



**University of  
Zurich<sup>UZH</sup>**

<http://r-bayesian-networks.org/>  
[gilles.kratzer@gmail.com](mailto:gilles.kratzer@gmail.com)  
[reinhard.furrer@math.uzh.ch](mailto:reinhard.furrer@math.uzh.ch)

**GILLES KRATZER, PHD**

**PROF. REINHARD FURRER**

**USER! CONFERENCE, ONLINE 07.07.2021**

---

# ADVANCED BAYESIAN NETWORK MODELLING

## *Outline of the talk*

- ▶ Heuristic search
- ▶ MCMC over structures
  
- ▶ *Other advanced methods/features:*
  - ▶ *Scoring system*
  - ▶ *Tunable parameter prior*
  - ▶ *Structural prior*
  - ▶ *Data separation*
  - ▶ *Covariate adjustment*
  - ▶ *Likelihood contribution*

Find maximum a posteriori score

✓ Exact search

▶ Heuristic search

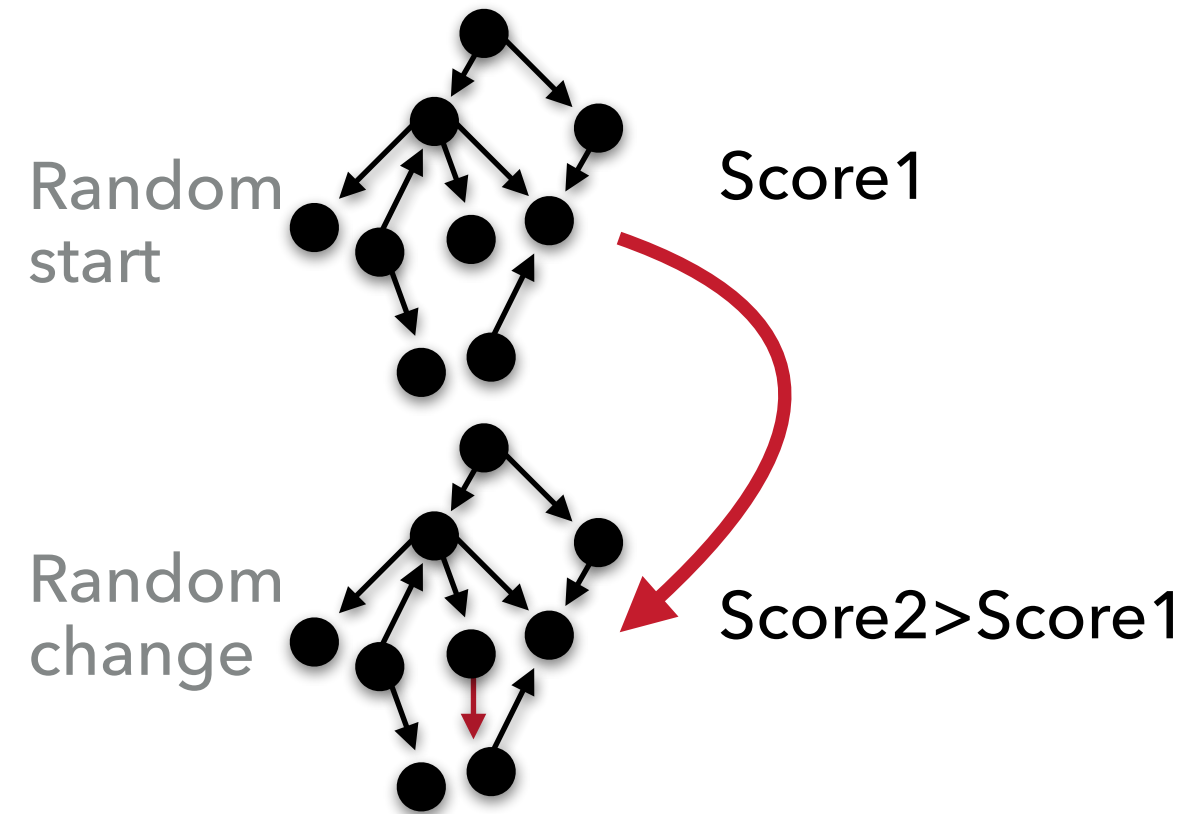
▶ MCMC over structures

# HEURISTIC SEARCH

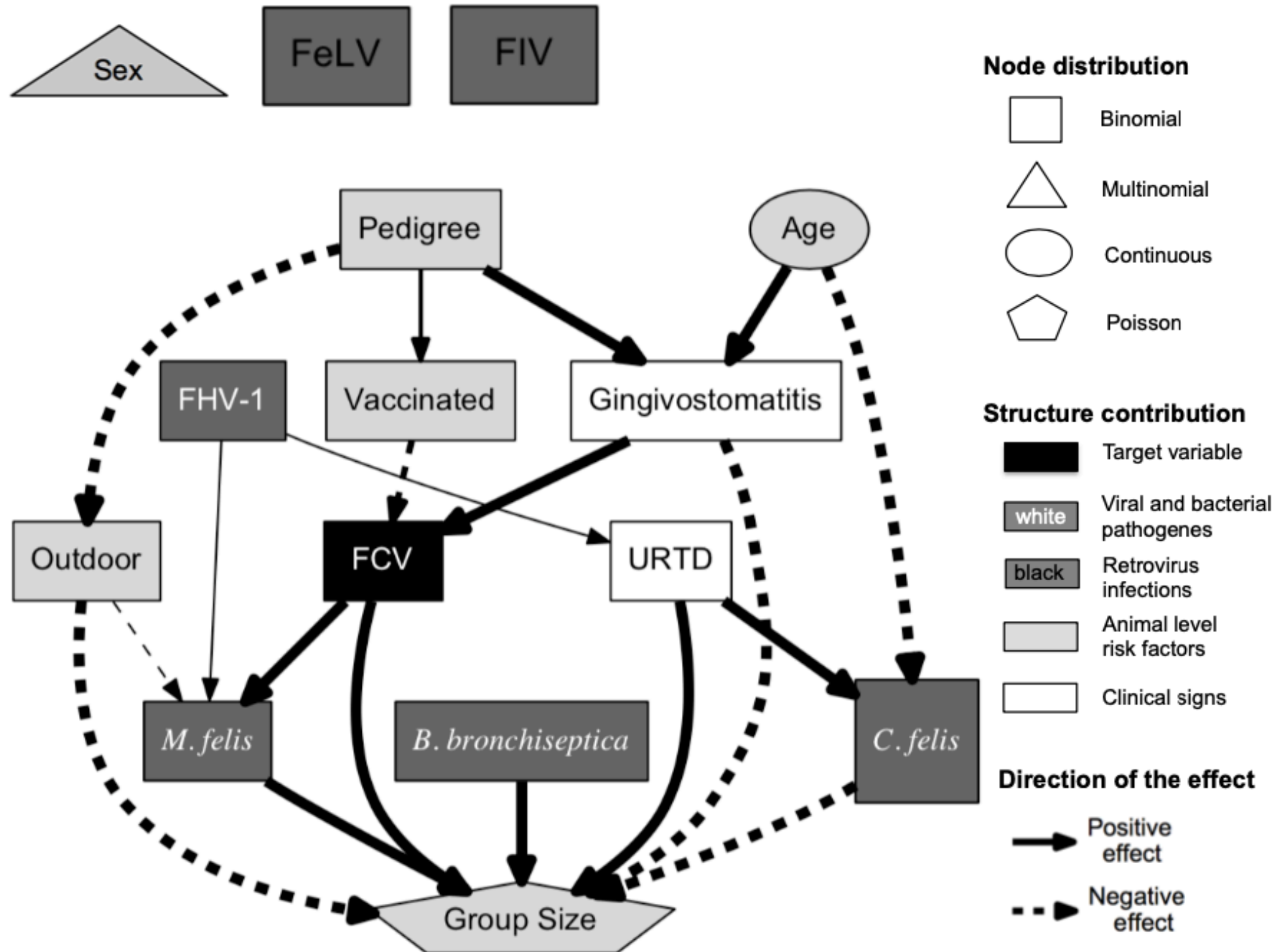
---

## *Heuristic search: Greedy Hill-Climbing*

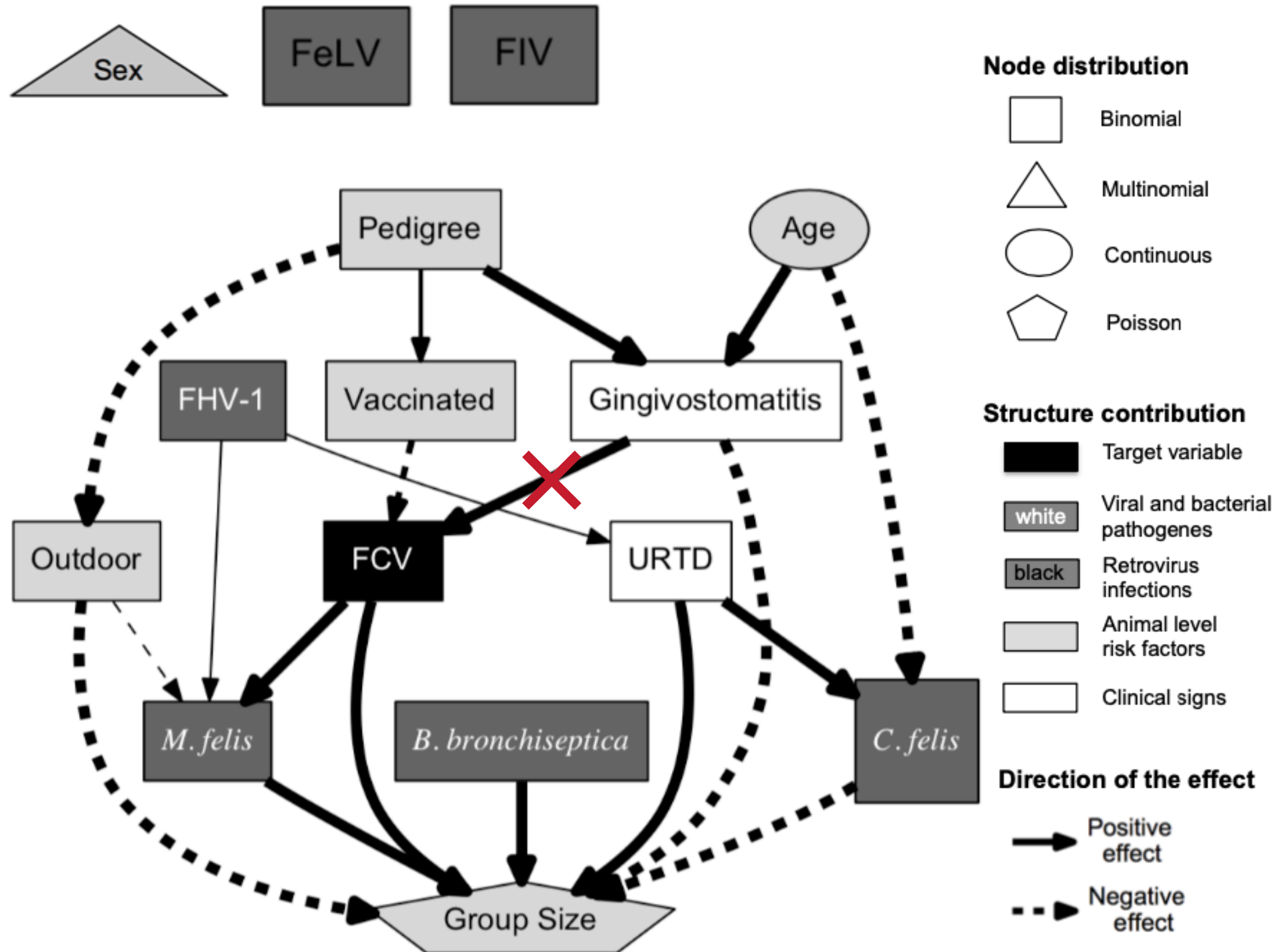
- ▶ Simplest heuristic local search
  - ▶ Start with a given network
    - ▶ empty network
    - ▶ best tree
    - ▶ a random network
  - ▶ At each iteration
    - ▶ Evaluate all possible changes
    - ▶ Apply change that leads to best improvement in score
    - ▶ Reiterate
  - ▶ Stop when no modification improves score
- ▶ *Pitfalls:*
  - ▶ Local Maxima
  - ▶ Plateaus
- ▶ *Solution:*
  - ▶ Tabu
  - ▶ Random restart
  - ▶ Simulated annealing



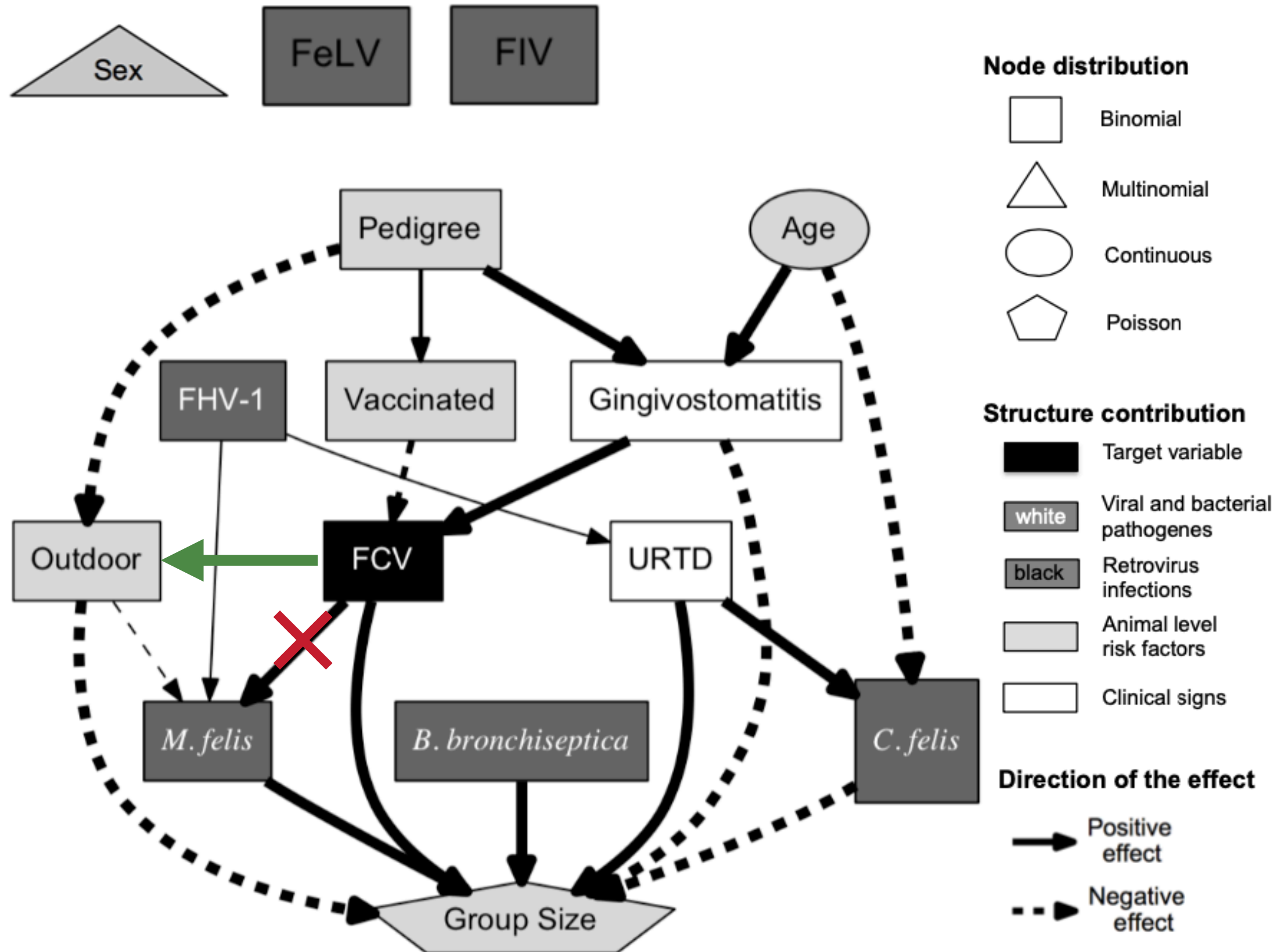
# BAYESIAN NETWORK OF THE “CAT FLU”



# BAYESIAN NETWORK OF THE “CAT FLU”

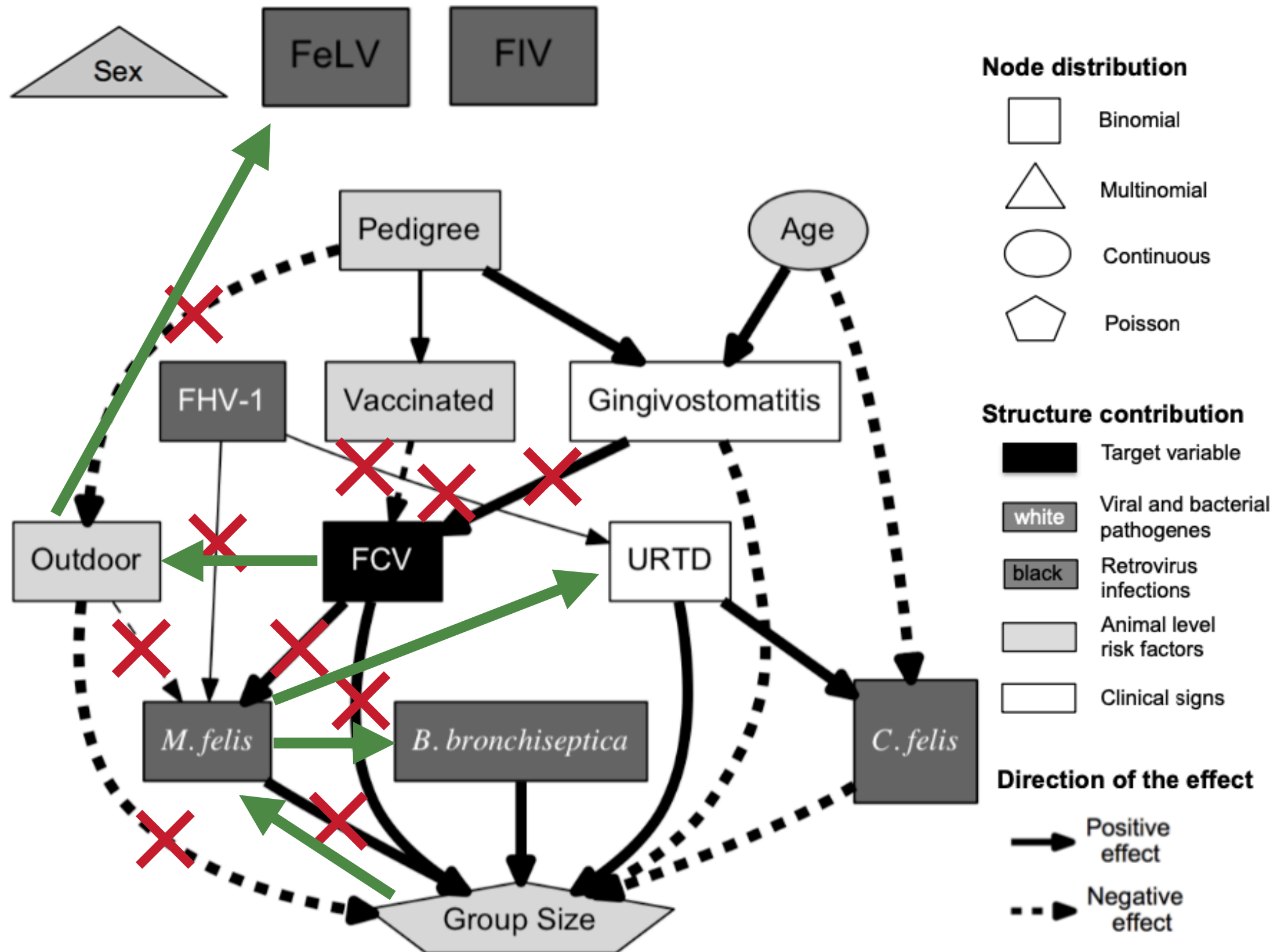


# BAYESIAN NETWORK OF THE "CAT FLU"

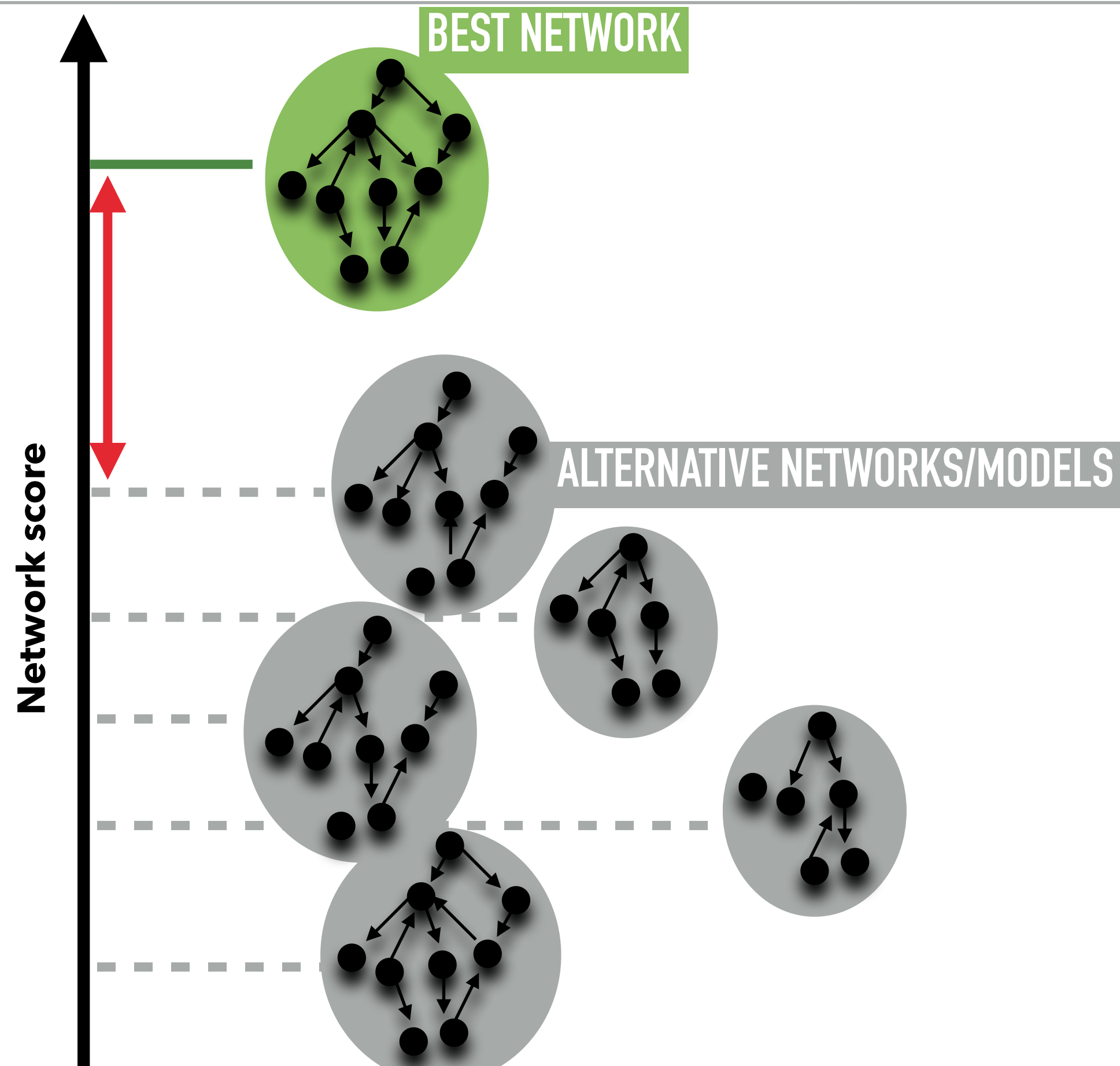




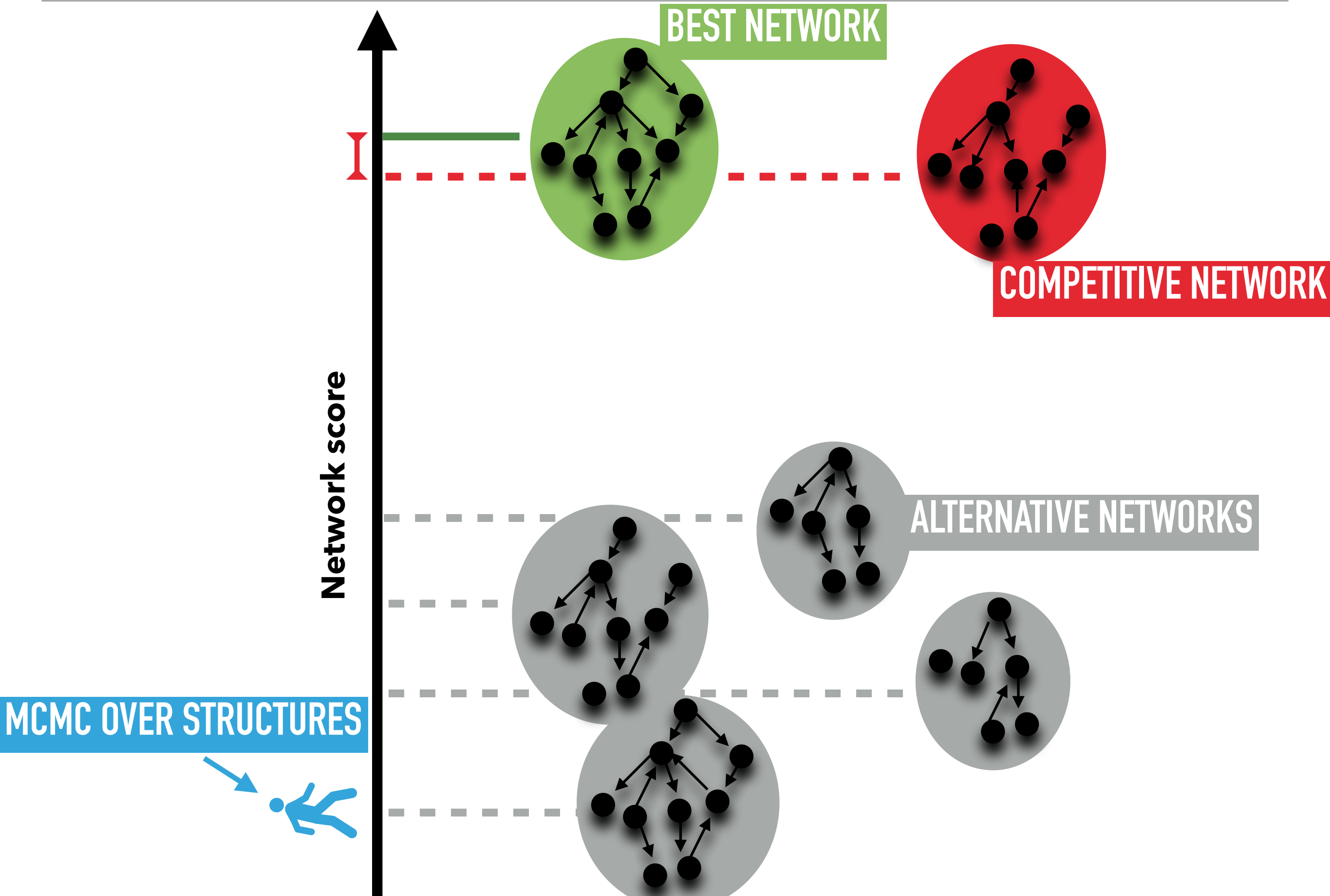
# BAYESIAN NETWORK OF THE "CAT FLU"



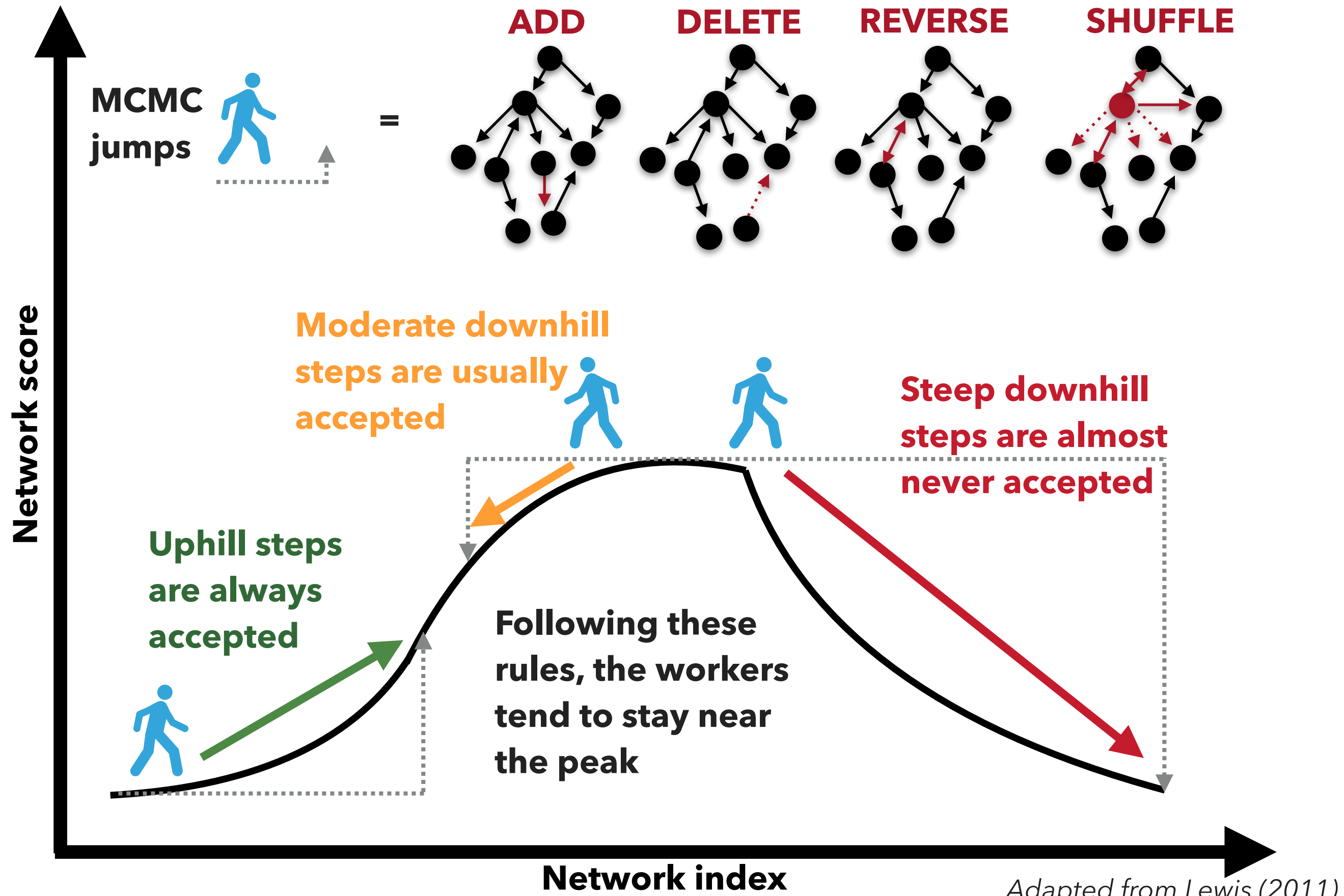




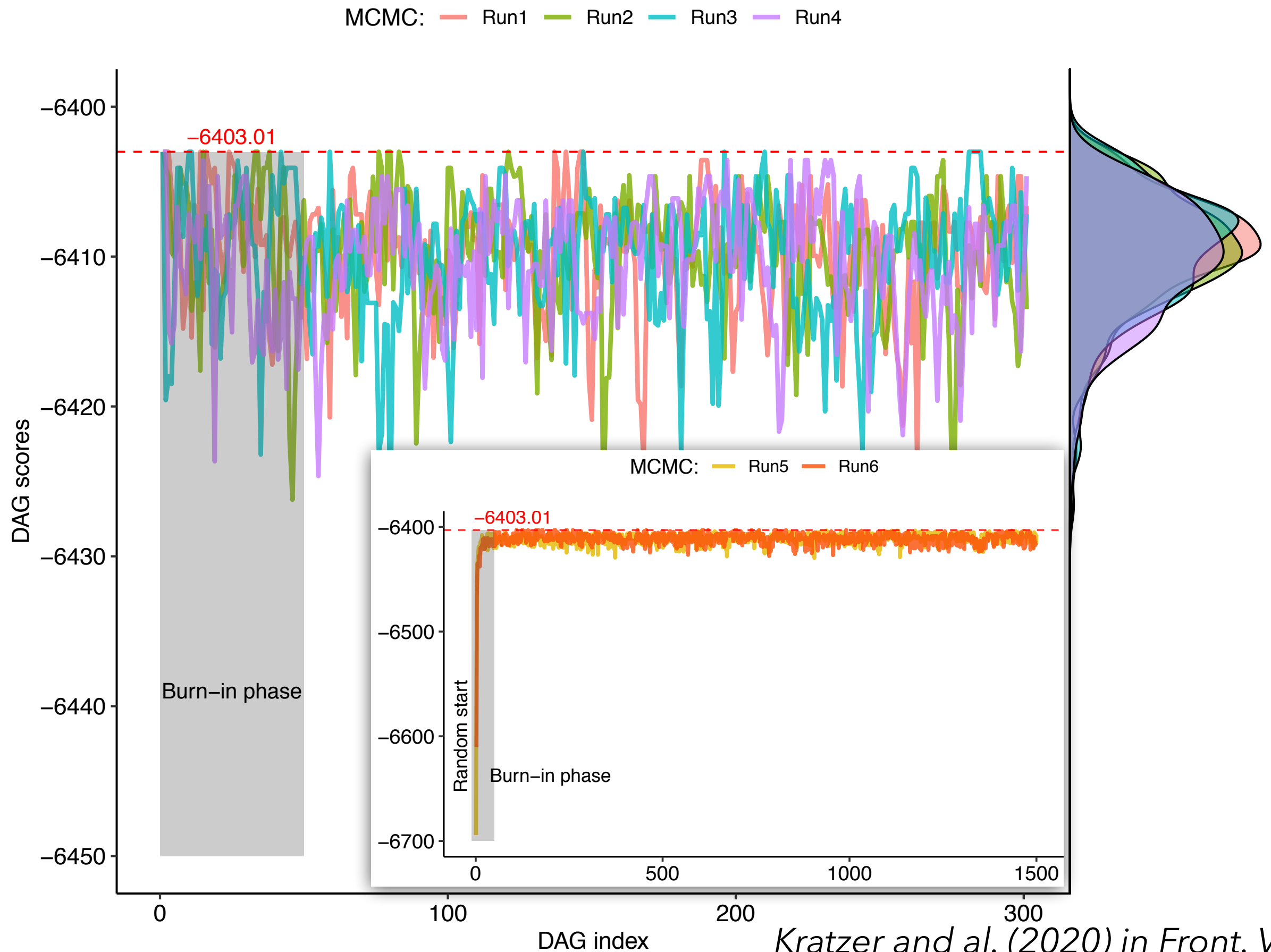
# MCMCABN - HOW TO CHOOSE THE BEST NETWORK?



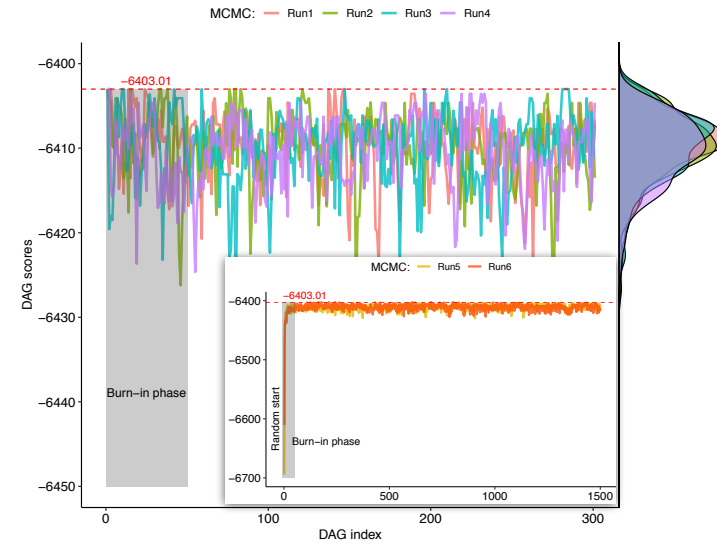
# MCMC TO COMPUTE THE POSTERIOR DISTRIBUTION OF NETWORK



# MCMC OVER STRUCTURE - POSTERIOR DISTRIBUTION

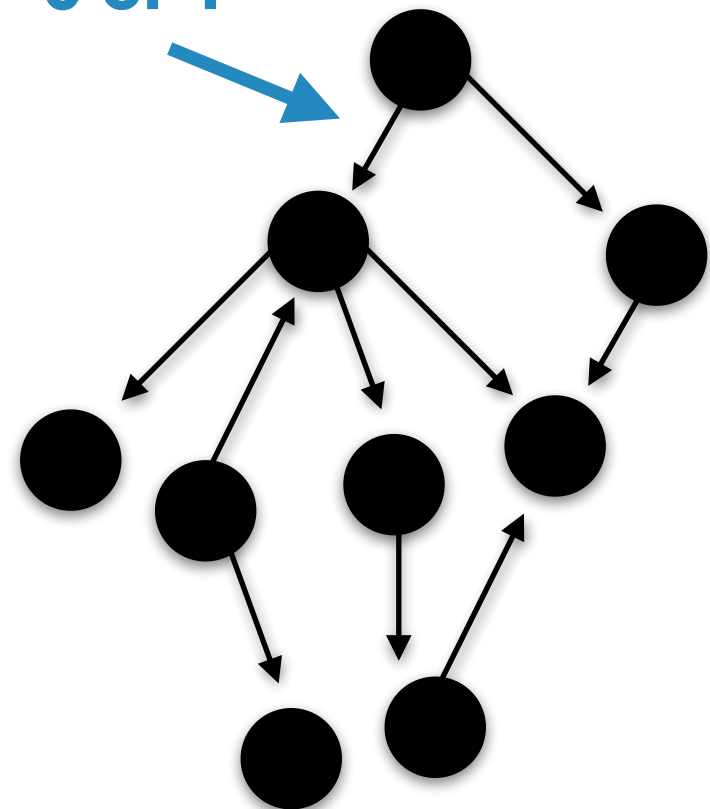


# WHY DO WE PERFORM MCMC OVER STRUCTURES?

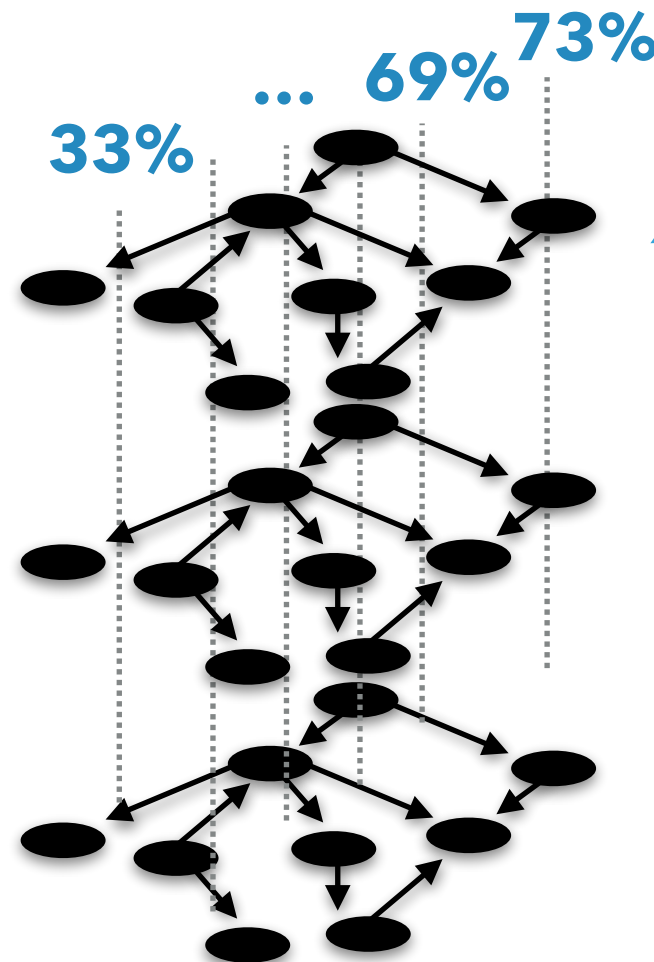


**Best Unique Bayesian Network**

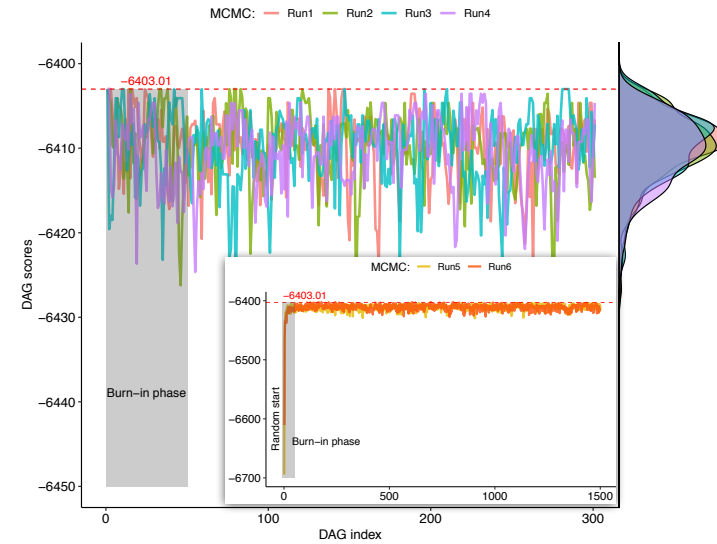
0 or 1



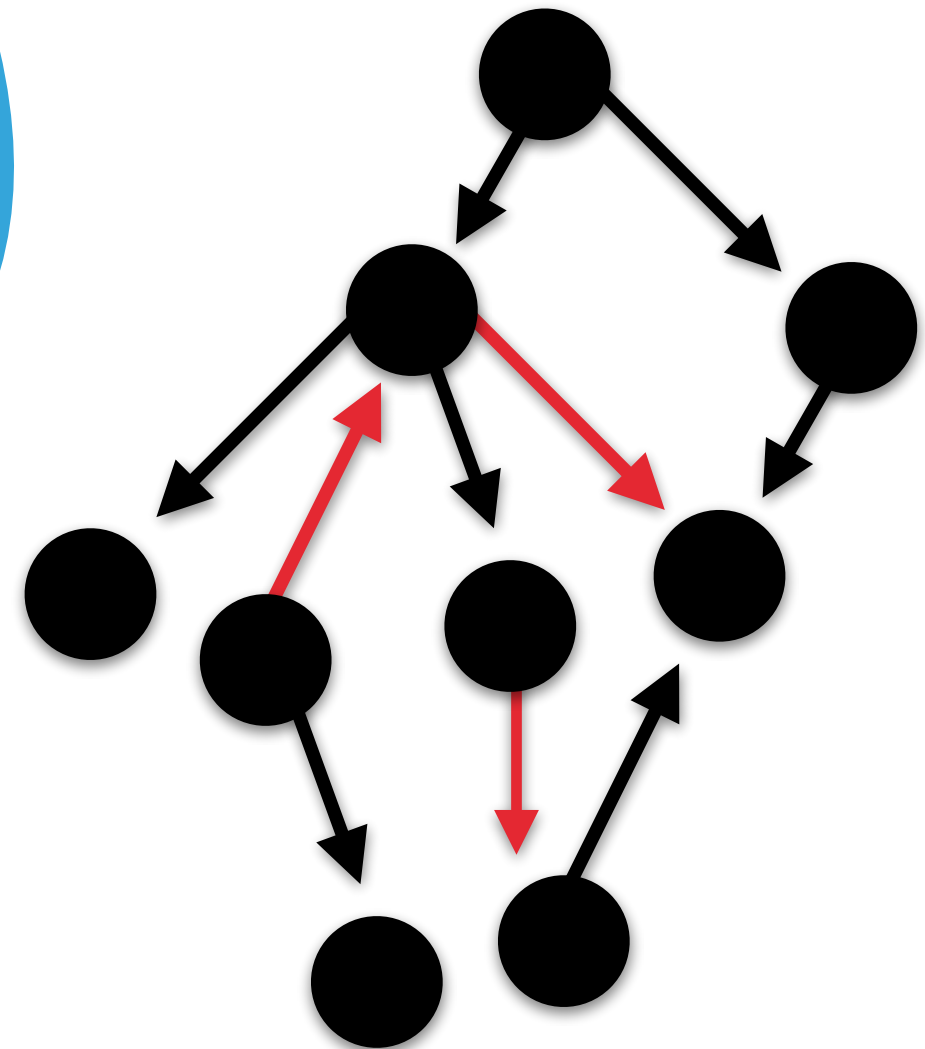
**Counting prevalence of each ARC**



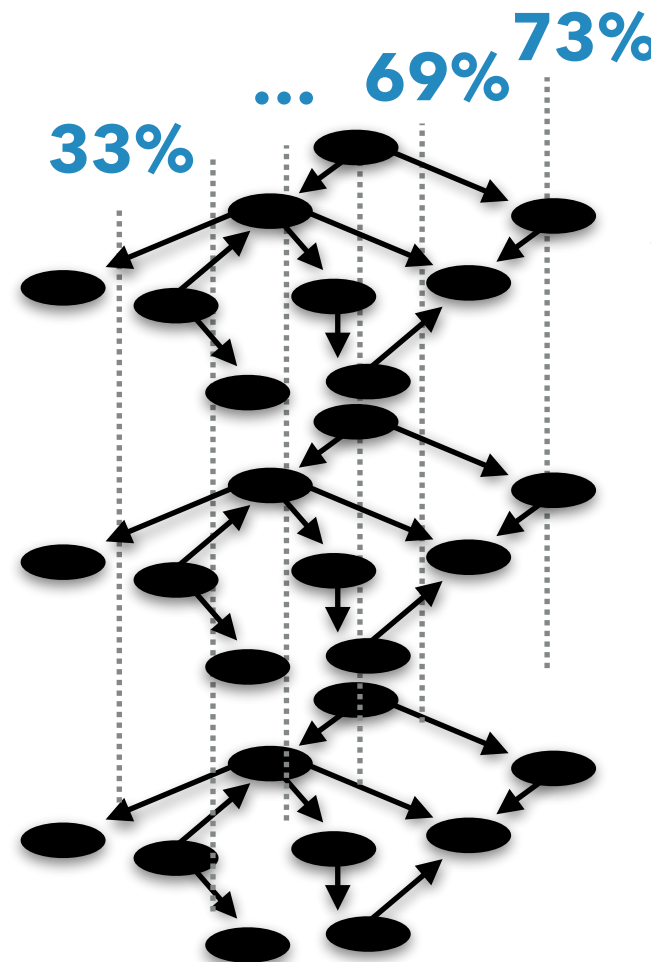
# WHY DO WE PERFORM MCMC OVER STRUCTURES?



## Structural queries

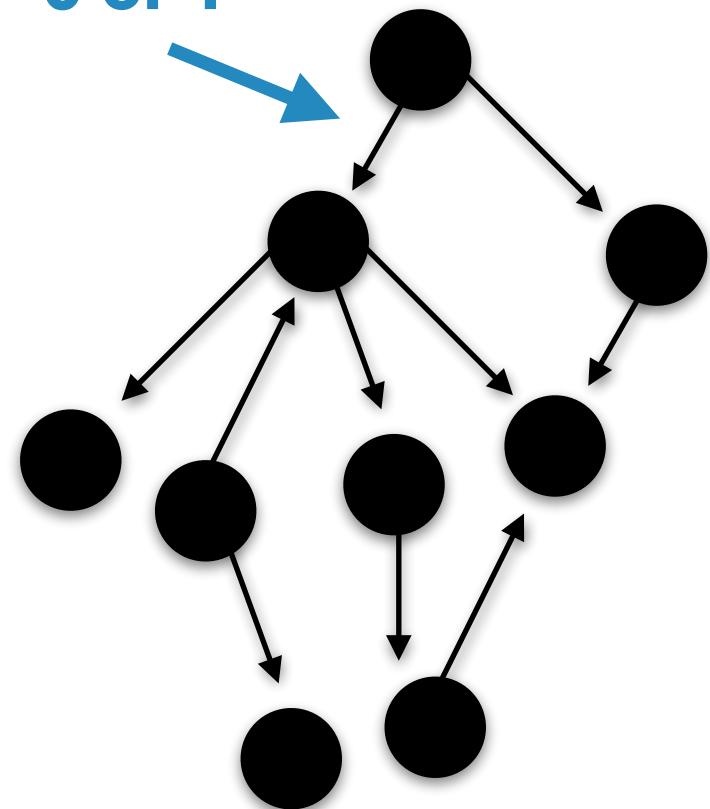


## Counting prevalence of each ARC



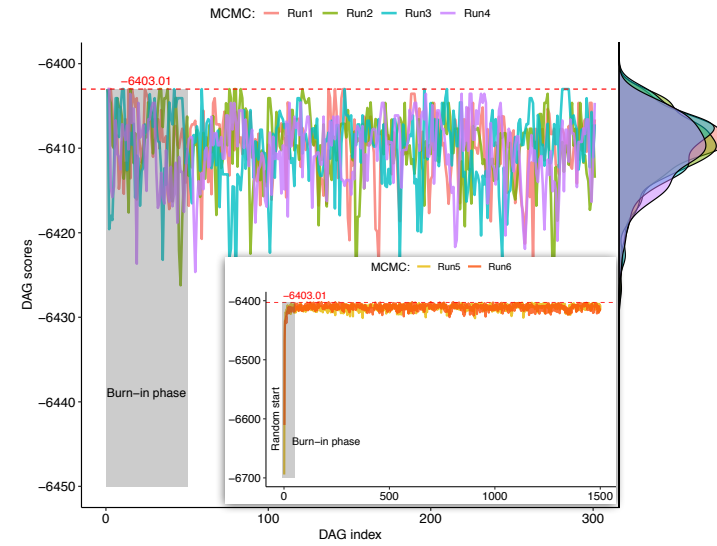
## Best Unique Bayesian Network

0 or 1

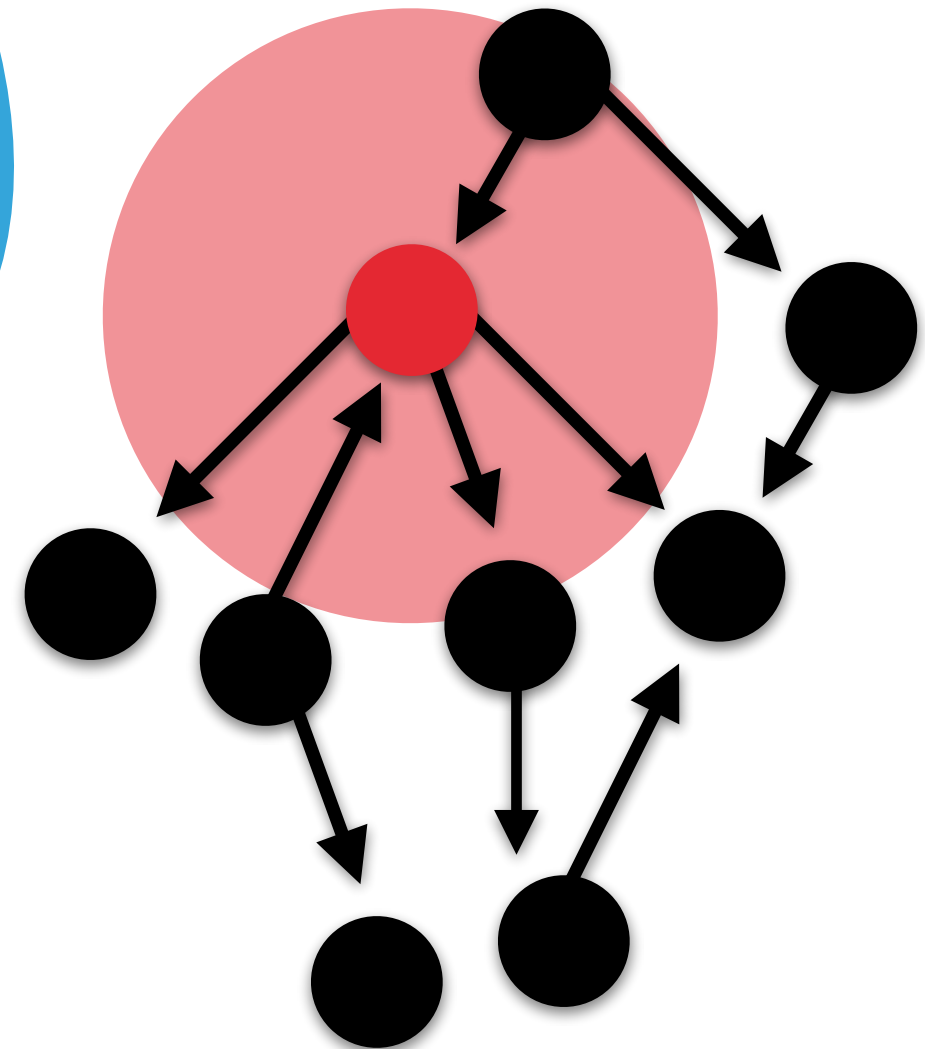




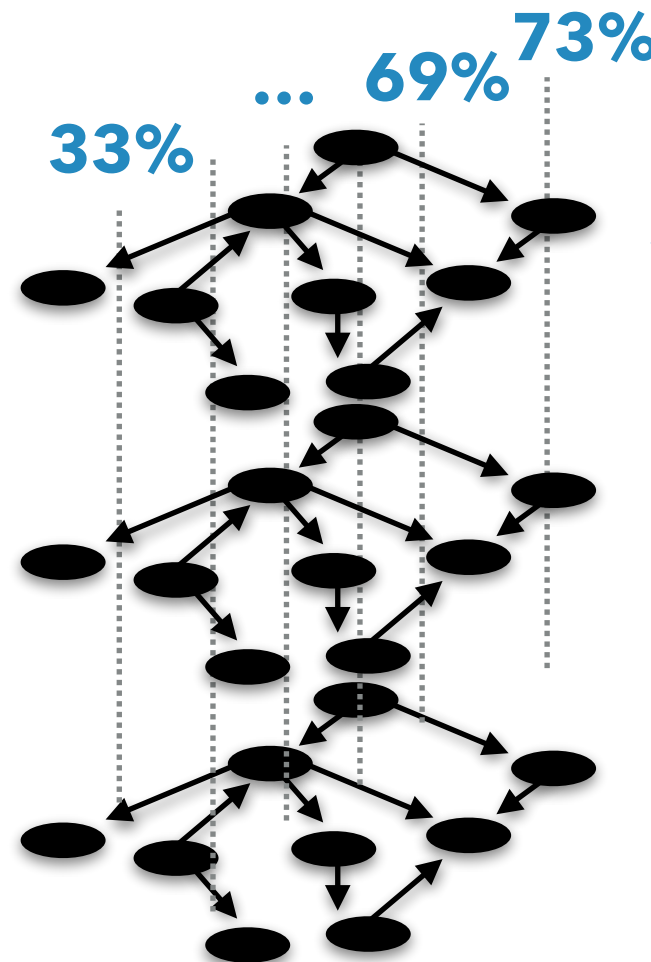
# WHY DO WE PERFORM MCMC OVER STRUCTURES?



## Structural queries

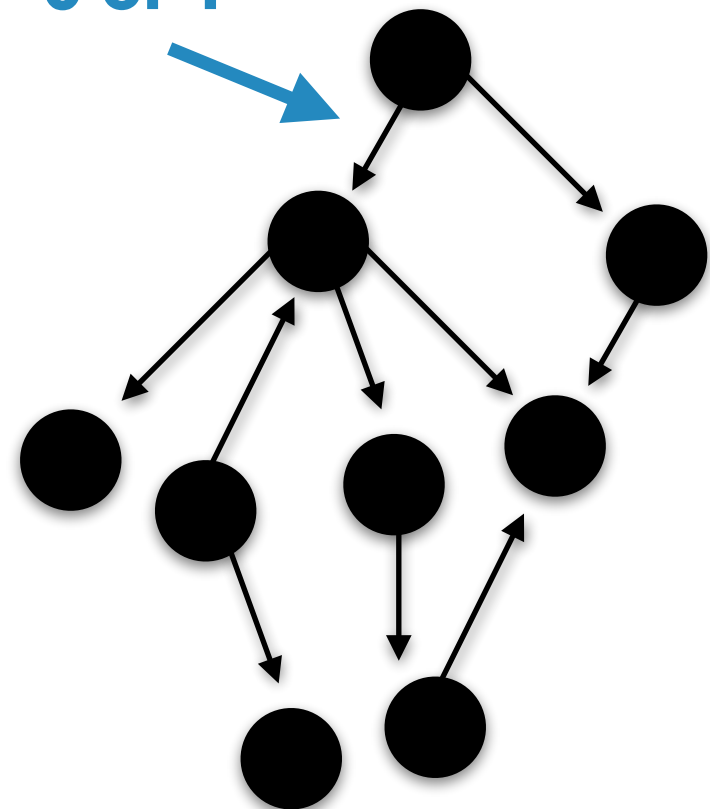


## Counting prevalence of each ARC



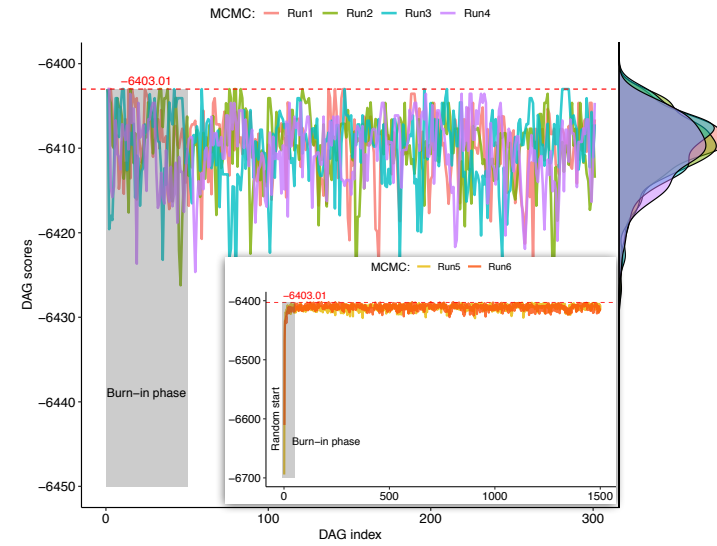
## Best Unique Bayesian Network

0 or 1

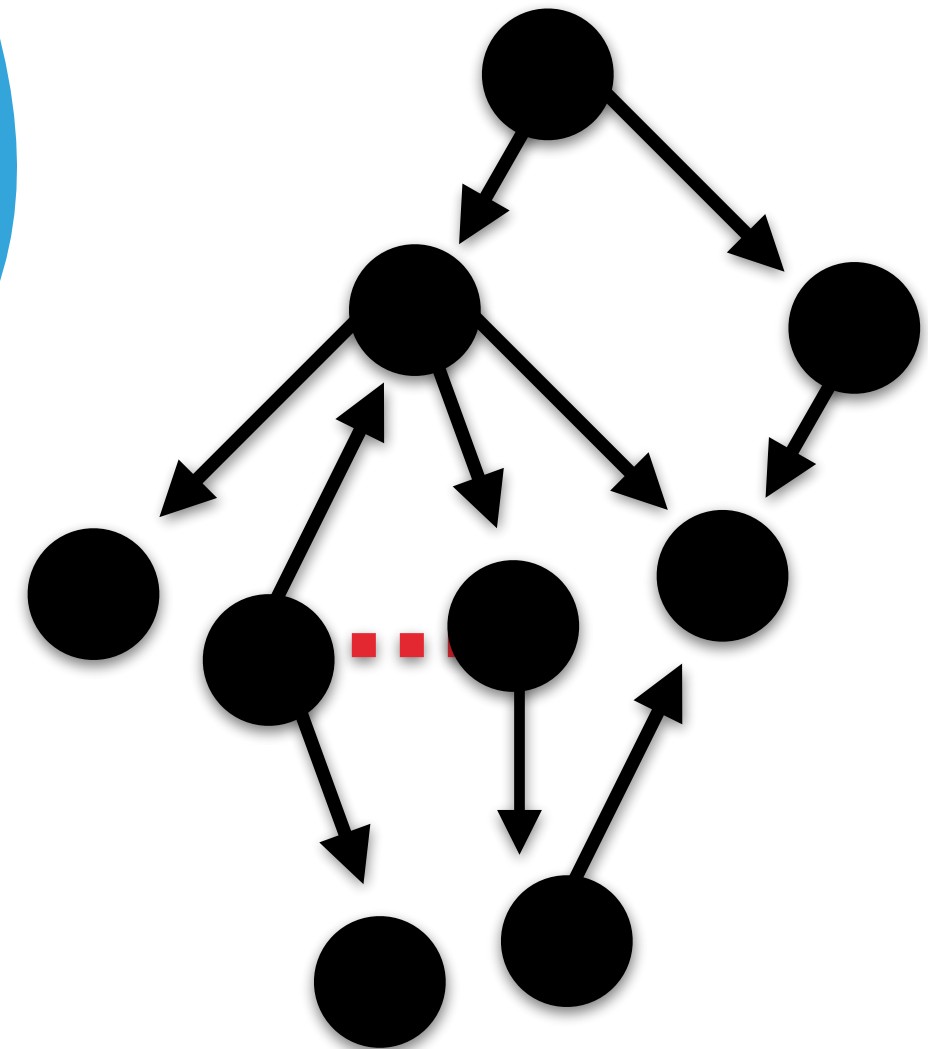




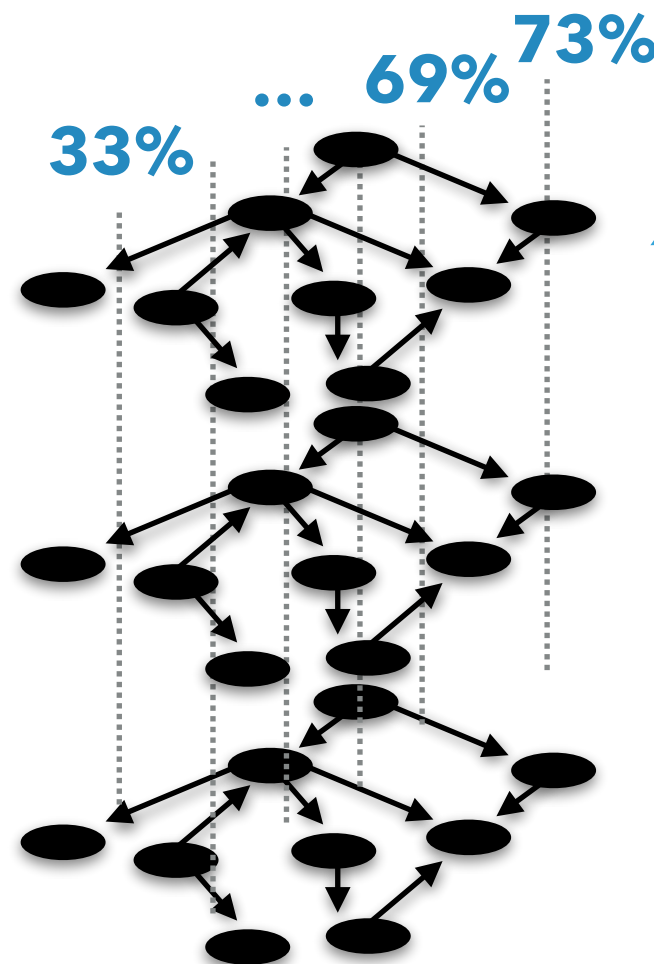
# WHY DO WE PERFORM MCMC OVER STRUCTURES?



## Structural queries

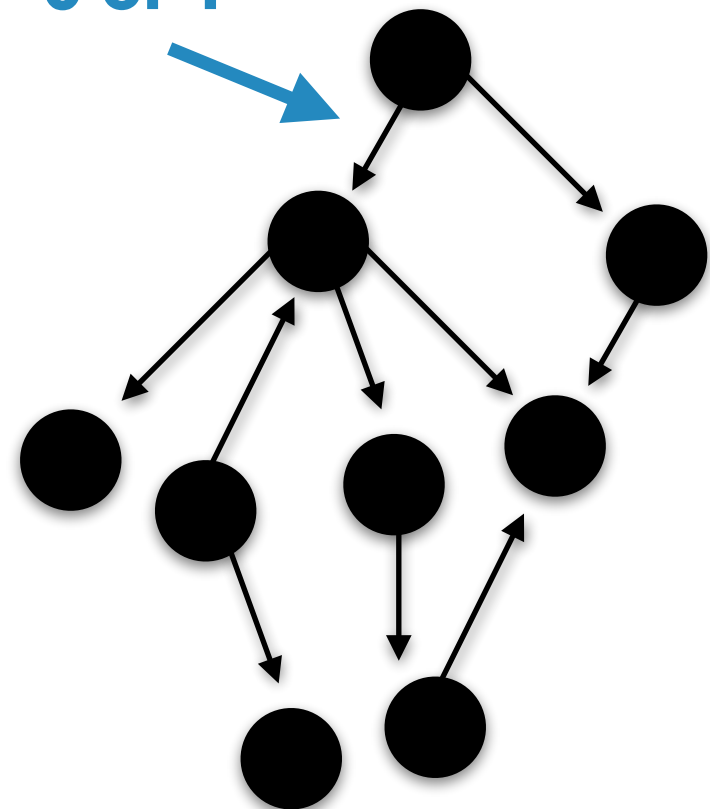


## Counting prevalence of each ARC



## Best Unique Bayesian Network

0 or 1



# MCMC OVER STRUCTURES

---

## *MCMC over structures*

- ▶ Selecting the most probable structure
- ▶ Controlling for overfitting
- ▶ Sampling the landscape of high scoring structures
  - ▶ In applied perspective avoid reducing the richness of BN modelling to only **one** structure
  - ▶ Quantify the marginal impact of relationships by marginalising out over structures

# MCMC OVER STRUCTURES

---

## *MCMC over structures*

- ▶ Selecting the most probable structure
- ▶ Controlling for overfitting
- ▶ Sampling the landscape of high scoring structures
  - ▶ In applied perspective avoid reducing the richness of BN modelling to only **one** structure
  - ▶ Quantify the marginal impact of relationships by marginalising out over structures

## **Question?**

- ▶ Why in a typical regression analysis (GLM) based on observational data we usually present only the best fitting model?



# HANDS-ON EXERCICES