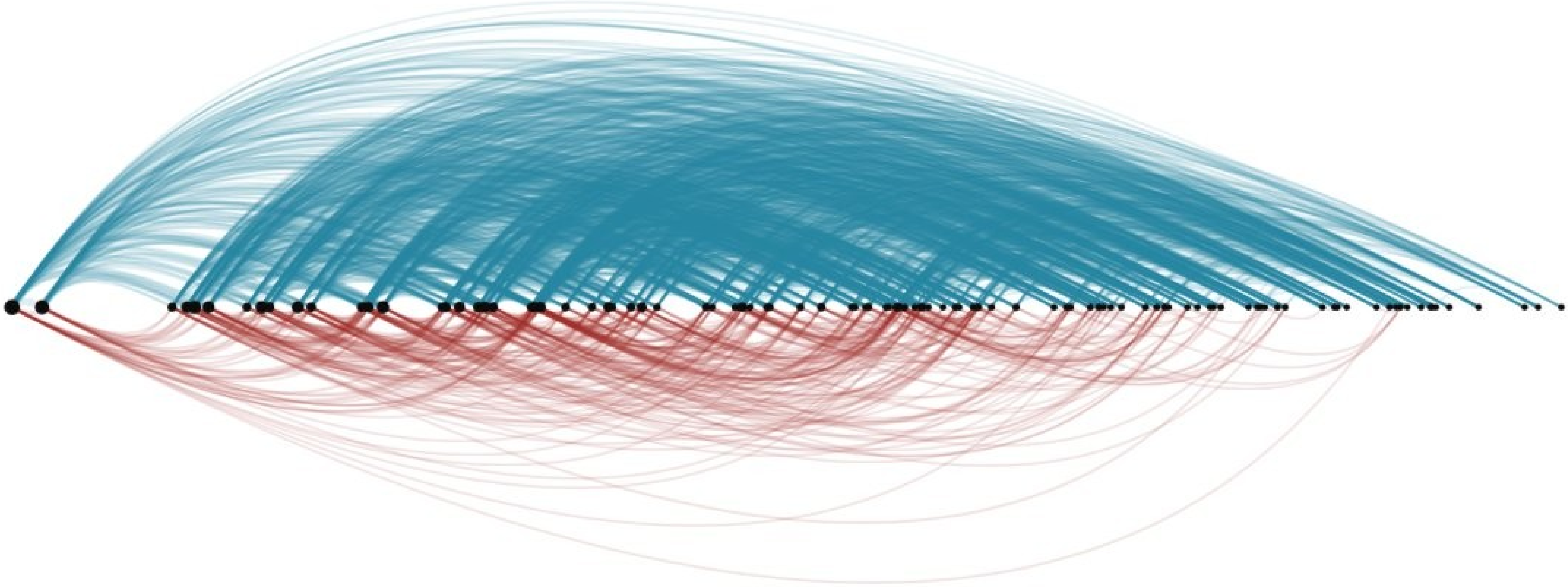
Hierarchical community structure in networks

Pre-prints available

arXiv:2009.07196
arXiv:2009.07525

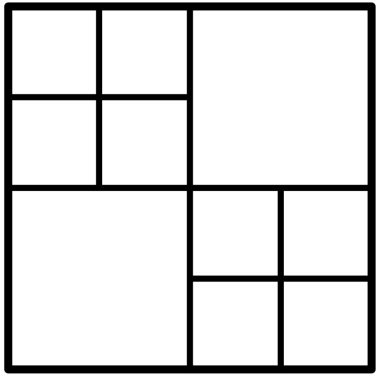
Leto Peel
Maastricht University
@PiratePeel

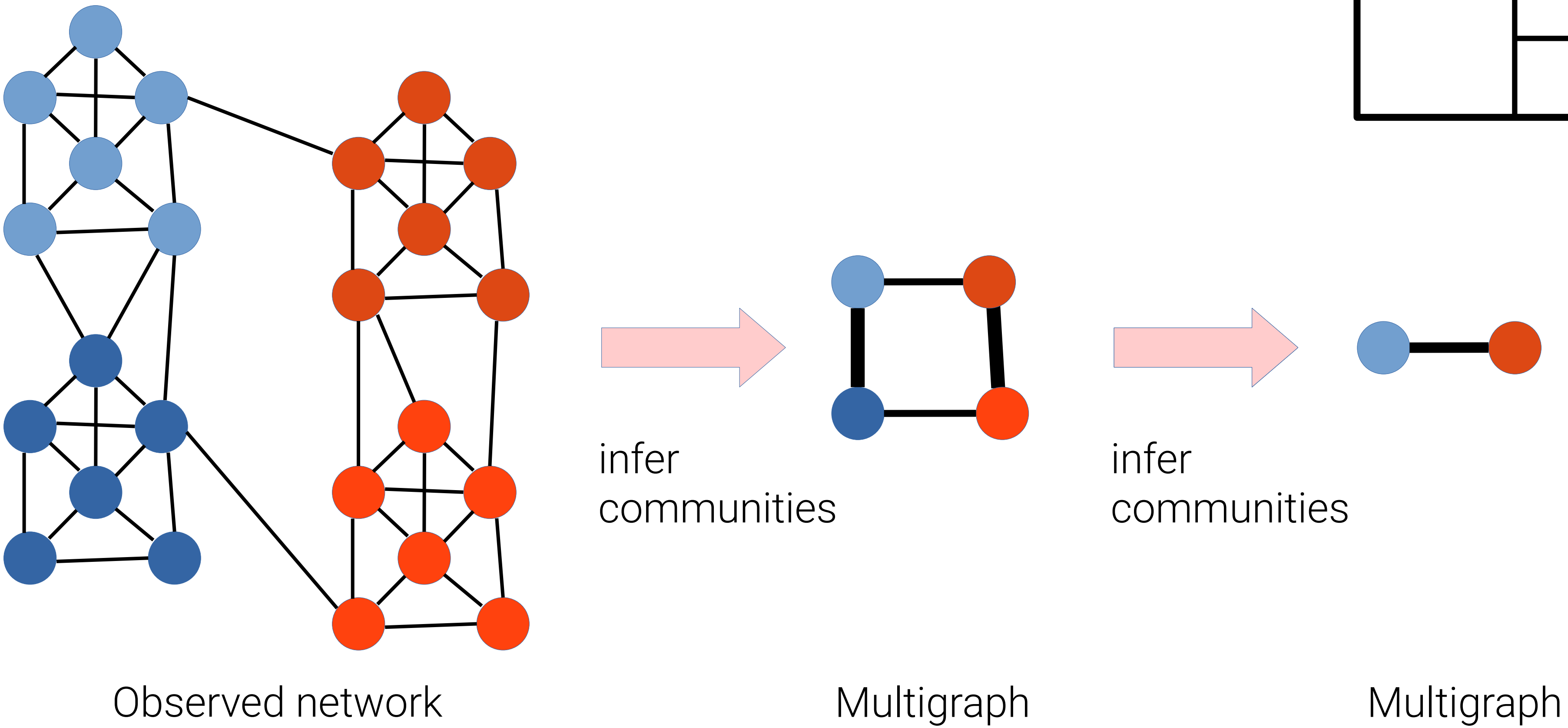


Hierarchies?

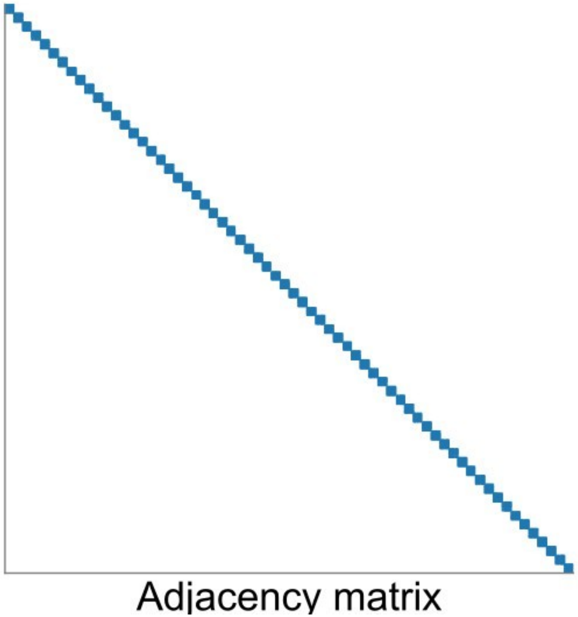
Hierarchies?

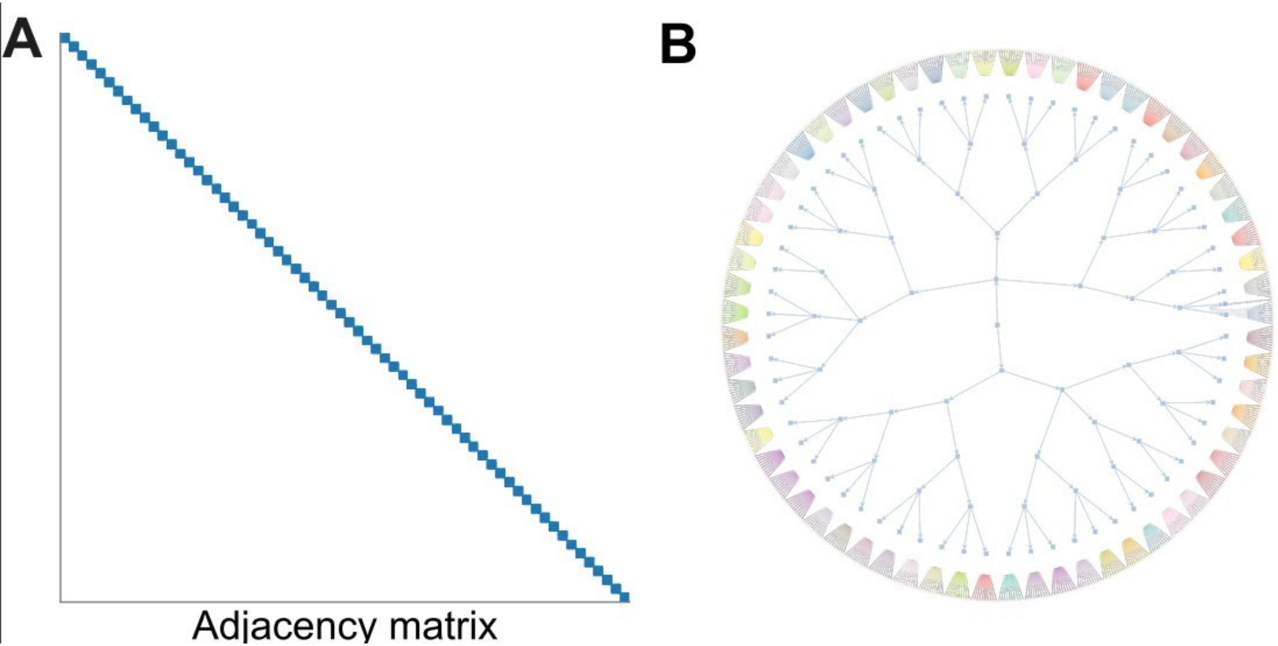
Building the hierarchy

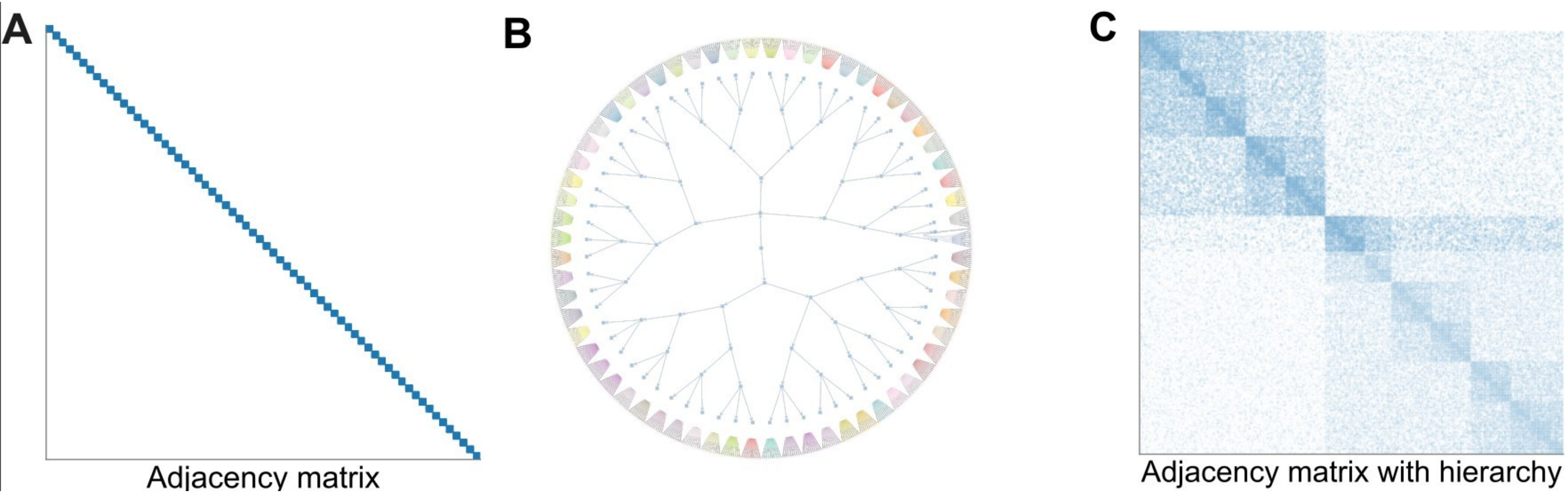




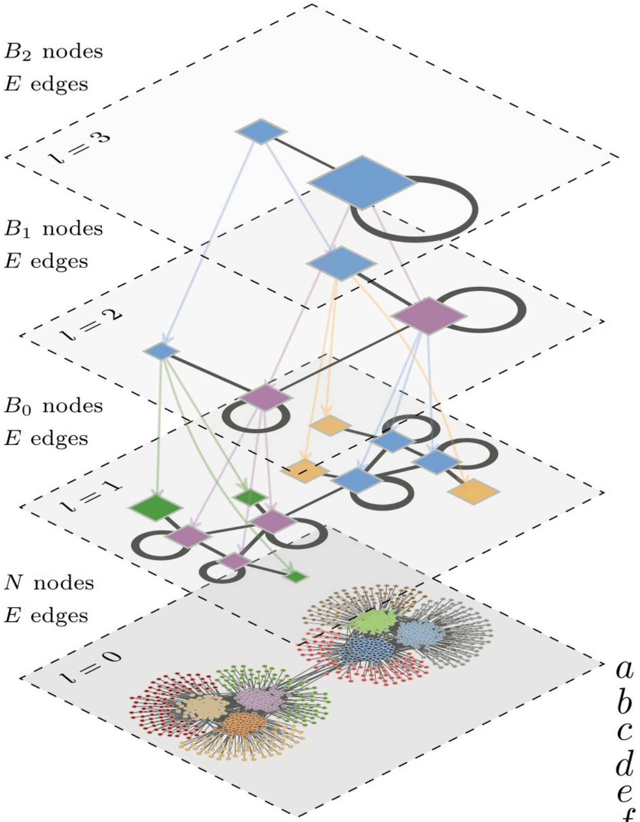
A



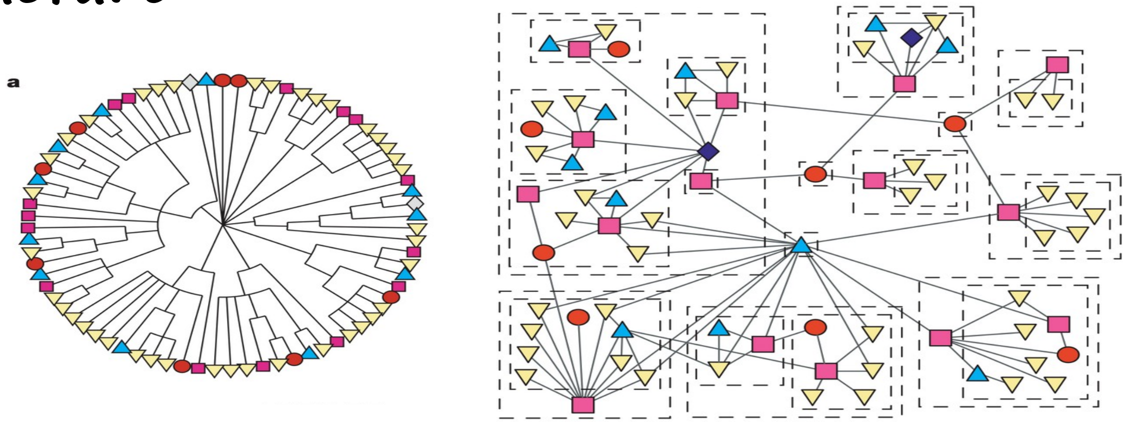

A
Adjacency matrix
B


A
Adjacency matrix
B
C
Adjacency matrix with hierarchy

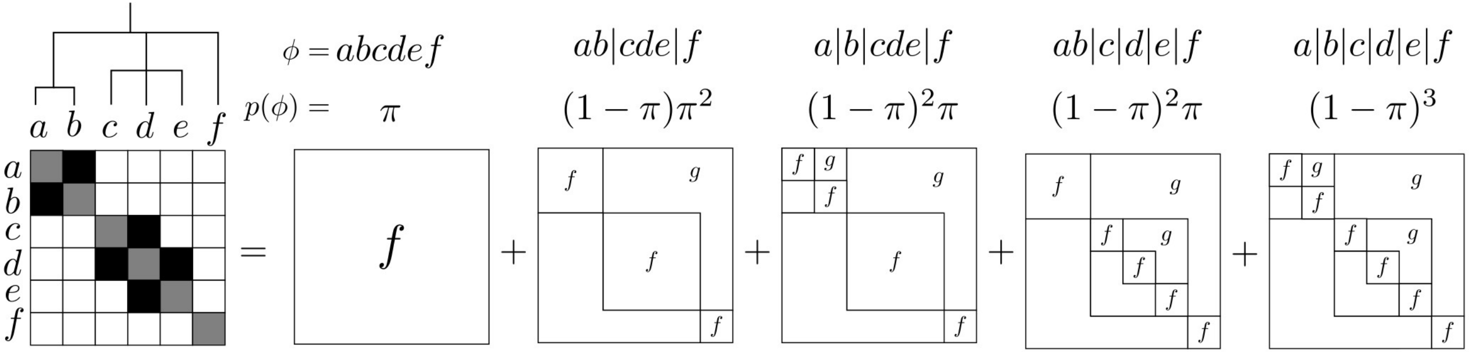
Inferring hierarchical structure



[Peixoto 2014]



[Clauset, Moore & Newman 2008]



[Blundell & Teh 2013]

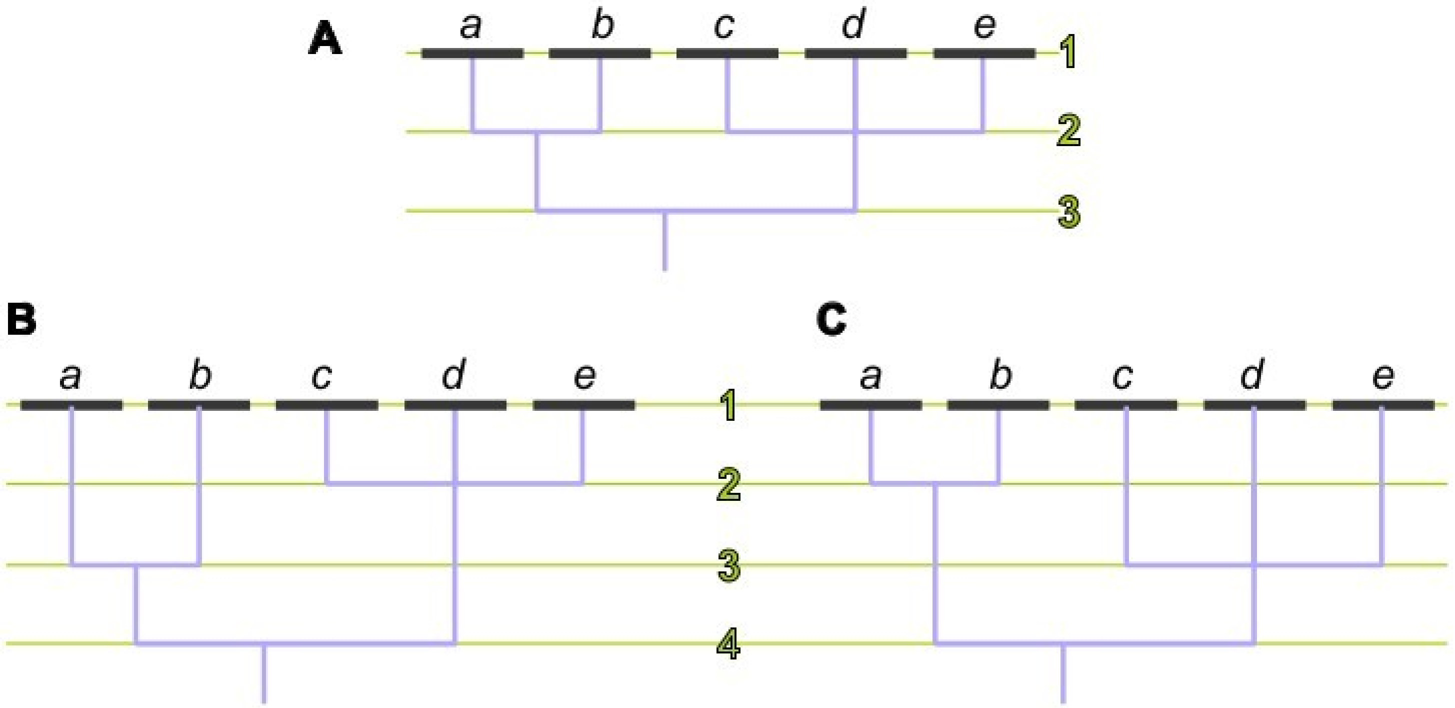
There's no free lunch!

(especially true for virtual conferences!)

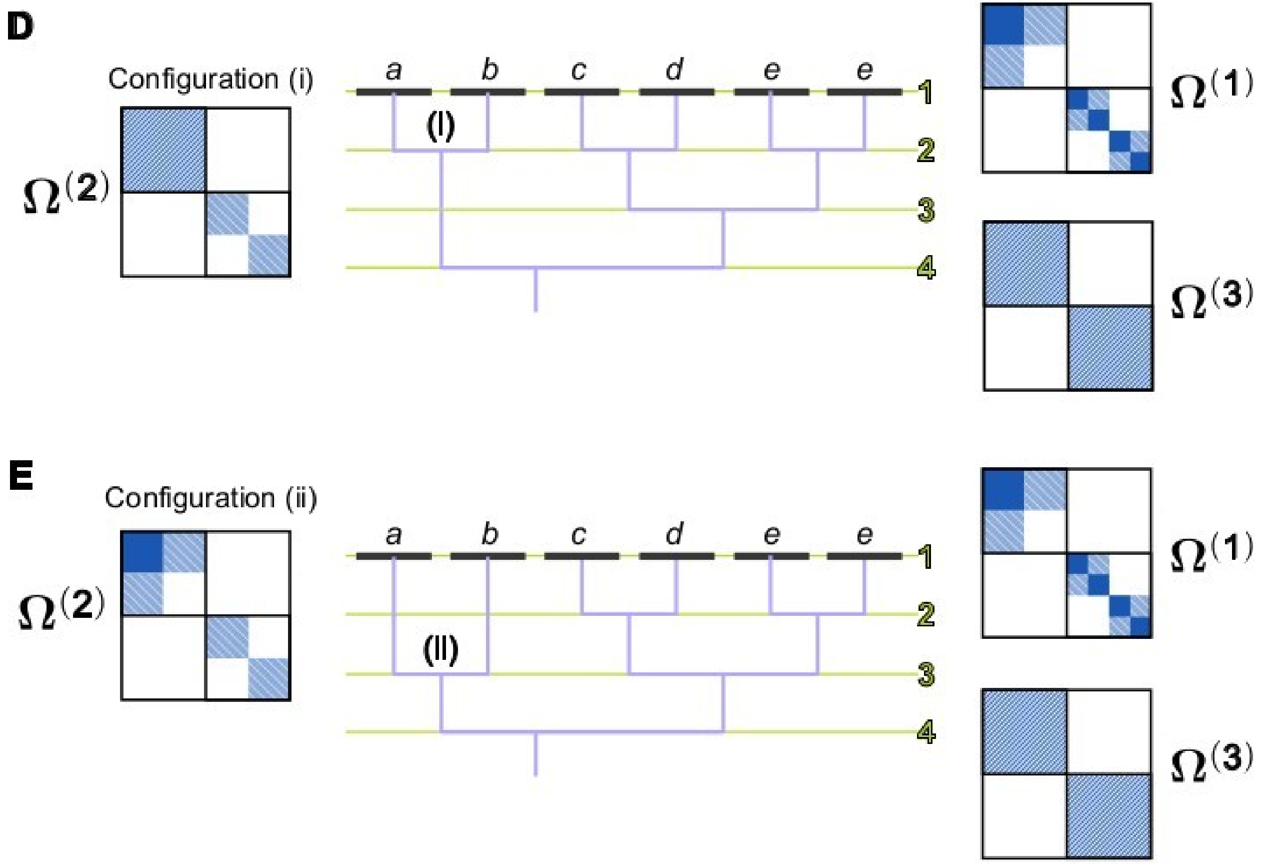
When are hierarchies identifiable?

Assigning branches to levels

Assigning branches to levels

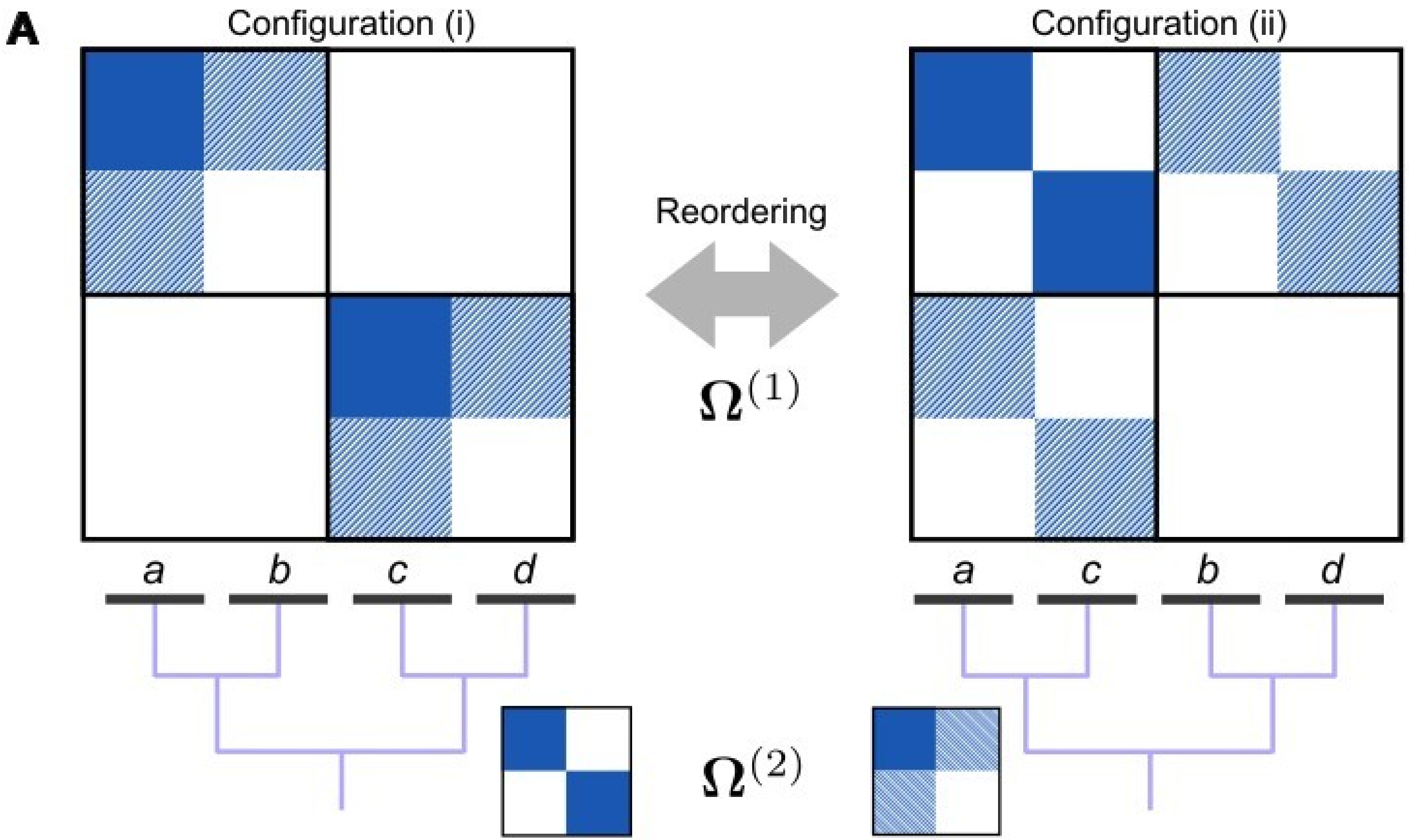


Assigning branches to levels

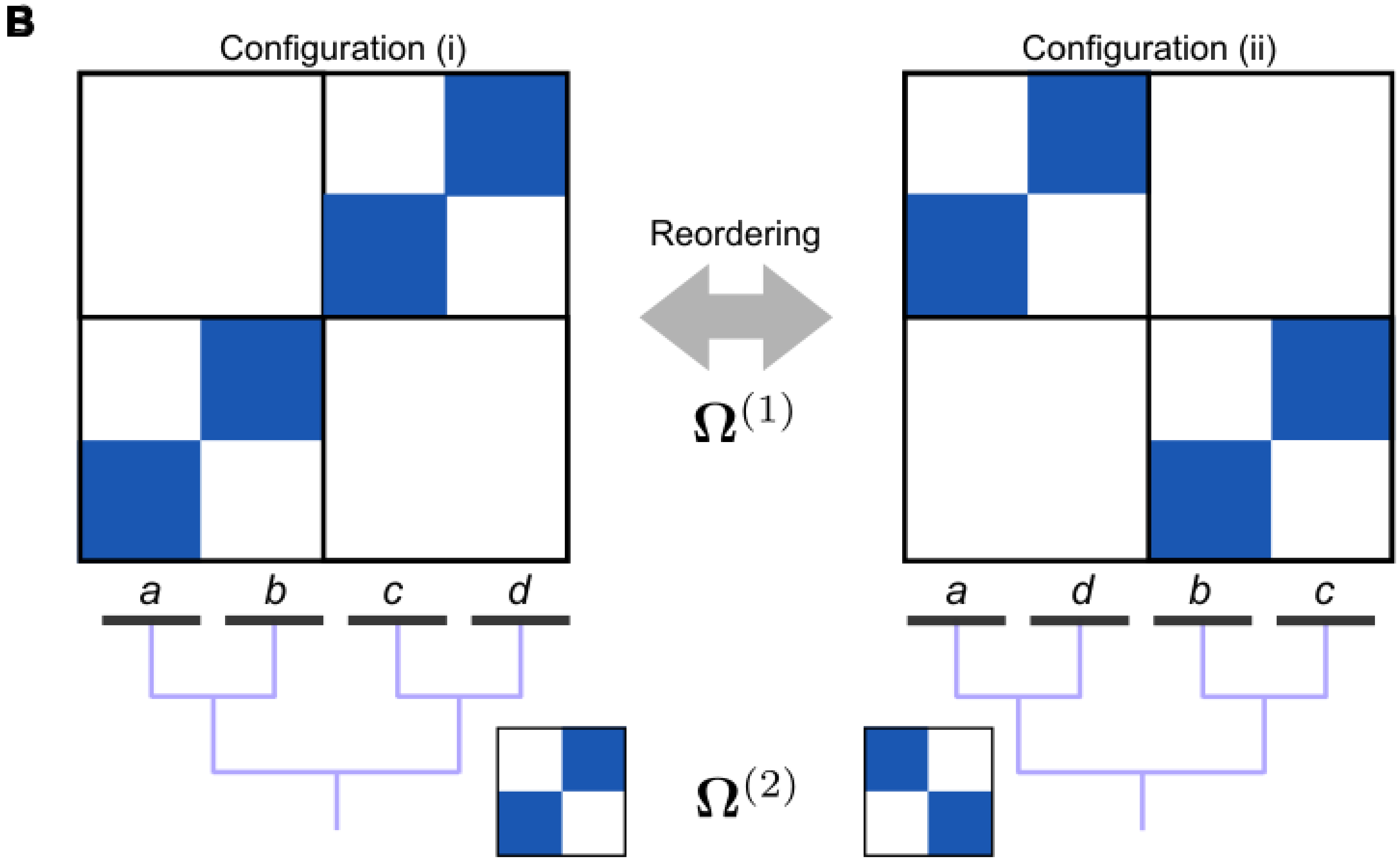


Order of agglomeration

Order of agglomeration

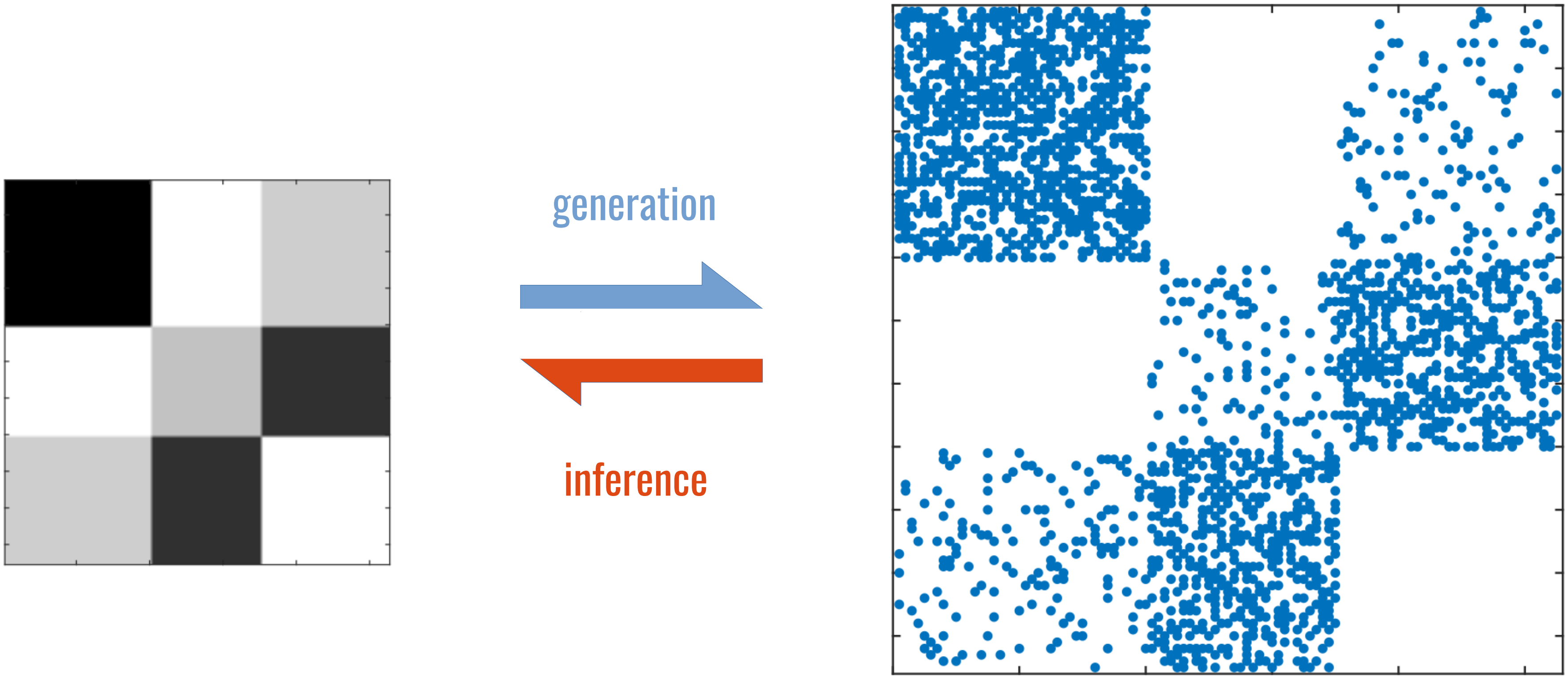


Order of agglomeration



Number of levels

The Stochastic Blockmodel (SBM)



Mixing Matrix

Adjacency Matrix

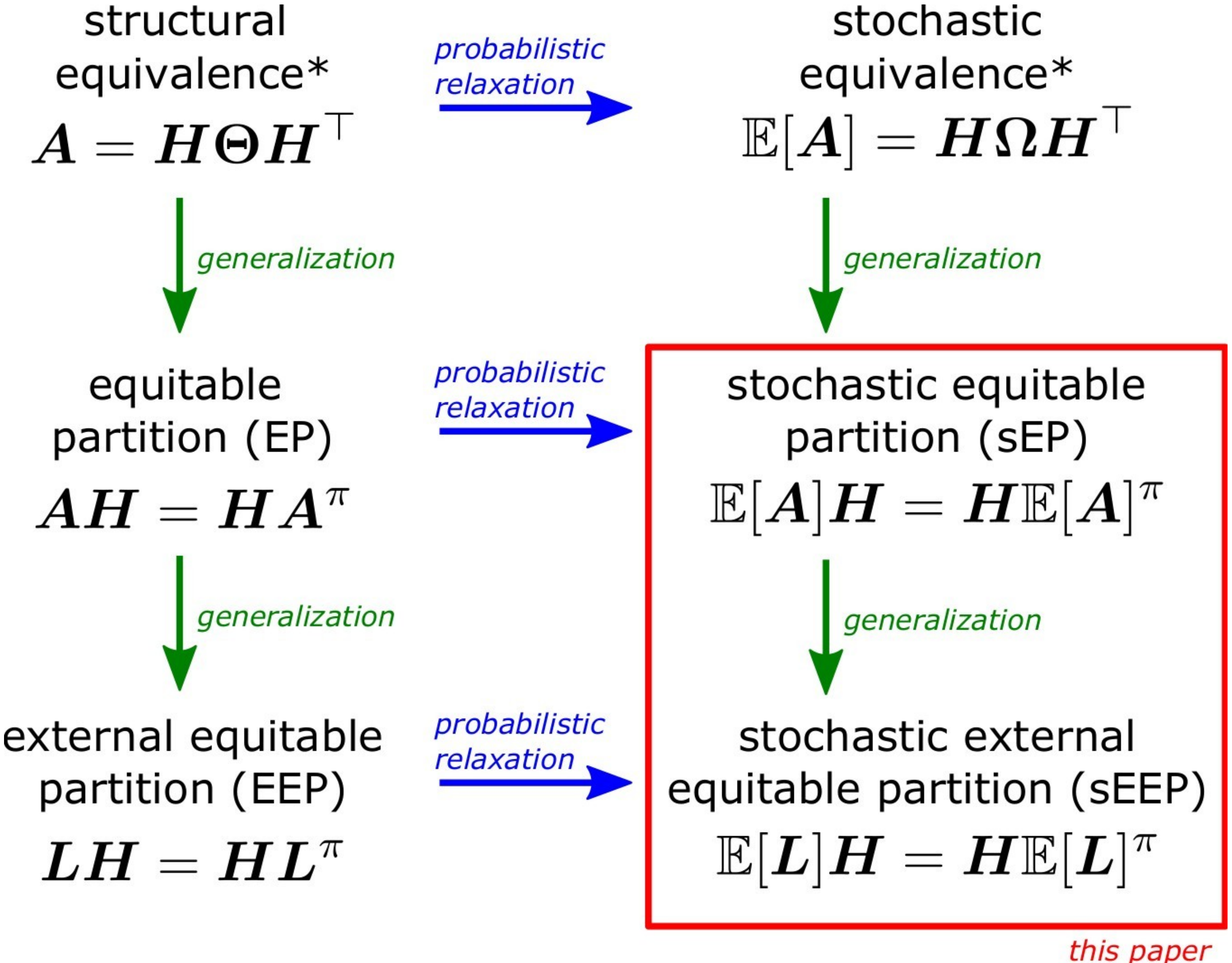
Spectral methods of inference

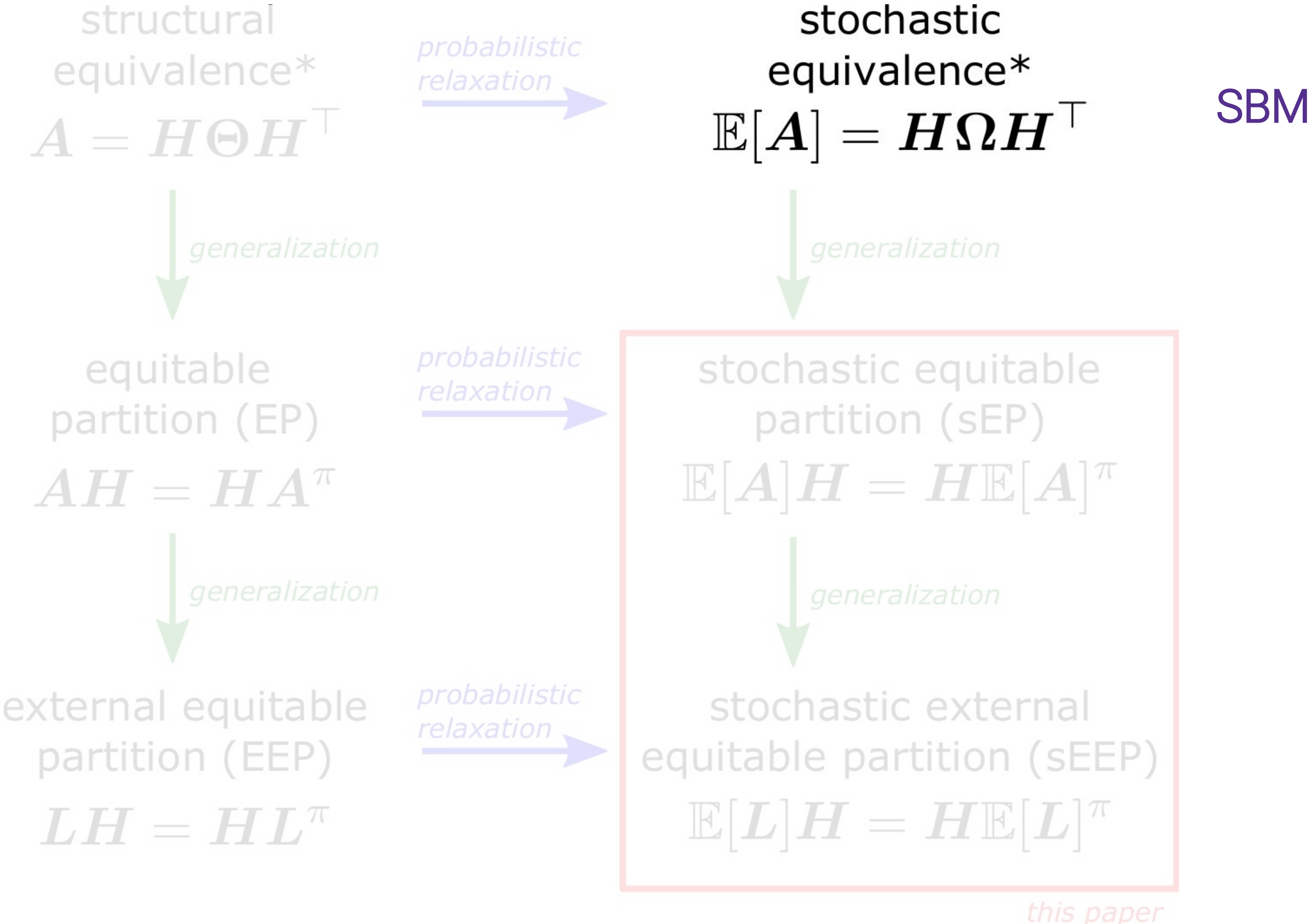
Adjacency matrix [Donath and Hoffmann 1972]
Laplacian [Fiedler 1973]
Modularity Matrix -- e.g., [Newman 2006], [Nadakuditi & Newman 2012]
Non-Backtracking matrix --- e.g., [Krzakala et al 2013]
Regularized spectral clustering --- e.g., [Rohe 2011], [Le et al 2016]
Bethe Hessian --- [Saade et al. 2014], [Le & Levina 2015]

Spectral methods of inference

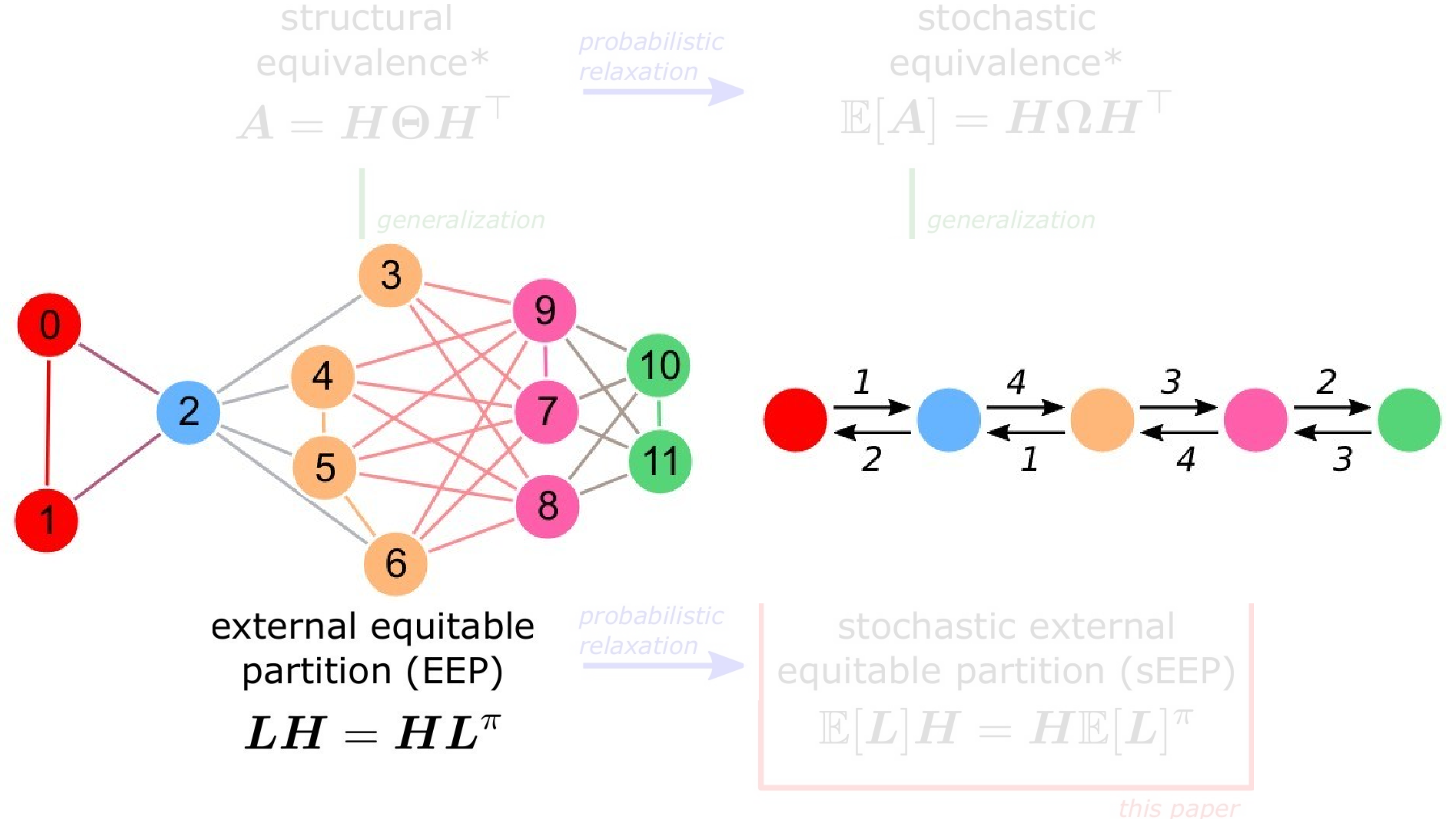
Adjacency matrix [Donath and Hoffmann 1972]
Laplacian [Fiedler 1973]
Modularity Matrix -- e.g., [Newman 2006], [Nadakuditi & Newman 2012]
Non-Backtracking matrix --- e.g., [Krzakala et al 2013]
Regularized spectral clustering --- e.g., [Rohe 2011], [Le et al 2016]
Bethe Hessian --- [Saade et al. 2014], [Le & Levina 2015]

SBM


structural equivalence*
$A = H\Theta H^\top$
probabilistic relaxation
stochastic equivalence*
$\mathbb{E}[A] = H\Omega H^\top$
generalization
generalization
equitable partition (EP)
$AH = HA^\pi$
probabilistic relaxation
stochastic equitable partition (sEP)
$\mathbb{E}[A]H = H\mathbb{E}[A]^\pi$
generalization
generalization
external equitable partition (EEP)
$LH = HL^\pi$
probabilistic relaxation
stochastic external equitable partition (sEEP)
$\mathbb{E}[L]H = H\mathbb{E}[L]^\pi$
this paper



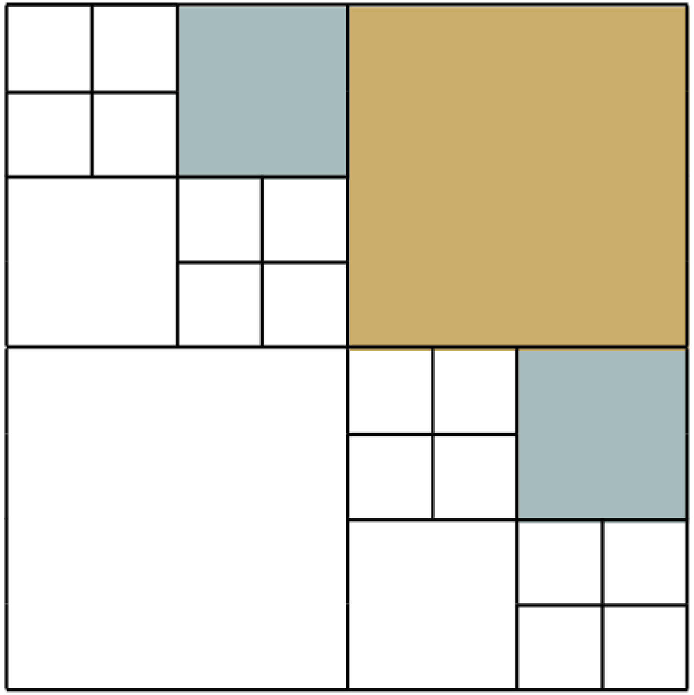
structural equivalence*
$A = H\Theta H^\top$
probabilistic relaxation
stochastic equivalence*
$\mathbb{E}[A] = H\Omega H^\top$
SBM
generalization
generalization
equitable partition (EP)
$AH = HA^\pi$
probabilistic relaxation
stochastic equitable partition (sEP)
$\mathbb{E}[A]H = H\mathbb{E}[A]^\pi$
generalization
generalization
external equitable partition (EEP)
$LH = HL^\pi$
probabilistic relaxation
stochastic external equitable partition (sEEP)
$\mathbb{E}[L]H = H\mathbb{E}[L]^\pi$
this paper



structural equivalence*
$A = H\Theta H^\top$
probabilistic relaxation
stochastic equivalence*
$\mathbb{E}[A] = H\Omega H^\top$
generalization
generalization
0
1
2
3
4
5
6
7
8
9
10
11
1
2
4
1
3
4
2
3
external equitable partition (EEP)
$LH = HL^\pi$
probabilistic relaxation
stochastic external equitable partition (sEEP)
$\mathbb{E}[L]H = H\mathbb{E}[L]^\pi$
this paper

Generate a simple hierarchy...

$$\mathbb{E}[A] =$$

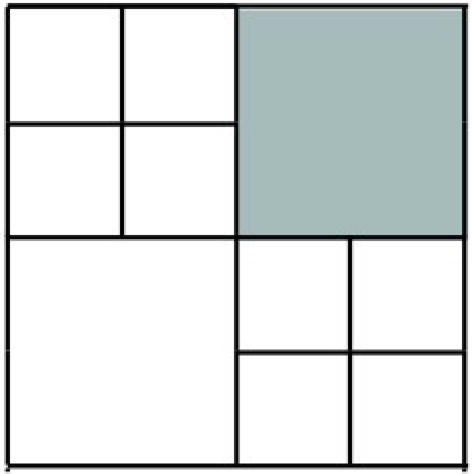


.

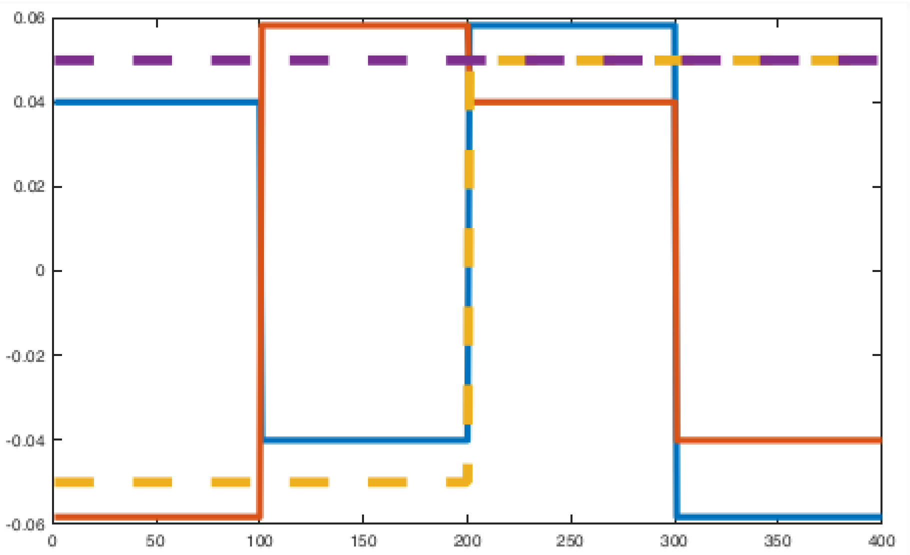
Constant probability between groups at each hierarchical level

Spectral properties

$$\mathbb{E}[A]$$

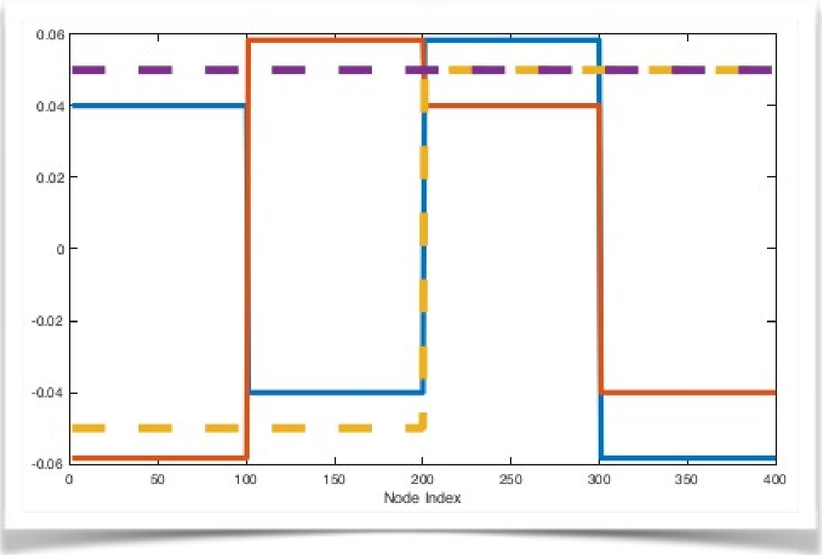


First 4 Eigenvectors of the Laplacian

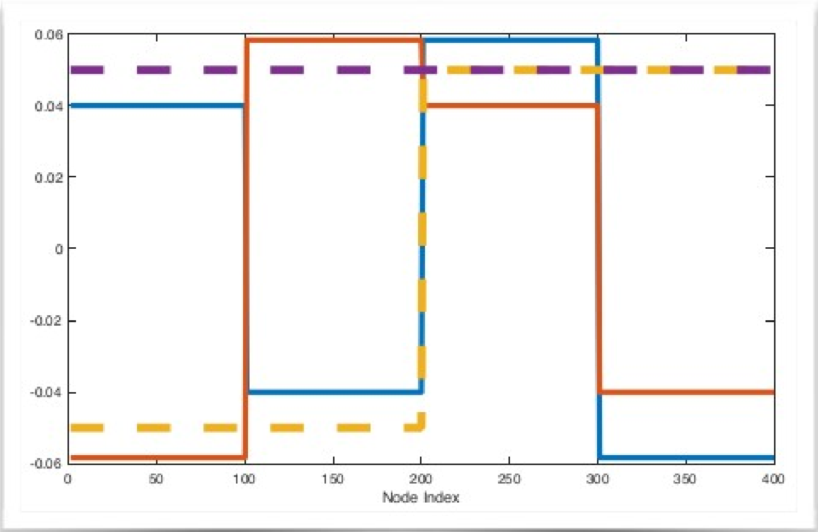


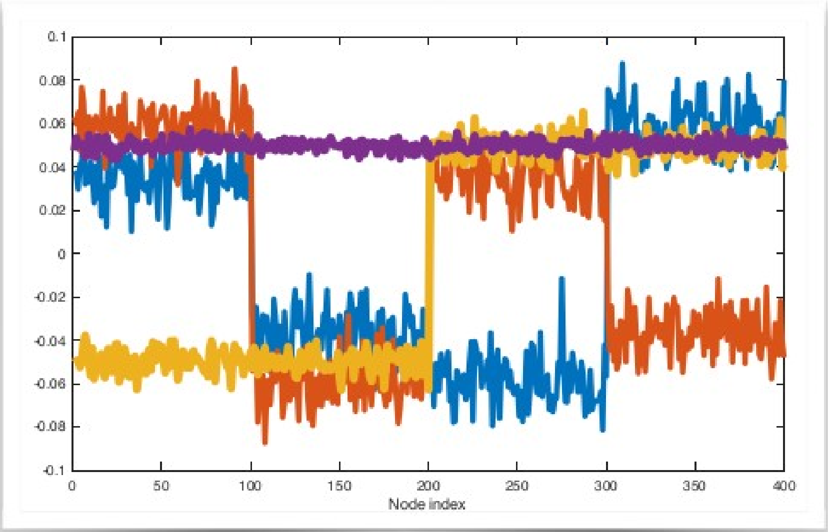
Node index

If we could just “see” the expected adjacency matrix, then we could just look for constant eigenvectors

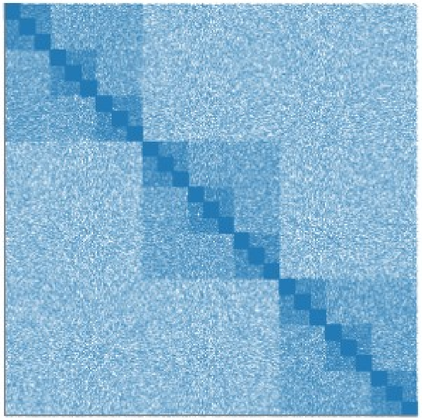


If we could just “see” the expected adjacency matrix, then we could just look for constant eigenvectors



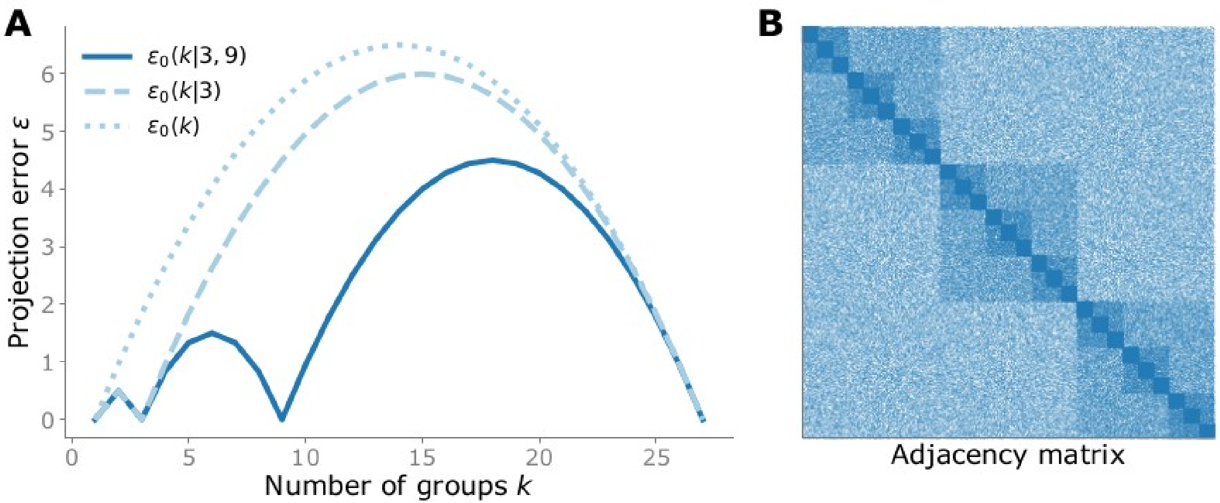


Expected projection error

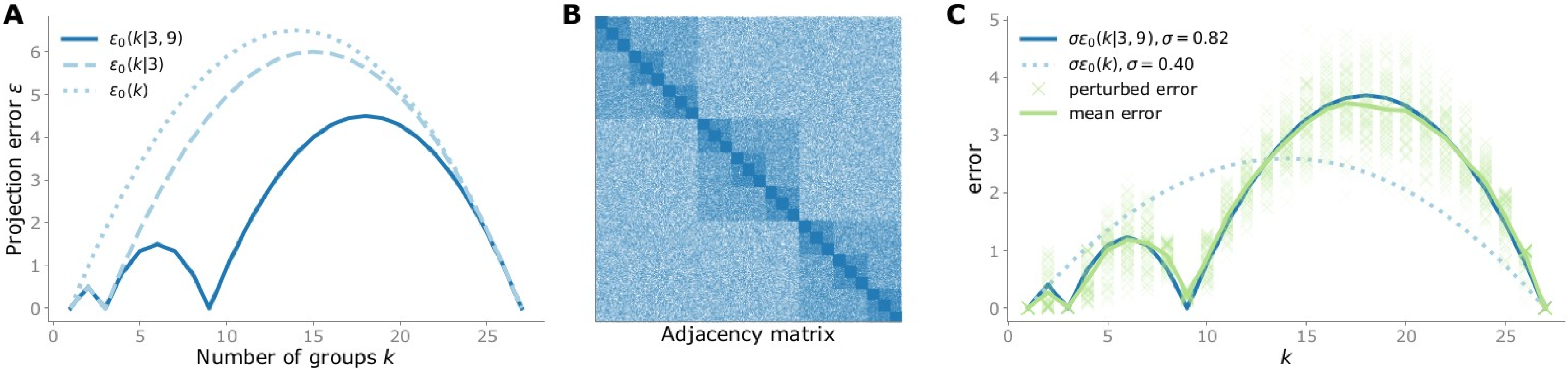


Adjacency matrix

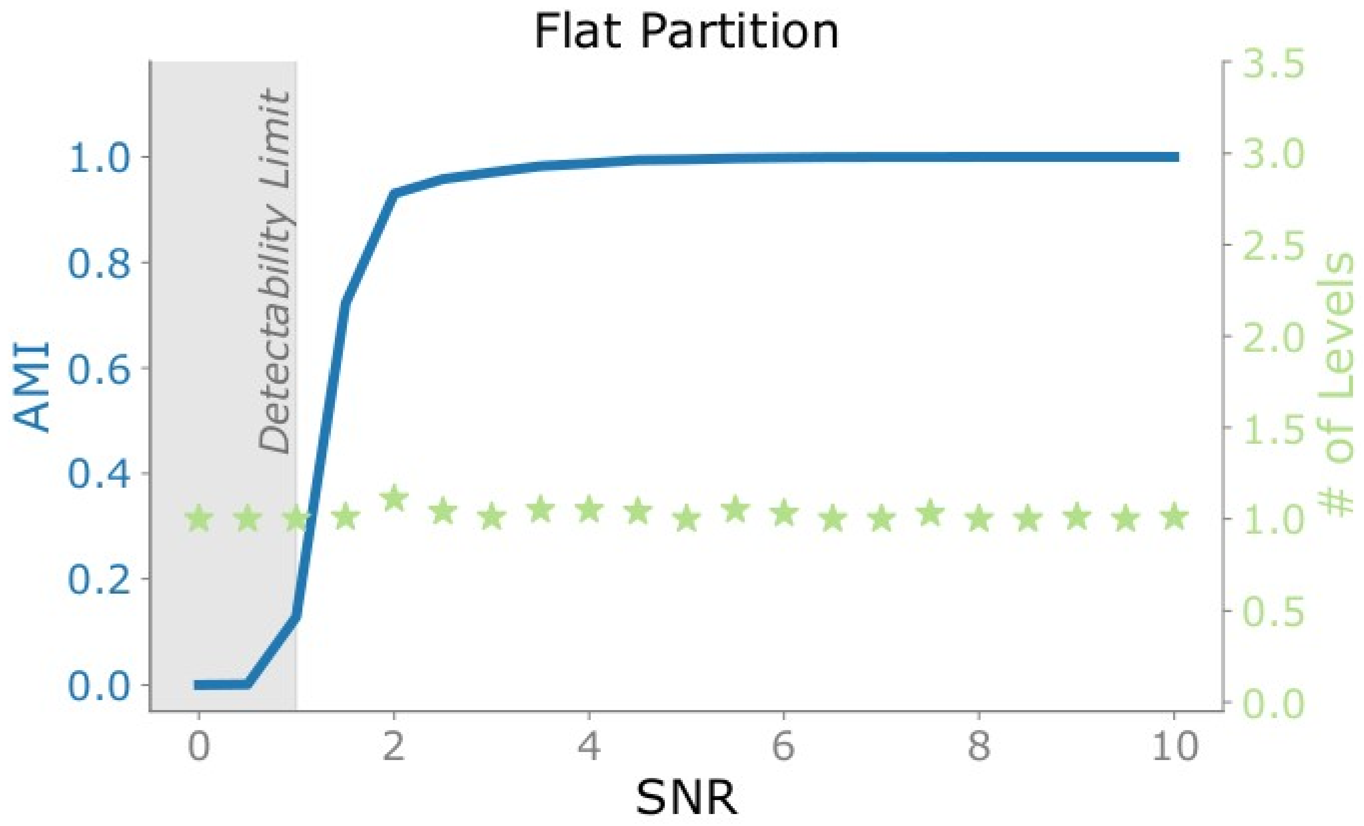
Expected projection error



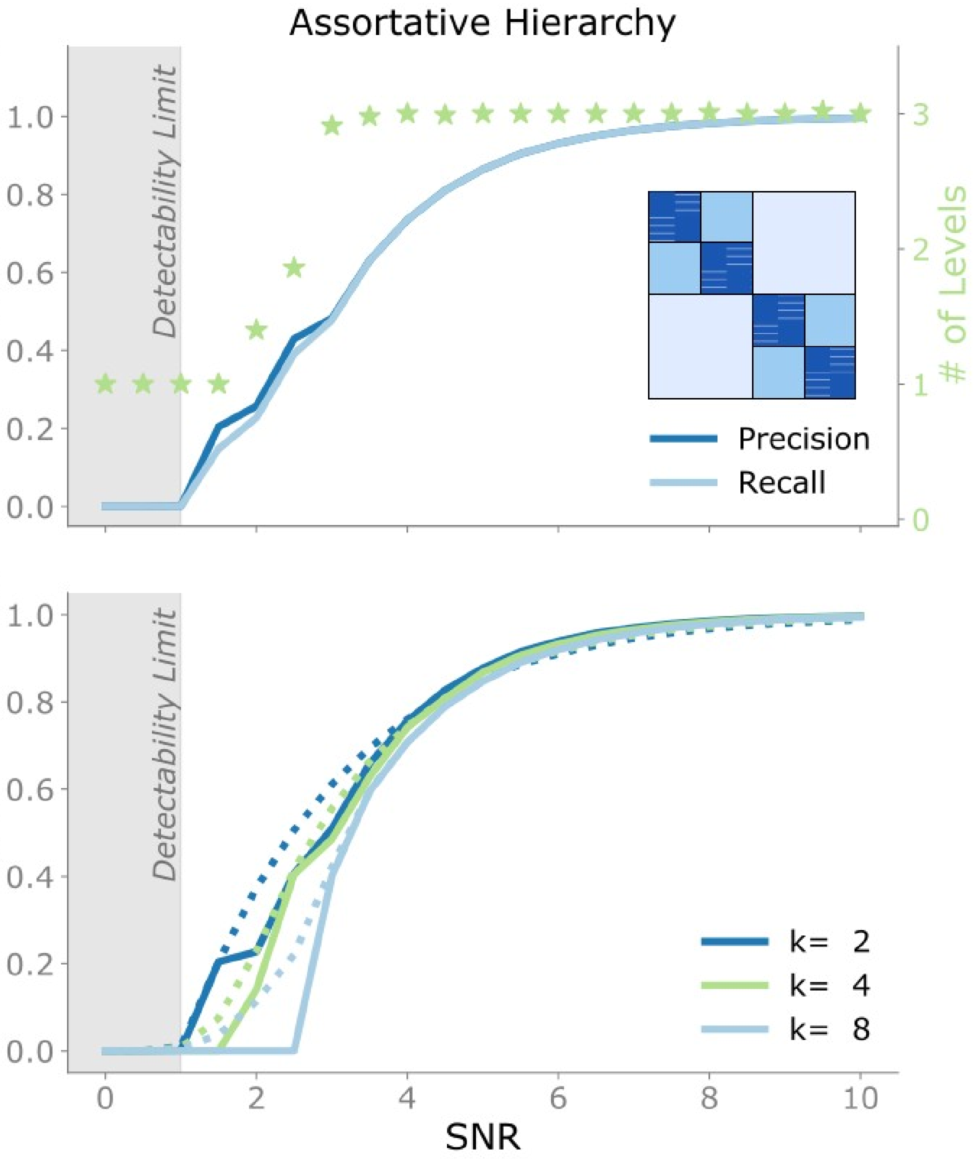
Expected projection error



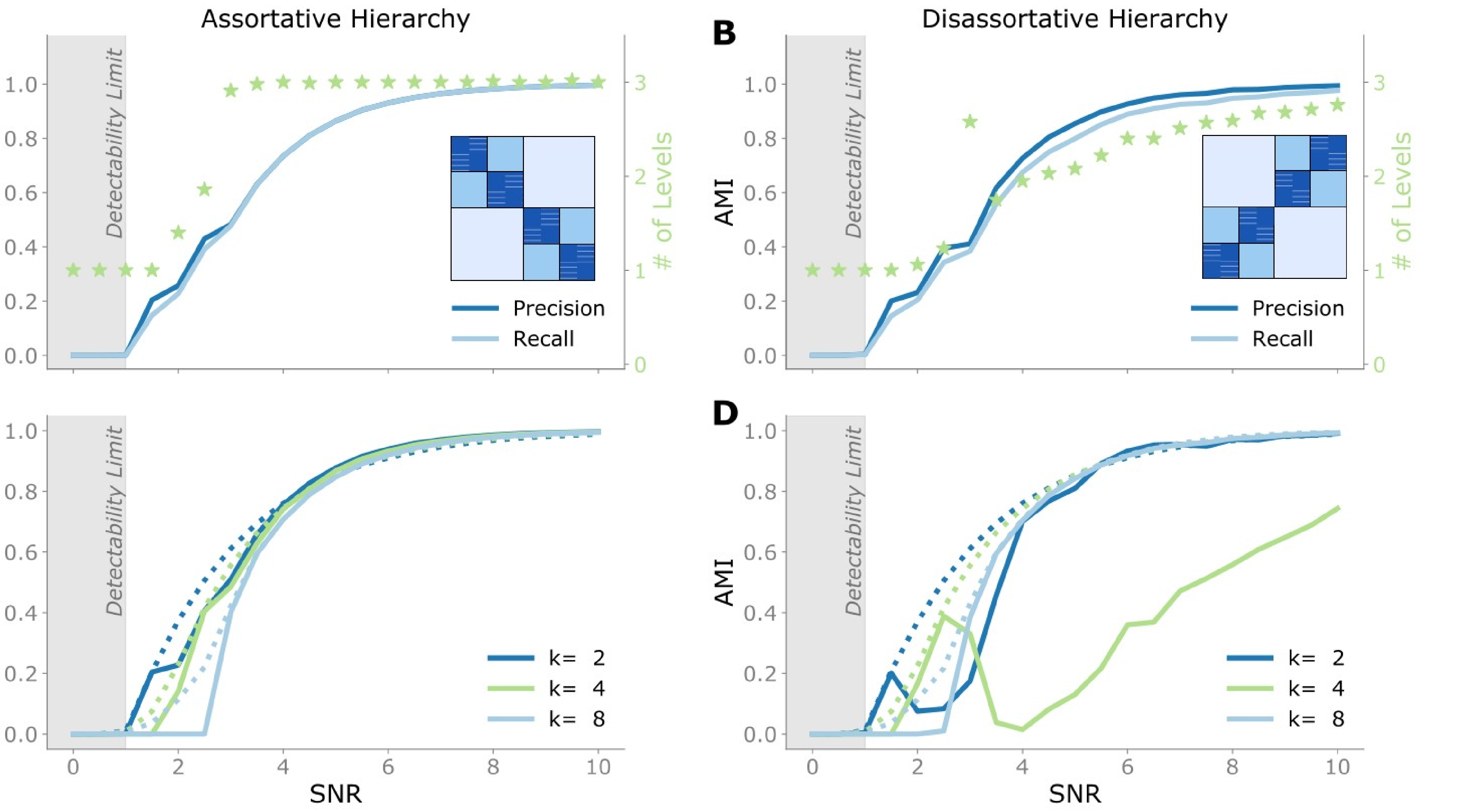
It's fast, but does it work?



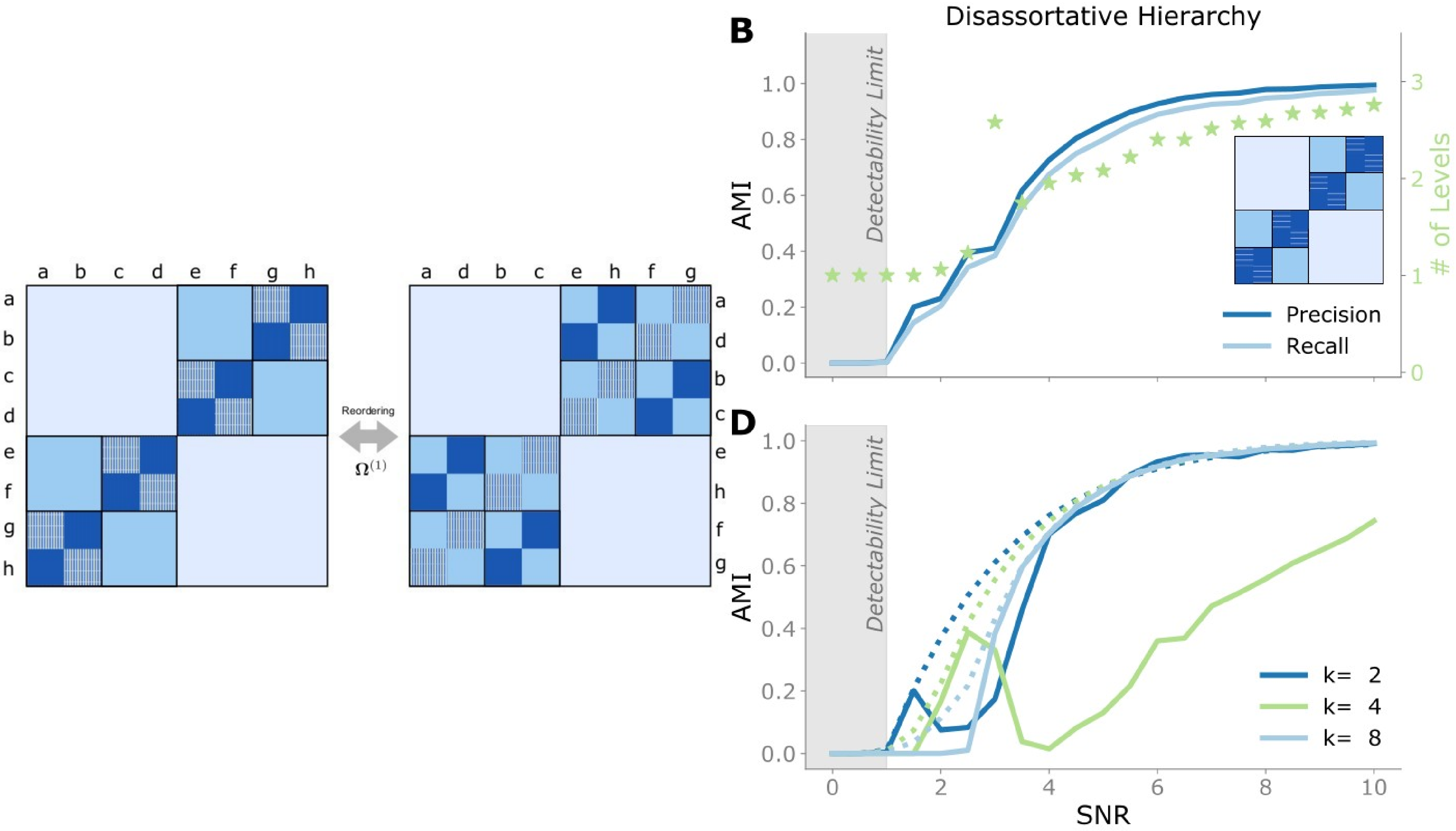
(Signal-to-noise ratio)



Assortative Hierarchy
Detectability Limit
# of Levels
Precision
Recall
Detectability Limit
k= 2
k= 4
k= 8
SNR



Assortative Hierarchy
Disassortative Hierarchy
B
D
Detectability Limit
Precision
Recall
# of Levels
AMI
k= 2
k= 4
k= 8
SNR



a b c d e f g h
a b c d e f g h
Reordering
$\Omega^{(1)}$
a d b c e h f g
a d b c e h f g
Disassortative Hierarchy
B
D
Detectability Limit
Detectability Limit
AMI
AMI
# of Levels
SNR
Precision
Recall
k= 2
k= 4
k= 8

So what method should I use?

Take home questions...

Take home questions...

- When do we have enough data to support a hierarchy?

Take home questions...

- When do we have enough data to support a hierarchy?
- How do we account for identifiability issues when measuring performance?

Take home questions...

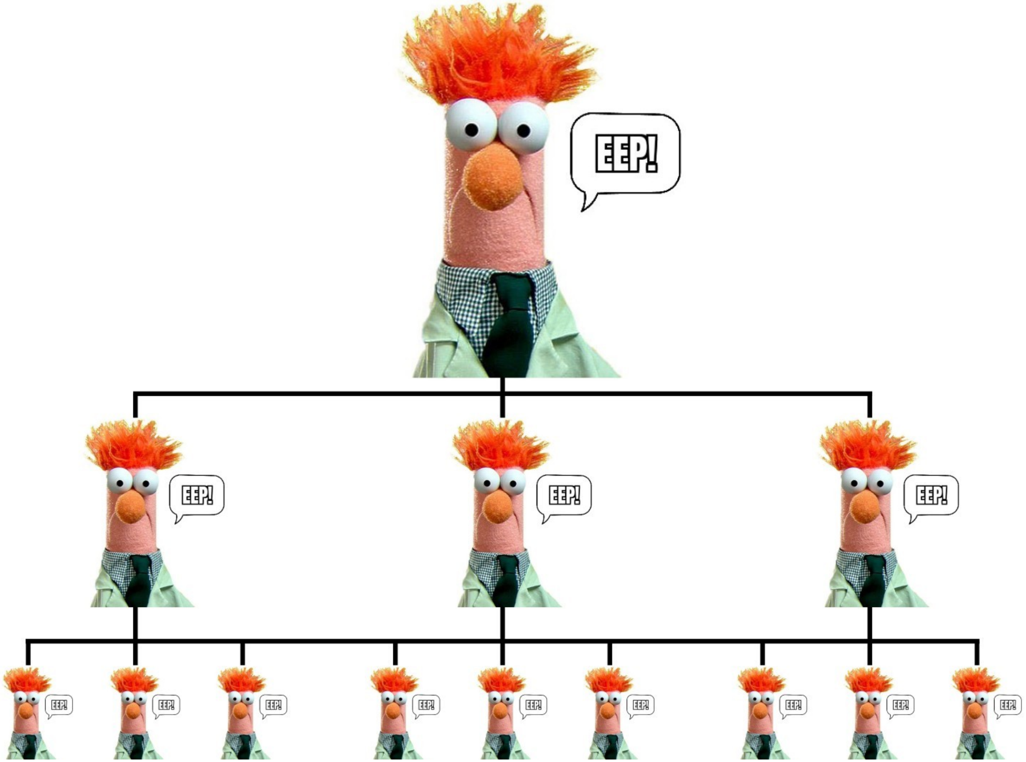
- When do we have enough data to support a hierarchy?
- How do we account for identifiability issues when measuring performance?
- Can we stop overloading terms? (community, hierarchy, higher-order, etc.)

In collaboration with

Michael Schaub
(RWTH Aachen)

Pre-prints available

arXiv:2009.07196
arXiv:2009.07525

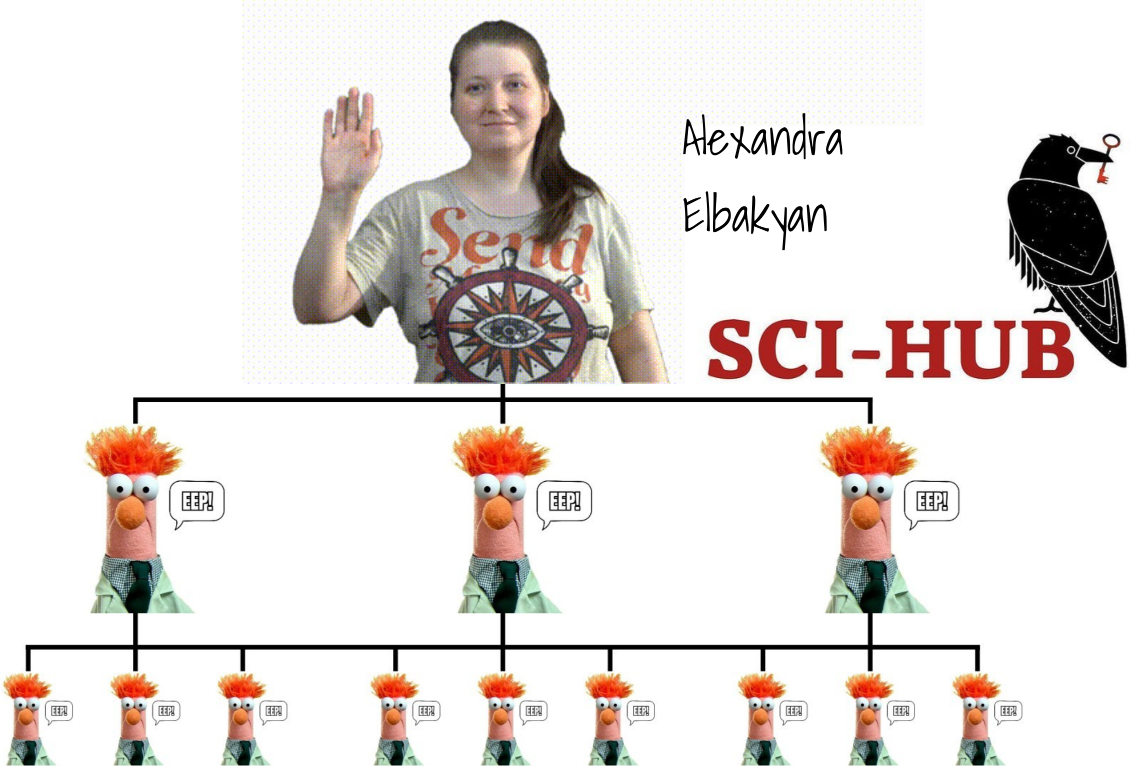


"It's EEPs all the way down"

Contact:
l.peel@maastrichtuniversity.nl
@PiratePeel

In collaboration with

Michael Schaub
(RWTH Aachen)



"It's EEPs all the way down"

Pre-prints available

arXiv:2009.07196
arXiv:2009.07525

Contact:
l.peel@maastrichtuniversity.nl
@PiratePeel