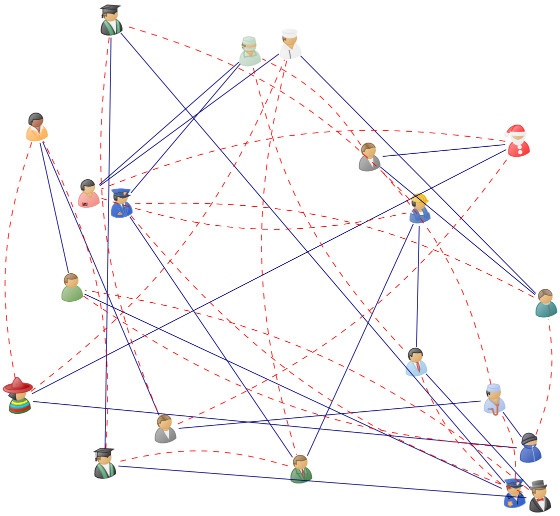
ASSESSING FRUSTRATION IN REAL-WORLD SIGNED NETWORKS: TOWARDS A STATISTICAL THEORY OF BALANCE

Anna Gallo[1], Diego Garlaschelli[1], Renaud Lambiotte[2], Fabio Saracco[3], Tiziano Squartini[1]

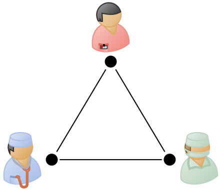
1 — IMT School for Advanced Studies Lucca (Italy)
2 — Mathematical Institute, University of Oxford (United Kingdom)
3 — 'Enrico Fermi' Research Center (CREF), Rome (Italy)

ETH Workshop - ETH Zurich
Signed Relations and Structural Balance in Complex Systems: From Data to Models
15 - 17 May 2024

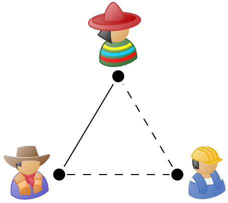


From a historical perspective, the interest in the study of signed networks is rooted into the BALANCE THEORY, stating that social agents tend to avoid the formation of *unbalanced*, or *frustrated* connections.

A friend of a friend will be a friend.



A friend of an enemy will be an enemy.



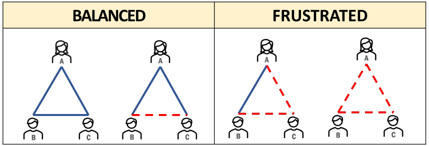
F. Heider, *The Journal of Psychology* (1946).
F. Harary, *Michigan Mathematical Journal* (1953).
D. Cartwright, F. Harary, *Psychological Review* (1956).

Structural balance theory:

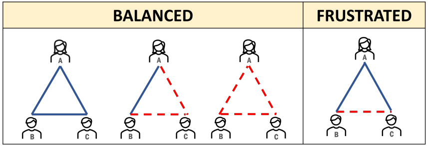
LOCAL PERSPECTIVE

Structural balance theory: **local perspective**

Structural strong balance theory:



Structural weak balance theory:



SSBT → real-world networks tend to avoid triads with an odd number of negative links;
SWBT → real-world networks tend to avoid triads with a single negative link.

These statements can be supported only after a comparison with a benchmark:

Hypothesis: any undirected, signed network has been generated by a probabilistic model

$\Longrightarrow$ we expect some degree of statistical noise to affect any configuration we may want to study.

F. Heider, *The Journal of Psychology* (1946).
F. Harary, *Michigan Mathematical Journal* (1953).
D. Cartwright, F. Harary, *Psychological Review* (1956).

- The position of the edges is fixed, while their signs are shuffled.
  J. Leskovec, D. Huttenlocher, J. Kleinberg, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2010).
  A. Kirkley, G. T. Cantwell, M. E. Newman, *Physical Review E* (2019).

- Signed version of the Local Rewiring Algorithm.
  H. Saiz, J. Gomez-Gardenes, P. Nuche, A. Giron, Y. Pueyo, C. L. Alados, *Ecography* (2017).

- Balanced Signed Chung-Lu model (BSCL).
  T. Derr, C. Aggarwal, J. Tang, *Proceedings of the 27th ACM International Conference on Information and Knowledge Management* (2018).

- Null models defined constraining structural properties of signed networks within the ERG framework.
  J. Lerner, *Social Networks* (2016).
  C. Fritz, M. Mehrl, P. W. Thurner, et al., *arXiv:2205.13411* (2022).

A. G., D. Garlaschelli, R. Lambiotte, F. Saracco, T. Squartini, *Communications Physics* (2024)

We formally extend the **ERGMs** framework to **binary, undirected, signed networks.**

The maximisation of the **Shannon entropy**,

$$S[P] = -\sum_{\mathbf{A}\in\mathbb{A}} P(\mathbf{A}) \ln P(\mathbf{A})$$

leads to identify the functional form of the probability distribution $P(\mathbf{A})$

$$P(\mathbf{A}) = \frac{e^{-H(\mathbf{A})}}{\sum_{\mathbf{A}\in\mathbb{A}} e^{-H(\mathbf{A})}} = \frac{e^{-\sum_{i=1}^{M}\theta_i C_i(\mathbf{A})}}{\sum_{i=1}^{M} e^{-\sum_{i=1}^{M}\theta_i C_i(\mathbf{A})}},$$

that preserves a set of empirical constraints on average.

J. Park and M. E. J. Newman, *Physics Review E* (2004).

T. Squartini, D. Garlaschelli, *New Journal of Physics* (2011).

T. Squartini, D. Garlaschelli, *Springer* (2017).

The **ERGMs** can be employed for

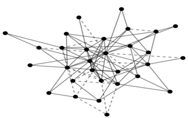
- reconstructing a network;
- detecting statistically significant patterns in networks.

We employ our benchmarks to compare the empirical abundance of short cycles (→ triangles) with its expected value, on a bunch of real-world systems.

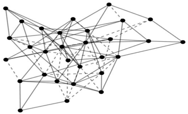
T. Squartini, D. Garlaschelli, **Maximum-Entropy Networks**, *Springer* (2017).

G. Cimini, T. Squartini, F. Saracco, D. Garlaschelli, A. Gabrielli, G. Caldarelli, **The statistical physics of real-world networks**, *Nature Reviews Physics* (2019).

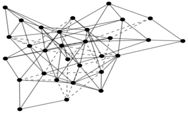
$$H(\mathbf{A}) = \alpha L^+(\mathbf{A}) + \beta L^-(\mathbf{A})$$



Realisation of a SRGM



Real network



Realisation of a SRGM-FT

Signed Random Graph Model

$$p^- = \frac{e^{-\beta}}{1 + e^{-\alpha} + e^{-\beta}} = \frac{2L^-}{N(N-1)},$$

$$p^+ = \frac{e^{-\alpha}}{1 + e^{-\alpha} + e^{-\beta}} = \frac{2L^+}{N(N-1)},$$

$$p^0 = 1 - p^- - p^+.$$

Signed Random Graph Model with Fixed Topology

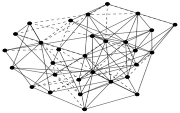
$$p^- = \frac{e^{-\beta}}{e^{-\alpha} + e^{-\beta}} = \frac{L^-}{L^*},$$

$$p^+ = \frac{e^{-\alpha}}{e^{-\alpha} + e^{-\beta}} = \frac{L^+}{L^*}$$

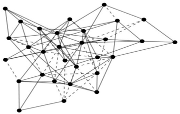
$$p^+ + p^- = 1.$$

A. G., D. Garlaschelli, R. Lambiotte, F. Saracco, T. Squartini, *Communications Physics* (2024)

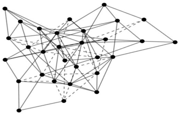
$$H(\mathbf{A}) = \sum_{i=1}^{N}(\alpha_i k_i^{+}(\mathbf{A}) + \beta_i k_i^{-}(\mathbf{A}))$$



Realisation of a SCM



Real network



Realisation of a SCM-FT

Signed Configuration Model

$$p_{ij}^{-} = \frac{e^{-(\beta_i+\beta_j)}}{1+e^{-(\alpha_i+\alpha_j)}+e^{-(\beta_i+\beta_j)}},$$

$$p_{ij}^{+} = \frac{e^{-(\alpha_i+\alpha_j)}}{1+e^{-(\alpha_i+\alpha_j)}+e^{-(\beta_i+\beta_j)}},$$

$$p_{ij}^{0} = 1 - p_{ij}^{-} - p_{ij}^{+}.$$

Signed Configuration Model with Fixed Topology

$$p_{ij}^{-} = \frac{e^{-(\beta_i+\beta_j)}}{e^{-(\alpha_i+\alpha_j)}+e^{-(\beta_i+\beta_j)}},$$

$$p_{ij}^{+} = \frac{e^{-(\alpha_i+\alpha_j)}}{e^{-(\alpha_i+\alpha_j)}+e^{-(\beta_i+\beta_j)}},$$

$$p_{ij}^{-} + p_{ij}^{+} = 1.$$

A. G., D. Garlaschelli, R. Lambiotte, F. Saracco, T. Squartini, *Communications Physics* (2024)

SCM

$$k_i^+(\mathbf{A}^*) = \sum_{\substack{j=1 \\ (j \neq i)}}^{N} p_{ij}^+ = \langle k_i^+ \rangle \quad \forall\, i,$$

$$k_i^-(\mathbf{A}^*) = \sum_{\substack{j=1 \\ (j \neq i)}}^{N} p_{ij}^- = \langle k_i^- \rangle \quad \forall\, i.$$

SCM-FT

$$k_i^+(\mathbf{A}^*) = \sum_{\substack{j=1 \\ (j \neq i)}}^{N} |a_{ij}^*| p_{ij}^+ = \langle k_i^+ \rangle \quad \forall\, i,$$

$$k_i^-(\mathbf{A}^*) = \sum_{\substack{j=1 \\ (j \neq i)}}^{N} |a_{ij}^*| p_{ij}^- = \langle k_i^- \rangle \quad \forall\, i.$$

	SCM			SCM-FT		
	MAE	MRE	Time (s)	MAE	MRE	Time (s)
CoW, 1946-49	$\simeq 6.4 \cdot 10^{-2}$	$\simeq 1.4 \cdot 10^{-2}$	$\simeq 0.02$	$\simeq 7.1 \cdot 10^{-2}$	$\simeq 3.9 \cdot 10^{-3}$	$\simeq 0.01$
CoW, 1950-53	$\simeq 7.6 \cdot 10^{-2}$	$\simeq 2.1 \cdot 10^{-2}$	$\simeq 0.02$	$\simeq 6.0 \cdot 10^{-2}$	$\simeq 3.2 \cdot 10^{-2}$	$\simeq 0.01$
CoW, 1990-93	$\simeq 9.8 \cdot 10^{-2}$	$\simeq 2.7 \cdot 10^{-2}$	$\simeq 0.12$	$\simeq 9.0 \cdot 10^{-2}$	$\simeq 3.1 \cdot 10^{-2}$	$\simeq 0.07$
CoW, 1994-97	$\simeq 9.8 \cdot 10^{-2}$	$\simeq 2.8 \cdot 10^{-2}$	$\simeq 0.08$	$\simeq 5.3 \cdot 10^{-2}$	$\simeq 2.1 \cdot 10^{-2}$	$\simeq 0.06$

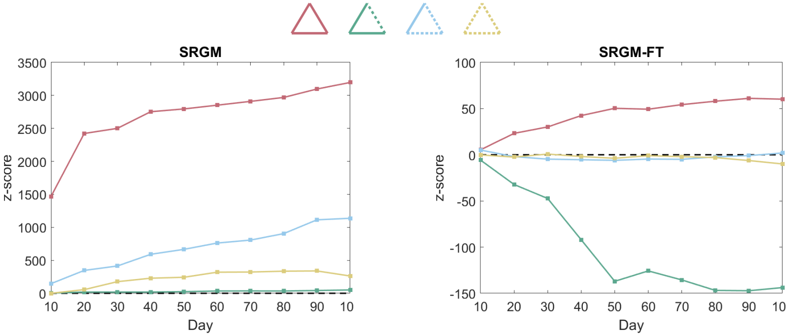
Table: Performance of the fixed-point algorithm to solve the systems of equations defining the SCM and the SCM-FT on some snapshots of the Correlates of Wars dataset.

For further information about the numerical methods to solve the ERGMs, see N. Vallarano, M. Bruno, E. Marchese, G. Trapani, F. Saracco, G. Cimini, M. Zanon, T. Squartini, *Sci. Rep.* (2021) and the toolbox **NEMtropy** on Github *(https://pypi.org/project/NEMtropy/)*.

Results · Massive Multiplayer Online Game (MMOG)

$$z_{\triangle} = \frac{N_{\triangle}(\mathbf{A}^*) - \langle N_{\triangle} \rangle}{\sigma[N_{\triangle}]} \rightarrow \begin{cases} > 0 & \Longrightarrow \text{tendency of the } \triangle \text{ to be over-represented} \\ < 0 & \Longrightarrow \text{tendency of the } \triangle \text{ to be under-represented} \end{cases}$$

in the data with respect to the null model.



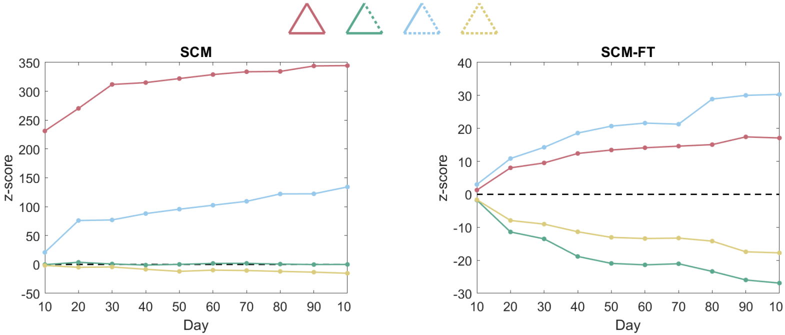
In *Gallo et al.* (2024) a closed form expressions for the z-scores is provided.

M. Szell, R. Lambiotte, S. Thurner, *Proceedings of the National Academy of Sciences* (2010).
A. G., D. Garlaschelli, R. Lambiotte, F. Saracco, T. Squartini, *Communications Physics* (2024)

Results · Massive Multiplayer Online Game (MMOG)

$$z_\triangle = \frac{N_\triangle(\mathbf{A}^*) - \langle N_\triangle \rangle}{\sigma[N_\triangle]} \rightarrow \begin{cases} > 0 & \Longrightarrow \text{ tendency of the } \triangle \text{ to be over-represented} \\ < 0 & \Longrightarrow \text{ tendency of the } \triangle \text{ to be under-represented} \end{cases}$$

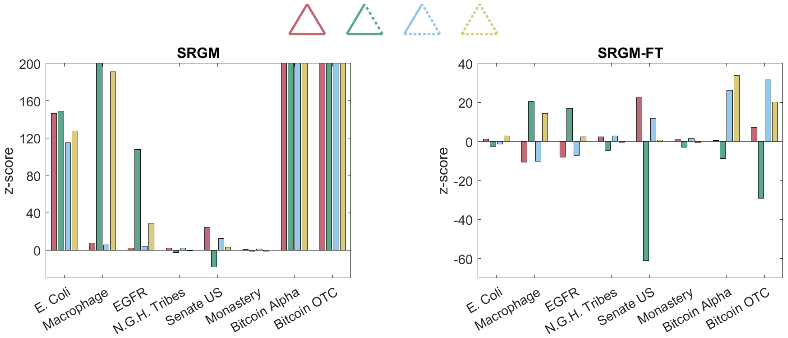
in the data with respect to the null model.



In *Gallo et al.* (2024) a closed form expressions for the z-scores is provided.

M. Szell, R. Lambiotte, S. Thurner, *Proceedings of the National Academy of Sciences* (2010).
A. G., D. Garlaschelli, R. Lambiotte, F. Saracco, T. Squartini, *Communications Physics* (2024)

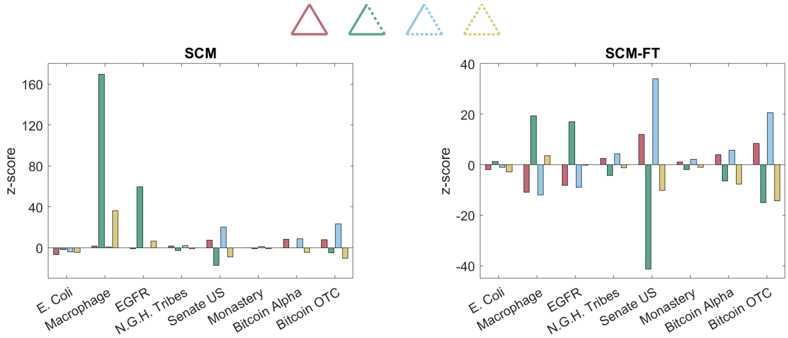
$$z_\triangle = \frac{N_\triangle(\mathbf{A}^*) - \langle N_\triangle \rangle}{\sigma[N_\triangle]} \rightarrow \begin{cases} > 0 & \Longrightarrow \text{tendency of the } \triangle \text{ to be over-represented} \\ < 0 & \Longrightarrow \text{tendency of the } \triangle \text{ to be under-represented} \end{cases}$$ in the data with respect to the null model.



In *Gallo et al.* (2024) a closed form expressions for the z-scores is provided.

S. Aref, M. C. Wilson, *Journal of Complex Networks* (2019).
A. G., D. Garlaschelli, R. Lambiotte, F. Saracco, T. Squartini, *Communications Physics* (2024)

$$z_\triangle = \frac{N_\triangle(\mathbf{A}^*) - \langle N_\triangle \rangle}{\sigma[N_\triangle]} \rightarrow \begin{cases} > 0 & \Longrightarrow \text{tendency of the } \triangle \text{ to be over-represented} \\ < 0 & \Longrightarrow \text{tendency of the } \triangle \text{ to be under-represented} \end{cases}$$ in the data with respect to the null model.



In *Gallo et al.* (2024) a closed form expressions for the z-scores is provided.

S. Aref, M. C. Wilson, *Journal of Complex Networks* (2019).
A. G., D. Garlaschelli, R. Lambiotte, F. Saracco, T. Squartini, *Communications Physics* (2024).

An unambiguous conclusion about the theory best explaining the phenomenon cannot be reached, since it turns out to depend on a number of factors:

- homogeneous null models favour the WBT;
- heterogeneous null model favour the SBT;
- agents, that can only choose *how* interact, strongly avoid to engage in frustrated relationships (fixed topology benchmarks);
- agents, that can choose *with whom* and *how* interact, seem to be more tolerant (free topology benchmarks).

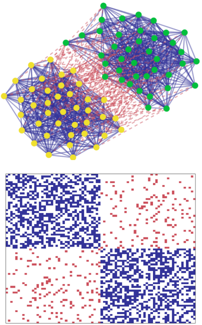
Structural balance theory:

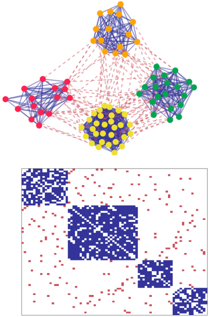
MESOSCOPIC PERSPECTIVE

Balance theory

STRONG

WEAK





● → positive links   ● → negative links

F. Harary, *Michigan Mathematical Journal* (1953).

D. Cartwright, F. Harary, *Psychological Review* (1956).

Frustration index or line index of balance:

$$F(\boldsymbol{\sigma}) = \sum_{i=1}^{N} \sum_{j(>i)} a_{ij}^{-} \delta_{\sigma_i,\sigma_j} + \sum_{i=1}^{N} \sum_{j(>i)} a_{ij}^{+} (1 - \delta_{\sigma_i,\sigma_j})$$

$$= L_{\bullet}^{-} + L_{\circ}^{+}.$$

Theorem 1: $F(\boldsymbol{\sigma}) = 0 \iff$ the partition $\boldsymbol{\sigma}$ is k-balanced.

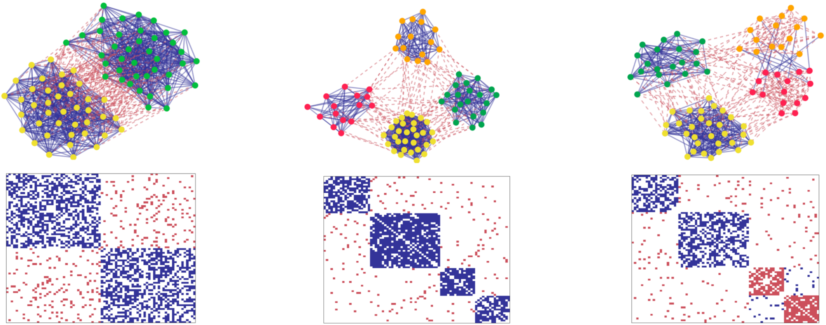
	$F(\boldsymbol{\sigma})$
Fraternity	1
N.G.H. Tribes	2
Slovenian Parliament	2
Monastery	5
US Senate	247
CoW, 1946-49	12
CoW, 1950-53	11
CoW, 1954-57	27
CoW, 1958-61	25
EGFR	189
Macrophage	316
Bitcoin Alpha	1399
Bitcoin OTC	3259

According to the F-test, none of the listed, real-world networks turns out to satisfy the TBT.

A generalized theory of balance: the **relaxed balance theory**

Balance theory

STRONG | WEAK | RELAXED



● → positive links ● → negative links

F. Harary, *Michigan Mathematical Journal* (1953).

D. Cartwright, F. Harary, *Psychological Review* (1956).

P. Doreian, A. Mrvar, *Social Networks* (2009).

Both the TBT and the RBT present limitations:

- the TBT quickly dismisses most real graphs as frustrated;
- the RBT lacks a proper mathematisation since a score function as $F(\sigma)$ is missing.

How to overcome such limitations?

Recasting the theory of balance within a statistical framework.

P. Doreian, A. Mrvar, *Social Networks* (2009).
A. G., D. Garlaschelli, T. Squartini — *ArXiv: 2404.15914* (2024).

Let us assume the presence of a probabilistic model behind the appearance of any signed configuration G and let k be the number of modules.

Then, the TBT can be rephrased by requiring $p_{rr}^{-} = 0,\ r = 1 \dots k$ and $p_{rs}^{+} = 0,\ \forall\, r < s$.

Let us assume the presence of a probabilistic model behind the appearance of any signed configuration G and let k be the number of modules.

Then, the TBT can be rephrased by requiring $p_{rr}^{-}=0,\ r=1\dots k$ and $p_{rs}^{+}=0,\ \forall\, r<s$.

Softening these positions leads us to define the following inference scheme:

- if G satisfies

$$p_{rr}^{+}>p_{rr}^{-},\quad r=1\dots k\quad\text{and}\quad p_{rs}^{+}<p_{rs}^{-},\quad\forall\, r<s,$$

  then G obeys the statistical variant of the TBT;
- otherwise, it obeys the statistical variant of the RBT.

A. G., D. Garlaschelli, T. Squartini — *ArXiv: 2404.15914* (2024).

To test our inference scheme we need ① a signed, generative model, and ② a proper score function to optimize.

① Signed Stochastic Block Model.

$$\mathcal{L}_{\mathrm{SSBM}}=\prod_{r=1}^{k}(p_{rr}^{+})^{L_{rr}^{+}}(p_{rr}^{-})^{L_{rr}^{-}}(1-p_{rr}^{+}-p_{rr}^{-})^{\binom{N_r}{2}-L_{rr}}\prod_{r=1}^{k}\prod_{s(>r)}(p_{rs}^{+})^{L_{rs}^{+}}(p_{rs}^{-})^{L_{rs}^{-}}(1-p_{rs}^{+}-p_{rs}^{-})^{N_r N_s-L_{rs}}.$$

② Bayesian Information Criterion.

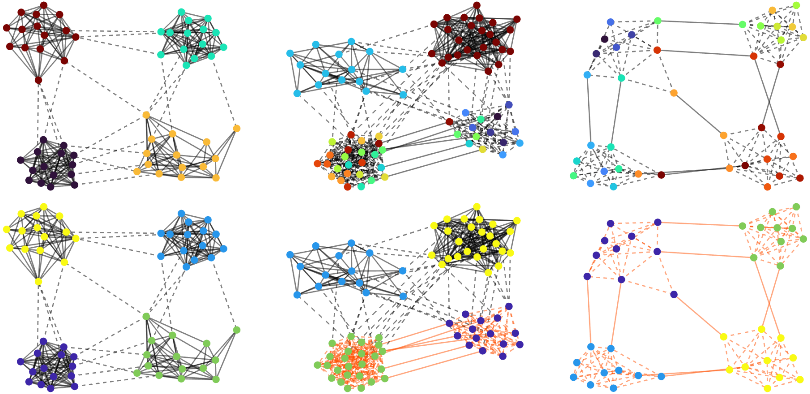
$$\mathrm{BIC}=\kappa_{\mathrm{SSBM}}\ln n-2\ln\mathcal{L}_{\mathrm{SSBM}},$$

where $\kappa_{\mathrm{SSBM}}=k(k+1)$ and $n=N(N-1)/2$.

A. G., D. Garlaschelli, T. Squartini — *ArXiv: 2404.15914* (2024).

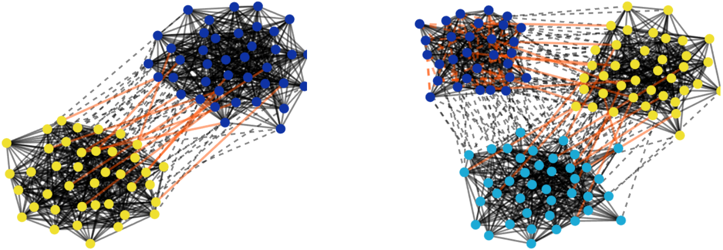
Partitions recovered minimizing F →

Partitions recovered minimizing BIC →



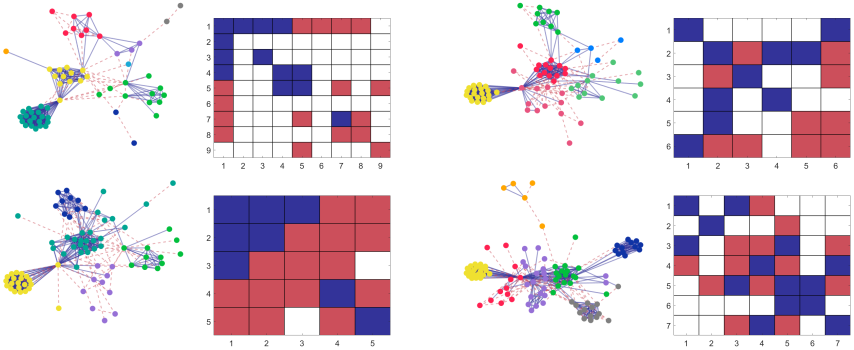
A. G., D. Garlaschelli, T. Squartini — *ArXiv: 2404.15914* (2024).

The F-test may return configurations that are neither traditionally nor relaxedly balanced:



→ Thanks to our statistical framework they are classified as balanced according to the statistical variant of the TBT.

A. G., D. Garlaschelli, T. Squartini — *ArXiv: 2404.15914* (2024).



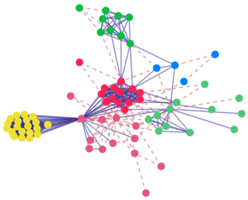
Minimisation of BIC on four snapshots of the Correlates of Wars dataset.
A generic block, indexed as rs, is coloured in blue if $L^+_{rs} > L^-_{rs}$, in red if $L^+_{rs} < L^-_{rs}$ and in white if $L^+_{rs} = L^-_{rs}$.

→ Thanks to our statistical framework, we conclude that the partitions above obey the statistical variant of the RBT.

P. Doreian, A. Mrvar, *Journal of Social Structure* (2015).
A. G., D. Garlaschelli, T. Squartini — *ArXiv: 2404.15914* (2024).

→ Mesoscopic perspective:

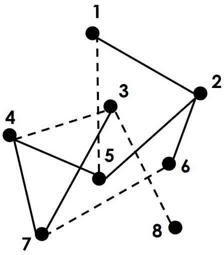
- most real-world networks are quickly dismissed as frustrated, according to the TBT, by the F -test;
- the RBT attempts to overcome such a limitation but lacks a proper mathematisation;
- recasting the BT within a statistical framework allows to achieve both goals (i.e. overcoming the limitations of the TBT and formalising the RBT);
- our contribution reconciles the generally contradictory results one gets when combining a purely structure-based community detection with a purely sign-based one.



Thanks for your attention!

anna.gallo@imtlucca.it

- A. G., D. Garlaschelli, R. Lambiotte, F. Saracco, T. Squartini, *Communications Physics* (2024)
- A. G., D. Garlaschelli, T. Squartini — *ArXiv: 2404.15914* (2024).



Signed adjacency matrix

$\mathbf{A} \equiv \{a_{ij}\}_{i,j}^{N}$ with $a_{ij} \in \{-1, 0, +1\}$

We introduce the auxiliary variables

$$a_{ij}^{+} = \begin{cases} 1, & \text{if } a_{ij} > 0, \\ 0, & \text{otherwise;} \end{cases} \qquad a_{ij}^{-} = \begin{cases} 1, & \text{if } a_{ij} < 0, \\ 0, & \text{otherwise,} \end{cases}$$

which allow to write

$$\mathbf{A} = \mathbf{A}^{+} - \mathbf{A}^{-} \quad \text{and} \quad |\mathbf{A}| = \mathbf{A}^{+} + \mathbf{A}^{-}$$

where $\mathbf{A}^{+} \equiv \{a_{ij}^{+}\}_{i,j}^{N}$ and $\mathbf{A}^{-} \equiv \{a_{ij}^{-}\}_{i,j}^{N}$.

Positive edges

$$L^{+} = \sum_{i=1}^{N} \sum_{\substack{j=1 \\ (j>i)}}^{N} a_{ij}^{+}$$

Negative edges

$$L^{-} = \sum_{i=1}^{N} \sum_{\substack{j=1 \\ (j>i)}}^{N} a_{ij}^{-}$$

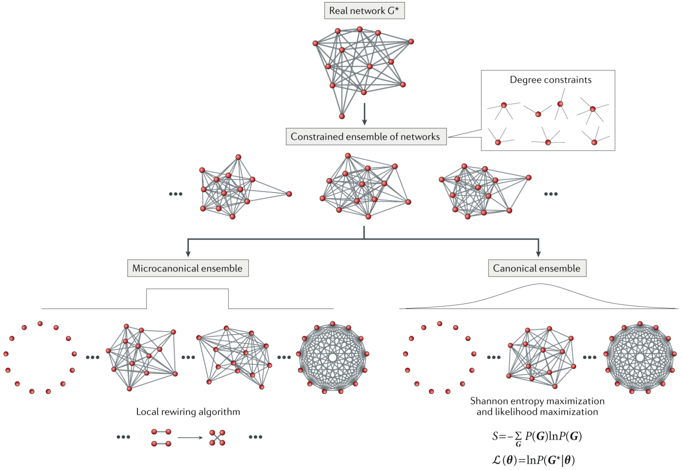
Positive degree

$$k_i^{+} = \sum_{\substack{j=1 \\ (j\neq i)}}^{N} a_{ij}^{+}$$

Negative degree

$$k_i^{-} = \sum_{\substack{j=1 \\ (j\neq i)}}^{N} a_{ij}^{-}$$

The solution of the maximization problem depends on whether we require the constraints to be soft or hard.



G. Cimini, T. Squartini, F. Saracco, D. Garlaschelli, A. Gabrielli, G. Caldarelli, **The statistical physics of real-world networks**, *Nature Reviews Physics* (2019).

Let us notice that we can write

$$P_{\text{SRGM}}(\mathbf{A}) = P_{\text{RGM}}(\mathbf{A}) \cdot \frac{P_{\text{SRGM}}(\mathbf{A})}{P_{\text{RGM}}(\mathbf{A})}$$

$$= p^L(1-p)^{\binom{N}{2}-L} \cdot \frac{(p^-)^{L^-}(p^+)^{L^+}(1-p^- - p^+)^{\binom{N}{2}-L^- -L^+}}{p^L(1-p)^{\binom{N}{2}-L}}$$

$$= p^L(1-p)^{\binom{N}{2}-L} \cdot \frac{(p^-)^{L^-}(p^+)^{L^+}(1-p^- - p^+)^{\binom{N}{2}-L^- -L^+}}{p^{L^-}p^{L^+}(1-p)^{\binom{N}{2}-L^- -L^+}}$$

$$= p^L(1-p)^{\binom{N}{2}-L} \cdot \left(\frac{p^-}{p}\right)^{L^-}\left(\frac{p^+}{p}\right)^{L^+}\left(\frac{1-p^- - p^+}{1-p}\right)^{\binom{N}{2}-L^- -L^+}.$$

Since the RGM induced by the SRGM satisfies the relationship $p \equiv p^- + p^+$, one has

$$P_{\text{SRGM}}(\mathbf{A}) = p^L(1-p)^{\binom{N}{2}-L} \cdot \left(\frac{p^-}{p}\right)^{L^-}\left(\frac{p^+}{p}\right)^{L^+}$$

$$= P_{\text{RGM}}(\mathbf{A}) \cdot P_{\text{SRGM-FT}}(\mathbf{A})$$

Let us notice that we can write

$$P_{\text{SCM}}(\mathbf{A}) = P_{\text{FM}}(\mathbf{A}) \cdot \frac{P_{\text{SCM}}(\mathbf{A})}{P_{\text{FM}}(\mathbf{A})}$$

$$= \prod_{i=1}^{N} \prod_{\substack{j=1\\(j>i)}}^{N} p_{ij}^{a_{ij}} (1-p_{ij})^{1-a_{ij}} \cdot \frac{\prod_{i=1}^{N} \prod_{\substack{j=1\\(j>i)}}^{N} (p_{ij}^{-})^{a_{ij}^{-}} (p_{ij}^{+})^{a_{ij}^{+}} (1-p_{ij}^{-}-p_{ij}^{+})^{1-a_{ij}^{-}-a_{ij}^{+}}}{\prod_{i=1}^{N} \prod_{\substack{j=1\\(j>i)}}^{N} p_{ij}^{a_{ij}^{-}+a_{ij}^{+}} (1-p_{ij})^{1-a_{ij}^{-}-a_{ij}^{+}}}$$

$$= \prod_{i=1}^{N} \prod_{\substack{j=1\\(j>i)}}^{N} p_{ij}^{a_{ij}} (1-p_{ij})^{1-a_{ij}} \cdot \prod_{i=1}^{N} \prod_{\substack{j=1\\(j>i)}}^{N} \left(\frac{p_{ij}^{-}}{p_{ij}}\right)^{a_{ij}^{-}} \left(\frac{p_{ij}^{+}}{p_{ij}}\right)^{a_{ij}^{+}} \left(\frac{1-p_{ij}^{-}-p_{ij}^{+}}{1-p_{ij}}\right)^{1-a_{ij}^{-}-a_{ij}^{+}},$$

where $P_{\text{FM}}(\mathbf{A})$ indicates the probability distribution of any factorizable (null) model. Upon requiring $p_{ij} \equiv p_{ij}^{-} + p_{ij}^{+}$, the FM turns out to be an *induced CM* and this let us to obtain

$$P_{\text{SCM}}(\mathbf{A}) = \prod_{i=1}^{N} \prod_{\substack{j=1\\(j>i)}}^{N} p_{ij}^{a_{ij}} (1-p_{ij})^{1-a_{ij}} \cdot \prod_{i=1}^{N} \prod_{\substack{j=1\\(j>i)}}^{N} \left(\frac{p_{ij}^{-}}{p_{ij}}\right)^{a_{ij}^{-}} \left(\frac{p_{ij}^{+}}{p_{ij}}\right)^{a_{ij}^{+}}$$

	SCM			**SCM-FT**		
	MAE	MRE	Time (s)	MAE	MRE	Time (s)
MMOG, Day 10	$\simeq 1.2 \cdot 10^{-1}$	$\simeq 2.1 \cdot 10^{-2}$	$\simeq 7$	$\simeq 1.7 \cdot 10^{-2}$	$\simeq 1.0 \cdot 10^{-2}$	$\simeq 5$
MMOG, Day 20	$\simeq 1.4 \cdot 10^{-1}$	$\simeq 2.6 \cdot 10^{-2}$	$\simeq 9$	$\simeq 3.1 \cdot 10^{-2}$	$\simeq 1.5 \cdot 10^{-2}$	$\simeq 21$
MMOG, Day 30	$\simeq 2.0 \cdot 10^{-1}$	$\simeq 3.1 \cdot 10^{-2}$	$\simeq 10$	$\simeq 6.5 \cdot 10^{-2}$	$\simeq 2.3 \cdot 10^{-2}$	$\simeq 34$
MMOG, Day 40	$\simeq 2.0 \cdot 10^{-1}$	$\simeq 3.2 \cdot 10^{-2}$	$\simeq 12$	$\simeq 8.4 \cdot 10^{-2}$	$\simeq 2.2 \cdot 10^{-2}$	$\simeq 35$
MMOG, Day 50	$\simeq 1.9 \cdot 10^{-1}$	$\simeq 3.3 \cdot 10^{-2}$	$\simeq 14$	$\simeq 1.2 \cdot 10^{-1}$	$\simeq 4.2 \cdot 10^{-2}$	$\simeq 54$
MMOG, Day 60	$\simeq 2.0 \cdot 10^{-1}$	$\simeq 3.0 \cdot 10^{-2}$	$\simeq 18$	$\simeq 1.1 \cdot 10^{-1}$	$\simeq 3.8 \cdot 10^{-2}$	$\simeq 80$
MMOG, Day 70	$\simeq 2.2 \cdot 10^{-1}$	$\simeq 3.2 \cdot 10^{-2}$	$\simeq 19$	$\simeq 9.1 \cdot 10^{-2}$	$\simeq 2.8 \cdot 10^{-2}$	$\simeq 85$
MMOG, Day 80	$\simeq 1.9 \cdot 10^{-1}$	$\simeq 3.0 \cdot 10^{-2}$	$\simeq 22$	$\simeq 1.1 \cdot 10^{-1}$	$\simeq 5.2 \cdot 10^{-2}$	$\simeq 109$
MMOG, Day 90	$\simeq 2.0 \cdot 10^{-1}$	$\simeq 3.4 \cdot 10^{-2}$	$\simeq 25$	$\simeq 1.3 \cdot 10^{-1}$	$\simeq 4.5 \cdot 10^{-2}$	$\simeq 120$
MMOG, Day 100	$\simeq 2.1 \cdot 10^{-1}$	$\simeq 3.2 \cdot 10^{-2}$	$\simeq 27$	$\simeq 1.1 \cdot 10^{-1}$	$\simeq 4.7 \cdot 10^{-2}$	$\simeq 138$

Table: Performance of the fixed-point algorithm to solve the systems of equations defining the SCM and the SCM-FT on the snapshots of the MMOG dataset.

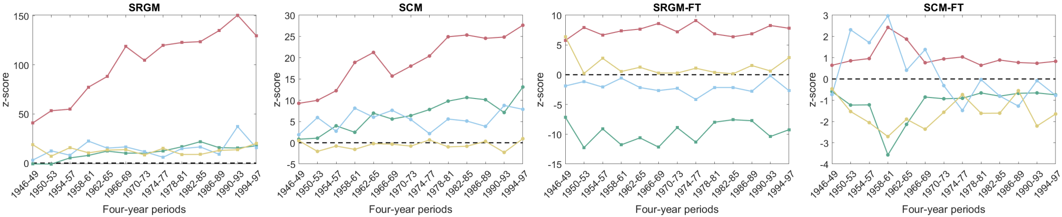
For further information about the numerical methods to solve the ERGMs, see N. Vallarano, M. Bruno, E. Marchese, G. Trapani, F. Saracco, G. Cimini, M. Zanon, T. Squartini, *Sci. Rep.* (2021) and the toolbox **NEMtropy** on Github *(https://pypi.org/project/NEMtropy/)*.

	SCM			SCM-FT		
	MAE	MRE	Time (s)	MAE	MRE	Time (s)
N.G.H. Tribes	$\simeq 3.2 \cdot 10^{-2}$	$\simeq 1.1 \cdot 10^{-2}$	$\simeq 0.005$	$\simeq 2.7 \cdot 10^{-2}$	$\simeq 1.1 \cdot 10^{-2}$	$\simeq 0.0004$
Monastery	$\simeq 2.4 \cdot 10^{-2}$	$\simeq 1.4 \cdot 10^{-2}$	$\simeq 0.006$	$\simeq 1.2 \cdot 10^{-2}$	$\simeq 9.3 \cdot 10^{-3}$	$\simeq 0.006$
Senate US	$\simeq 1.0 \cdot 10^{-1}$	$\simeq 1.1 \cdot 10^{-2}$	$\simeq 0.06$	$\simeq 5.9 \cdot 10^{-2}$	$\simeq 1.0 \cdot 10^{-2}$	$\simeq 0.12$
EGFR	$\simeq 5.8 \cdot 10^{-2}$	$\simeq 2.1 \cdot 10^{-2}$	$\simeq 0.54$	$\simeq 3.8 \cdot 10^{-2}$	$\simeq 1.1 \cdot 10^{-2}$	$\simeq 0.35$
Macrophage	$\simeq 8.1 \cdot 10^{-2}$	$\simeq 2.1 \cdot 10^{-2}$	$\simeq 2$	$\simeq 2.6 \cdot 10^{-2}$	$\simeq 1.5 \cdot 10^{-2}$	$\simeq 1.8$
E. Coli	$\simeq 1.9 \cdot 10^{-1}$	$\simeq 2.2 \cdot 10^{-2}$	$\simeq 13$	$\simeq 1.3 \cdot 10^{-2}$	$\simeq 1.2 \cdot 10^{-2}$	$\simeq 10$
Bitcoin Alpha	$\simeq 2.1 \cdot 10^{-1}$	$\simeq 3.1 \cdot 10^{-2}$	$\simeq 76$	$\simeq 5.5 \cdot 10^{-2}$	$\simeq 2.2 \cdot 10^{-2}$	$\simeq 138$
Bitcoin OTC	$\simeq 2.5 \cdot 10^{-1}$	$\simeq 4.0 \cdot 10^{-2}$	$\simeq 24$	$\simeq 1.1 \cdot 10^{-1}$	$\simeq 3.5 \cdot 10^{-2}$	$\simeq 267$

Table: Performance of the fixed-point algorithm to solve the systems of equations defining the SCM and the SCM-FT on a bunch of real-world networks.

For further information about the numerical methods to solve the ERGMs, see N. Vallarano, M. Bruno, E. Marchese, G. Trapani, F. Saracco, G. Cimini, M. Zanon, T. Squartini, *Sci. Rep.* (2021) and the toolbox **NEMtropy** on Github *(https://pypi.org/project/NEMtropy/)*.

$$z_\triangle = \frac{N_\triangle(\mathbf{A}^*) - \langle N_\triangle \rangle}{\sigma[N_\triangle]} \rightarrow \begin{cases} > 0 & \Longrightarrow \text{tendency of the } \triangle \text{ to be over-represented} \\ < 0 & \Longrightarrow \text{tendency of the } \triangle \text{ to be under-represented} \end{cases} \text{ in the data with respect to the null model.}$$



In *Gallo et al.* (2024) a closed form expressions for the z-scores is provided.

P. Doreian, A. Mrvar, *Journal of Social Structure* (2015).

A. G., D. Garlaschelli, R. Lambiotte, F. Saracco, T. Squartini, *Communications Physics* (2024).

$$Q = \sum_{i=1}^{N} \sum_{\substack{j=1\\(j>i)}}^{N} [a_{ij}^{*} - \langle a_{ij} \rangle] \delta_{c_i c_j} = \sum_{i=1}^{N} \sum_{\substack{j=1\\(j>i)}}^{N} [(a_{ij}^{+})^{*} - (a_{ij}^{-})^{*} - \langle a_{ij}^{+} \rangle + \langle a_{ij}^{-} \rangle] \delta_{c_i c_j}$$

$$= \sum_{i=1}^{N} \sum_{\substack{j=1\\(j>i)}}^{N} (a_{ij}^{+})^{*} \delta_{c_i c_j} - \sum_{i=1}^{N} \sum_{\substack{j=1\\(j>i)}}^{N} (a_{ij}^{-})^{*} \delta_{c_i c_j} - \sum_{i=1}^{N} \sum_{\substack{j=1\\(j>i)}}^{N} \langle a_{ij}^{+} \rangle \delta_{c_i c_j} + \sum_{i=1}^{N} \sum_{\substack{j=1\\(j>i)}}^{N} \langle a_{ij}^{-} \rangle \delta_{c_i c_j}$$

$$= \sum_{i=1}^{N} \sum_{\substack{j=1\\(j>i)}}^{N} (a_{ij}^{+})^{*} \delta_{c_i c_j} - \sum_{i=1}^{N} \sum_{\substack{j=1\\(j>i)}}^{N} (a_{ij}^{-})^{*} \delta_{c_i c_j} - \sum_{i=1}^{N} \sum_{\substack{j=1\\(j>i)}}^{N} p_{ij}^{+} \delta_{c_i c_j} + \sum_{i=1}^{N} \sum_{\substack{j=1\\(j>i)}}^{N} p_{ij}^{-} \delta_{c_i c_j}$$

$$= L_{\bullet}^{+} - L_{\bullet}^{-} - \langle L_{\bullet}^{+} \rangle + \langle L_{\bullet}^{-} \rangle = L^{+} - L_{\circ}^{+} - L_{\bullet}^{-} - \langle L^{+} - L_{\circ}^{+} \rangle + \langle L_{\bullet}^{-} \rangle$$

$$= -(L_{\circ}^{+} + L_{\bullet}^{-}) + \langle L_{\circ}^{+} + L_{\bullet}^{-} \rangle + L^{+} - \langle L^{+} \rangle = -[(L_{\circ}^{+} + L_{\bullet}^{-}) - \langle L_{\circ}^{+} + L_{\bullet}^{-} \rangle] + L^{+} - \langle L^{+} \rangle$$

where

$$L_{\bullet}^{+} = \sum_{i=1}^{N} \sum_{\substack{j=1\\(j>i)}}^{N} (a_{ij}^{+})^{*} \delta_{c_i c_j} = \sum_{i=1}^{N} \sum_{\substack{j=1\\(j>i)}}^{N} (a_{ij}^{+})^{*} - \sum_{i=1}^{N} \sum_{\substack{j=1\\(j>i)}}^{N} (a_{ij}^{+})^{*} (1 - \delta_{c_i c_j}) = L^{+} - L_{\circ}^{+}$$