Comparison of blockmodeling approaches for dynamic networks with newcomers and departure nodes by Monte Carlo simulation

Marjan Cugmas & Aleš Žiberna

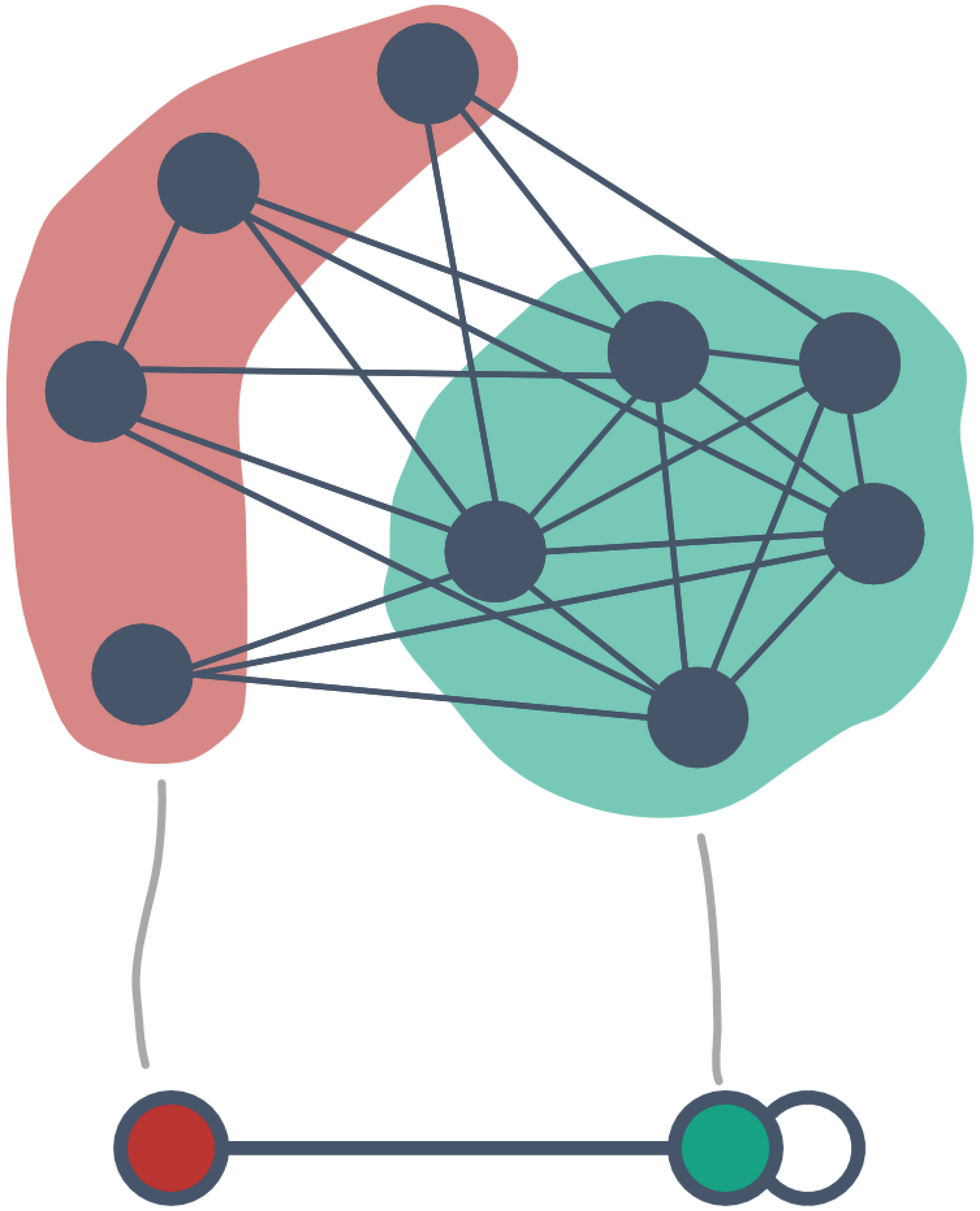
Faculty of Social Sciences, University of Ljubljana

EUSN 2023

University *of Ljubljana*
Faculty *of Social Sciences*

Blockmodeling

With blockmodeling we can study the relationships between the units.



Blockmodeling is clustering approach for reducing large, potentially incoherent network to a smaller, comprehensible structure that is easier to interpret.

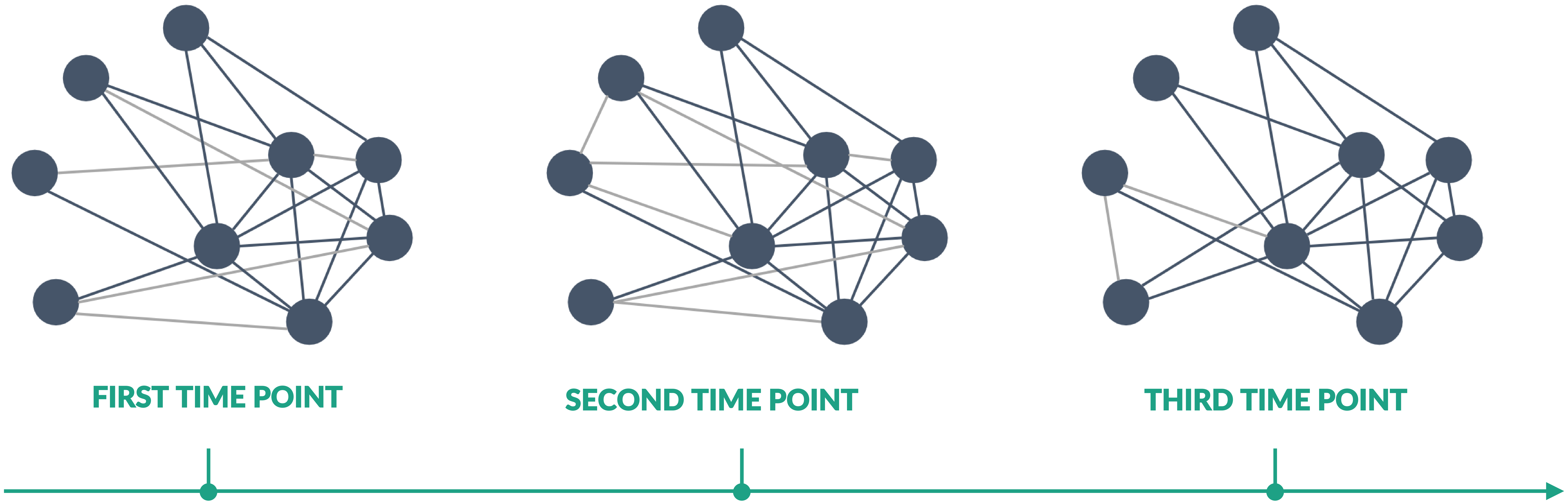
The result of blockmodeling is a partition of equivalent (according to their links in the network) nodes and an image matrix representing the links between and within the obtained clusters.

The term block refers to the links between two clusters and within one cluster.

Dynamic networks

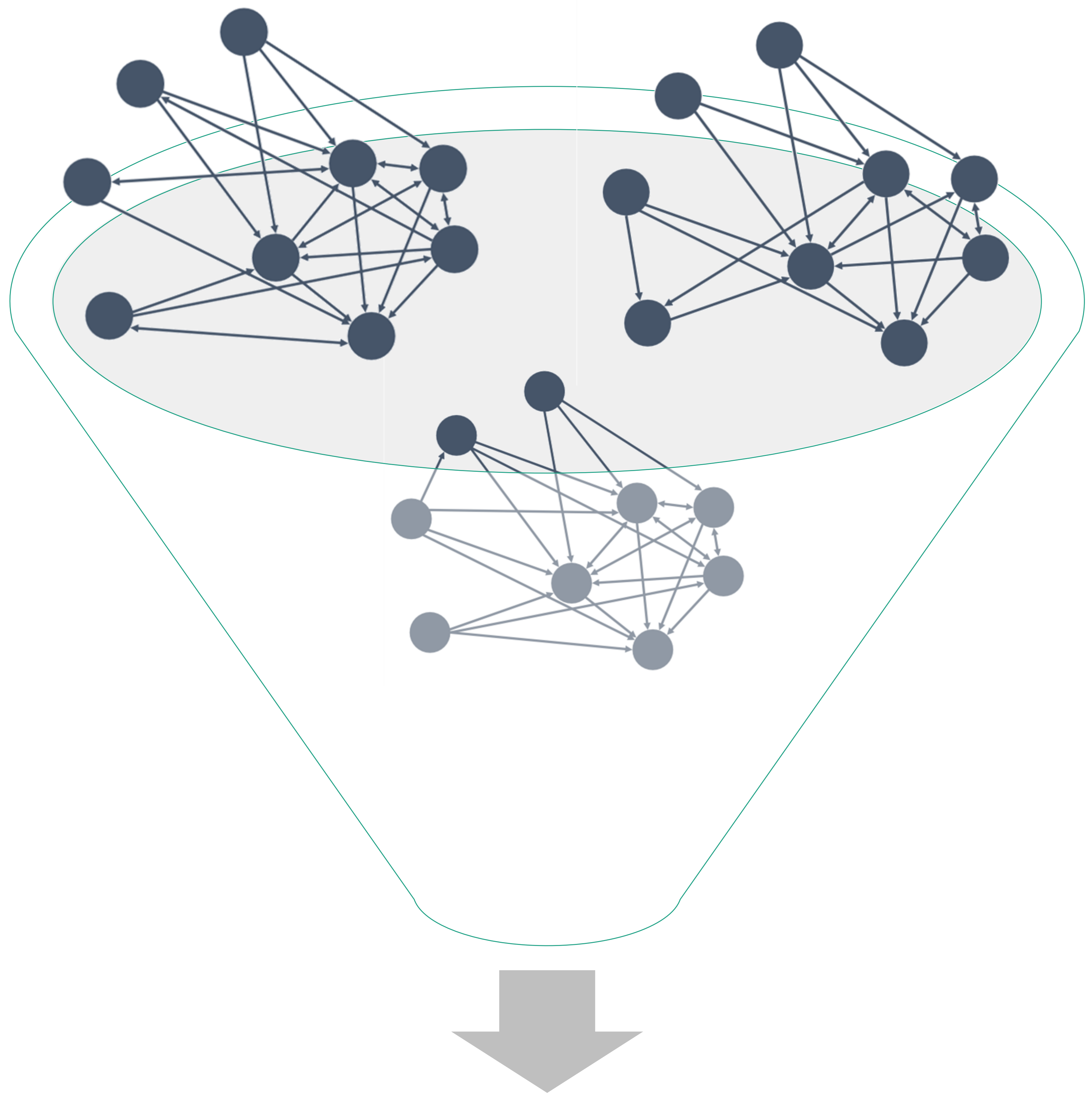
Several types of dynamic networks exists. Here, the focus is on networks, measured at multiple points in time.

- Snapshot networks: most of nodes are present at all time points and the same type of relations is measured.
- Example: a survey of mutual friendships among high school students in February, March and April.



Blockmodeling of dynamic networks

The idea is to take advantage of the fact that consecutively observed networks are dependent.



Partitions for each time point.

Validity

Considering dependency between the ties from different time points can increase validity of the results.

The aim

To identify equivalent groups and ties among the identified groups for each time point and to study how they change in time.

Recent development

Most approaches for blockmodeling dynamic networks were developed in recent years.

Previous work

The two approaches have the highest performance in a wide range of scarious.


Social Networks 73 (2023) 7–19
Contents lists available at ScienceDirect
Social Networks
journal homepage: www.elsevier.com/locate/socnet
Approaches to blockmodeling dynamic networks: A Monte Carlo simulation study
Marjan Cugmas*, Aleš Žiberna
ARTICLE INFO
ABSTRACT
Introduction

We have studied the performance of various blockmodeling algorithms that can be applied on dynamic networks, including approaches for multilevel networks, linked networks, and dynamic networks.

FACTORS

Network size, blockmodel type, stability of partitions in time, and local network mechanisms affect the performance of different blockmodeling approaches for temporal networks.

APPROACHES

Using dynamic blockmodeling approaches is usually superior to using separate blockmodeling approaches. The most affective are

- **Dynamic Stochastic Blockmodel** (Matias and Miele 2017) (if blockmodel type does not change);
- **Stochastic Blockmodel for Multipartite Networks** (Bar-Hen et al. 2020) (if blockmodel type change).

LIMITATIONS

Simulations were done for asymmetric networks and without incomers or outgoers.

Cugmas, M., & Žiberna, A. (2023). Approaches to blockmodeling dynamic networks: A Monte Carlo simulation study. *Social Networks*, 73, 7-19.

The aim

Addressed by Monte Carlo simulations.

Empirically compare blockmodeling approaches on symmetric networks with incomers and outgoers.

Evaluate sensitivity to the basic network characteristics.

Propose guidelines for choosing blockmodeling approaches.

1

NETWORKS WITH DIFFERENT PROPERTIES

Different network characteristics are considered, such as network size, blockmodel type, etc.

2

NETWORKS RESEMBLE REAL WORD NETWORKS

The networks are generated by considering local network mechanisms which makes them closer to the real-world networks.

3

KNOWN BLOCKMODEL TYPES AND PARTITIONS

The networks are generated such that blockmodel types and partitions are known. Both can change in time.

Approaches for dynamic networks

Different approaches implemented in R, Python or MATLAB are considered.

DSBM

Matias & Miele (2016)

Statistical clustering of temporal networks through a dynamic stochastic block model

SBMfMLV

Chabert-Liddell (2022)

A stochastic block model approach for the analysis of multilevel networks /…/

BSBMfLN

Peixoto (2020)

Bayesian stochastic blockmodeling

SBMfMPN

Bar-Hen et al. (2020)

Block models for generalized multipartite networks

SBMfLN

Škulj & Žiberna (2021)

Stochastic blockmodeling for linked networks

KMfLN

Žiberna (2020)

K-means-based algorithm for blockmodeling linked networks

Stochastic blockmodeling: assume an underlying statistical model and estimate it by maximizing some likelihood-based measure.
A model enables statistical inference.

Deterministic blockmodeling: iterative algorithm search for homogenous blocks in term of tie values.

Conditional cluster probabilities: cluster probabilities in a current time point depend on cluster membership in a previous time point(s).

Linked and multipartite networks: a collection of at least two one-mode networks and one two-mode network linking these one-mode networks. In the context of dynamic networks, the two-mode networks "link" the same units from different time points. Such network is blockmodeled as a single network (with the restriction that nodes from different one-mode networks can not mix).

Within group ties probabilities are fixed in time.

Like DSBM but without some restrictions (e.g., fixed number of groups, fixed diagonal blocks).

Assumes Poisson distribution of links.

They enable weighting different parts (e.g., one-mode and two -mode) of a network.

Considered factors

Detailed descriptions follow on the next slides.

NETWORK SIZE

Small (48 nodes) and
large (96 nodes) networks.

INCOMERS & OUTGOERS

Nodes can join or leave the network anytime.

GROUPS' STABILITY

Nodes can change group membership.

MECHANISMS

Inconsistencies are generated randomly or by local mechanisms.

BLOCKMODEL TYPES

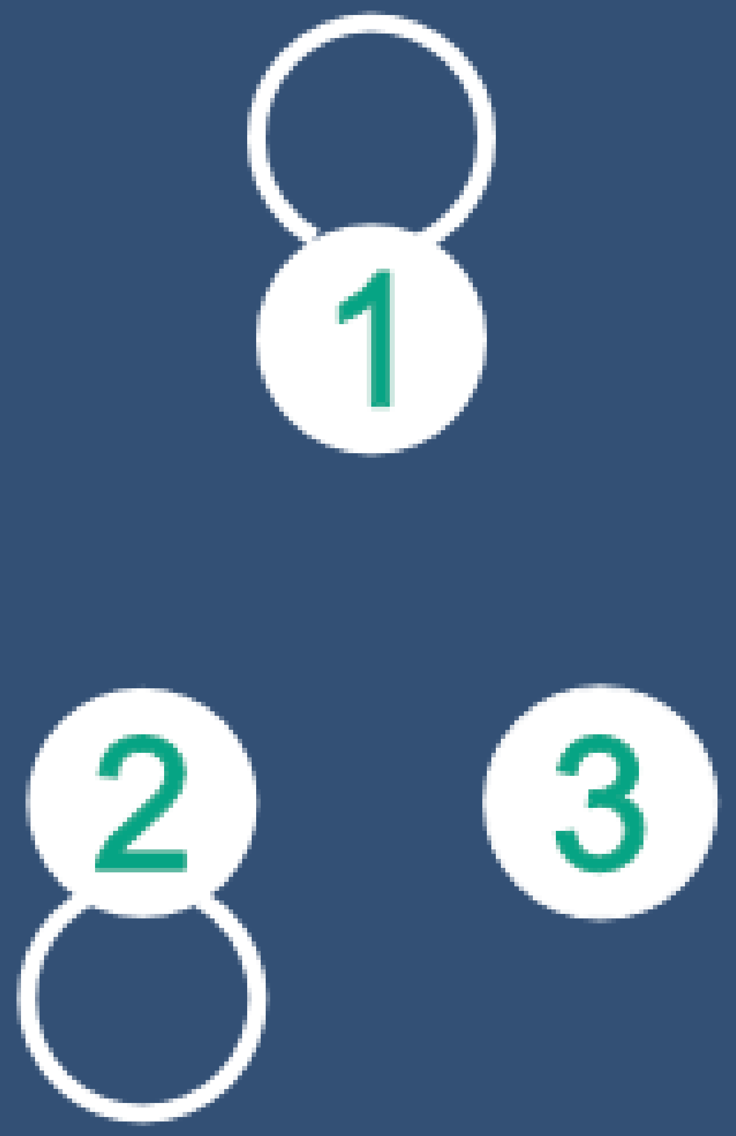
They remain the same or change in time.

Three groups are in all generated networks.

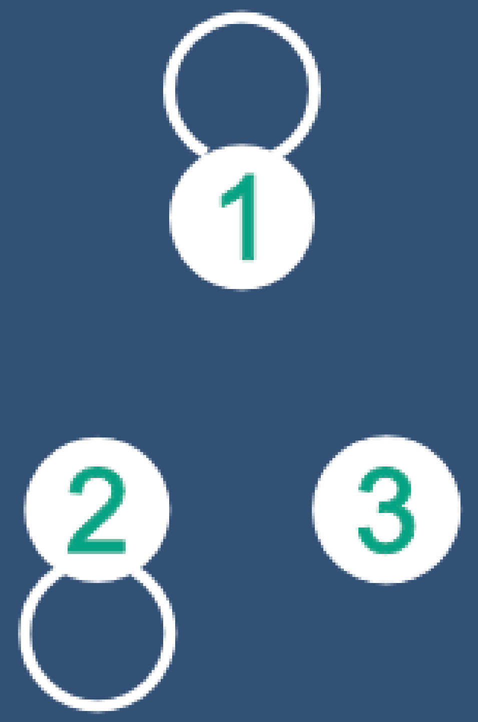
BLOCKMODEL TYPES

Different blockmodel types and different transitions (changes) of blockmodel types are considered.
Blocks with density around 5% are considered null, and blocks with density around 25% are considered complete.

INITIAL BLOCKMODEL TYPE

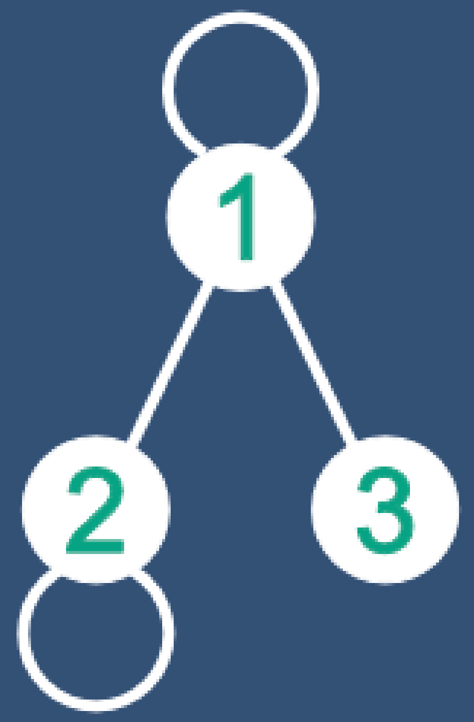


FINAL BLOCKMODEL TYPE



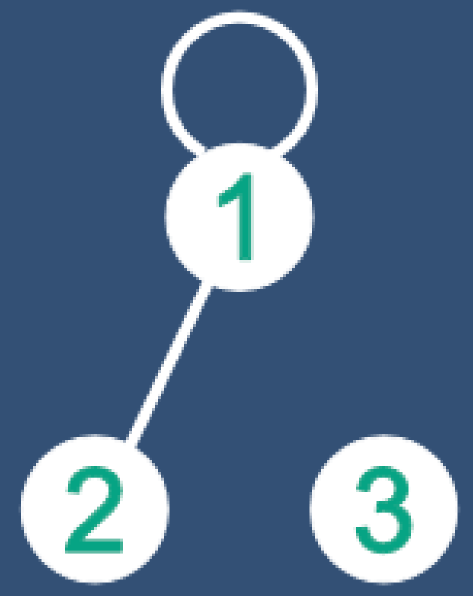
A

Nothing changes.



B

Two off-diagonal blocks change.



C

One diagonal and one off-diagonal blocks change.

INCOMERS & OUTGOERS

Some units can join (incomers) or leave (outgoers) the network.

When incomers join the network, they come in without any link to the others. A group to which they join is determined with probability that is proportional it's size.

All the units have the same probability to leave the network. The only two exceptions are incomers and the units from groups with less than 5 units.

Different shares of incomers and outgoers are considered.

- **0% incomers / 20% outgoers**
- **0% incomers / 0% outgoers**
- **20% incomers / 20% outgoers**
- **20% incomers / 0% outgoers**

GROUPS' STABILITY

In real networks, it is common for some units to change group membership.

This was simulated by randomly relocating a selected share of units between clusters at each successive time-points transitions.

This procedure has no effect on the cluster sizes.

0 % Completely stable

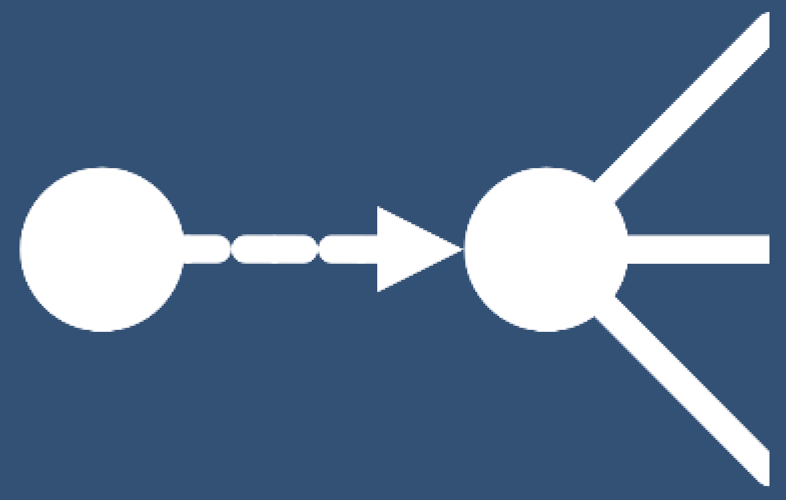
16 % Stable

33 % Unstable

100 % Completely unstable

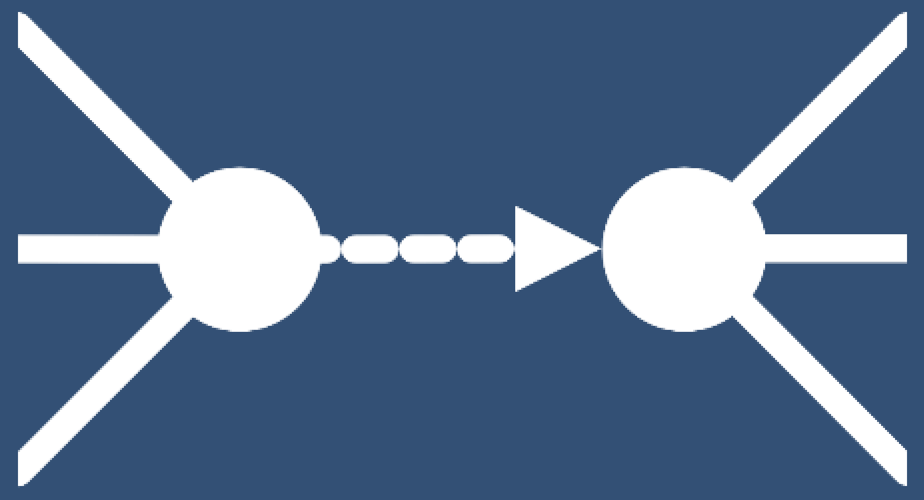
MECHANISMS

The links within blocks can be generated completely at random or based on the selected local network mechanisms (all mechanisms are assumed to have similar strengths reflected by the vector θ).



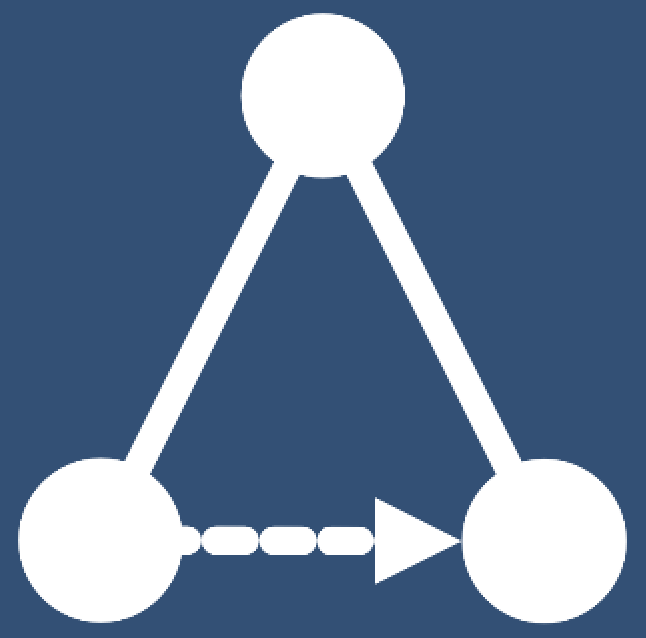
POPULARITY

Tendency to create links to those with the highest in-degree.



ASSORTATIVITY

Tendency to create links to those with similar in-degree.

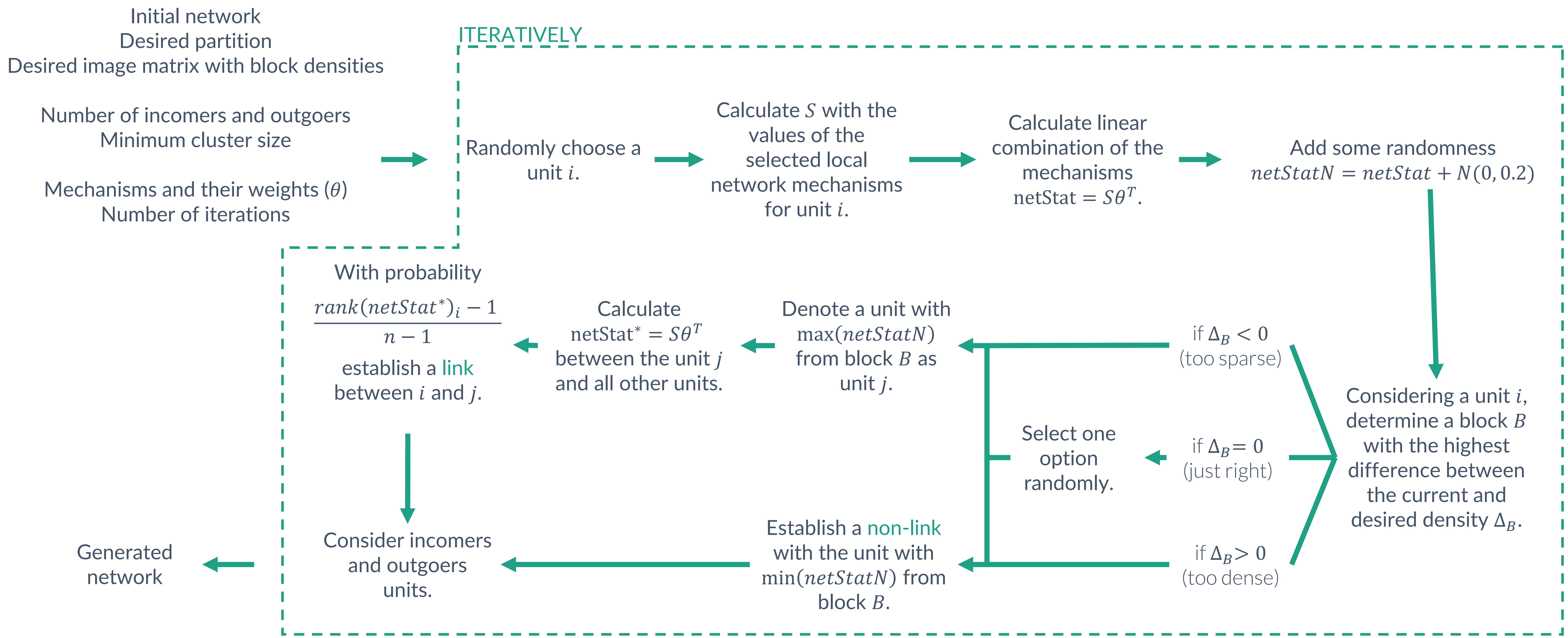


TRANSITIVITY

Tendency to create links to those who are "liked by a friend".

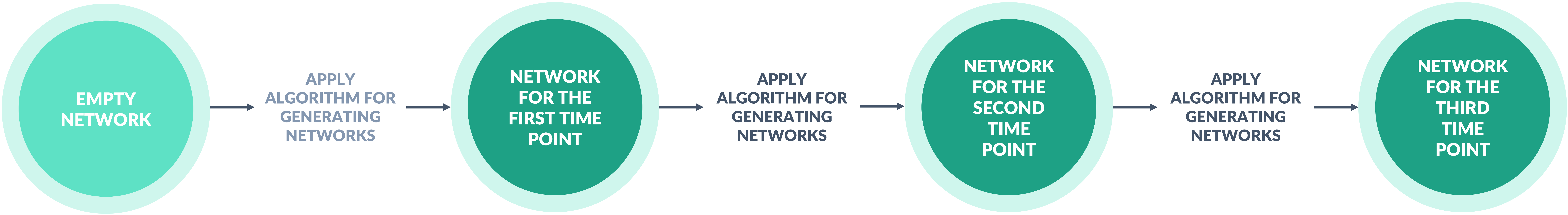
Generating networks

The number of iterations was set to 5,000.



Generating temporal networks

The algorithm for generating networks was used three times for each temporal network.



Separate blockmodeling approaches

Networks from each time points are blockmodeled separately.

SBM	KM	BSBM
Mariadassou et al. (2010)	Žiberna (2020)	Peixoto (2002)
Stochastic blockmodeling	**K-means based blockmodeling**	**Bayesian stochastic blockmodeling**
`stochBlockORP (StochBlock 0.1.2)`	`kmBlock (kmBlock 0.1.1)`	`minimize_blockmodel_dl mcmc_sweep (graph-tool 2.58)`
`rep = 1000`	`rep = 1000`	`deg_corr = False B_min=3 B_max=3`

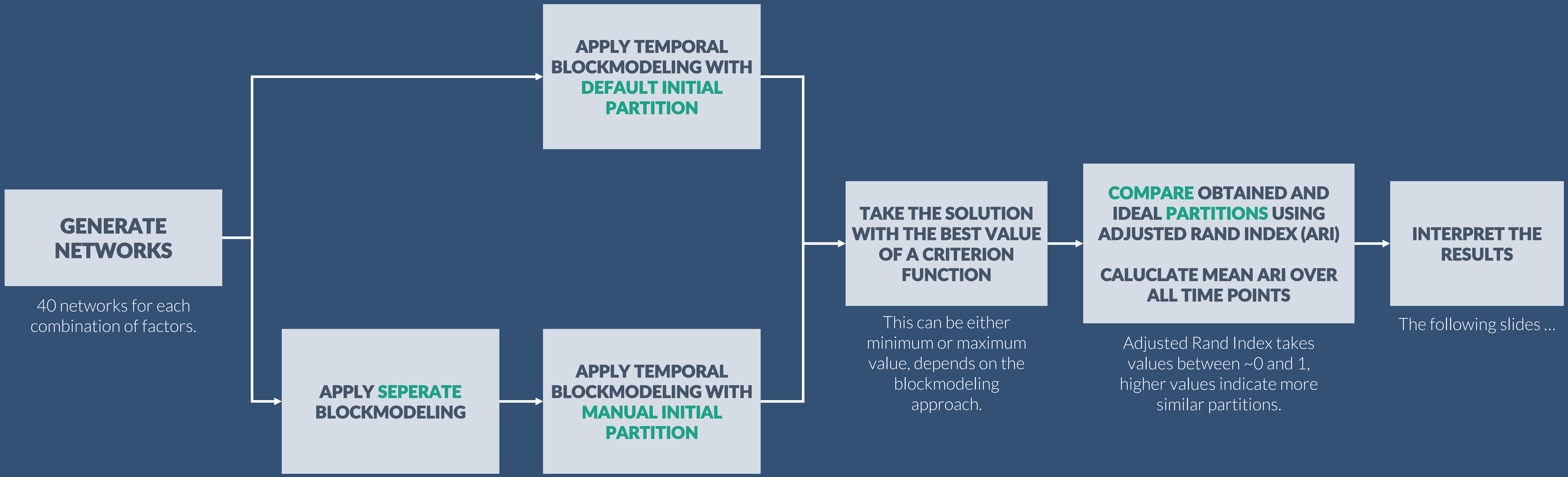
Dynamic blockmodeling

Default and manual initial partitions are considered.

DSBM	SBMfMLV	BSBMfLN	SBMfMPN	SBMfLN	KMfLN
Matias & Miele (2016)	Chabert-Liddell (2022)	Peixoto (2020)	Bar-Hen et al. (2020)	Škulj & Žiberna (2021)	Žiberna (2020)
Statistical clustering of temporal networks through a dynamic stochastic block model	**A stochastic block model approach for the analysis of multilevel networks** /…/	**Bayesian stochastic blockmodeling**	**Block models for generalized multipartite networks**	**Stochastic blockmodeling for linked networks**	**K-means-based algorithm for blockmodeling linked networks**
`select.dynsbm` `estimate.dysbm` `(dynsbm 0.7)`	`mlvsbm_create_generalized_network` `mlvsbm_estimate_generalized_network` `(MLVSBM 0.3.2)`	`minimize_blockmodel_dl` `mcmc_sweep` `(graph-tool 2.58)`	`multipartiteBMFixedModel` `(GREMLINS 0.2.0)`	`stochBlockORP` `(StochBlock 0.1.2)`	`kmBlockORPC` `(kmBlock 0.1.1)`
`iter.max = 20` `nstart = 25` `fixed.param=TRUE`	`directed = rep(FALSE, 3)` `distribution = rep("bernoulli", 3)`	`deg_corr = FALSE` `B_min=9` `B_max=9`	`maxiterVE = 200` `maxiterVEM = 200`	`rep = 1000`	`rep = 1000`
+ SBM 1. initial par. + SBM 4. initial par.	+ SBM initial par.	+ BSBM initial par.	+ SBM initial par.	+ SBM initial par.	+ KM initial par.

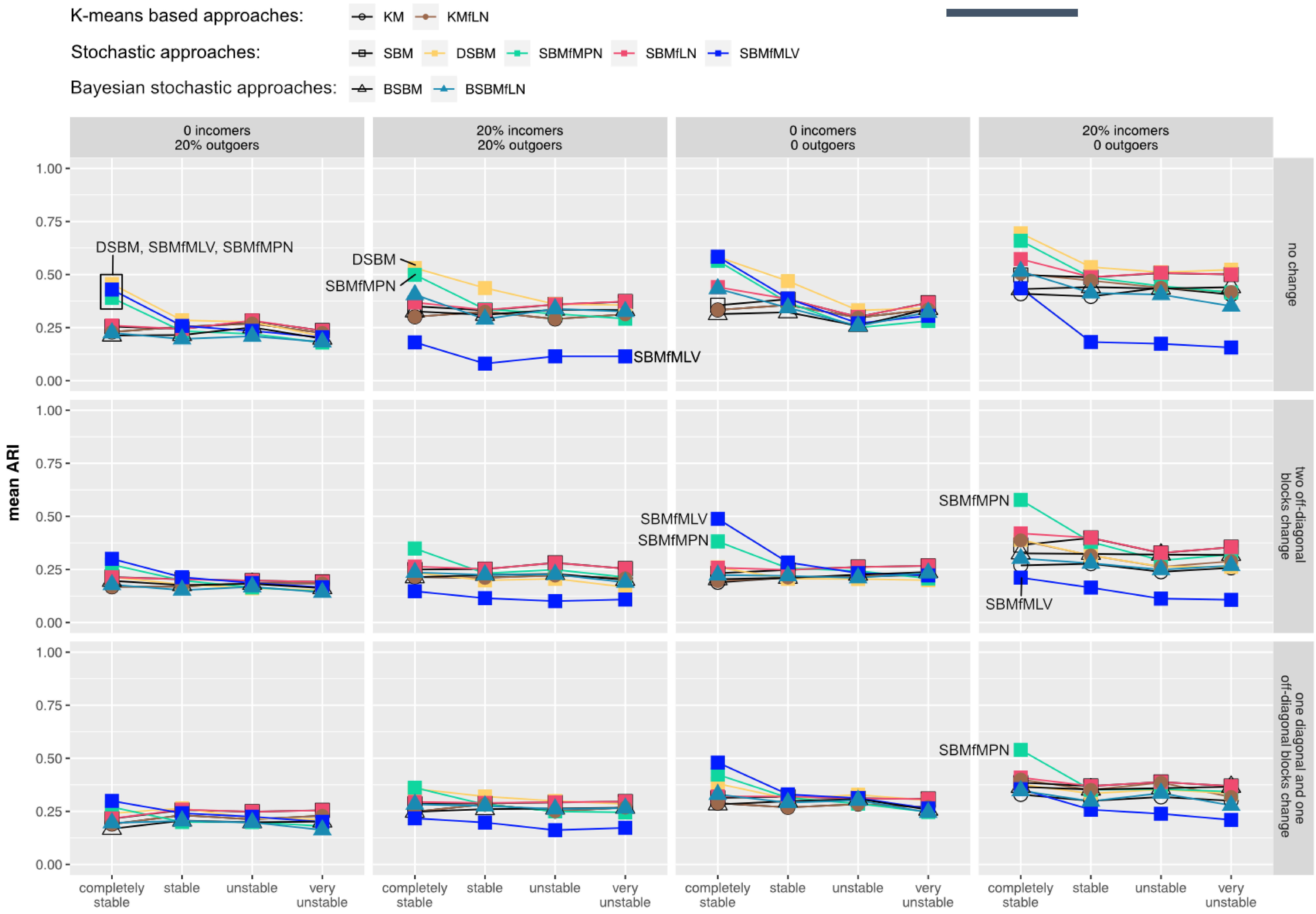
Summary of the simulation process

The results will be shown in the following slides.



Results

Small networks, mechanisms.



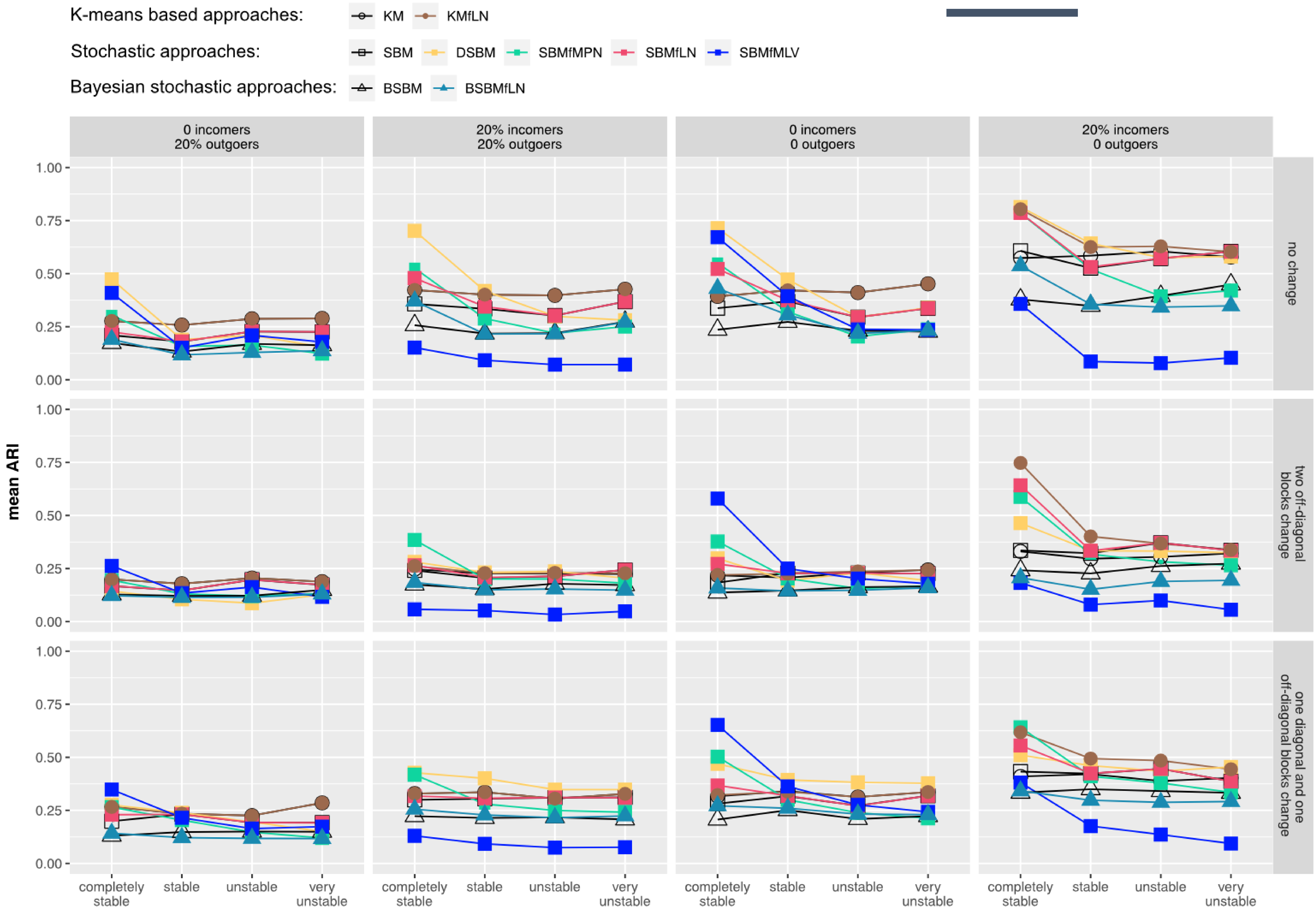
Separate blockmodeling is not affected by the stability of partitions, but it generally does not provide the best results.

Bellow are generally the most recommended approaches (in some cases they might not be the most optimal).

		Incomers / outgoers	
		No	Yes
BM type change	No	DSBM	
	Yes	SBMfMLV	SBMfMPN

Results

Small networks, random.



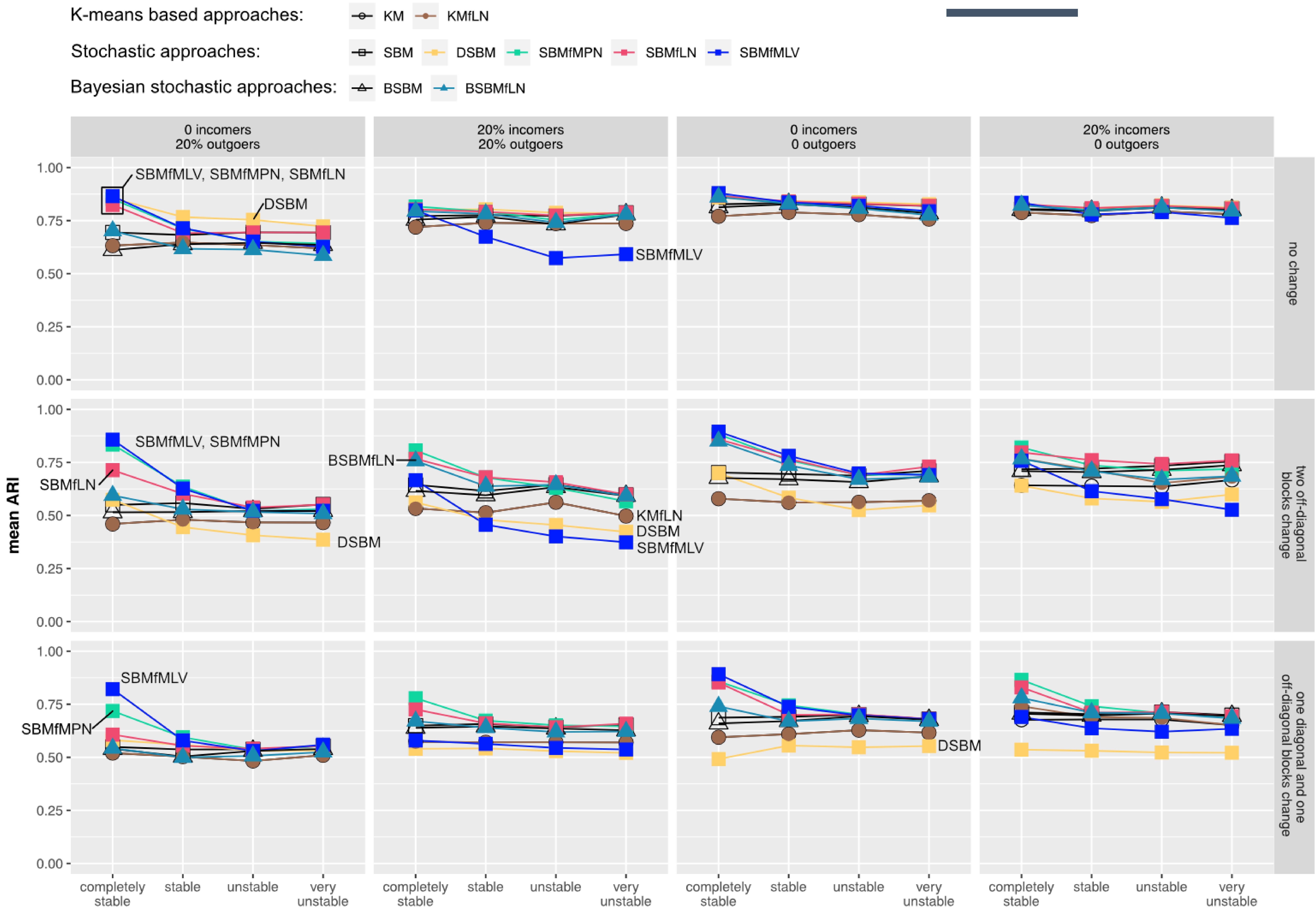
KMfLN improves in the case of random networks (especially in the case of unstable partitions) and outperform **SBMfMPN.**

Bellow are generally the most recommended approaches (in some cases they might not be the most optimal).

		Incomers / outgoers	
		No	Yes
BM type change	No	**DSBM** (stable partitions) **KMfLN** (unstable partitions)	
	Yes	**SBMfMLV** (stable) **KMfLN** (unstable)	**KMfLN**

Results

Large networks, mechanisms.



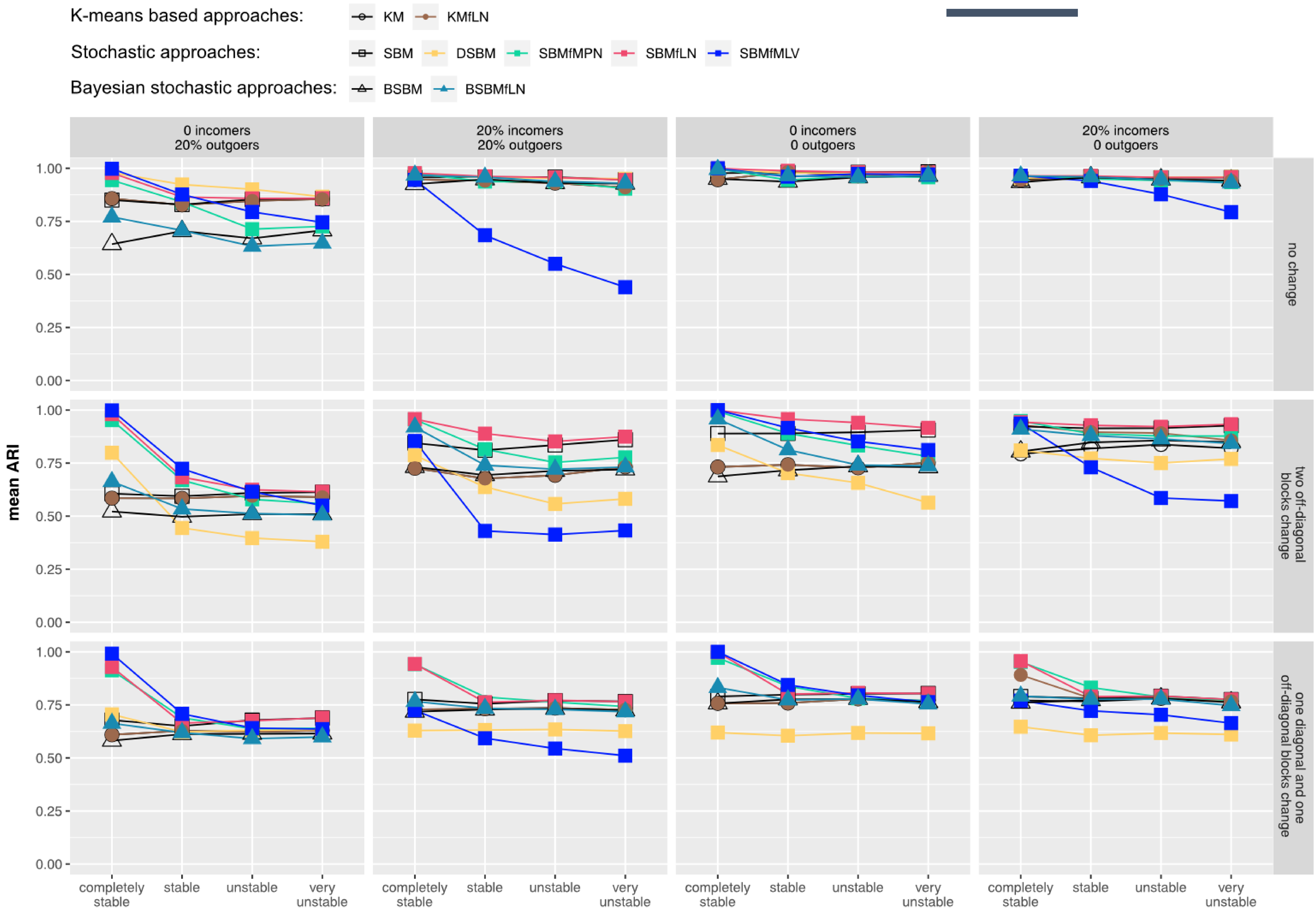
If the partitions are very unstable, temporal blockmodeling often does not outperform separate blockmodeling.

Bellow are generally the most recommended approaches (in some cases they might not be the most optimal).

		Incomers / outgoers	
		No	Yes
BM type change	No	DSBM	
	Yes	SBMfLN SBMfMLV SBMfMPN	SBMfMPN SBMfLN

Results

Large networks, random.



The **SBMfLN** in some cases (when two off-diagonal blocks changes) yield higher mean ARI, compared to **SBMfMPN**.

Bellow are generally the most recommended approaches (in some cases they might not be the most optimal).

		Incomers / outgoers	
		No	Yes
BM type change	No	DSBM	
	Yes	SBMfLN SBMfMLV SBMfMPN	SBMfLN SBMfMPN

CONCLUSIONS

FACTORS

Different network characteristics affect blockmodeling solutions.
The presence of incomers and outgoers has also a negative effect.

APPROACHES

Generally, the most worth considering approaches are:

- **Stochastic blockmodeling for multilevel networks (SBMfMLV)** (Chabert-Liddell 2022): a safe choice when there are no incomers and outgoers.
- **Stochastic blockmodeling for multipartite networks (SBMfMPN)** (Bar-Hen et al. 2020): a safe choice when there are incomers and outgoers. Might have some convergence issues.
- **K-means blockmodeling for linked networks (KMfLN)** (Žiberna 2020): generally effective but it can be outperformed by other approaches. Works best if the links within the blocks are randomly established.
- **Dynamic Stochastic Blockmodel (DSBM)** (Matias and Miele 2017): works very well if the blockmodel type does not change. Relatively (compared to other approaches) poor performance if the blockmodel type change and the network is large.

FUTURE WORK

To apply and evaluate these blockmodeling approaches on the real-world co-authorship networks.

GENERAL CONCLUSIONS

This study attempt to compare the efficiency of different blockmodeling approaches. Overall, several factors (network size, blocks' densities, local network mechanisms, etc.) affect efficiency of blockmodeling approaches.

→ Approaches not primarily developed for dynamic networks works well.
→ Using dynamic blockmodeling is usually better than separate blockmodeling.

01

PRIOR KNOWLEDGE & SEPARATE ANALYSES

Start with separate preliminary analyses to confirm your knowledge about the network.
Various factors can affect the efficiency of blockmodeling approaches.

02

TRY WITH DIFFERENT INITIAL PARTITIONS

Use different initial partitions (e.g., from separate analysis) and keep the solution with the best criterion value.

03

DSBM, SBMfMLV, SBMfMPN

Use DSBM for stable network structures,
otherwise consider SBMfMLV (without incomers/outgoers) or SBMfMPN (with incomers/outgoers).