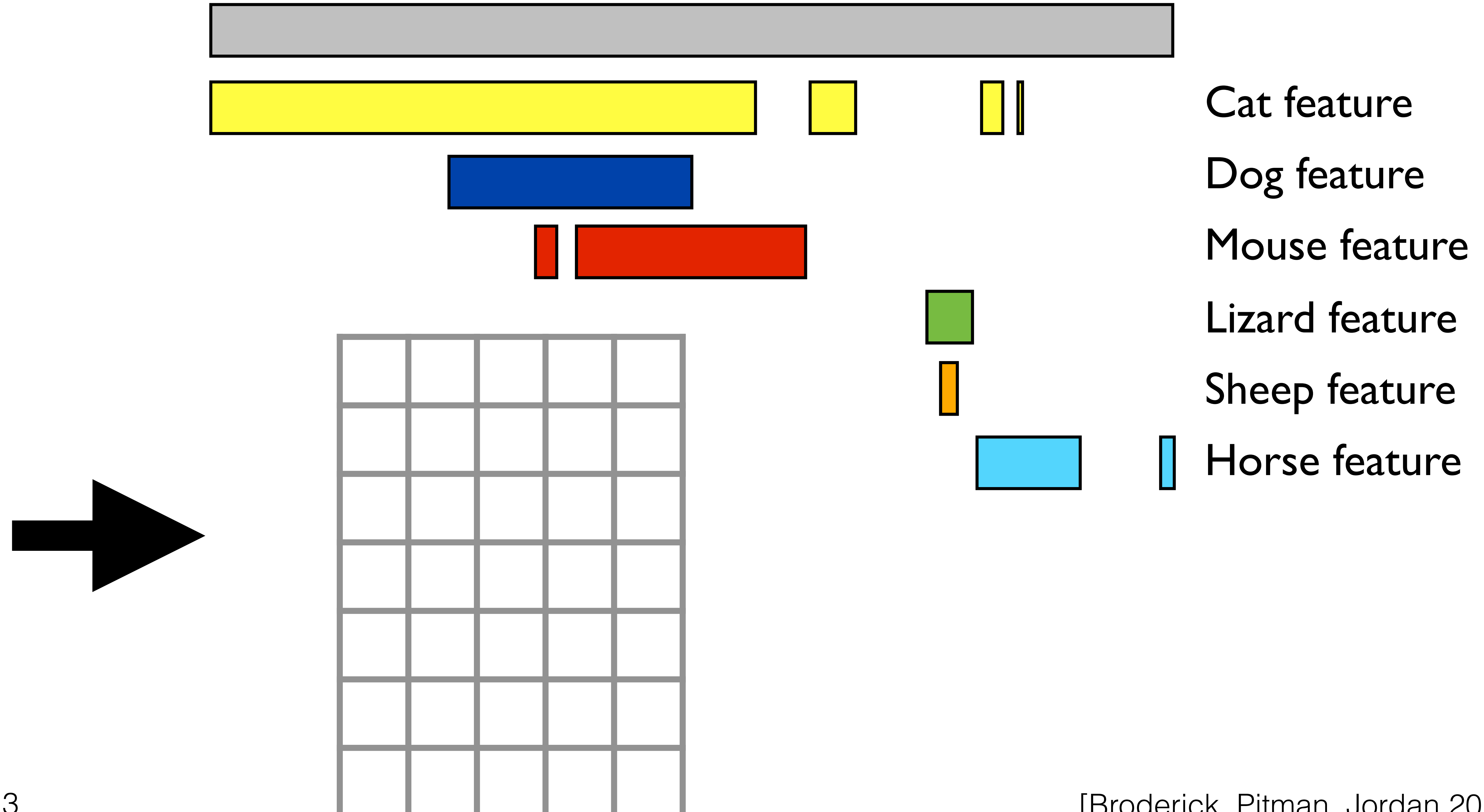




[Broderick, Pitman, Jordan 2013]

Paintboxes

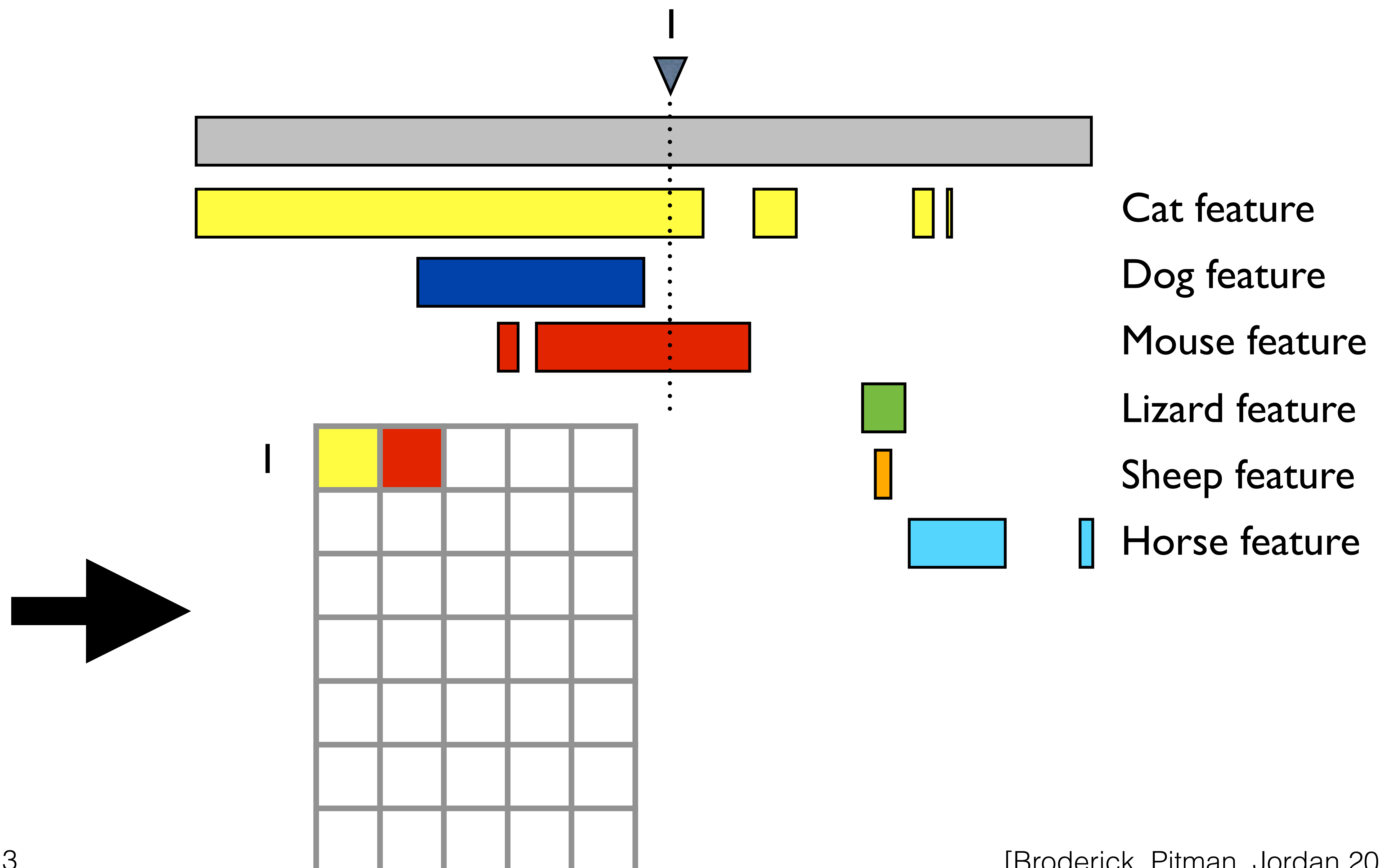
Exchangeable feature allocation: feature paintbox



[Broderick, Pitman, Jordan 2013]

Paintboxes

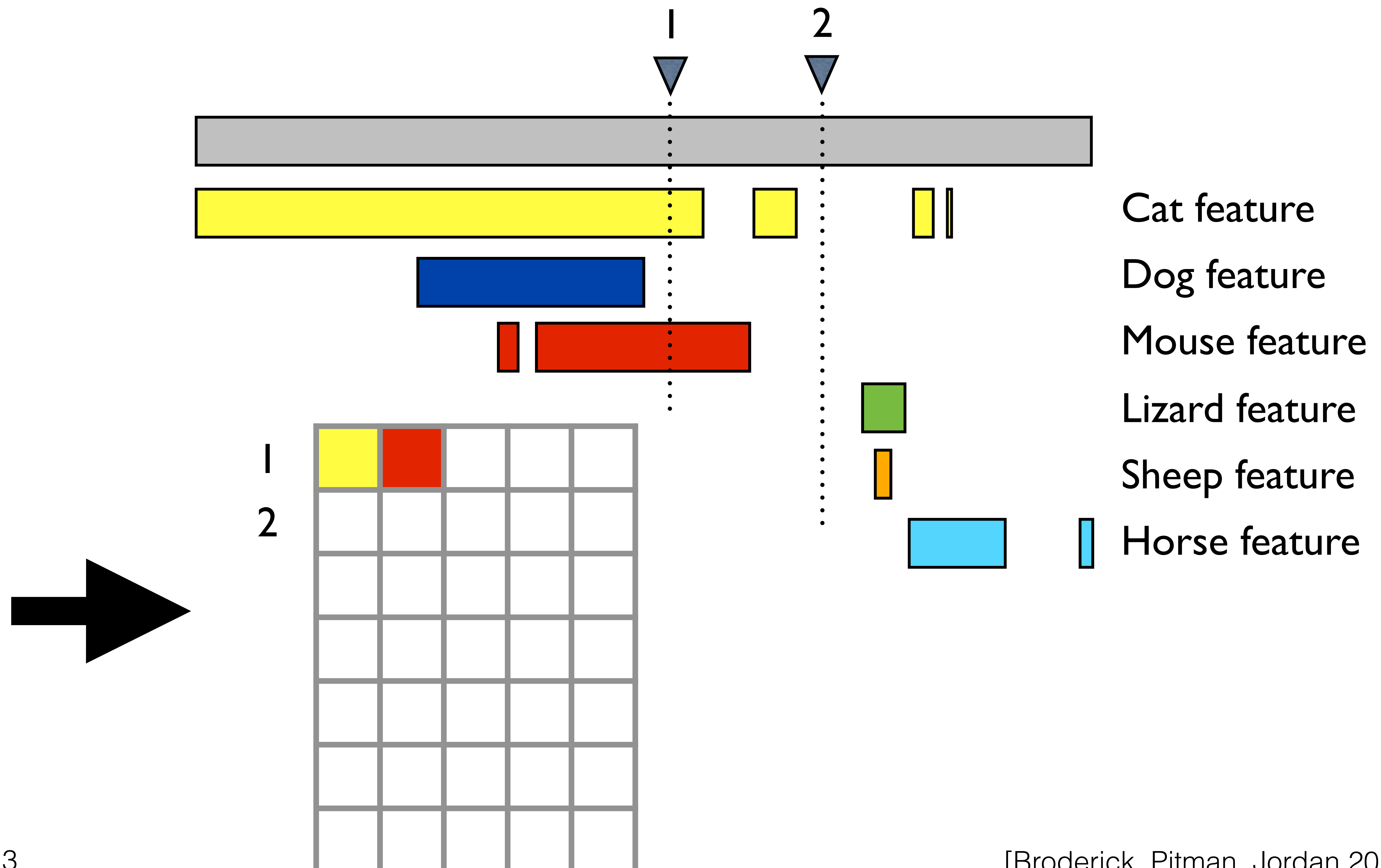
Exchangeable feature allocation: feature paintbox



[Broderick, Pitman, Jordan 2013]

Paintboxes

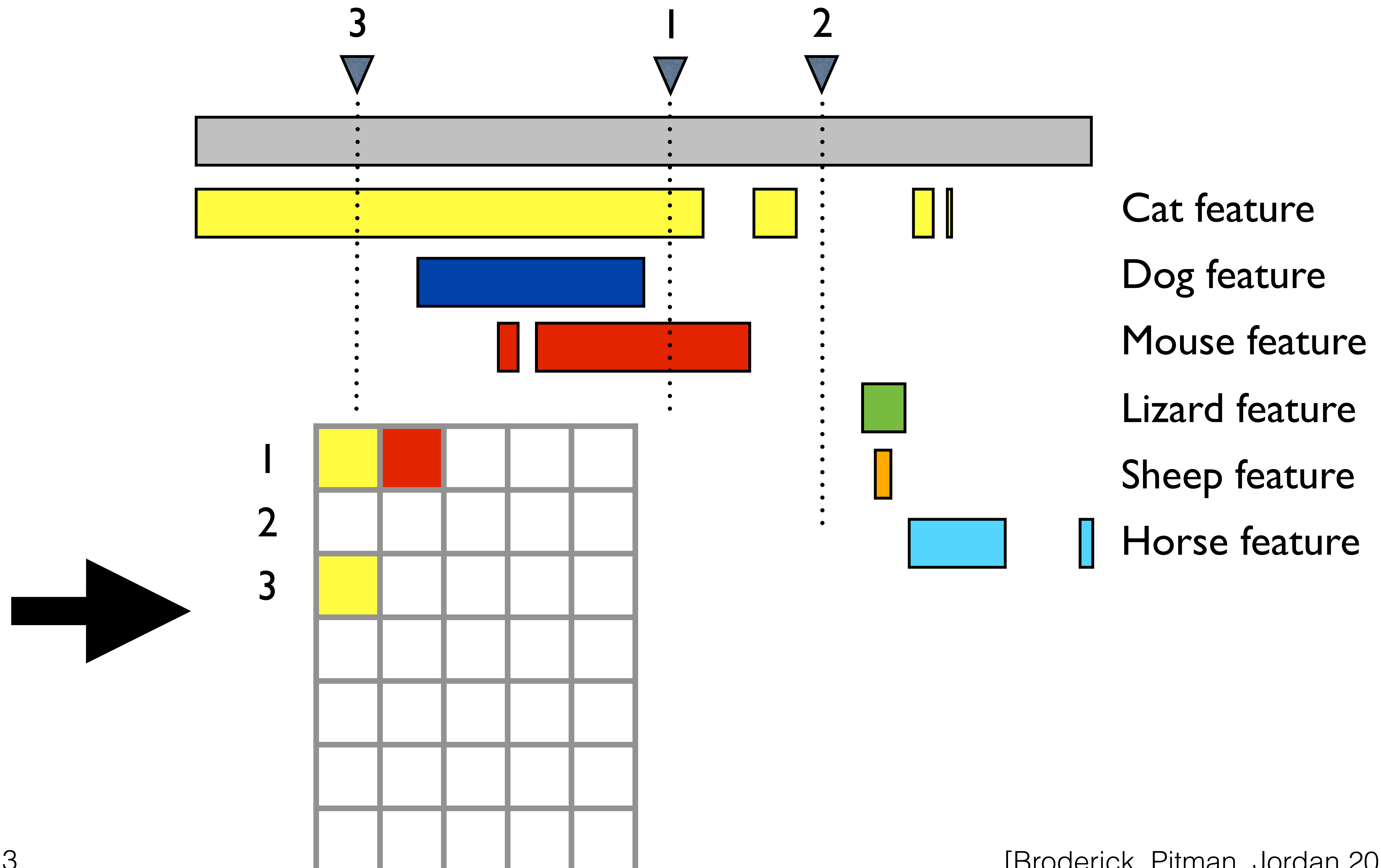
Exchangeable feature allocation: feature paintbox



[Broderick, Pitman, Jordan 2013]

Paintboxes

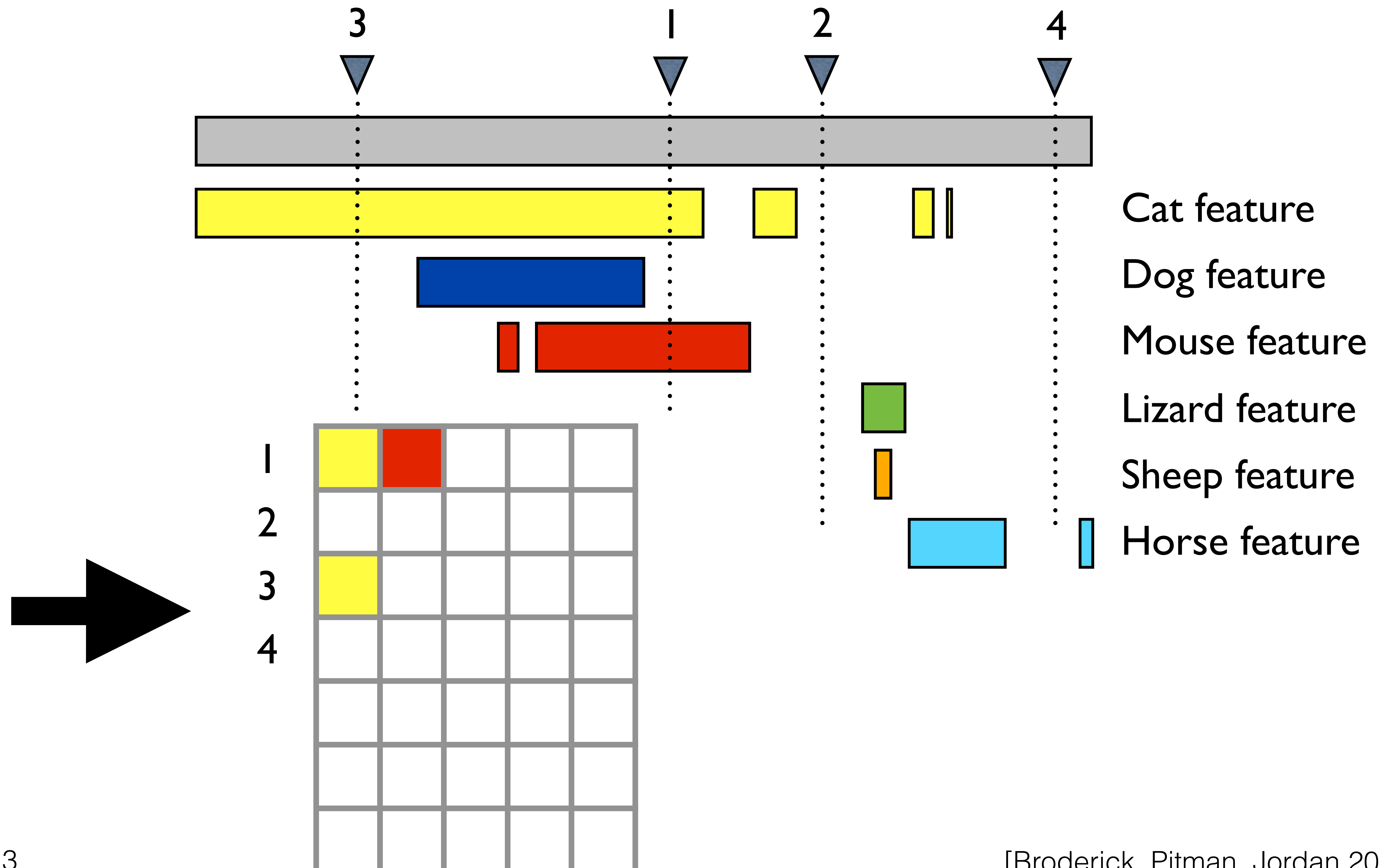
Exchangeable feature allocation: feature paintbox



[Broderick, Pitman, Jordan 2013]

Paintboxes

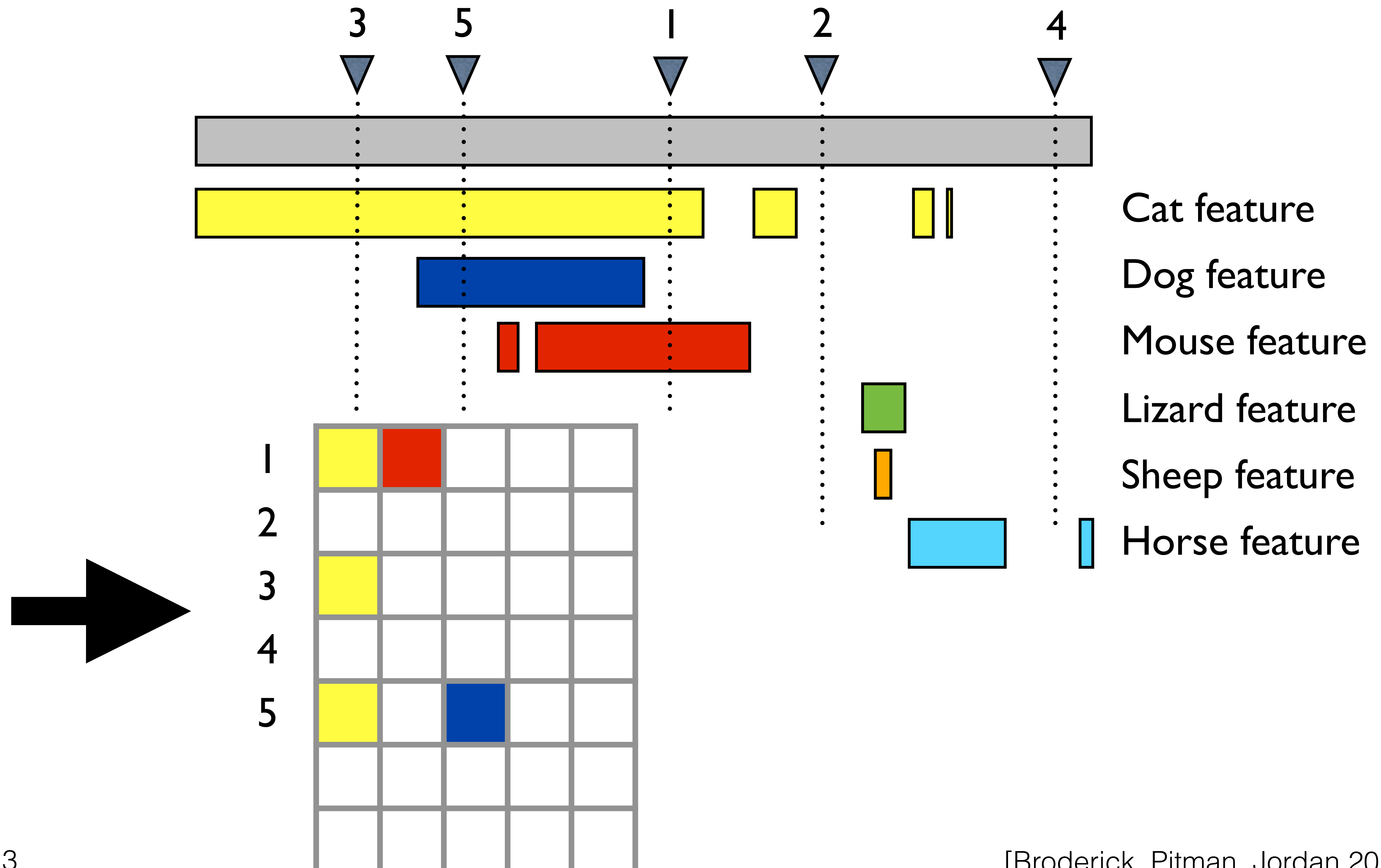
Exchangeable feature allocation: feature paintbox



[Broderick, Pitman, Jordan 2013]

Paintboxes

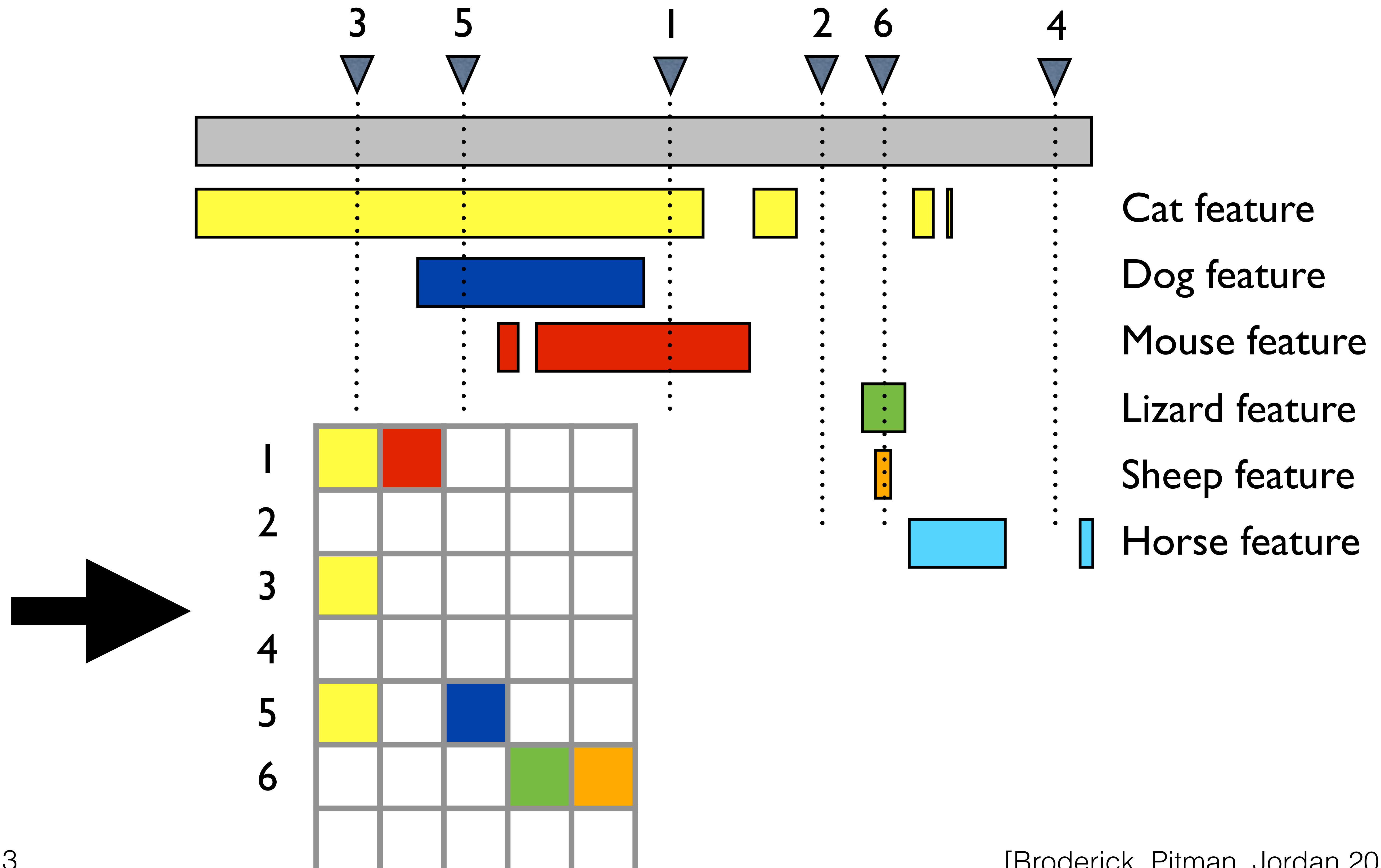
Exchangeable feature allocation: feature paintbox



[Broderick, Pitman, Jordan 2013]

Paintboxes

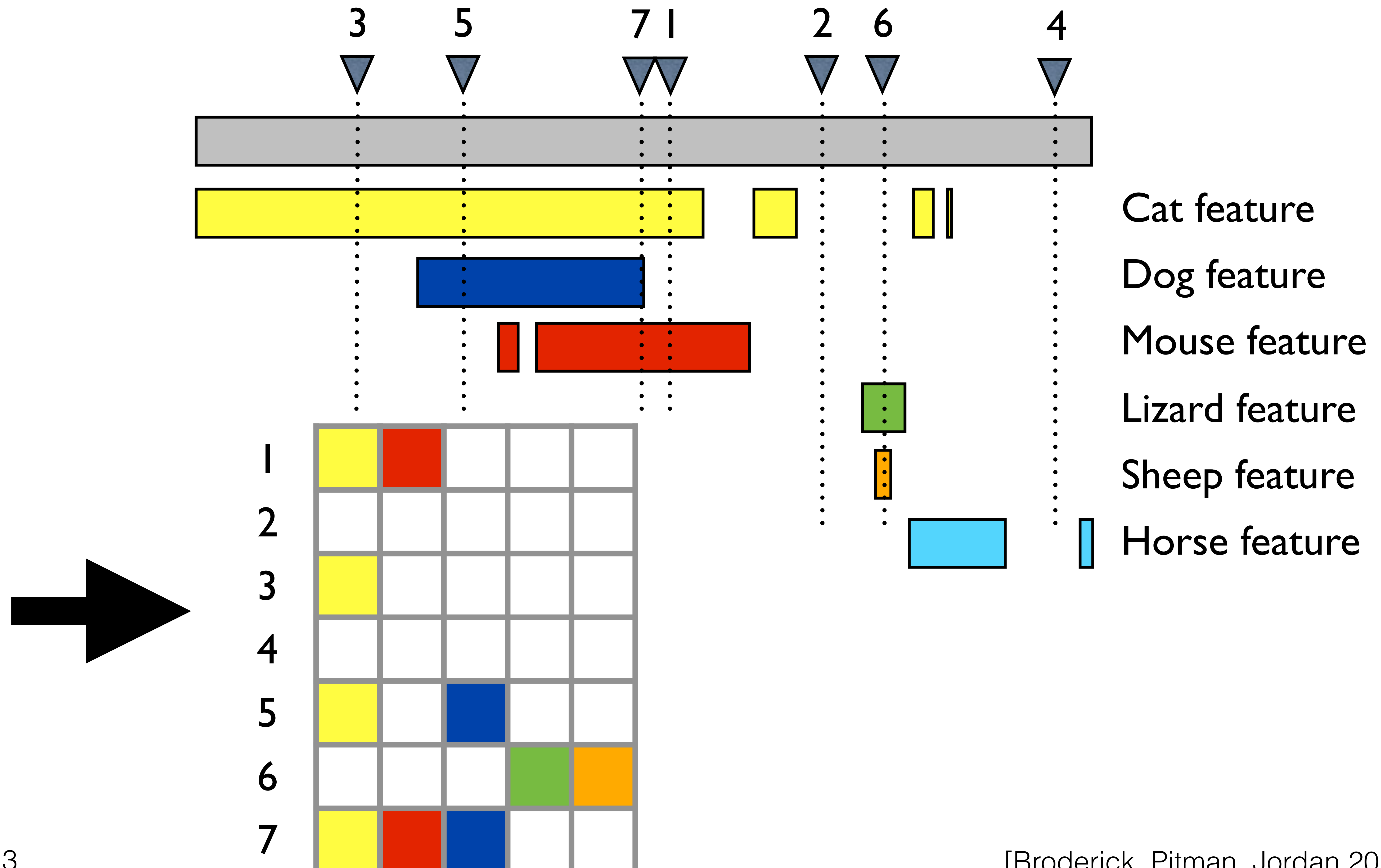
Exchangeable feature allocation: feature paintbox



[Broderick, Pitman, Jordan 2013]

Paintboxes

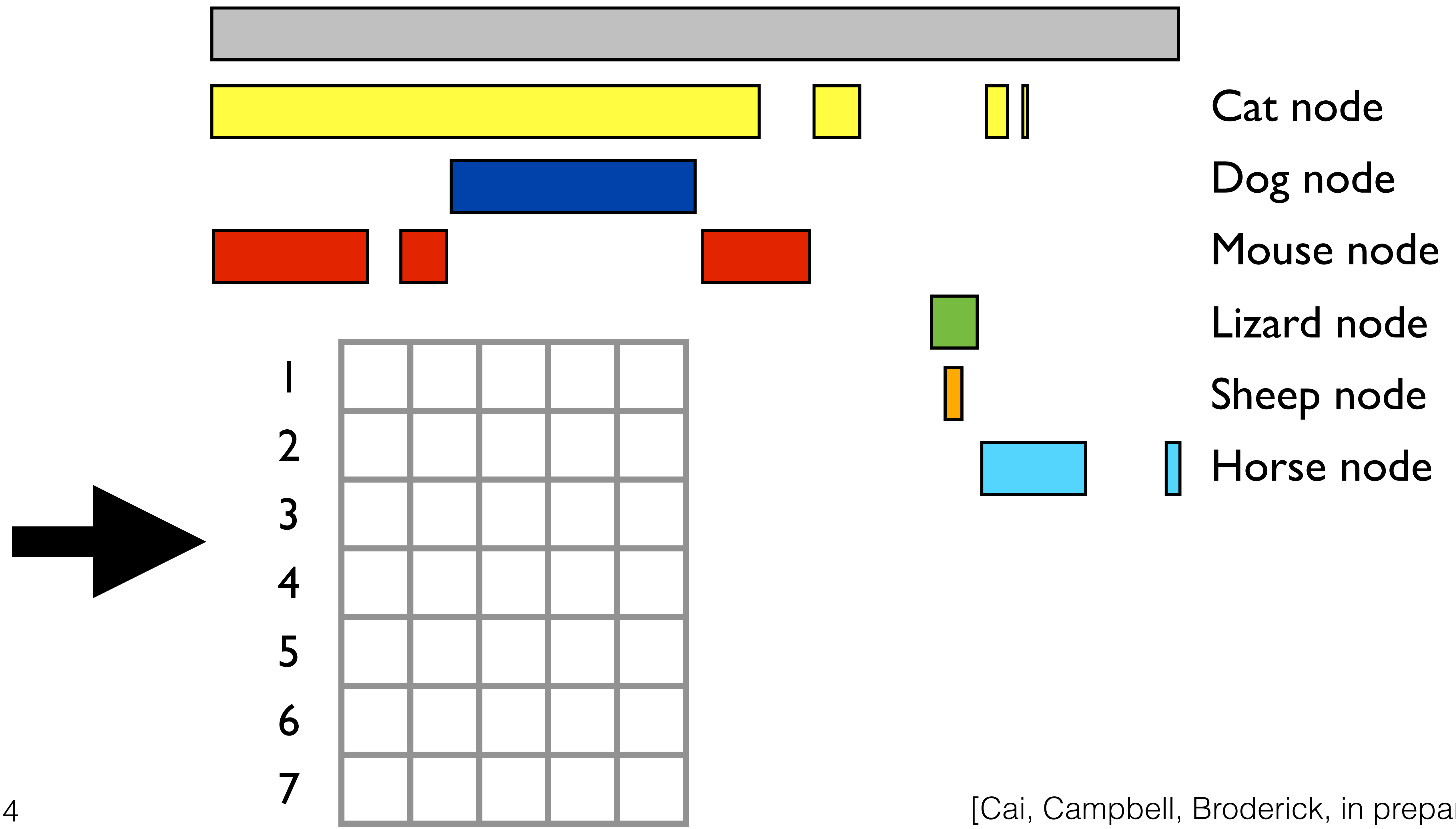
Exchangeable feature allocation: feature paintbox



[Broderick, Pitman, Jordan 2013]

Paintboxes

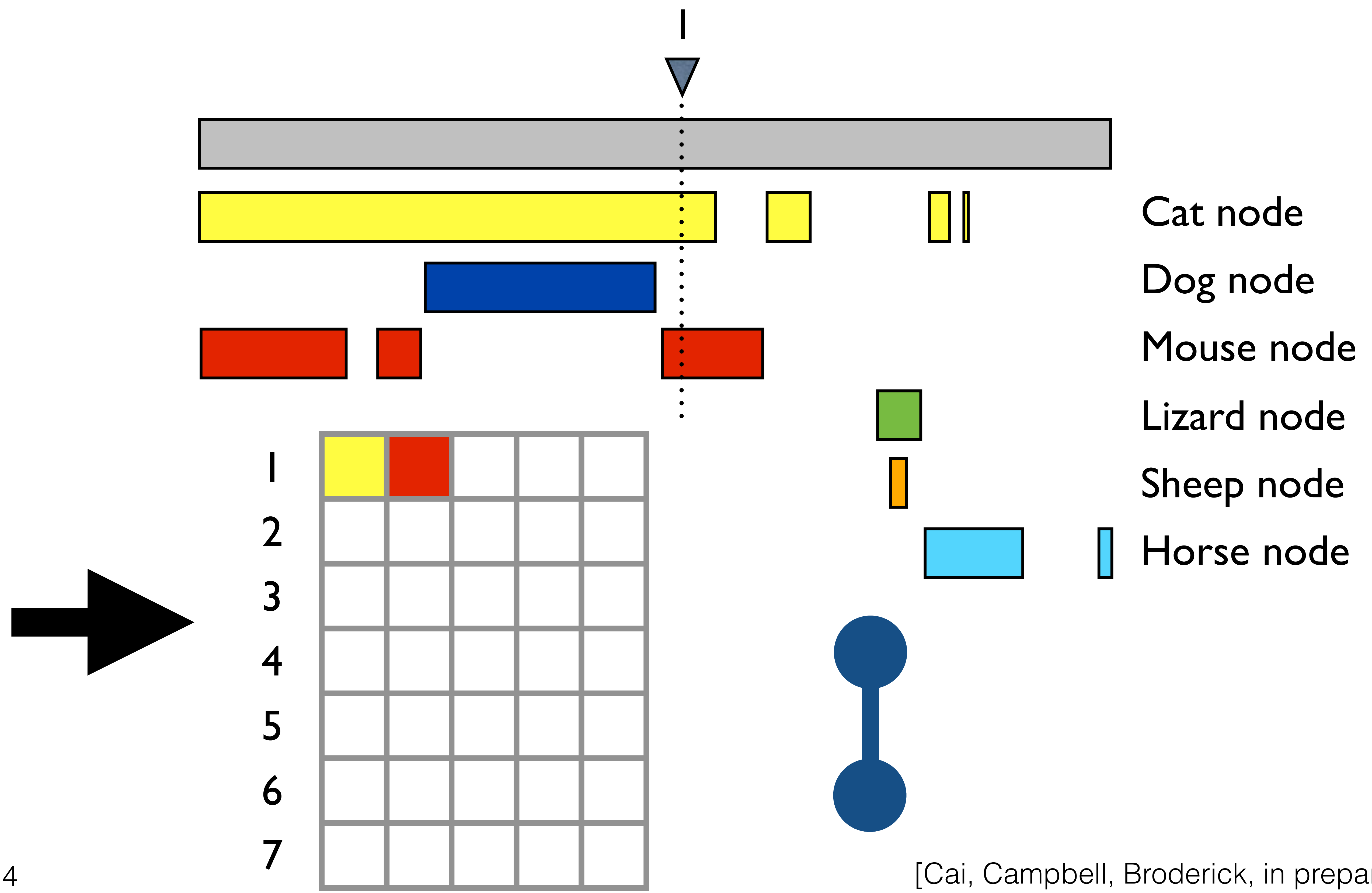
Edge-exchangeable graph



[Cai, Campbell, Broderick, in preparation]

Paintboxes

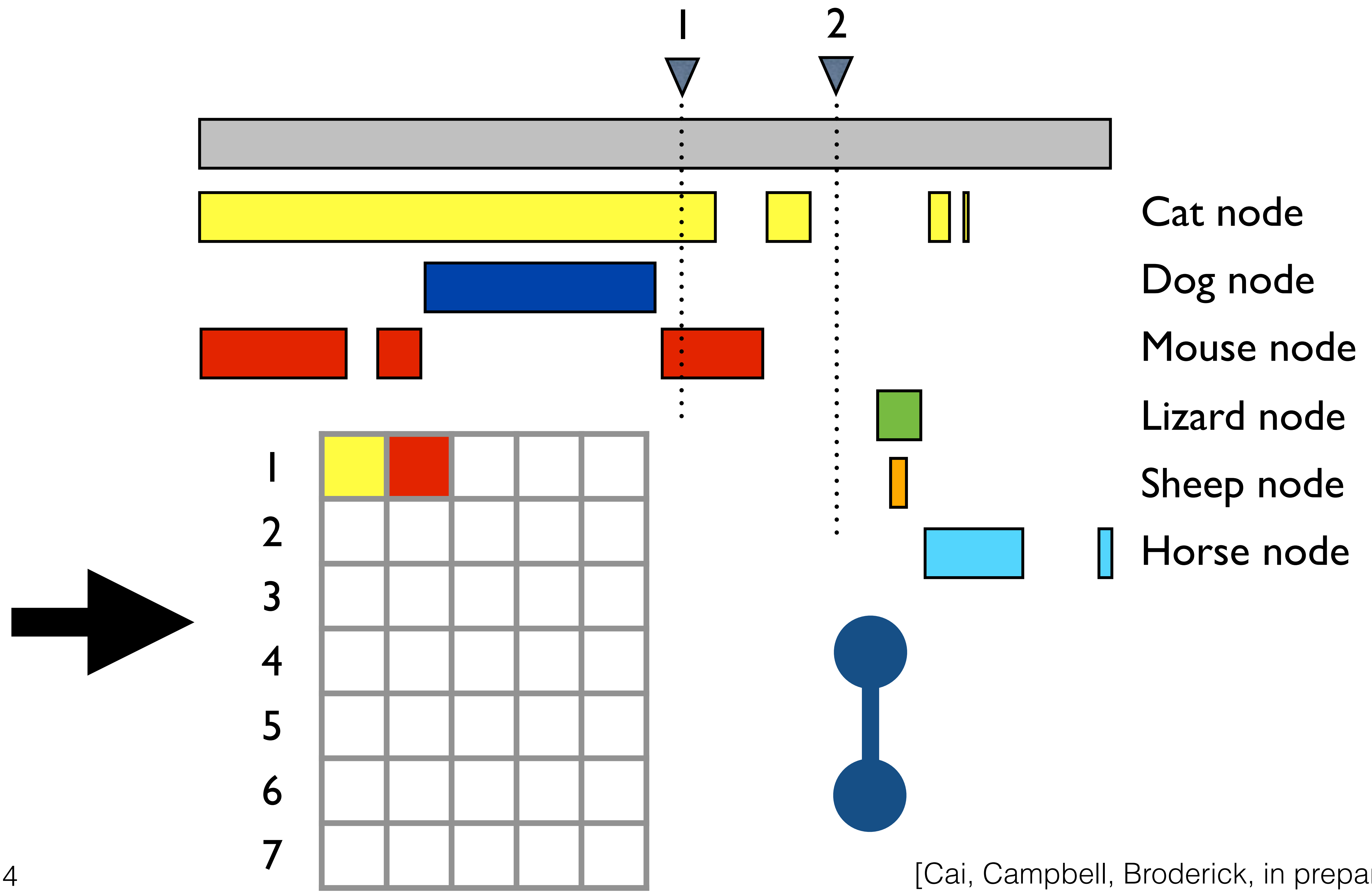
Edge-exchangeable graph



[Cai, Campbell, Broderick, in preparation]

Paintboxes

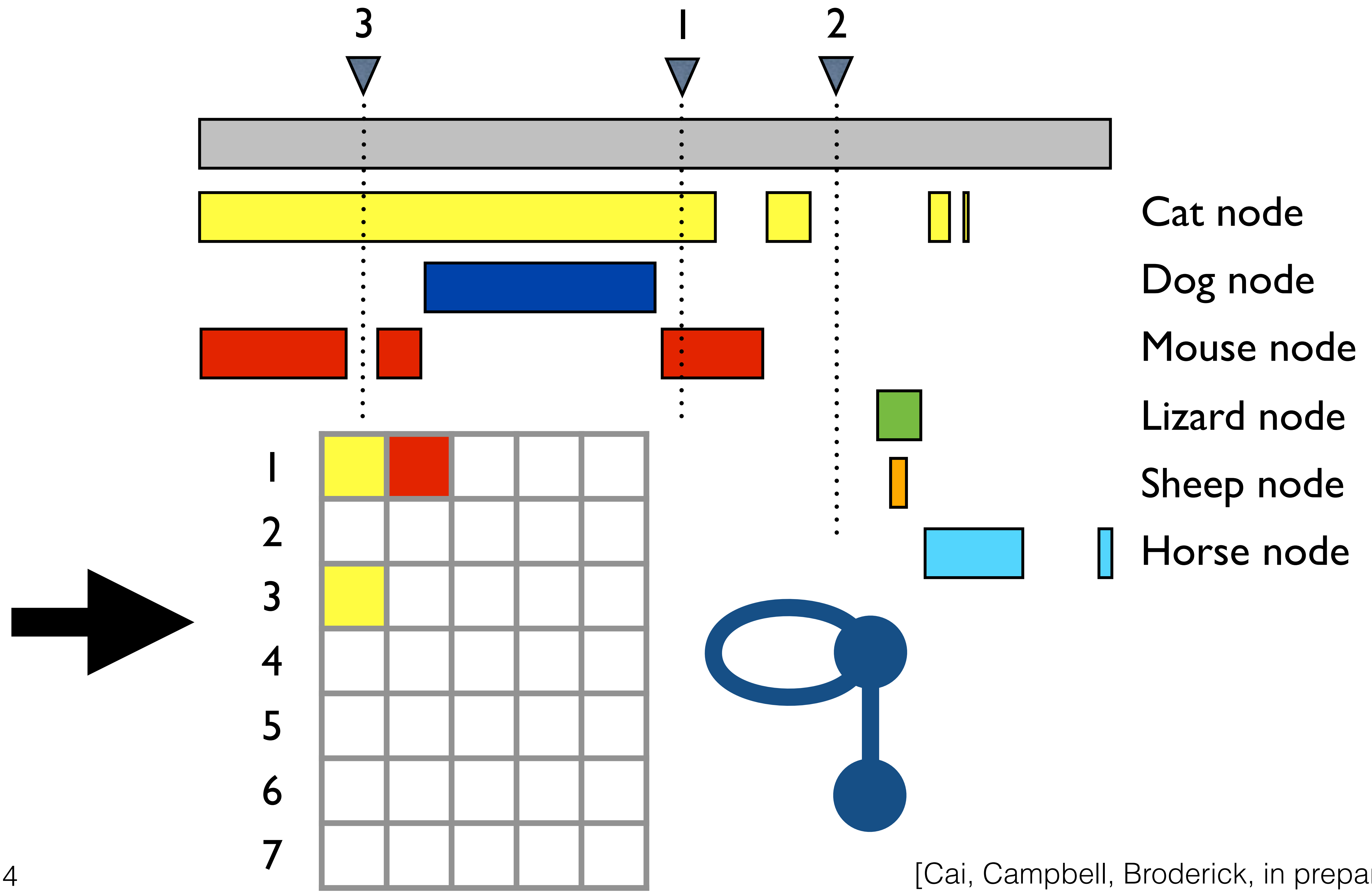
Edge-exchangeable graph



[Cai, Campbell, Broderick, in preparation]

Paintboxes

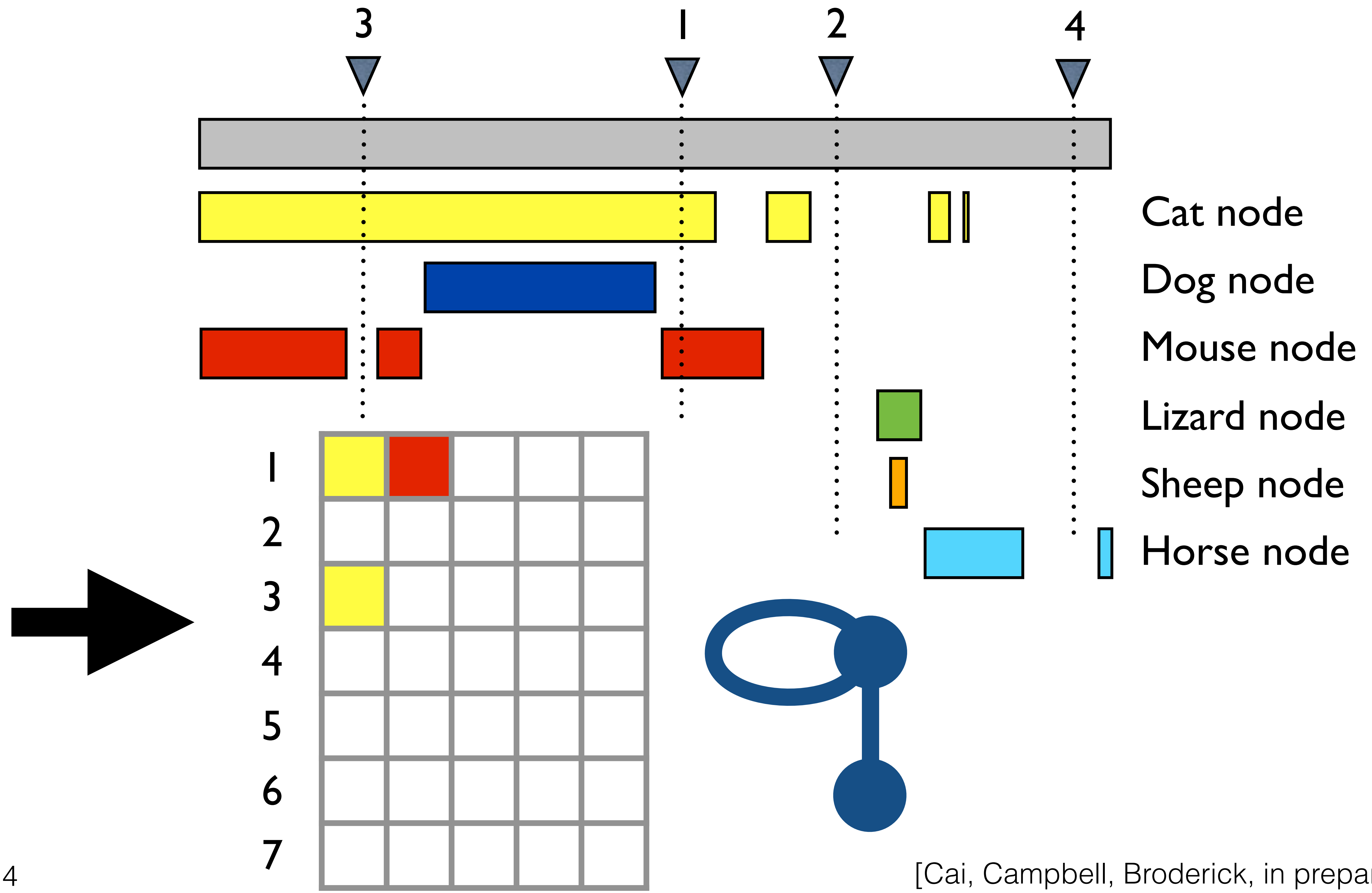
Edge-exchangeable graph



[Cai, Campbell, Broderick, in preparation]

Paintboxes

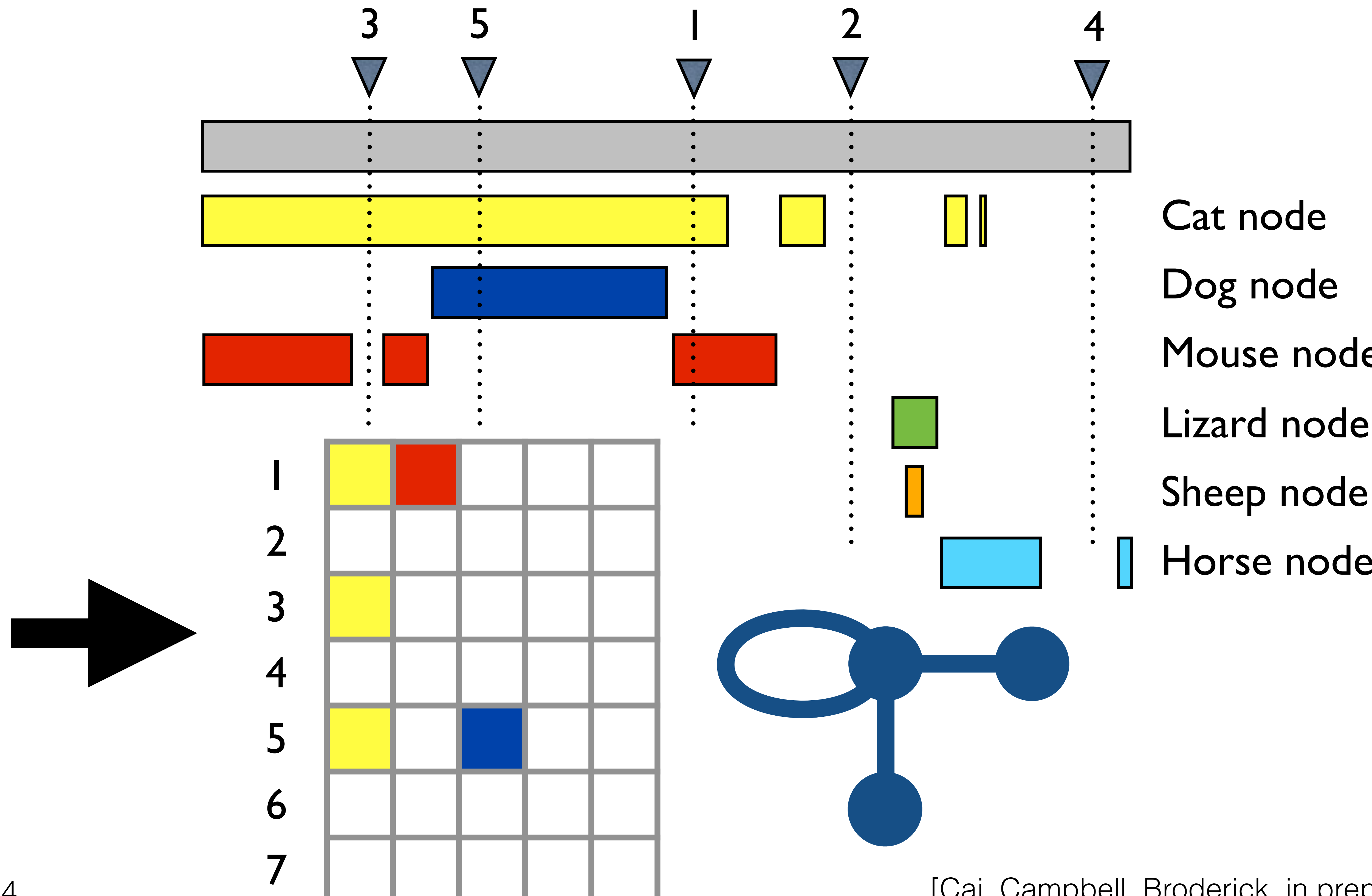
Edge-exchangeable graph



[Cai, Campbell, Broderick, in preparation]

Paintboxes

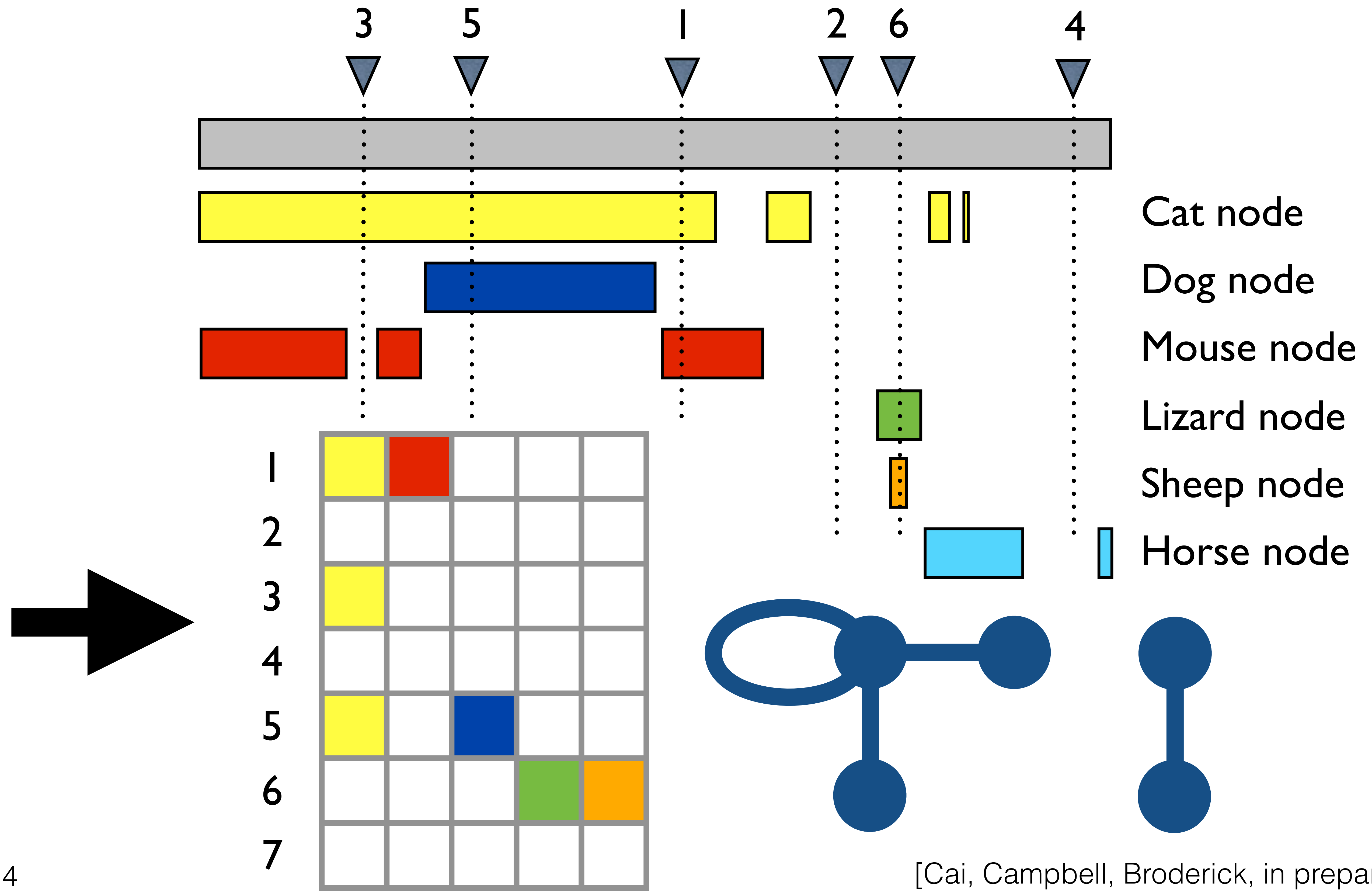
Edge-exchangeable graph



[Cai, Campbell, Broderick, in preparation]

Paintboxes

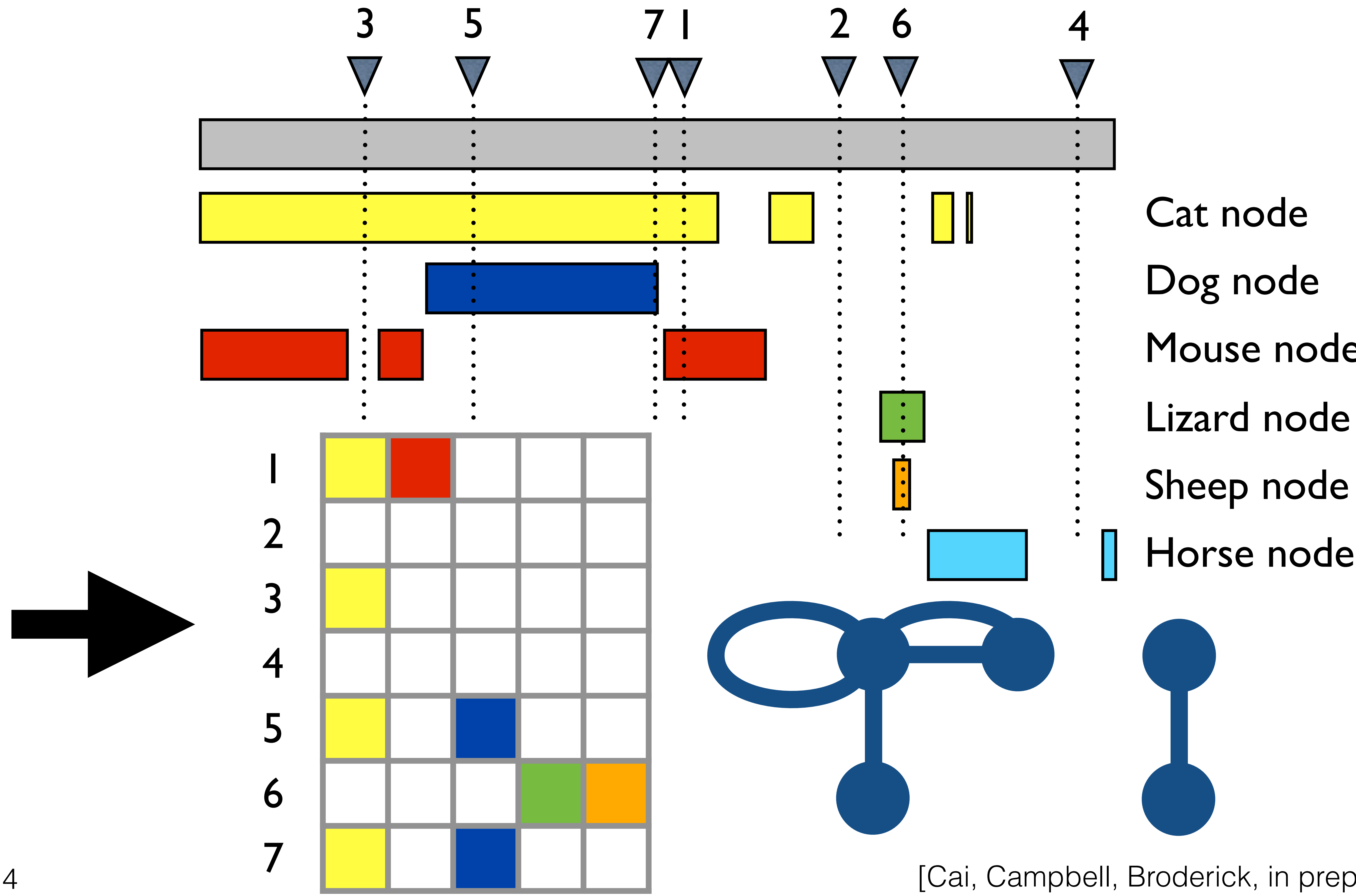
Edge-exchangeable graph



[Cai, Campbell, Broderick, in preparation]

Paintboxes

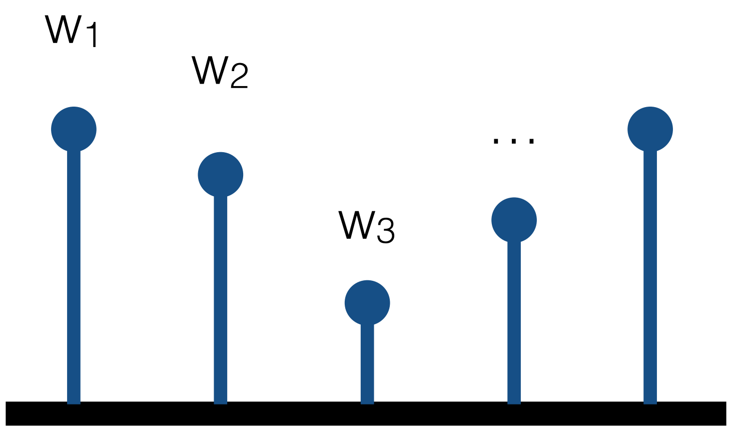
Edge-exchangeable graph



[Cai, Campbell, Broderick, in preparation]

Allows sparse graphs

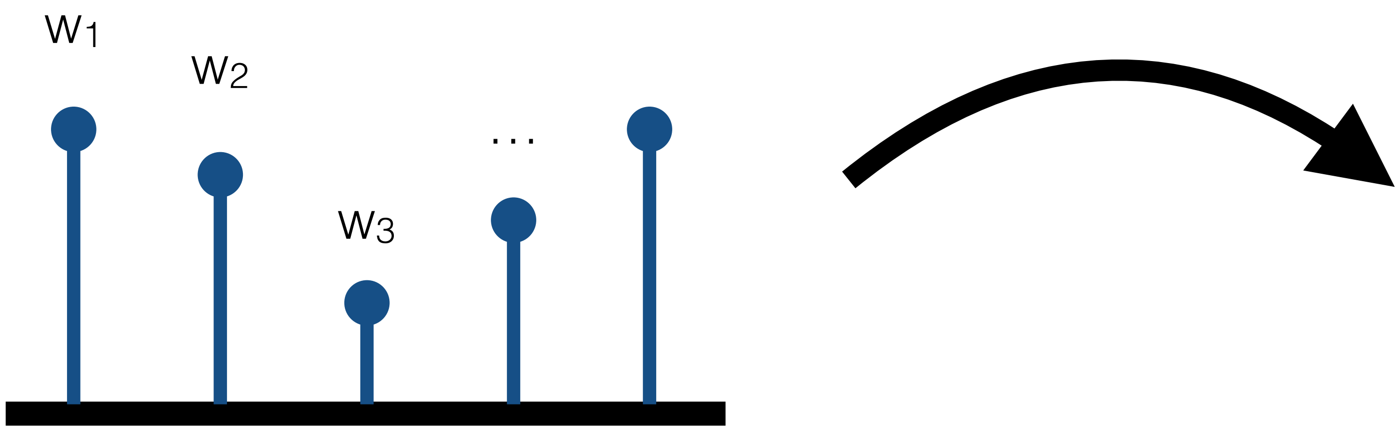
1. Draw discrete random measure $\sum_{j=1}^{\infty} w_j \delta_{\theta_j}$



[Cai, Campbell, Broderick, in preparation]

Allows sparse graphs

1. Draw discrete random measure $\sum_{j=1}^{\infty} w_j \delta_{\theta_j}$

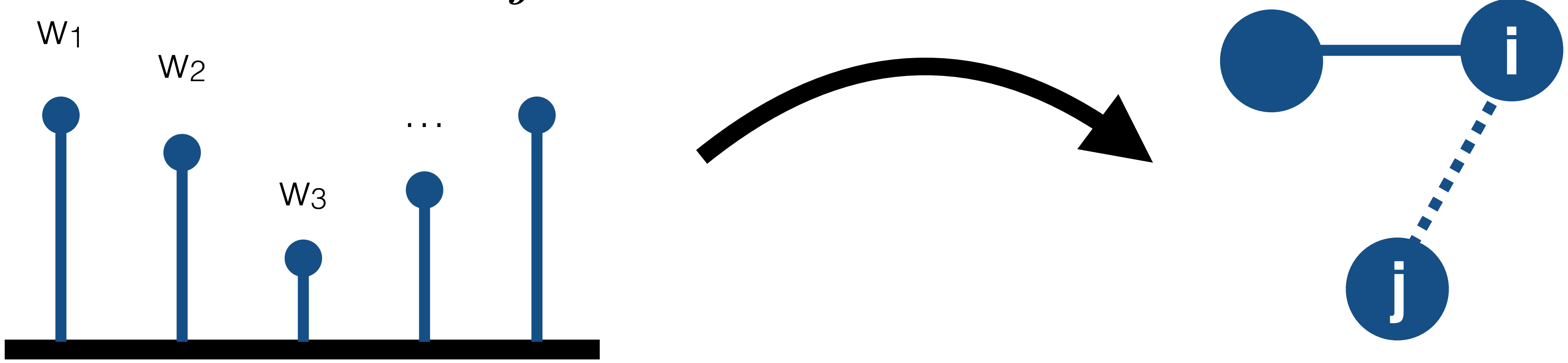


[Cai, Campbell, Broderick, in preparation]

Allows sparse graphs

1. Draw discrete random measure $\sum_{j=1}^{\infty} w_j \delta_{\theta_j}$

2. New edge {*i*,*j*} w.p. prop. to $w_i\, w_j$

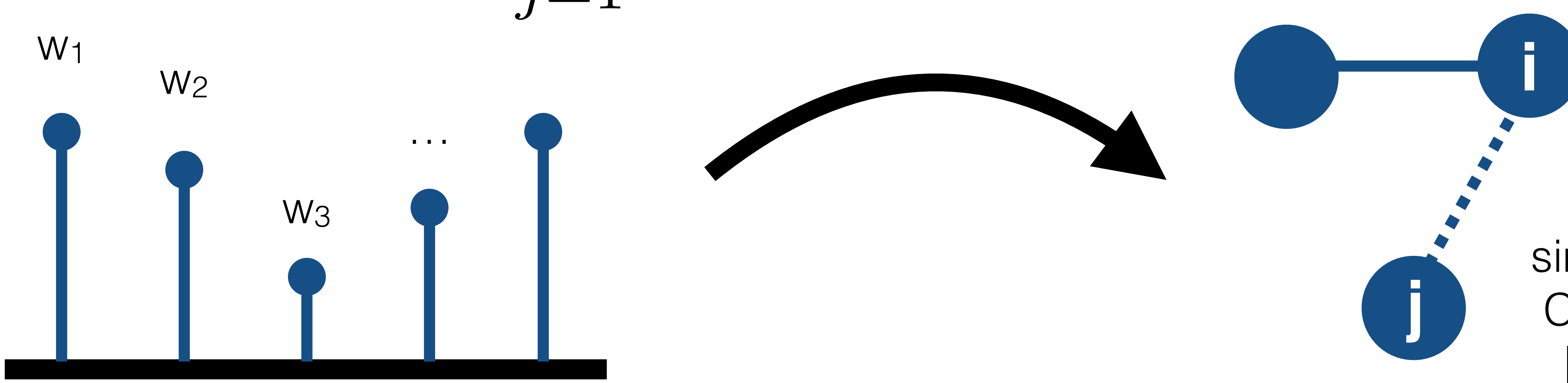


[Cai, Campbell, Broderick, in preparation]

Allows sparse graphs

1. Draw discrete random measure $\sum_{j=1}^{\infty} w_j \delta_{\theta_j}$

2. New edge $\{i,j\}$ w.p. prop. to $w_i\, w_j$



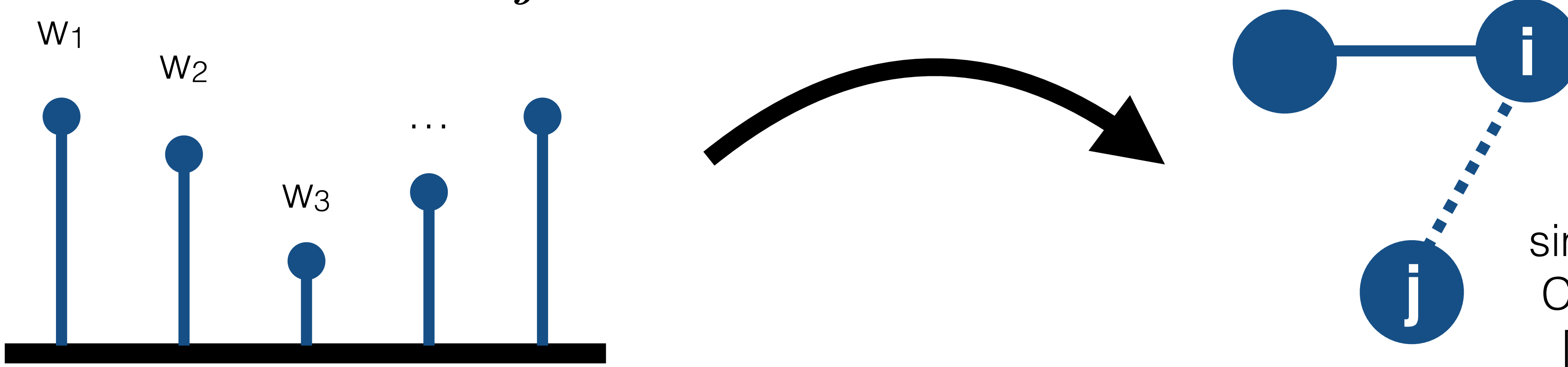
similar model using a CRM considered by [Caron, Fox 2015]

[Cai, Campbell, Broderick, in preparation]

Allows sparse graphs

1. Draw discrete random measure $\sum_{j=1}^{\infty} w_j \delta_{\theta_j}$

2. New edge $\{i,j\}$ w.p. prop. to $w_i\, w_j$



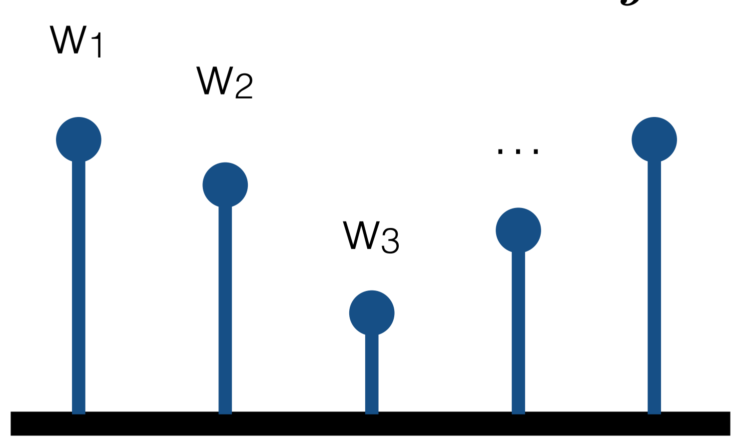
similar model using a CRM considered by [Caron, Fox 2015]

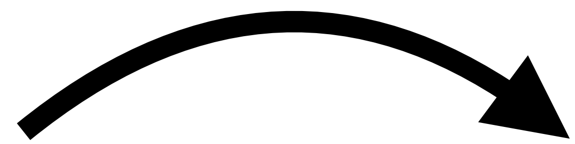
- "Graph frequency model"

[Cai, Campbell, Broderick, in preparation]

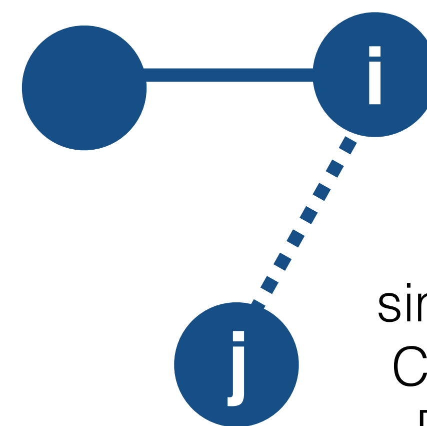
Allows sparse graphs

1. Draw discrete random measure $\sum_{j=1}^{\infty} w_j \delta_{\theta_j}$



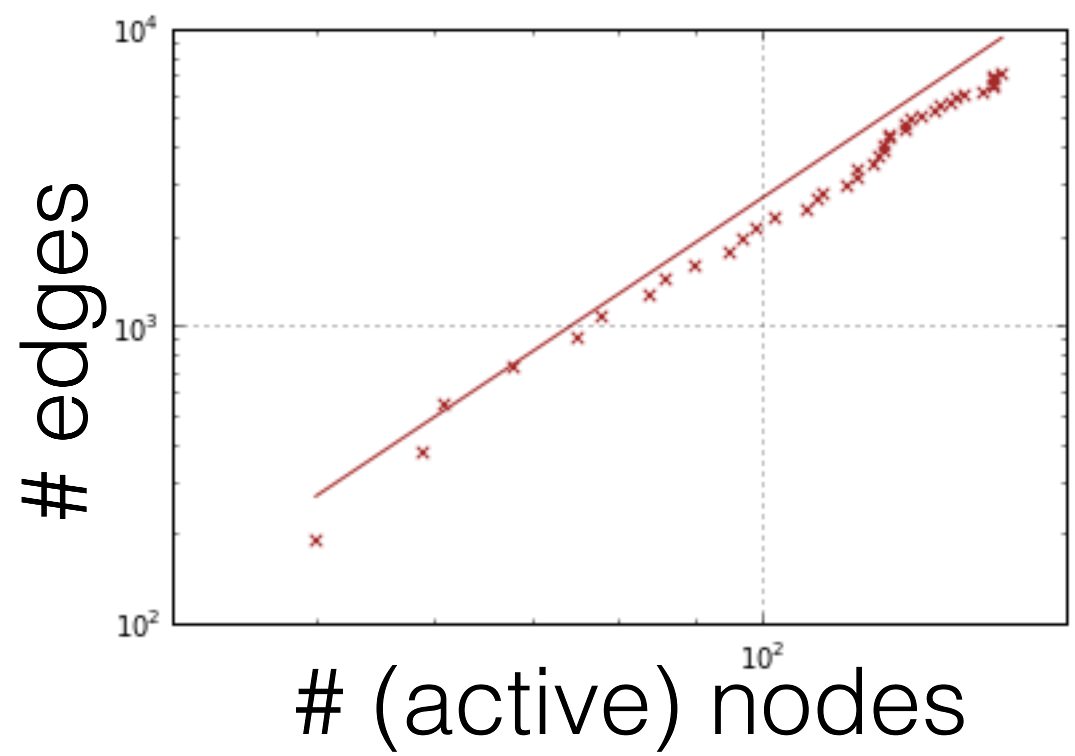


2. New edge $\{i,j\}$ w.p. prop. to $w_i\, w_j$



similar model using a CRM considered by [Caron, Fox 2015]

- "Graph frequency model"



[Cai, Campbell, Broderick, in preparation]

Conclusions

Conclusions

- Defined edge-exchangeable graph sequence

Conclusions

- Defined edge-exchangeable graph sequence
- Gave a de Finetti-style characterization

Conclusions

- Defined edge-exchangeable graph sequence
- Gave a de Finetti-style characterization
- Showed it allowed power laws

Conclusions

- Defined edge-exchangeable graph sequence
- Gave a de Finetti-style characterization
- Showed it allowed power laws
- Characterization of “dust”

Conclusions

- Defined edge-exchangeable graph sequence
- Gave a de Finetti-style characterization
- Showed it allowed power laws
- Characterization of “dust”

Up Next

Conclusions

- Defined edge-exchangeable graph sequence
- Gave a de Finetti-style characterization
- Showed it allowed power laws
- Characterization of “dust”

Up Next

- Characterize sparse, edge-exchangeable graphs

Conclusions

- Defined edge-exchangeable graph sequence
- Gave a de Finetti-style characterization
- Showed it allowed power laws
- Characterization of “dust”

Up Next

- Characterize sparse, edge-exchangeable graphs
- Characterize the different types of power laws (type I, type II, type III vs. nodes, edges, triangles)

Conclusions

- Defined edge-exchangeable graph sequence
- Gave a de Finetti-style characterization
- Showed it allowed power laws
- Characterization of “dust”

Up Next

- Characterize sparse, edge-exchangeable graphs
- Characterize the different types of power laws (type I, type II, type III vs. nodes, edges, triangles)
- Prove an EPPF-style representation

Conclusions

- Defined edge-exchangeable graph sequence
- Gave a de Finetti-style characterization
- Showed it allowed power laws
- Characterization of “dust”

Up Next

- Characterize sparse, edge-exchangeable graphs
- Characterize the different types of power laws (type I, type II, type III vs. nodes, edges, triangles)
- Prove an EPPF-style representation
- Truncation approximations

[Campbell, Huggins, How, Broderick 2016]

References

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

Broderick, T, and Cai, D. Edge-exchangeable graphs and sparsity. *NIPS 2015 Workshop on Networks in the Social and Information Sciences*. 2015.

Cai D, Broderick T. Completely random measures for modeling power laws in sparse graphs. *NIPS 2015 Workshop on Networks in the Social and Information Sciences*. 2015.

Campbell T*, Huggins JH*, How J, and Broderick T. Truncated random measures. Under review. Preprint on arXiv:1603.00861