



- Many networks describing complex systems are directed:
 the interactions between elements are not symmetric.
- Directed networks exhibit *trophic coherence q:* there exists a well-defined hierarchy of vertices.
- We will focus on the study of the following aspects related to q:
 - Looplessness
 - Non normality
 - Strong connectivity

We will use many types of dataset: from genes, metabolites, species, neurons, words, computers, and trading nation

Looplessness in networks is linked to trophic coherence

Feedback is a fundamental process in dynamical systems that occurs when the output of an element is coupled to its input.

Definitions:

Directed configuration ensemble:

set of directed graphs with a given number of nodes and degree sequence.



which is the subset of graphs from the directed configuration ensemble that satisfy the constraint that the proportion of neighbors connected to basal nodes is exactly kin Lb /L.

Coherence Ensemble:

ensemble of directed graphs that not only have given in- and out-degree distributions (as in the directed configuration ensemble) but also given trophic coherence

Consider the directed, unweighted graph given by N x N adjacency matrix A which has L edges:

Branching factor

$$lpha = rac{\langle k^{in} k^{out}
angle}{\langle k
angle}.$$



Basal node

one with in degree equal to zero

n has L edges: Inspiration:
Page rank equation → r = cMr +
(1-c)q

Trophic level
$$s_i = 1 + rac{1}{k_i^{in}} \sum_j a_{ij} \; s_j.$$

$$s_i = 1 \, (orall i \, ext{such that} \, k_i^{in} \, = 0)$$
 .

Looplessness in networks is linked to trophic coherence

Definitions:



Trophic difference

$$x_{ij} = s_i - s_j$$

Stdev. of distribution of Trophic difference

Trophic coherence

$$q = \sqrt{rac{1}{L} \sum_{ij} a_{ij} x_{ij}^2 - 1}.$$

Maximal coherence q = 0

Loop exponent

$$au = \ln\,lpha + rac{1}{2\,lpha^2} - rac{1}{2q^2},$$
 $au < exttt{0 loopless regime} \ au > exttt{0 loopful regime}$



Basal ensemble

Expectations for the trophic coherence:

$$\widetilde{q}' = \sqrt{rac{L}{L_B} - 1}$$

Expectation of branching factor

$$\widetilde{\alpha} = \frac{L - L_B}{N - B}$$



Coherence ensemble $\lambda_1 = e^{\tau}$.

$$\overline{\lambda_1} = e^{\tau}$$
.

Expected proportion of paths that are cycles

$$\overline{c}_{\nu} = \frac{\tilde{\alpha}}{L} \frac{\tilde{q}}{q} \exp \left[\frac{\nu}{2} \left(\frac{1}{\tilde{q}^2} - \frac{1}{q^2} \right) \right].$$

Looplessness in networks is linked to trophic coherence

Directed Acyclic Graphs

 $\tilde{n_{\nu}} \approx L\alpha^{\nu-1}$ Total number of paths

Probability that a graph randomly chosen from the coherence ensemble will have exactly m_{ν} cycles.

 $\Rightarrow p(m_{\nu}) = {\widetilde{n}_{\nu} \choose m} \bar{c}_{\nu}^{m_{\nu}} (1 - \bar{c}_{\nu})^{\widetilde{n}_{\nu} - m_{\nu}}$

Probability that a network from the coherence ensemble would have no directed cycles of length greater or equal to n:

$$P_n = \prod_{\nu=-n}^{\infty} p(m_{\nu} = 0).$$

Probability that a network drawn randomly from this ensemble would be acyclic:

$$P_{acyclic} = \prod_{
u=2}^{\infty} p\left(m_{
u} \ = 0
ight) \quad = \prod_{
u=2}^{\infty} \left\{1 - rac{\widetilde{lpha}}{L} rac{\widetilde{q}}{q} \mathrm{exp}\left[rac{
u}{2} \left(rac{1}{\widetilde{q}^2} - rac{1}{q^2}
ight)
ight]
ight\}^{Llpha^{
u}-1}.$$

Taking logarithms and considering graphs with sufficiently negative **T**:



Non Normality matrix

An N x N matrix A is said to be normal if its adjacency matrix commutes with its transpose or conversely it is non-normal if $AA^T \neq A^TA$

- A large deviation from normality to indicate a network with a well-defined directionality. This is also what occurs in trophically coherent networks.
- We can quantify a network being highly non-normal by using the Hermici's departure from normality:

$$D_{\mathrm{F}} = \sqrt{\|A\|_F^2 - \sum_i |\lambda_i|^2},$$

In order to compare matrices of different sizes, they use the normalised version:

$$d_{\mathrm{F}} = \frac{D_{\mathrm{F}}}{\|A\|_{\mathrm{F}}}.$$

A normal matrix will have $d_F =$ 0. and d_F is closer to 1 the more A departs from normality

Theorem. The expected deviation from normality, $\overline{d_F}$, for digraphs drawn from the coherence ensemble tends to 1 with increasing trophic coherence. That is,

$$\lim_{q\to 0} \overline{d_{\rm F}} = 1.$$

Furthermore.

$$\overline{d_{\mathrm{F}}} > \sqrt{1 - \frac{1}{\langle k \rangle}}$$

Coherent networks are highly non normal.

for digraphs in the τ < 0 regime, where $\langle k \rangle$ is the mean degree.

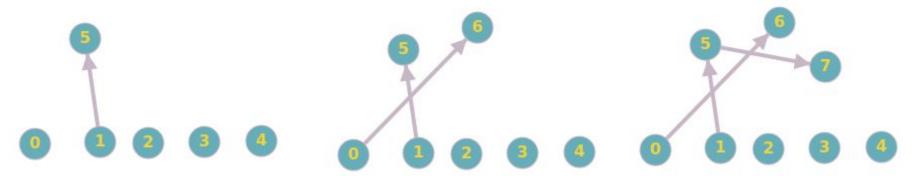
Generalised Preferential-Prey model

Alternative to PPM (networks generated always acyclic).

Possibility to tune trophic coherence.

Algorithm:

- Input: N total nodes; B basal nodes; <k> mean degree or L total number of edges; T temp. (related to incoherence q)
- Begin with B basal nodes in directed graph. Add (N-B) non-basal nodes one by one. When each node is added connect it with directed edge with direction from B to Non basal node at first addition while remaining nodes added have in-connections from both basal and non-basal nodes.. Meaning added node preys on nodes already present.



Generalised Preferential-Prey model

Algorithm:



) (N-B) edges are added \rightarrow (L-N+B) edges are to be added. Put all possible combinations of edges in a list.

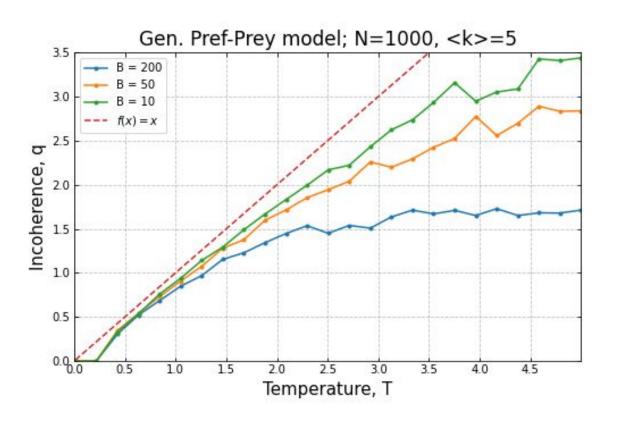


- Choose an edge randomly from list of all possible edges.
- If probability of connection from j to i, $\mathcal{U}(0,1) < P_{ji} = e^{\frac{-(x_{ij}-1)^2}{2T^2}}$ then add edge from j to i and remove that particular edge from list of all possible edges.
- $igcup_{igcep}$ The trophic level values 's' will be updated in the next iteration.

"Gaussian distribution of distances x, is found to be a good fit to empirical data on several kinds of networks."

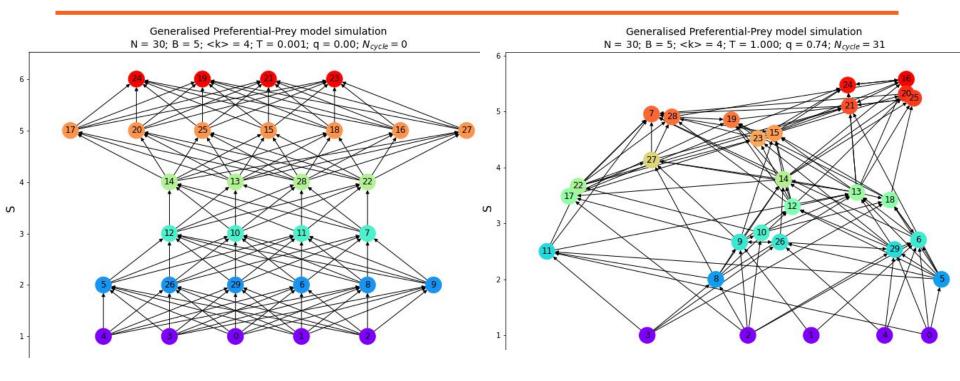


Monotonic relation btw. q and T: Results



- "Temperature" parameter
 T tunes the degree of
 trophic coherence.
- $T = 0 \rightarrow \text{maximally}$ coherent networks(q = 0).
- Incoherence q, increasing monotonically with T.
- Low T $\rightarrow q \approx T$
- High T → coherence saturates.

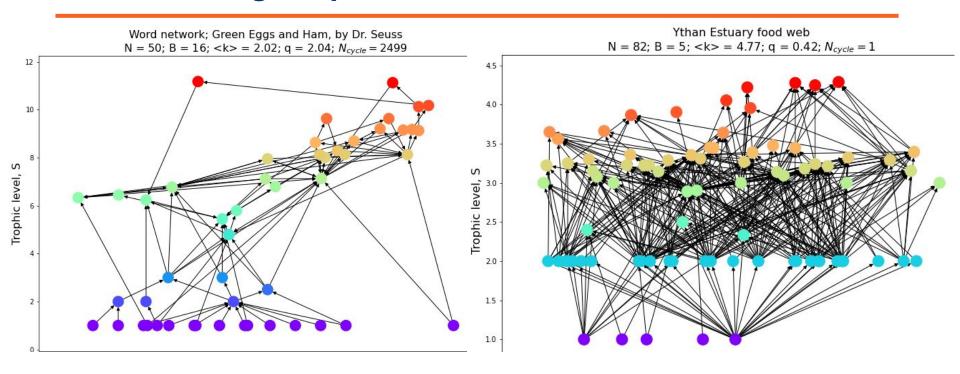
Simulation: Results



Color representation of node:
 VIBGYOR; node with higher S values
 → orange/red

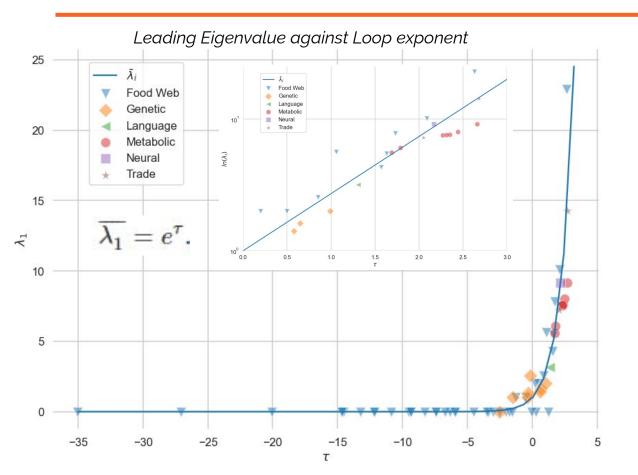
Increase in $T \rightarrow$ Increase in number of cycles; highly cyclic

Visualizing Trophic Levels of Real networks



- If word 1 appears immediately before word 2 in at least one sentence in the text, a directed edge (arrow) is placed from word 1 to word 2.
- The height of each word is proportional to its trophic level.

Looplessness: Results



The coherence ensemble expected value provides a good estimate of almost all of the empirical values.

$$\lambda(T < 0) = 0.22 \pm 0.54$$

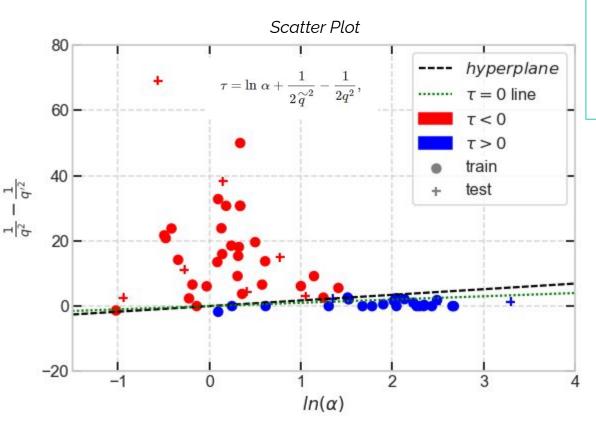
 $\lambda(T > 0) = 6.13 \pm 4.86$

- Self-loops(cannibalism in food webs) are removed.
- Total 61 networks:
- $T < 0 \rightarrow 36$;
- $T > 0 \rightarrow 25$

Looplessness: Results

	q/q'	s/s'	α/α'	τ	λ_1	τ < 0	Acyclic
Food	0.44 <u>+</u> 0.17	0.77 ± 0.32	1.01 ± 0.23	-6.23 ± 7.96	1.54 ± 4.04		31/42
Genetic	0.99 <u>+</u> 0.05	1.0 ± 0.01	1.19 ± 0.32	-0.34 ± 1.11	1.36 ± 0.71	5/8	1/8
Language	1.01 ± 0.0	1.16 ± 0.0	1.55 ± 0.0	1.31 ± 0.0	3.17 ± 0.0	0/1	0/1
Metabolic	1.81 ± 0.1	1.93 ± 0.12	3.98 <u>+</u> 0.96	2.22 ± 0.32	7.36 ± 1.11	0/7	0/7
Neural	0.42 + 0.0	0.39 ± 0.0	1.42 ± 0.0	2.17 ± 0.0	9.15 ± 0.0	0/1	0/1
Trade	1/1 ± 0.04	1.05 ± 0.08	1.19 ± 0.09	2.37 ± 0.32	10.78 ± 3.53	0/2	0/2
•	IIIICalitiv	gnificantly acoherent	Positive corr in and out-de			uracy of · liction: 8	

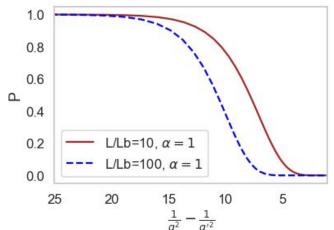
Looplesness: Results

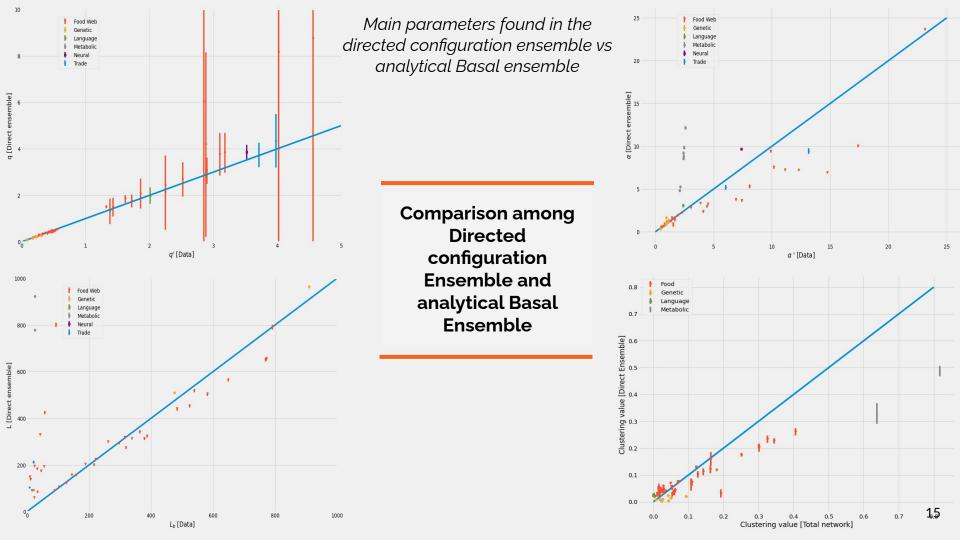


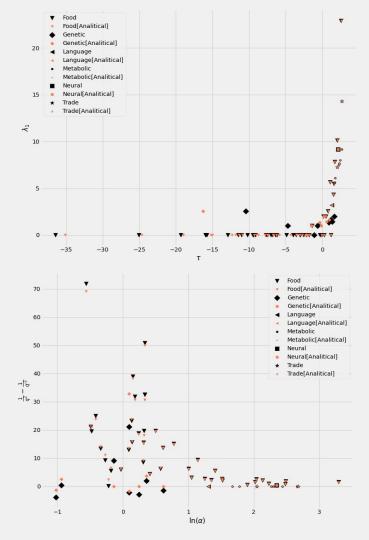
Classification also according to **Perceptron**:

Train-Test split: 85% and 15%
Train accuracy: 100%
Test accuracy: 100%

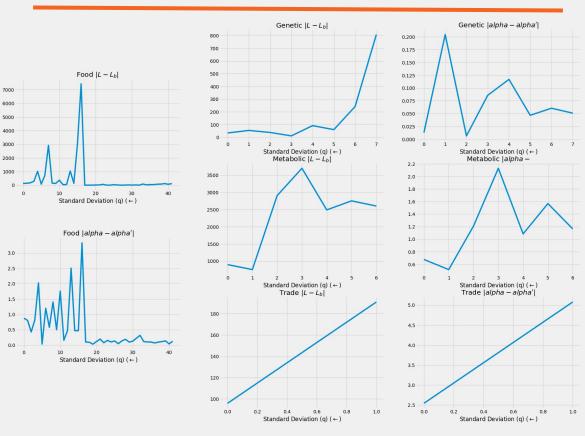
Probability of being acyclic, P(acyclic formula)



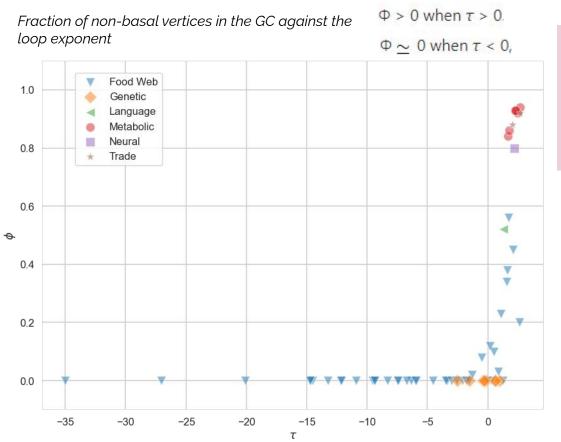




Comparison among ensembles



Strong Connectivity: Results



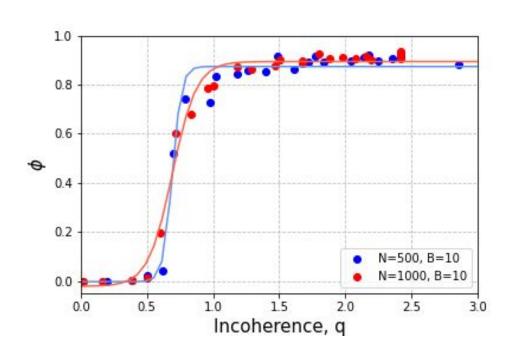
- Strongly connected reach any vertex from any other vertex along directed path.
- SCC necessarily contain long cycles.
- **T < 0 loopless regime** → SCC is vanishingly small

		<k></k>	φ
<u>:</u>	net_AG	2.38	0.93
apc	net_CE	2.44	0.92
Metabolic	net_CQ	2.05	0.84
_	net_CT	2.11	0.86
	net_MJ	2.40	0.93
	net_SC	2.54	0.94
	net_TH	2.43	0.93

	<k></k>	φ	
benguela	6.76	0.10	
Berwicktxt	3.12	0.00	
Blackrocktxt	4.36	0.00	
bridge	4.16	0.08	
Broadtxt	6.00	0.00	
broom	2.58	0.00	
canton	6.82	0.00	
Catlins	2.29	0.00	
cayman_islands	15.55	0.00	

Strong Connectivity: Results

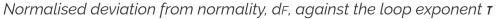
Fraction of non-basal vertices in the GC against q for the Predator-Prey model

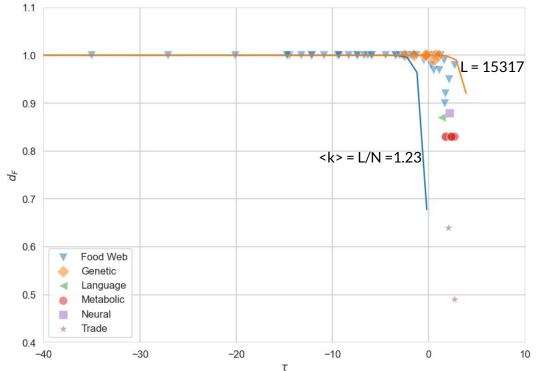


• Specific trophic coherence

 Gen-PPM displays continuous (i.e. second order) phase transition in strong connectivity with coherence

Non-normalized: Results



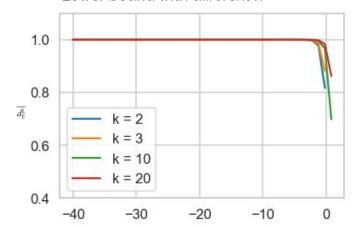


Upper and lower bounds on dF

$$\overline{d_{\rm F}^{\rm L}} = \sqrt{1 - \frac{N}{L}} {\rm e}^{2\tau},$$

$$\overline{d_{\rm F}^{\rm U}} = \sqrt{1 - \frac{1}{L} {\rm e}^{2\tau}}$$

Lower bound with different k



Real networks with small or negative au are indeed highly non-normal

What was done:

Reproduced the results from the paper exactly.

4

Visualised the real complex systems network dataset in accordance with S.

2

Implemented Generalised Preferential-Prey model; Investigated the relation between T and q.

5

Used Perceptron algorithm to re-classify networks with $\tau < 0$ and $\tau > 0$

3

Visualised the simulated network in accordance with trophic level S.

6

.Compare the results obtained from Directed Configuration Ensemble with the results obtained analytically.

Conclusion



Directed networks can belong to either of two regimes which is determine by the sign of τ , which is a function only of trophic coherence and the in- and out-degree sequences:

Loopful

Loopless



Edges are not strongly aligned with a global direction.



Edges are organized according to a global direction



Trophically incoherent



Networks are highly coherent



Small deviations



Non-normal



Large strongly connected components



Strongly connected components are vanishing



Leading eigenvalues of the datasets follow the relation $\lambda_1 = e^{\tau}$.



Metabolic networks are highly incoherent and have positive in-out degree correlation in comparison wrt other networks.



Increase in q, P(acyclic) decreases.



Trophic coherence is parameter that is related with the non-normality and the GC are related.



Gen-PPM displays continuous (i.e. second order) phase transition in strong connectivity with coherence,



In principle, Direct configuration ensemble lead to similar results but not the same

Future Work

No. of trials-1000 - obtain mean and stdev. for each point. - q vs T and Non basal fraction vs q plots.

Took 1.5 hours for each point; Better parallelizing methods could be implemented since even Dask parallel computing was taking long time.

Generalization to weighted matrices

Only Generalized Pref-prey model implemented; Various other models could also be checked->; cascade (CM), generalized niche (GNM), niche (NM), nested hierarchy(NHM), minimum potential niche (MPNM), and PPM.

Investigate more on implementing basal ensemble with its own class of constraints.

Study network properties such as Robustness or assortativity to better understand the trophic coherence.

Zebrafish Connectome - In a neural network, information enters through the **sensory neurons(Basal nodes)**, is processed through various kinds of **inter-neurons**, and finally reaches the **motor neurons**. Using trophic levels we can determine vertex function.

References and Github

- 1. Dataset of all the networks can be found at: https://www.samuel-johnson.org/data
- 2. Johnson S, Jones NS. Looplessness in networks is linked to trophic coherence. Proc Natl Acad Sci U S A. 2017 May 30;114(22):5618-5623. doi: 10.1073/pnas.1613786114. Epub 2017 May 16. PMID: 28512222; PMCID: PMC5465891.
 - S I Appendix: <u>https://www.pnas.org/content/pnas/suppl/2017/05/15/1613786114.DCSupplementall/pnas.1613786114.sapp.pdf</u>
- 3. Johnson, Samuel. (2019). **Digraphs are different: Why directionality matters in complex systems.** https://arxiv.org/abs/1908.07025v1
- 4. Klaise J, Johnson S. From neurons to epidemics: How trophic coherence affects spreading processes. Chaos. 2016 Jun;26(6):065310. doi: 10.1063/1.4953160. PMID: 27368799.
- 5. Python code, datasets, csv files and figures of this project can be found at: https://github.com/deliriarte/QLS-Project

be	
Ben	
Black	
Bi	
	Food Web
cayman_i	1 000 Web
ches	
CO	
Cowee	
Cowe	
Dempster	
Dempster	
Dempster	

N

В

Lb

s/s' 11 acy 29.0 2.0 9.0 6.76 0.69 4.56 0.15 0.17 0.50 4.59 0.97 0.10 enguela 0.66 2.00 wicktxt 77.0 35.0 216.0 3.12 0.18 0.33 0.53 1.02 1.06 -12.210.00 0.61 1.00 0.00 True rocktxt 86.0 49.0 338.0 4.36 0.19 0.33 0.56 1.01 1.27 -9.510.00 1.27 1.00 0.00 True bridge 25.0 8.0 34.0 4.16 0.53 1.43 0.37 0.68 0.69 -0.521.00 2.86 0.99 0.08 False 53.0 -20.10roadtxt 94.0 524.0 6.00 0.14 0.28 0.49 1.03 1.19 0.00 1.16 1.00 0.00 True 85.0 1.0 24.0 0.40 2.85 0.14 0.30 2.72 2.58 1.17 -2.080.00 1.00 0.00 True broom 54.0 649.0 0.15 0.27 0.57 1.01 1.22 -14.52 1.20 canton 102.0 6.82 0.00 1.00 0.00 True 14.0 89.0 0.20 1.00 0.62 Catlins 48.0 2.29 0.49 0.41 0.96 -10.900.00 1.00 0.00 True 242.0 10.0 339.0 15.55 0.77 3.18 0.24 0.29 0.51 1.22 0.00 7.52 1.00 0.00 True islands 31.0 5.0 23.0 2.16 1.38 0.32 0.89 -1.811.51 apeake 0.45 0.73 0.00 1.00 0.00 True 29.0 3.0 33.0 8.38 1.20 2.52 0.48 0.47 0.83 1.63 5.48 6.70 achella 0.90 0.38 False ta17txt 71.0 38.0 128.0 2.08 0.24 0.40 0.60 1.00 1.25 -5.940.00 0.76 1.00 0.00 True eta1txt 58.0 28.0 103.0 2.17 0.30 0.47 0.64 1.00 1.08 -3.390.00 0.83 1.00 0.00 True 46.0 0.21 0.37 0.57 1.01 -7.42rsAutxt 83.0 364.0 4.99 1.03 0.00 1.37 1.00 0.00 True rs Sptxt 93.0 50.0 484.0 5.78 0.13 0.33 0.37 1.02 1.11 -27.07 0.00 1.40 1.00 0.00 True 107.0 50.0 792.0 9.02 0.27 0.57 1.05 1.03 -3.510.00 3.14 0.00 rs Sutxt 0.47 1.00 True el verde 155.0 28.0 248.0 9.72 1.01 2.25 0.45 0.49 1.20 2.09 10.12 11.93 0.95 0.45 False 48.0 298.0 4.19 0.20 0.43 0.99 1.10 -9.350.00 1.65 1.00 Germantxt 0.470.00 True

a/a'

τ

Food Web

grass	61.0	8.0	16.0	1.59	0.40	2.25	0.18	0.42	0.63	-3.03	0.00	0.97	1.00	0.00	True
Healytxt	96.0	47.0	539.0	6.60	0.22	0.42	0.53	1.03	1.12	-6.34	0.00	2.17	1.00	0.00	True
Kyeburntxt	98.0	58.0	582.0	6.42	0.18	0.28	0.62	1.02	1.18	-9.39	0.00	1.38	1.00	0.00	True
LilKyeburntxt	78.0	42.0	315.0	4.81	0.23	0.44	0.53	1.01	1.10	-5.97	0.00	1.84	1.00	0.00	True
little_rock	92.0	12.0	93.0	10.70	0.67	3.10	0.22	0.25	0.76	1.06	5.66	8.45	0.97	0.23	False
Lough_Hyne	349.0	49.0	1410.0	14.62	0.60	1.62	0.37	0.59	0.63	0.85	2.56	7.70	1.00	0.03	False
Martinstxt	105.0	48.0	262.0	3.27	0.32	0.56	0.58	0.99	1.26	-2.56	0.00	1.78	1.00	0.00	True
Narrowdaletxt	71.0	28.0	127.0	2.17	0.23	0.46	0.50	0.98	1.14	-7.45	0.00	0.71	1.00	0.00	True
NorthColtxt	78.0	25.0	188.0	3.09	0.28	0.53	0.52	0.98	1.36	-4.52	0.00	1.36	1.00	0.00	True
Powdertxt	78.0	32.0	221.0	3.44	0.22	0.46	0.47	0.99	1.12	-8.32	0.00	1.15	1.00	0.00	True
reef	50.0	3.0	57.0	10.70	0.94	2.90	0.32	0.36	0.92	1.73	7.80	9.32	0.92	0.56	False
shelf	79.0	2.0	39.0	17.44	0.73	5.86	0.12	0.10	0.69	1.57	4.32	12.07	0.99	0.34	False
skipwith	25.0	1.0	11.0	7.56	0.61	4.02	0.15	0.16	0.61	0.20	2.00	4.54	0.98	0.12	False
stony	109.0	61.0	769.0	7.59	0.15	0.27	0.55	1.03	1.16	-14.66	0.00	1.40	1.00	0.00	True
Stonytxt	112.0	63.0	772.0	7.41	0.15	0.27	0.55	1.03	1.18	-14.72	0.00	1.40	1.00	0.00	True
st_marks	48.0	6.0	55.0	4.54	0.63	1.72	0.37	0.63	0.99	0.26	0.00	3.85	1.00	0.00	True
st_martin	42.0	6.0	46.0	4.88	0.59	1.86	0.32	0.54	0.79	-0.05	0.00	3.48	1.00	0.00	True
SuttonAutxt	80.0	49.0	319.0	4.19	0.15	0.22	0.66	1.02	1.28	-13.27	0.00	0.66	1.00	0.00	True
SuttonSptxt	74.0	50.0	379.0	5.28	0.10	0.18	0.56	1.01	1.15	-35.01	0.00	0.57	1.00	0.00	True
Sutton Sutxt	87.0	63.0	387.0	4.87	0.28	0.31	0.89	1.08	0.52	-1.59	0.00	0.80	1.00	0.00	True

	Troy	77.0	40.0	144	0 2.35	0.19	0.51	0.37	0.99	1.14	4 -12.	16 (0.00	1.14	1.00	0.00	True
Food Web	Venlawtxt	66.0	30.0	158.	.0 2.83	0.23	0.43	0.53	1.01	1.3	5 -6.	72 (0.00	1.09	1.00	0.00	True
1000 Web	Weddel_sea	483.0	61.0	5562	0 31.71	0.72	1.32	0.55	0.75	1.1	7 2.	63 22	2.91	26.99	0.98	0.20	False
	Ythan96	82.0	5.0	42.	.0 4.77	0.42	2.88	0.15	0.28	0.9) -1.	32	1.00	4.10	1.00	0.02	False
			N	В	Lb	<k></k>	q	q'	q/q'	s/s'	α/α'	τ	λ_1	α	df	φ	асу
	net_cance	er 404	19.0	3967.0	11623.0	2.89	0.08	0.09	1.00	1.00	1.07	-0.16	2.54	1.10	1.00	0.0	False
	net_co	li 41	8.0	312.0	475.0	1.24	0.27	0.30	0.88	0.99	0.94	-2.54	0.00	0.39	1.00	0.0	True
	net_e_co	li 147	70.0	1316.0	2769.0	1.98	0.23	0.22	1.03	1.00	1.21	0.65	1.62	1.10	1.00	0.0	False
Genetic	net_m_tuberculosi	s 162	24.0	1542.0	3084.0	1.95	0.17	0.17	1.02	1.00	1.24	0.99	2.00	1.28	1.00	0.0	False
	net_non_cance	er 407	71.0	4004.0	8407.0	2.08	0.08	0.08	0.99	1.00	0.99	-1.54	1.00	0.87	1.00	0.0	False
	net_p_aeruginos	a 69	1.0	606.0	910.0	1.43	0.30	0.30	1.00	1.03	1.94	0.58	1.41	1.85	0.99	0.0	False
	net_s_cerevisia	e 293	33.0	2764.0	5964.0	2.10	0.17	0.18	0.98	1.00	1.29	-0.38	1.00	1.43	1.00	0.0	False
	net_yeas	st 68	88.0	557.0	1021.0	1.57	0.25	0.24	1.04	1.00	0.81	-0.31	1.32	0.36	1.00	0.0	False

50.0 16.0 20.0 2.02 2.04 2.01 1.01 1.16

q/q' s/s' α/α'

1.55 1.31 3.17 3.69 0.87 0.52 False

N

net_green_eggs

Language

М	leta	hol	lic
1 * 1	Cla	$\mathcal{O}_{\mathcal{O}}$	LIC

	N	В	Lb	<k></k>	q	q'	q/q'	s/s'	α/α'	τ	λ_1	α	df	φ	acy
net_AG	1267.0	36.0	55.0	2.38	13.79	7.33	1.88	2.06	4.34	2.35	7.62	10.43	0.83	0.93	False
net_CE	1172.0	40.0	55.0	2.44	13.29	7.15	1.86	2.04	4.60	2.44	8.00	11.41	0.83	0.92	False
net_CQ	386.0	20.0	25.0	2.05	8.98	5.54	1.62	1.70	2.55	1.69	5.57	5.34	0.83	0.84	False
net_CT	446.0	19.0	25.0	2.11	11.77	6.05	1.94	2.02	2.77	1.79	6.07	5.93	0.83	0.86	False
net_MJ	1081.0	32.0	56.0	2.40	12.47	6.73	1.85	1.98	4.00	2.27	7.53	9.65	0.83	0.93	False
net_SC	1510.0	43.0	53.0	2.54	14.61	8.45	1.73	1.82	5.54	2.66	9.15	14.29	0.83	0.94	False
net_TH	1111.0	30.0	56.0	2.43	12.17	6.88	1.77	1.90	4.08	2.31	7.59	9.99	0.83	0.93	False

Neural

	N	В	Lb	<k></k>	q	q'	q/q'	s/s'	α/α'	τ	λ_1	α	df	φ	acy
net_celegans_neural	297.0	3.0	175.0	7.9	1.49	3.52	0.42	0.39	1.42	2.17	9.15	10.49	0.88	0.8	False

N B Lb <k> q q' q/q' s/s' α/α' τ $λ_1$ α df φ acy net_trade_basic 24.0 2.0 21.0 12.92 4.24 3.71 1.14 1.14 1.10 2.68 14.30 14.51 0.49 0.92 False net_trade_minerals 24.0 3.0 8.0 5.62 4.21 3.98 1.06 0.97 1.28 2.05 7.25 7.74 0.64 0.88 False

Trade