Phys. Rev. E 97, 022315 (2018)

https://github.com/tatsuro-kawamoto/graphBIX

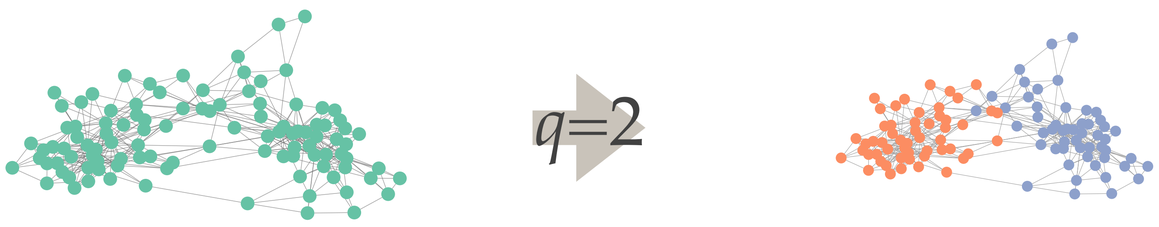
# Assessment of the number of clusters in community detection using the alluvial diagram

Tatsuro Kawamoto (AIST, Japan),
Yoshiyuki Kabashima (Tokyo Tech., Japan)

## Community detection

For various numbers of clusters, perform a Bayesian inference using the stochastic block model

MCMC, Belief Propagation, Variational Bayes, etc.




Then, determine the number of clusters $q$ that most efficiently describes the network.

## 2 step procedure to determine q

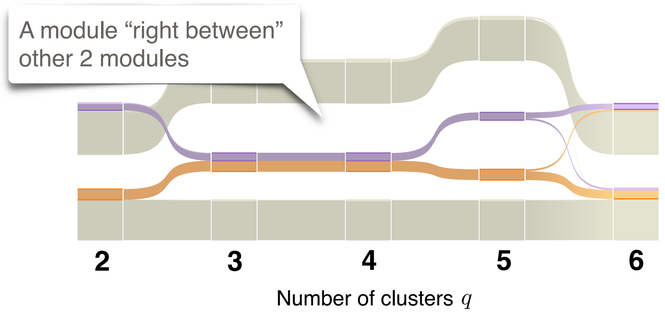
1. Estimate the “range” of appropriate number of clusters using **information criteria**.
2. Finer inspection of the way the network is actually partitioned.

For the 2nd step,

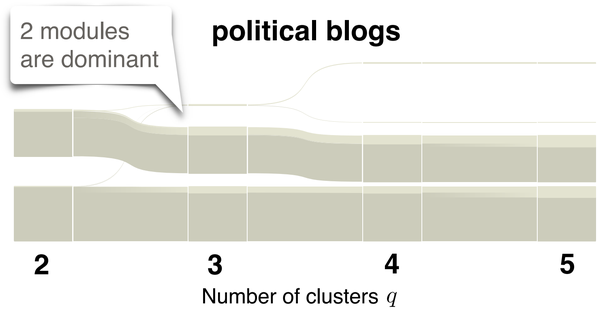
## Alluvial diagrams are very useful.

1. Generate a “.smap” file @ https://github.com/tatsuro-kawamoto/graphBIX
2. Generate the alluvial diagram from the “.smap” file @ http://www.mapequation.org/

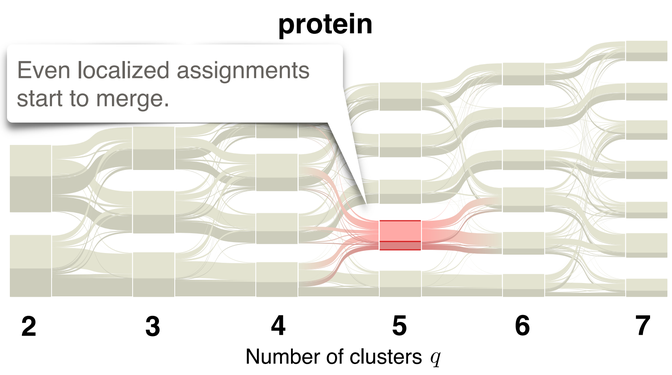
**political books**



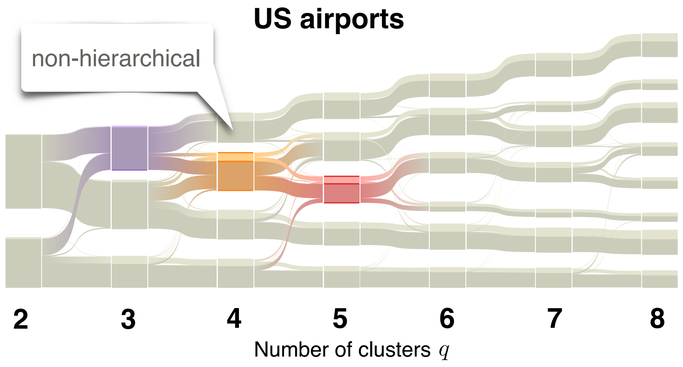

**political blogs**



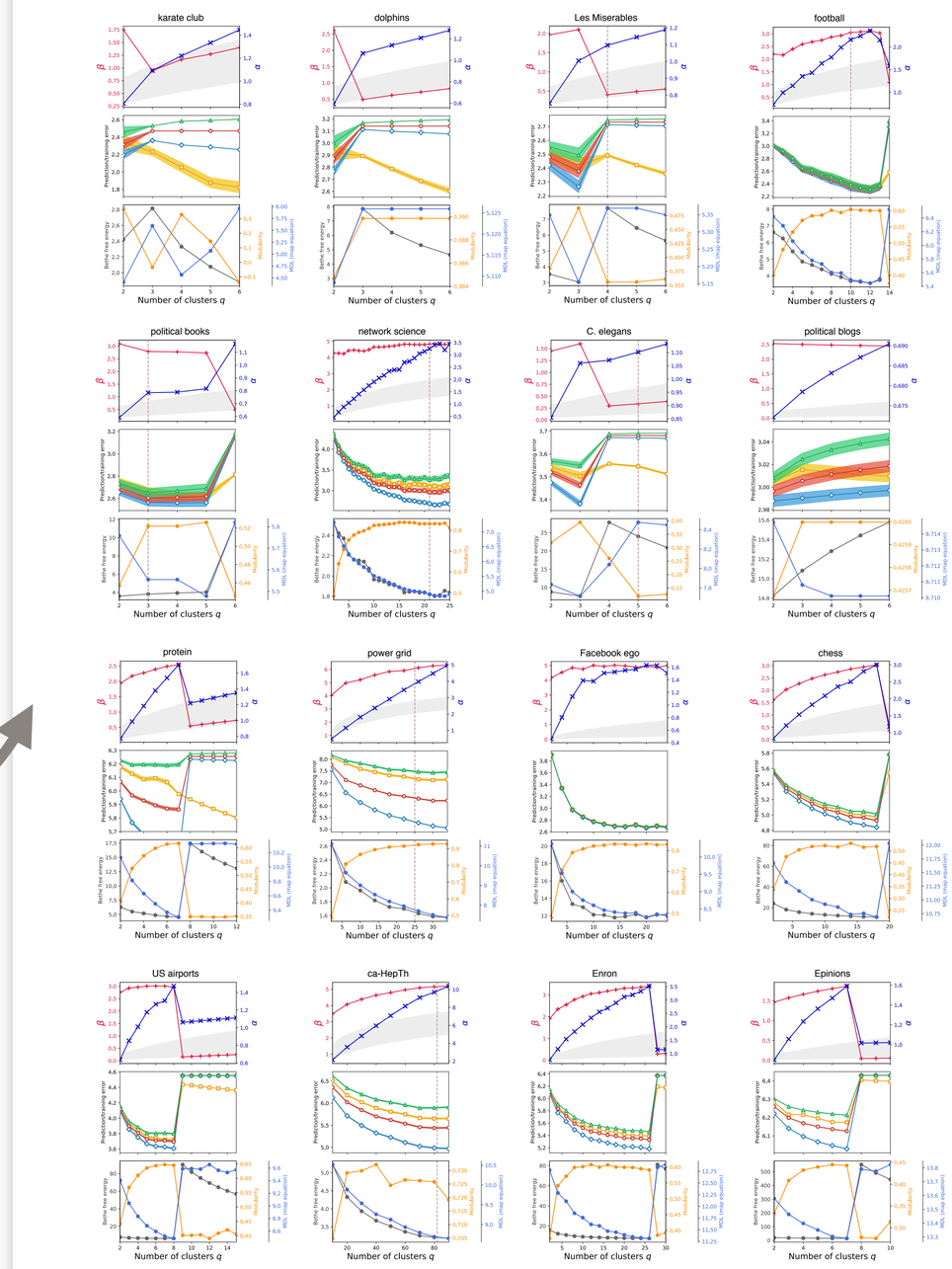

**protein**




**US airports**




### Assessment on various real-world networks




- Estimated model parameters (top)
- Prediction errors (middle)
- Modularity, Map equation, Bethe free energy (bottom)

**Alluvial diagram is an extension of the Sankey diagram**

- Bundles represent clusters.
- Variation of cluster assignments for different (input) number of clusters.
- Statistical significance of assignments can be represented by color gradient.

## Appropriateness of partitions can be assessed visually.

[original paper of the alluvial diagram] M. Rosvall, C. T. Bergstrom, PLoS ONE 5(1): e8694 (2010).
[Application for the cluster assessment] T. Kawamoto, Y. Kabashima, Phys. Rev. E 97, 022315 (2018).

- The partition may mostly consists of insignificant cluster assignments.
- Candidate partitions may be all similar (e.g., only the difference of tiny clusters).
- The partitions may constitute a non-hierarchical structure.

A more interesting example

Human face recognition (by human)