Interplay between *k*-core and communities in networks

Alessio Cardillo (@a_cardillo)

Dept. Computer Science & Mathematics – Universitat Rovira i Virgili
Tarragona (Spain)

CompleNet Live 2021
Wednesday, 26th May 2021

University of BRISTOL

UNIVERSITAT ROVIRA i VIRGILI


SODA
80¢

nature physics

PERSPECTIVE

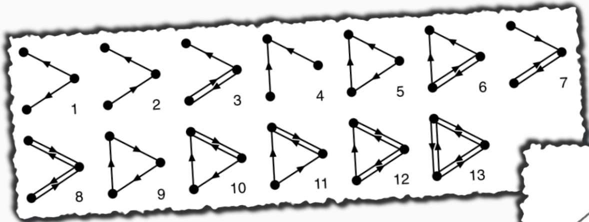
https://doi.org/10.1038/s41567-019-0459-y

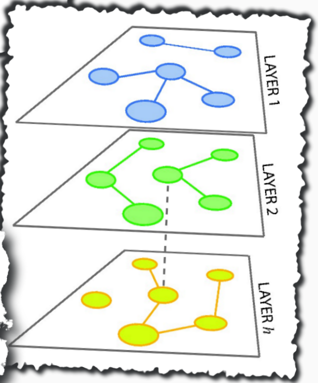
From networks to optimal higher-order models of complex systems

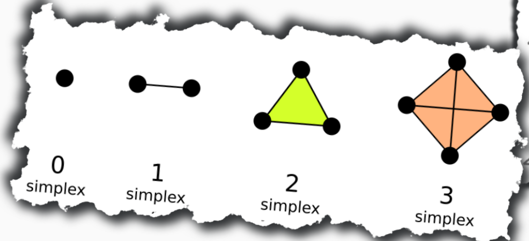
Renaud Lambiotte[1], Martin Rosvall[2] and Ingo Scholtes[3*]

Rich data are revealing that complex dependencies between the nodes of a network may not be captured by models based on pairwise interactions. Higher-order network models go beyond these limitations, offering new perspectives for understanding complex systems.

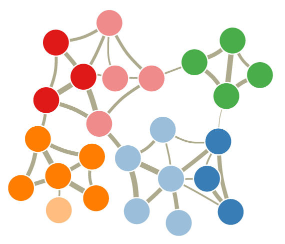
- R. Lambiotte, M. Rosvall, & I. Scholtes, *Nature Physics*, **15**, 313–320 (2019).


1
2
3
4
5
6
7
8
9
10
11
12
13

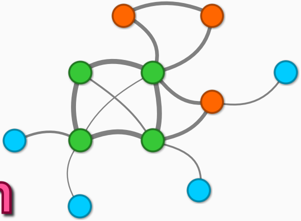

LAYER 1
LAYER 2
LAYER h


0 simplex
1 simplex
2 simplex
3 simplex

High-order structures



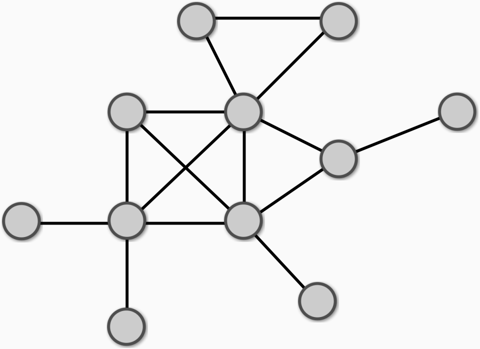
Communities



k-core decomposition

- S. Fortunato, & D. Hric, *Physics Reports*, **659**, 1–44, (2016).
- Y.-X. Kong, G.-Y. Shi, R.-J. Wu, & Y.-C. Zhang, *Physics Reports*, **832**, 1–32, (2019).

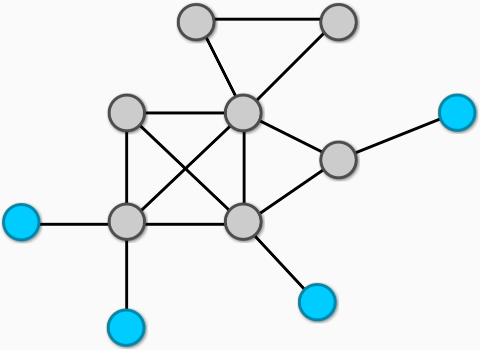
The *k*-core in a nutshell



- S. B. Seidman, Social Networks, **5**, 269–287, (1983).
- M. Kitsak *et al.*, Nature Physics, **6**, 888–893, (2010).

The *k*-core in a nutshell

k-shell 1 $(k_\mathrm{s} = 1)$

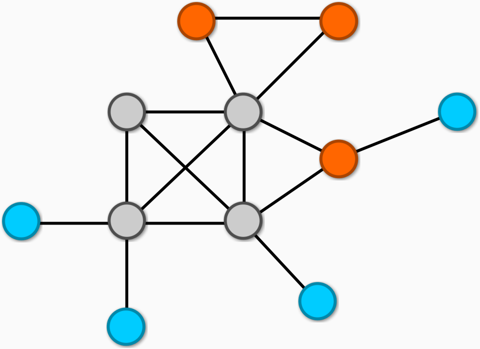


- S. B. Seidman, Social Networks, **5**, 269–287, (1983).
- M. Kitsak *et al.*, Nature Physics, **6**, 888–893, (2010).

The k-core in a nutshell

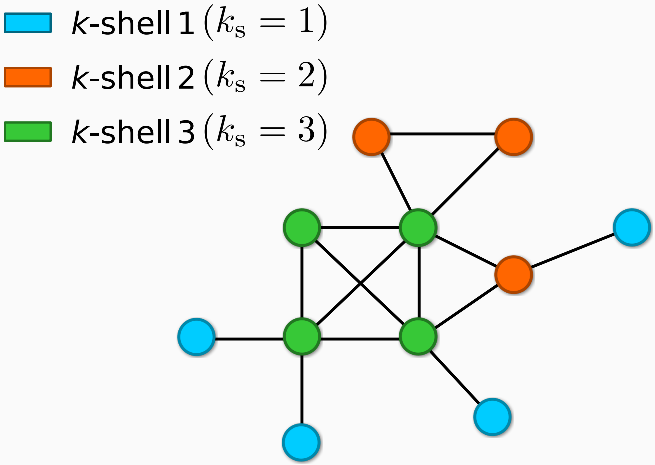
k-shell 1 $(k_{\mathrm{s}} = 1)$

k-shell 2 $(k_{\mathrm{s}} = 2)$



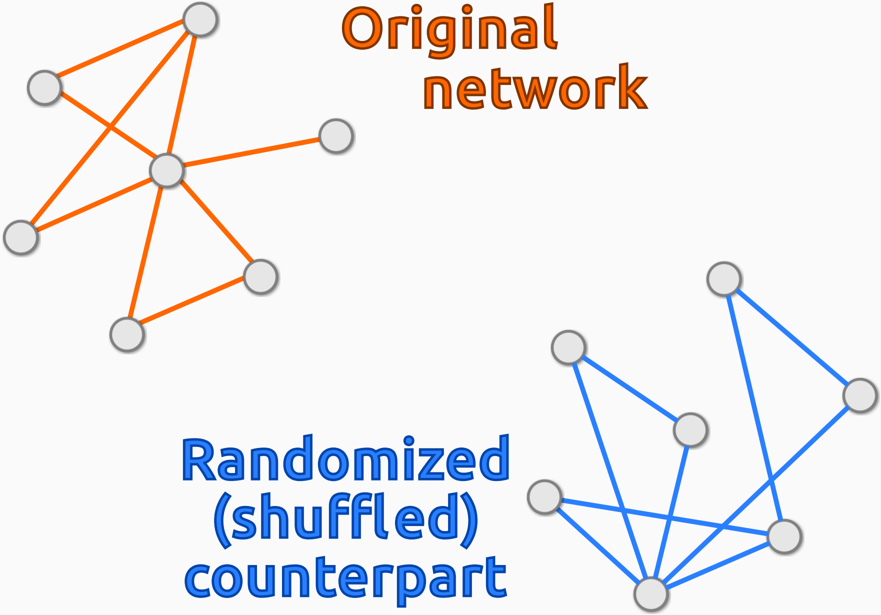
- S. B. Seidman, Social Networks, **5**, 269–287, (1983).
- M. Kitsak *et al.*, Nature Physics, **6**, 888–893, (2010).

The k-core in a nutshell



- S. B. Seidman, Social Networks, **5**, 269–287, (1983).
- M. Kitsak *et al.*, Nature Physics, **6**, 888–893, (2010).

Are the *k-core* and the communities two faces of the same coin?


Original
network
Randomized
(shuffled)
counterpart

Data

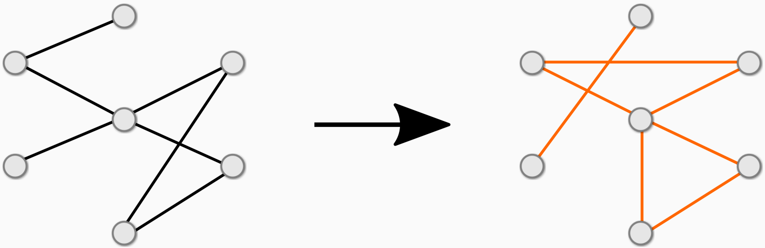
SYSTEM'S TYPE	No.
Social network	9
Web-blogs	1
Scientific collaborations	2
Email	1
Linguistics	1
Air transportation	1
TOTAL	15

Results

Shuffling method 1: `deg`

Definition

The `deg` method returns a network where the **degree sequence** is **preserved**.



- B. K. Fosdick, D. B. Larremore, J. Nishimura, and J. Ugander, SIAM Review, **60**, 315–355, (2018).

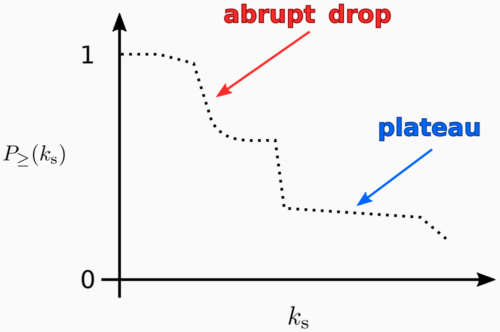
Shuffling method 1: deg

Survival probability

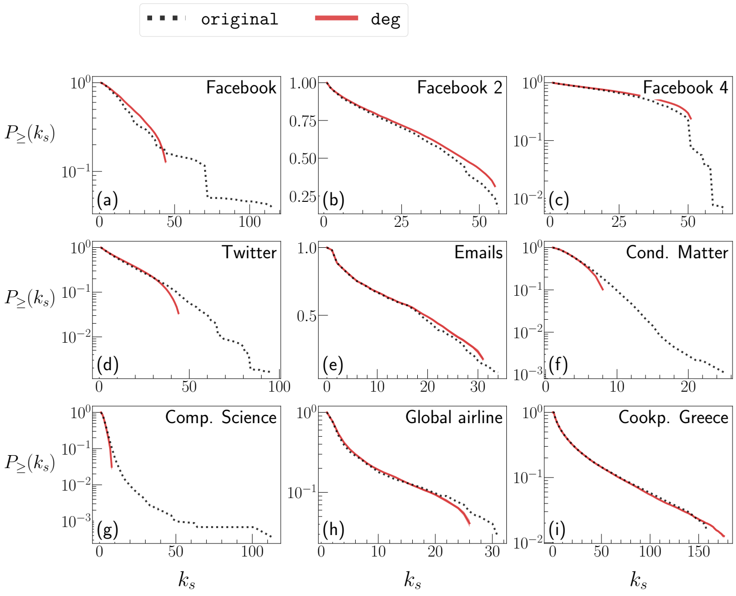
$$P_{\geq}(k_s) = \sum_{x=k_s}^{\max(k_s)} p(x)$$

with

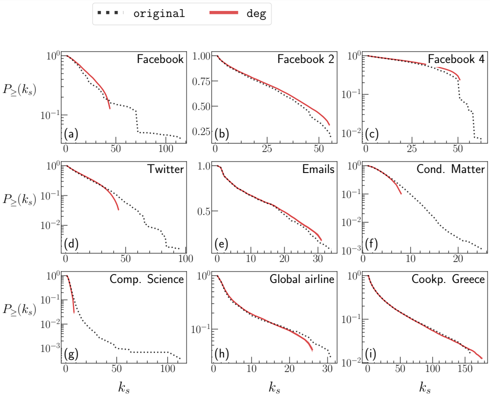
$$p(x) = \frac{N'(k_s = x)}{N} .$$



Shuffling method 1: deg



Shuffling method 1: deg



The degree alone **does not** reproduce the properties of the *k*-shell decomposition!

- I. Alvarez-Hamelin *et al.*, Networks & Heterogeneous Media, **3**, 371–393, (2008).

Shuffling methods 2 & 3: `commA` and `commB`

Shuffling methods 2 & 3: `commA` and `commB`

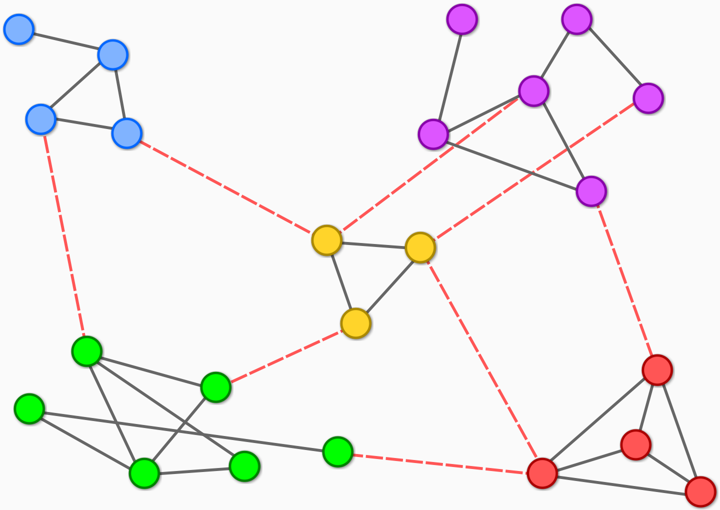
Definition

Beside the degree, the `commA` and `commB` shuffling preserve also the nodes' **community membership**.

Note

As the community membership **depends on the algorithm used**, we find the communities using both the **Louvain** (`Lvn`) and the **Stochastic Block Model** (`SBM`).

Shuffling methods 2 & 3: `commA` and `commB`



Shuffling methods 2 & 3: `commA` and `commB`

The `commA` method

It preserves, **for each community**, the total number of edges within each community (**intra-**), and between each pair of communities (**inter-**).

The `commB` method

It preserves, **for each node**, the numbers of **intra-** and **inter-** community edges.

Shuffling methods 2 & 3: `commA` and `commB`

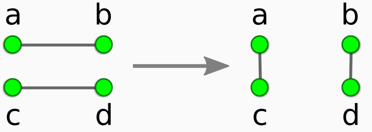
The `commA` method

It preserves, **for each community**, the total number of edges within each community (**intra-**), and between each pair of communities (**inter-**).

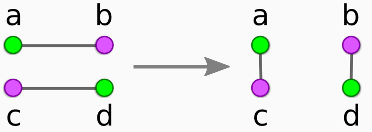
The `commB` method

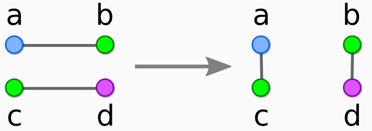
It preserves, **for each node**, the numbers of **intra-** and **inter-** community edges.

intra-community

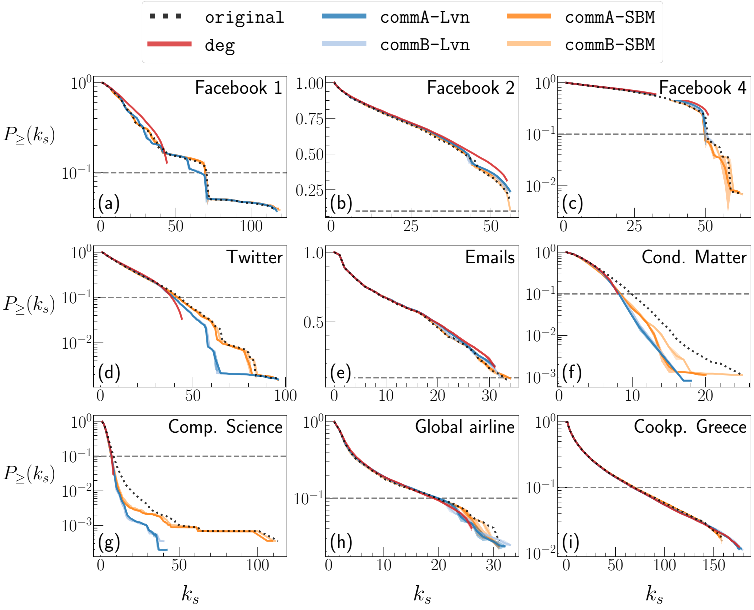


inter-community





Survival distribution



Comparing k-shell decompositions

For the whole network, we compute:

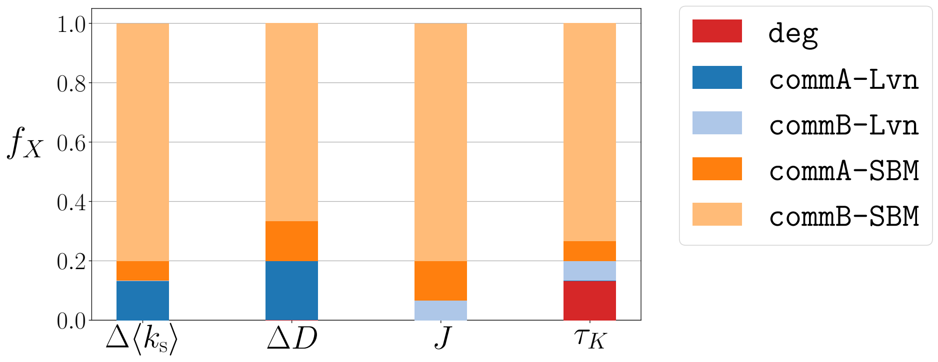
- **Average k_s index**: $\langle k_s \rangle$.
- **Degeneracy**: $D = \max(k_s)$.

Moreover, for the innermost k-shells only, we compute:

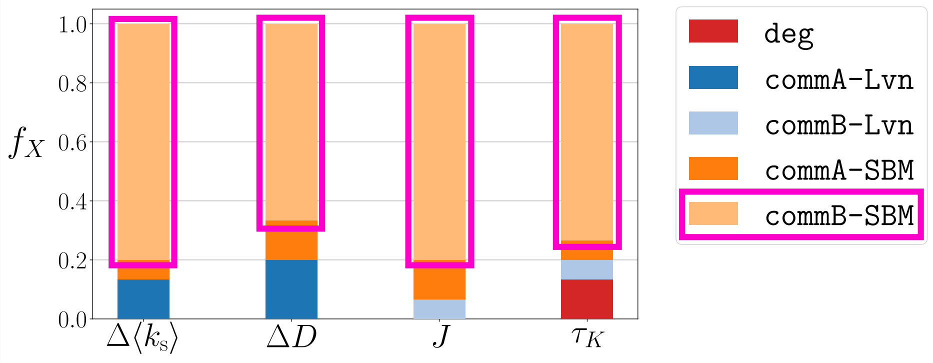
- **Jaccard score**: $J(\mathcal{X}, \mathcal{Y}) \in [0, 1]$.
- **Generalized Kendall's**: $\tau_K \in [-1, 1]$ and quantifies the inconsistency between ranked (k-shell index) lists.

- R. Fagin, R. Kumar, & D. Sivakumar, SIAM Journal on Discrete Mathematics, **17**, 134–160, (2003).

Comparing k-shell decompositions



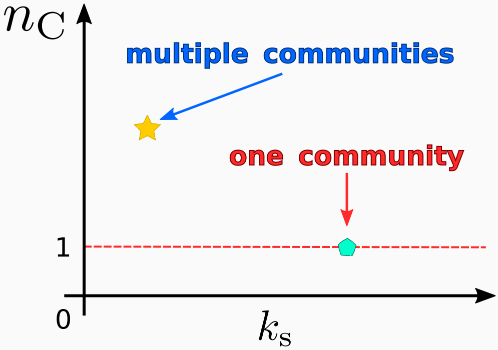
Comparing k-shell decompositions



Overlap between communities & *k*-core

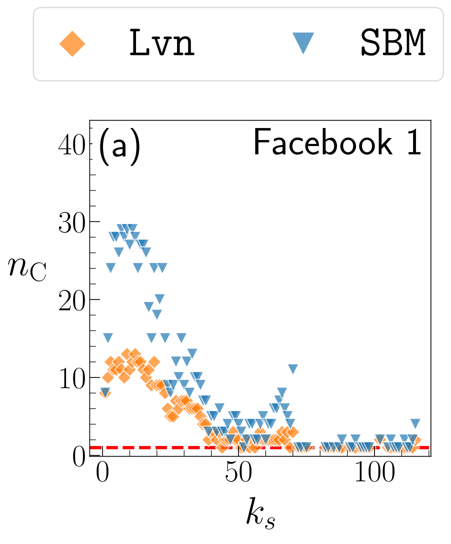
To what extent the *k*-core and the communities **overlap**?

Overlap between communities & *k*-core



We count the **number of distinct communities**, n_C, to which the nodes of a *k*-shell belong to.

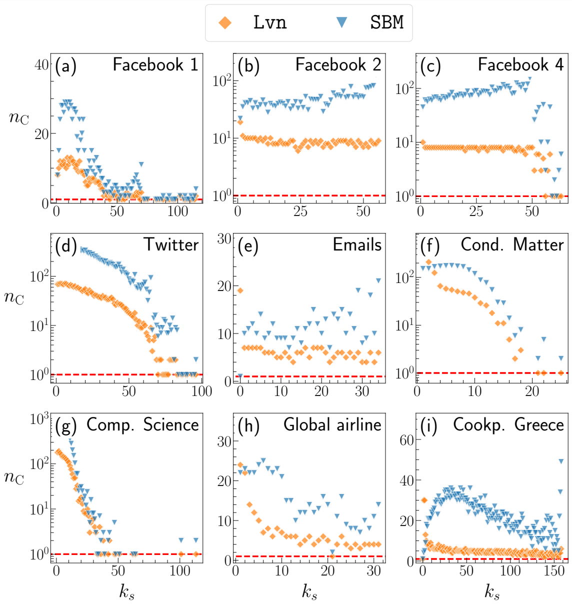
Overlap between communities & *k*-core



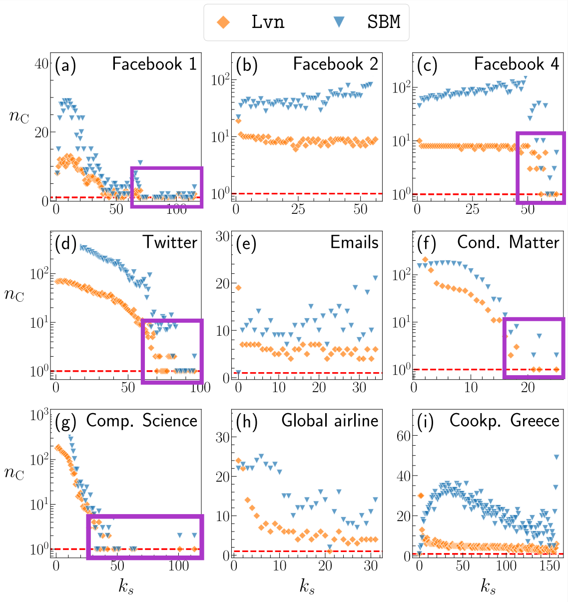
For $k_s \geq \tilde{k}_s$, the nodes of the *k*-core **belong to the same community**.

- S. Osat, F. Radicchi, & F. Papadopoulos, Physical Review Research, **2**, 023176, (2020).

Overlap between communities & k-core

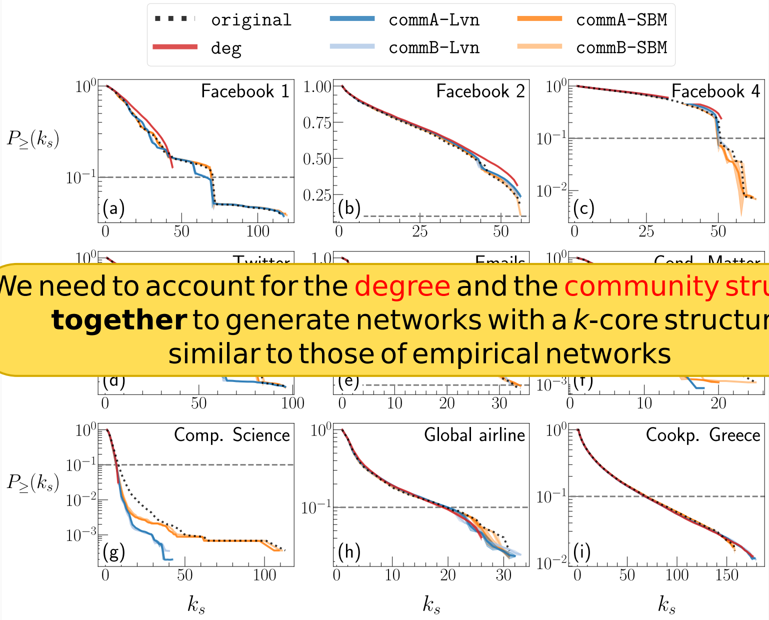


Overlap between communities & k-core



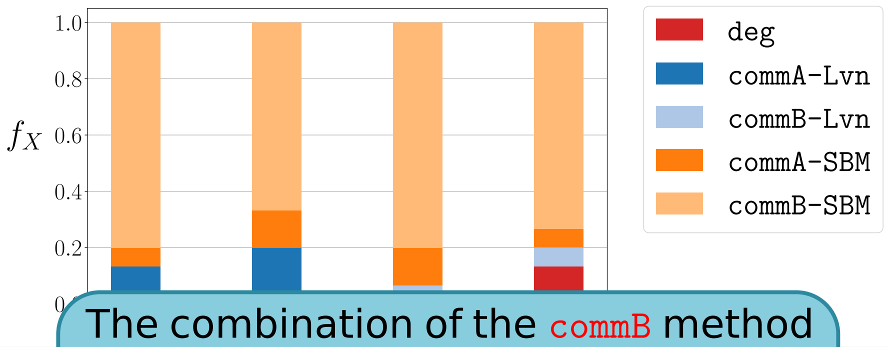
Summing up . . .

Take home messages



We need to account for the degree and the community structure **together** to generate networks with a *k*-core structure similar to those of empirical networks

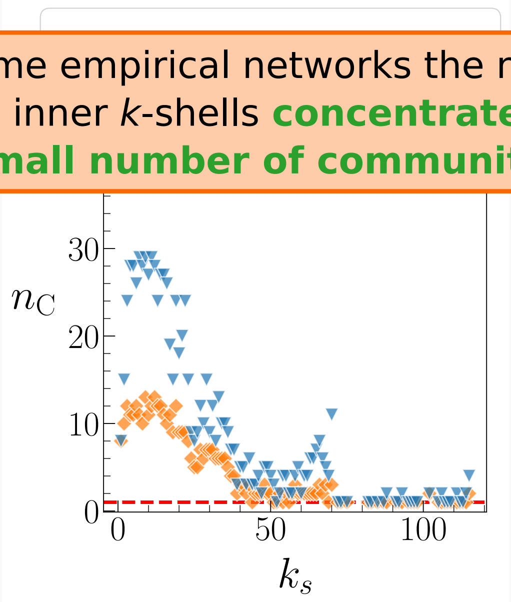
Take home messages



The combination of the `commB` method and the `SBM` communities reproduces best the *k*-core's empirical features

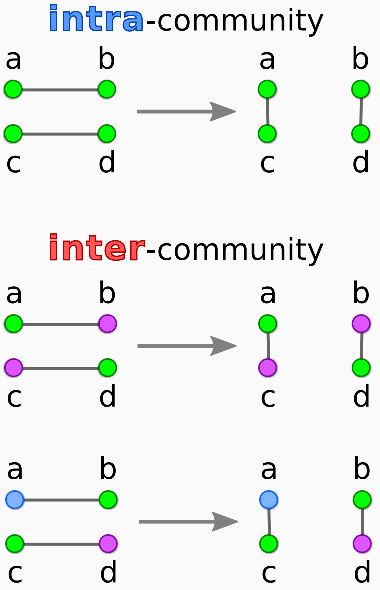
Take home messages

In some empirical networks the nodes of the inner *k*-shells **concentrate into a small number of communities**



Take home messages

The "*community aware*" rewiring method could be used to asses whether a property of a network **is a direct expression of its community structure or not**



Acknowledgements

Irene Malvestio

(Polytech. of Milan / Univ. of Bristol)

Naoki Masuda

(SUNY Buffalo / Univ. of Bristol)

scientific reports

Explore content ⌄ Journal information ⌄ Publish with us ⌄

nature › scientific reports › articles › article

Article | Open Access | Published: 07 September 2020

Interplay between k-core and community structure in complex networks

Irene Malvestio, Alessio Cardillo & Naoki Masuda

Scientific Reports **10**, Article number: 14702 (2020) | Cite this article

10.1038/s41598-020-71426-8

alessio.cardillo@urv.cat

http://www.bifi.es/~cardillo/

@a_cardillo

Extra content

Datasets – topological characteristics

Data set	N	L	$\langle k \rangle$	k_{max}	$\langle k_s \rangle$	D	N_c^{Lvn}	Q^{Lvn}	N_c^{SBM}	Q^{SBM}
Facebook 1	4039	88234	43.691	1045	26.880	115	16	0.835	62	0.551
Facebook 2	6386	217662	68.168	930	35.712	56	19	0.419	198	0.158
Facebook 3	2235	90954	81.391	467	44.508	63	8	0.436	87	0.139
Facebook 4	11247	351358	62.480	415	32.413	63	10	0.438	274	0.193
Facebook 5	27737	1034802	74.615	2555	38.681	81	18	0.470	547	0.172
Twitter	81306	1342296	33.018	3383	17.762	96	73	0.808	510	0.511
Web-blogs	1490	16715	22.436	351	12.154	36	275	0.426	17	0.076
Emails	1005	16064	31.968	345	17.063	34	26	0.410	33	0.232
Cond. Matter	23133	93439	8.078	279	4.900	25	619	0.730	203	0.633
Comp. Science	317080	1049866	6.622	343	4.215	113	209	0.822	676	0.726
Global airline	3376	19179	11.362	248	6.123	31	26	0.665	40	0.311
Words	146005	656999	9.000	1008	5.289	31	378	0.759	548	0.583
Cookpad Greece	32235	745178	46.234	8196	23.709	158	40	0.166	76	0.020
Cookpad Spain	122158	1749751	28.647	12637	14.547	162	262	0.270	90	0.035
Cookpad UK	13758	47525	6.909	1880	3.558	33	199	0.350	8	0.114

Indicators of comparison

$$\langle k_s \rangle = \frac{1}{N} \sum_i k_s(i)$$

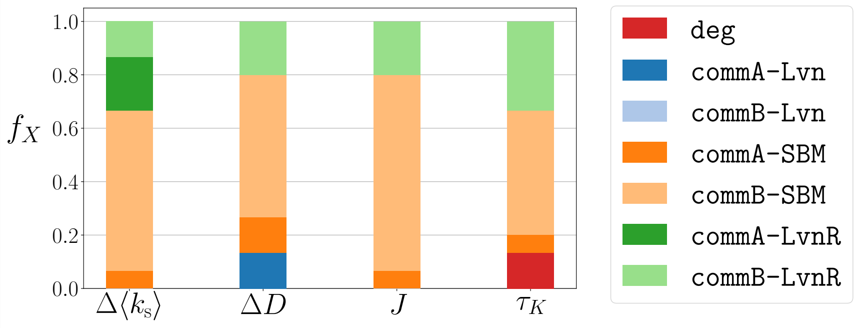
$$D = \max_G k_s$$

$$\Delta X = \frac{|X_G - X_{G'}|}{X_G}, \quad \text{where} \quad X \in \left\{ \langle k_s \rangle, D \right\}$$

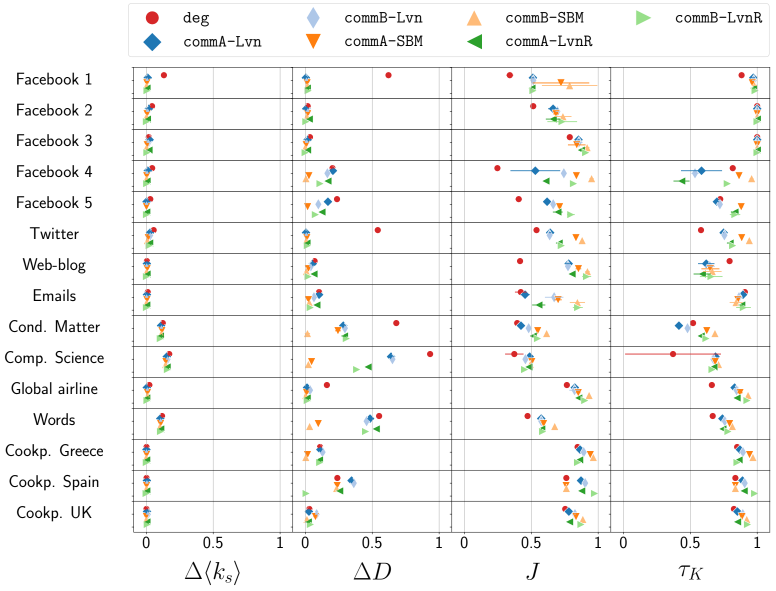
$$J(\mathcal{A}, \mathcal{B}) = \frac{|\mathcal{A} \cap \mathcal{B}|}{|\mathcal{A} \cup \mathcal{B}|}$$

$$\tau_K(\mathcal{X}, \mathcal{Y}) = 1 - \frac{1}{m_A m_B} \sum_{z_1, z_2 \in \mathcal{A} \cup \mathcal{B}} K_{z_1, z_2}(\mathcal{X}, \mathcal{Y})$$

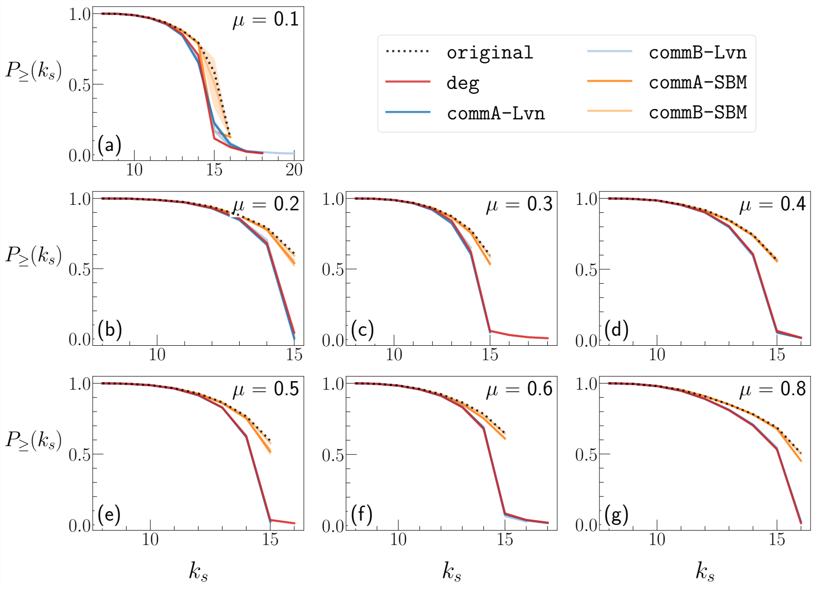
Louvain & resolution limit



Louvain & resolution limit



LFR benchmark



LFR benchmark