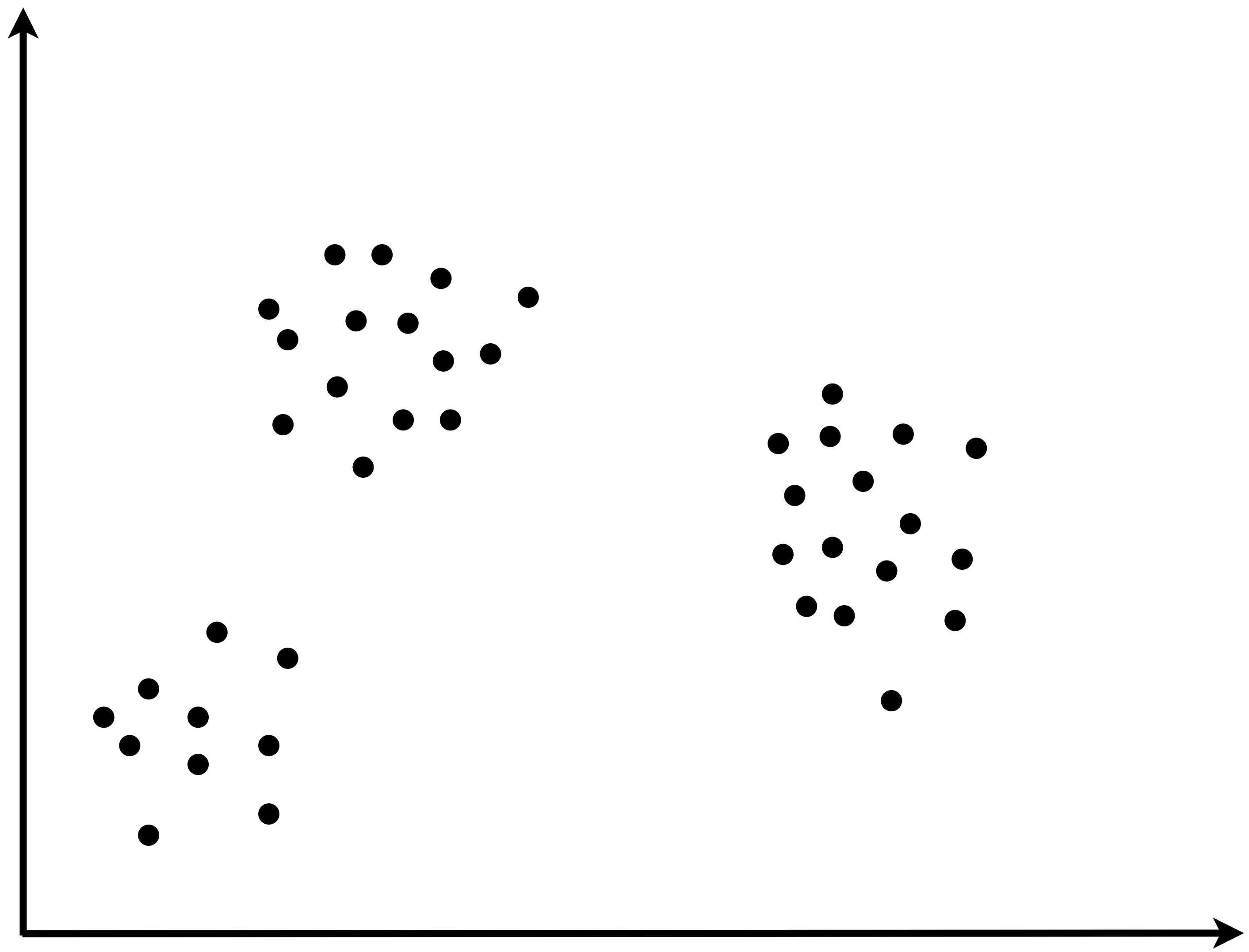
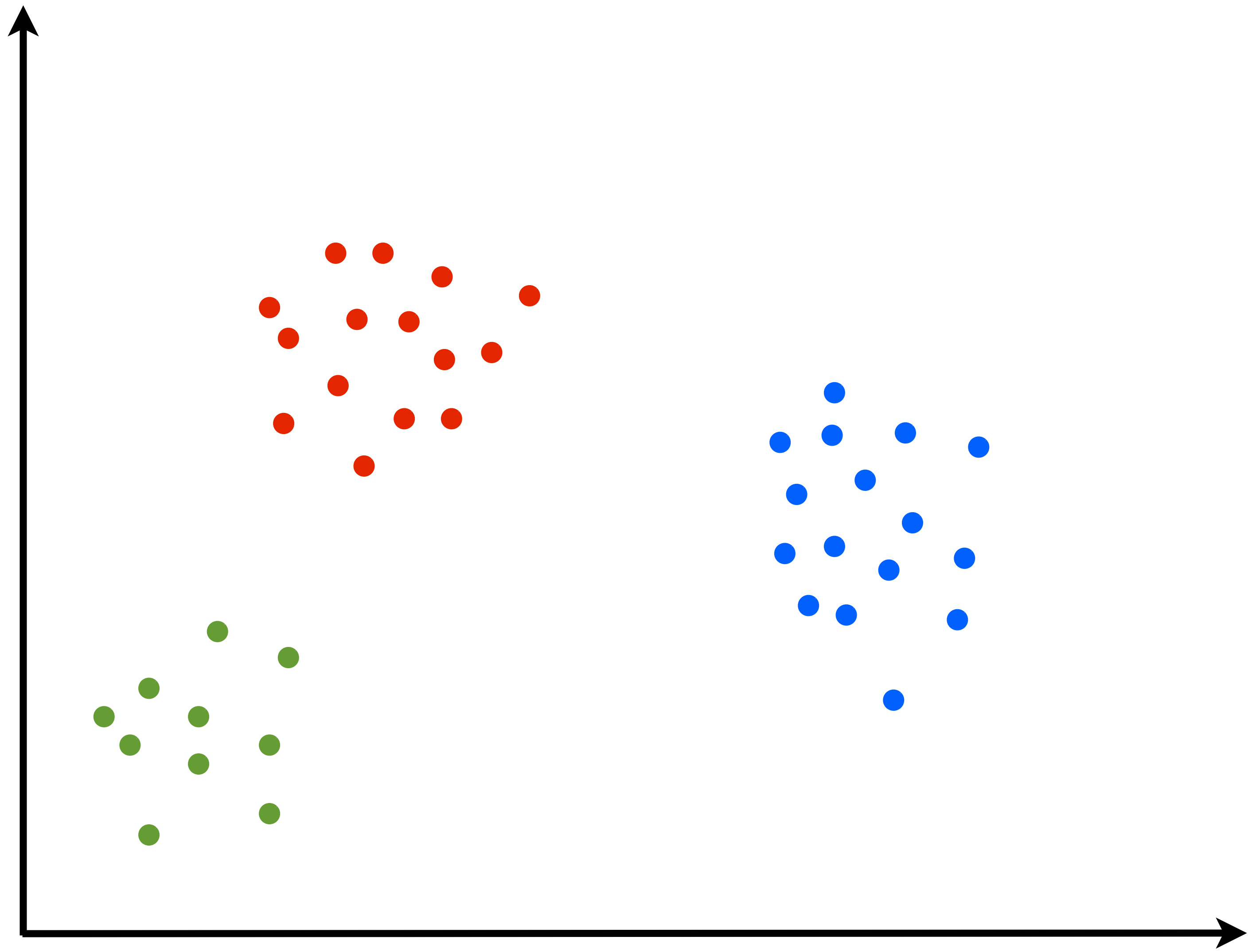
Feature allocations, probability functions, and paintboxes

Tamara Broderick, Jim Pitman, Michael I. Jordan
UC Berkeley

Clustering/Partition

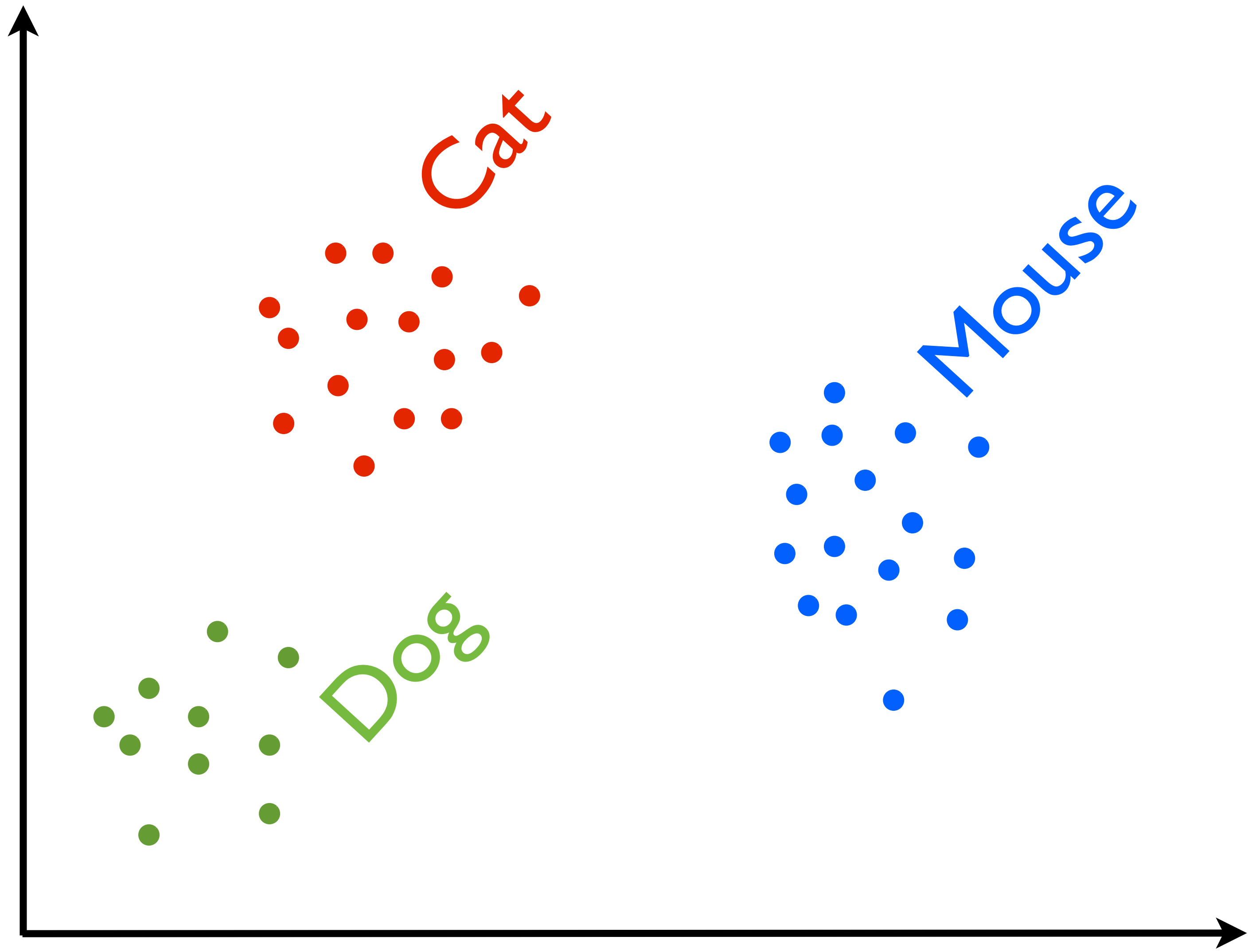


Clustering/Partition



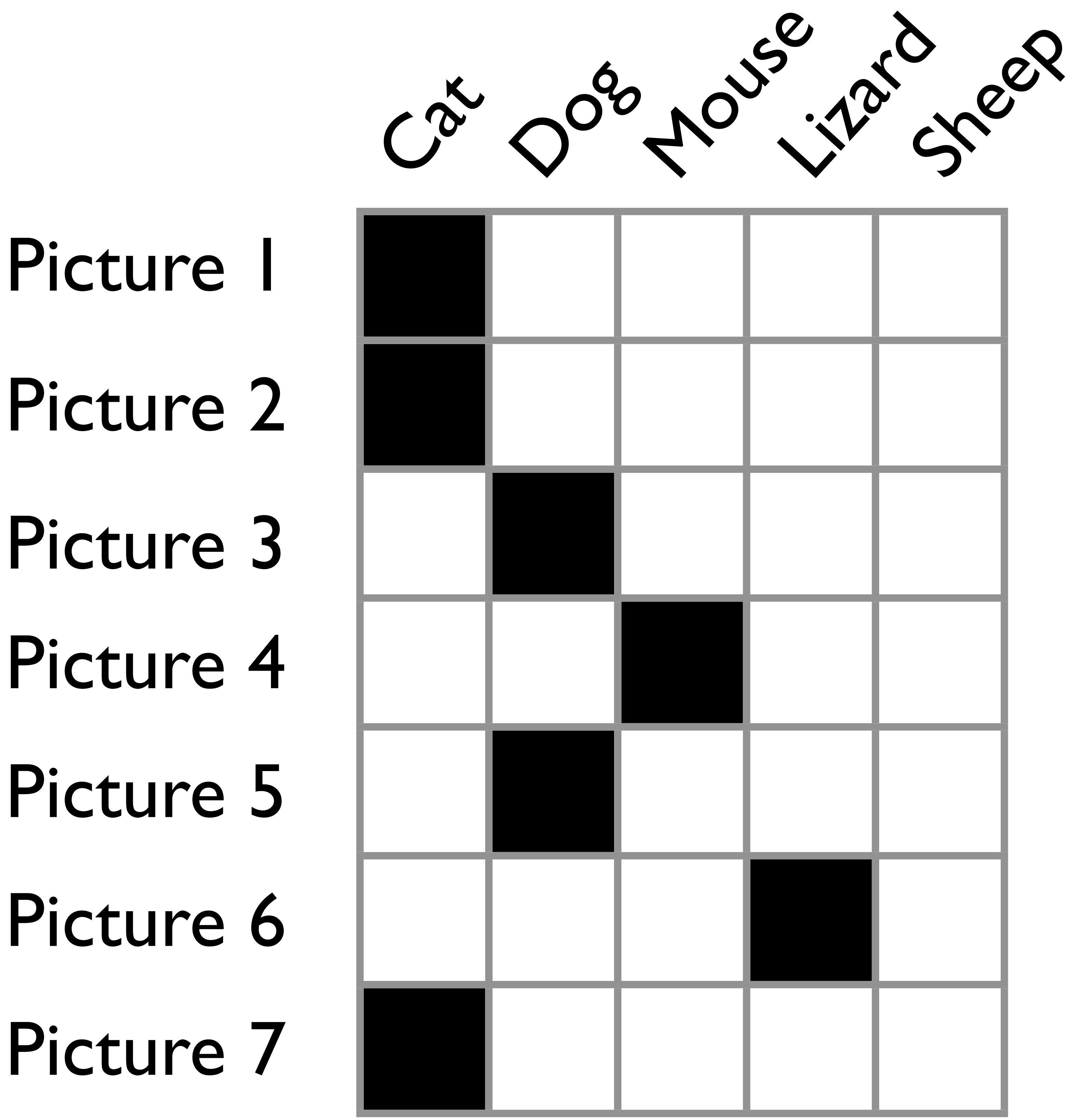
"clusters",
"classes",
"blocks (of a partition)"

Clustering/Partition

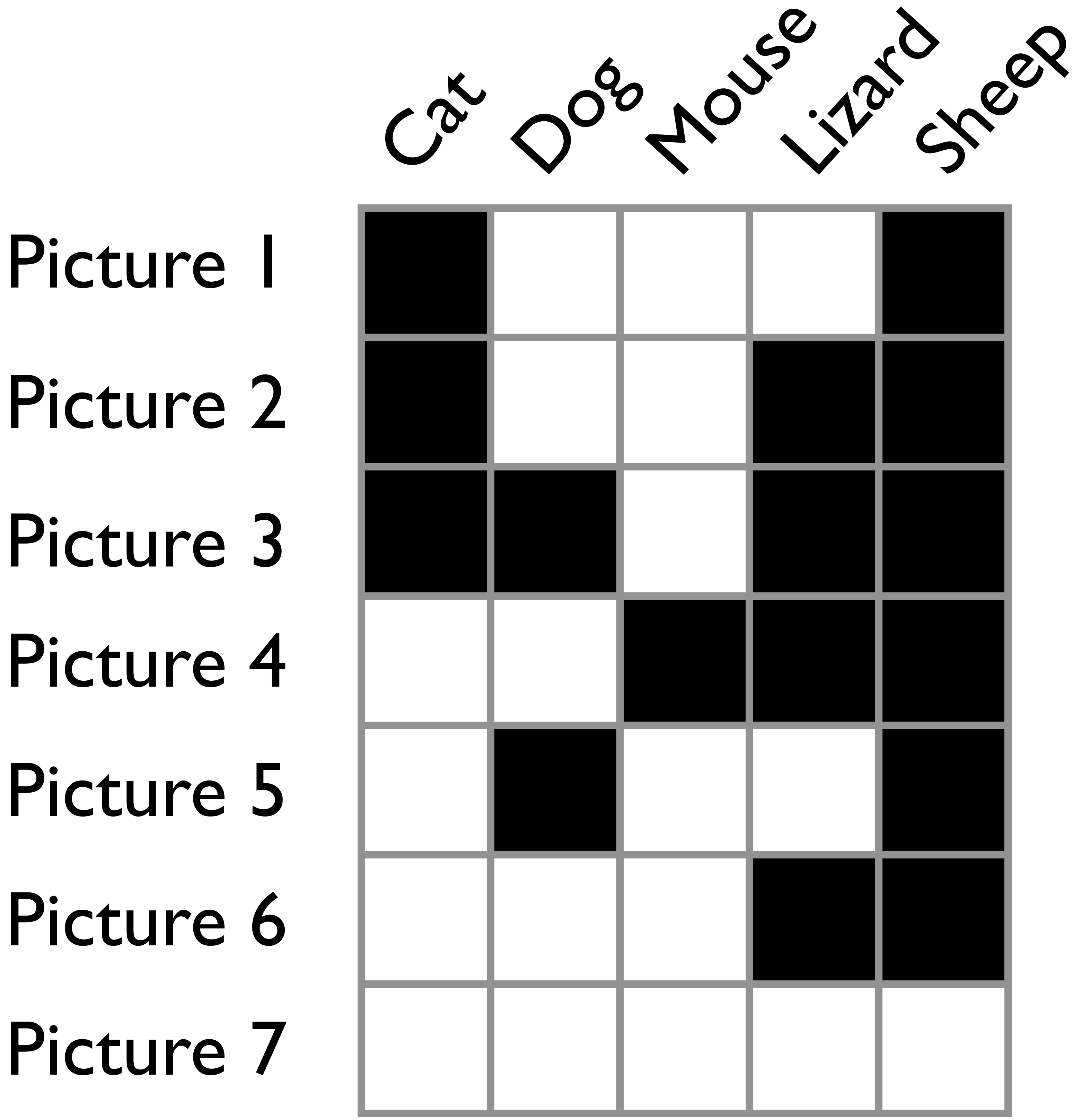


“clusters”,
“classes”,
“blocks (of a partition)”

Clustering/Partition



Latent feature allocation



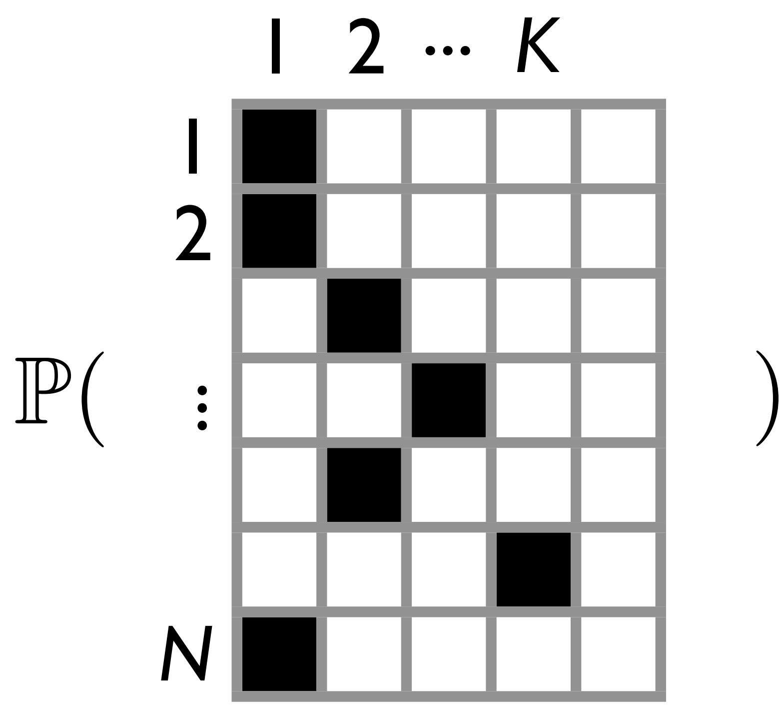
“features”,
“topics”

- Exchangeable
- Finite # of features per data point

Characterizations

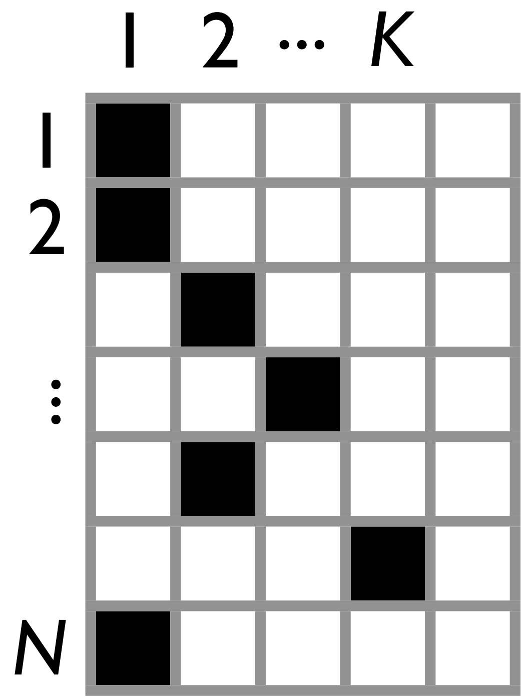
- Exchangeable cluster distributions are characterized
- What about exchangeable feature distributions?

Exchangeable probability functions



Exchangeable probability functions

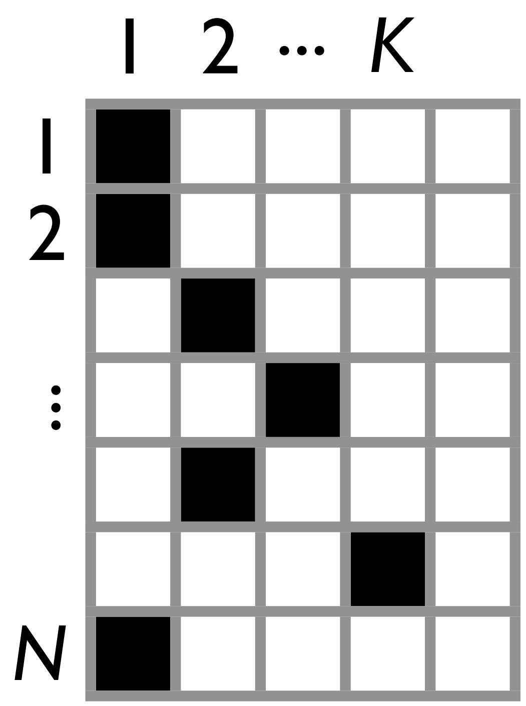
$\mathbb{P}($



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

Exchangeable probability functions

Size of *K*th cluster

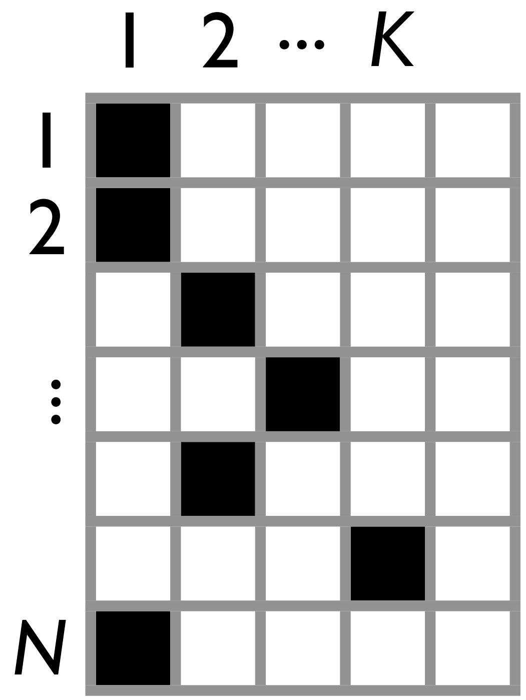
$\mathbb{P}($ 

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

Exchangeable probability functions

Exchangeable partition probability function (EPPF)

$\mathbb{P}($



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

[Pitman 1995]

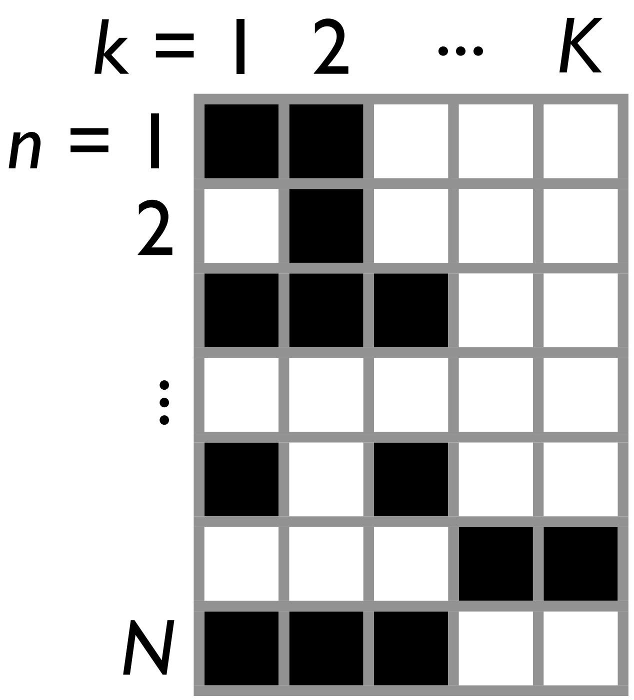
Exchangeable probability functions

"Exchangeable feature probability function" (EFPF)?

Example: Indian buffet process

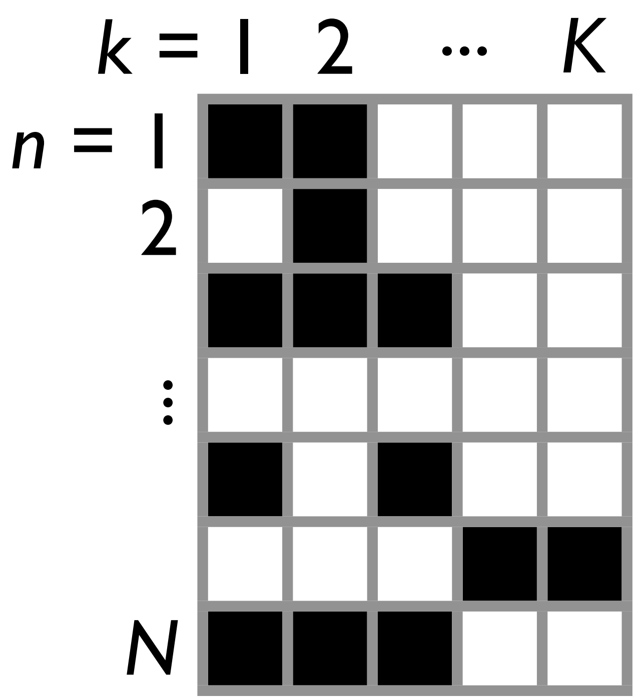
[Griffiths, Ghahramani 2006]

Example: Indian buffet process



[Griffiths, Ghahramani 2006]

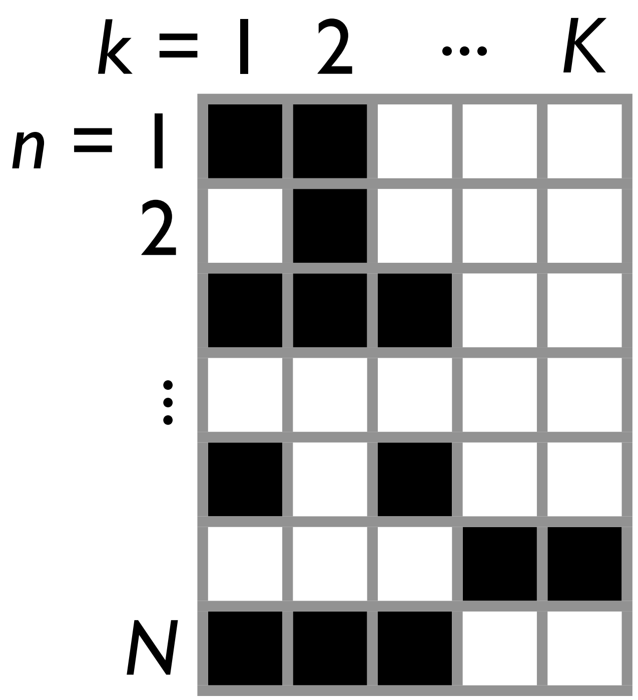
Example: Indian buffet process



For n = 1, 2, ..., N

[Griffiths, Ghahramani 2006]

Example: Indian buffet process

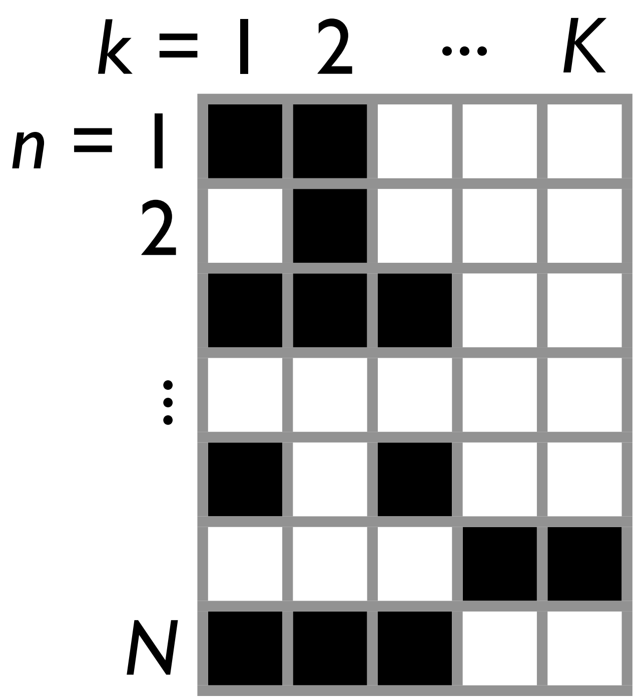


For n = 1, 2, ..., N

1. Data point n has an existing feature k that has already occurred $S_{n-1,k}$ times with probability $\frac{S_{n-1,k}}{\theta + n - 1}$

[Griffiths, Ghahramani 2006]

Example: Indian buffet process

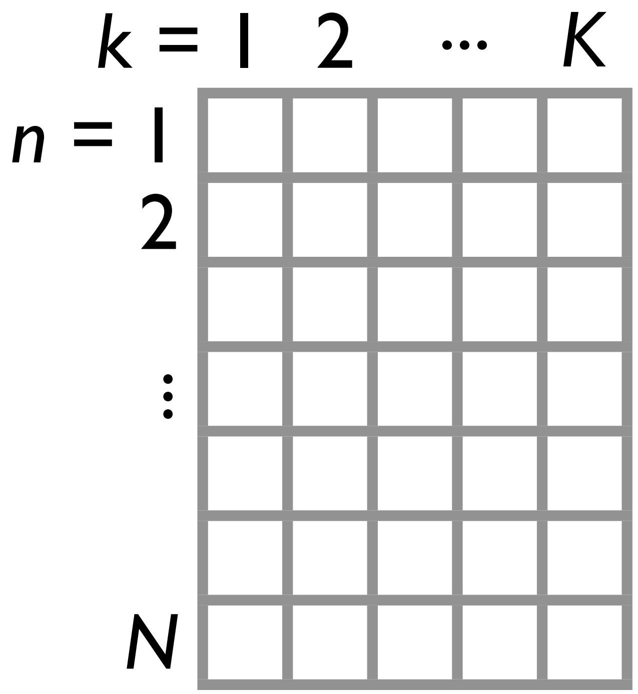


For n = 1, 2, ..., N

1. Data point n has an existing feature k that has already occurred $S_{n-1,k}$ times with probability $\dfrac{S_{n-1,k}}{\theta + n - 1}$
2. Number of new features for data point n: $K_n^+ = \mathrm{Poisson}\left(\gamma \dfrac{\theta}{\theta + n - 1}\right)$

[Griffiths, Ghahramani 2006]

Example: Indian buffet process

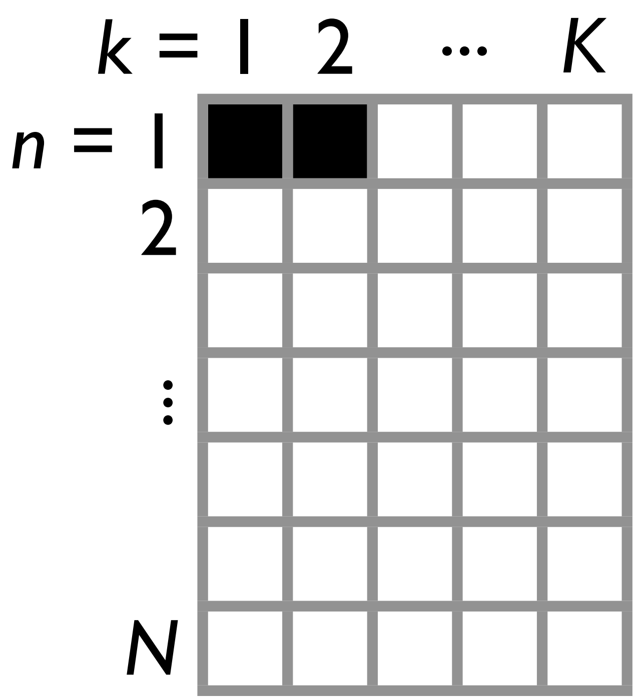


For n = 1, 2, ..., N

1. Data point n has an existing feature k that has already occurred $S_{n-1,k}$ times with probability $$\frac{S_{n-1,k}}{\theta + n - 1}$$
2. Number of new features for data point n: $$K_n^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

[Griffiths, Ghahramani 2006]

Example: Indian buffet process



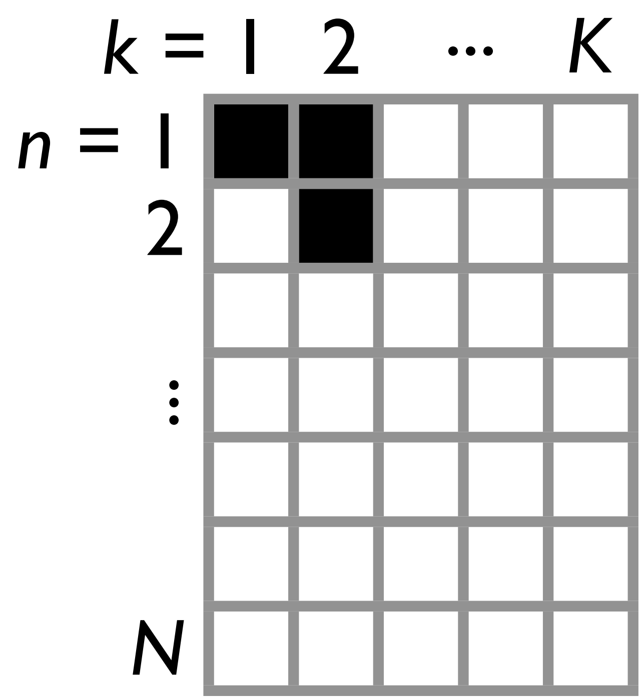
For n = 1, 2, ..., N

1. Data point n has an existing feature k that has already occurred $S_{n-1,k}$ times with probability $\dfrac{S_{n-1,k}}{\theta + n - 1}$

2. Number of new features for data point n: $K_n^+ = \mathrm{Poisson}\left(\gamma \dfrac{\theta}{\theta + n - 1}\right)$

[Griffiths, Ghahramani 2006]

Example: Indian buffet process



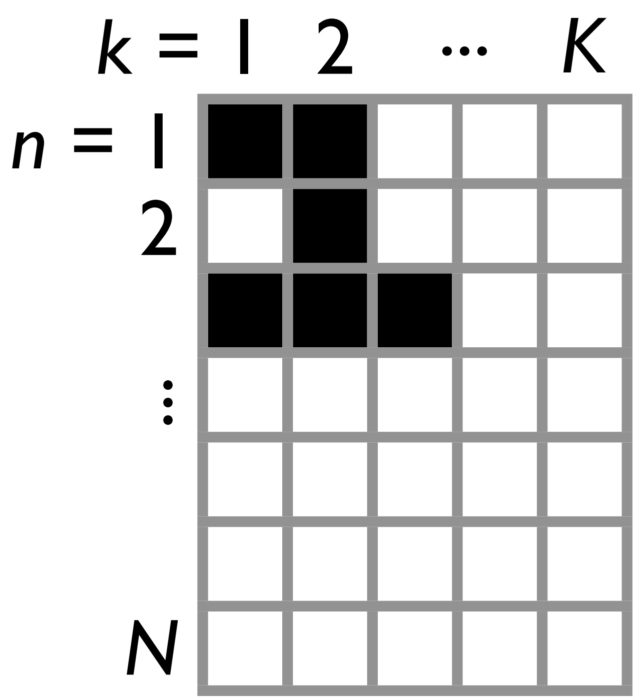
For n = 1, 2, ..., N

1. Data point n has an existing feature k that has already occurred $S_{n-1,k}$ times with probability $\dfrac{S_{n-1,k}}{\theta + n - 1}$

2. Number of new features for data point n: $K_n^+ = \mathrm{Poisson}\left(\gamma \dfrac{\theta}{\theta + n - 1}\right)$

[Griffiths, Ghahramani 2006]

Example: Indian buffet process



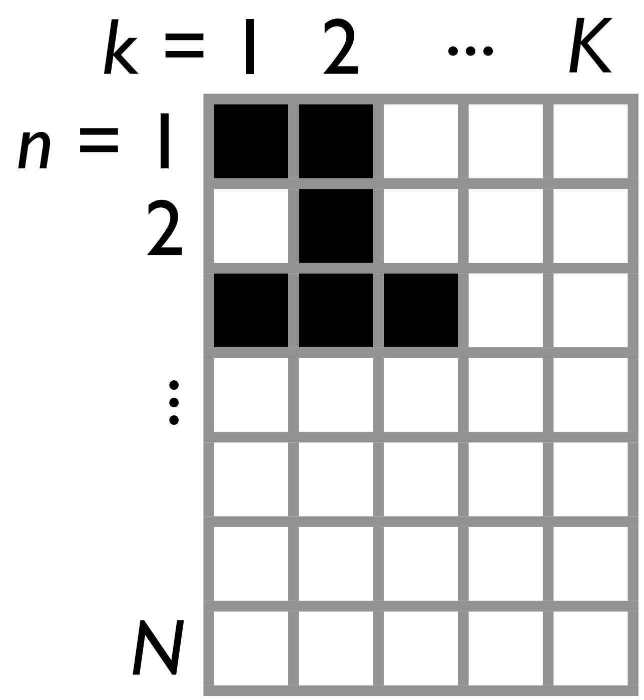
For n = 1, 2, ..., N

1. Data point n has an existing feature k that has already occurred $S_{n-1,k}$ times with probability $\dfrac{S_{n-1,k}}{\theta + n - 1}$

2. Number of new features for data point n: $K_n^+ = \mathrm{Poisson}\left(\gamma \dfrac{\theta}{\theta + n - 1}\right)$

[Griffiths, Ghahramani 2006]

Example: Indian buffet process

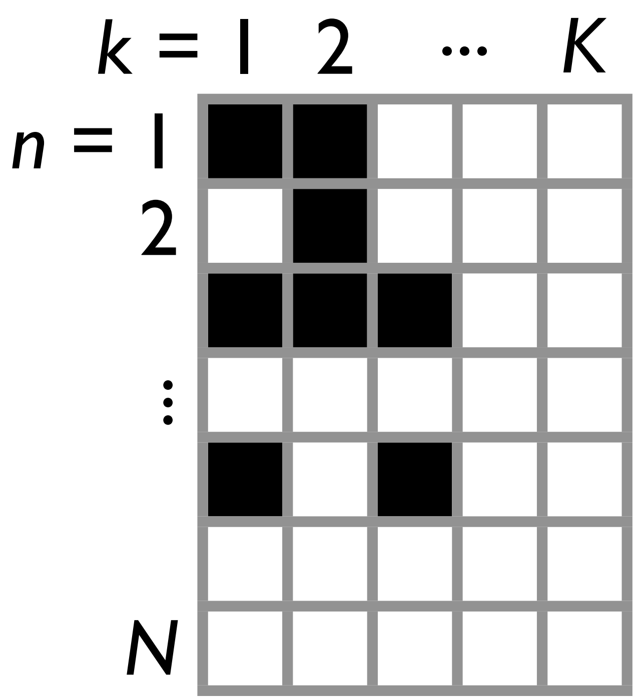


For n = 1, 2, ..., N

1. Data point n has an existing feature k that has already occurred $S_{n-1,k}$ times with probability $\dfrac{S_{n-1,k}}{\theta + n - 1}$
2. Number of new features for data point n: $K_n^+ = \text{Poisson}\left(\gamma \dfrac{\theta}{\theta + n - 1}\right)$

[Griffiths, Ghahramani 2006]

Example: Indian buffet process

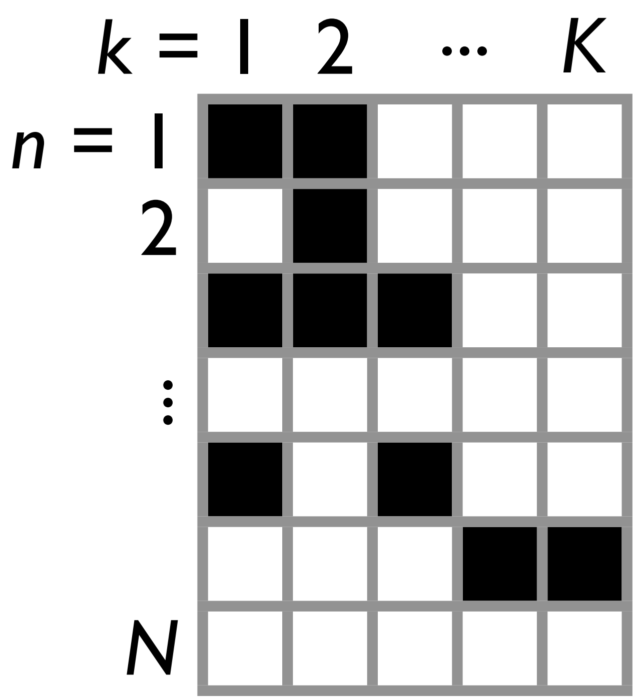


For n = 1, 2, ..., N

1. Data point n has an existing feature k that has already occurred $S_{n-1,k}$ times with probability $\dfrac{S_{n-1,k}}{\theta + n - 1}$
2. Number of new features for data point n: $K_n^+ = \text{Poisson}\left(\gamma \dfrac{\theta}{\theta + n - 1}\right)$

[Griffiths, Ghahramani 2006]

Example: Indian buffet process



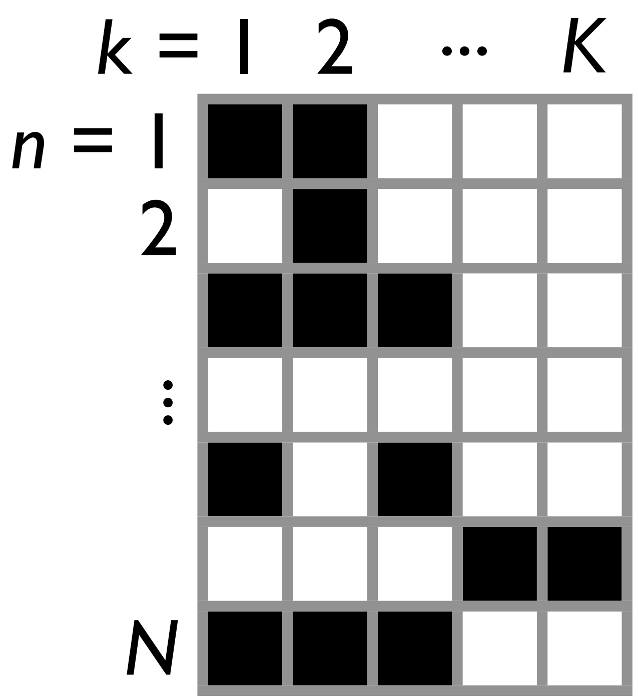
For n = 1, 2, ..., N

1. Data point n has an existing feature k that has already occurred $S_{n-1,k}$ times with probability $\dfrac{S_{n-1,k}}{\theta + n - 1}$

2. Number of new features for data point n: $K_n^+ = \mathrm{Poisson}\left(\gamma \dfrac{\theta}{\theta + n - 1}\right)$

[Griffiths, Ghahramani 2006]

Example: Indian buffet process



For n = 1, 2, ..., N

1. Data point n has an existing feature k that has already occurred $S_{n-1,k}$ times with probability $\dfrac{S_{n-1,k}}{\theta + n - 1}$
2. Number of new features for data point n: $K_n^+ = \mathrm{Poisson}\left(\gamma \dfrac{\theta}{\theta + n - 1}\right)$

[Griffiths, Ghahramani 2006]

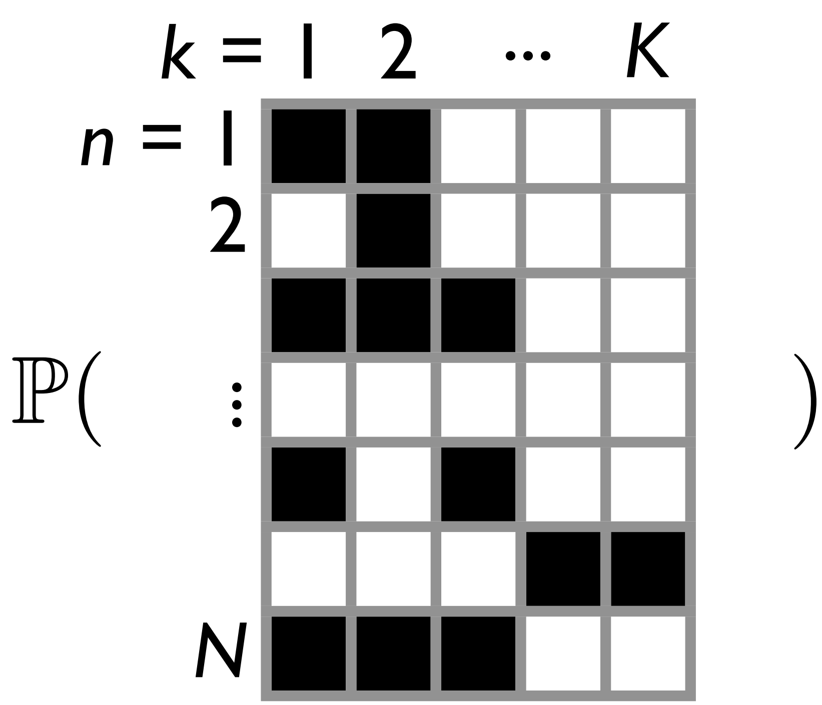
Exchangeable probability functions

"Exchangeable feature probability function" (EFPF)?

Exchangeable probability functions

"Exchangeable feature probability function" (EFPF)?

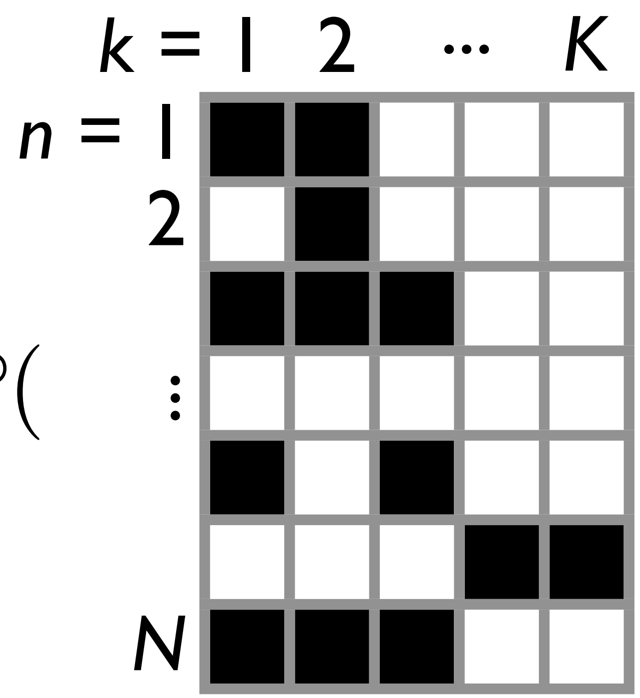
Example: Indian buffet process (IBP)



Exchangeable probability functions

"Exchangeable feature probability function" (EFPF)?

Example: Indian buffet process (IBP)

$\mathbb{P}($


$)$

$$= \frac{1}{K_N!}(\theta\gamma)^{K_N} \exp\left(-\theta\gamma \sum_{n=1}^{N} (\theta + n - 1)^{-1}\right) \prod_{k=1}^{K_N} \frac{\Gamma(S_{N,k})\Gamma(N - S_{N,k} + \theta)}{\Gamma(N + \theta)}$$

Exchangeable probability functions

"Exchangeable feature probability function" (EFPF)?

Example: Indian buffet process (IBP)

$$\mathbb{P}\left(\begin{array}{r|ccccc} & k=1 & 2 & & \cdots & K \\ \hline n=1 & \blacksquare & \blacksquare & \square & \square & \square \\ 2 & \square & \blacksquare & \square & \square & \square \\ & \blacksquare & \blacksquare & \blacksquare & \square & \square \\ \vdots & \square & \square & \square & \square & \square \\ & \blacksquare & \square & \blacksquare & \square & \square \\ & \square & \square & \square & \blacksquare & \blacksquare \\ N & \blacksquare & \blacksquare & \blacksquare & \square & \square \end{array}\right)$$

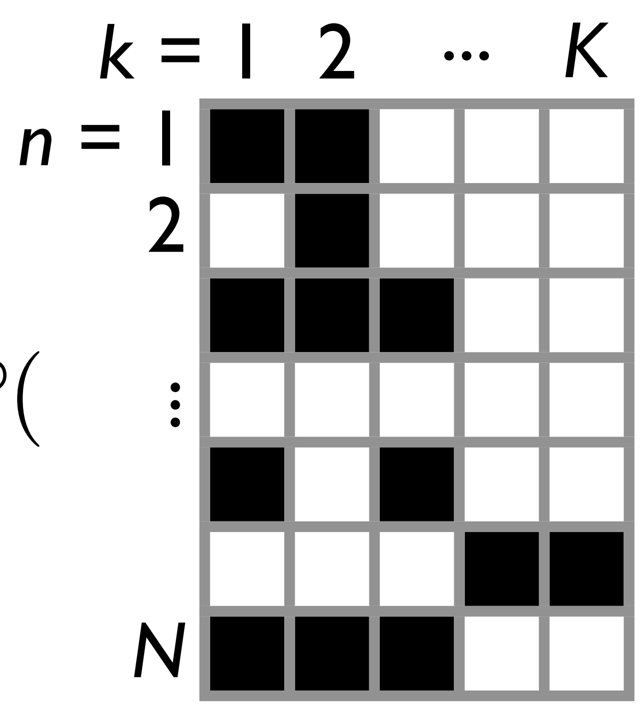
Size of *k*th feature

$$= \frac{1}{K_N!}(\theta\gamma)^{K_N} \exp\left(-\theta\gamma \sum_{n=1}^{N} (\theta + n - 1)^{-1}\right) \prod_{k=1}^{K_N} \frac{\Gamma(S_{N,k})\Gamma(N - S_{N,k} + \theta)}{\Gamma(N + \theta)}$$

Exchangeable probability functions

"Exchangeable feature probability function" (EFPF)?

Example: Indian buffet process (IBP)

$\mathbb{P}($  $)$

Size of *k*th feature

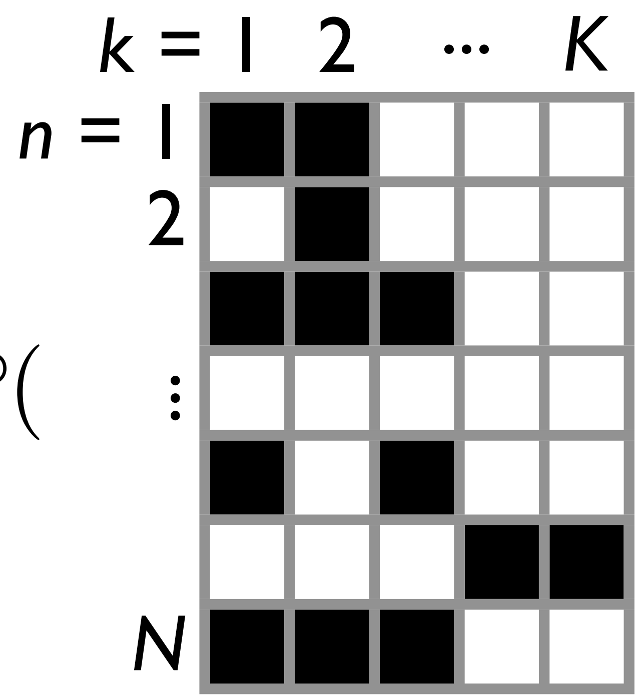
Number of features

$$= \frac{1}{K_N!}(\theta\gamma)^{K_N} \exp\left(-\theta\gamma \sum_{n=1}^{N} (\theta + n - 1)^{-1}\right) \prod_{k=1}^{K_N} \frac{\Gamma(S_{N,k})\Gamma(N - S_{N,k} + \theta)}{\Gamma(N + \theta)}$$

Exchangeable probability functions

"Exchangeable feature probability function" (EFPF)?

Example: Indian buffet process (IBP)

$\mathbb{P}($  $)$

Number of data points

Size of *k*th feature

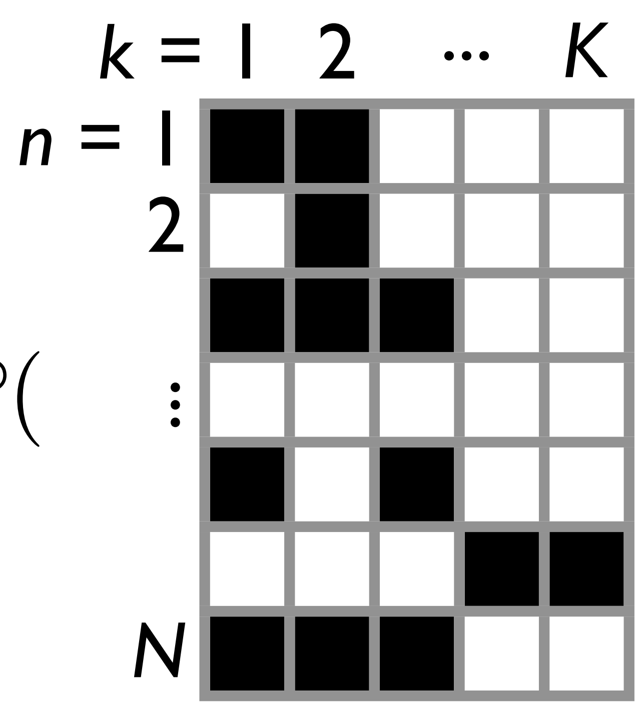
Number of features

$$= \frac{1}{K_N!}(\theta\gamma)^{K_N} \exp\left(-\theta\gamma \sum_{n=1}^{N} (\theta + n - 1)^{-1}\right) \prod_{k=1}^{K_N} \frac{\Gamma(S_{N,k})\Gamma(N - S_{N,k} + \theta)}{\Gamma(N+\theta)}$$

Exchangeable probability functions

"Exchangeable feature probability function" (EFPF)?

Example: Indian buffet process (IBP)

$\mathbb{P}($  $)$

Number of data points

Size of *k*th feature

Number of features

$$= \frac{1}{K_N!}(\theta\gamma)^{K_N} \exp\left(-\theta\gamma \sum_{n=1}^{N} (\theta + n - 1)^{-1}\right) \prod_{k=1}^{K_N} \frac{\Gamma(S_{N,k})\Gamma(N - S_{N,k} + \theta)}{\Gamma(N+\theta)}$$

$$= p(N; S_{N,1}, S_{N,2}, \ldots, S_{N,K})$$

[Broderick, Jordan, Pitman 2012]

Exchangeable probability functions

"Exchangeable feature probability function" (EFPF)?

Example: Indian buffet process (IBP)

$$
\mathbb{P}\left(
\begin{array}{c|ccccc}
 & k=1 & 2 & & \cdots & K \\
\hline
n=1 & 1 & 1 & 0 & 0 & 0 \\
2 & 0 & 1 & 0 & 0 & 0 \\
 & 1 & 1 & 1 & 0 & 0 \\
\vdots & 0 & 0 & 0 & 0 & 0 \\
 & 1 & 0 & 1 & 0 & 0 \\
 & 0 & 0 & 0 & 1 & 1 \\
N & 1 & 1 & 1 & 0 & 0
\end{array}
\right)
$$

Number of data points

Size of kth feature

Number of features

$$= \frac{1}{K_N!}(\theta\gamma)^{K_N} \exp\left(-\theta\gamma\sum_{n=1}^{N}(\theta+n-1)^{-1}\right)\prod_{k=1}^{K_N}\frac{\Gamma(S_{N,k})\Gamma(N-S_{N,k}+\theta)}{\Gamma(N+\theta)}$$

$$= p(N; S_{N,1}, S_{N,2}, \ldots, S_{N,K})$$

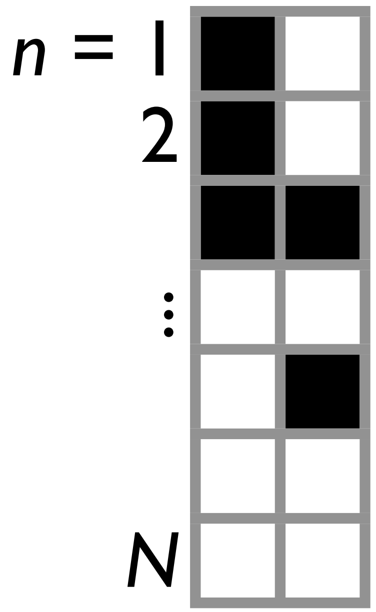
"EFPF"

[Broderick, Jordan, Pitman 2012]

Exchangeable probability functions

"Exchangeable feature probability function" (EFPF)?

Counterexample

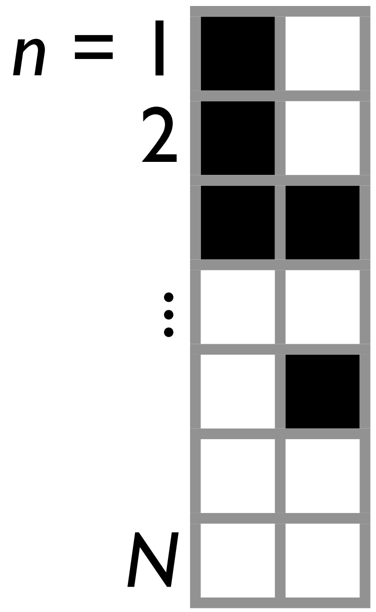


[Broderick, Jordan, Pitman 2012]

Exchangeable probability functions

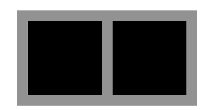
"Exchangeable feature probability function" (EFPF)?

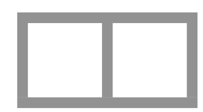
Counterexample



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

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

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

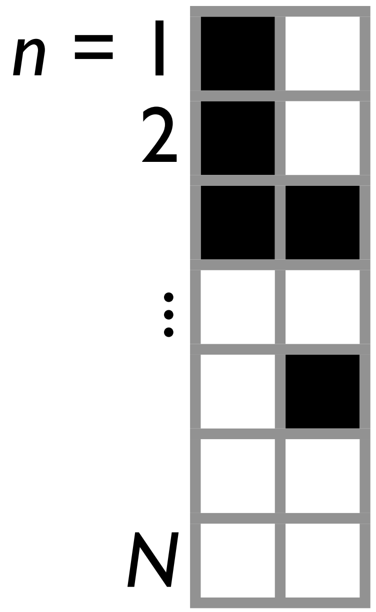
$\mathbb{P}(\text{row} =$  $) = p_4$

[Broderick, Jordan, Pitman 2012]

Exchangeable probability functions

"Exchangeable feature probability function" (EFPF)?

Counterexample

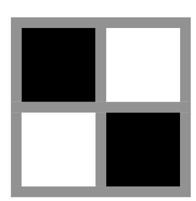


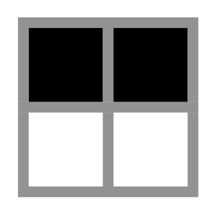
$$\mathbb{P}(\text{row} = \blacksquare\square) = p_1$$

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

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

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



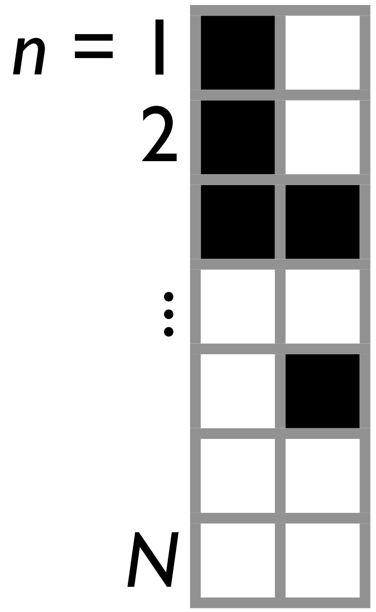


[Broderick, Jordan, Pitman 2012]

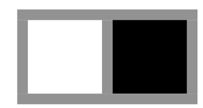
Exchangeable probability functions

"Exchangeable feature probability function" (EFPF)?

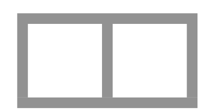
Counterexample

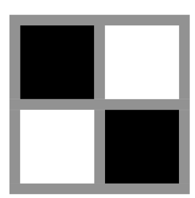
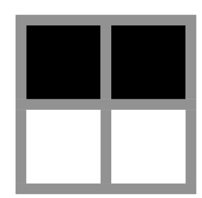


$\mathbb{P}(\mathrm{row} =$  $) = p_1$

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

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

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

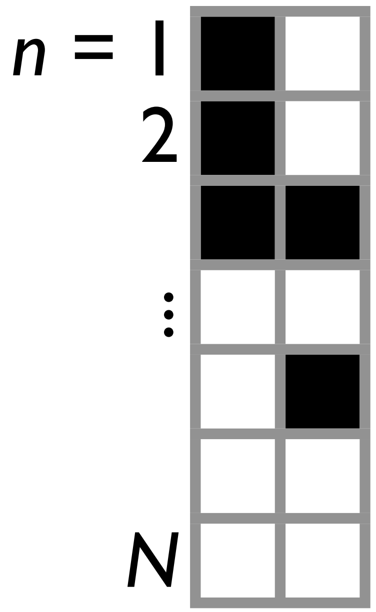
$\mathbb{P}($  $)$ $\mathbb{P}($  $)$

[Broderick, Jordan, Pitman 2012]

Exchangeable probability functions

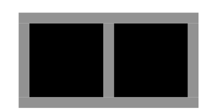
"Exchangeable feature probability function" (EFPF)?

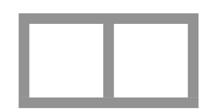
Counterexample

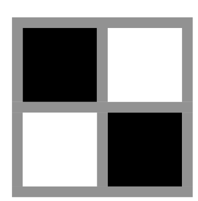
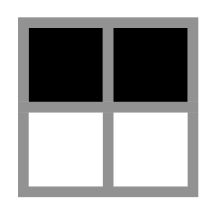


$\mathbb{P}(\text{row} =$  $) = p_1$

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

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

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

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

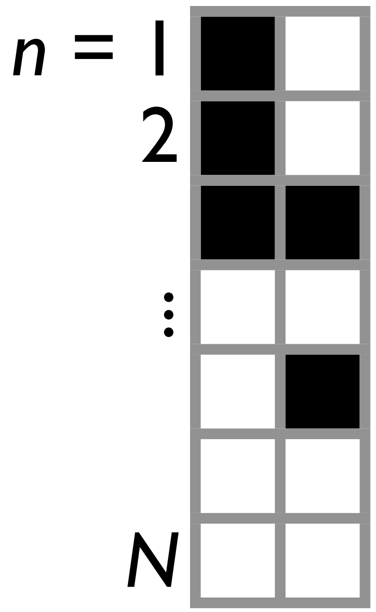
$p_1 p_2$ $p_3 p_4$

[Broderick, Jordan, Pitman 2012]

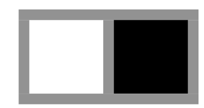
Exchangeable probability functions

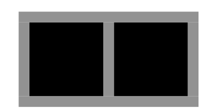
"Exchangeable feature probability function" (EFPF)?

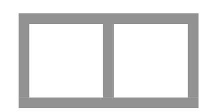
Counterexample

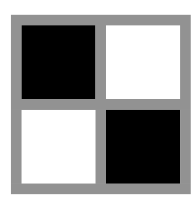
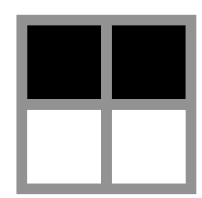


ℙ(row = ) = p_1

ℙ(row = ) = p_2

ℙ(row = ) = p_3

ℙ(row = ) = p_4

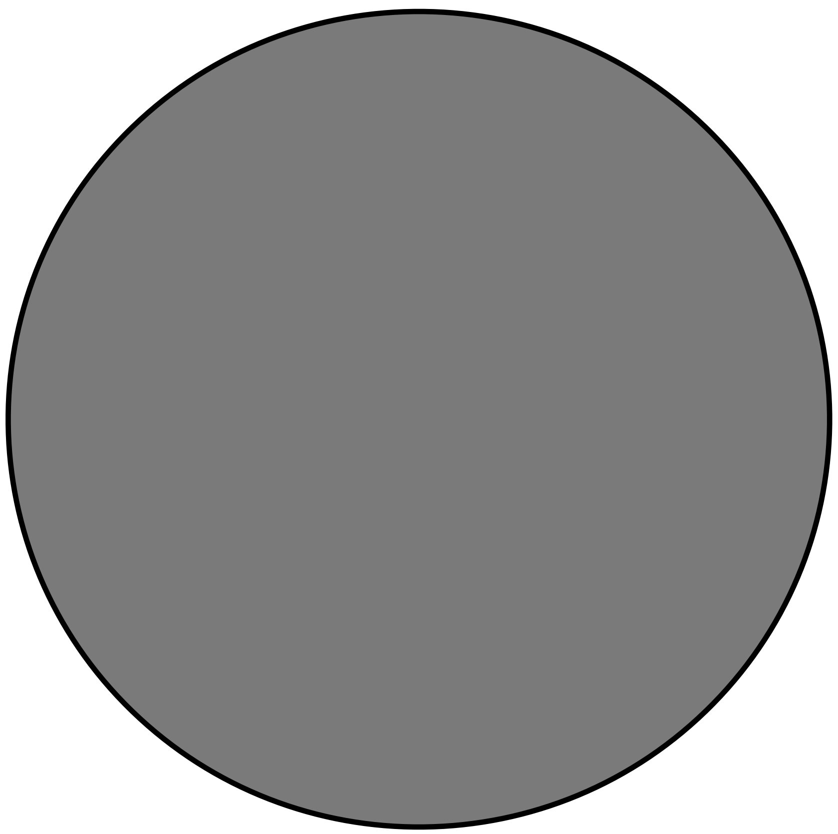
ℙ() ≠ ℙ()

$$p_1 p_2 \neq p_3 p_4$$

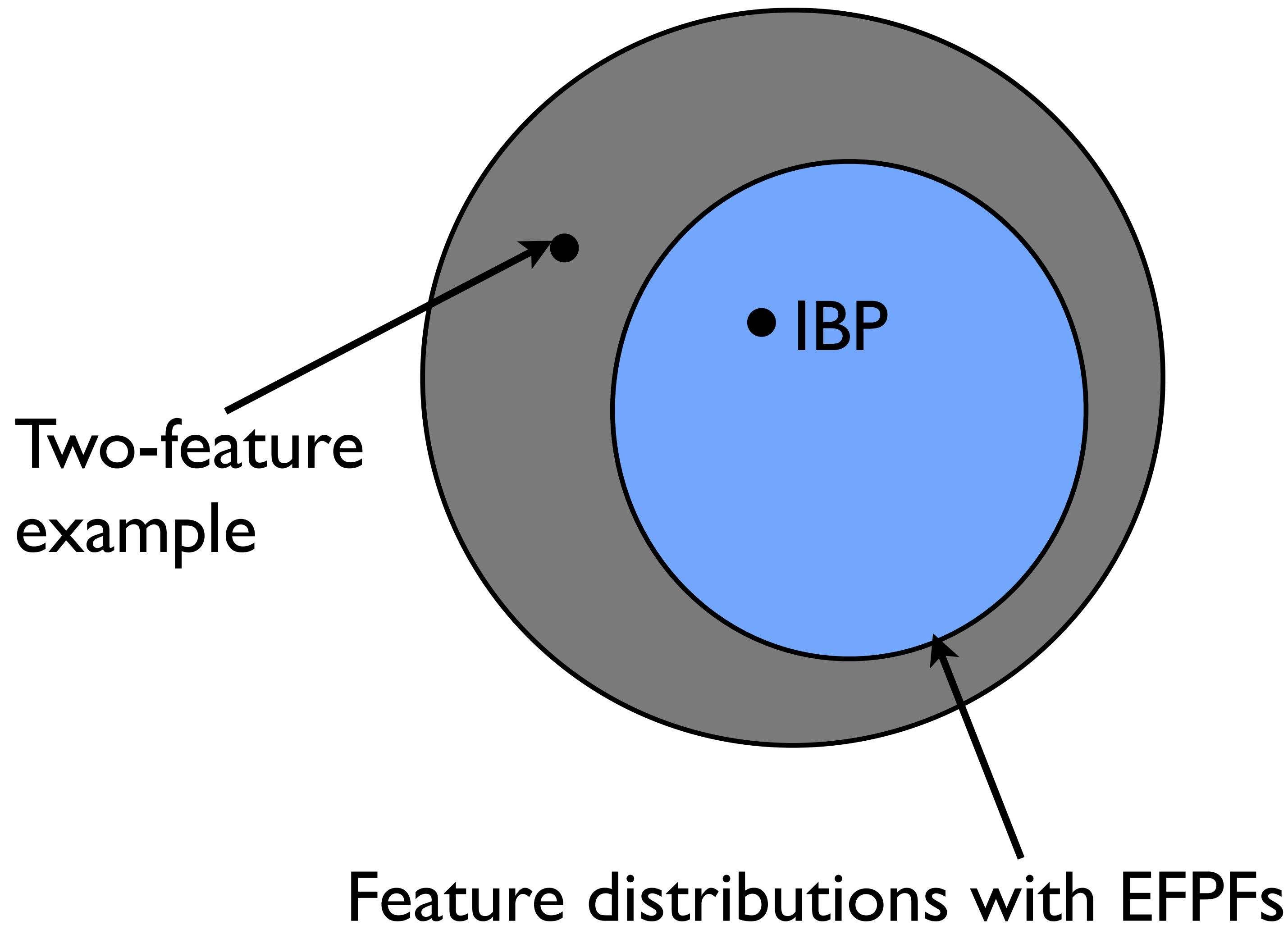
[Broderick, Jordan, Pitman 2012]

Exchangeable probability functions

Exchangeable cluster distributions
= Cluster distributions with EPPFs



Exchangeable feature distributions



[Broderick, Jordan, Pitman 2012]

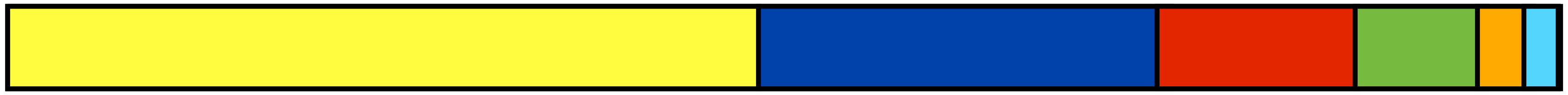
Paintboxes

Exchangeable partition: Kingman paintbox

[Kingman 1978]

Paintboxes

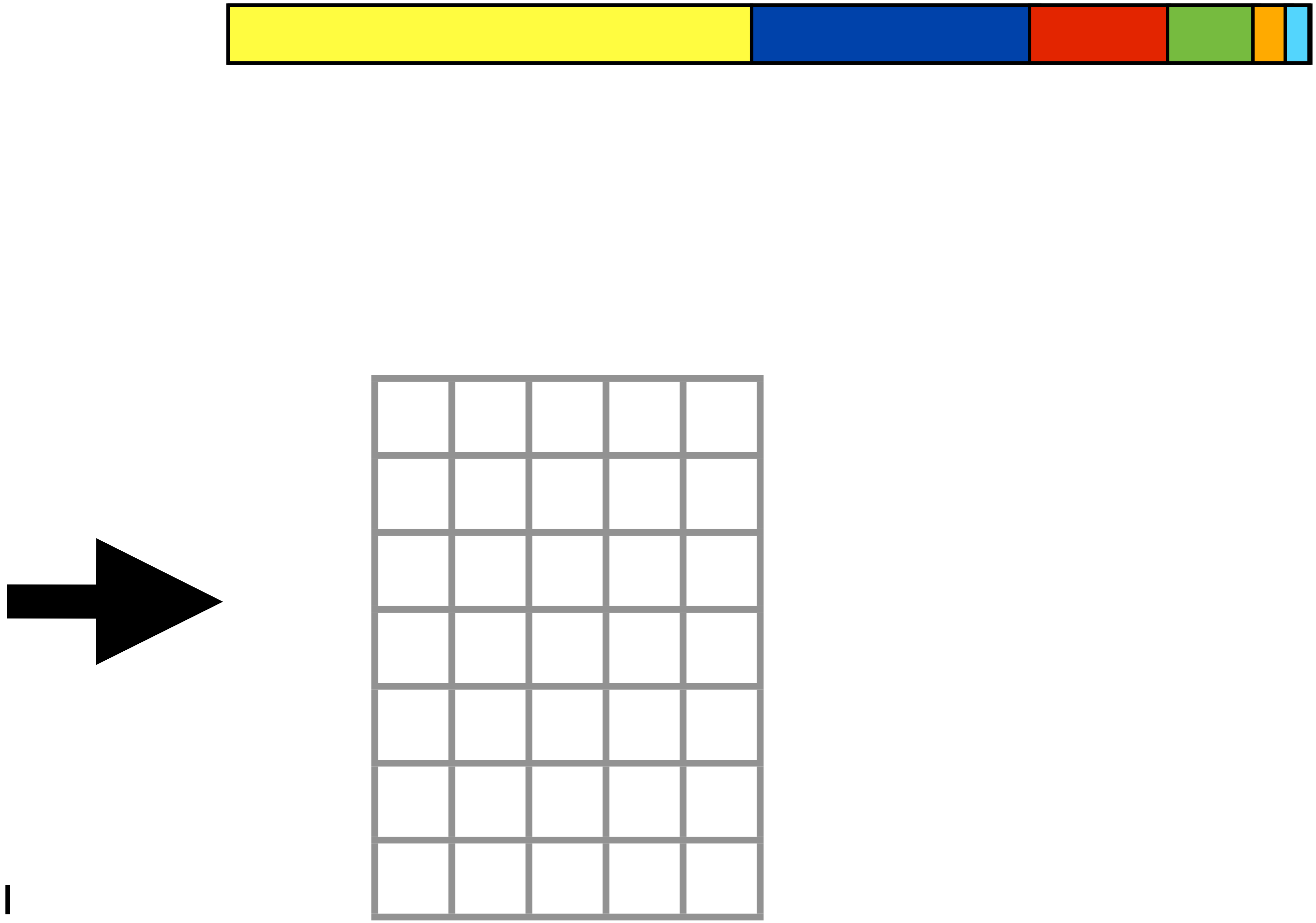
Exchangeable partition: Kingman paintbox



[Kingman 1978]

Paintboxes

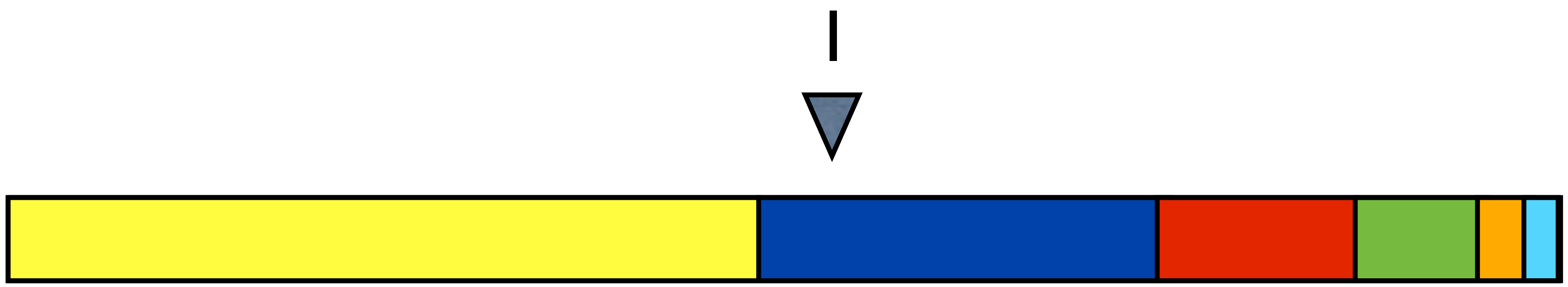
Exchangeable partition: Kingman paintbox

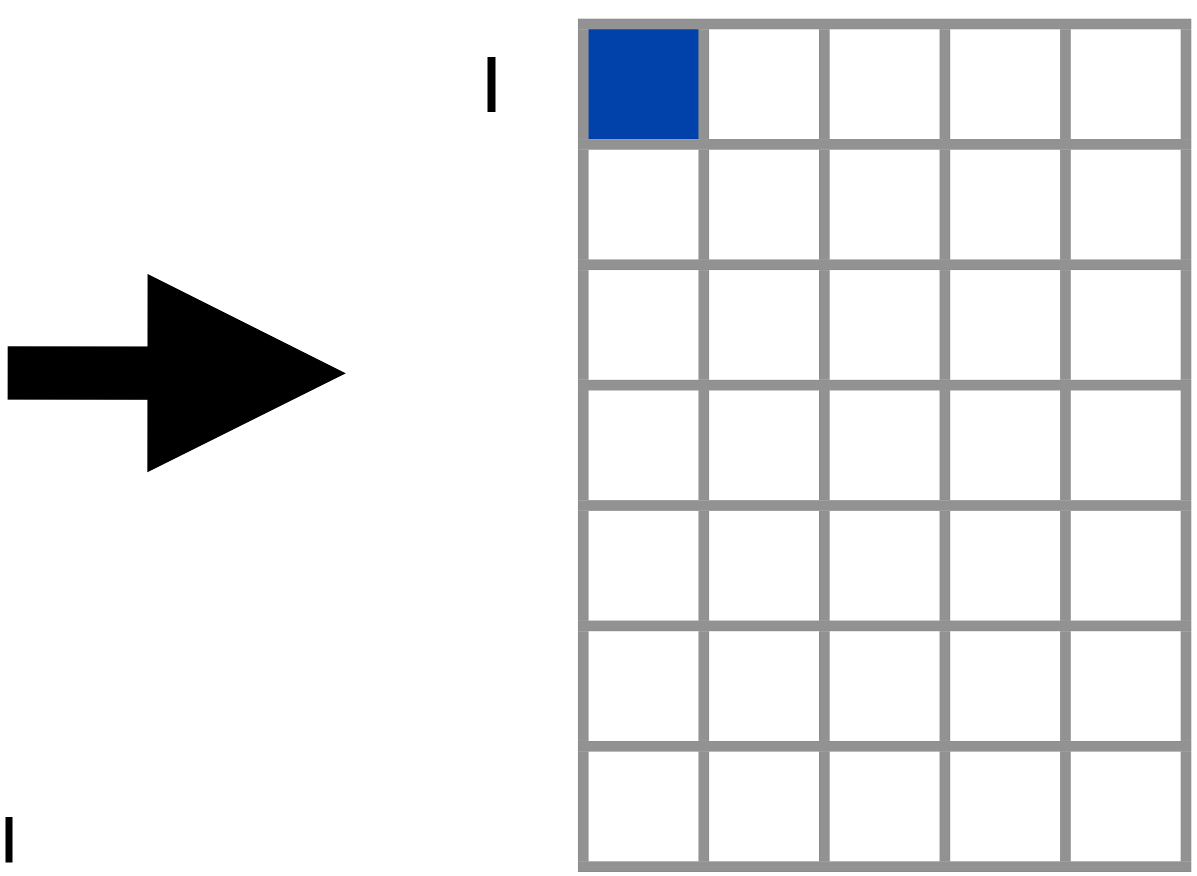


[Kingman 1978]

Paintboxes

Exchangeable partition: Kingman paintbox

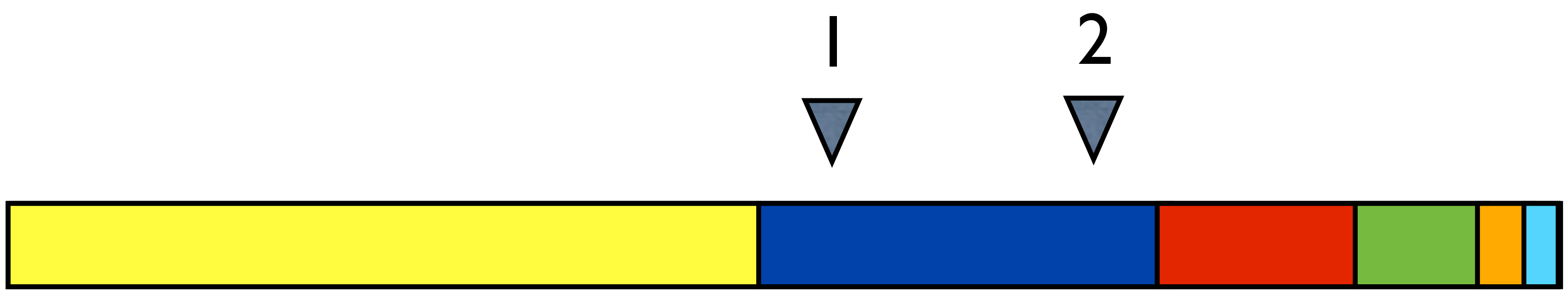


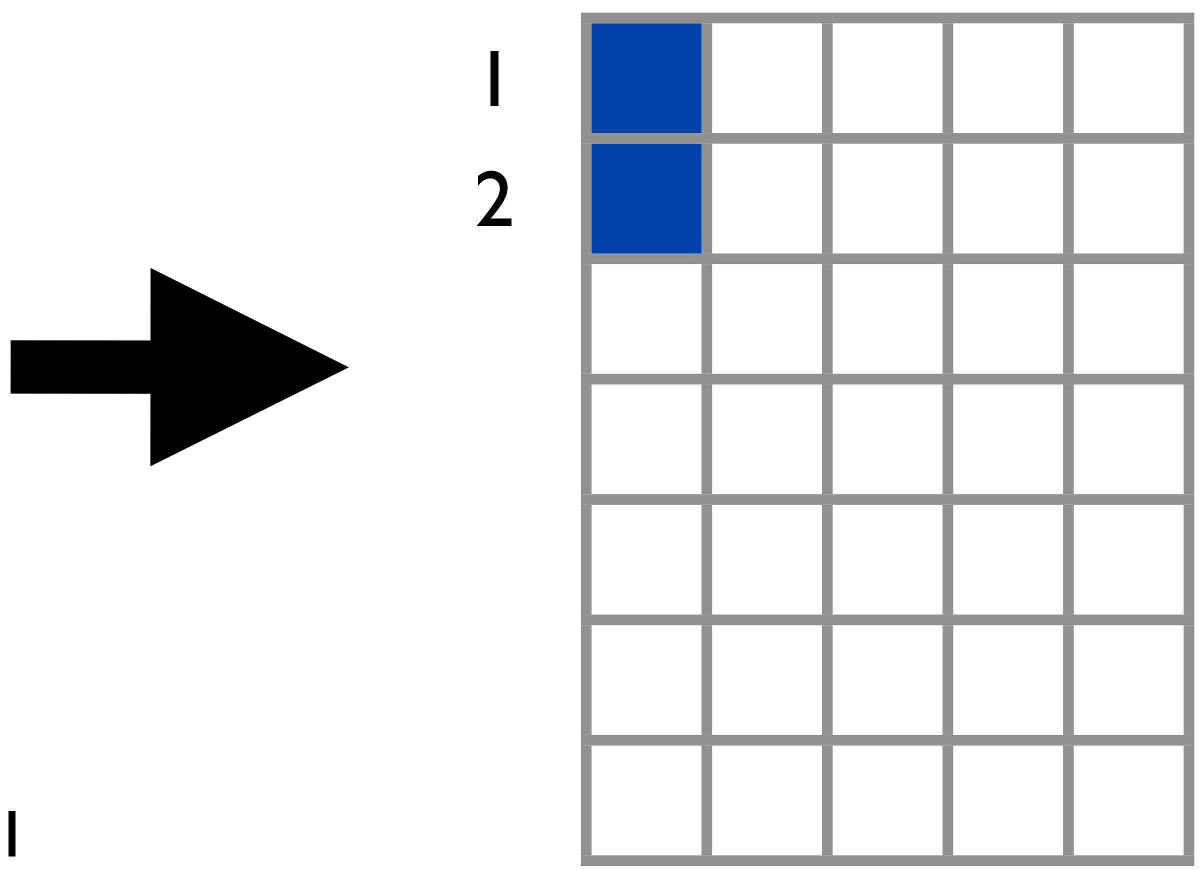


[Kingman 1978]

Paintboxes

Exchangeable partition: Kingman paintbox

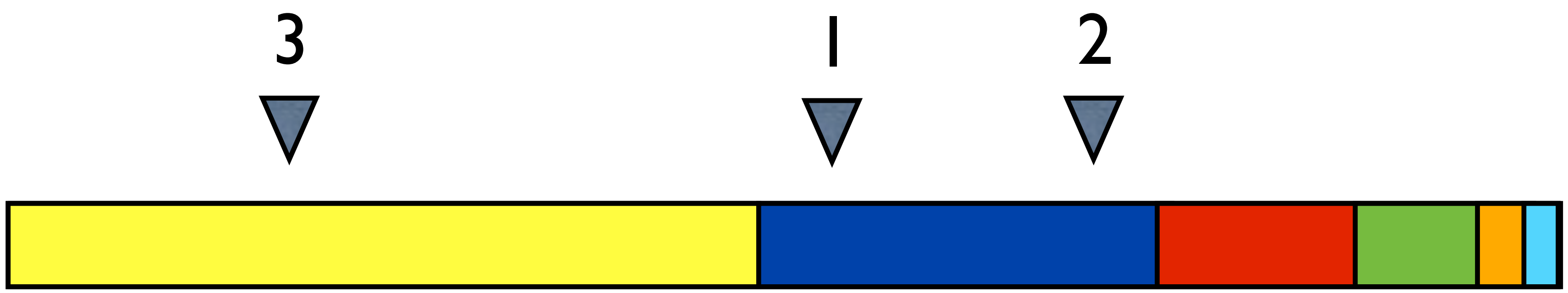


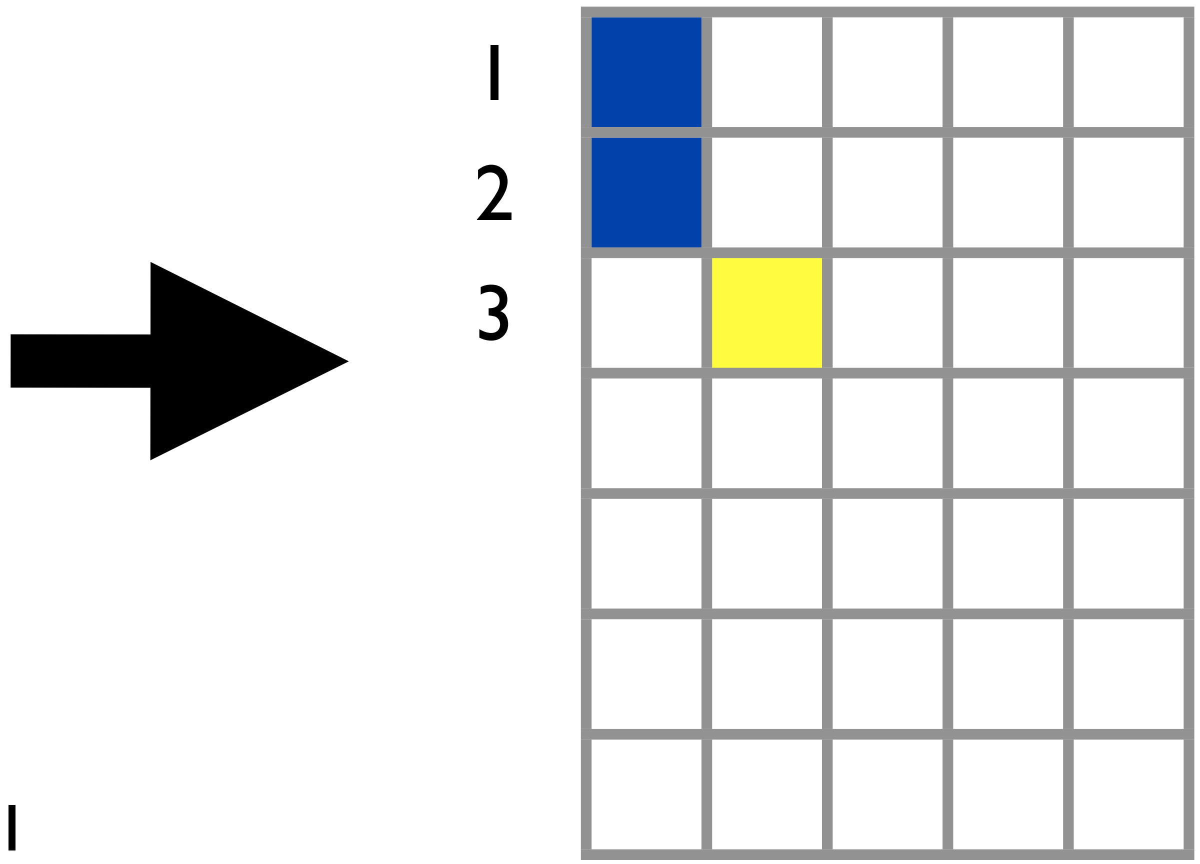


[Kingman 1978]

Paintboxes

Exchangeable partition: Kingman paintbox

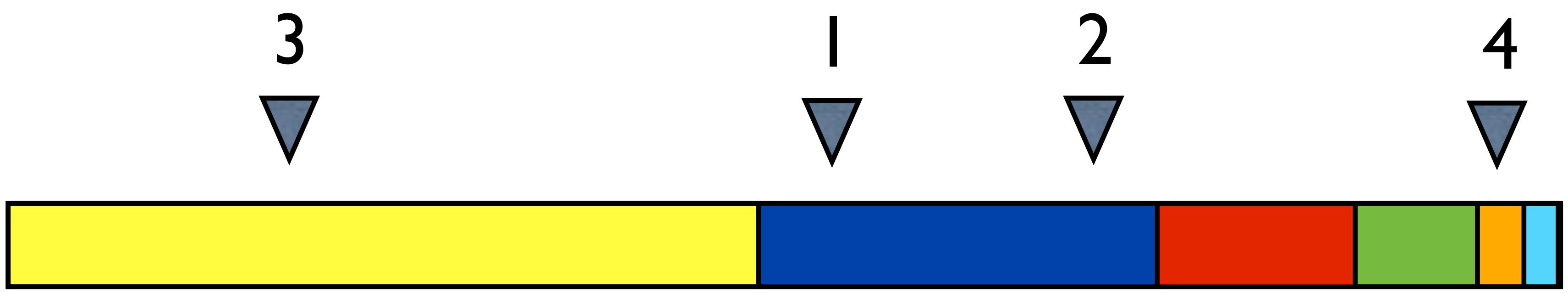


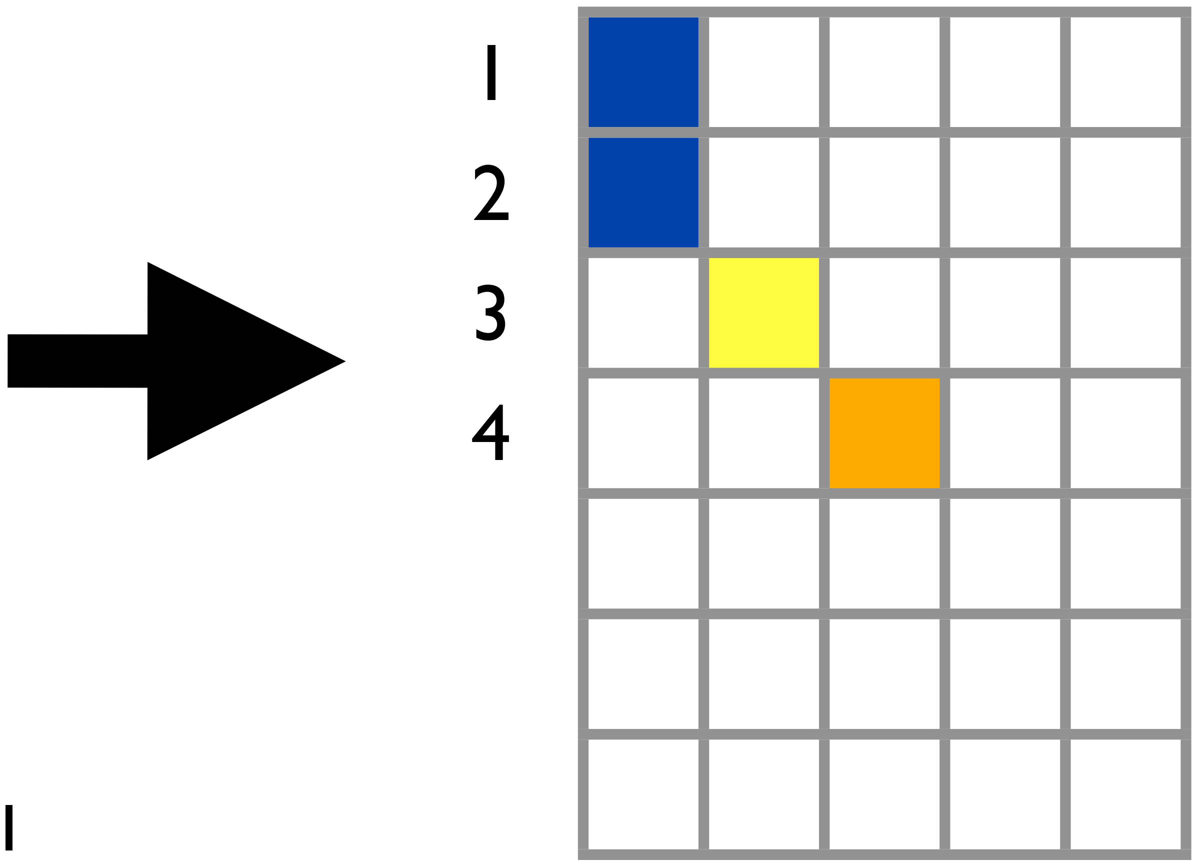


[Kingman 1978]

Paintboxes

Exchangeable partition: Kingman paintbox

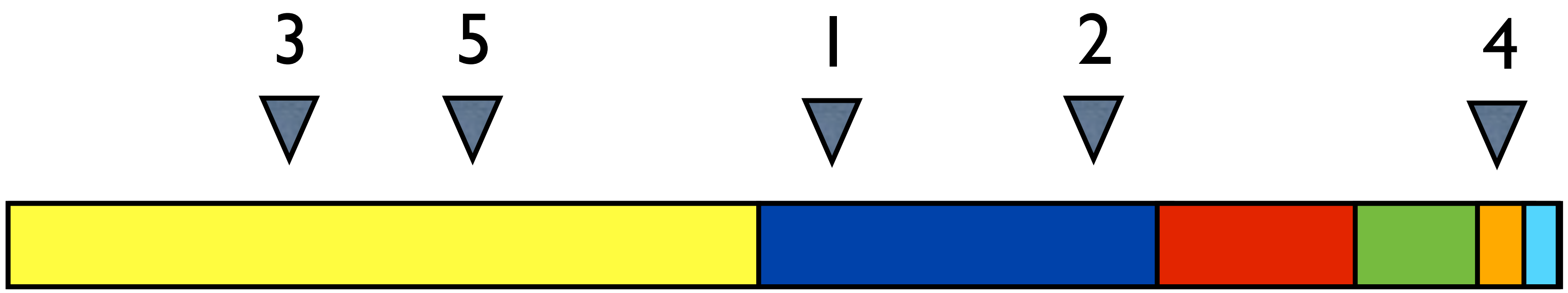


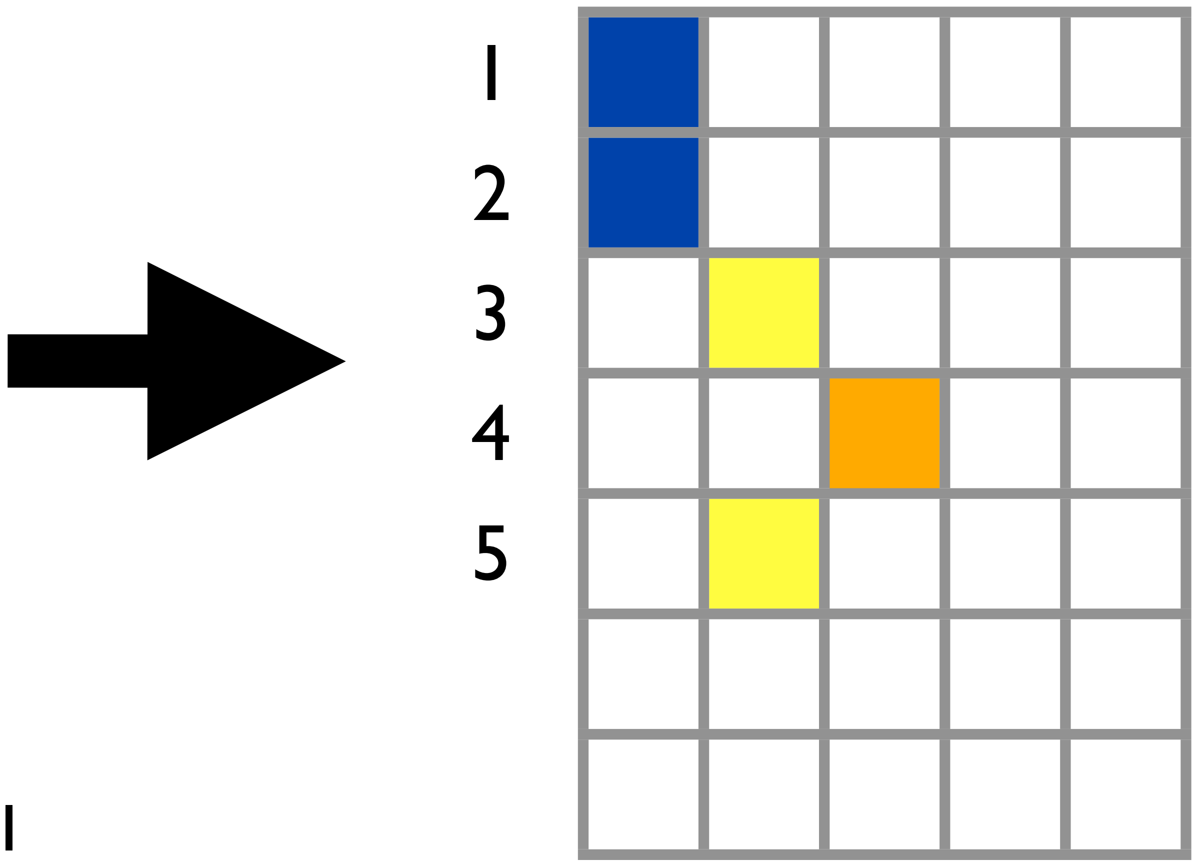


[Kingman 1978]

Paintboxes

Exchangeable partition: Kingman paintbox

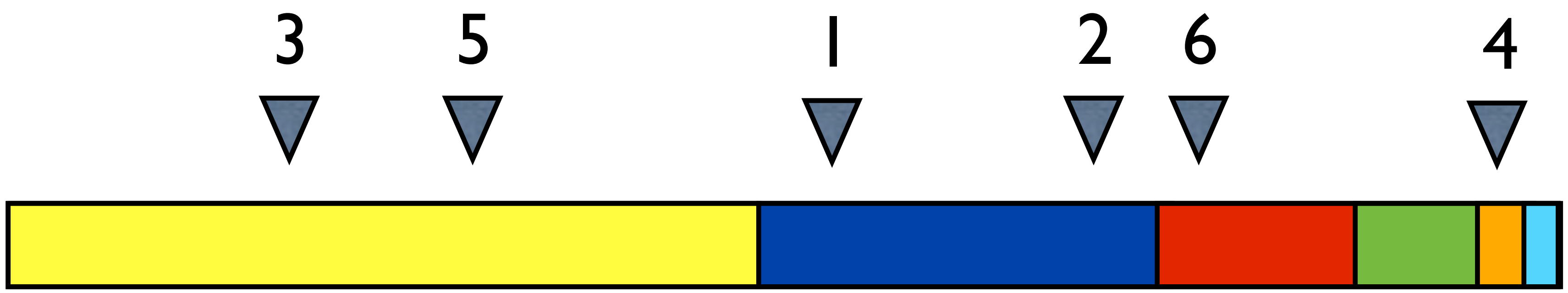


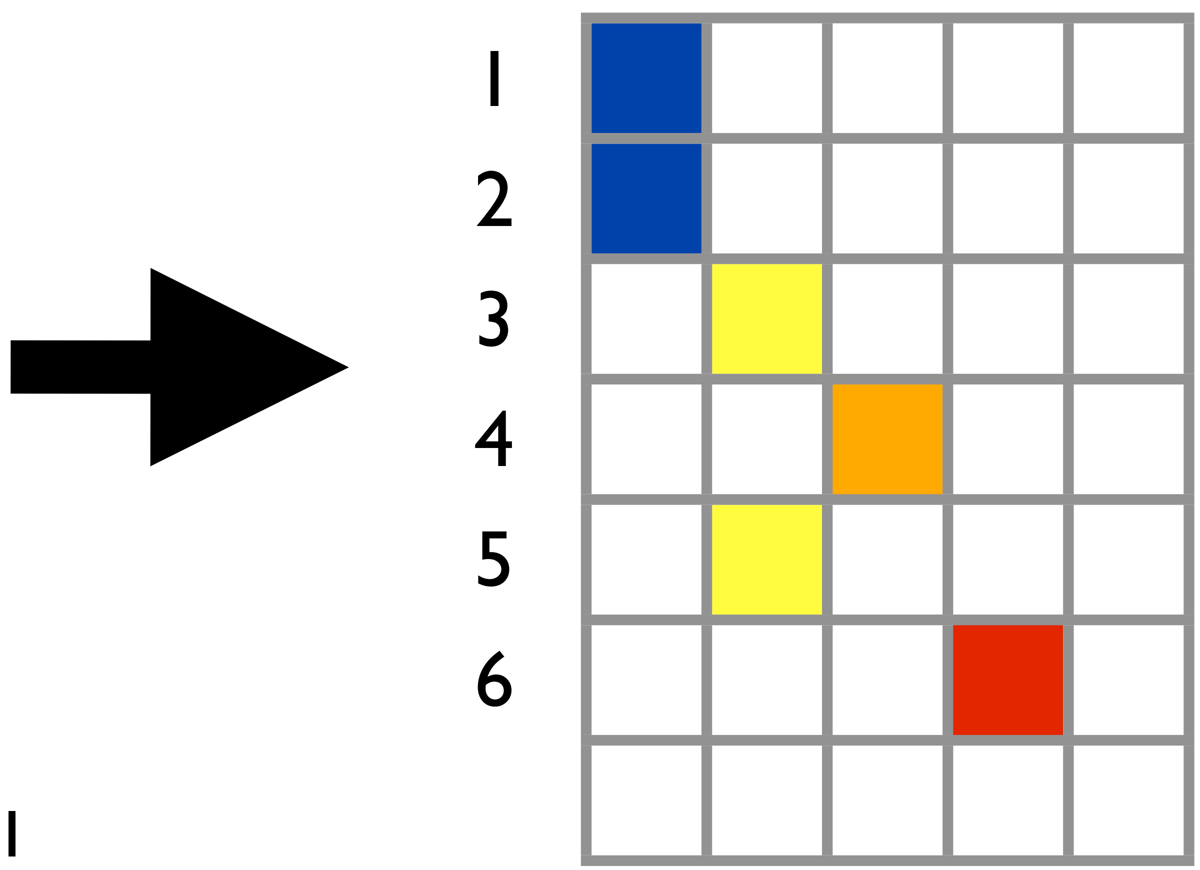


[Kingman 1978]

Paintboxes

Exchangeable partition: Kingman paintbox

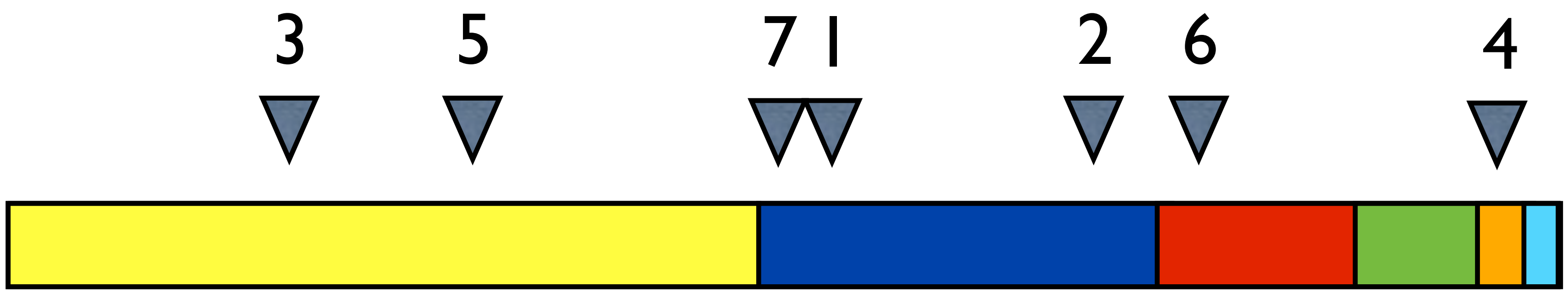


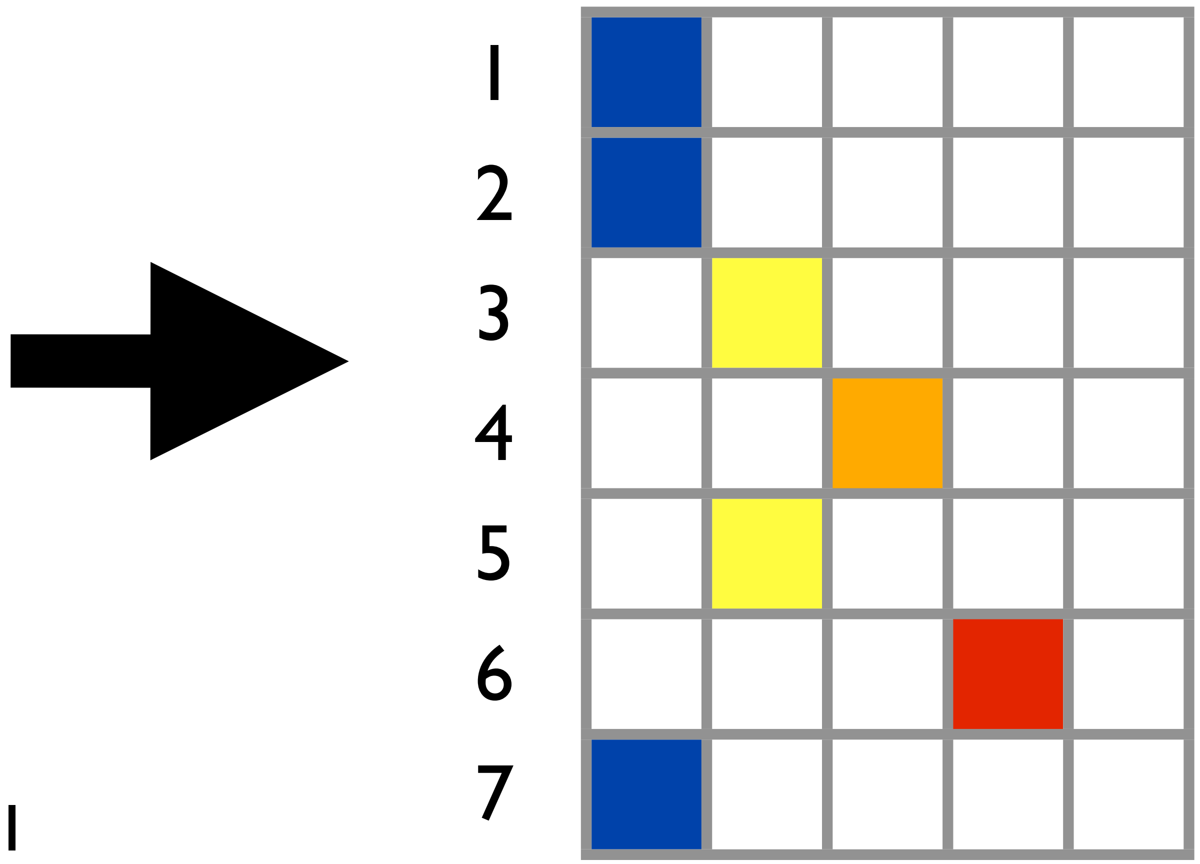


[Kingman 1978]

Paintboxes

Exchangeable partition: Kingman paintbox

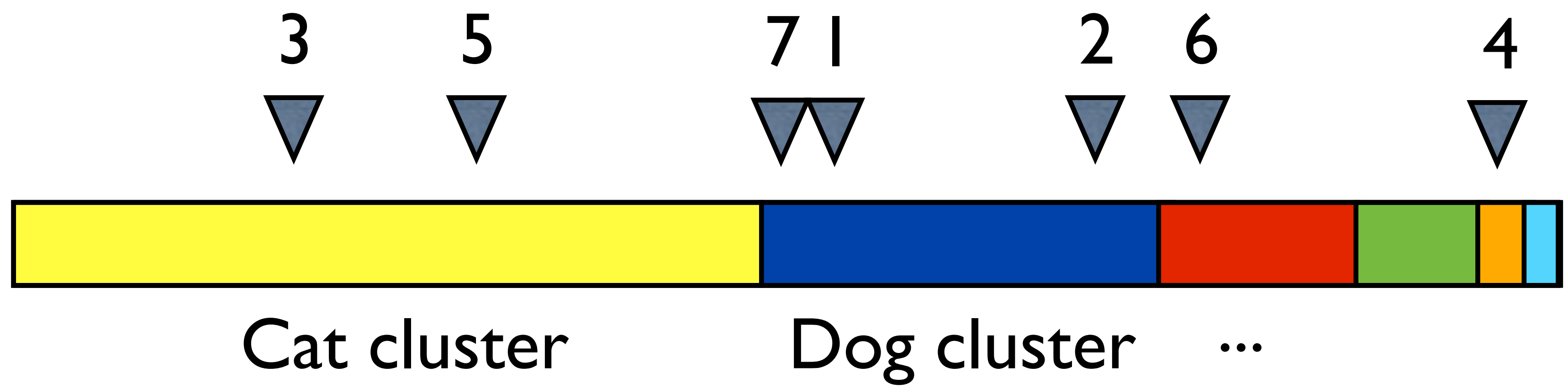


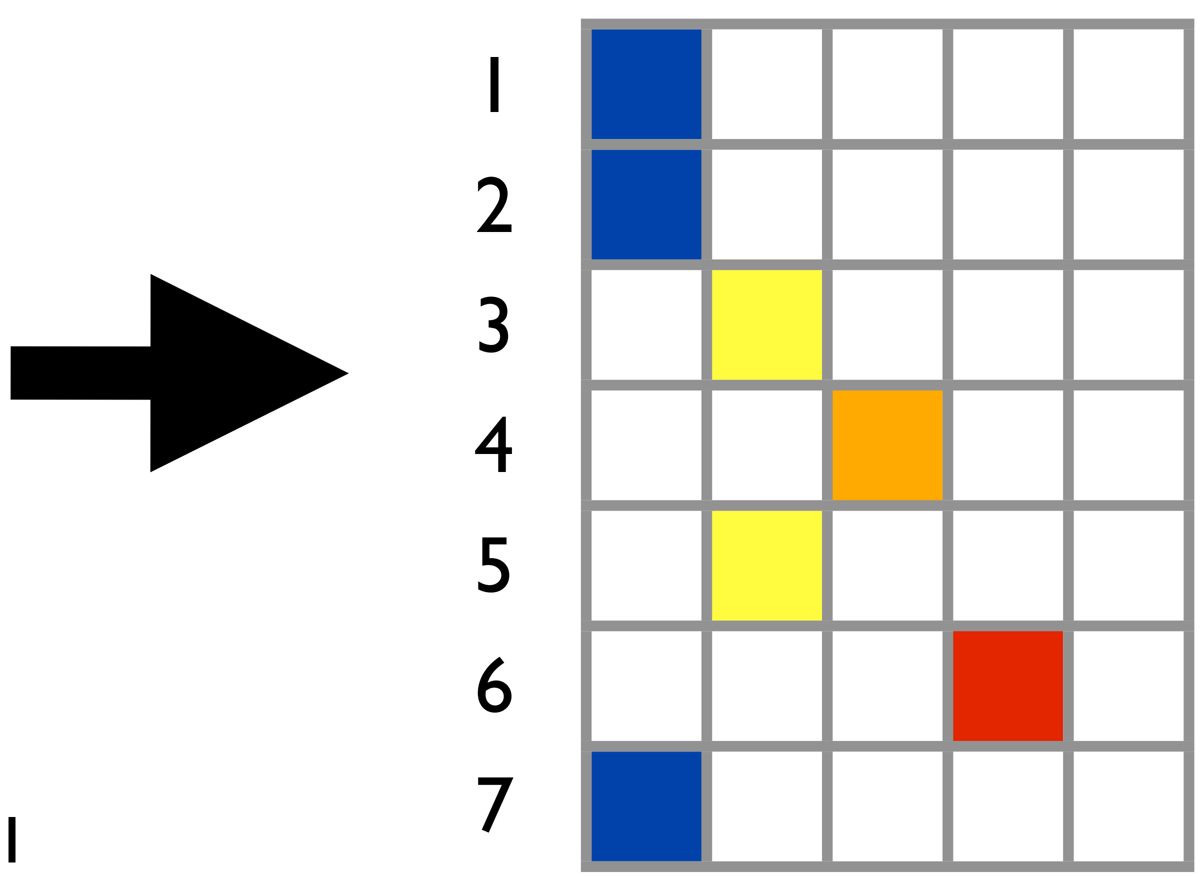


[Kingman 1978]

Paintboxes

Exchangeable partition: Kingman paintbox

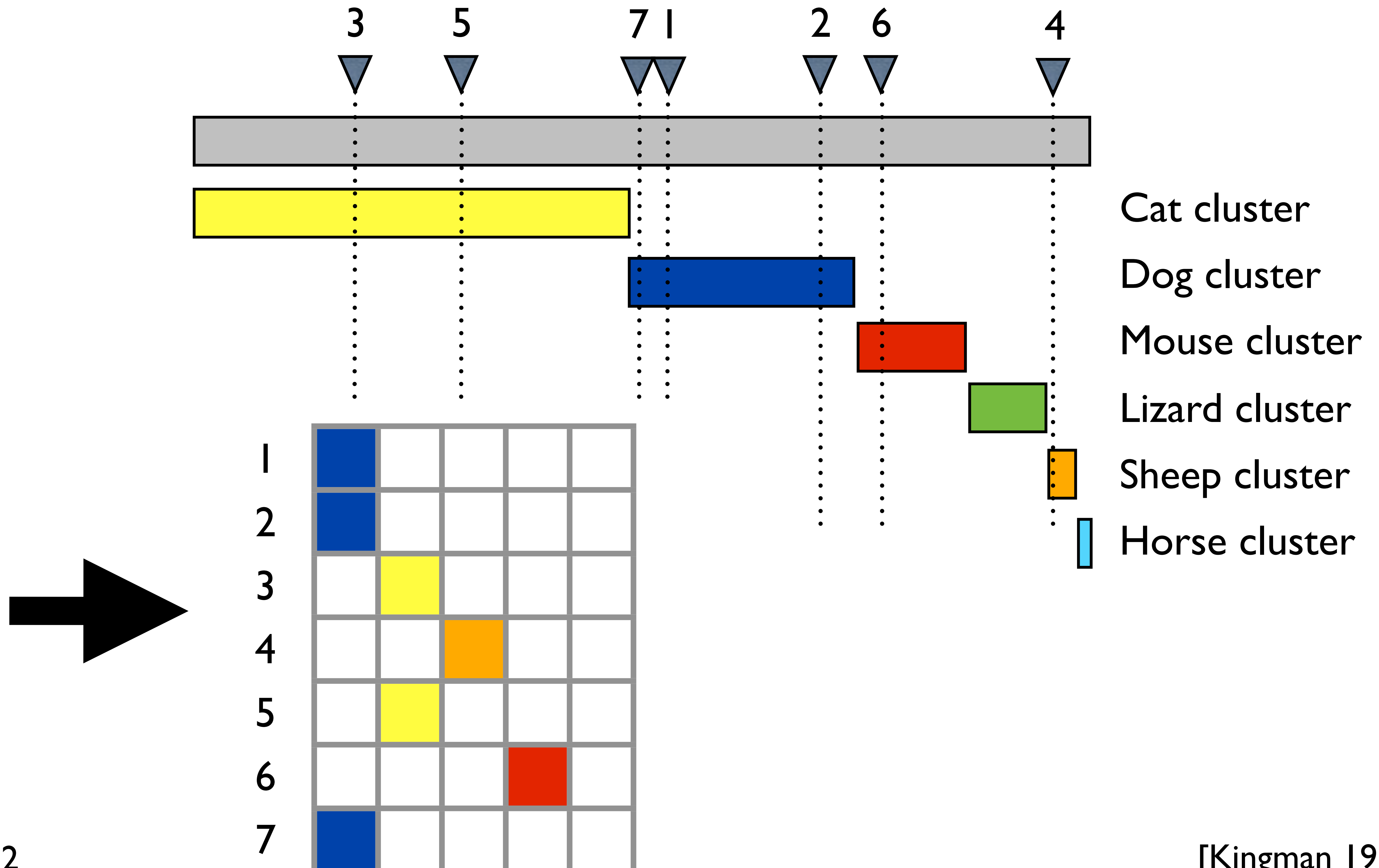




[Kingman 1978]

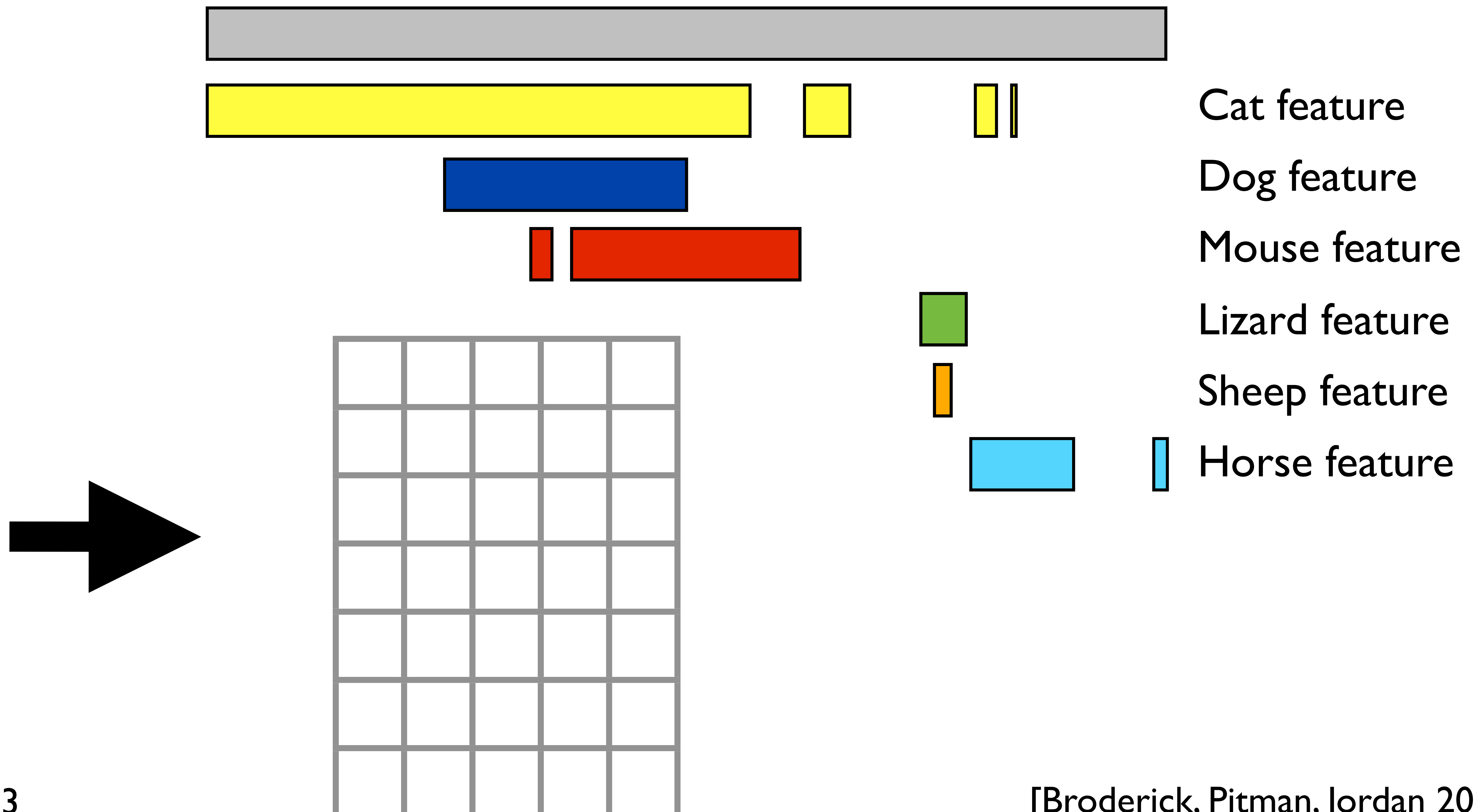
Paintboxes

Exchangeable partition: Kingman paintbox



[Kingman 1978]

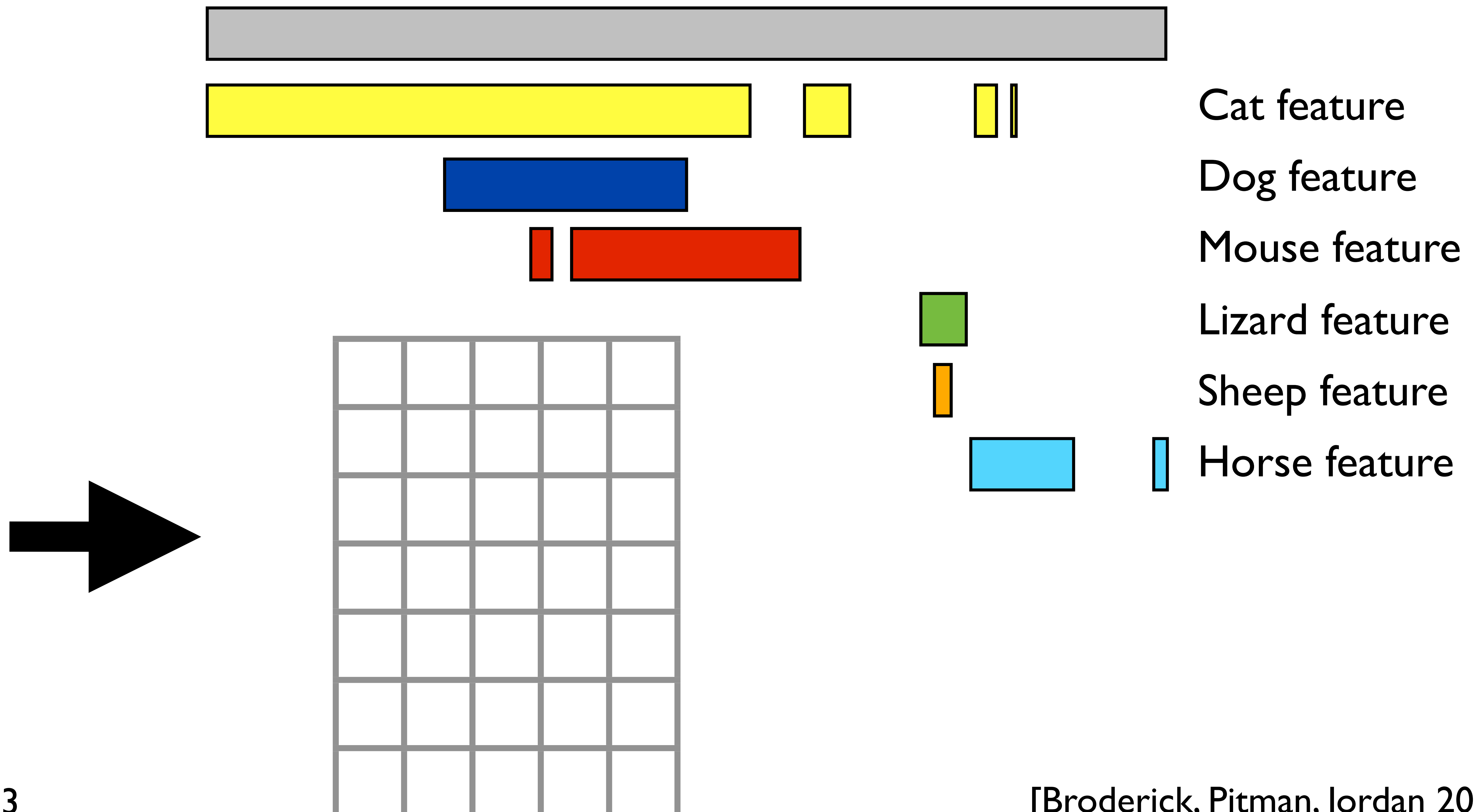
Paintboxes



[Broderick, Pitman, Jordan 2013]

Paintboxes

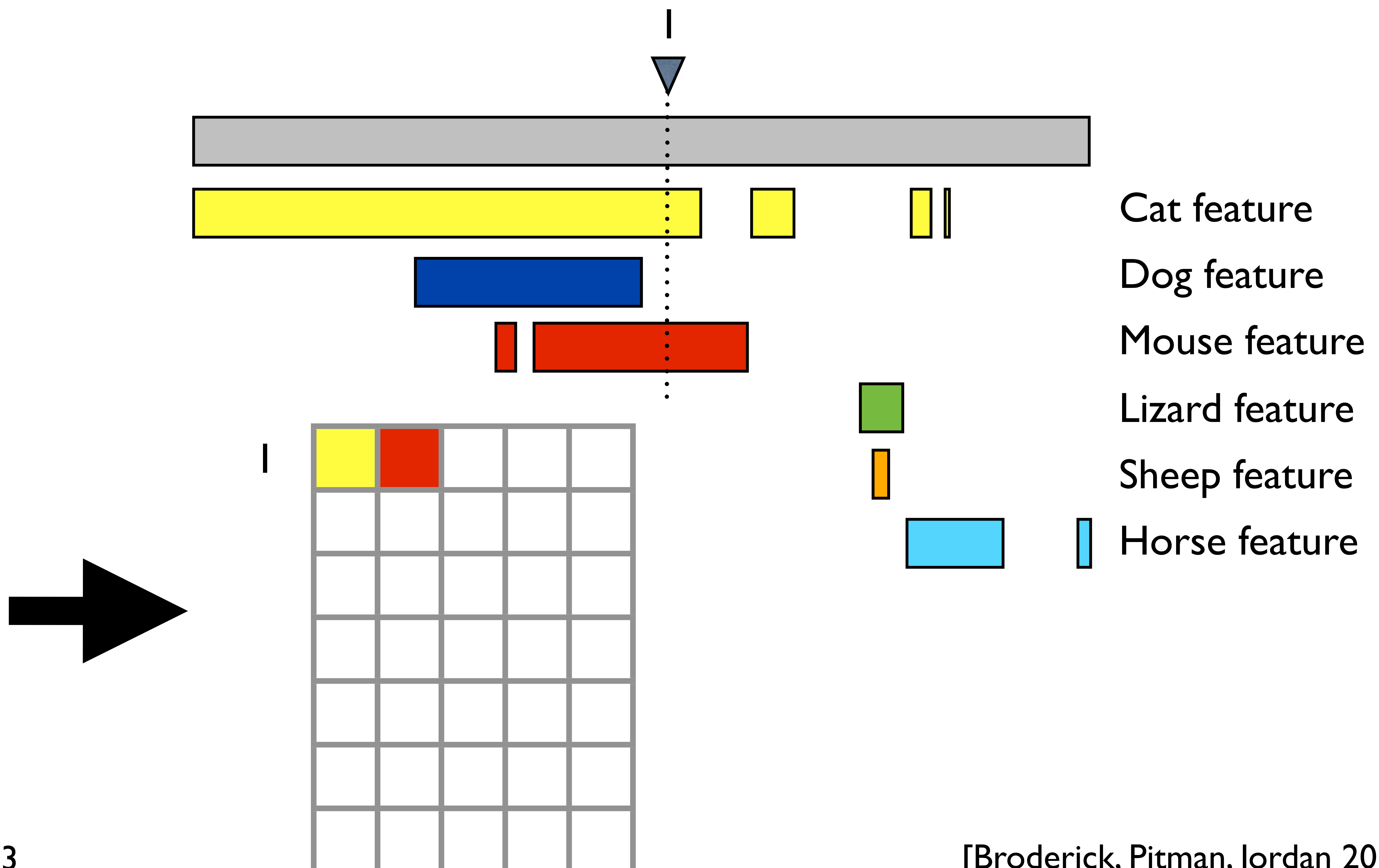
Exchangeable feature allocation: feature paintbox



[Broderick, Pitman, Jordan 2013]

Paintboxes

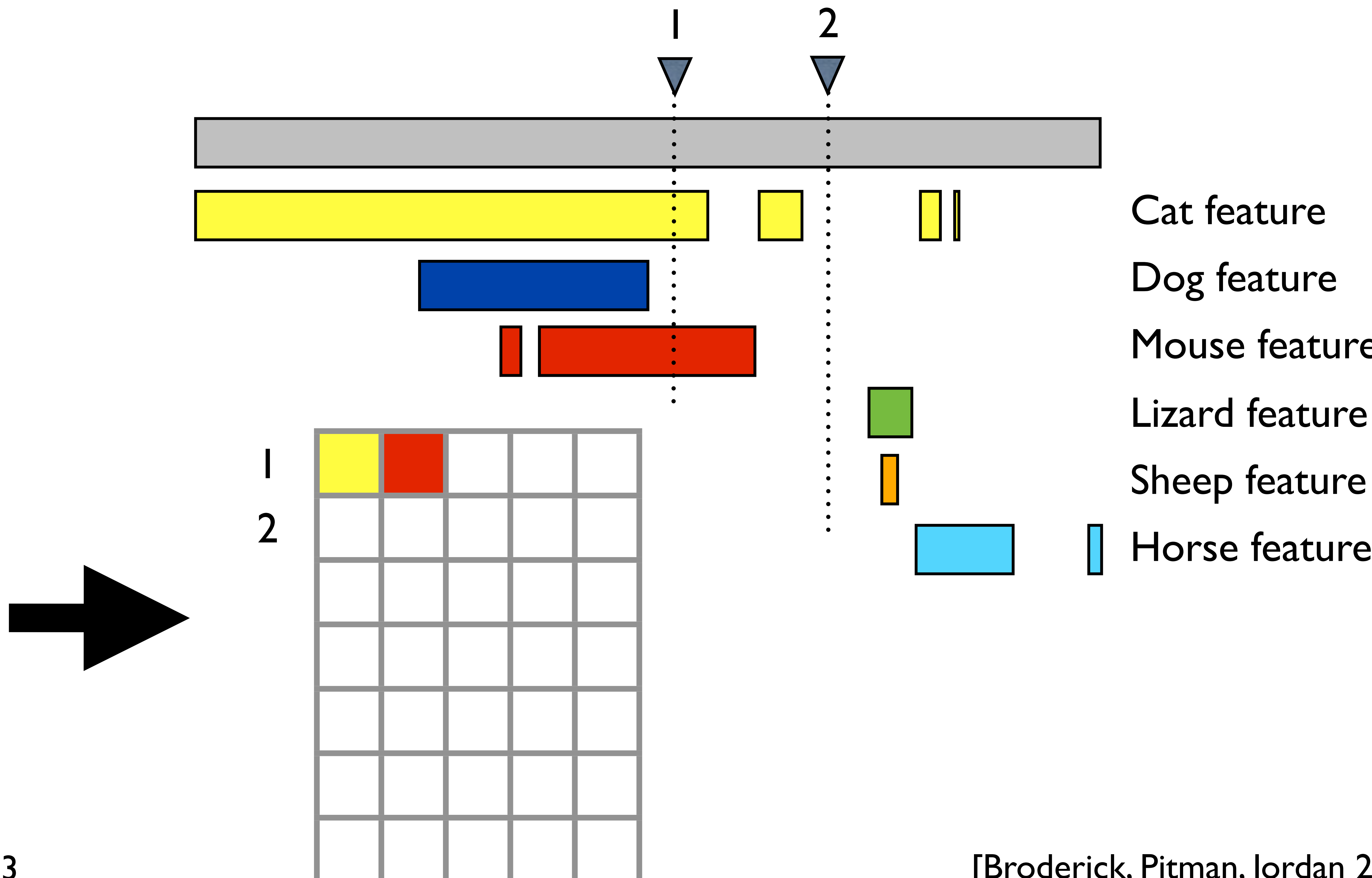
Exchangeable feature allocation: feature paintbox



[Broderick, Pitman, Jordan 2013]

Paintboxes

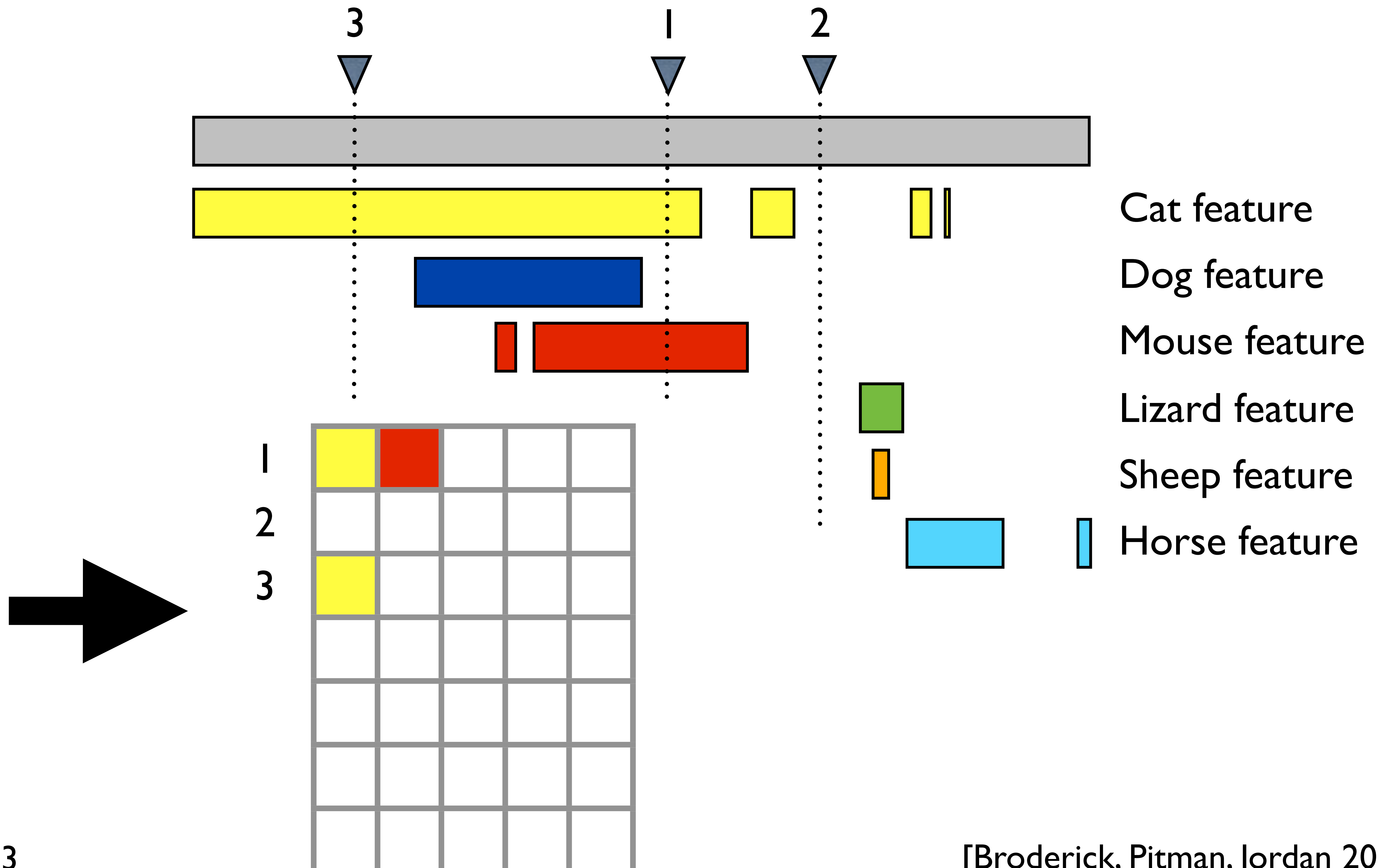
Exchangeable feature allocation: feature paintbox



[Broderick, Pitman, Jordan 2013]

Paintboxes

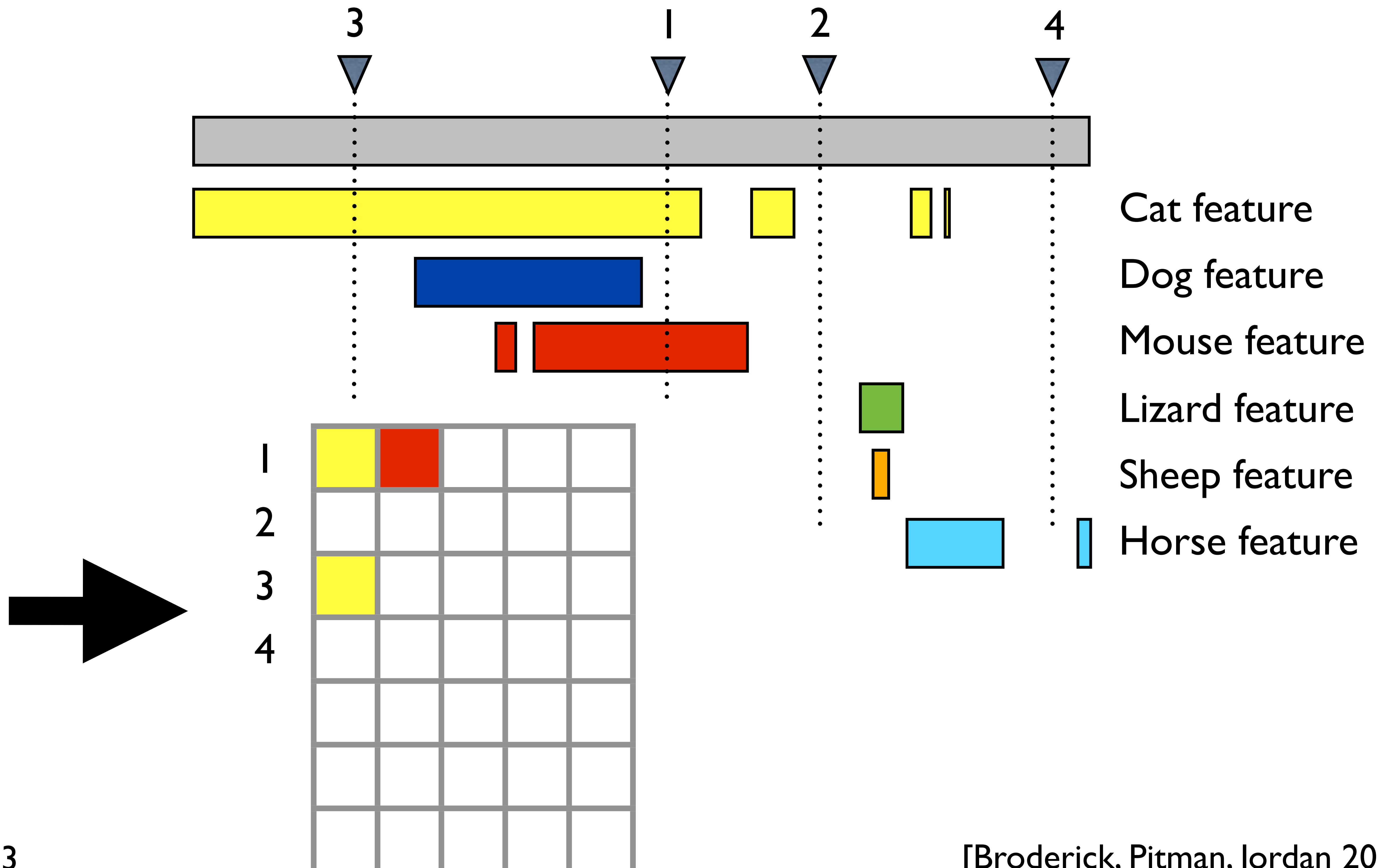
Exchangeable feature allocation: feature paintbox



[Broderick, Pitman, Jordan 2013]

Paintboxes

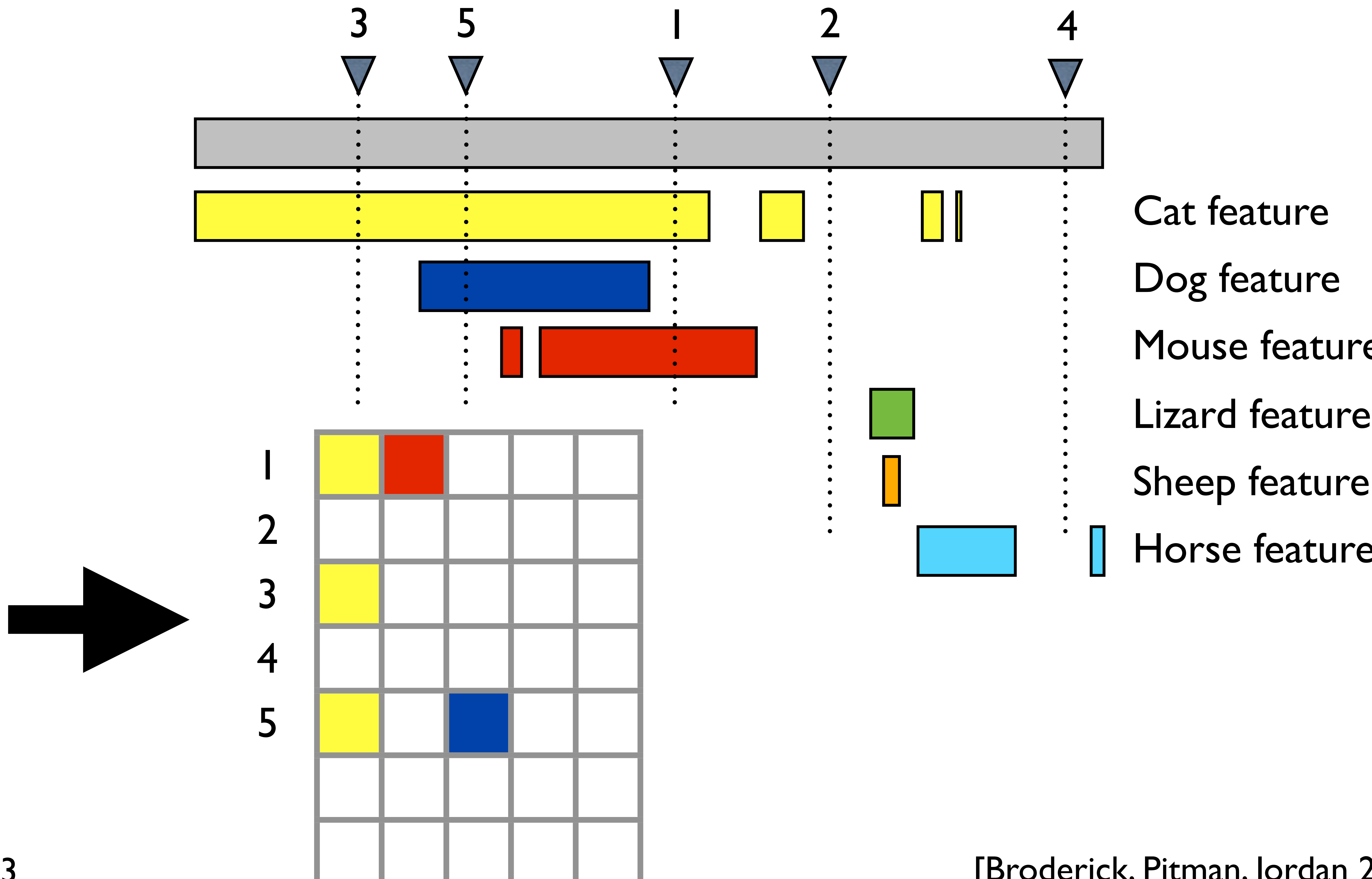
Exchangeable feature allocation: feature paintbox



[Broderick, Pitman, Jordan 2013]

Paintboxes

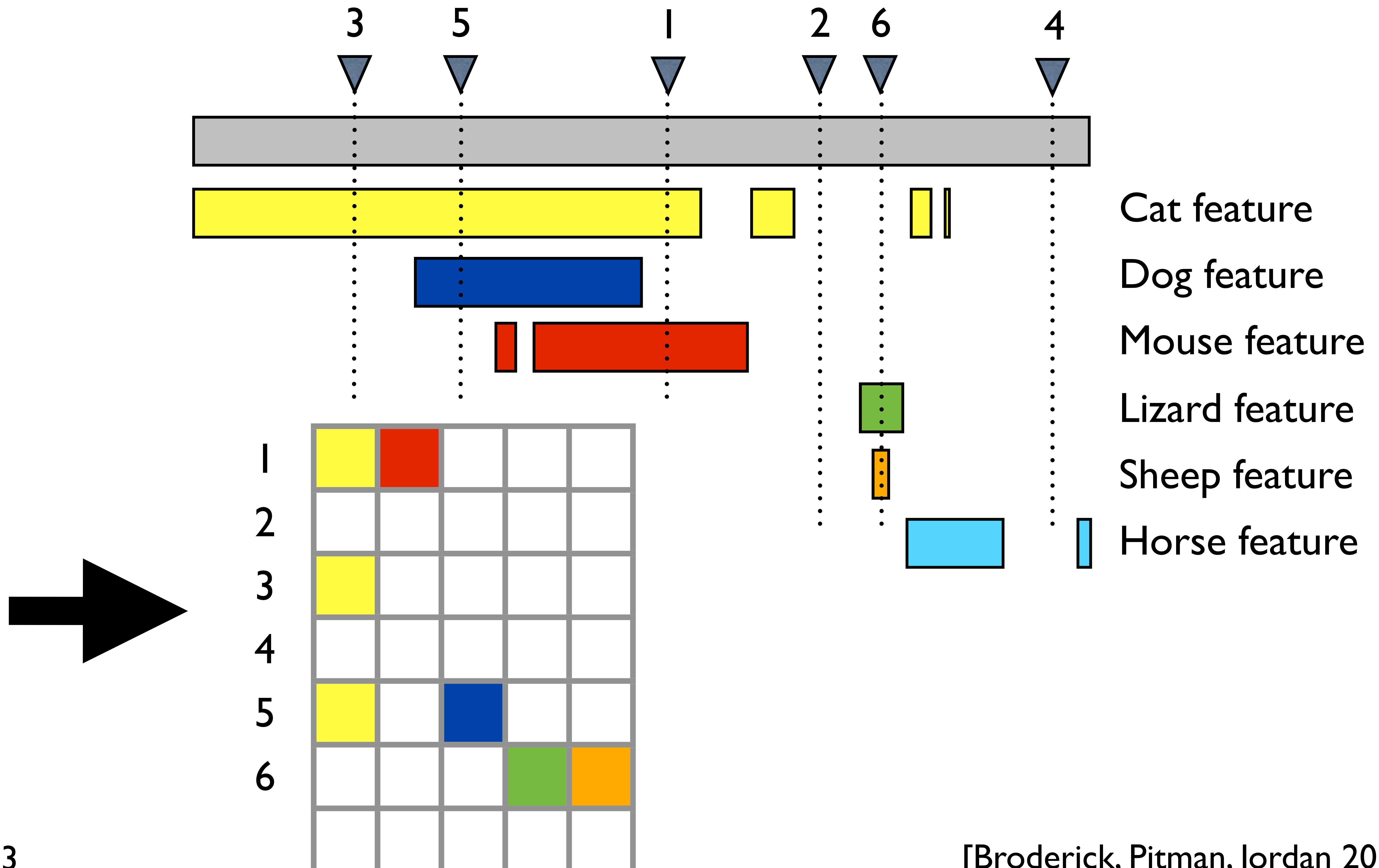
Exchangeable feature allocation: feature paintbox



[Broderick, Pitman, Jordan 2013]

Paintboxes

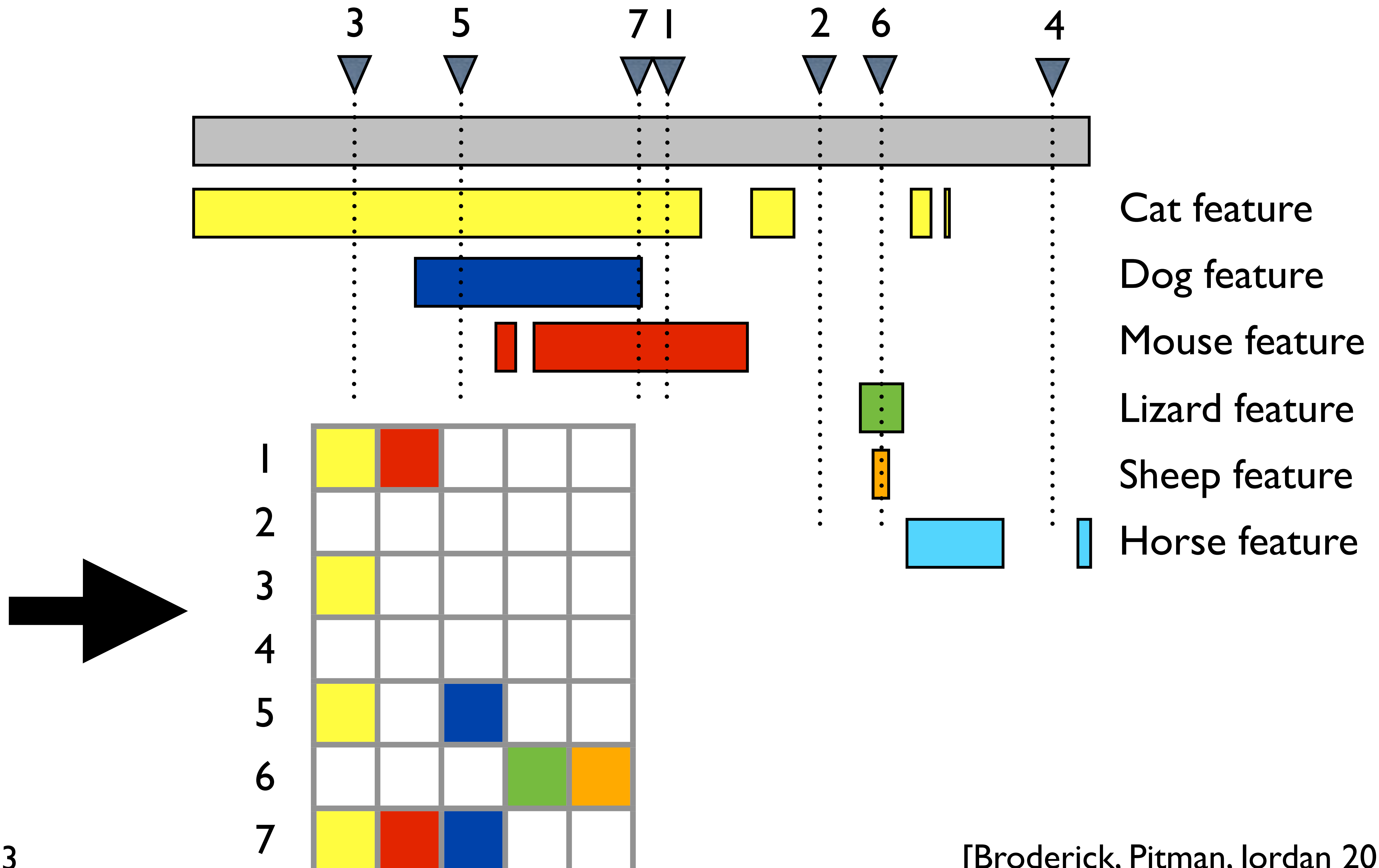
Exchangeable feature allocation: feature paintbox



[Broderick, Pitman, Jordan 2013]

Paintboxes

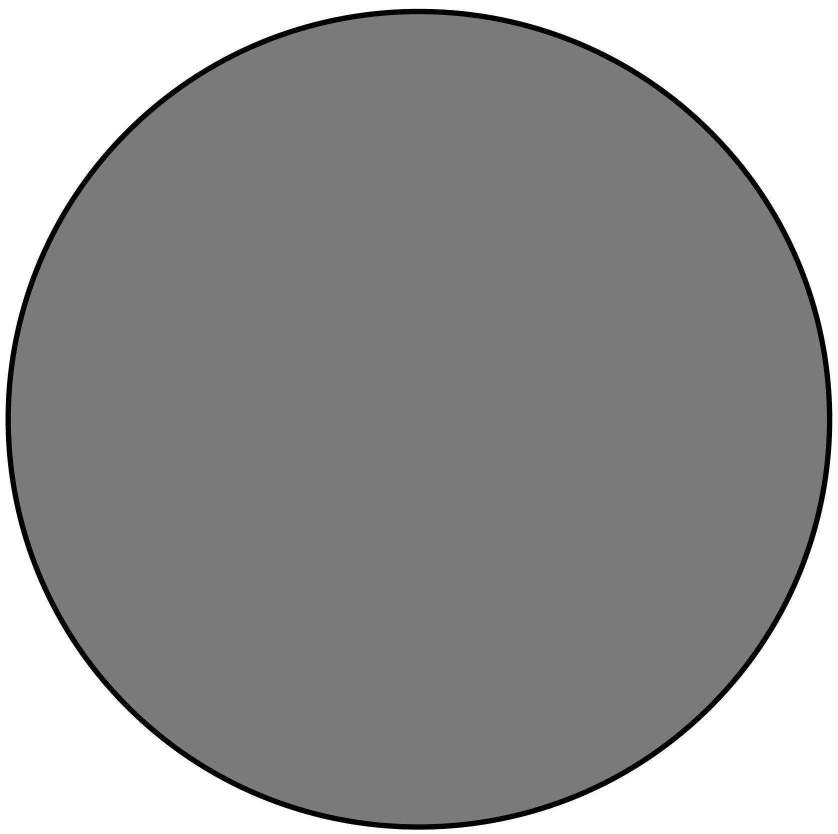
Exchangeable feature allocation: feature paintbox



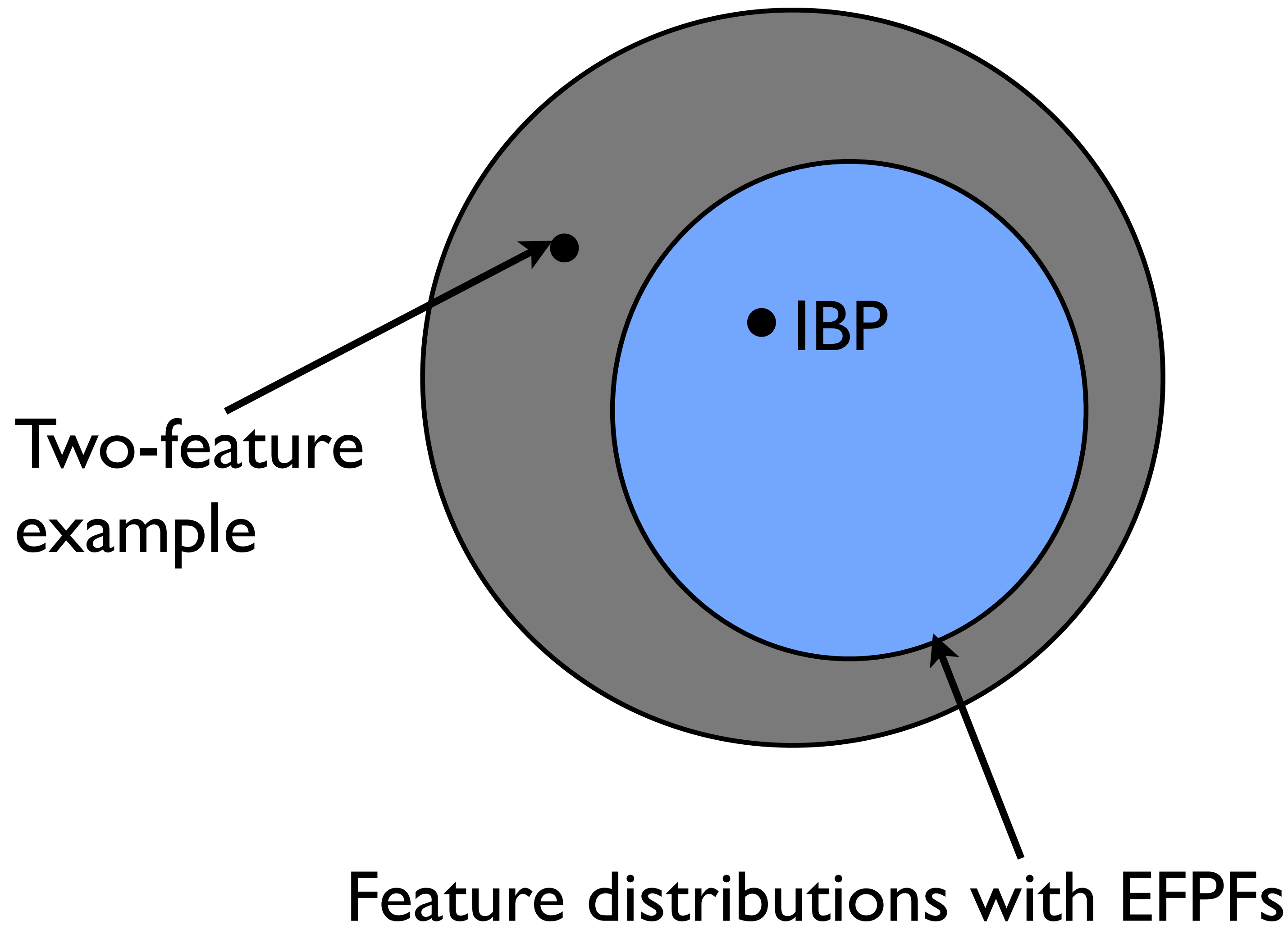
[Broderick, Pitman, Jordan 2013]

Paintboxes

Exchangeable cluster distributions
= Cluster distributions with EPPFs



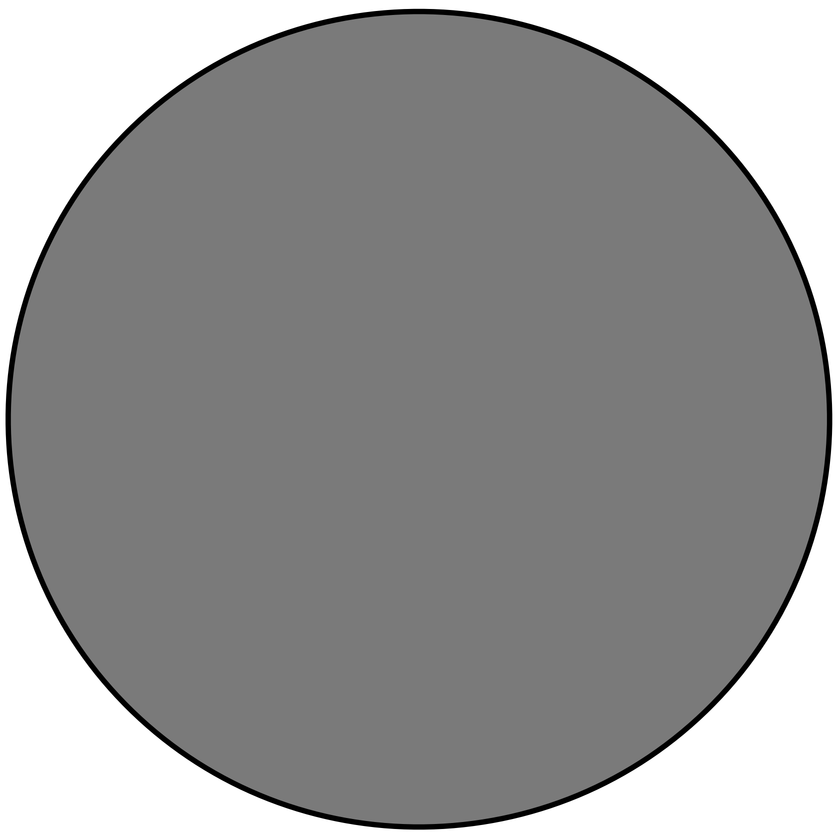
Exchangeable feature distributions



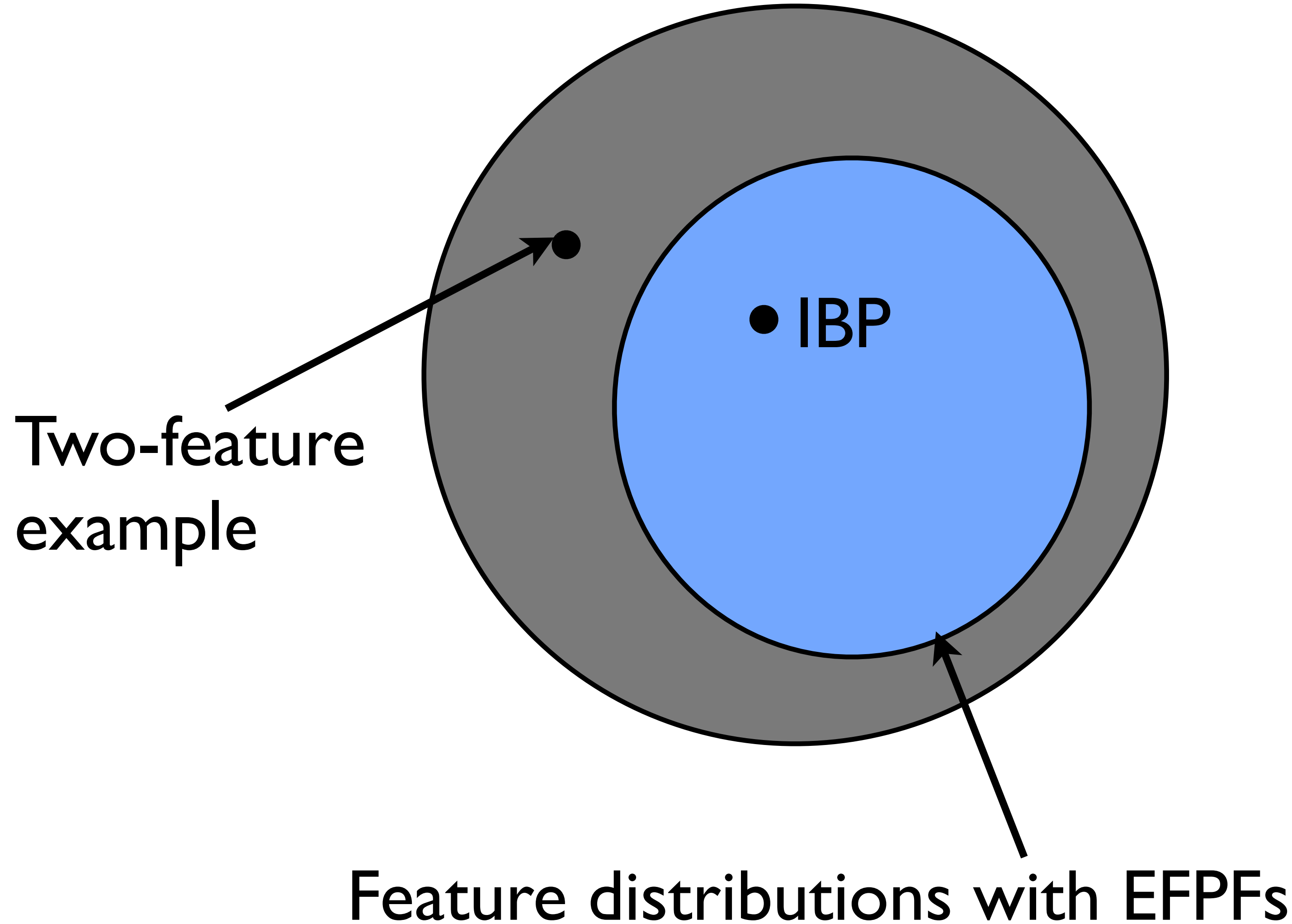
[Broderick, Pitman, Jordan 2013]

Paintboxes

Exchangeable cluster distributions
= Cluster distributions with EPPFs
= Kingman paintbox partitions



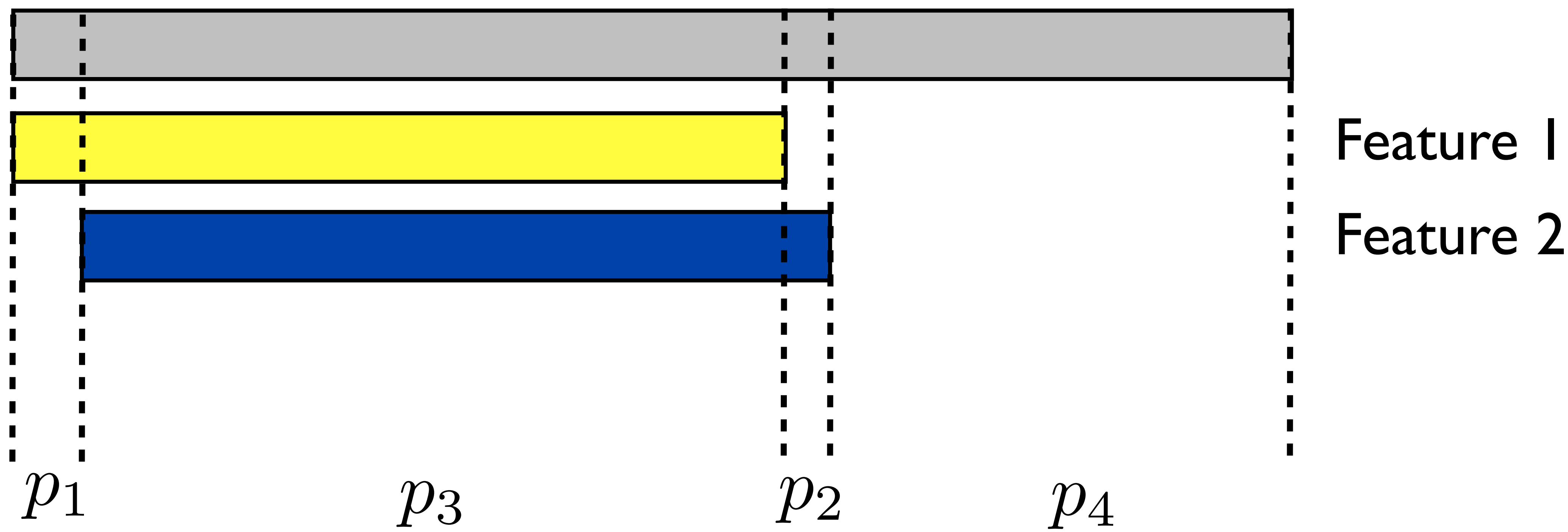
Exchangeable feature distributions
= Feature paintbox allocations



[Broderick, Pitman, Jordan 2013]

Paintboxes

Two feature example



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

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

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

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

Paintboxes

Indian buffet process: beta feature frequencies

[Thibaux, Jordan 2007]

Paintboxes

Indian buffet process: beta feature frequencies

For m = 1, 2, ...
1. Draw $K_m^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + m - 1}\right)$

[Thibaux, Jordan 2007]

Paintboxes

Indian buffet process: beta feature frequencies

For *m* = 1, 2, ...

1. Draw $K_m^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + m - 1}\right)$

   Set $K_m = \sum_{j=1}^{m} K_j^+$

2. For k = $K_{m-1}, \ldots, K_m$

   Draw a frequency of size

$$q_k \sim \mathrm{Beta}(1, \theta + m - 1)$$

[Thibaux, Jordan 2007]

Paintboxes

Indian buffet process: beta feature frequencies

For m = 1, 2, ...

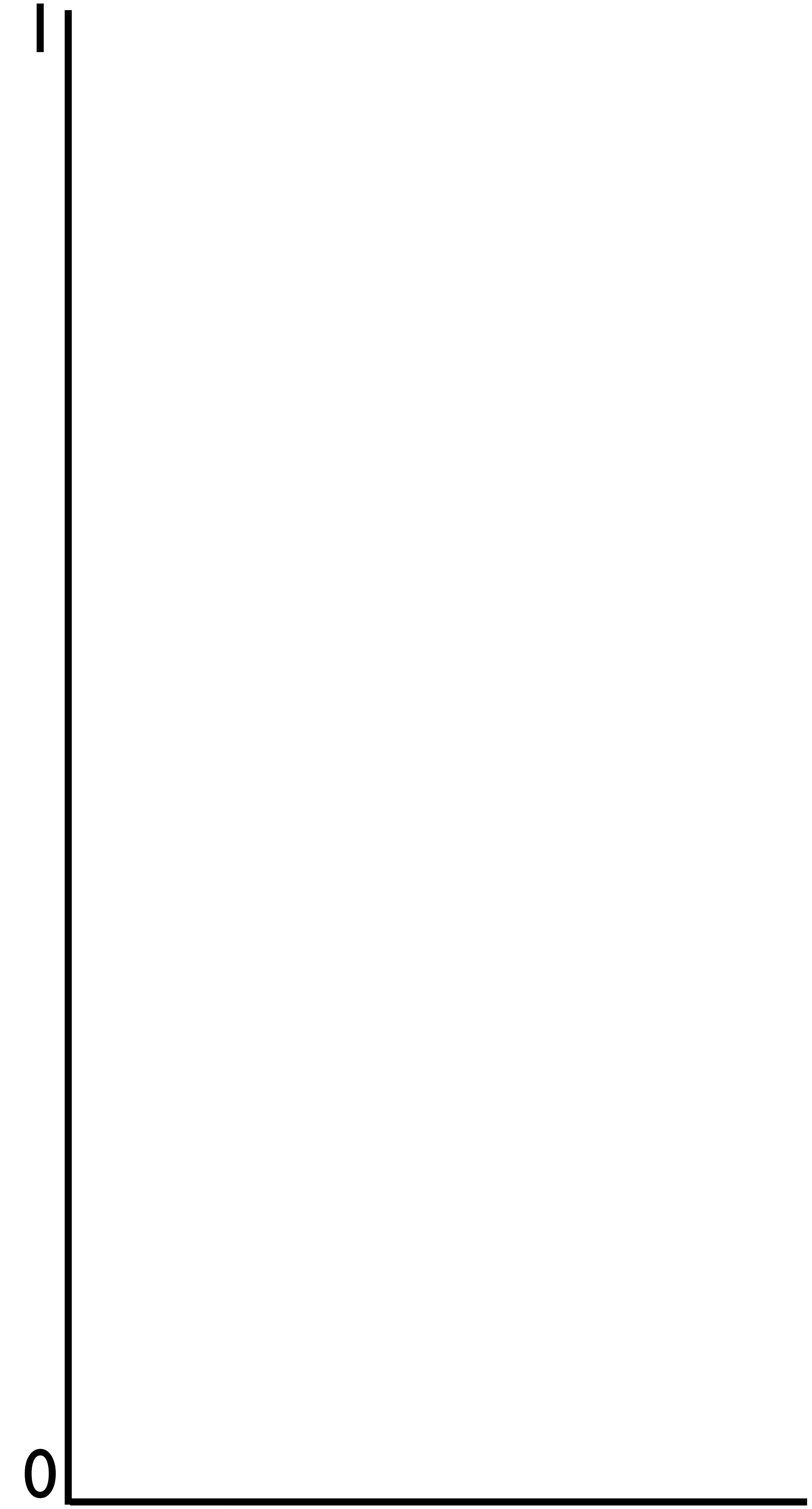
1. Draw $K_m^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + m - 1}\right)$

   Set $K_m = \sum_{j=1}^{m} K_j^+$

2. For k = $K_{m-1}, \ldots, K_m$

   Draw a frequency of size

$$q_k \sim \mathrm{Beta}(1, \theta + m - 1)$$



[Thibaux, Jordan 2007]

Paintboxes

Indian buffet process: beta feature frequencies

For m = 1, 2, ...

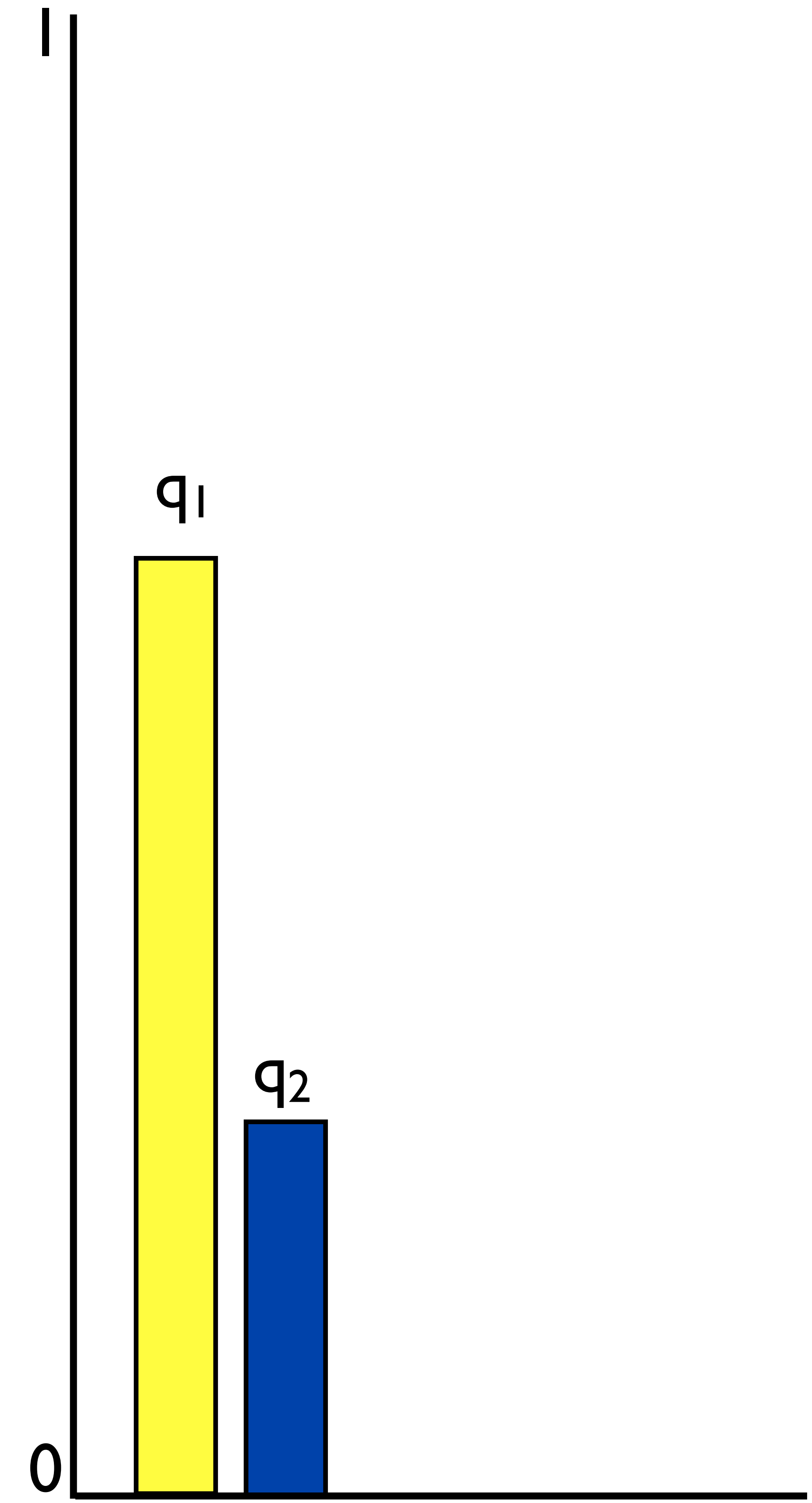
1. Draw $K_m^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + m - 1}\right)$

   Set $K_m = \sum_{j=1}^{m} K_j^+$

2. For k = $K_{m-1}, \ldots, K_m$

   Draw a frequency of size

   $$q_k \sim \mathrm{Beta}(1, \theta + m - 1)$$



[Thibaux, Jordan 2007]

Paintboxes

Indian buffet process: beta feature frequencies

For *m* = 1, 2, ...

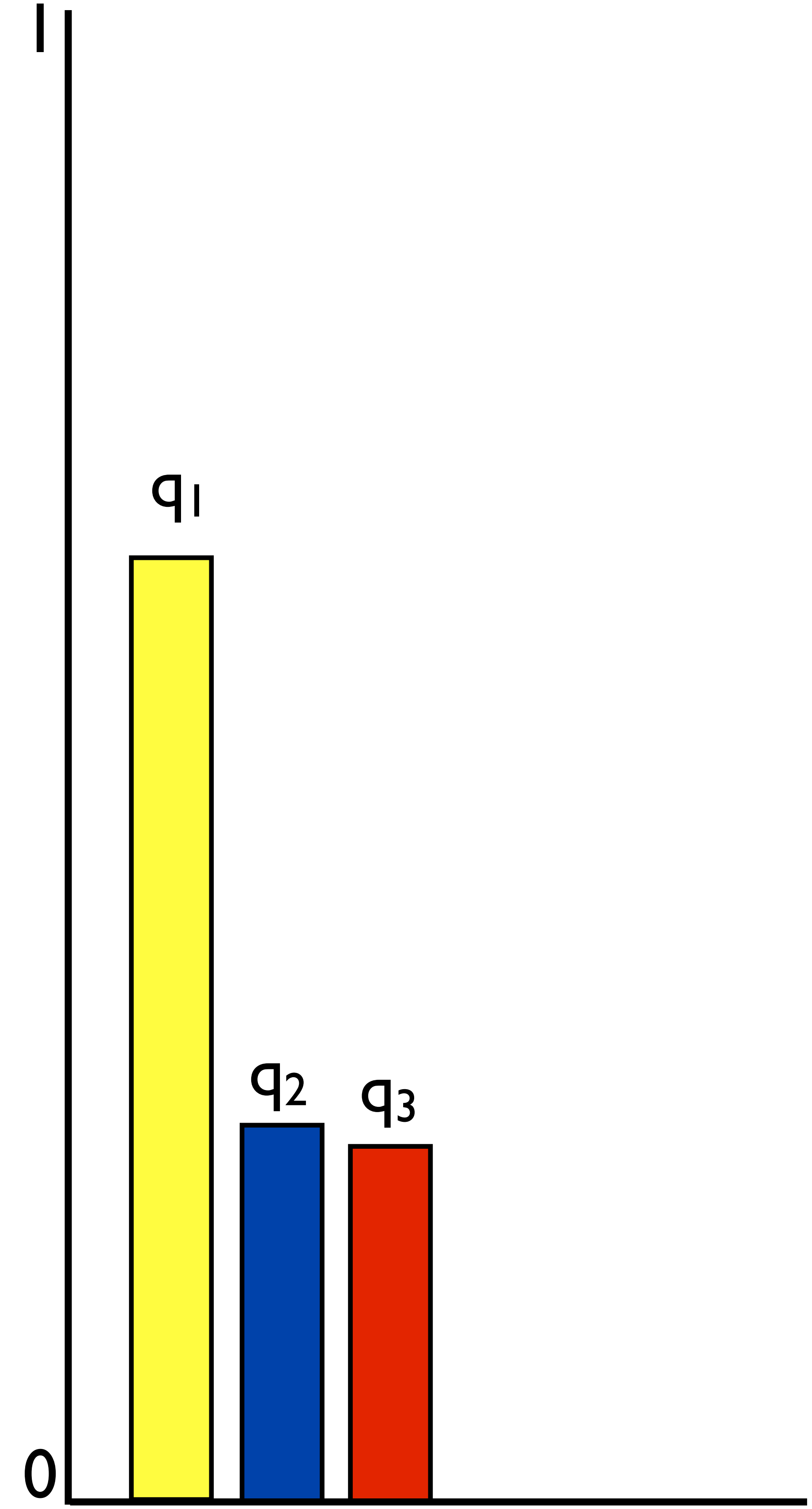
1. Draw $K_m^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + m - 1}\right)$

   Set $K_m = \sum_{j=1}^{m} K_j^+$

2. For k = $K_{m-1}, \ldots, K_m$

   Draw a frequency of size

$$q_k \sim \mathrm{Beta}(1, \theta + m - 1)$$



[Thibaux, Jordan 2007]

Paintboxes

Indian buffet process: beta feature frequencies

For m = 1, 2, ...

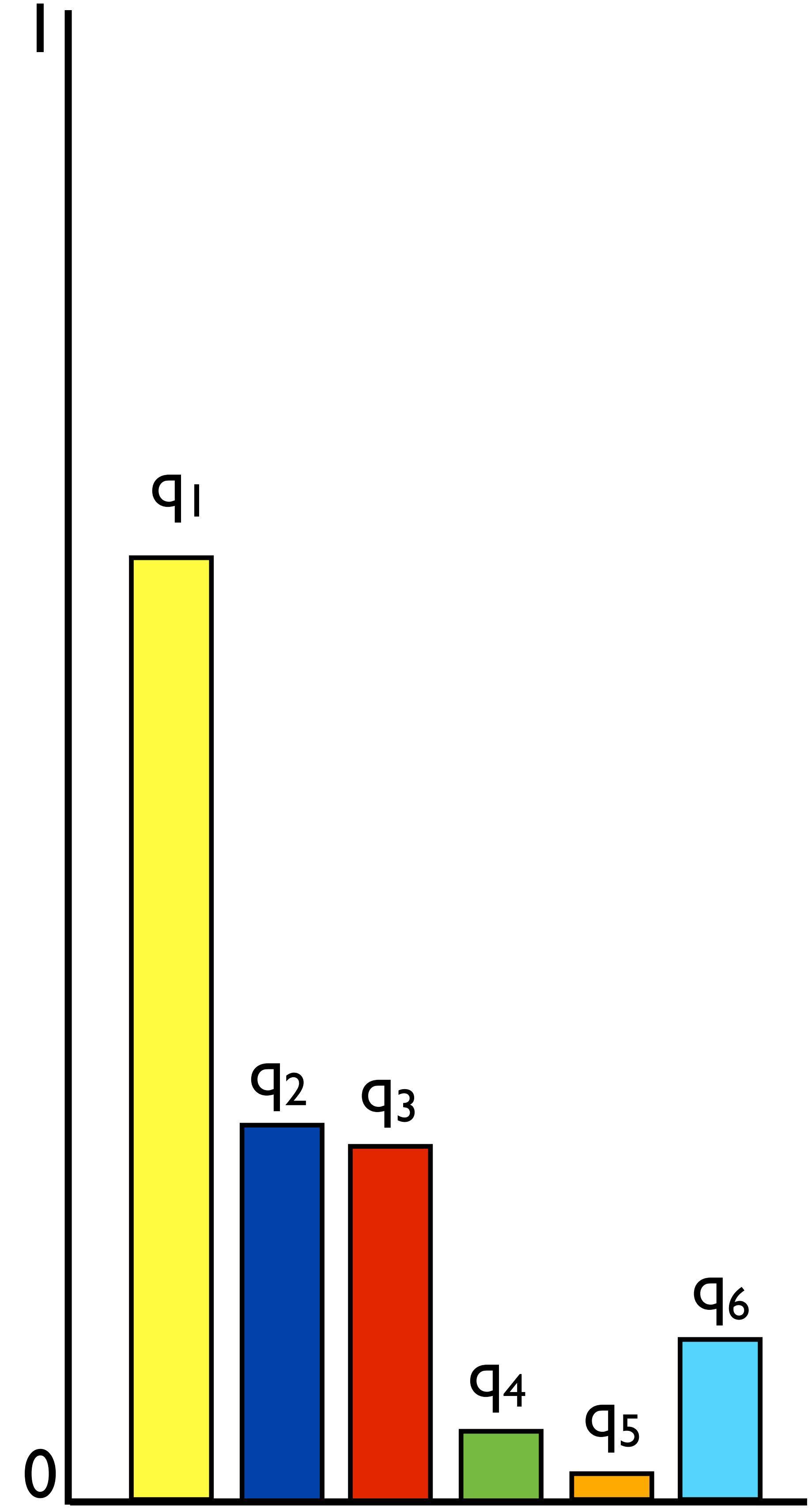
1. Draw $K_m^+ = \text{Poisson}\left(\gamma \frac{\theta}{\theta + m - 1}\right)$

   Set $K_m = \sum_{j=1}^{m} K_j^+$

2. For k = $K_{m-1}, \ldots, K_m$

   Draw a frequency of size

$$q_k \sim \text{Beta}(1, \theta + m - 1)$$



[Thibaux, Jordan 2007]

Paintboxes

Indian buffet process: beta feature frequencies

For *m* = 1, 2, ...

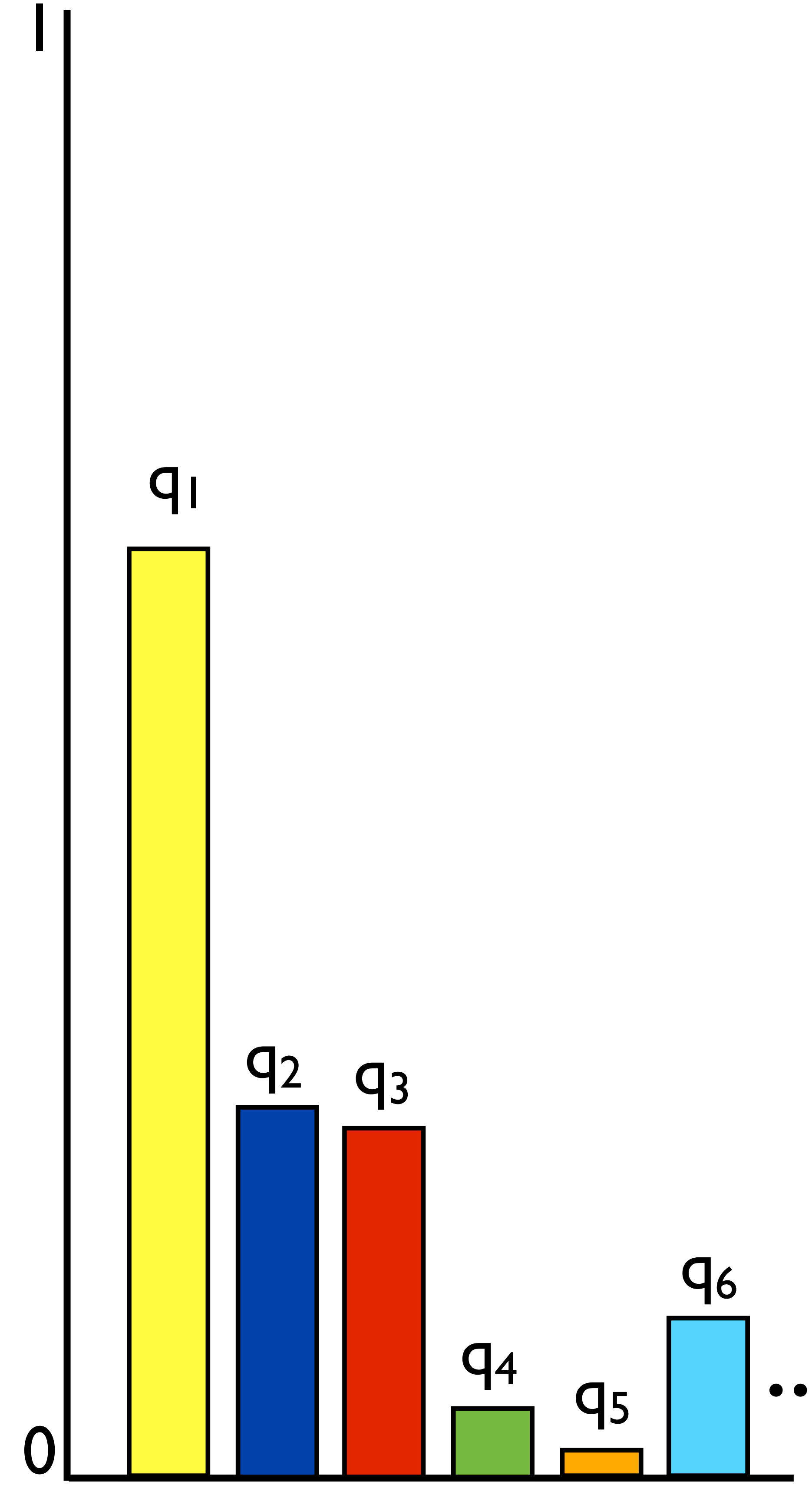
1. Draw $K_m^+ = \mathrm{Poisson}\left(\gamma \dfrac{\theta}{\theta + m - 1}\right)$

   Set $K_m = \sum_{j=1}^{m} K_j^+$

2. For k = $K_{m-1}, \ldots, K_m$

   Draw a frequency of size

$$q_k \sim \mathrm{Beta}(1, \theta + m - 1)$$



[Thibaux, Jordan 2007]

Paintboxes

Indian buffet process: beta feature frequencies

For m = 1, 2, ...

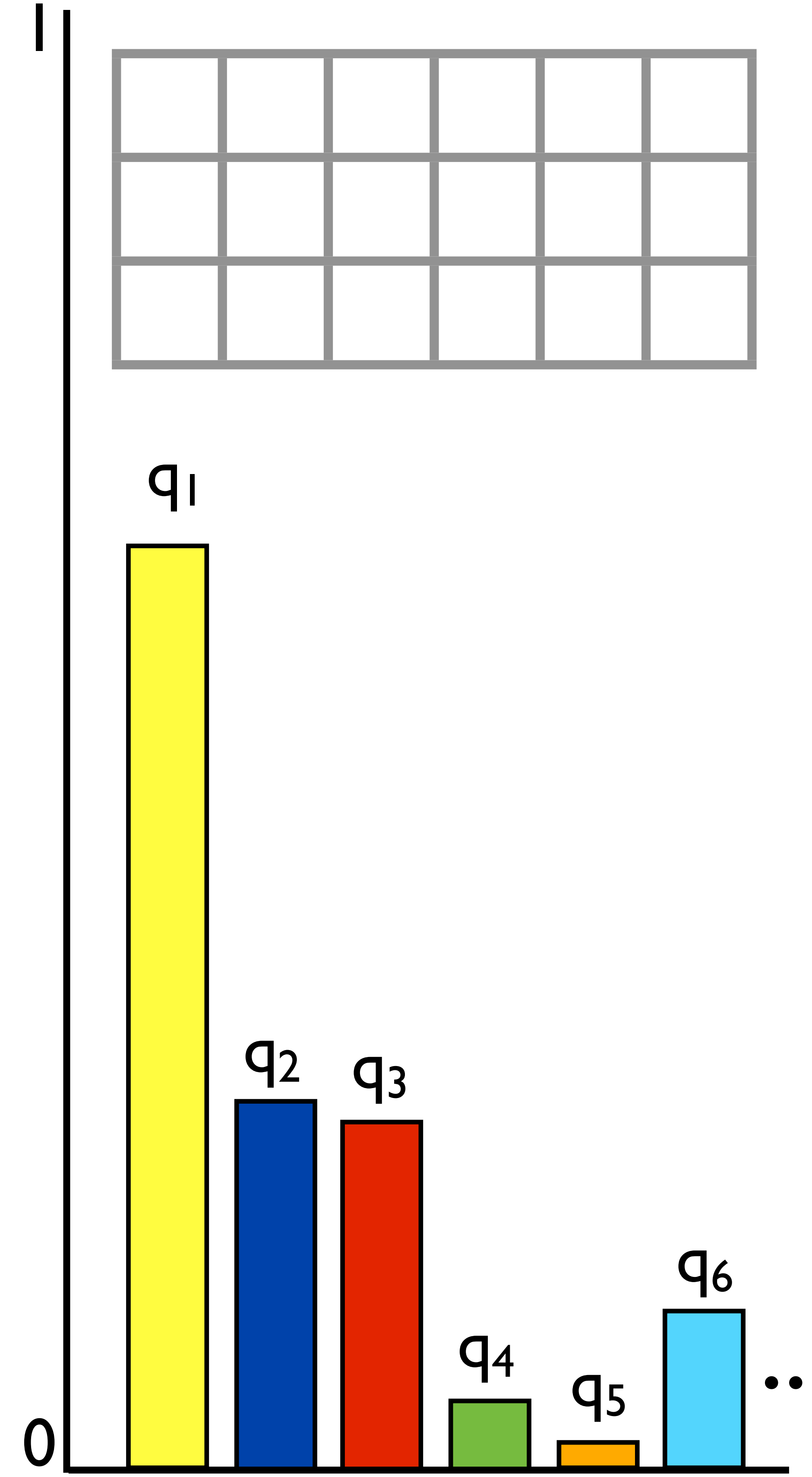
1. Draw $K_m^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + m - 1}\right)$

   Set $K_m = \sum_{j=1}^{m} K_j^+$

2. For k = $K_{m-1}, \ldots, K_m$

   Draw a frequency of size

$$q_k \sim \mathrm{Beta}(1, \theta + m - 1)$$



[Thibaux, Jordan 2007]

Paintboxes

Indian buffet process: beta feature frequencies

For m = 1, 2, ...

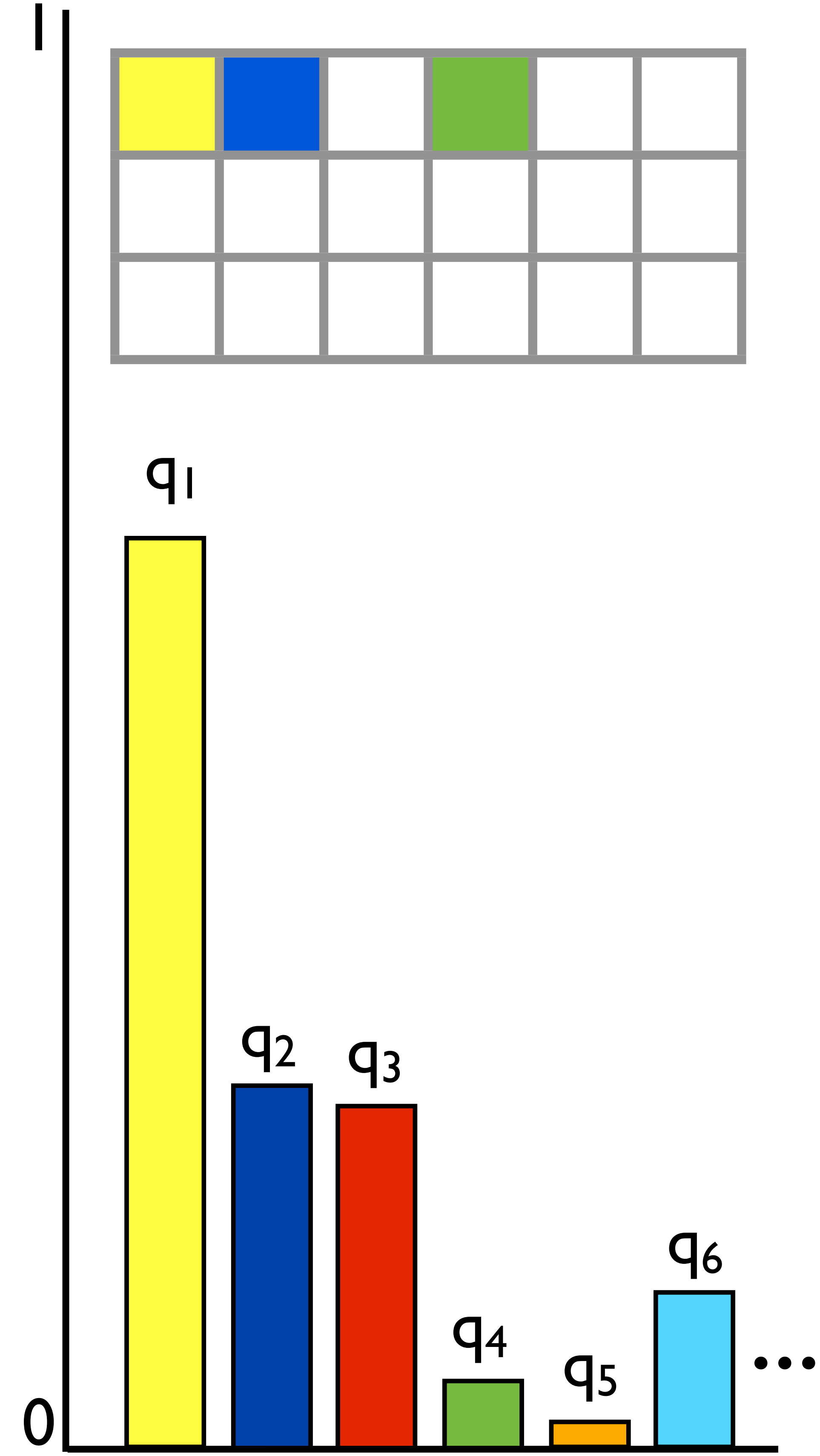
1. Draw $K_m^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + m - 1}\right)$

   Set $K_m = \sum_{j=1}^{m} K_j^+$

2. For k = $K_{m-1}, \ldots, K_m$

   Draw a frequency of size

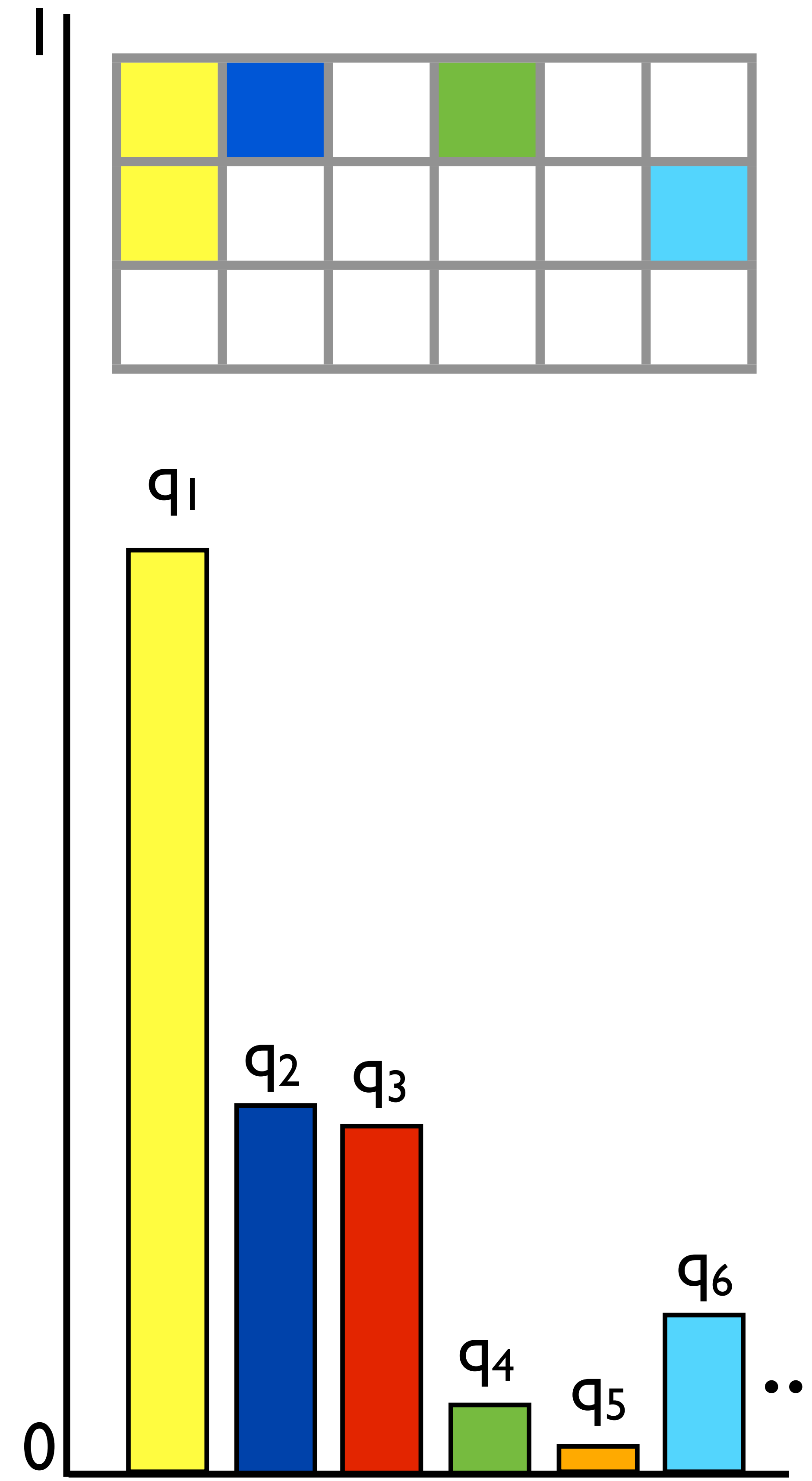
$$q_k \sim \mathrm{Beta}(1, \theta + m - 1)$$



[Thibaux, Jordan 2007]

Paintboxes

Indian buffet process: beta feature frequencies



For *m* = 1, 2, ...

1. Draw $K_m^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + m - 1}\right)$

   Set $K_m = \sum_{j=1}^{m} K_j^+$

2. For k = $K_{m-1}, \ldots, K_m$

   Draw a frequency of size

$$q_k \sim \mathrm{Beta}(1, \theta + m - 1)$$

[Thibaux, Jordan 2007]

Paintboxes

Indian buffet process: beta feature frequencies

For m = 1, 2, ...

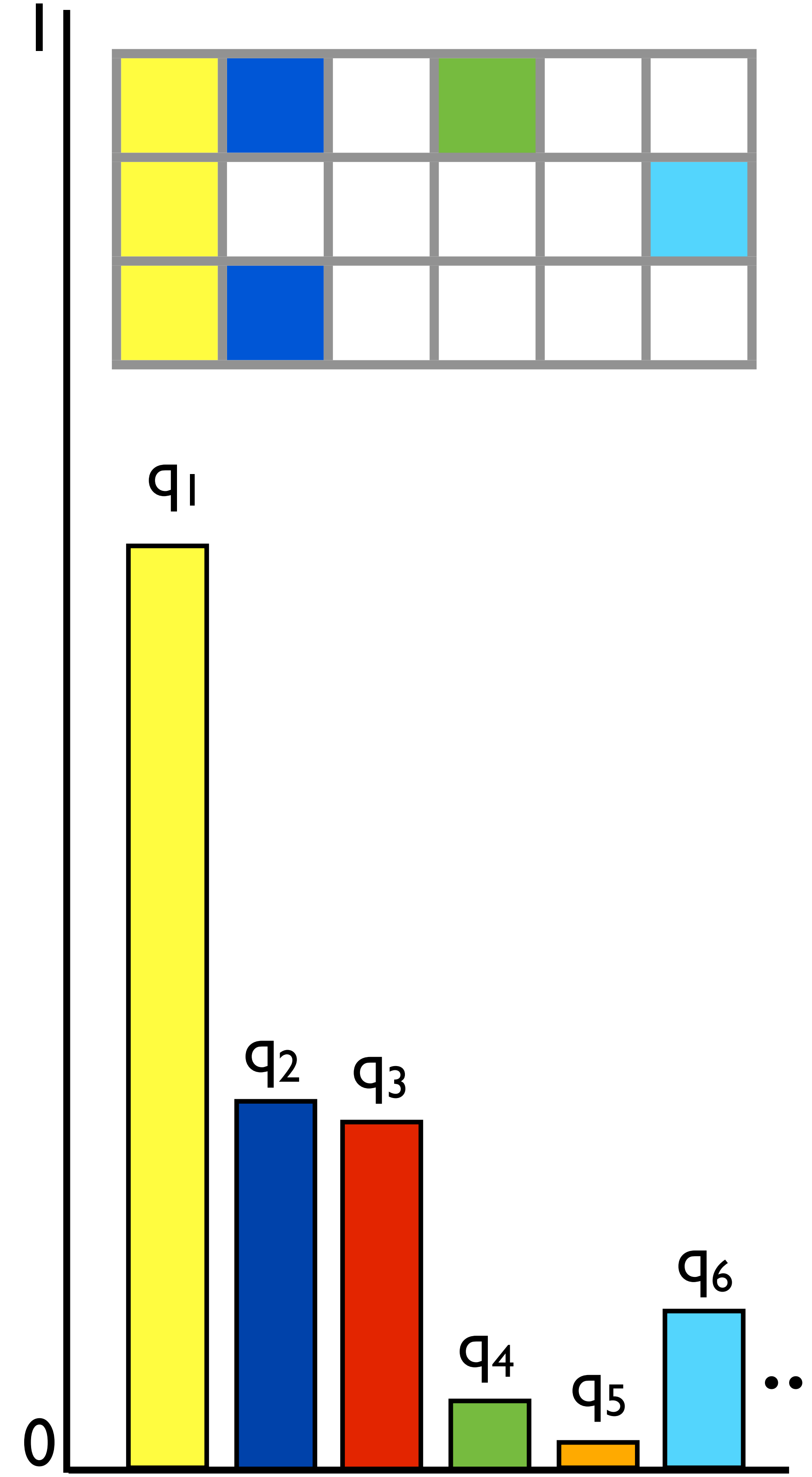
1. Draw $K_m^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + m - 1}\right)$

   Set $K_m = \sum_{j=1}^{m} K_j^+$

2. For k = $K_{m-1}, \ldots, K_m$

   Draw a frequency of size

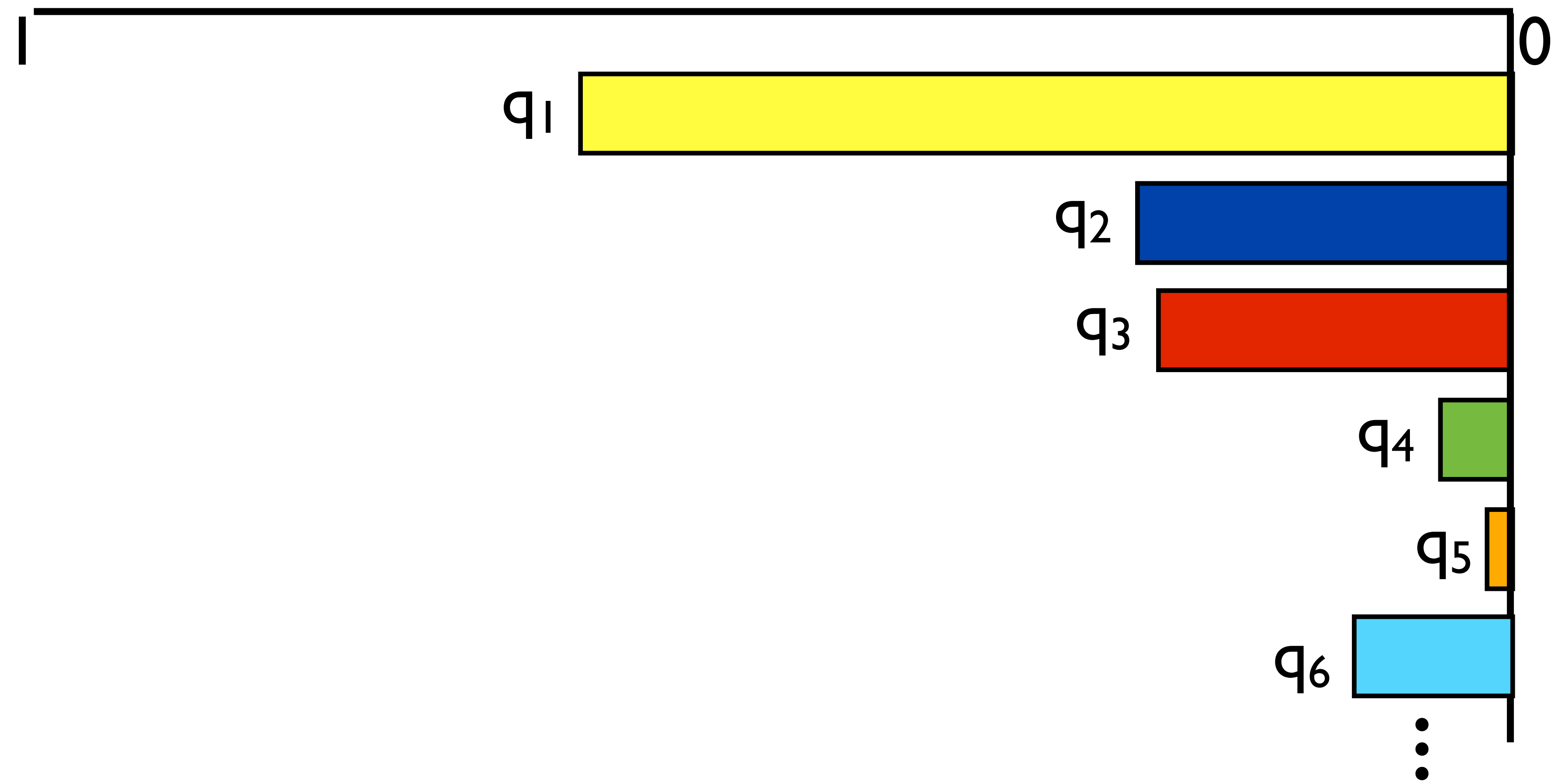
$$q_k \sim \mathrm{Beta}(1, \theta + m - 1)$$



[Thibaux, Jordan 2007]

Paintboxes

Indian buffet process: beta feature frequencies

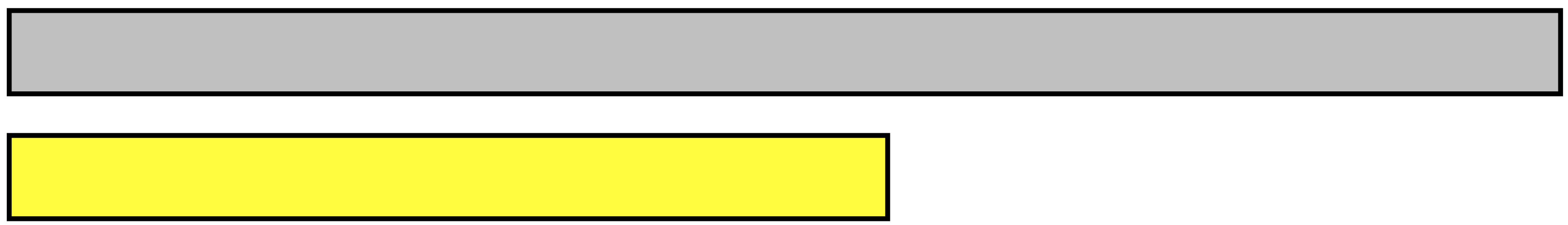


Paintboxes

Indian buffet process: beta feature frequencies

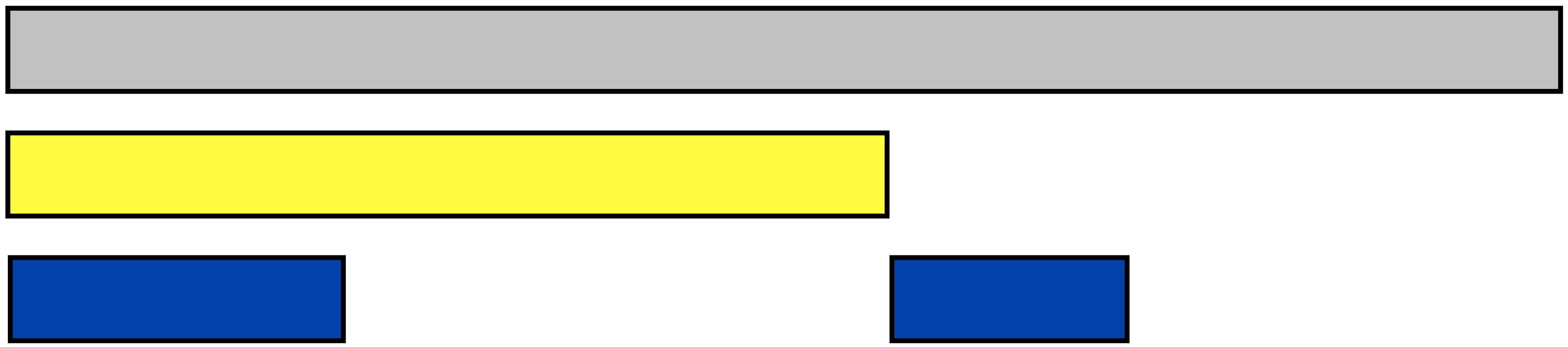
Paintboxes

Indian buffet process: beta feature frequencies



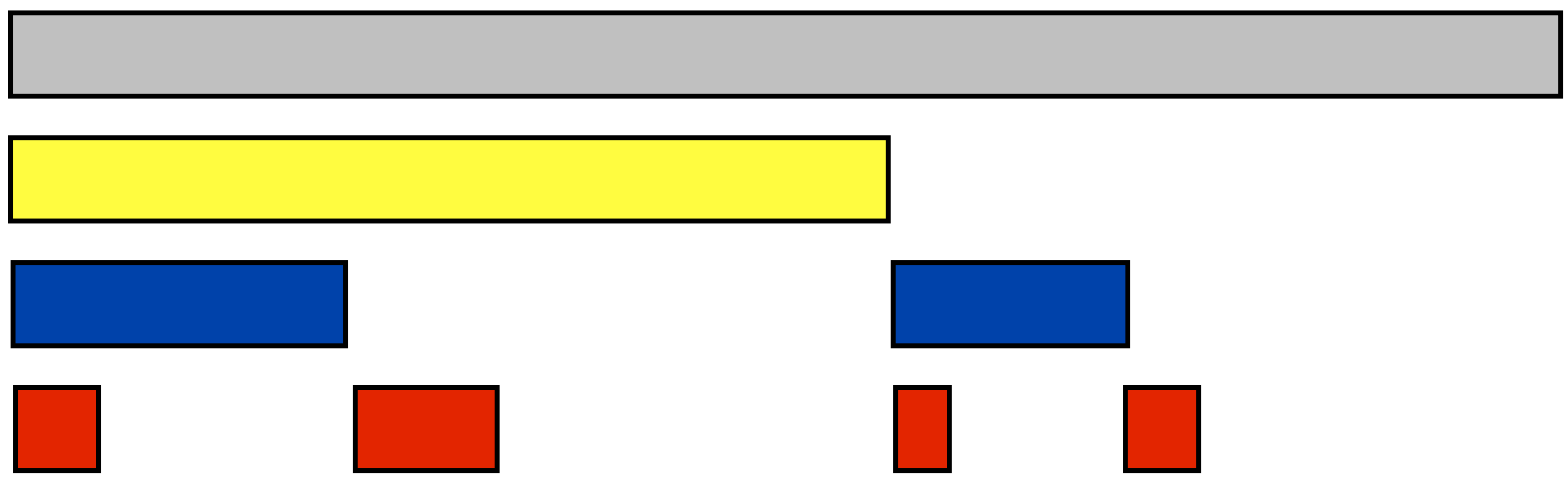
Paintboxes

Indian buffet process: beta feature frequencies



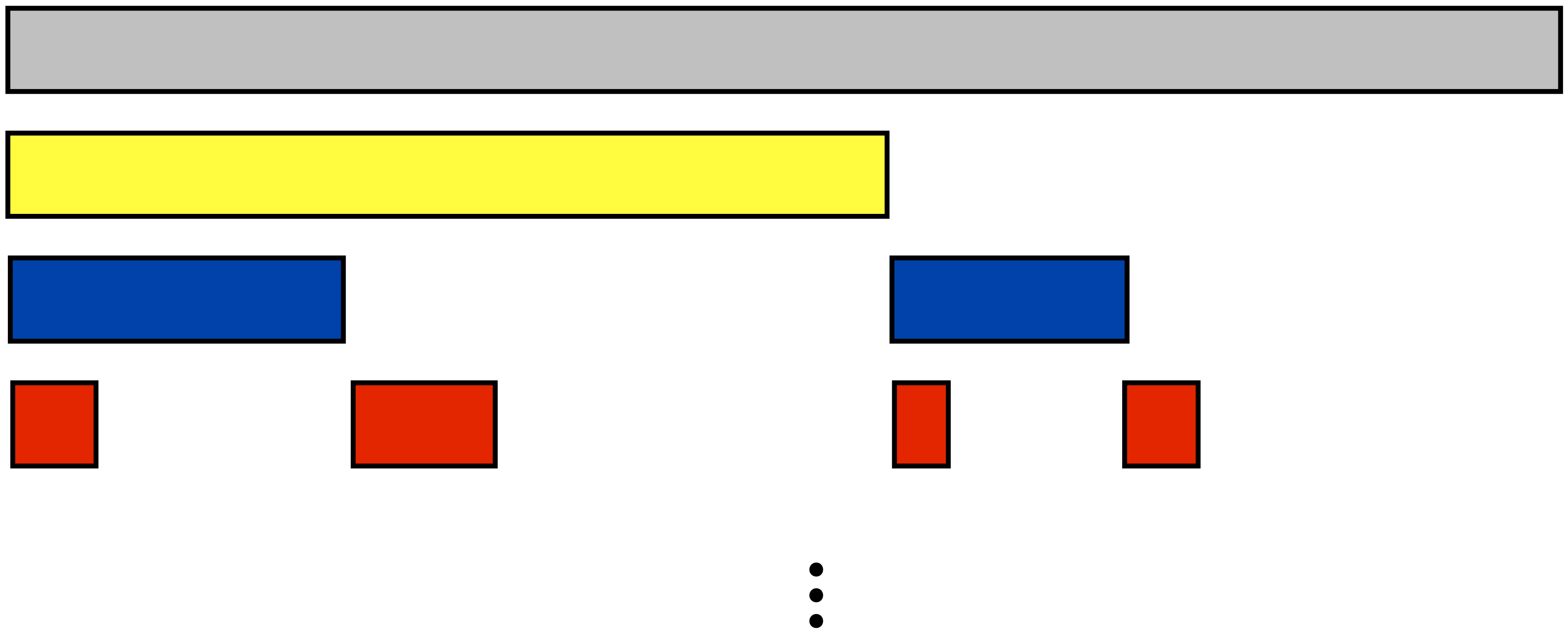
Paintboxes

Indian buffet process: beta feature frequencies

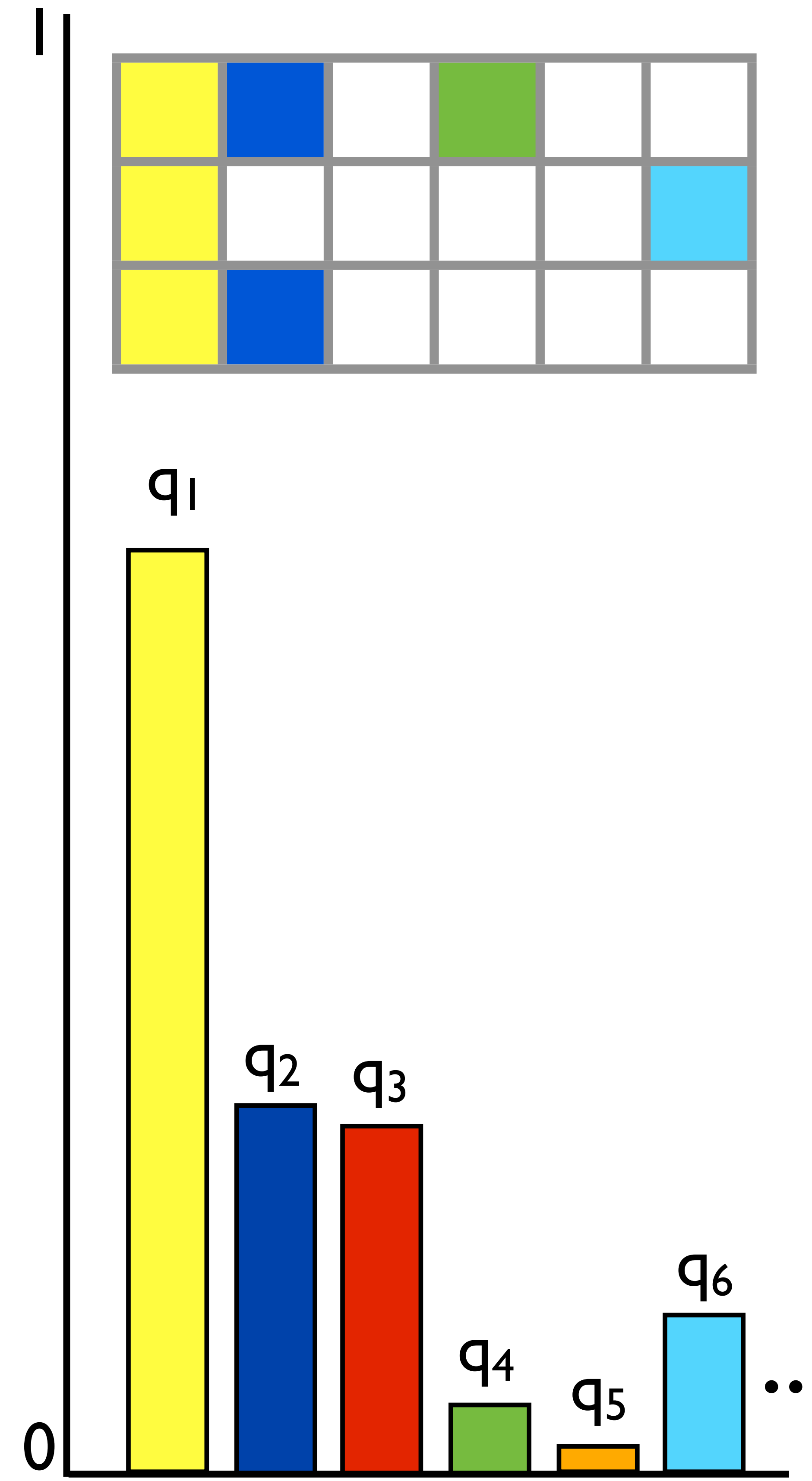


Paintboxes

Indian buffet process: beta feature frequencies

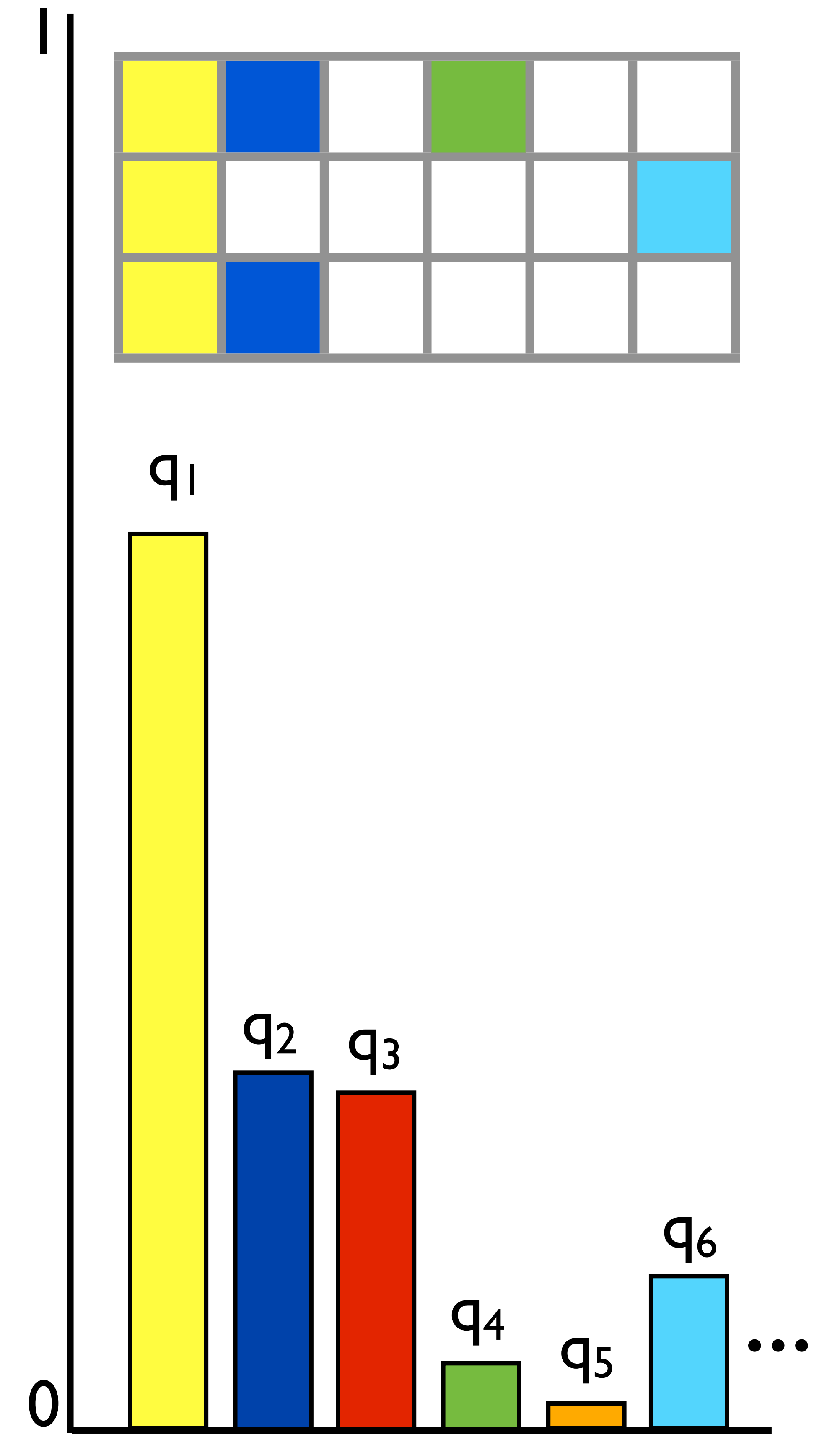


Paintboxes



Paintboxes

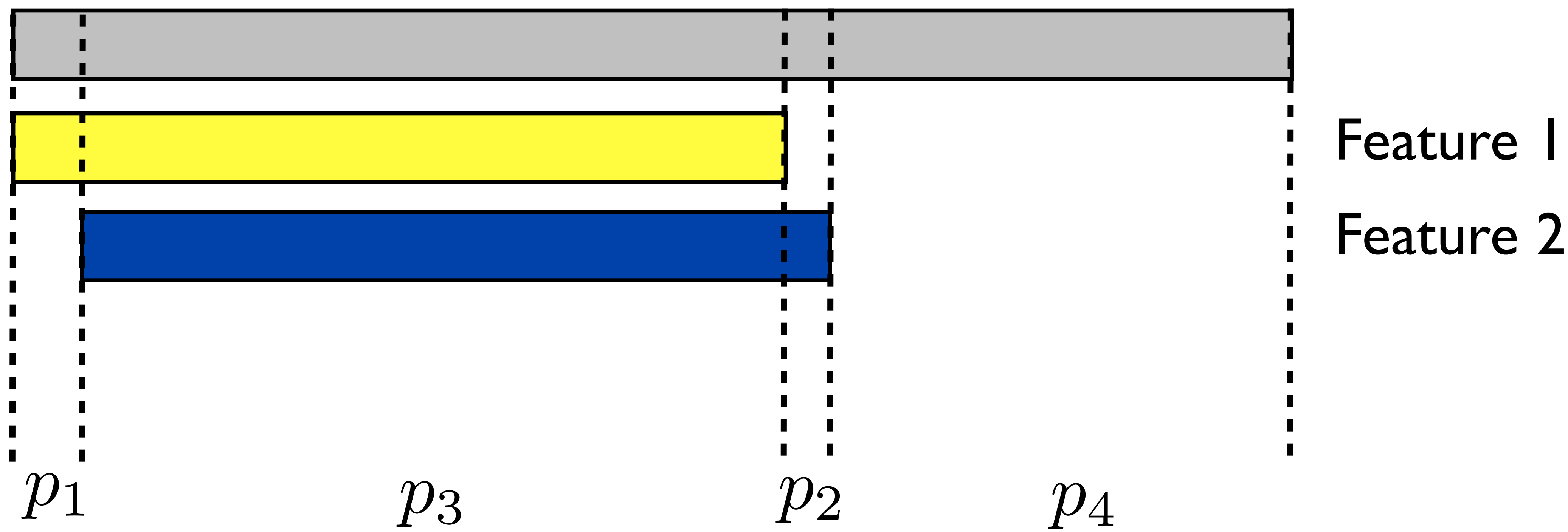
“Feature frequency models”



[Broderick, Pitman, Jordan 2013]

Paintboxes

Two feature example



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

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

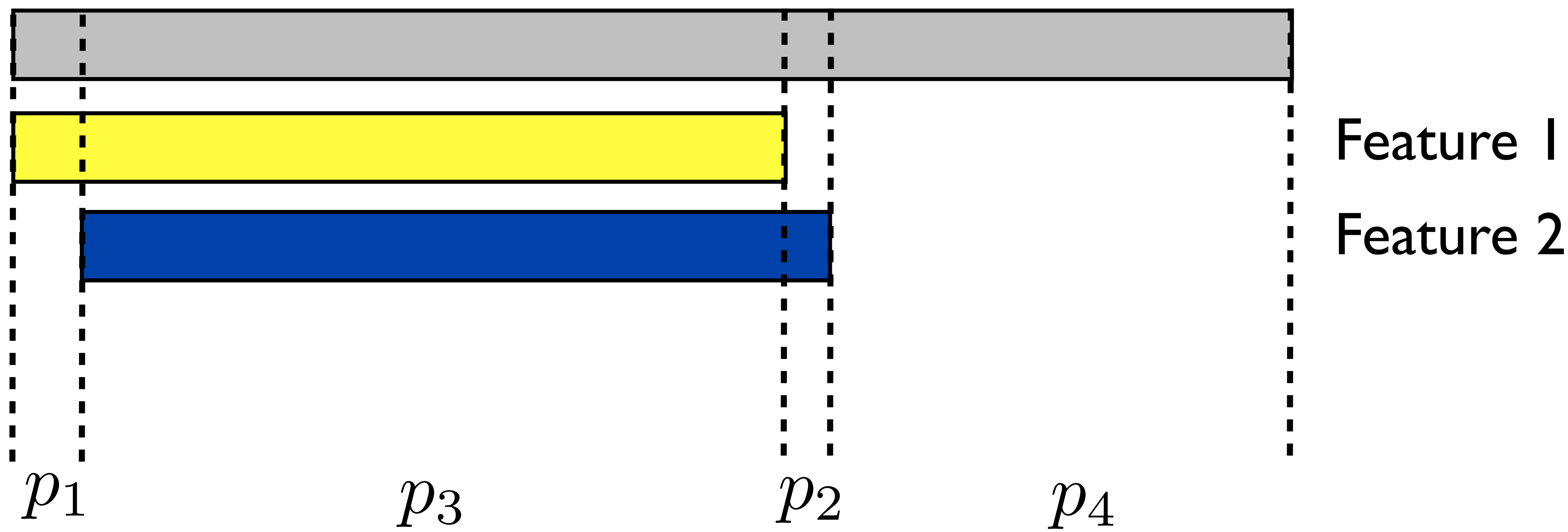
$$\mathbb{P}(\text{row} = \blacksquare\blacksquare) = p_3$$

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

Paintboxes

Two feature example

Not a feature frequency model



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

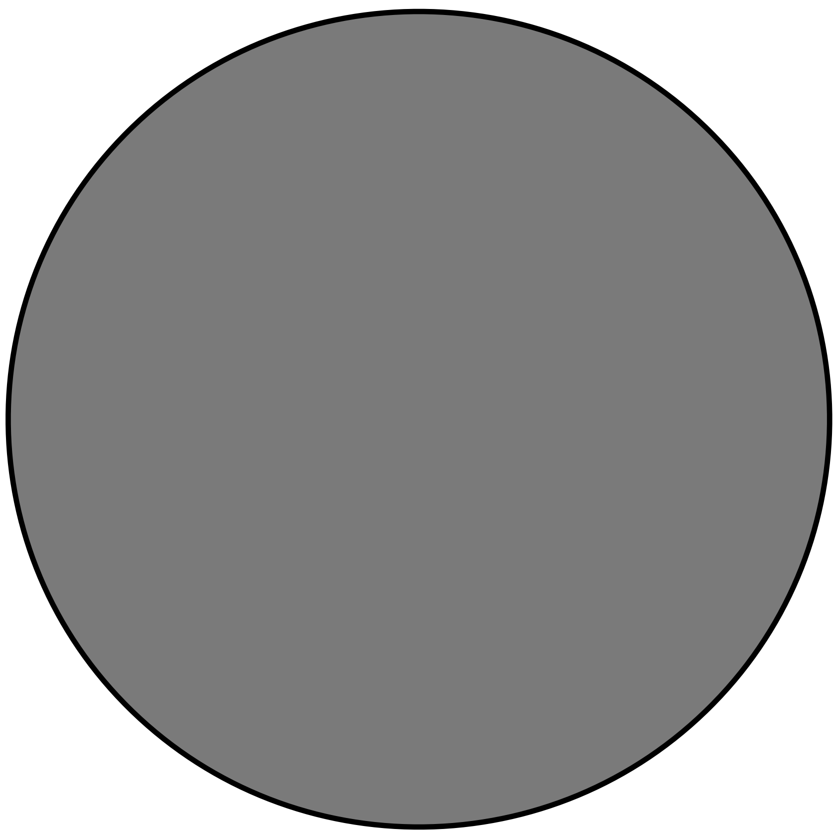
$$\mathbb{P}(\text{row} = \square\blacksquare) = p_2$$

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

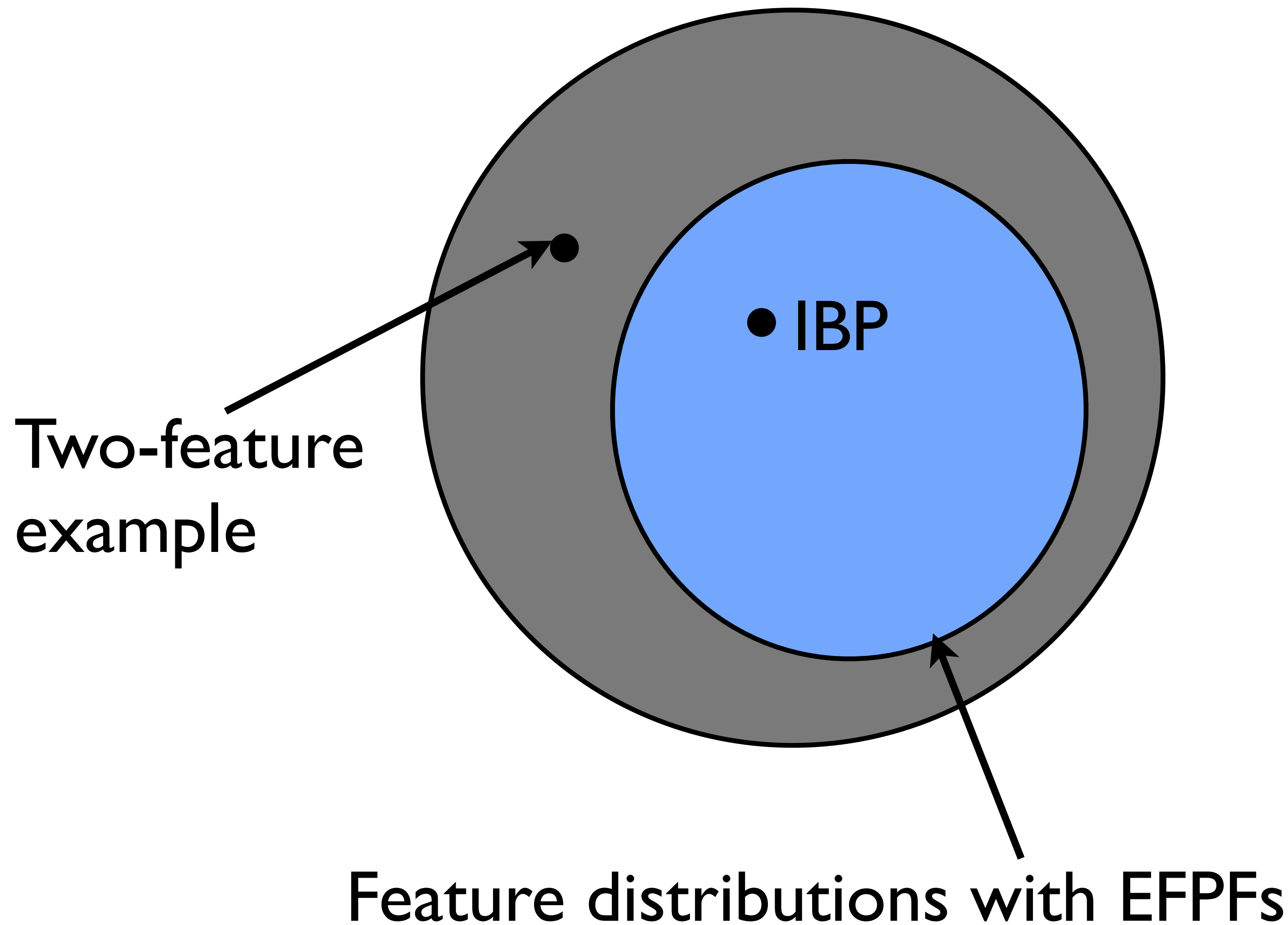
$$\mathbb{P}(\text{row} = \square\square) = p_4$$

Paintboxes

Exchangeable cluster distributions
= Cluster distributions with EPPFs
= Kingman paintbox partitions



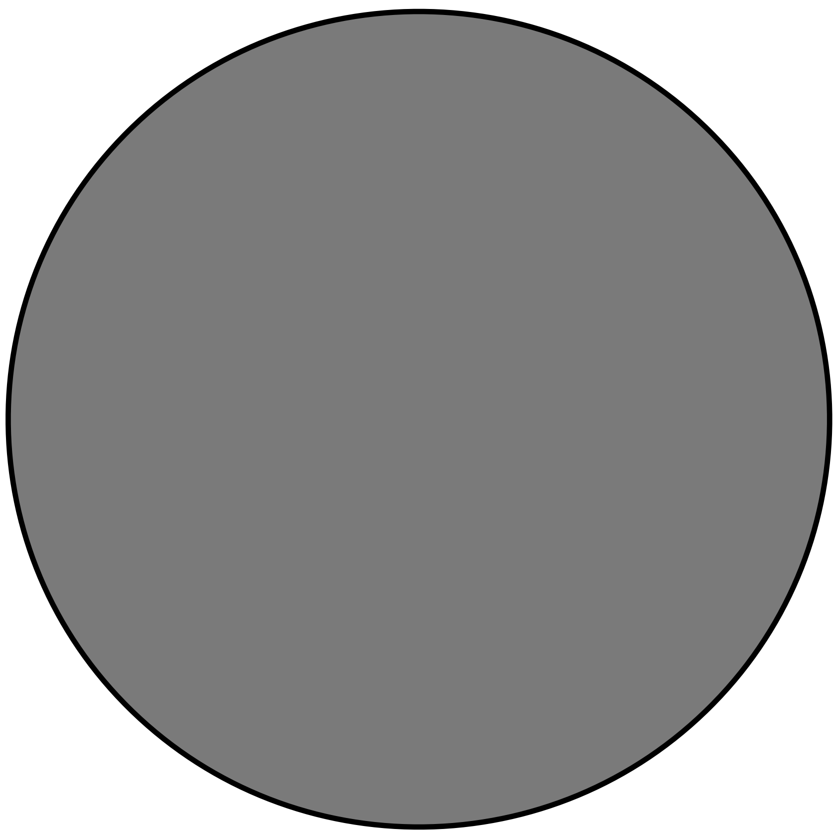
Exchangeable feature distributions
= Feature paintbox allocations



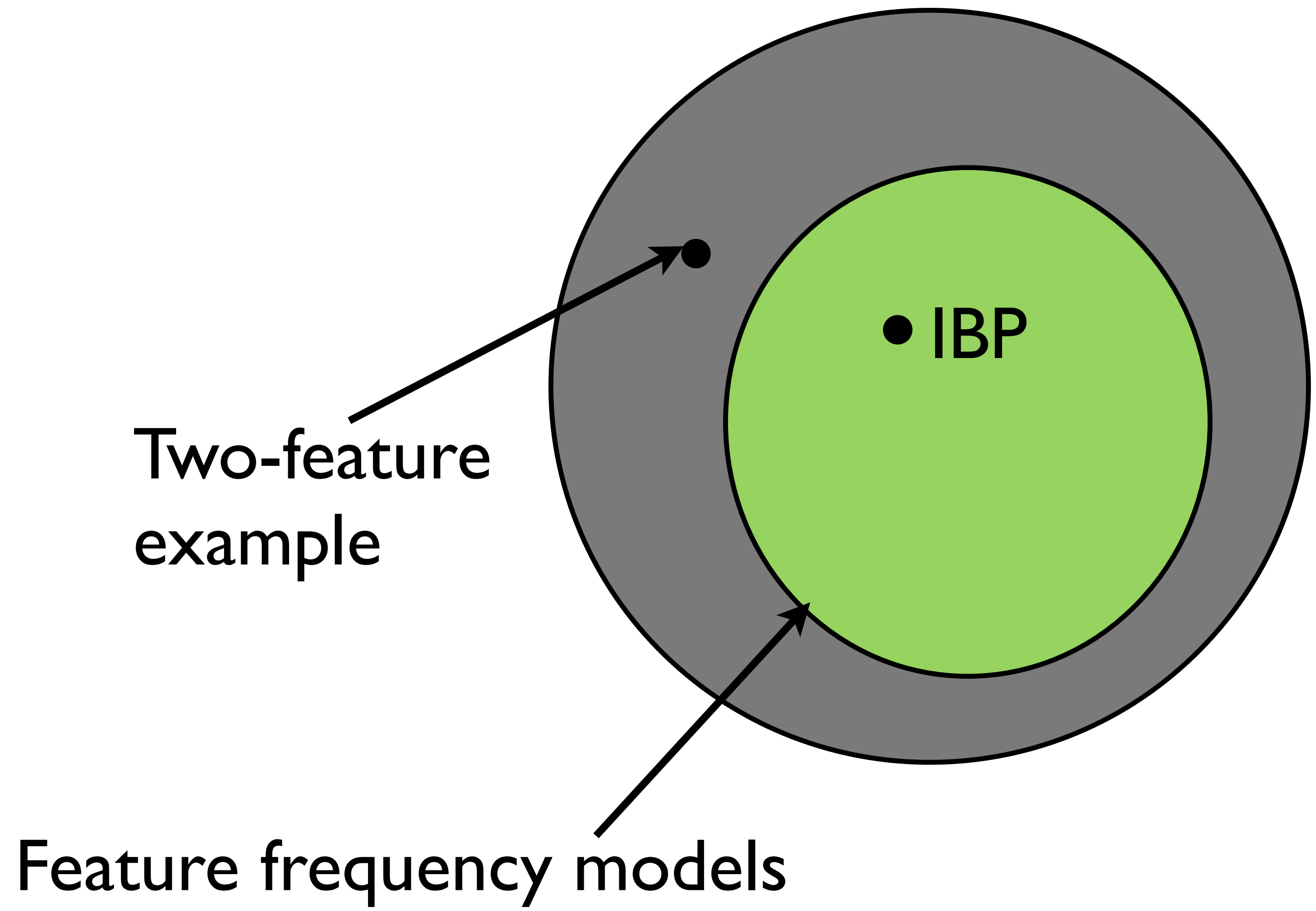
[Broderick, Pitman, Jordan 2013]

Feature frequency models: EFPFs?

Exchangeable cluster distributions
= Cluster distributions with EPPFs
= Kingman paintbox partitions



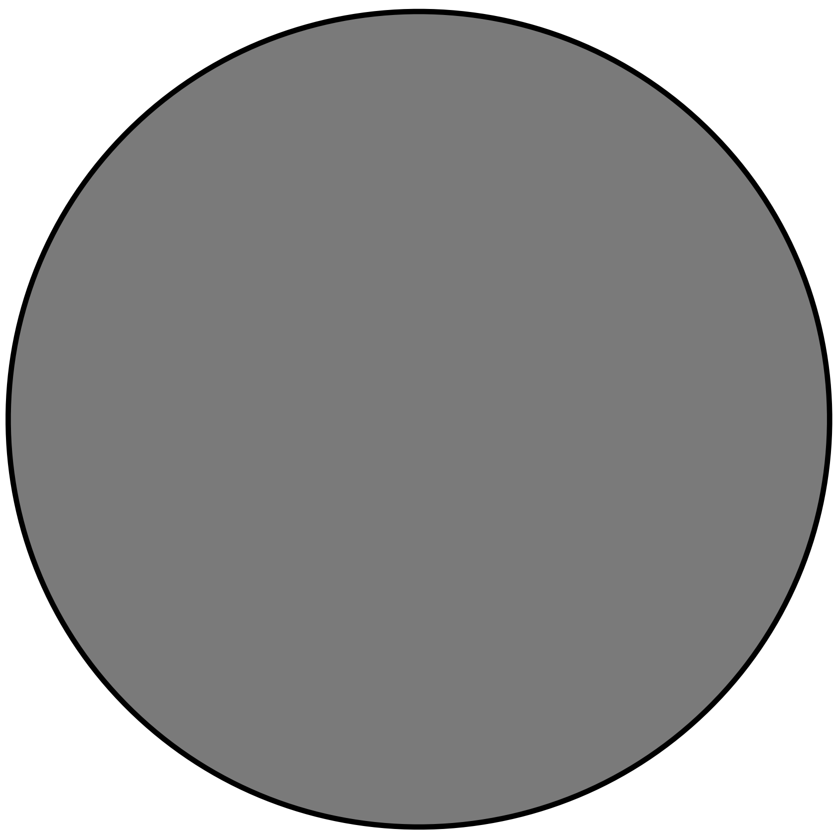
Exchangeable feature distributions
= Feature paintbox allocations



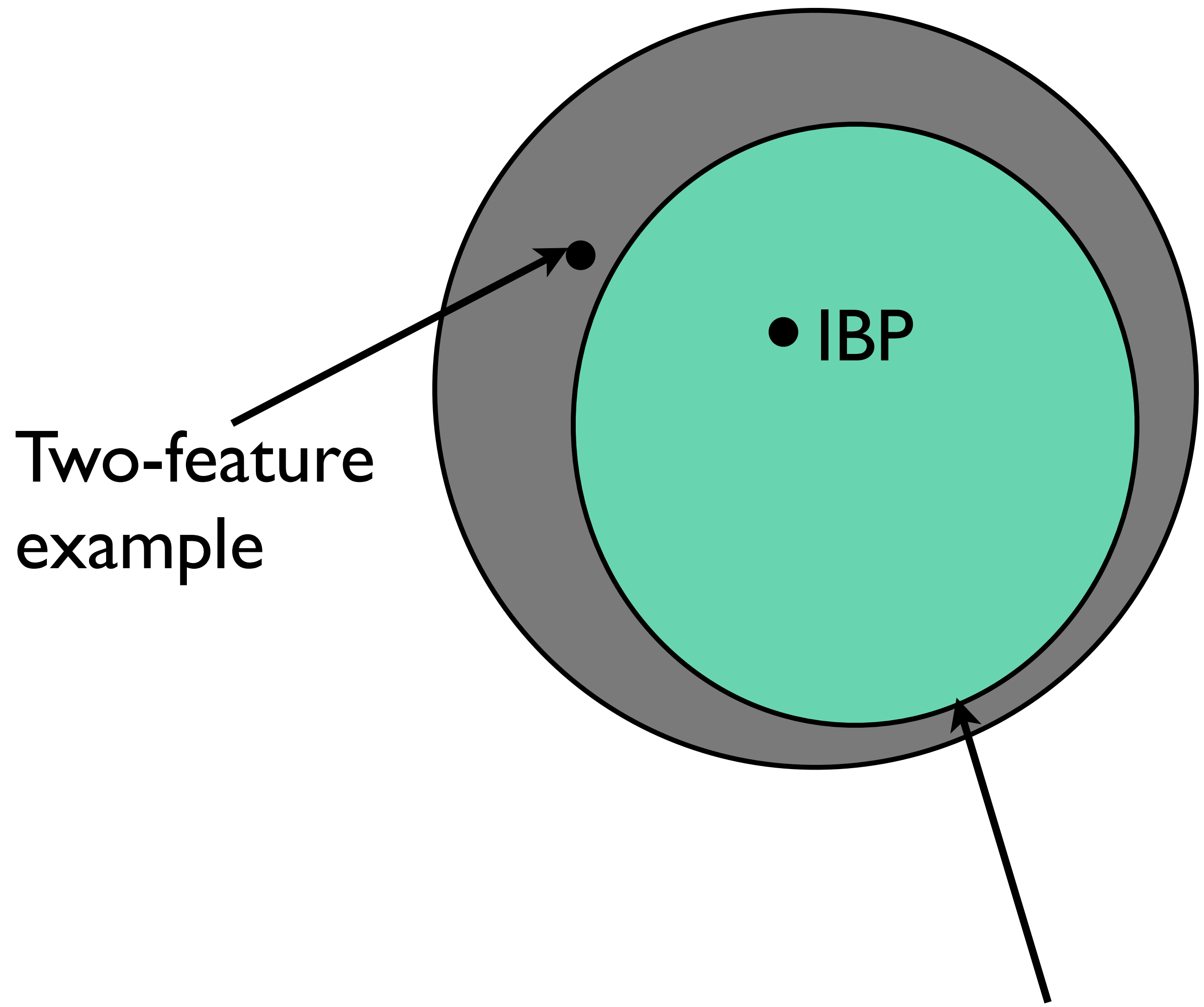
[Broderick, Pitman, Jordan 2013]

Distributions with EFPFs: frequencies?

Exchangeable cluster distributions
= Cluster distributions with EPPFs
= Kingman paintbox partitions



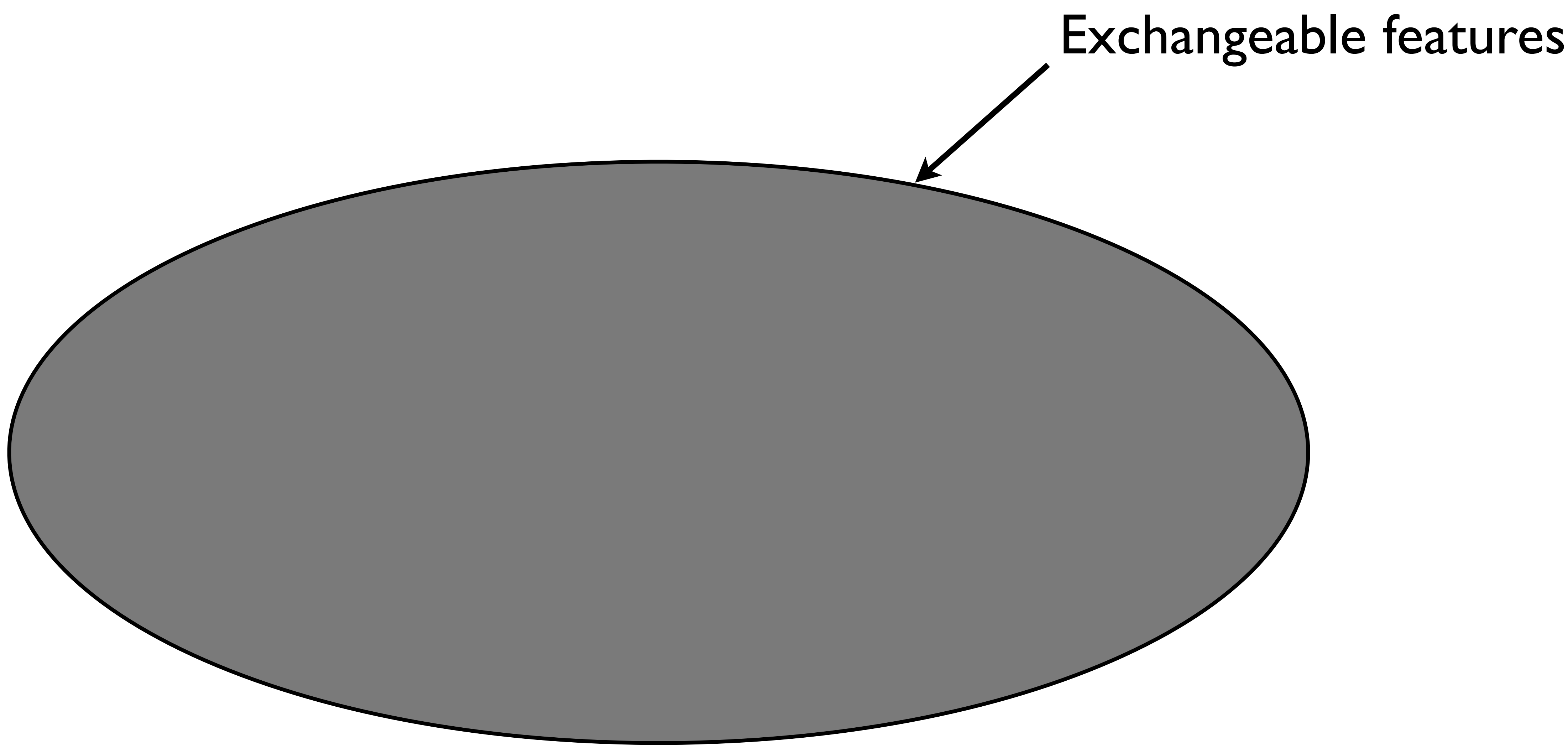
Exchangeable feature distributions
= Feature paintbox allocations



Feature distributions with EFPFs
= Feature frequency models

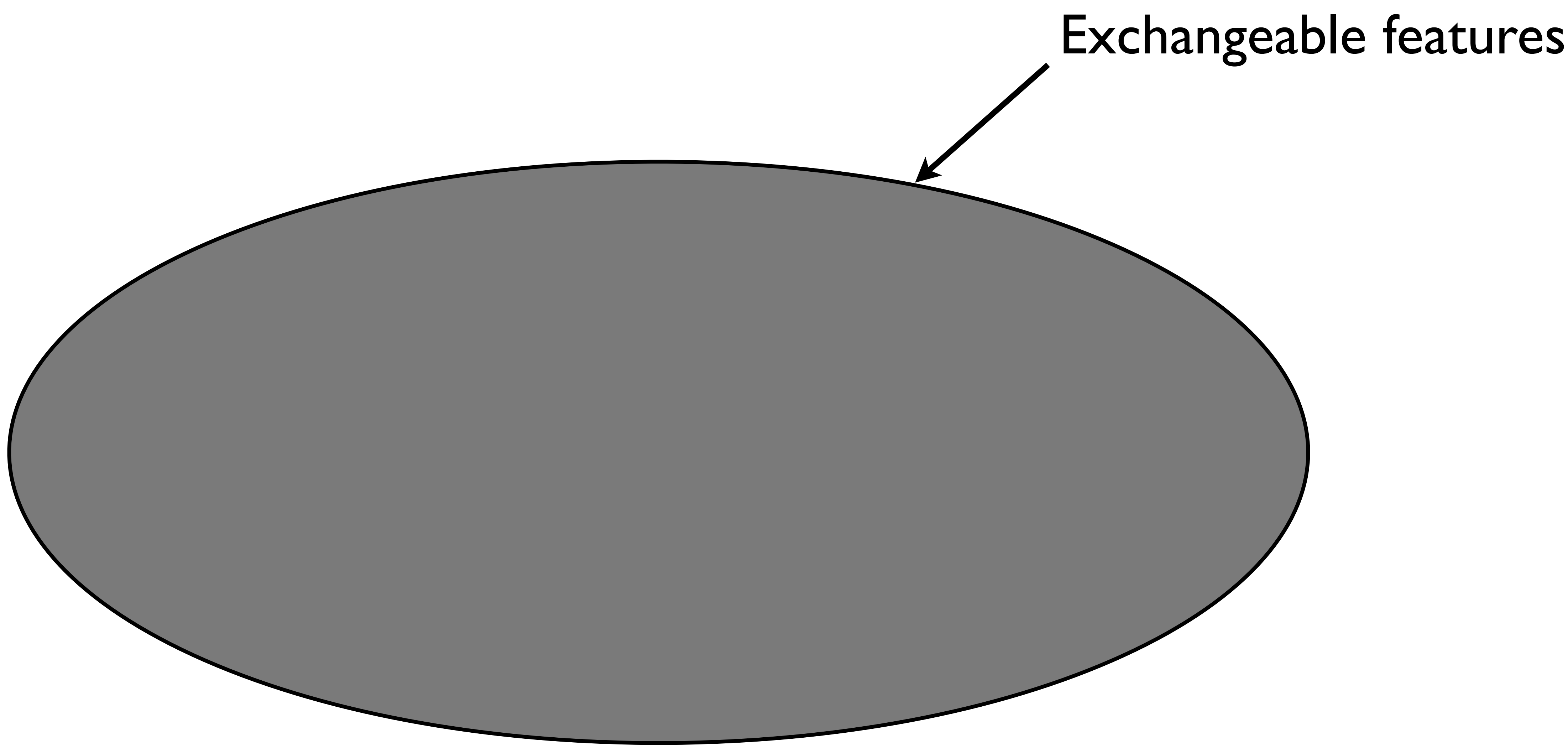
[Broderick, Pitman, Jordan 2013]

Conclusions



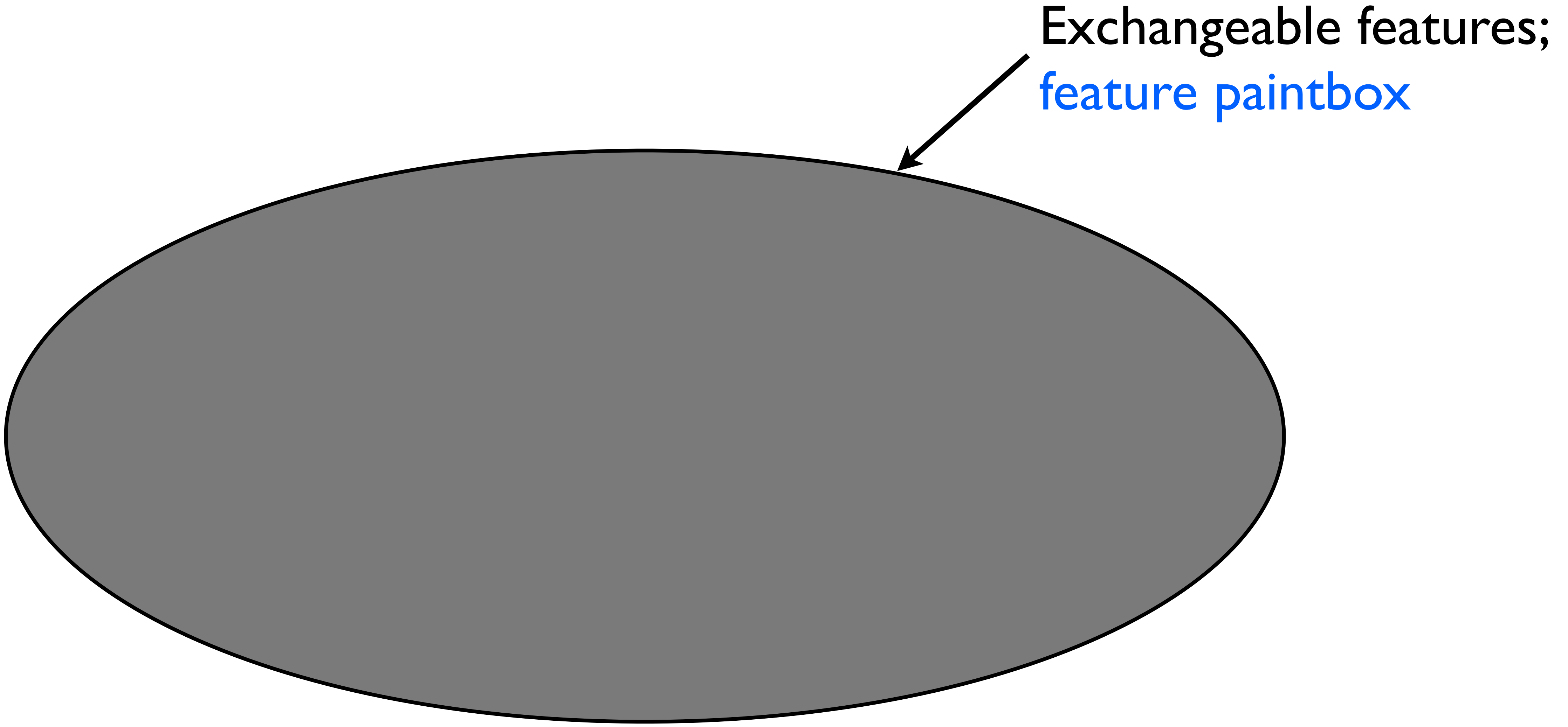
Conclusions

- Feature paintbox: characterization of exchangeable feature models



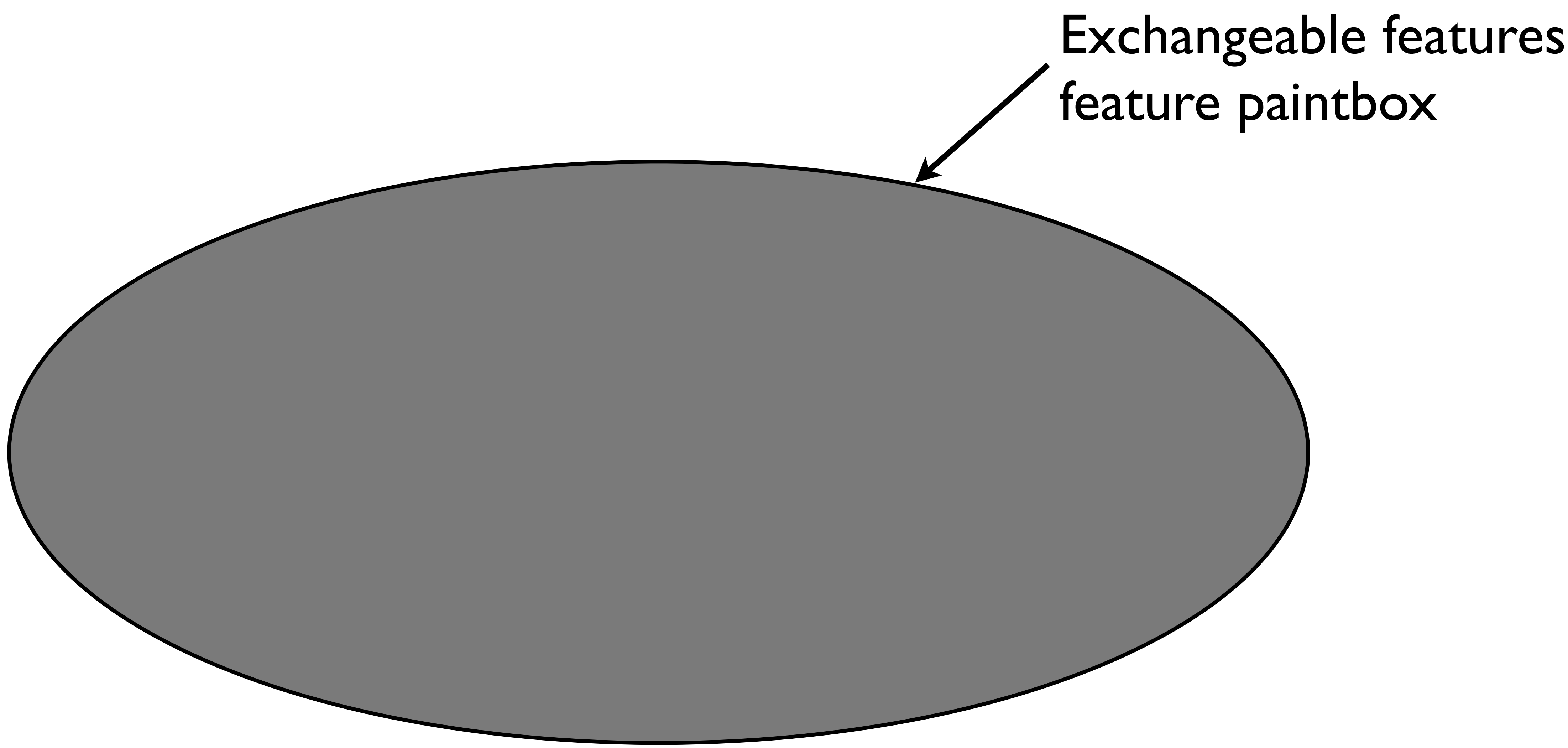
Conclusions

- Feature paintbox: characterization of exchangeable feature models



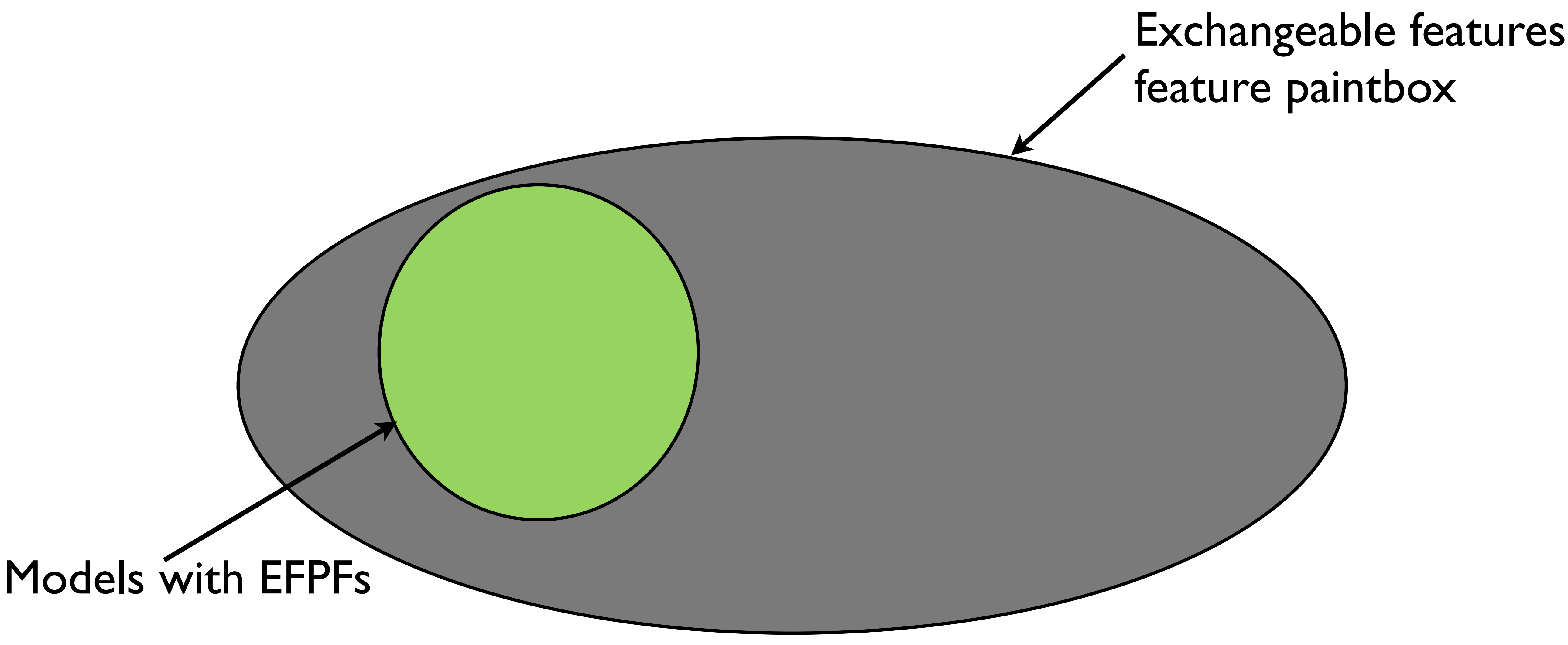
Conclusions

- Feature paintbox: characterization of exchangeable feature models
- Limits of clustering characterizations in feature case?



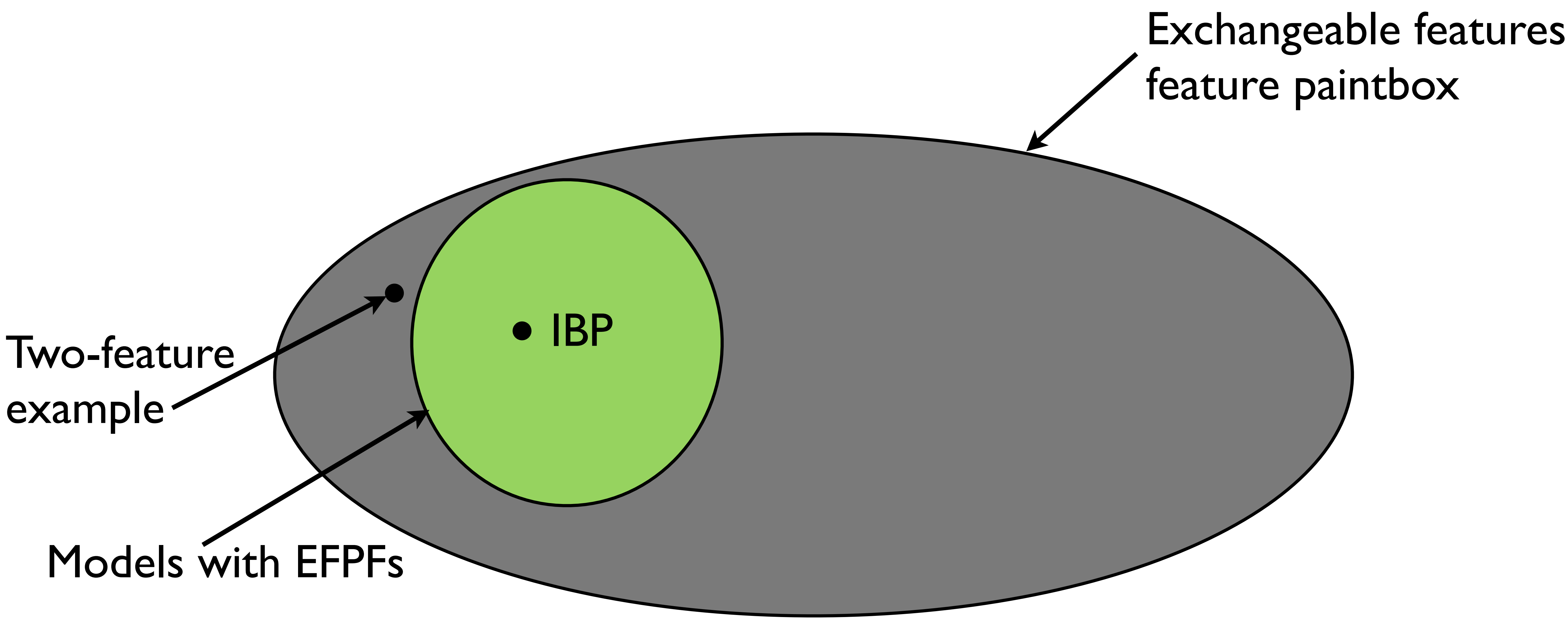
Conclusions

- Feature paintbox: characterization of exchangeable feature models
- Limits of clustering characterizations in feature case?



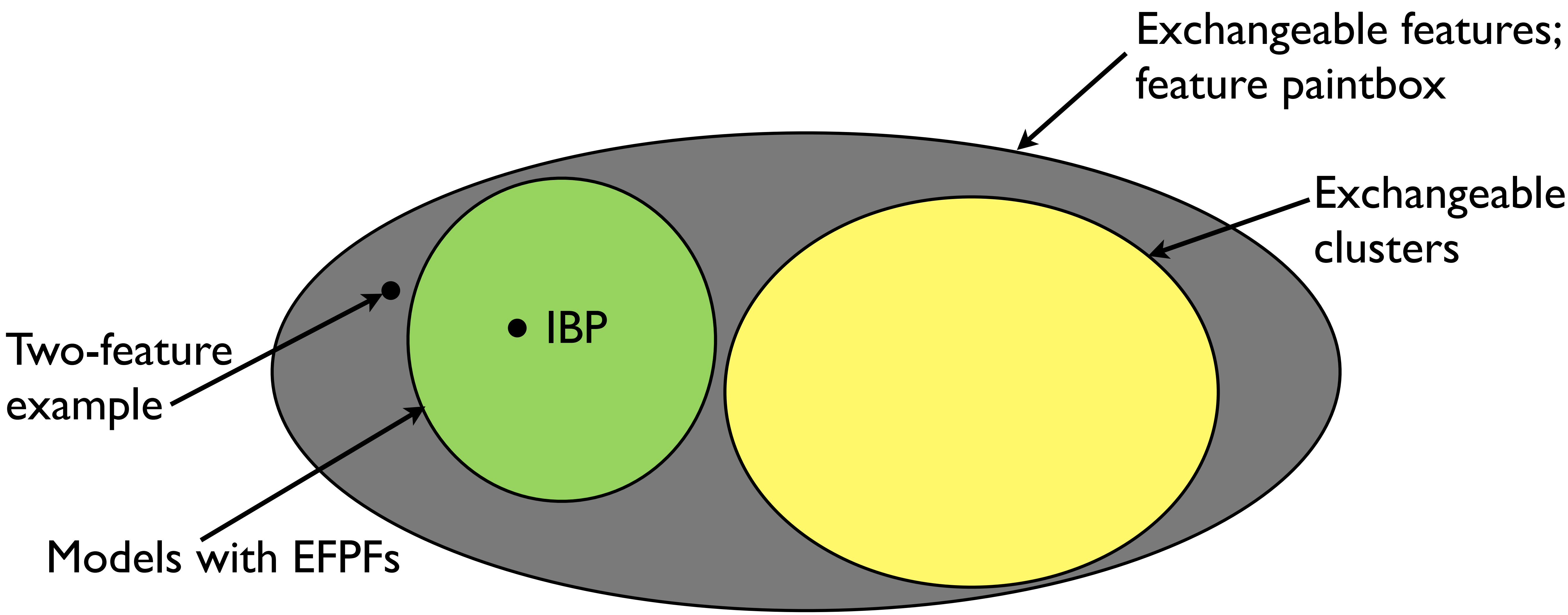
Conclusions

- Feature paintbox: characterization of exchangeable feature models
- Limits of clustering characterizations in feature case?



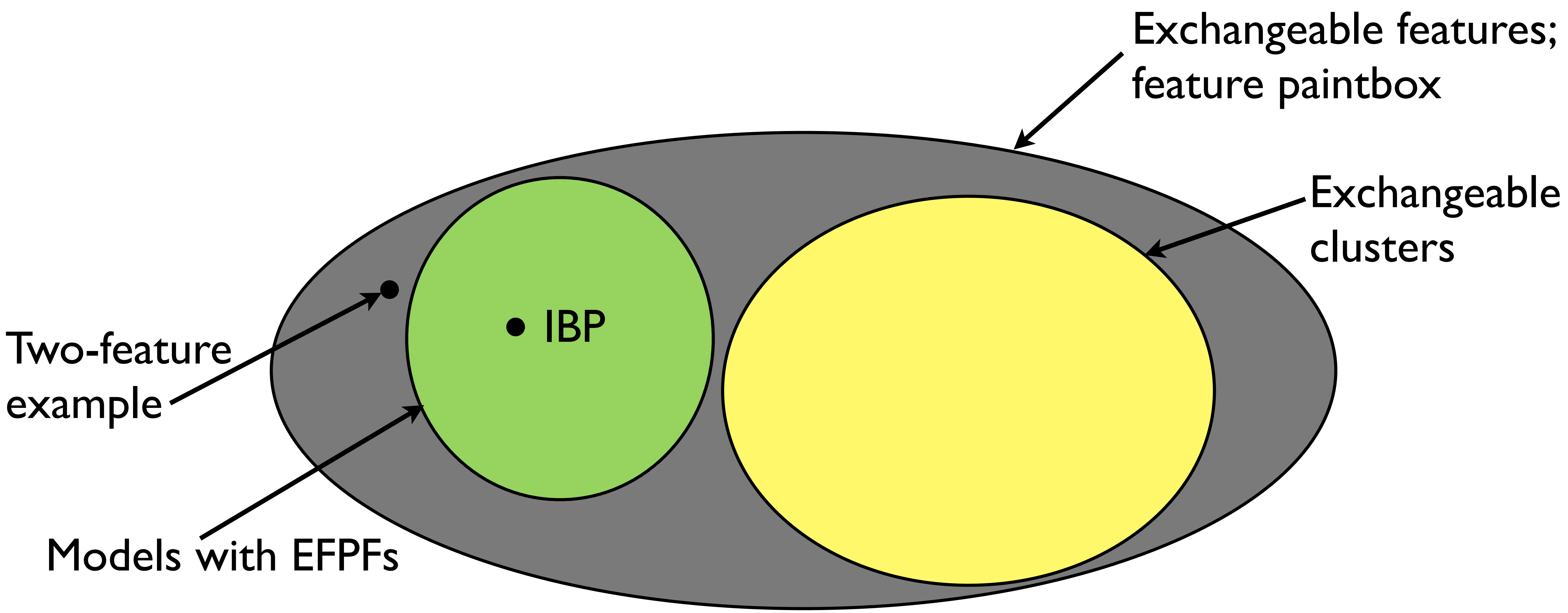
Conclusions

- Feature paintbox: characterization of exchangeable feature models
- Limits of clustering characterizations in feature case?



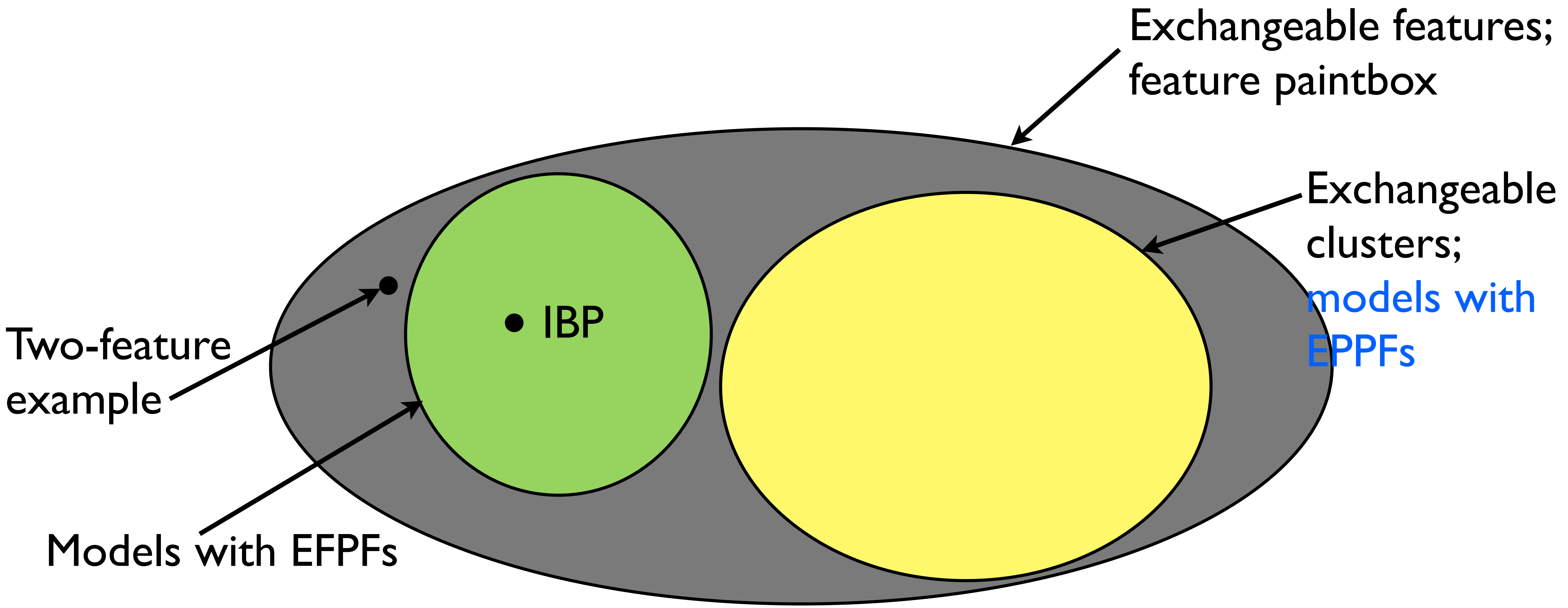
Conclusions

- Feature paintbox: characterization of exchangeable feature models
- Characterization of alternative correlation structure



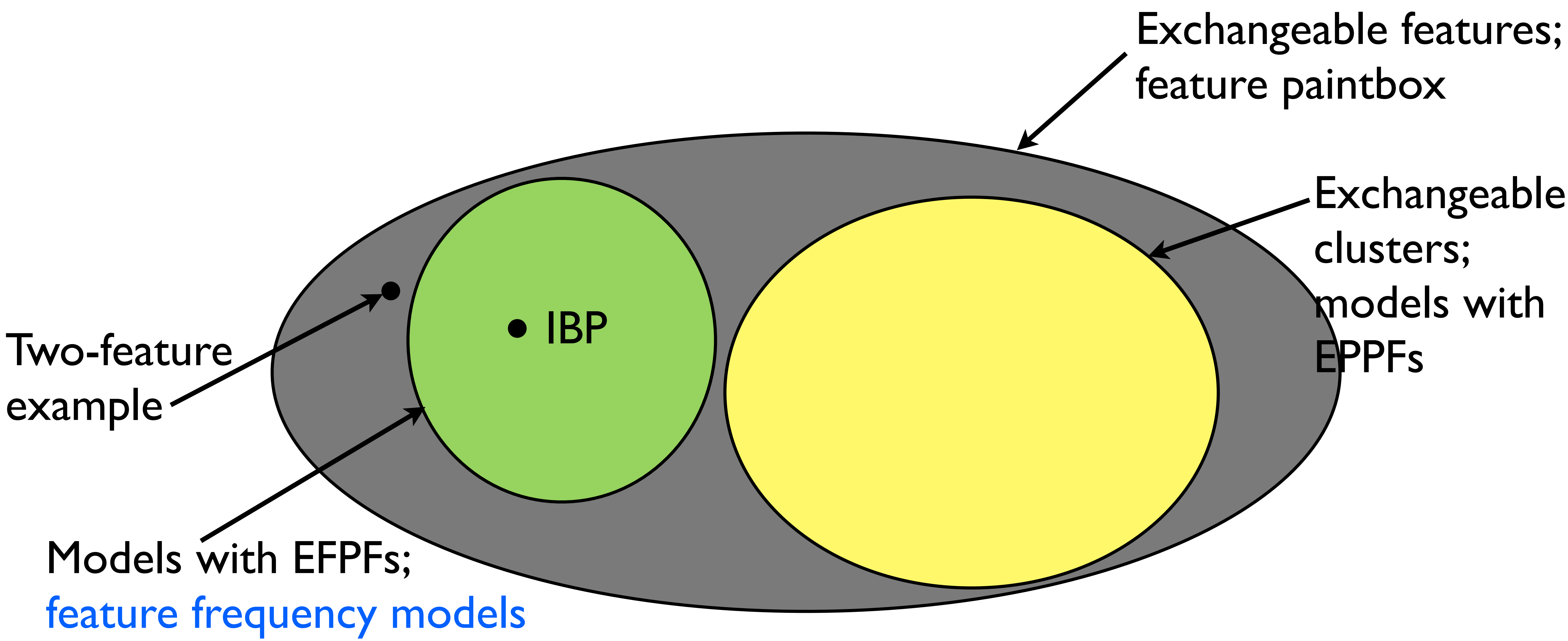
Conclusions

- Feature paintbox: characterization of exchangeable feature models
- Characterization of alternative correlation structure



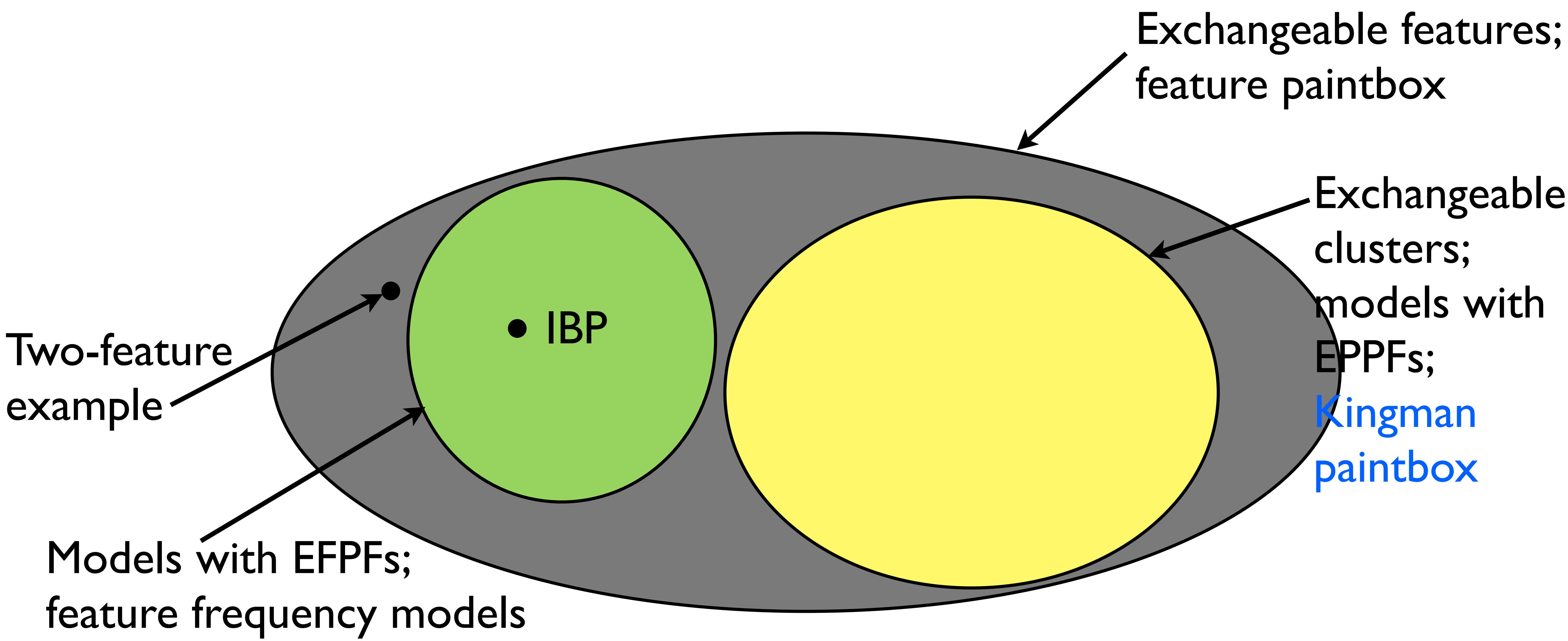
Conclusions

- Feature paintbox: characterization of exchangeable feature models
- Characterization of alternative correlation structure



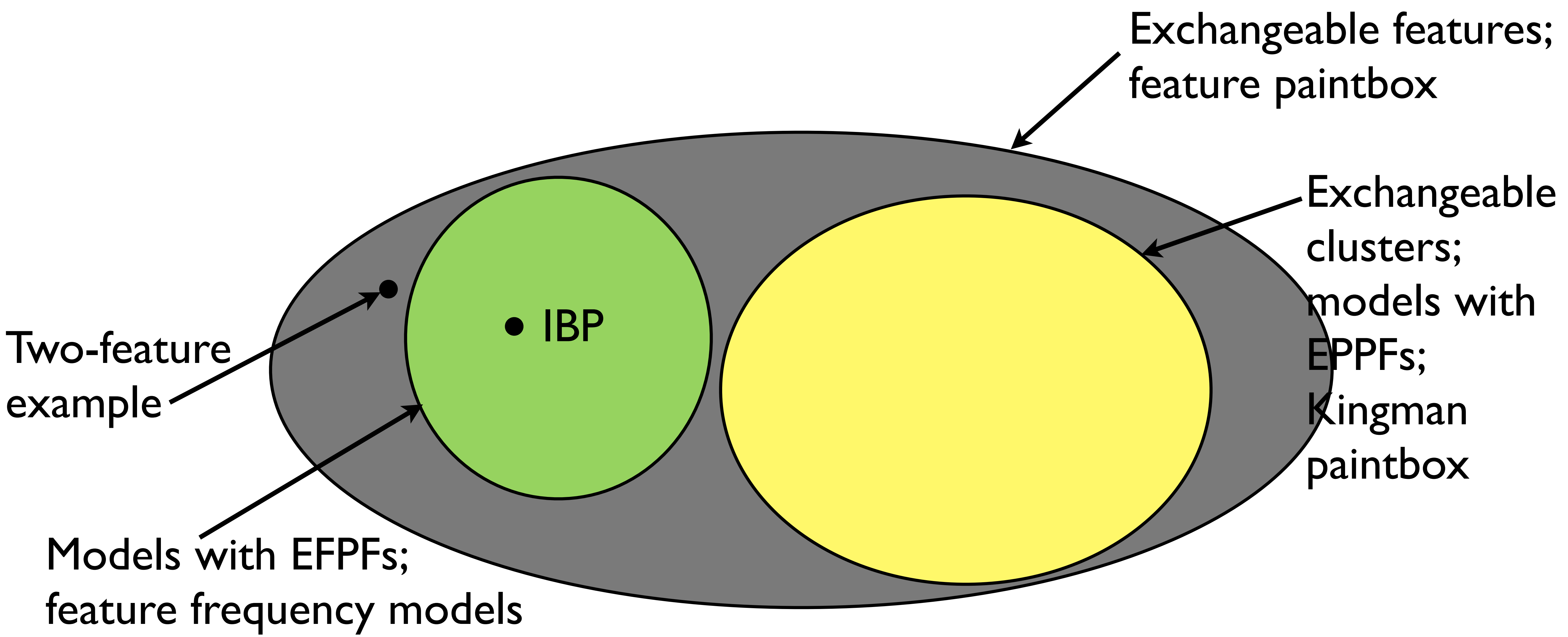
Conclusions

- Feature paintbox: characterization of exchangeable feature models
- Characterization of alternative correlation structure



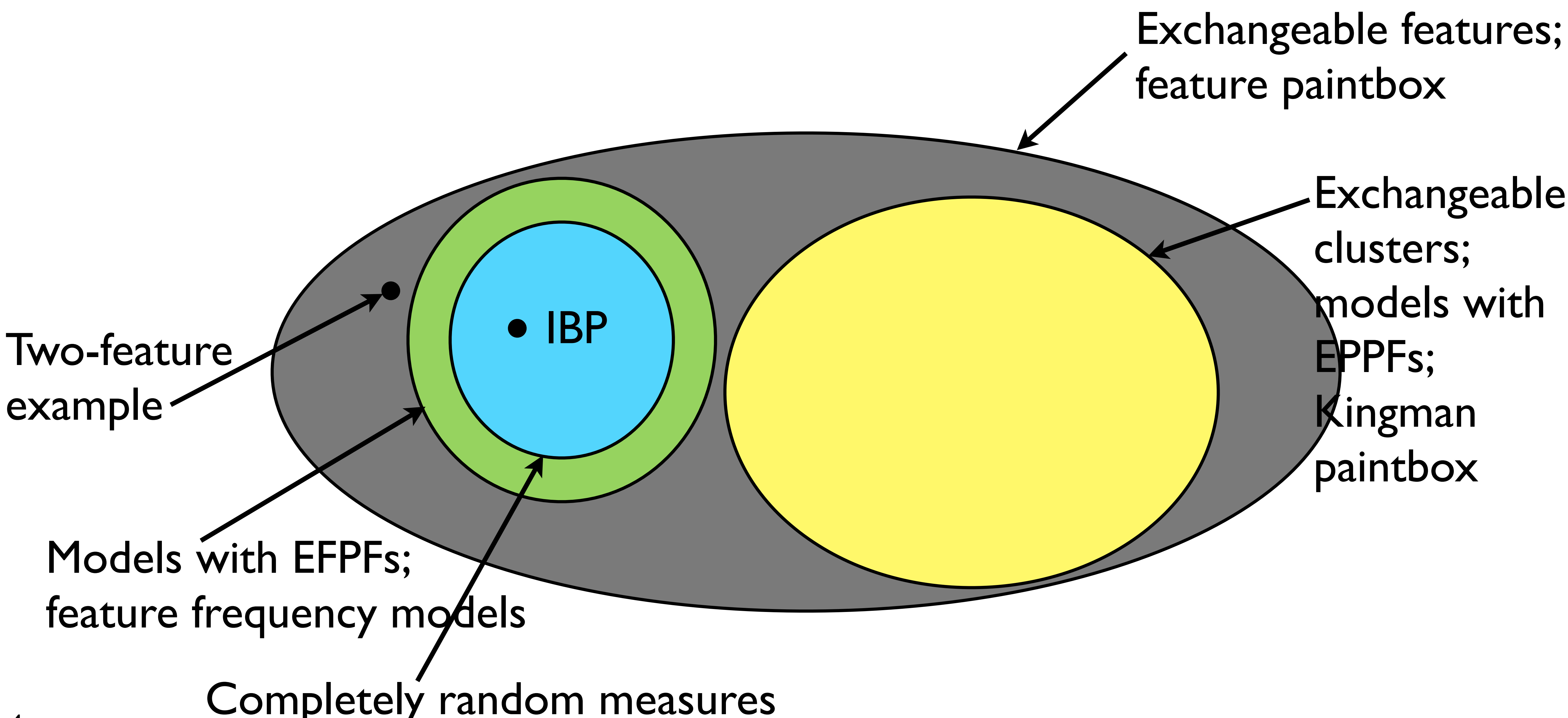
Conclusions

- Feature paintbox: characterization of exchangeable feature models
- Characterization of alternative correlation structure
- Remaining connections



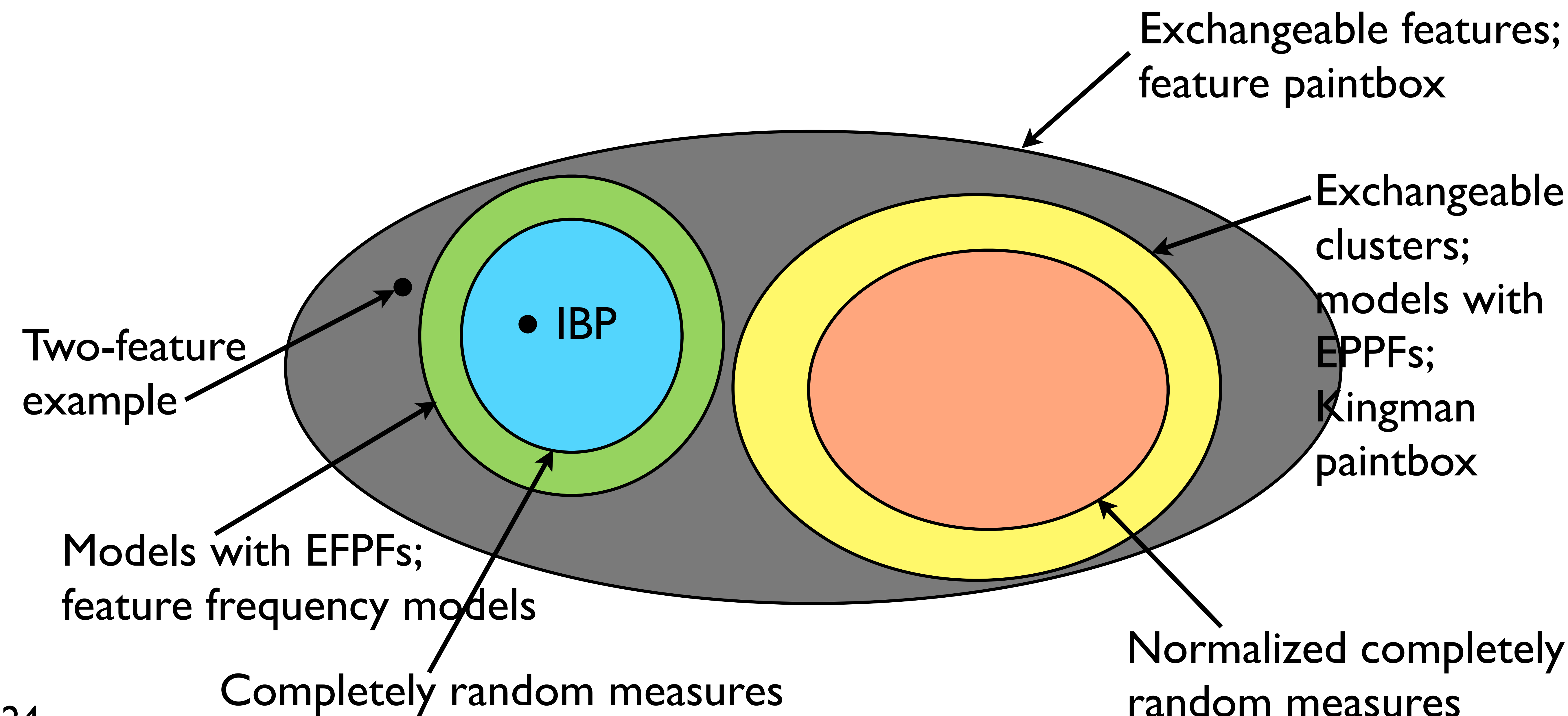
Conclusions

- Feature paintbox: characterization of exchangeable feature models
- Characterization of alternative correlation structure
- Remaining connections (CRMs)



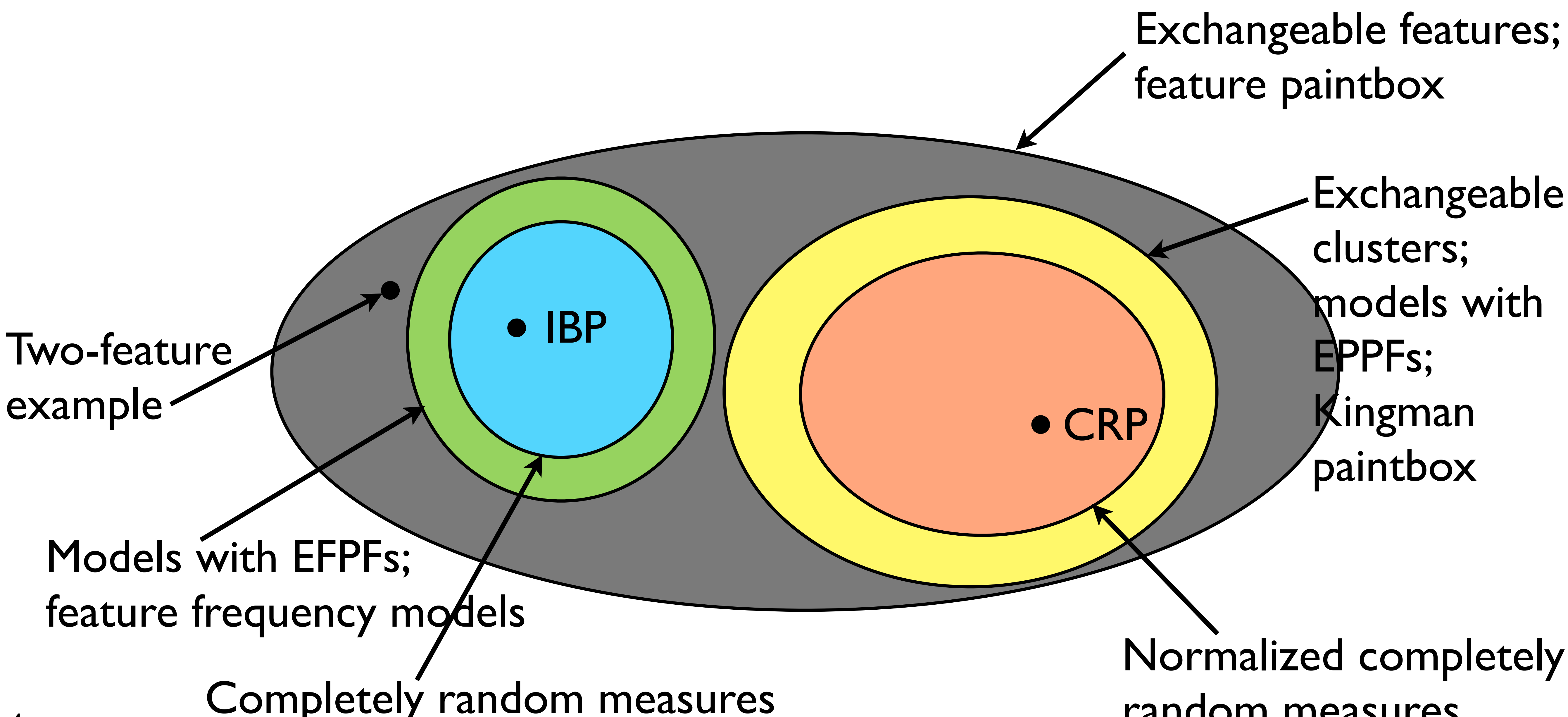
Conclusions

- Feature paintbox: characterization of exchangeable feature models
- Characterization of alternative correlation structure
- Remaining connections (CRMs)



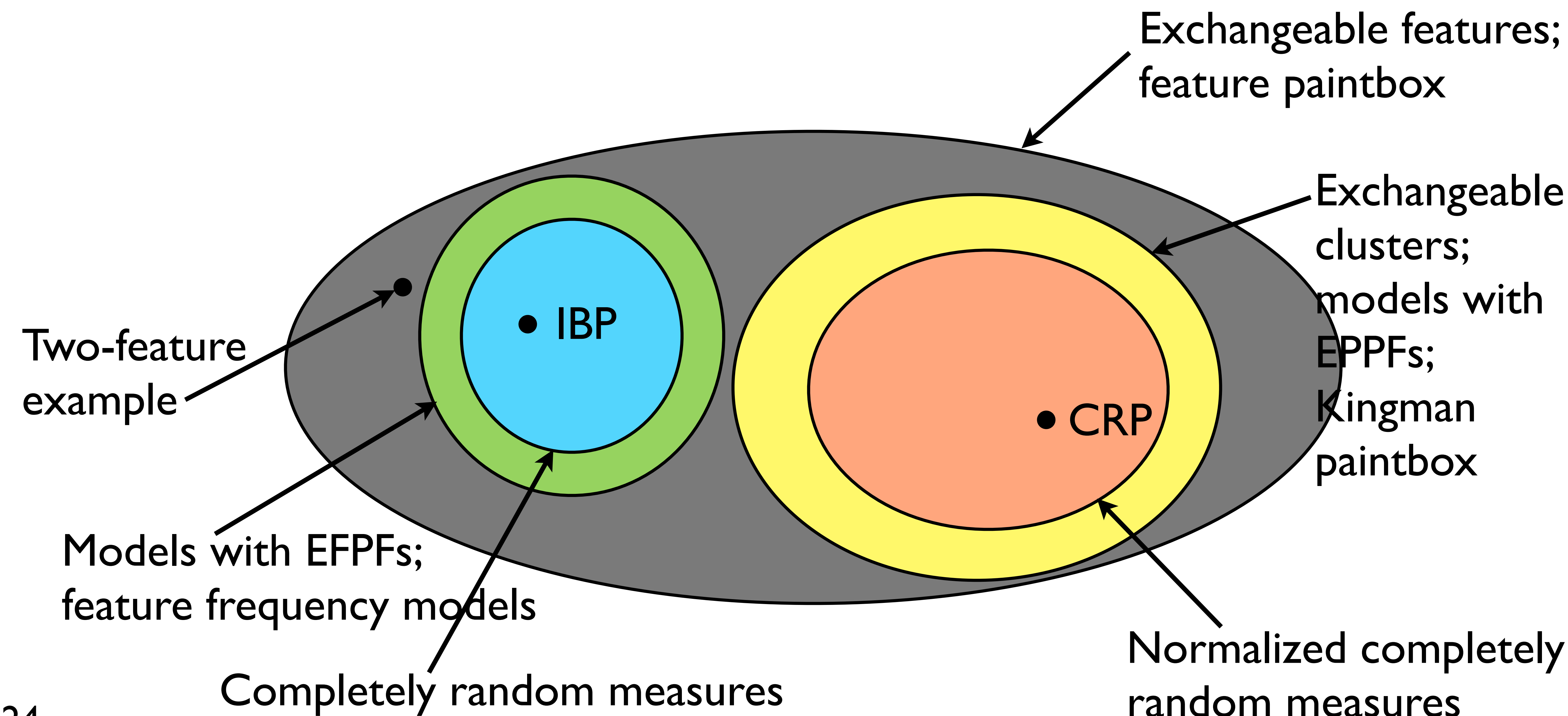
Conclusions

- Feature paintbox: characterization of exchangeable feature models
- Characterization of alternative correlation structure
- Remaining connections (CRMs)



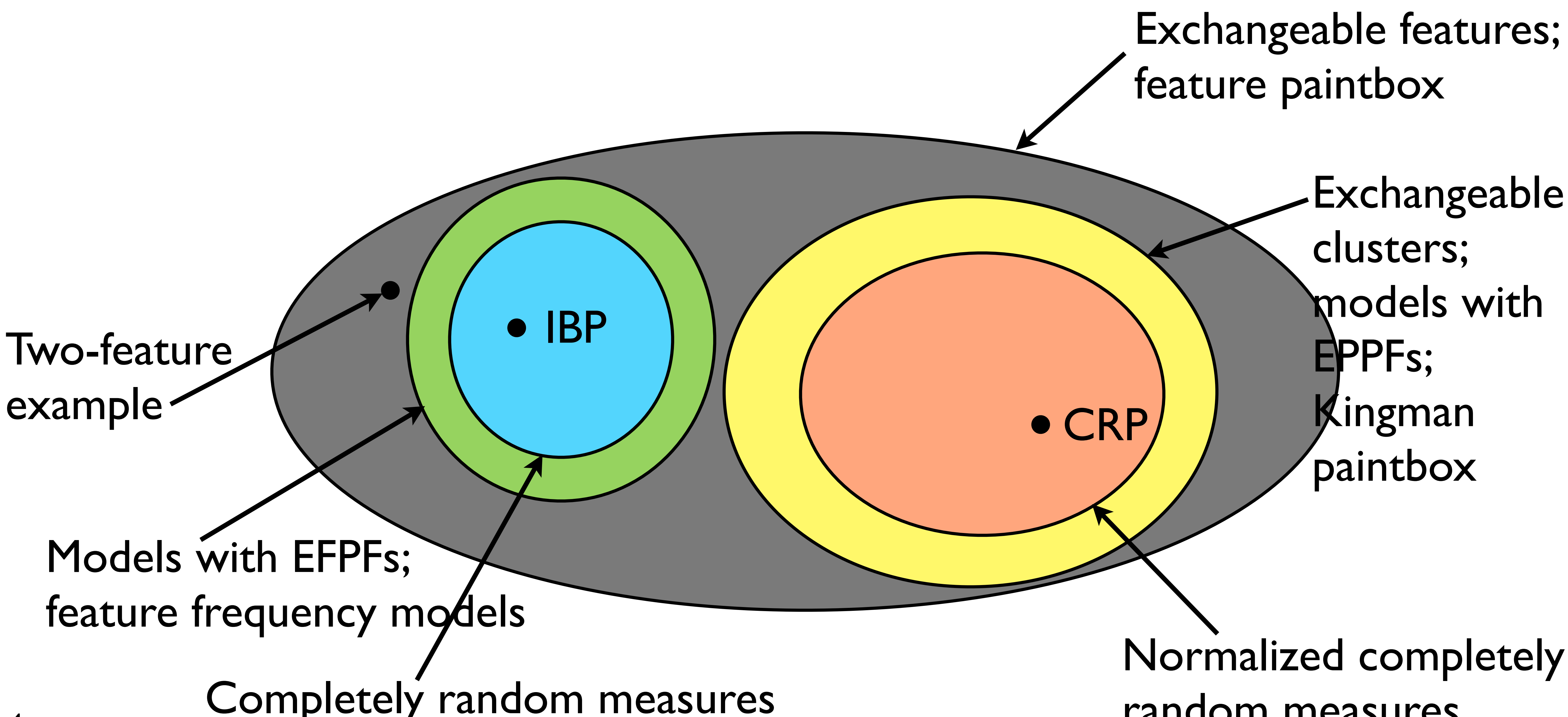
Conclusions

- Feature paintbox: characterization of exchangeable feature models
- Characterization of alternative correlation structure
- Remaining connections (CRMs, dust)



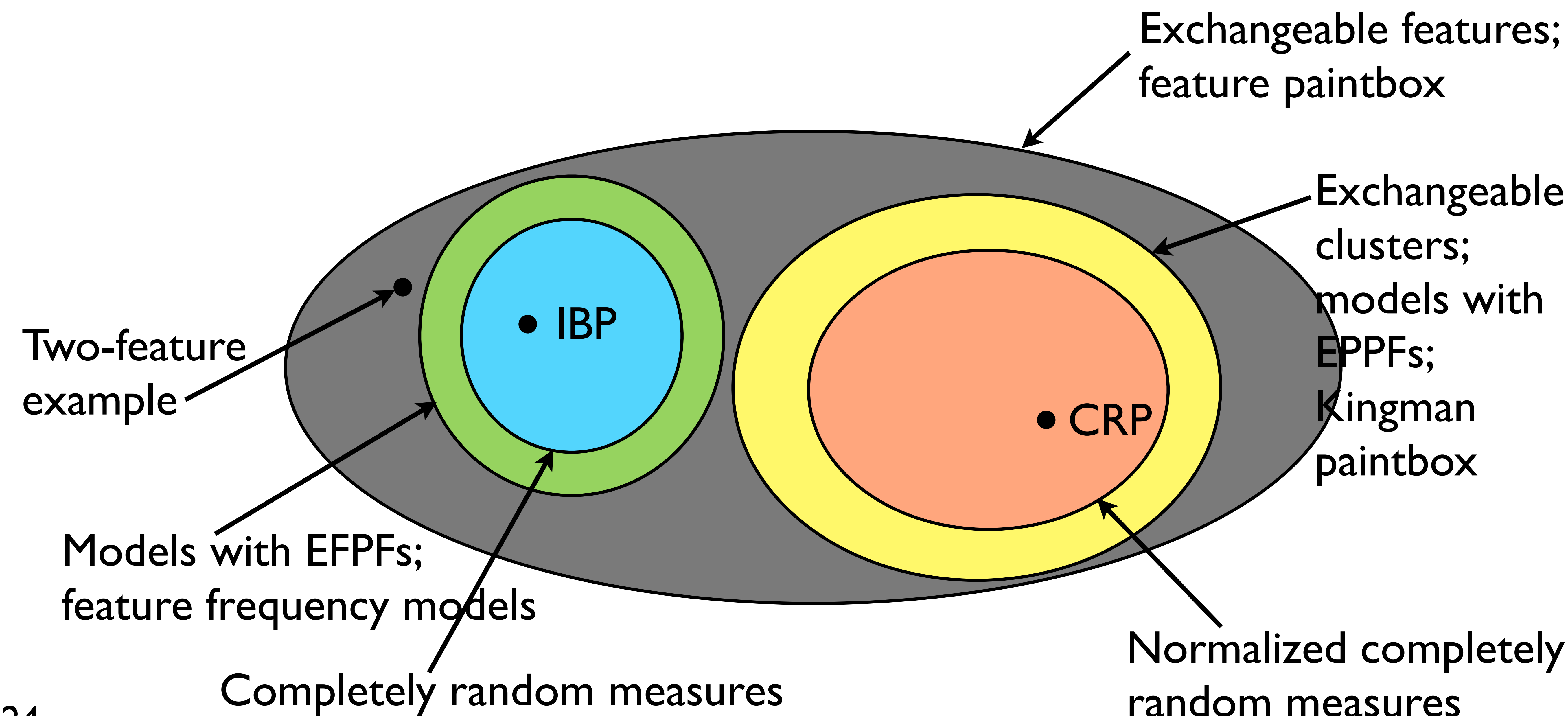
Conclusions

- Feature paintbox: characterization of exchangeable feature models
- Characterization of alternative correlation structure
- Remaining connections (CRMs, dust, etc)



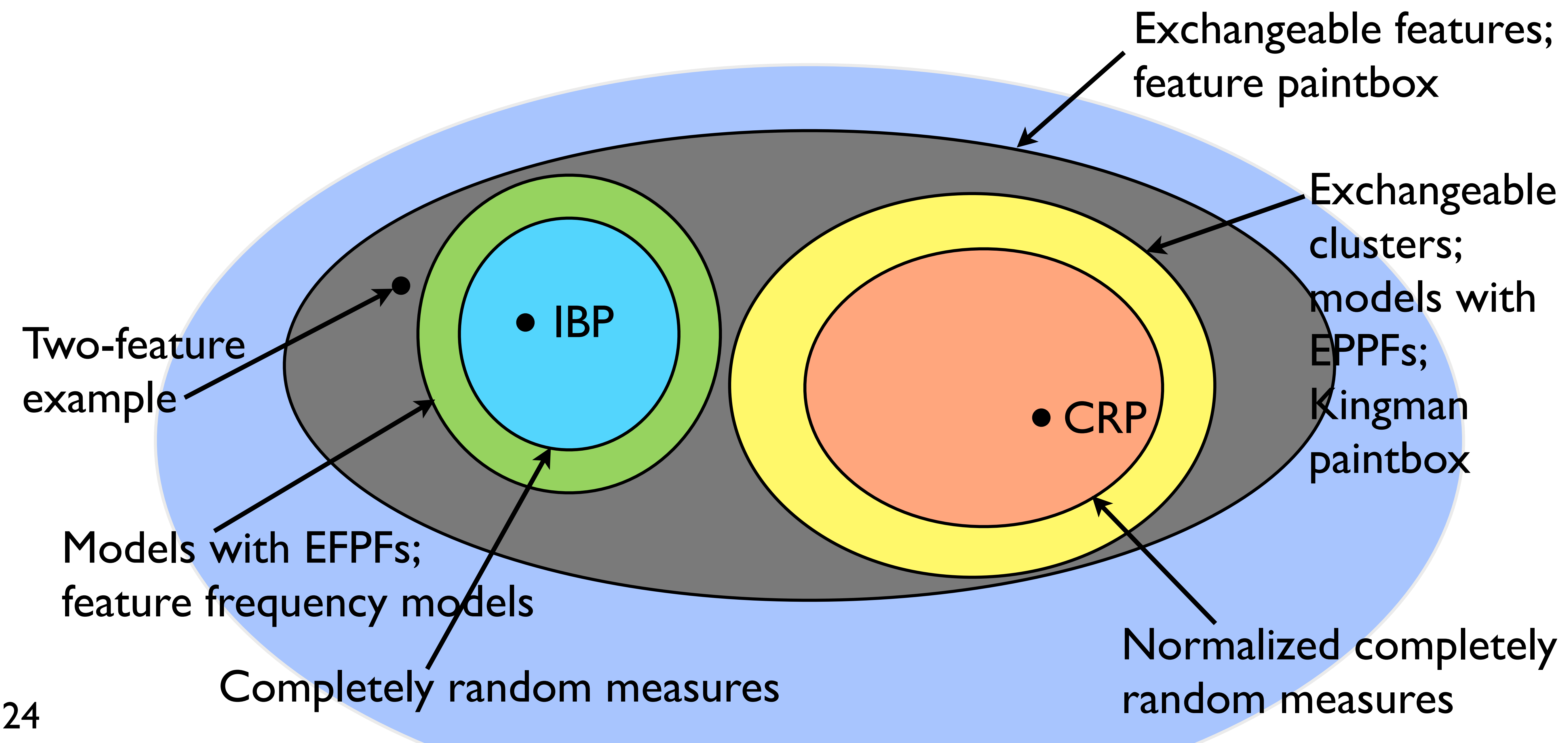
Conclusions

- Feature paintbox: characterization of exchangeable feature models
- Characterization of alternative correlation structure
- Remaining connections (CRMs, dust, etc)
- Other combinatorial structures



Conclusions

- Feature paintbox: characterization of exchangeable feature models
- Characterization of alternative correlation structure
- Remaining connections (CRMs, dust, etc)
- Other combinatorial structures



References

T. Broderick, J. Pitman, and M. I. Jordan. Feature allocations, probability functions, and paintboxes. *Bayesian Analysis, to appear.* Preprint arXiv:1301.6647, 2013.

T. Broderick, M. I. Jordan, and J. Pitman. Clusters and features from combinatorial stochastic processes. *Statistical Science, to appear.* Preprint arXiv: 1206.5862, 2012.

T. Broderick, L. Mackey, J. Paisley, and M. I. Jordan. Combinatorial clustering and the beta negative binomial process. *Preprint arXiv:1111.1802*, 2011.

T. Griffiths and Z. Ghahramani. Infinite latent feature models and the Indian buffet process. In *Advances in Neural Information Processing Systems*, 2006.

N. L. Hjort. Nonparametric bayes estimators based on beta processes in models for life history data. *Annals of Statistics*, 18(3):1259–1294, 1990.

Y. Kim. Nonparametric Bayesian estimators for counting processes. *Annals of Statistics*, 27(2):562–588, 1999.

J. F. C. Kingman. The representation of partition structures. *Journal of the London Mathematical Society*, 2(2):374, 1978.

J. Pitman. Exchangeable and partially exchangeable random partitions. *Probability Theory and Related Fields*, 102(2):145–158, 1995.

R. Thibaux and M. I. Jordan. Hierarchical beta processes and the Indian buffet process. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*, 2007.

M. Zhou, L. Hannah, D. Dunson, and L. Carin. Beta-negative binomial process and Poisson factor analysis. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*, 2012.