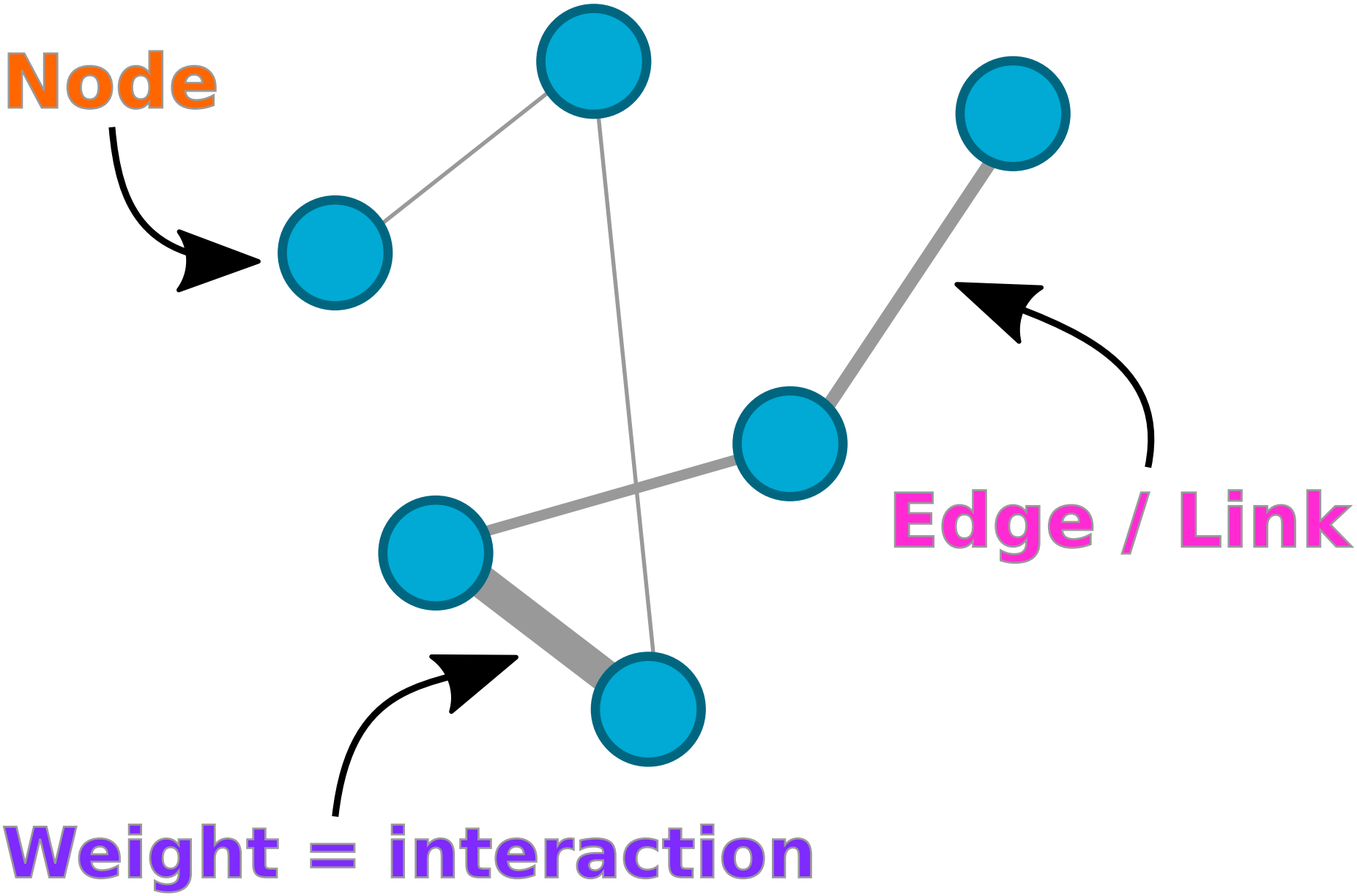
Enhanced extraction of weighted networks backbones

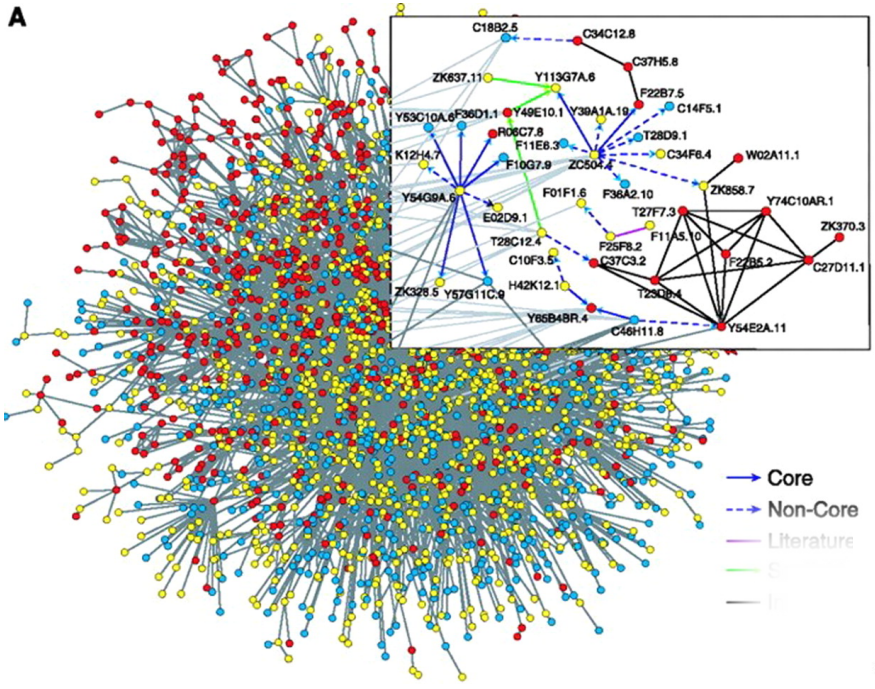
Alessio Cardillo

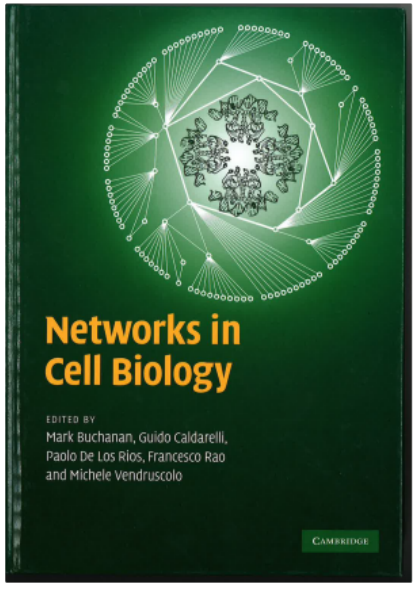
École polytechnique fédérale de Lausanne (EPFL), Lausanne, Switzerland

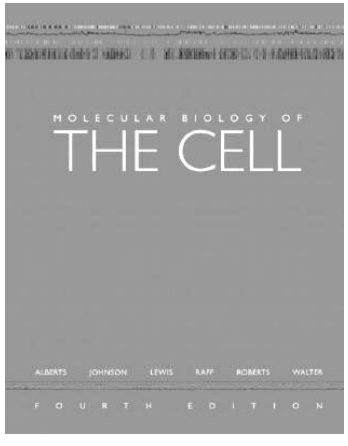
Wednesday 24 May 2017, Physics of Living Systems Meeting, EPFL, Switzerland

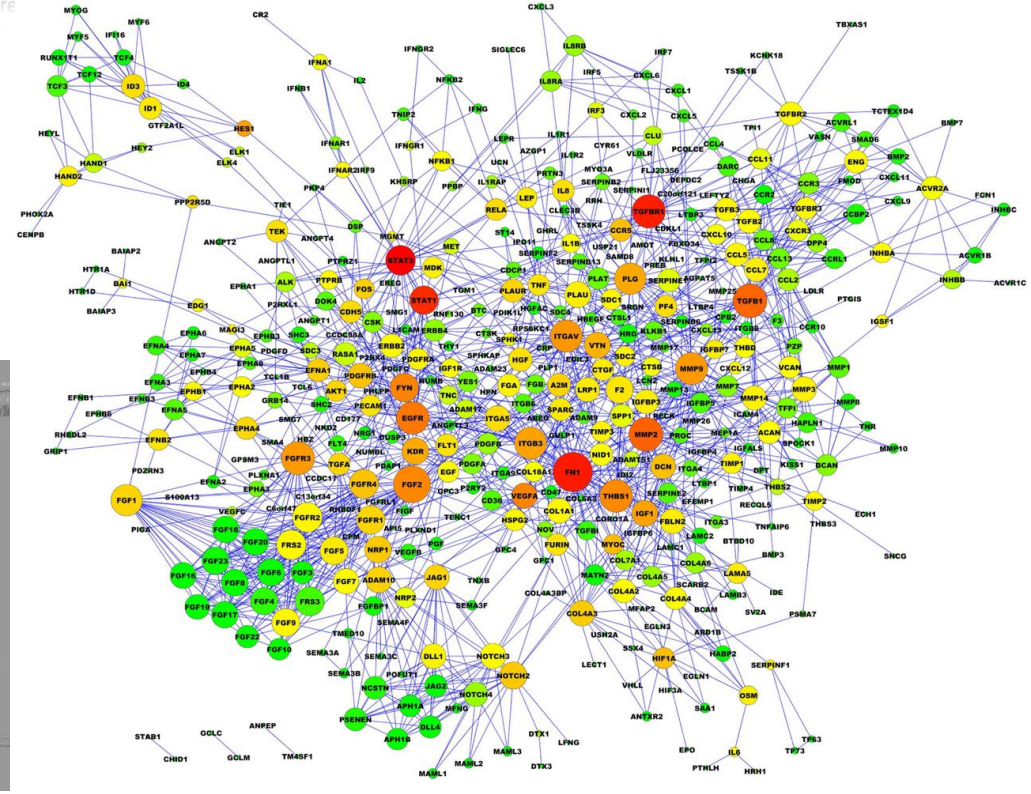

Node
Edge / Link
Weight = interaction

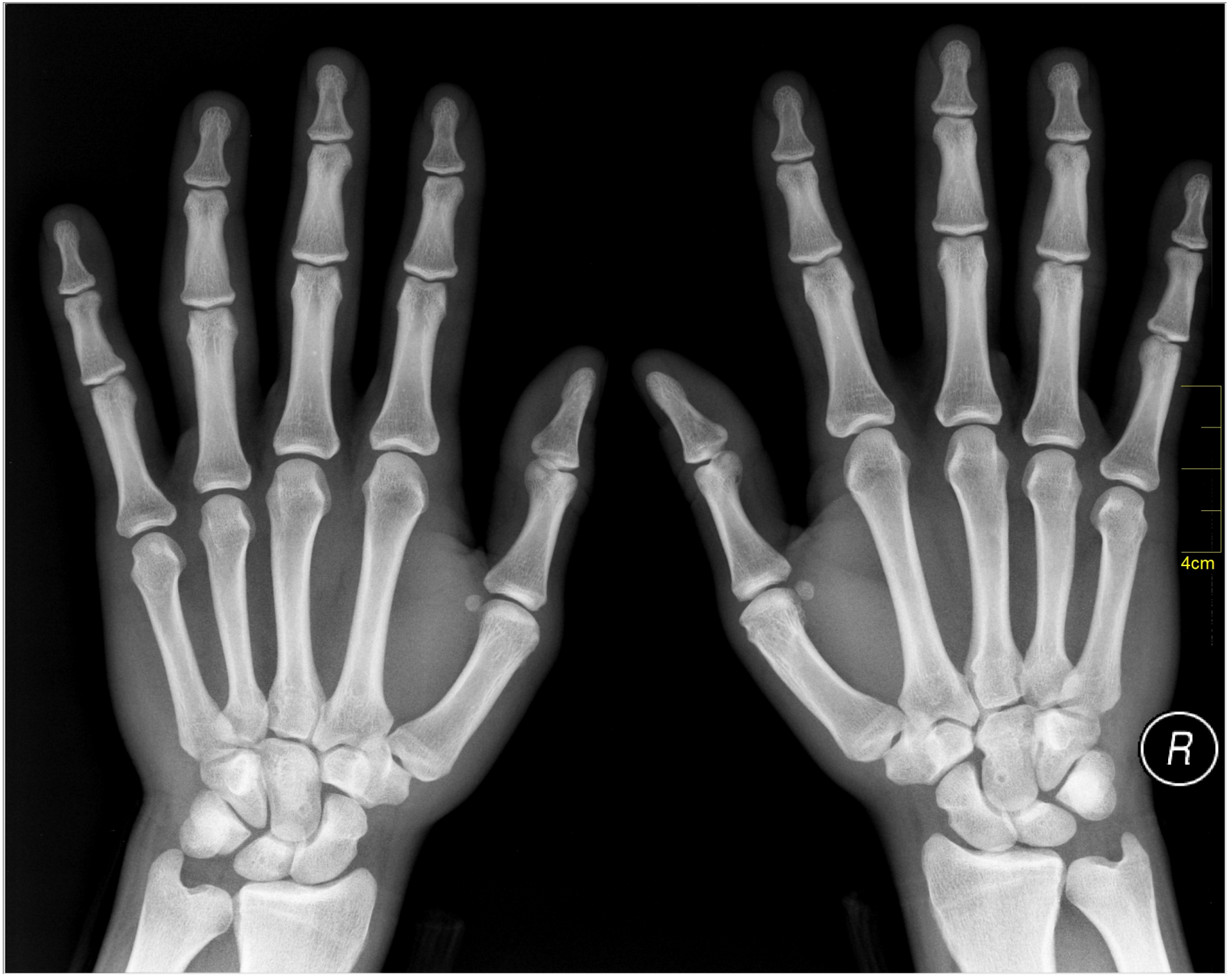
Biology and Networks









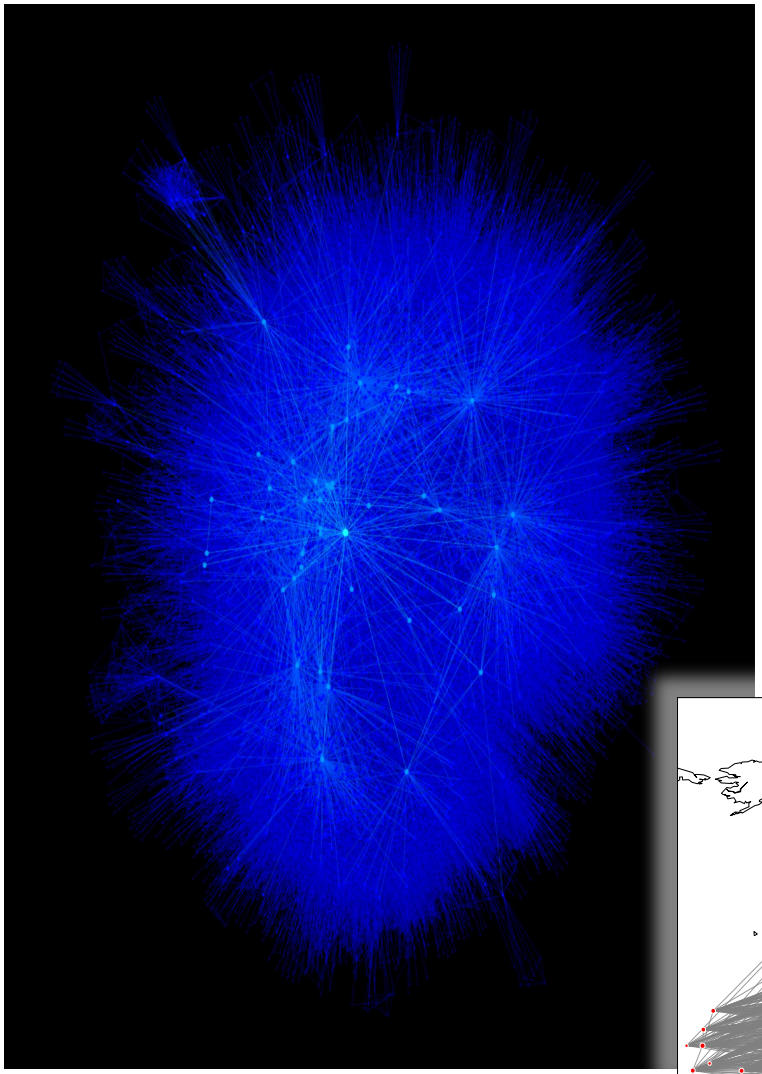

4cm
R

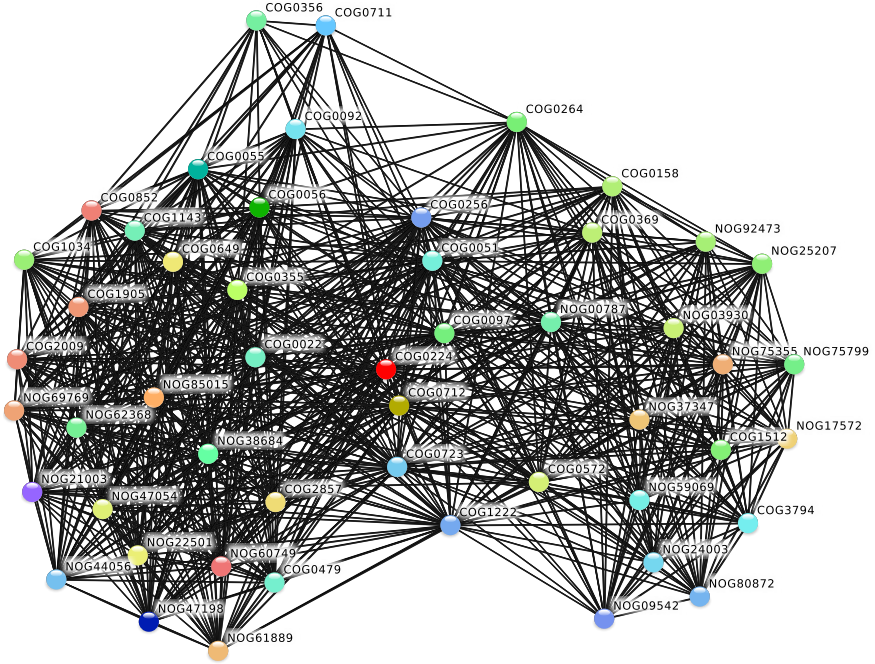
Motivation

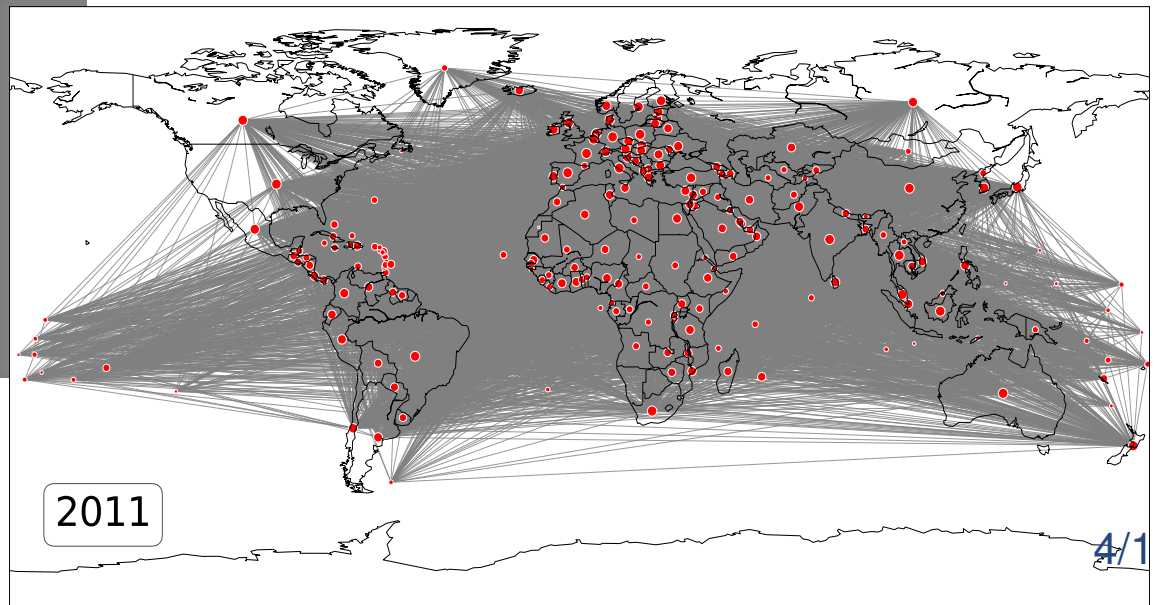

BIG
DATA
STORAGE
ANALYTICS
TECHNOLOGIES
INFORMATION
SYSTEMS
INCLUDE
SHARED
DISK
TIME
NETWORKS
SEARCH
RECORDS
COMPLEX
DATABASES
RESEARCH
SOFTWARE
USE
ONE
SINCE
USING
TYPES
WORKING
GARTNER
MASSIVELY
BIOLOGICAL
HUNDREDS
PERFORMANCE
RELATED
ALSO
CREATED
GENOMICS
ZETTABYTES
RELATIONAL
COMPUTING
TOOLS
SET
WITHIN
PROCESS
DEFINITION
RECONSIDER
OPPORTUNITIES
COMBAT
SIGNIFICANT
BIOGEOCHEMICAL
CONNECTOMICS
DESKTOP
CURRENTLY
FC
WORLD'S
TENS
CAPACITY
FORMS
PRESENTATIONS
PRACTITIONERS
NOW
ORGANIZATIONS
REQUIRING
BUSINESS
MOVING
UBIQUITOUS
PARALLEL
SAN
RADIO-FREQUENCY
SOLID
COMPLEXITY
WIRELESS
TOLERABLE
SIZE
NEEDED
QUALITIES
PETABYTES
INTERNET
PROCESSING
BURIED
LOGS
USED
DISTRIBUTED
CAPTURE
MAY
MANAGEMENT
MANAGE
GROW
SOCIAL
LARGE
STORE
DEFINING
CASE
EVERY
LARGER
COST
CONTINUES
TERABYTES
MPP
CITATION
INDEXING
DIFFICULTY
TARGET
ABILITY
SETS
EXAMPLES
APPLIED
SENSOR
ARCHIVES
AMOUNT
DESCRIBING
ELAPSED
CURRENT
THOUGHT

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Motivation







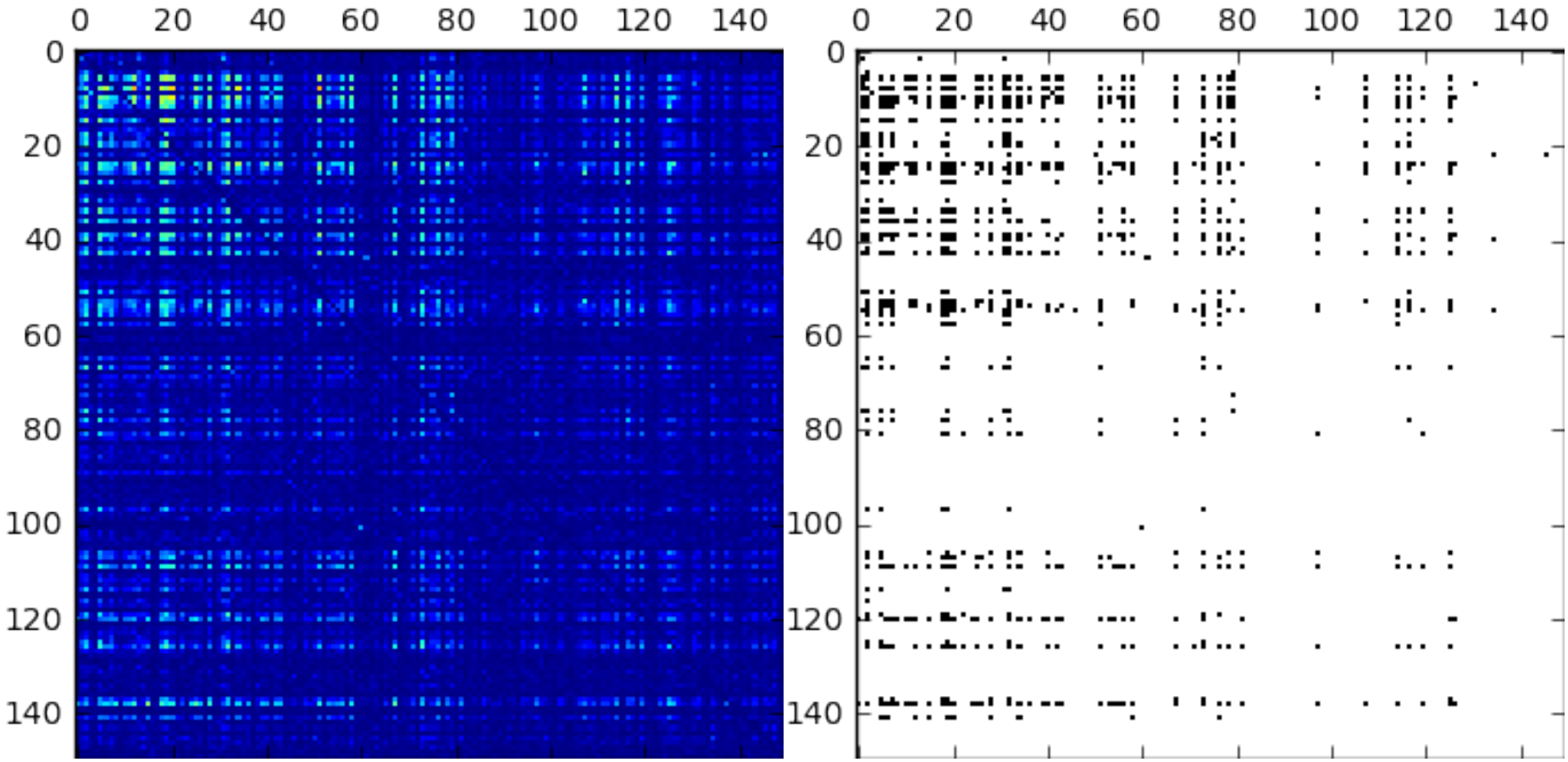
Motivation

Question:

What can we learn from a complex system whose network representation is **noisy** and/or extremely **dense**?

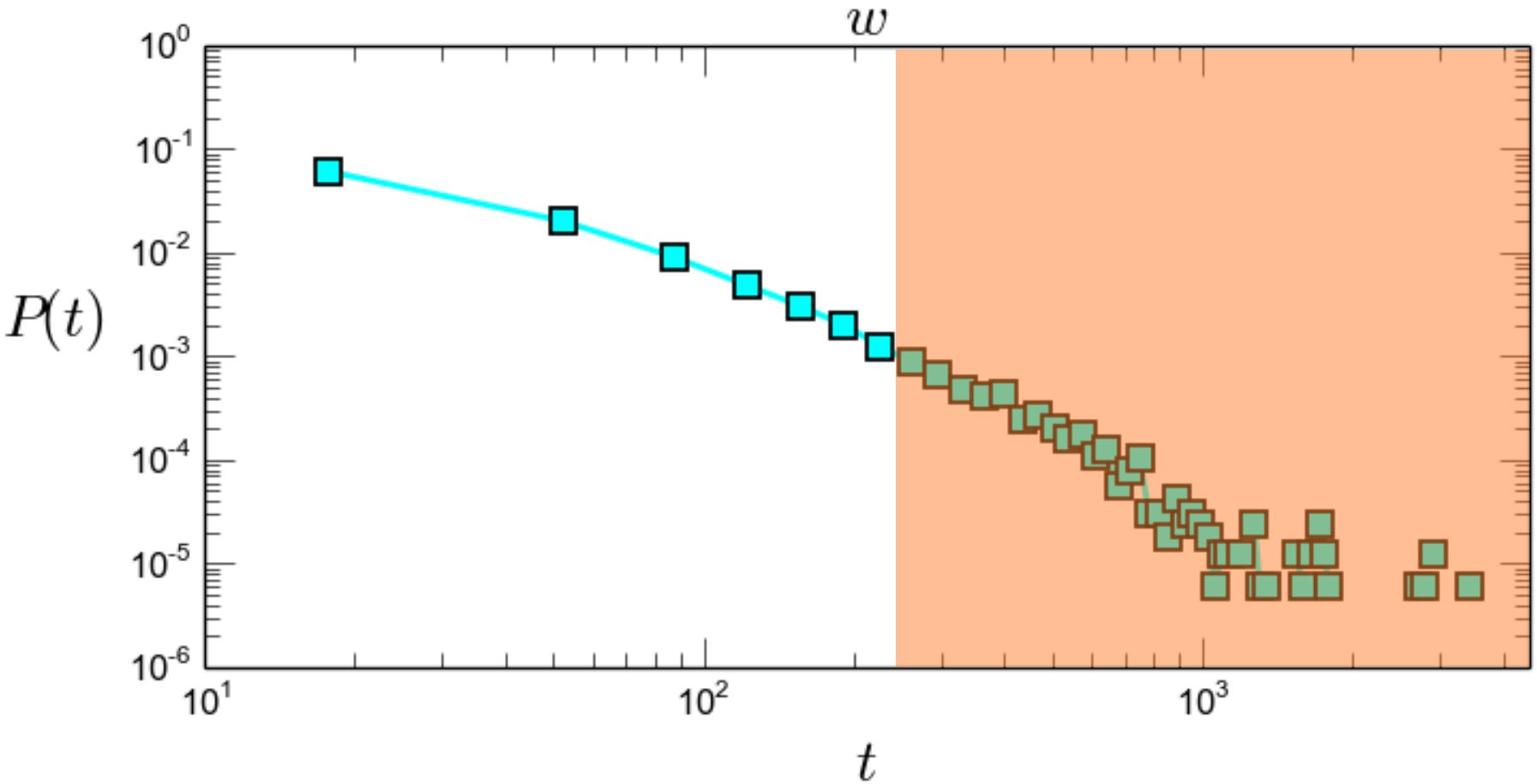

BRACE YOURSELVES
FILTERING IS COMING
imgflip.com

Thresholding



- De Vico Fallani F., *et al. A Topological Criterion for Filtering Information in Complex Brain Networks*. PLoS Comp. Bio. **13** e1005305 (2017).

Thresholding



- Granovetter, M. S. *The Strength of Weak Ties*. Am. Jour. Soc., **78**, 1360 (1973).

ECM-filter

Main Features

1. Based on the comparison between the observed network and a null model one.

ECM-filter

Main Features

1. Based on the comparison between the observed network and a null model one.
2. Null model: maximum-entropy canonical ensemble of networks satisfying given constraints.

Main Features

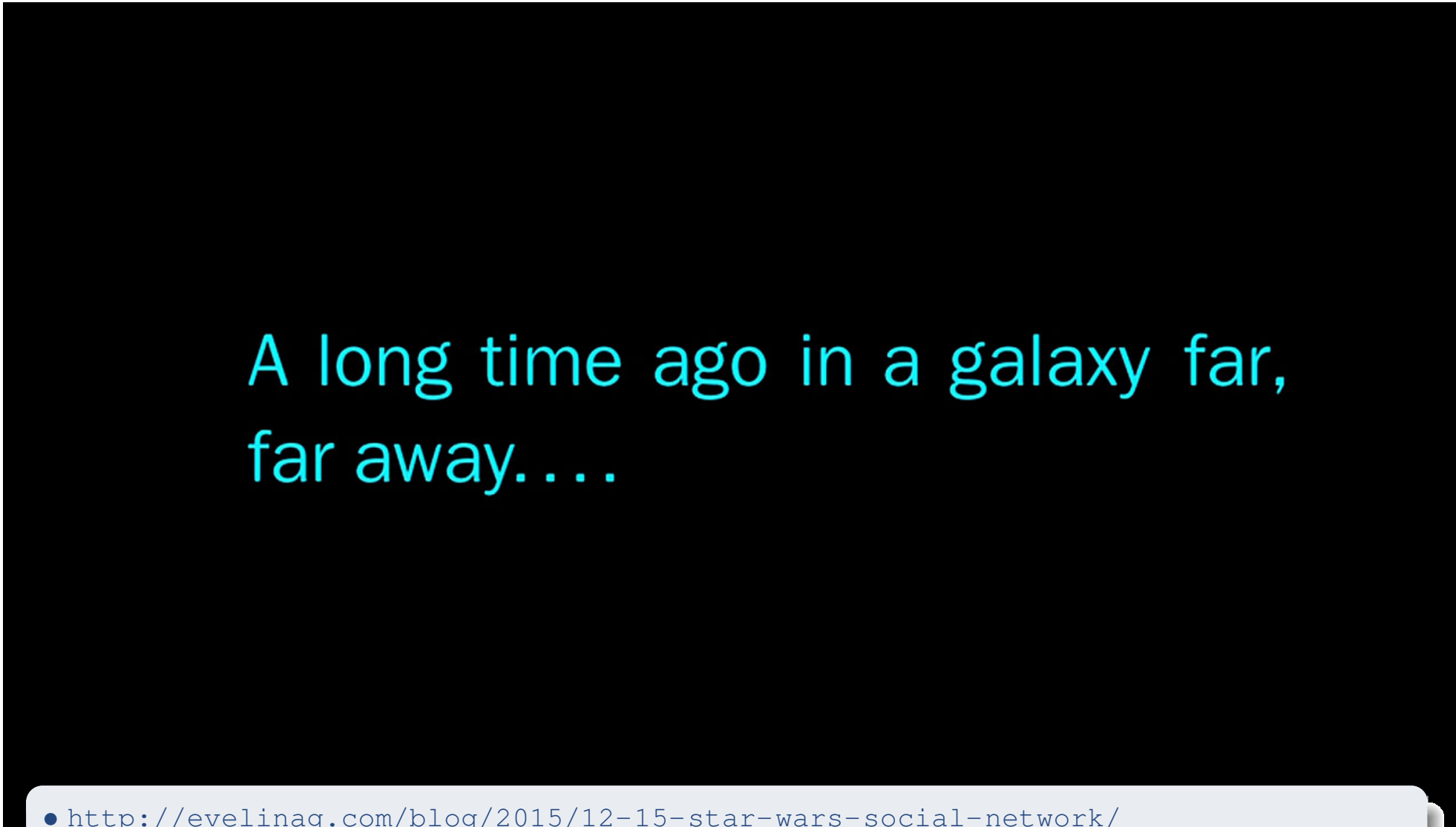
1. Based on the comparison between the observed network and a null model one.
2. Null model: maximum-entropy canonical ensemble of networks satisfying given constraints.
3. Two versions: **local** (focus on links) and **global** (focus on entire network).

• Mastrandrea, R., *et al. Enhanced reconstruction of weighted networks from strengths and degrees.* New Jour. Phys., **16**, 043022. (2014).

Results

A serious example . . .

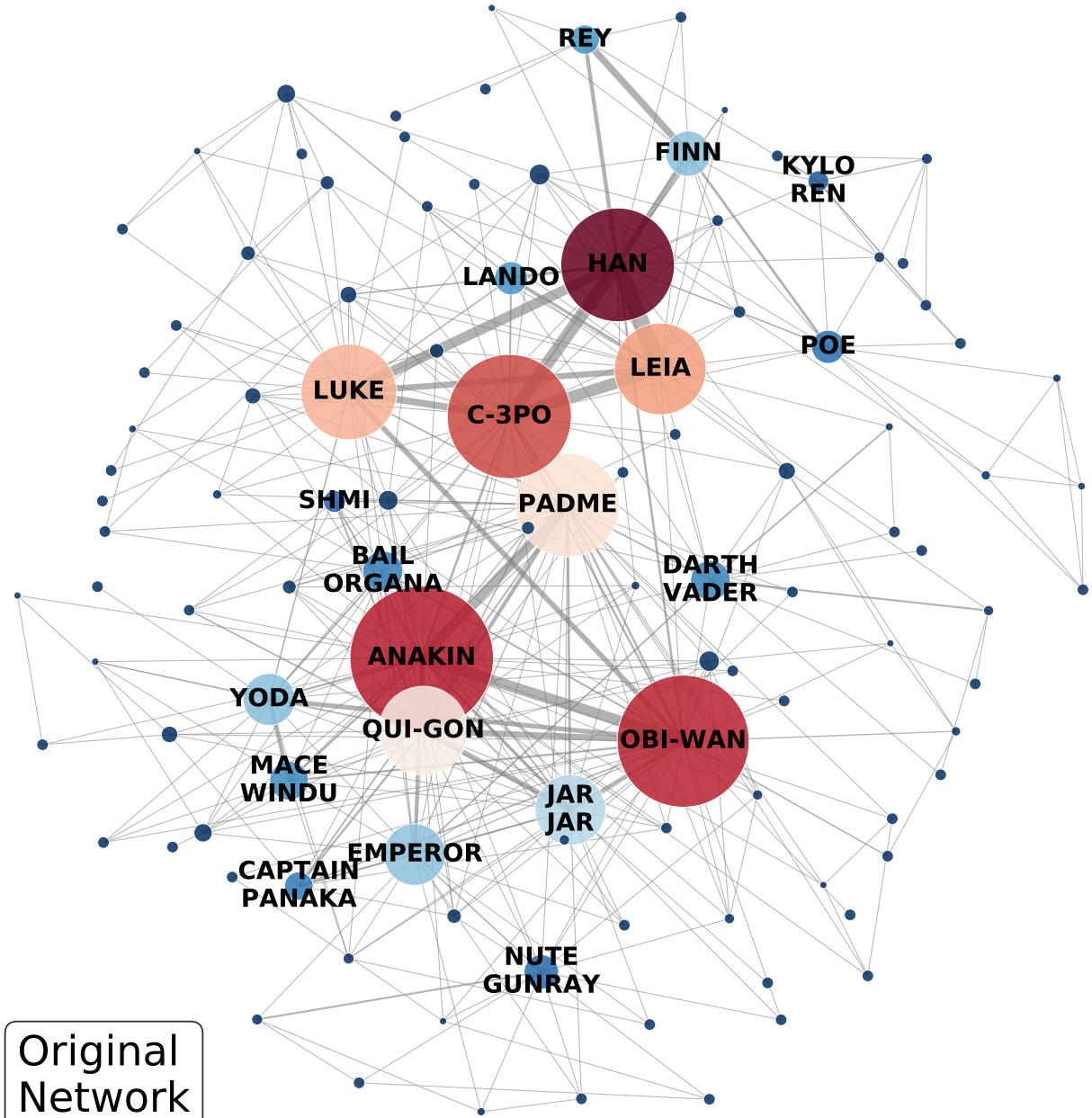
A serious example . . .



- http://evelinag.com/blog/2015/12-15-star-wars-social-network/

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

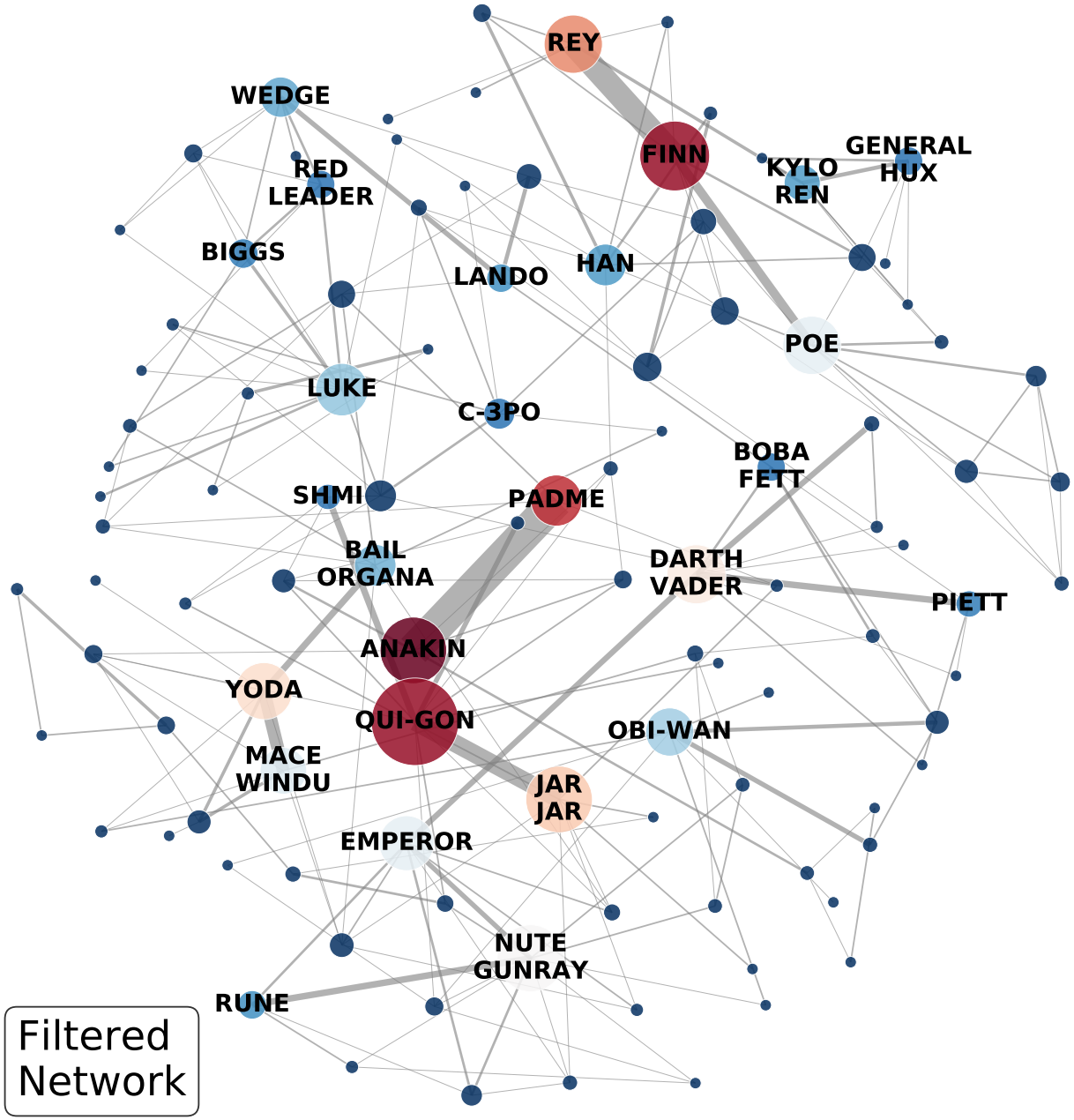
A serious example . . .



Original Network

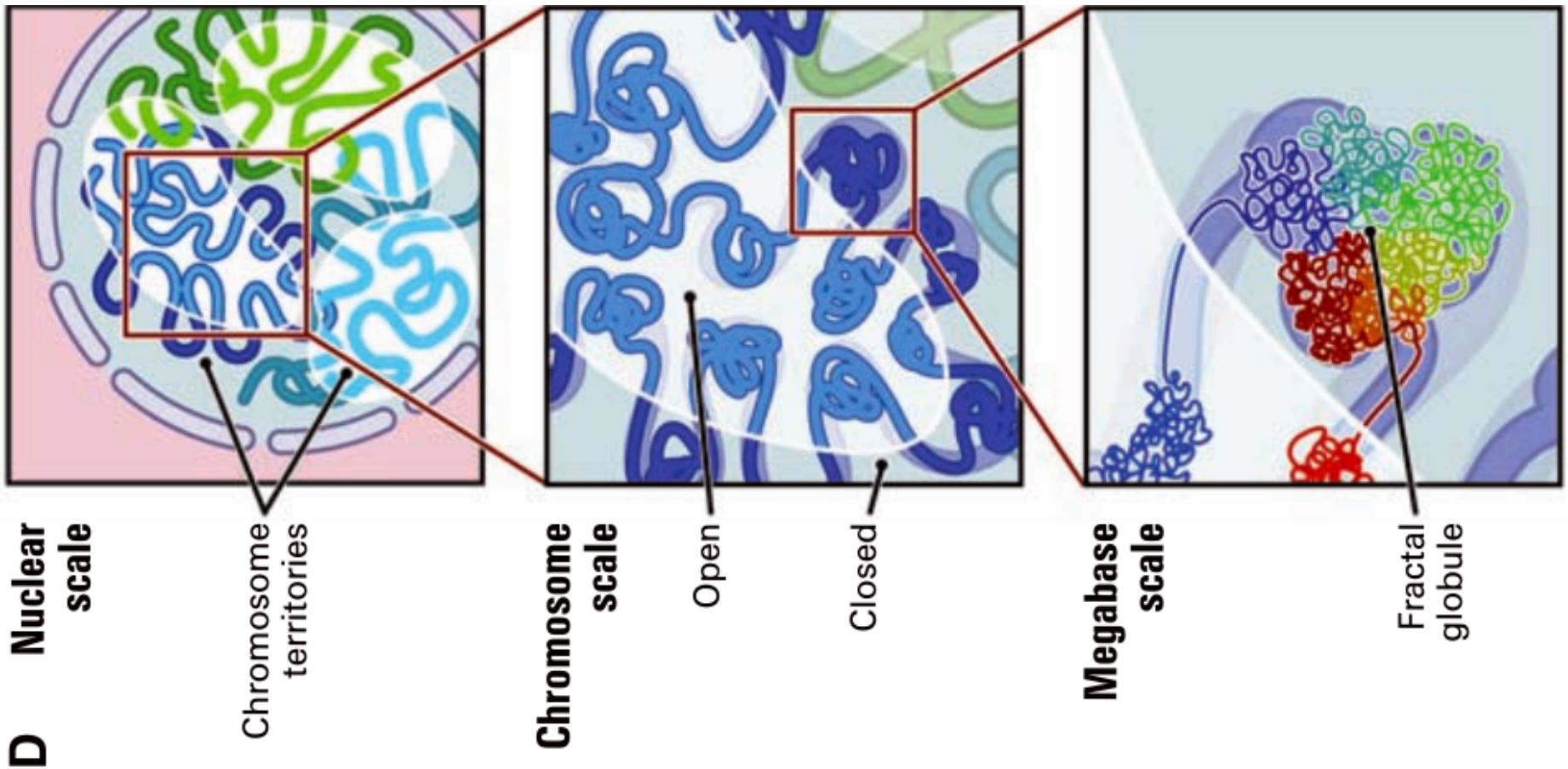
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

A serious example . . .

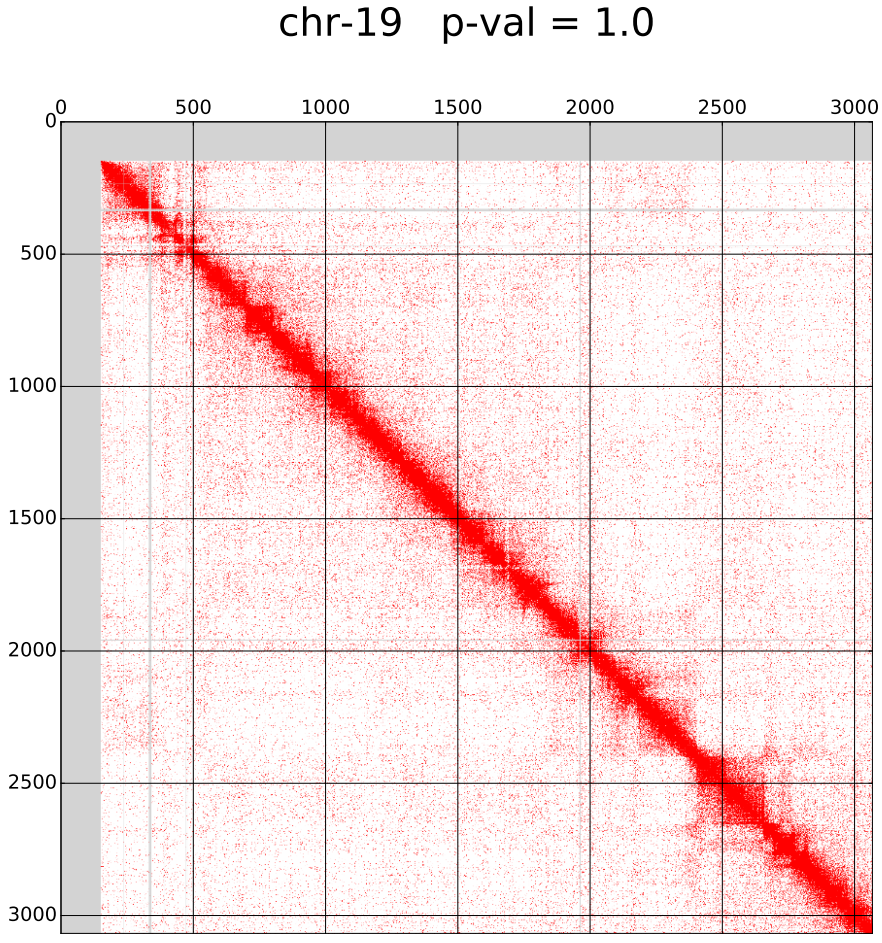


Filtered Network

. . . and a less serious one



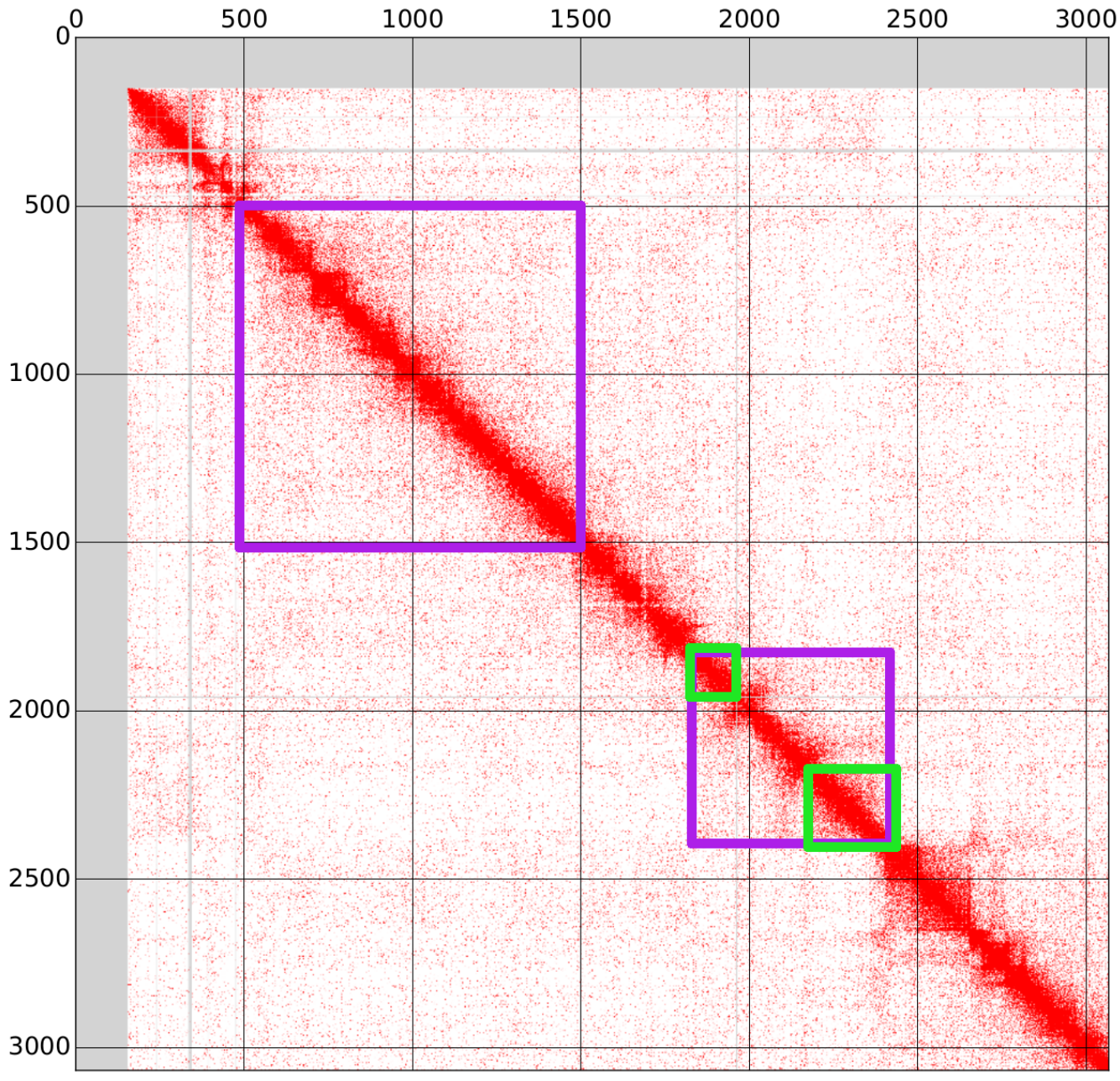
. . . and a less serious one

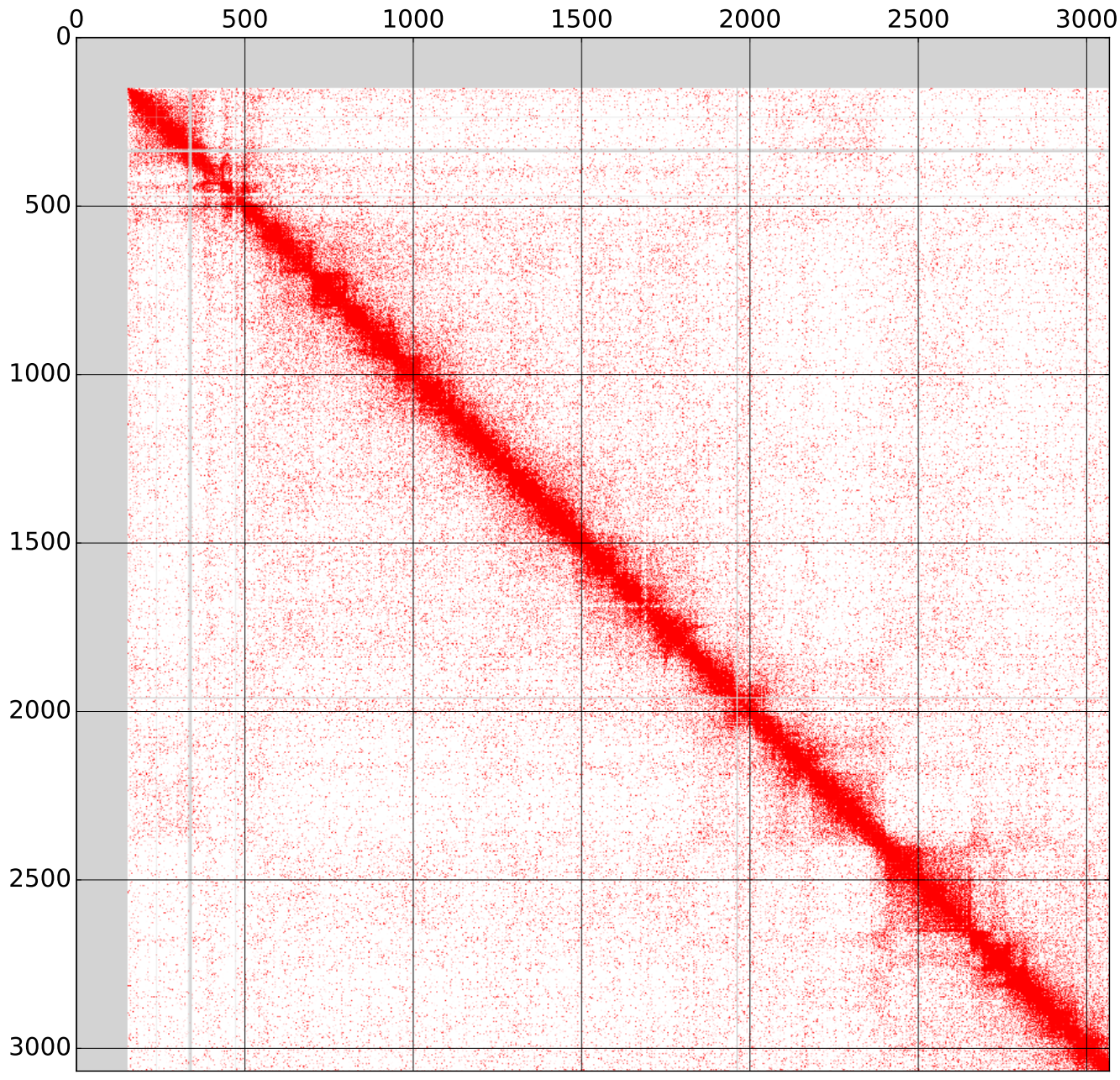


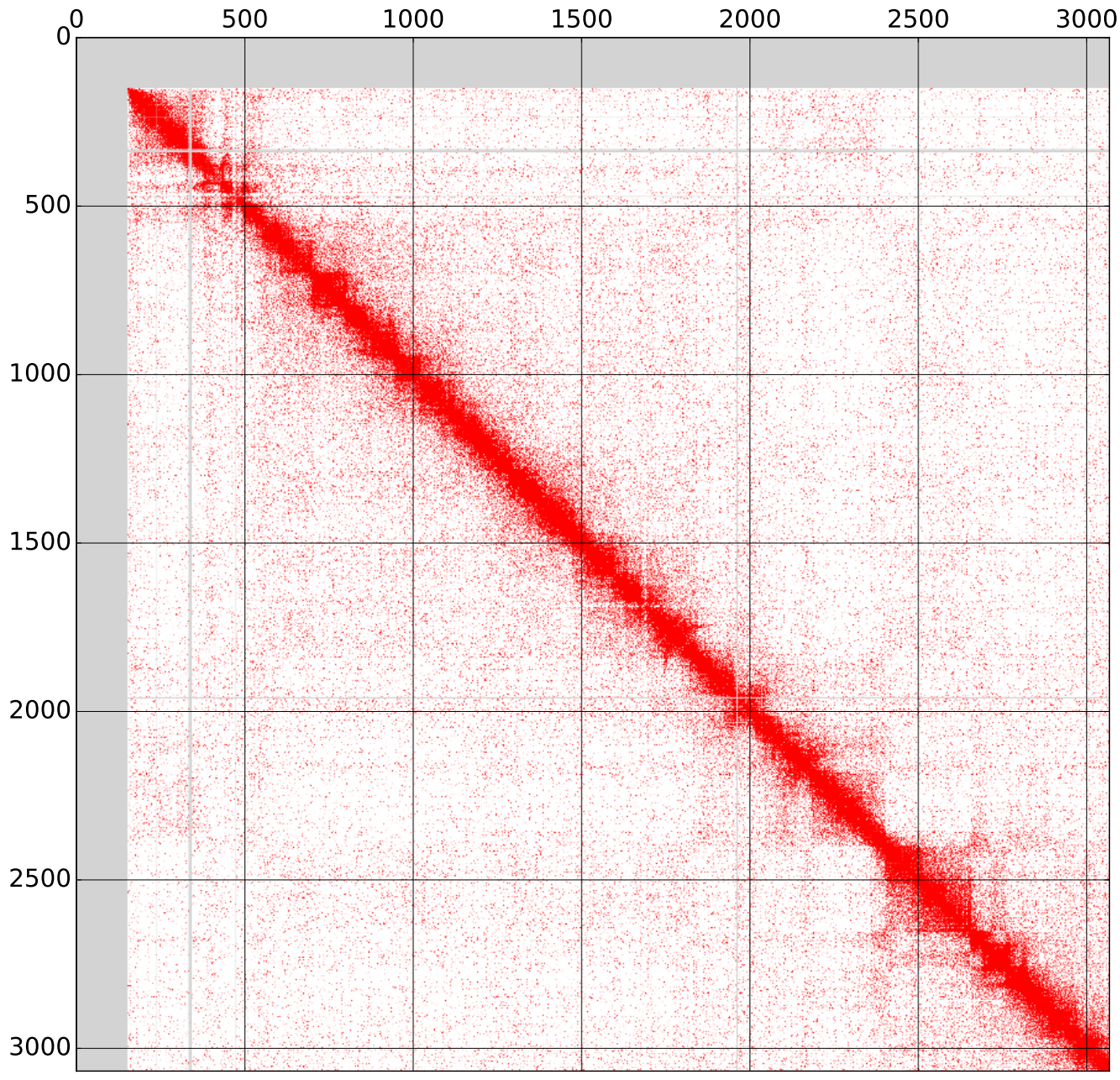
Data

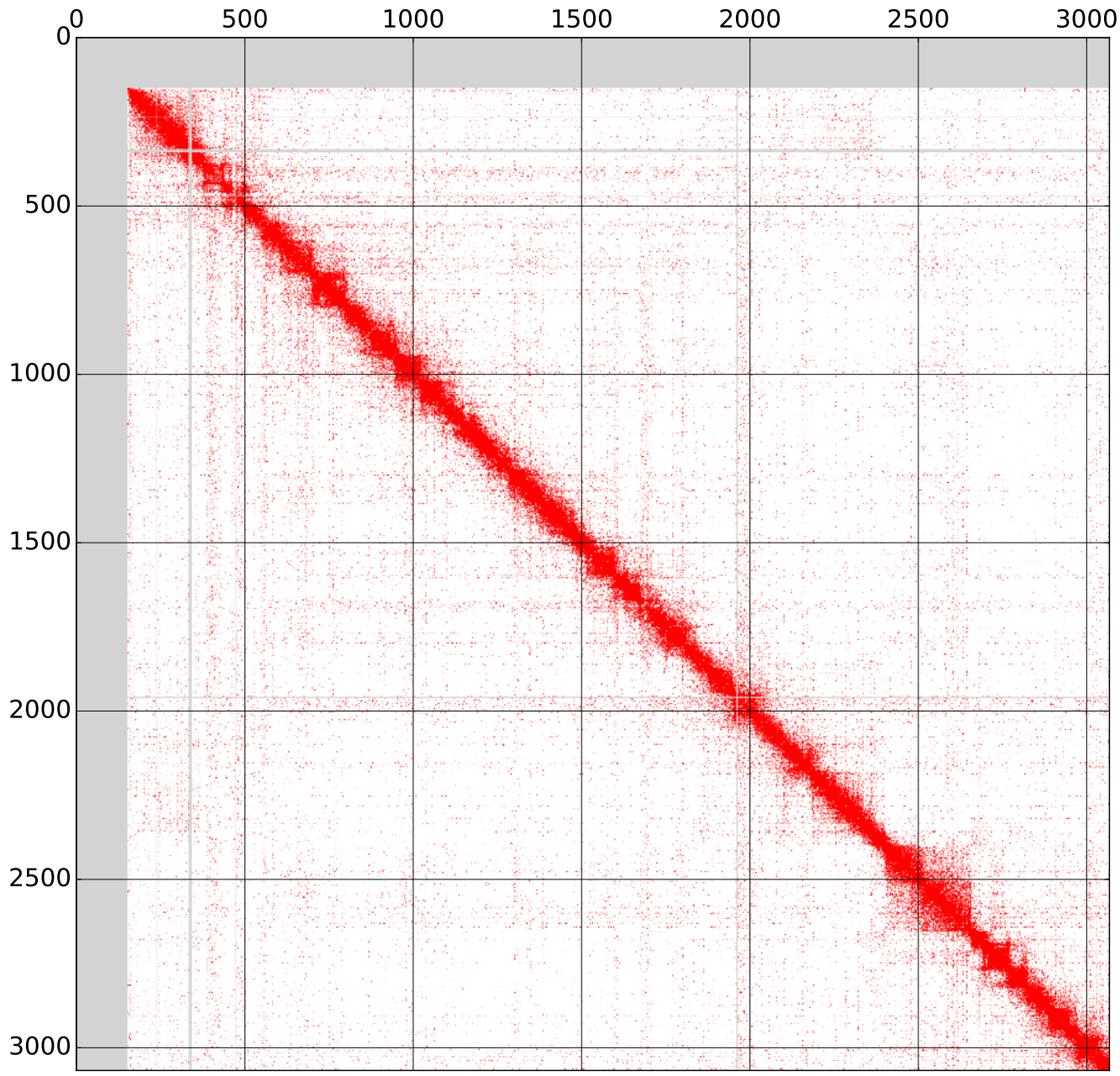
- Mouse (20 chr)
- Hi-C / cis-interactions
- Res. 20 Kbp
- Network sizes: $N \in [3068, 9868]$

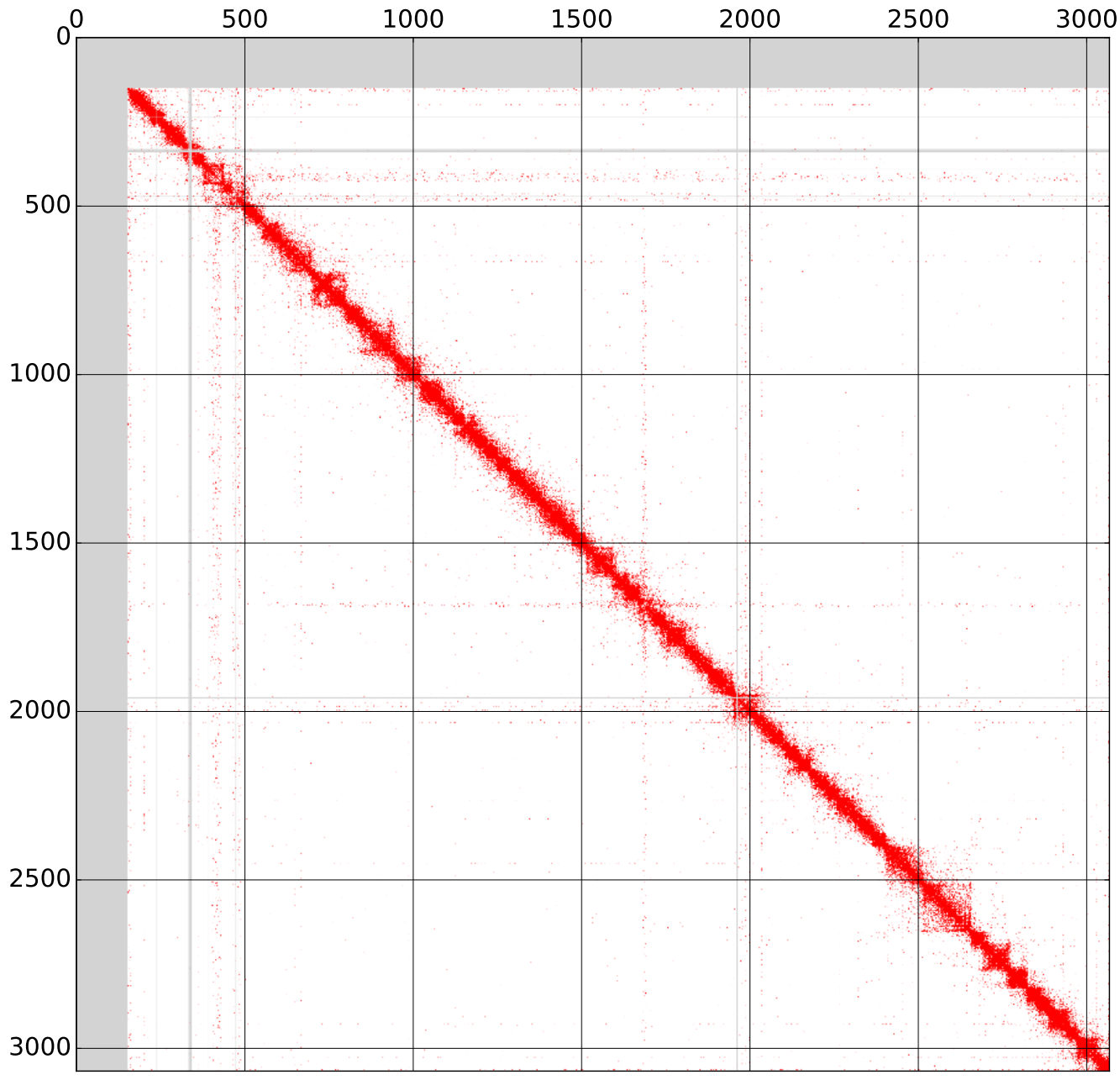
- Imakaev, M. *et al.* Nature Methods, **9**, 999–1003. (2012).
- Giorgetti, L., *et al.* Nature, **535**, 575–579. (2016).


chr-19 p-val = 1.0
0
500
1000
1500
2000
2500
3000


chr-19 p-val = 1.0
0
500
1000
1500
2000
2500
3000
0
500
1000
1500
2000
2500
3000


chr-19 p-val = 0.25
0
500
1000
1500
2000
2500
3000


chr-19 p-val = 0.1
0
500
1000
1500
2000
2500
3000
0
500
1000
1500
2000
2500
3000

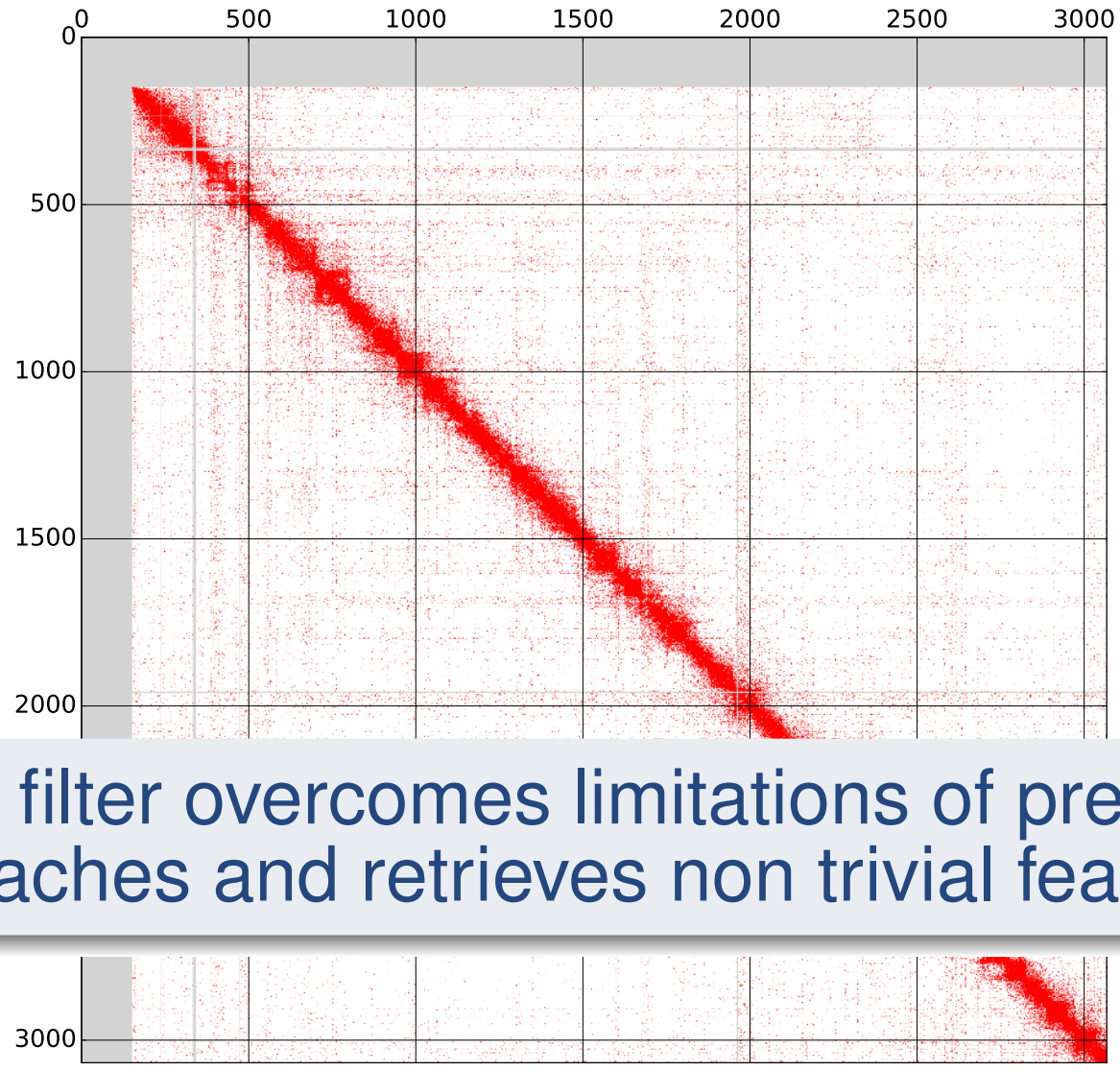

chr-19 p-val = 0.05
0
500
1000
1500
2000
2500
3000

Conclusions

Take home messages

Filtering is becoming a more and more required step to continue using networks in complex systems.

Take home messages



ECM filter overcomes limitations of previous approaches and retrieves non trivial features.

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Acknowledgements

Diego Garlaschelli

Valerio Gemmetto

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Acknowledgements

alessio.cardillo@epfl.ch

http://bifi.es/∼cardillo/

Wanna know more?

Stay tuned on the arXiv . . .

ECM-local: details

1. Generate the null model ensemble and compute:

$$q_{ij}(w) \equiv \frac{\left(x_i x_j\right)^{\Theta(w_{ij})} \left(y_i y_j\right)^{w_{ij}} \left(1 - y_i y_j\right)}{1 - y_i y_j + x_i x_j y_i y_j} \, .$$

ECM-local: details

1. Generate the null model ensemble and compute:

$$q_{ij}(w) \equiv \frac{\left(x_i x_j\right)^{\Theta(w_{ij})} \left(y_i y_j\right)^{w_{ij}} \left(1 - y_i y_j\right)}{1 - y_i y_j + x_i x_j y_i y_j} \, .$$

2. Associate for each link a p-value, γ, such that $P(w_{ij} > w_{ij}^{\star})$.

ECM-local: details

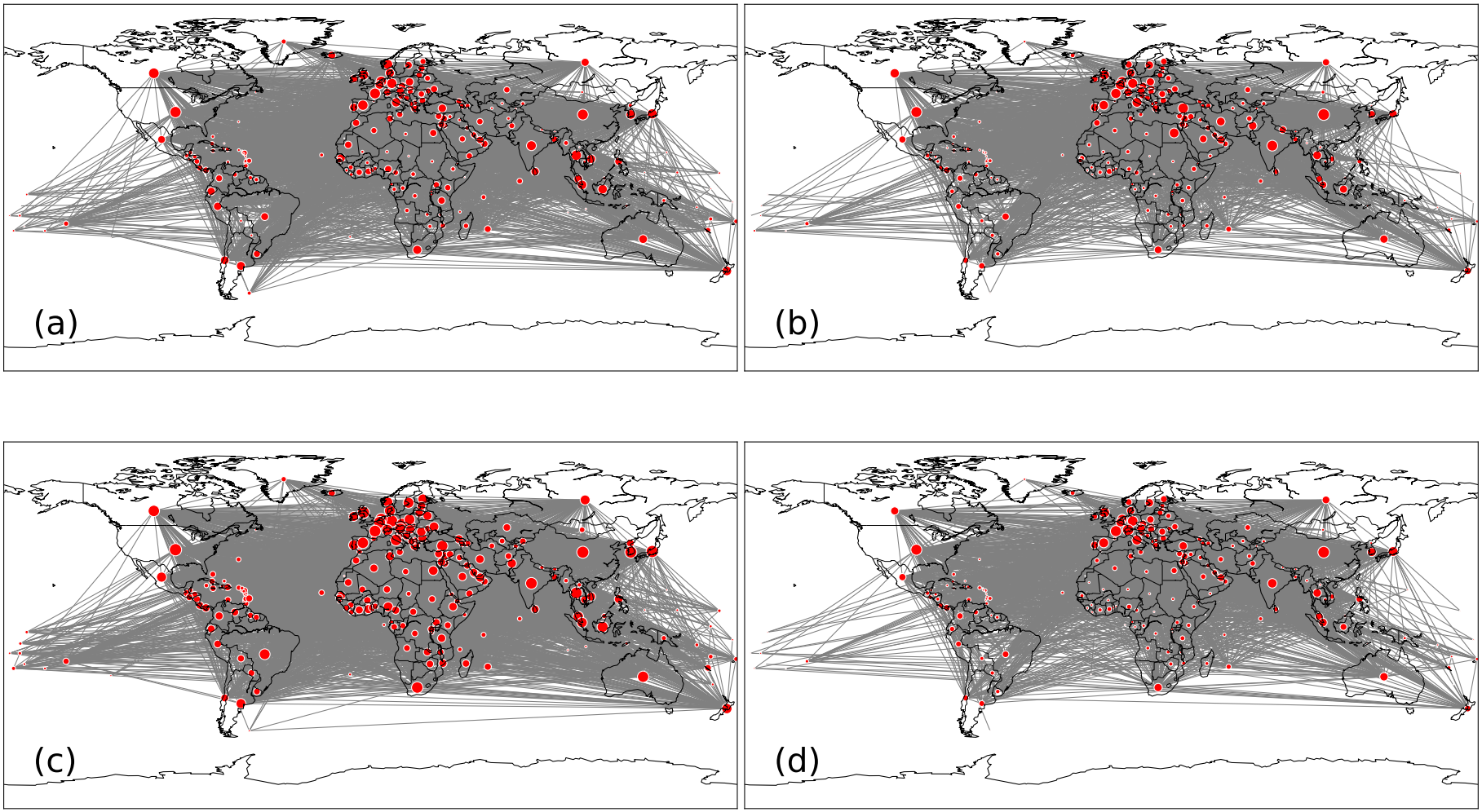
1. Generate the null model ensemble and compute:

$$q_{ij}(w) \equiv \frac{\left(x_i x_j\right)^{\Theta(w_{ij})} \left(y_i y_j\right)^{w_{ij}} \left(1 - y_i y_j\right)}{1 - y_i y_j + x_i x_j y_i y_j} \, .$$

2. Associate for each link a p-value, γ, such that $P(w_{ij} > w_{ij}^{\star})$.
3. Select a threshold $\widetilde{\gamma}$ and remove all the links with $\gamma_{ij} > \widetilde{\gamma}$.

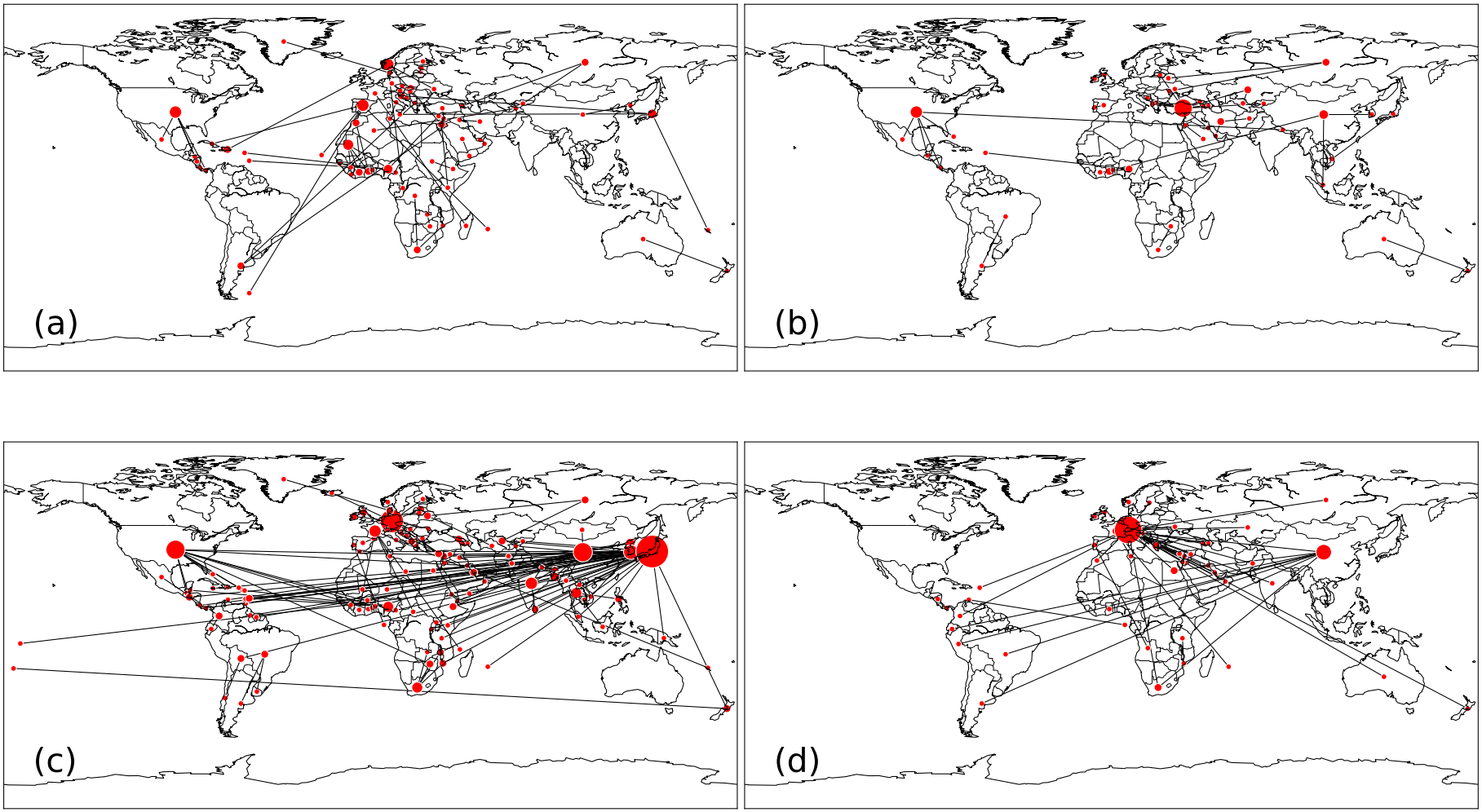
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Trade: single commodities



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Trade: single commodities



ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Trade: single commodities