Principled Network Skeletonization via Minimal Homology Bases

Marco Guerra[1], Alessandro De Gregorio[1], Giovanni Petri[2], Francesco Vaccarino[1,2]

[1] Department of Mathematical Sciences, Politecnico di Torino - [2] ISI Foundation, Torino

Homological Scaffold

- First introduced in [1] as a method for network skeletonization.
- Defined for a non-negatively weighted network as the weighted graph, on the same set of nodes as the original, composed of all the cycles that are generators of the 1-dimensional persistent homology of the associated clique complex, filtered by edge weight.
- Weighted by the number of times an edge belongs to a persistent cycle.

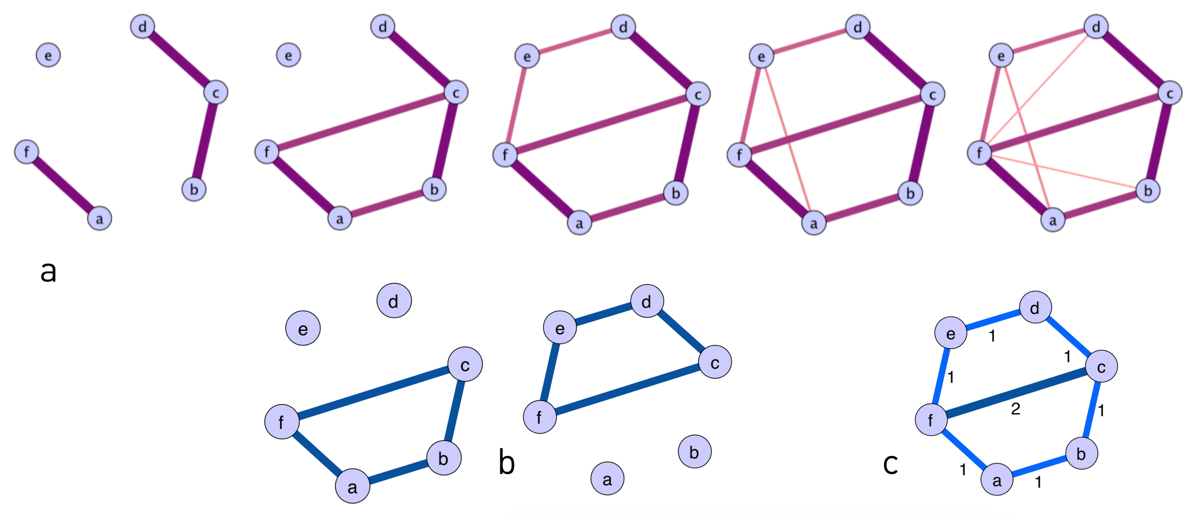

Fig.1 (Courtesy of [3]) A filtration of simplicial complexes (a), the generators of the persistent homology (b), and the resulting homological scaffold (c).

Addressed Problem

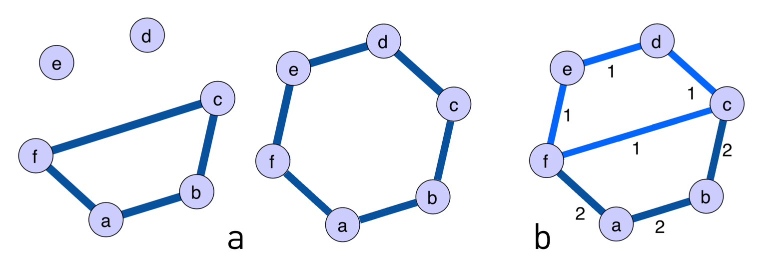

Fig.2 A different choice of generators (a) for the filtration in Fig.1 (a), leading to a different scaffold (b).

- The scaffold depends on a choice of representative cycles.
- Different choices are possible with the same filtration, and a choice of generators as in Fig.2 (a) would lead to the scaffold in Fig.2 (b) from the same input (weights differ).

The scaffold allows one to relate the mesoscopic (homological) structure of a network to its local constituents, highlighting which links are crucial to the global pattern.

On the other hand, the arbitrariness introduced by the choice of representative cycles makes this relation inaccurate.

Goal

To obtain a "canonical" homology basis, which allows us to define a scaffold in a principled, unique manner.

Method

Minimal Homology Bases

Recent work by Dey et al. [4] allows one to compute in polynomial time, for the 1-dimensional case, the minimal representatives of the homology basis cycles, i.e. a set

$$\{b_i\} = \underset{\text{Span}\{[b_i]\}=H_1}{\text{argmin}} \sum_i \mu(b_i)$$

where $\mu(b_i)$ is the length of cycle b_i, intended as the sum of its weights.

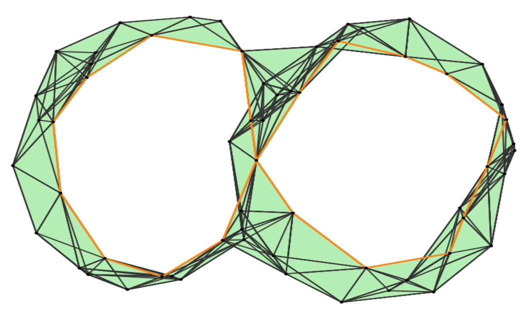

Fig.3 The shortest representatives of two generators of H_1.

Minimal Scaffold

- We define the minimal homological scaffold of a weighted network as the weighted graph on the same set of nodes, and having as edges those belonging to those cycles that, at some filtration step, are the minimal representative of a homology class.
- Weight is assigned for each time an edge belongs to a minimal cycle.

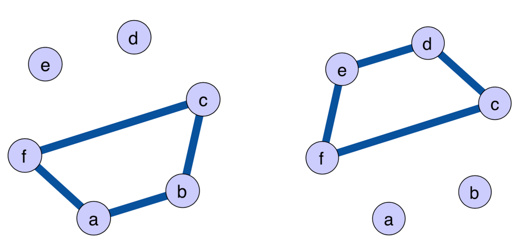

Fig.4 Given the filtration in Fig.1 (a), this is the only possible choice of **minimal** generators, leading necessarily to the scaffold in Fig.1 (c).

Computational Complexity

- Dey's algorithm has a worst-case complexity of order 9 in the number of points ([4]).
- Its computation is therefore largely more demanding than the generic case, for which superlinear algorithms are available.

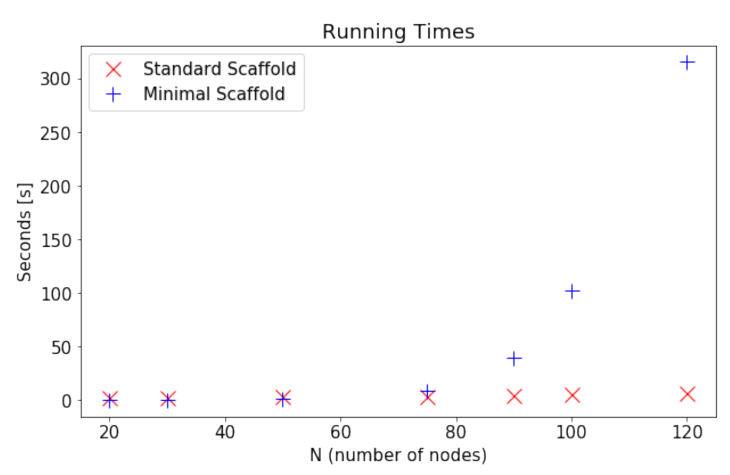

Fig.5 Scaffolds computation times for a family of Watts-Strogatz random networks [2] on a laptop computer (Params k=N/10, p=0.025).

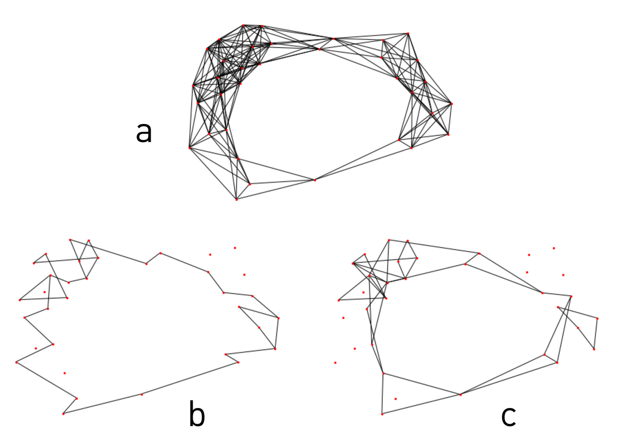

Fig.6 The standard (b) and minimal (c) scaffolds of a metric graph (a), whose weights are distances in the Euclidean plane.

Results and Comparison

The minimal scaffold is mathematically well-defined, unique up to few pathological cases, and reproduces tightly the geometry of the input.

At the same time, its computation is largely more demanding than the standard one.

Testing the minimal scaffold on real-world data and random models:

- It can precisely identify nodes of crucial importance for the global structure.
- Comparison: can the standard scaffold approximate the minimal?

Rank	Relative Strength	Rank	Relative Strength
1	32.188274	3	21.751803
2	24.498243	4	18.181431

Fig.7 The top 4 neurons of C.Elegans by node strength in the minimal scaffold (ratio on average strength).

	Spearman		Pearson	
Metric	Corr	p-val	Corr	p-val
Node Degree	0.953148	3.1842e-155	0.975559	3.4463e-196
Node Strength	0.772330	4.3712e-60	0.700653	3.7250e-45
Deg Centrality	0.953148	3.1842e-155	0.975559	3.4463e-196
Btw Centrality	0.952098	7.7348e-154	0.986412	1.8813e-233
Clsn Centrality	0.921274	5.1143e-123	0.960413	8.7695e-166
Eig Centrality	0.880711	9.5943e-98	0.858564	1.3911e-87
Clustering Coeff	0.412889	1.1778e-13	0.358577	1.9337e-10
Edge Weights	0.226321	1.3586e-09	0.086226	0.0224

Fig.8 For several network metrics the standard scaffold reproduces the statistical properties of the minimal one in C.Elegans. Others are unreliable, due to the different construction mechanisms.

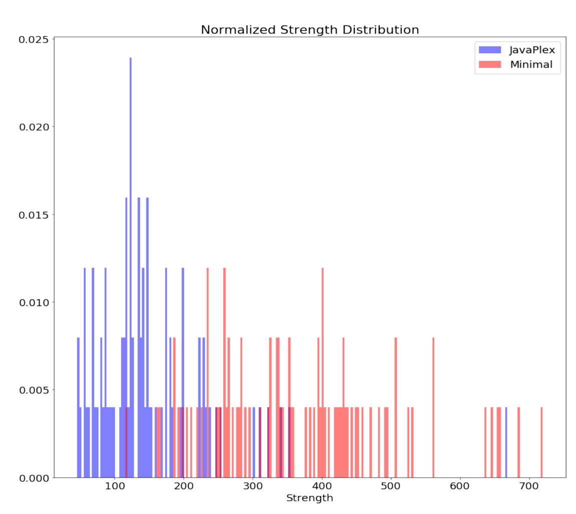

Fig.9 Node strength distribution of the scaffolds for the Watts-Strogatz random model (parameters as in [2]). The tight localization of cycles makes for a statistically larger fraction of node with high strength.

Conclusions and Future Directions

We have provided a new mathematically principled method for a topology-driven network skeletonization. We have observed that for many graph metrics the minimal scaffold has high correlation with the less computationally intensive, non-minimal scaffold which, therefore, is often a good enough approximation. We plan to further explore other approaches to canonicity in the choice of a homology basis.

References

[1] G. Petri, P. Expert, F. Turkheimer, R. Carhart-Harris, D. Nutt, P.J. Hellyer, F. Vaccarino: Homological scaffolds of brain functional networks. Journal of The Royal Society Interface (2014)

[2] A. Sizemore, C. Giusti, D.S. Bassett: Classification of weighted networks through mesoscale homological features. Journal of Complex Networks (2017)

[3] L.D. Lord, P. Expert, H.M. Fernandes, G. Petri, T.J. van Hartevelt, F. Vaccarino, G. Deco, F. Turkheimer, M.L. Kringelbach: Insights into Brain Architectures from the Homological Scaffolds of Functional Connectivity Networks. Frontiers in Systems Neuroscience (2016)

[4] T.K. Dey, T. Li, Y. Wang: Efficient algorithms for computing a minimal homology basis. Latin American Symposium on Theoretical Informatics - Springer (2018)

[5] C.J.A. Delfinado, H. Edelsbrunner: An incremental algorithm for betti numbers of simplicial complexes on the 3-sphere. Computer Aided Geometric Design (1995)

[6] H. Edelsbrunner, J. Harer: Computational Topology: An Introduction. Applied Mathematics, American Mathematical Society (2010)