



Is a single unique Bayesian network enough to accurately represent your data?

Gilles Kratzer¹, Reinhard Furrer^{1,2}

¹Department of Mathematics, ²Department of Computational Science; University of Zurich (Switzerland)

Contact: gilles.kratzer@math.uzh.ch

Motivation

- Bayesian network modelling output is usually a **unique** structure
- Need to account for the **uncertainty** in the structures
- Need for **Bayesian model averaging accounting** for prior structure knowledge

Summary

- Generate **MCMC samples** from the posterior distribution of the DAGs
- The samples can be **queried**
- Efficiently sample DAGs

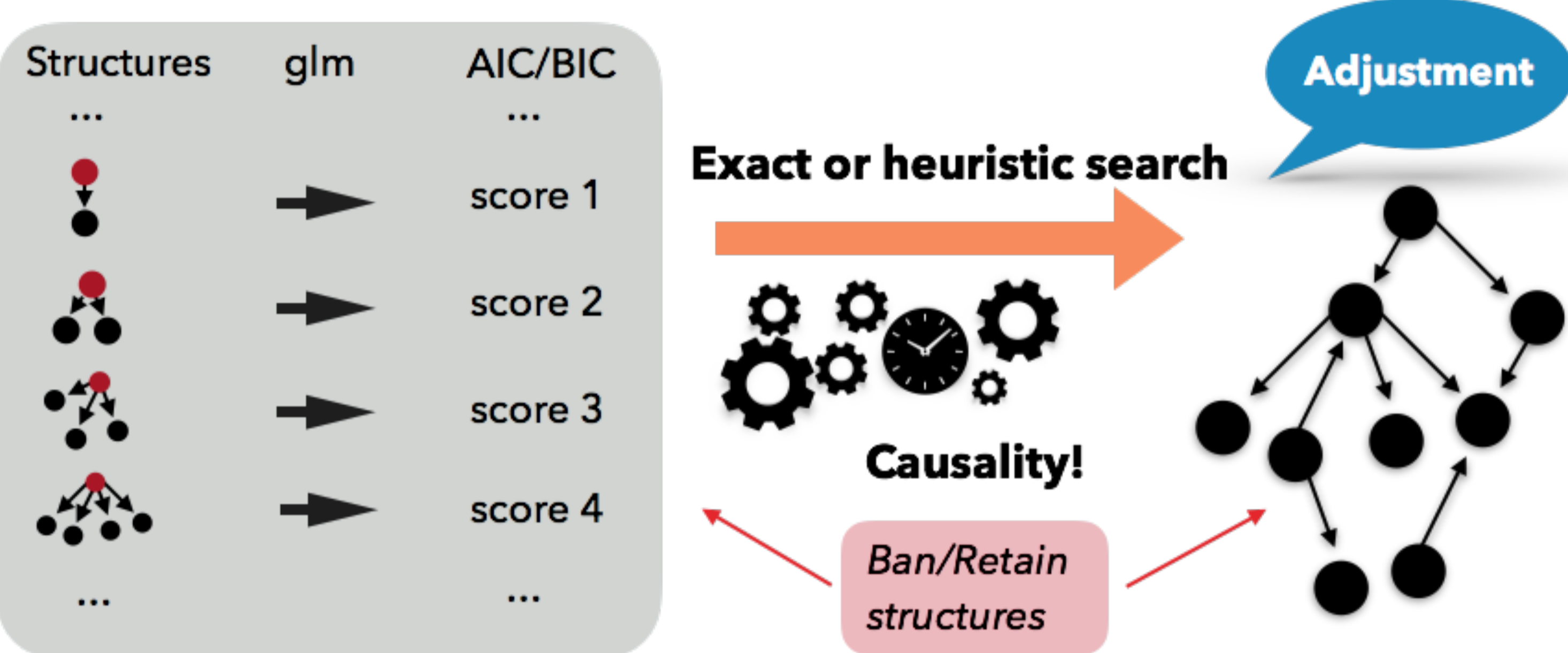
<https://www.math.uzh.ch/pages/mcmcabn/>

<https://CRAN.R-project.org/package=mcmcabn>

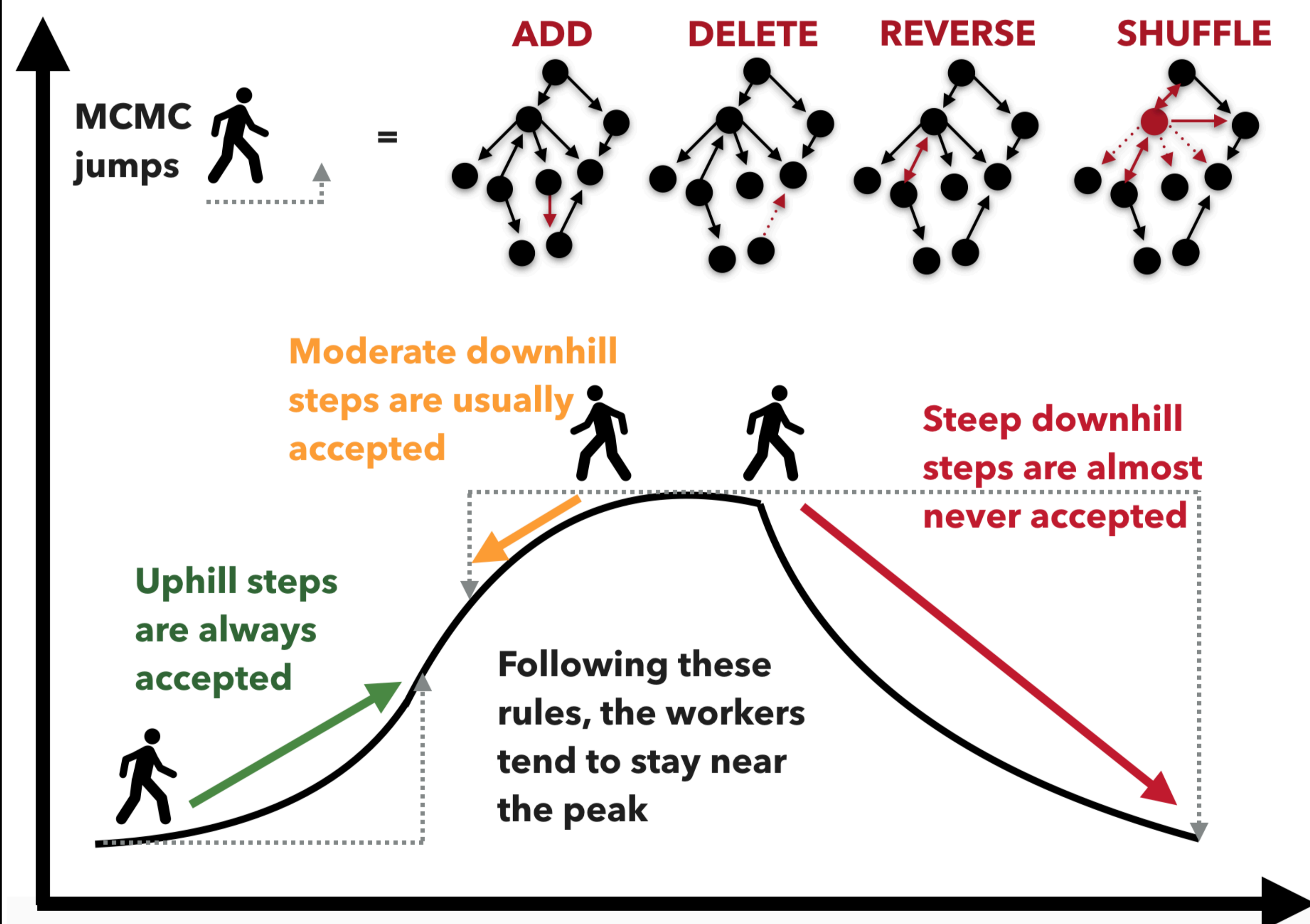
R package

- Structural MCMC sampler¹: **mcmcabn**
- Flexible** implementation
- Three** priors:
 - Uniformative
 - Penalized complexity
 - User defined

How to fit an Additive Bayesian Network from data with the search-and-score method²?



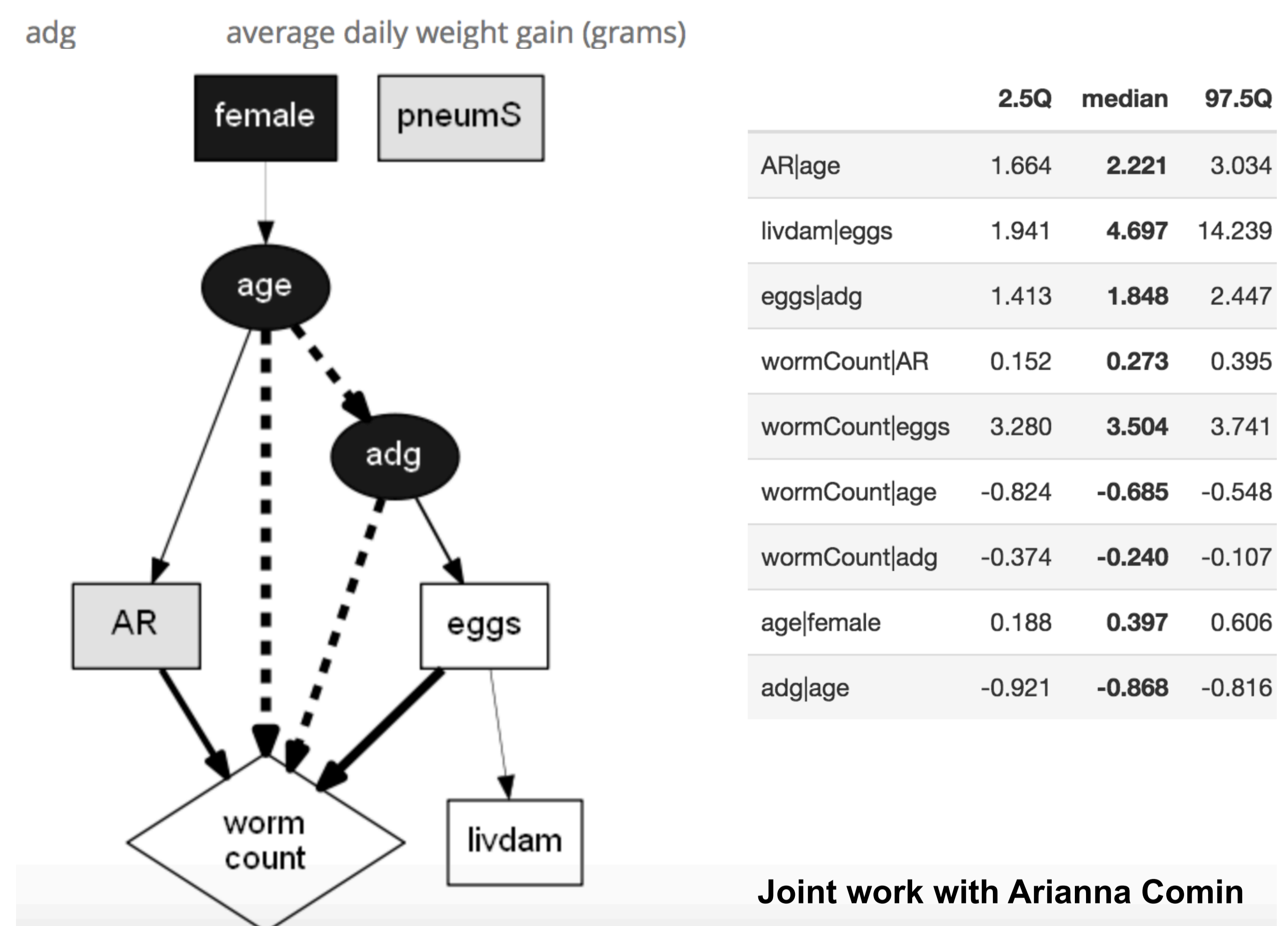
How to perform MCMC over structures?



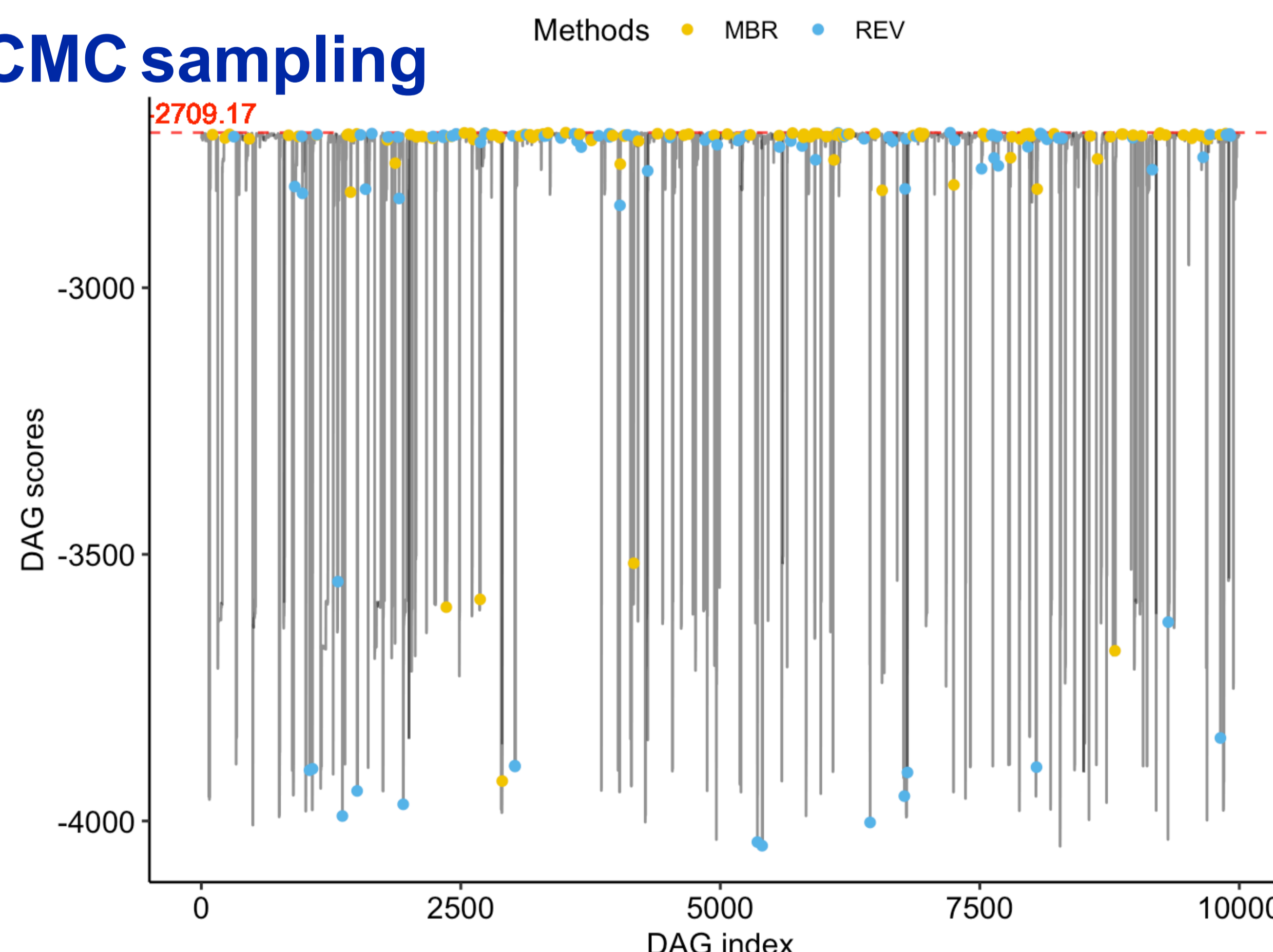
Illustrative example³

Adaptation of the **pig adg** dataset. 8 variables, n = 341 observations

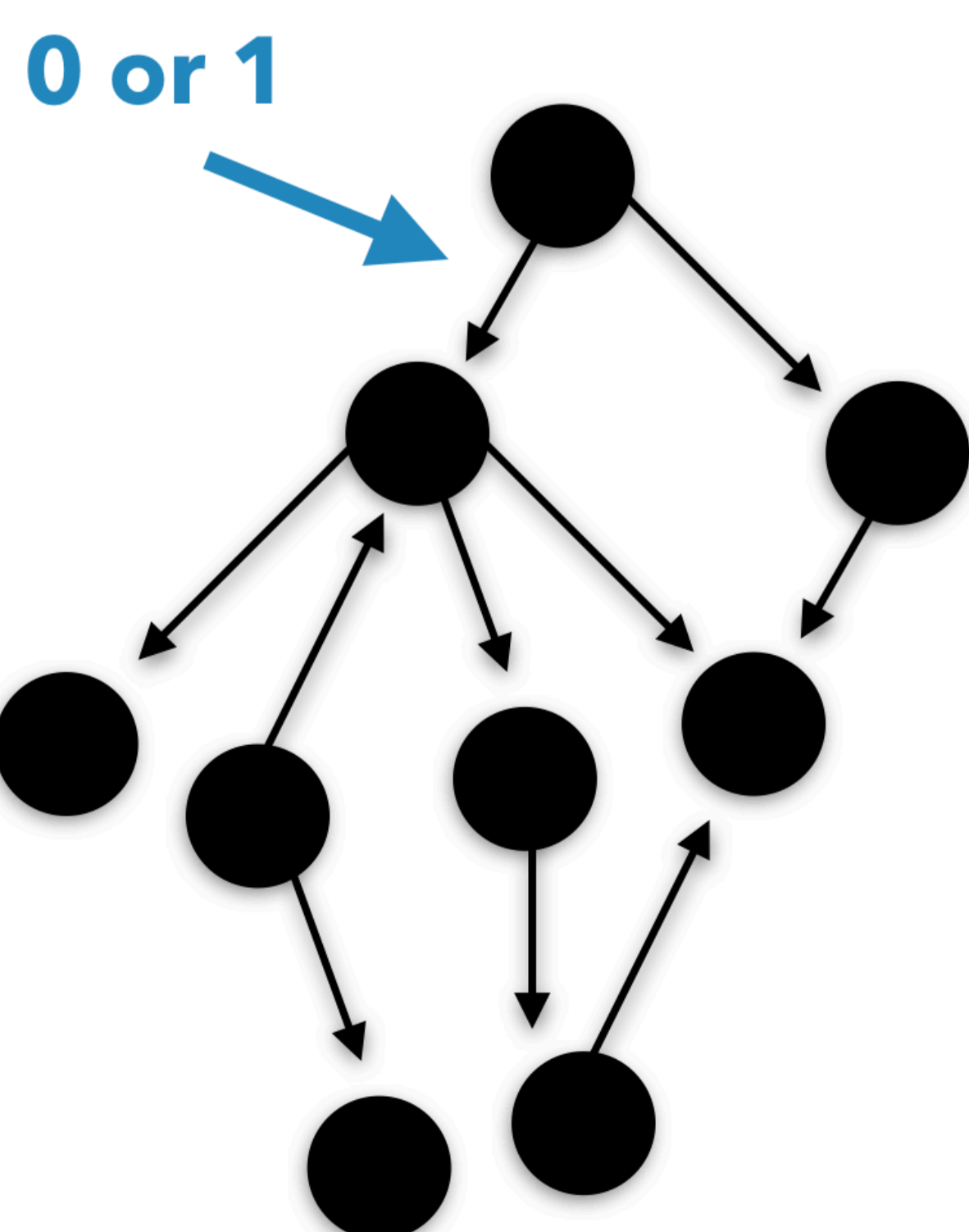
Variable	Meaning
AR	presence of atrophic rhinitis (0/1)
pneumS	presence of moderate to severe pneumonia (0/1)
female	sex of the pig (1=female, 0=castrated)
livdam	presence of liver damage (0/1)
eggs	presence of fecal/gastrointestinal nematode eggs at time of slaughter (0/1)
wormCount	count of nematodes in small intestine at time of slaughter (nr.)
age	days elapsed from birth to slaughter (days)



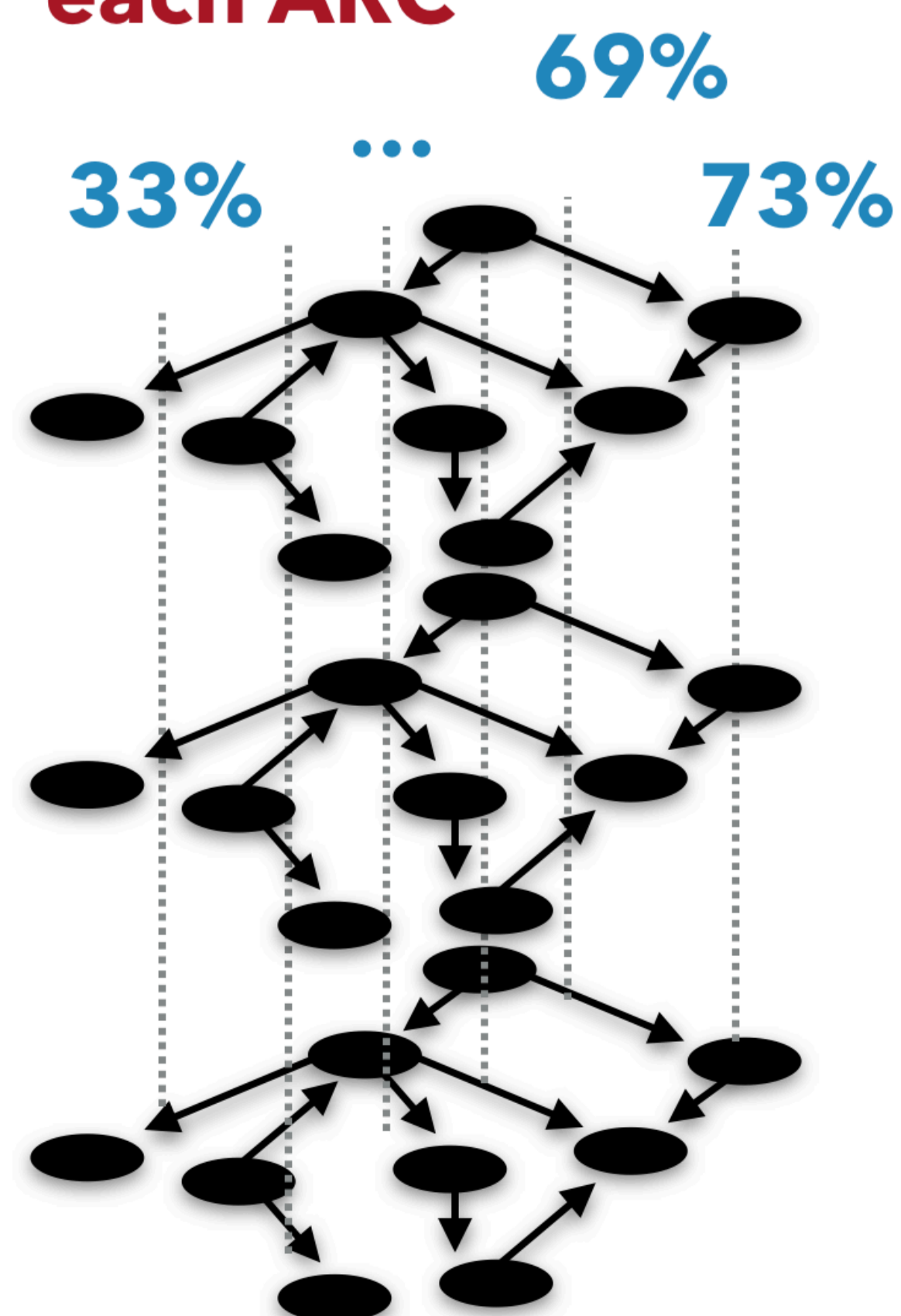
MCMC sampling



Best Unique Bayesian Network



Counting prevalence of each ARC



References

- Kratzer, G. and Furrer, R. (2016) "Is a single unique Bayesian network enough to accurately represent your data?". arXiv preprint arXiv:1902.06641.
- Lewis, F. I. et al. "Structure discovery in Bayesian networks: An analytical tool for analysing complex animal health data". Preventive Veterinary Medicine 100.2 (2011): 109-115.
- Dohoo, I. R., Martin, W., & Stryhn, H. (2003). Veterinary epidemiologic research (No. V413 DOHv). Charlottetown, Canada: AVC Incorporated.

