清華大學
Tsinghua University

Machine intelligence and network science for complex systems big data analysis

Carlo Vittorio Cannistraci

Center for Complex Network Intelligence (CCNI)

Tsinghua Laboratory of Brain and Intelligence

清华大学脑与智能实验室
Tsinghua Laboratory of Brain and Intelligence

Center for Complex Network Intelligence (CCNI)

Carlo Vittorio Cannistraci
Chair Professor and Chief Scientist
Period at THBI: 2020 - Now

清华大学脑与智能实验室
Tsinghua Laboratory of Brain and Intelligence

清華大學
Tsinghua University

Research philosophy

A transdisciplinary approach integrating information theory, machine learning and network science to investigate the physics of networked adaptive complex systems at different scales, from molecules to ecological and social systems, with a particular attention to brain/bio-inspired computing and complex big data (focus on: neuroscience, biomedicine and social science) pattern recognition analysis.

1. Theoretical topics

1.1 Network Geometry

1.2. Network topology & network automata as models of network self-organization

1.3. Complex network intelligence and Brain inspired computing for AI

1.4. Geometric Machine Learning

2. Applied topics

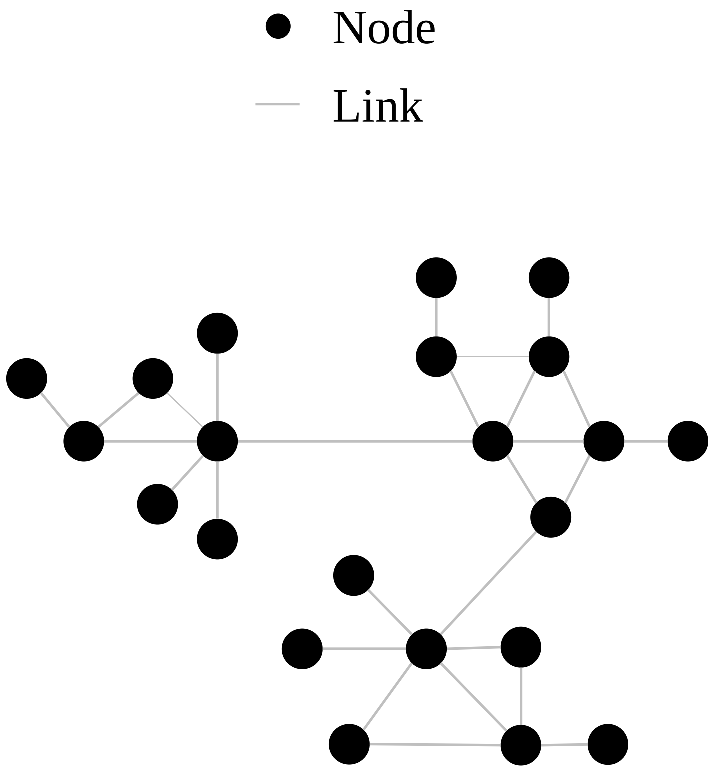
2.1 Brain Networks & Network neuroscience

2.2 Omic data analysis & design of multiomic biomarkers for precision medicine

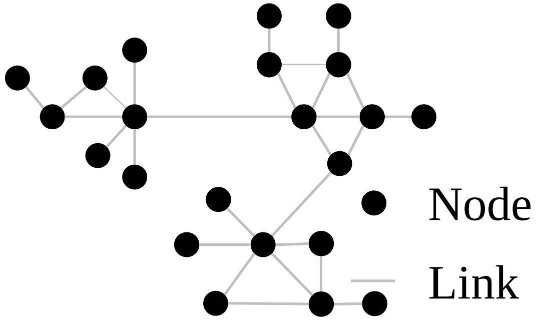
2.3 Neuromorphic and unconventional computing for AI

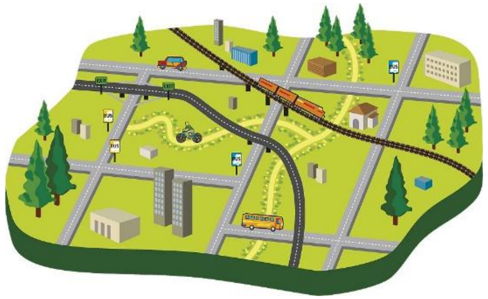
2.4 Social and Economical systems

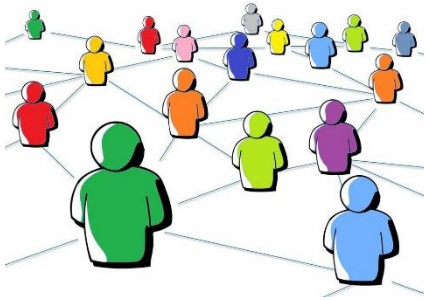
What is the difference between graph and network?

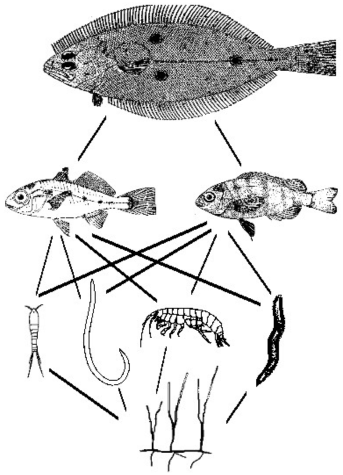


Complex Systems and Network Science







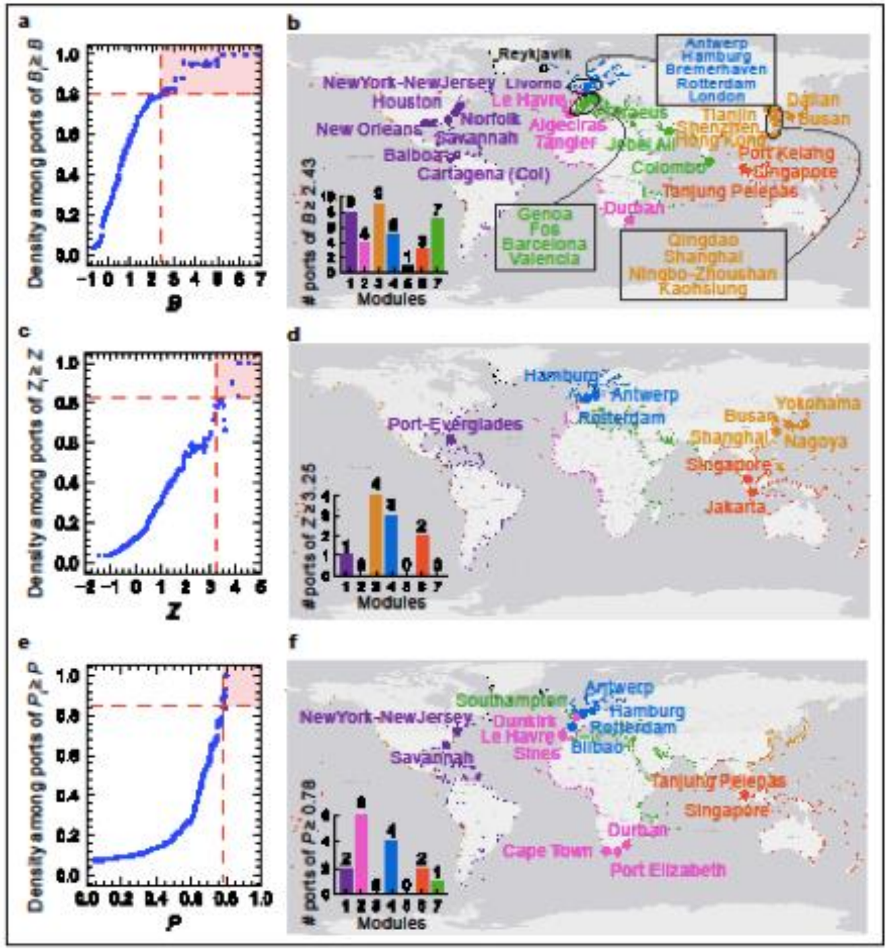


Modular gateway-ness connectivity and structural core organization in **maritime network science**.

Mengqiao Xu, Qian Pan, Alessandro Muscoloni, Haoxiang Xia and Carlo Vittorio Cannistraci.

Nature Communications 2020

The modular gateway-ness connectivity of maritime networks follows a core organization paradigm similar to brain networks



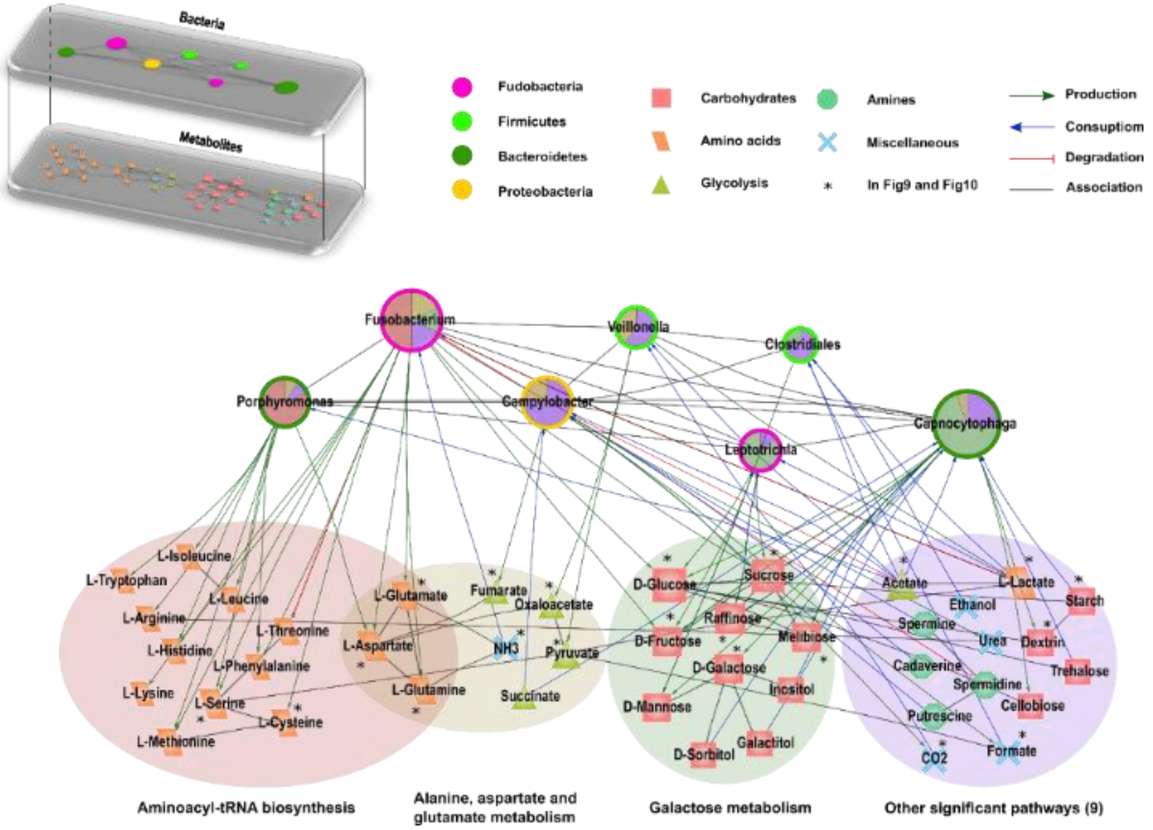
(2020)

Nonlinear machine learning pattern recognition and bacteria-metabolite multilayer network analysis of perturbed gastric microbiome.

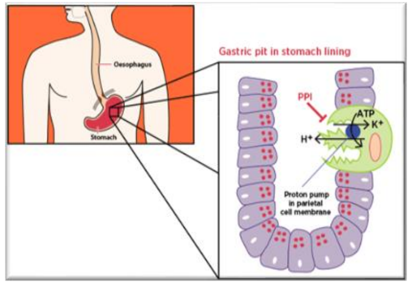
C Durán, S Ciucci, … and Carlo Vittorio Cannistraci

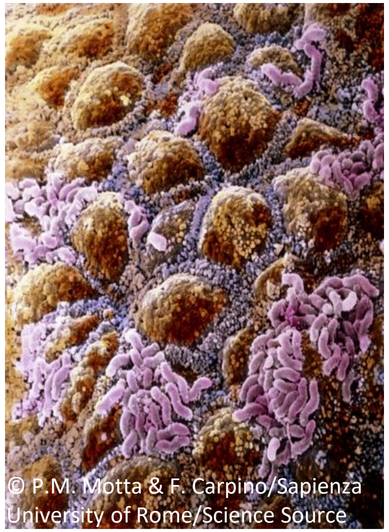
Nature communications 12 (1), 1-22, 2021

The discovered bacteria-metabolite network affected in gastric environment of dyspeptic patients.



(2021)





Gastric microbiota

nature communications

Article https://doi.org/10.1038/s41467-022-34634-6

Geometrical congruence, greedy navigability and myopic transfer in complex networks and brain connectomes

Received: 6 July 2020

Accepted: 1 November 2022

Published online: 27 November 2022

Check for updates

Carlo Vittorio Cannistraci[1,2,3,4,5,6] & Alessandro Muscoloni[1,4]

We introduce in network geometry a measure of *geometrical congruence (GC)* to evaluate the extent a network topology follows an underlying geometry. This requires finding all topological shortest-paths for each nonadjacent node pair in the network: a nontrivial computational task. Hence, we propose an optimized algorithm that reduces 26 years of worst scenario computation to one week parallel computing. Analysing artificial networks with patent geometry we discover that, different from current belief, hyperbolic networks do not show in general high GC and efficient greedy navigability (GN) with respect to the geodesics. The myopic transfer which rules GN works best only when degree-distribution power-law exponent is strictly close to two. Analysing real networks—whose geometry is often latent—GC overcomes GN as marker to differentiate phenotypical states in macroscale structural-MRI brain connectomes, suggesting connectomes might have a latent neurobiological geometry accounting for more information than the visible tridimensional Euclidean.

November 2022

OUTLINE of the talk

1. Introduction to network analysis and models

2. Network geometry, AI and applications

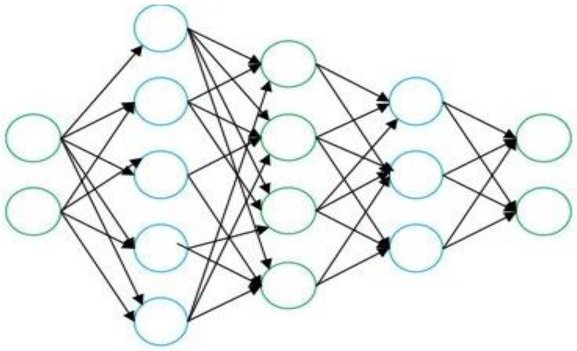
3. Tomorrow:

3.1 Network science for Sparse deep learning

3.2 Neuromorphic Computing

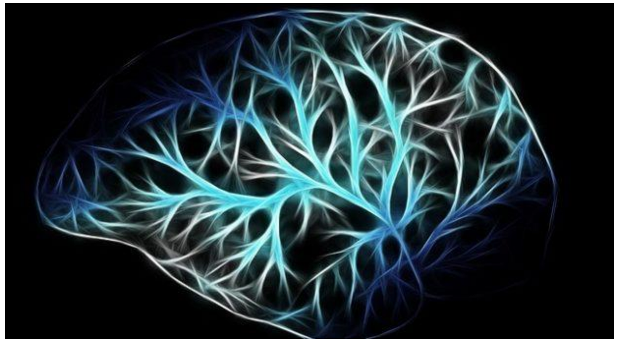
Crisis: I was a Master student in 2002

Artificial Neural Network (ANN)



Vs.

Brain Connectivity



Crisis: Why is brain connectivity sparse?

Yingtao Zhang

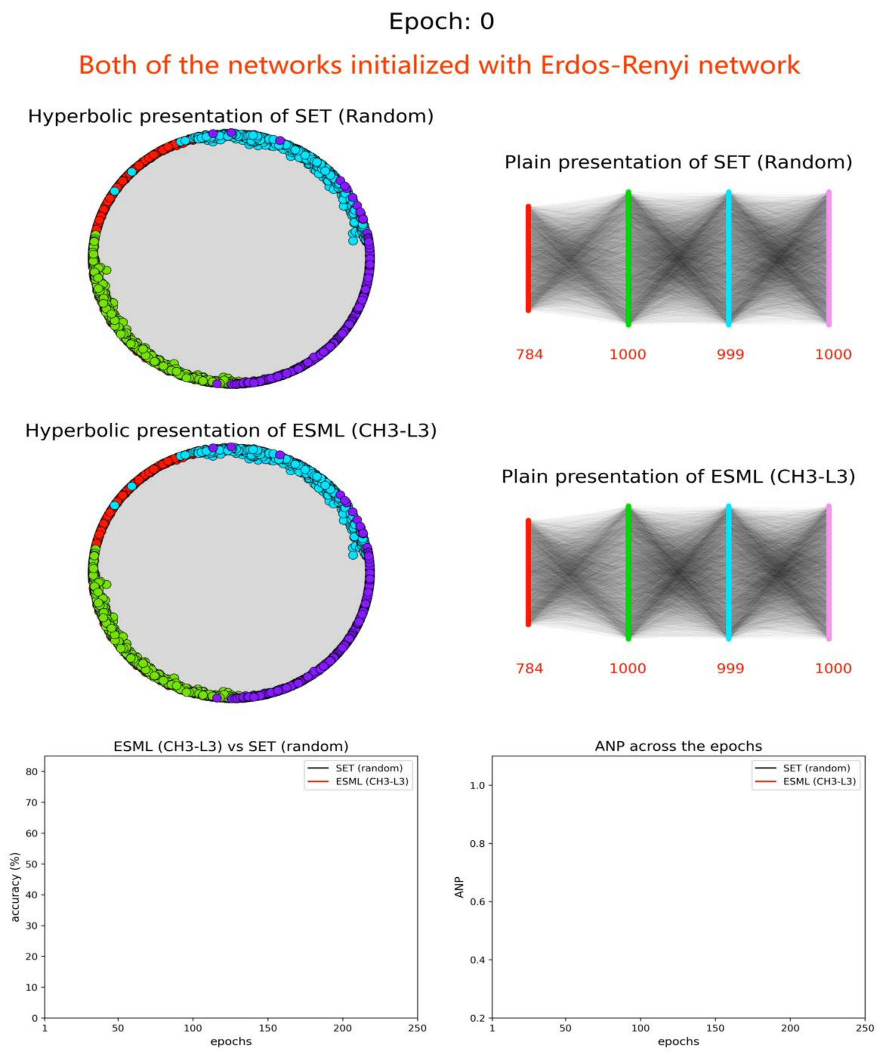
EPITOPOLOGICAL LEARNING AND CANNISTRACI-HEBB NETWORK SHAPE INTELLIGENCE BRAIN-INSPIRED THEORY FOR ULTRA-SPARSE ADVANTAGE IN DEEP LEARNING

Yingtao Zhang[1,2,3], **Jialin Zhao**[1,2,3], **Wenjing Wu**[1,2,3], **Alessandro Muscoloni**[1,2,4], & **Carlo Vittorio Cannistraci**[1,2,3,4] *
[1]Center for Complex Network Intelligence (CCNI)
[2]Tsinghua Laboratory of Brain and Intelligence (THBI)
[3]Department of Computer Science, [4]Department of Biomedical Engineering
Tsinghua University, Beijing, China.

ICLR2024 evaluation: avg. score 7.33, ranks 326/2261 accepted (in the top 15%)

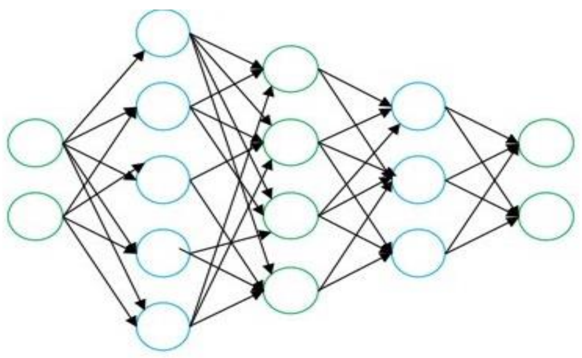
How the topology evolves during the epochs

https://www.youtube.com/watch?v=b5lLpOhb3BI



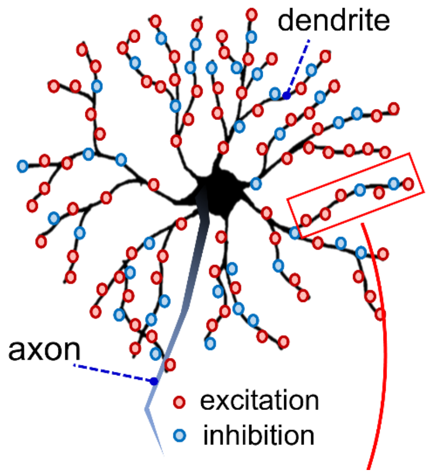
Crisis: I was a Master student in 2002

Artificial Neural Network (ANN)



Vs.

Brain Morphology



Crisis: Why is brain connectivity sparse (topology)?

Crisis: What is the contribution of morphology?

Preprints.org Instructions for Authors Statistics About FAQ Search here...

preprints.org > computer science and mathematics > artificial intelligence and machine learning > doi: 10.20944/preprints2

Preprint Article Version 1 Preserved in Portico This version is not peer-reviewed

Neuromorphic Dendritic Computation with Silent Synapses for Visual Motion Perception

Eunhye Baek *, Sen Song, Zhao Rong, Luping Shi *, Carlo Vittorio Cannistraci *

Version 1 : Received: 5 June 2023 / Approved: 6 June 2023 / Online: 6 June 2023 (10:04:05 CEST)

How to cite: Baek, E.; Song, S.; Rong, Z.; Shi, L.; Cannistraci, C.V. Neuromorphic Dendritic Computation with Silent Synapses for Visual Motion Perception. *Preprints* **2023**, 2023060438. https://doi.org/10.20944/preprints202306.0438.v1 Copy

Dr. E. Baek

Abstract

Most neuromorphic technologies use a point-neuron model, missing the spatiotemporal nature of neuronal computation performed in dendrites. Dendritic morphology and synaptic organization are structurally tailored for spatiotemporal information processing, enabling various computations like visual perception. Here, we report on a neuromorphic computational model termed 'dendristor', which integrates functional synaptic organization with dendritic tree-like morphology computation. The dendristor presents bioplausible nonlinear integration of excitatory and inhibitory synaptic inputs with silent synapses and diverse spatial distribution dependency. We show that the dendristor can emulate direction selectivity, which is the feature to react robustly to a preferred signal direction on the dendrite. We discover that silent synapses can remarkably enhance direction selectivity, turning out to be a crucial player in dendritic computation processing. Finally, we develop neuromorphic dendritic neural circuits that can emulate a cognitive function such as motion perception in the retina. Using dendritic morphology, we achieve visual perception of motion in 3D space by various mapping of spatial information on different dendritic branches. This neuromorphic dendritic computation innovates beyond current neuromorphic computation and provides solutions to explore new skylines in artificial intelligence, neurocomputation, and brain-inspired computing.

Nature Electronics 2024 Accepted

OUTLINE of the talk

1. Introduction to network analysis and models

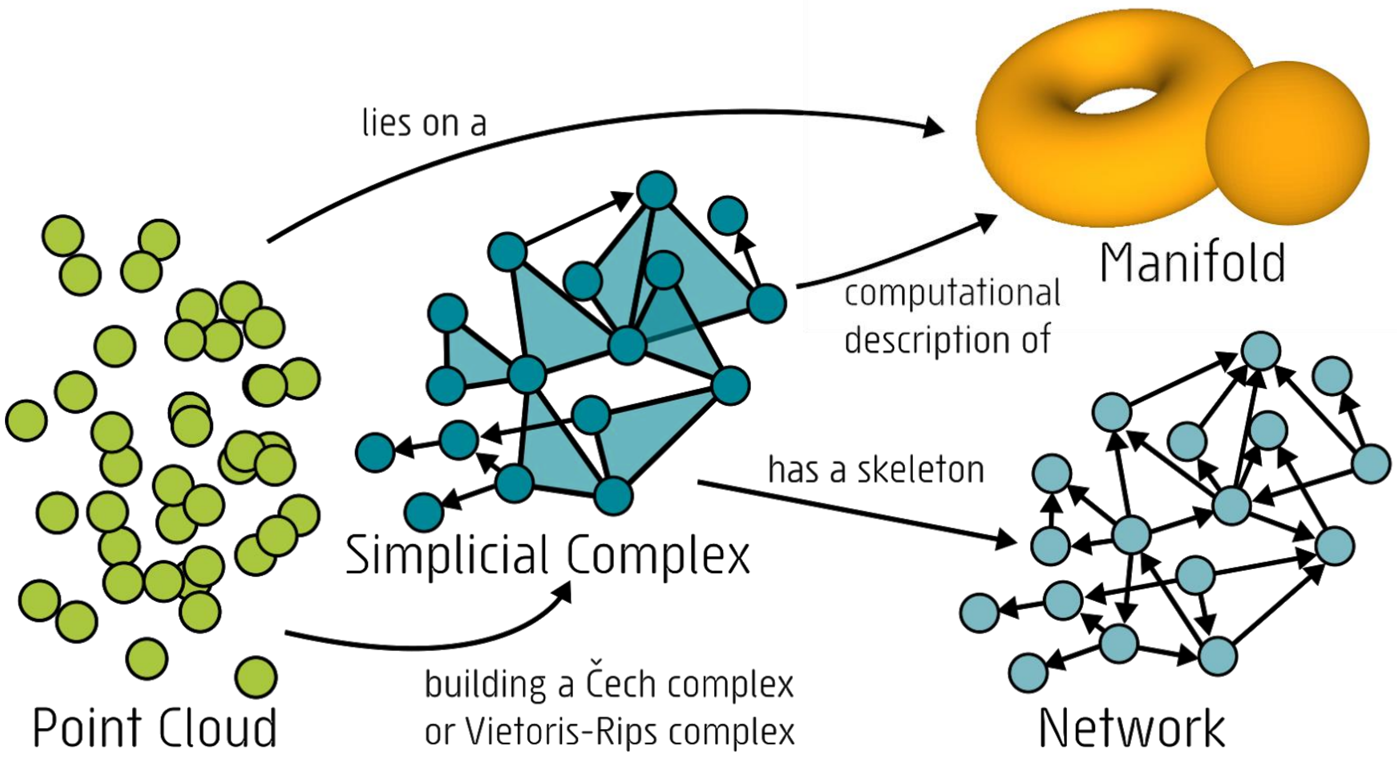
2. Network geometry, AI and applications

3. Tomorrow:

3.1 Network science for Sparse deep learning

3.2 Neuromorphic Computing

Artistic representation of the topics of today

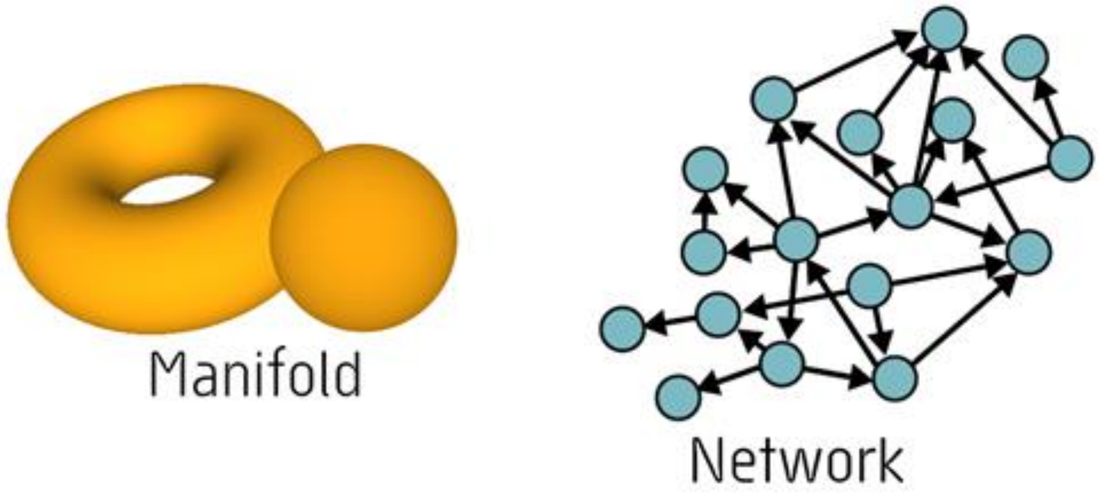


https://menchelab.com/higher-order-networks-and-the-topology-of-data

Artistic representation of the topics of today

Understanding the rule of association generating the networks

(direct problem)

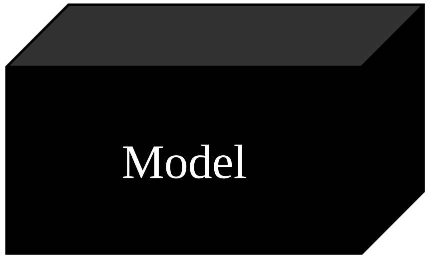


(inverse problem)

Given the network can we reverse the rules of association

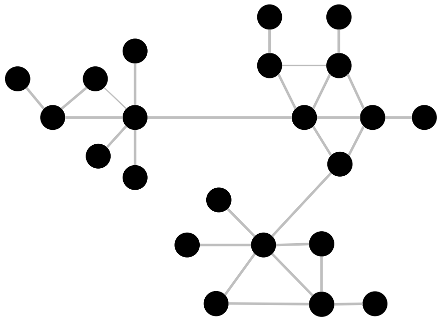
Generative models in Network Science

Input parameters →

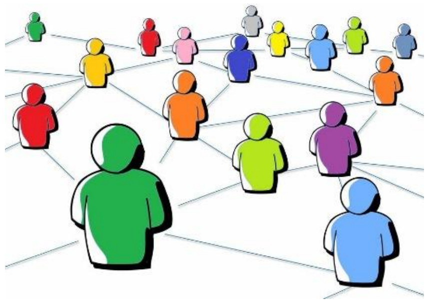


Analytical
or
Algorithical

→



≈

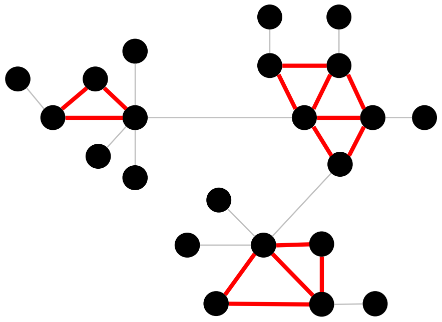


properties of real complex networks

Three basic properties of real complex networks

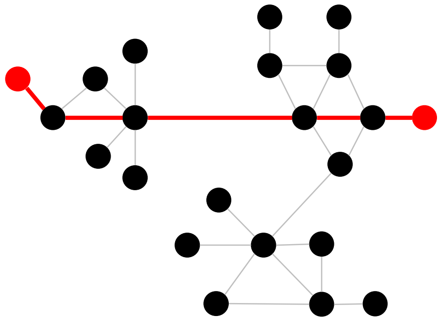
Average

Clustering Coefficient



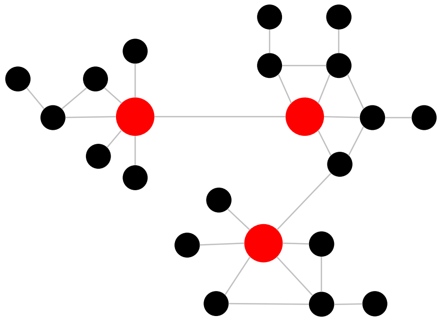
Average

Shortest path length

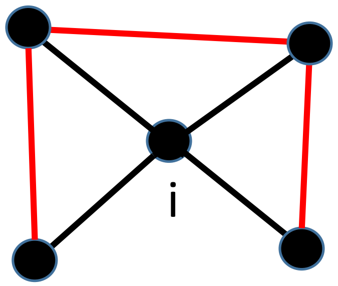


Average Degree

Average degree distribution



Average cluster coefficient



m = 4 nodes neighbors

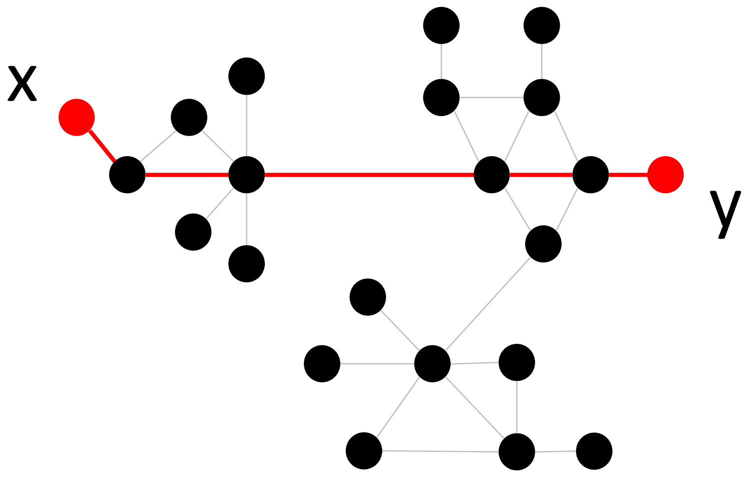
Cluster coefficient of a node (i) C(i)

$$C(i) = \frac{\text{\# Links between neighbors}}{\text{All possible Links between neighbors}} = \frac{k}{\binom{m}{2}} = \frac{k}{m*(m-1)/2}$$

For a network with N nodes

Average Clustrer Coefficinet = Σi C(i)/N

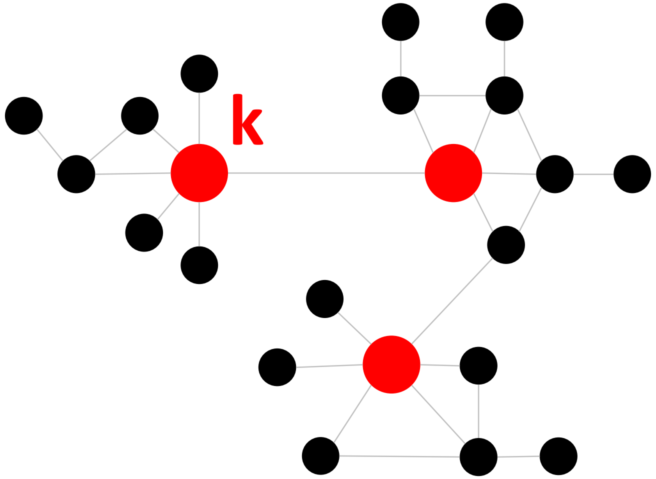
Average Shortest path length

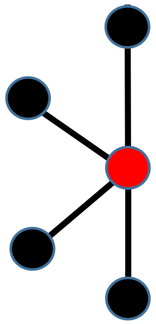


L= Σ SP(x,y)/N

Small wordness → L(N) ~ log (N)

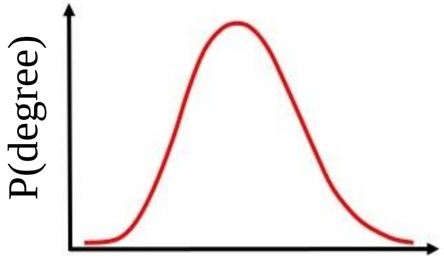
Average degree and degree distribution



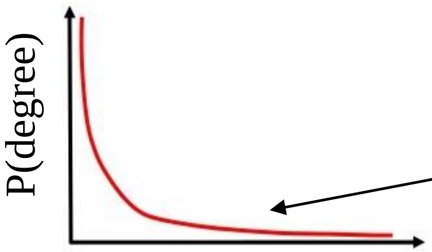


K (i) = # neighbors of i

binomial



Exponential or Power-law

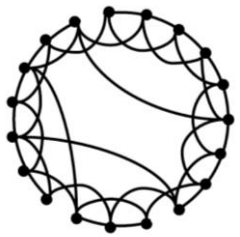


P(k) = exp(-k) [slim tale]

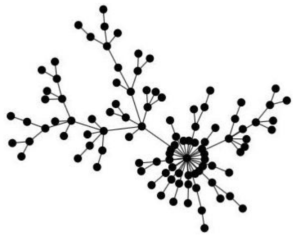
scale free →P(k) = k^(-ɣ) [fat tale]

Generative models

1959

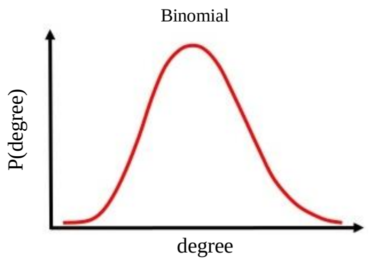


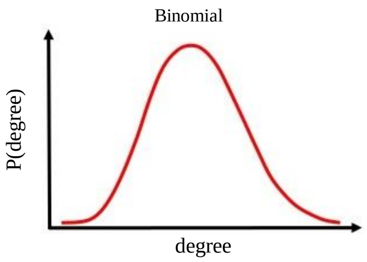
1998

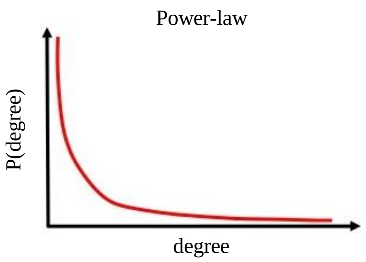


1999

	Erdős-Rényi	Watts-Strogatz	Barabási-Albert
Clustering	—	☑	—
Small-world	☑	☑	☑
Scale-free	—	—	☑

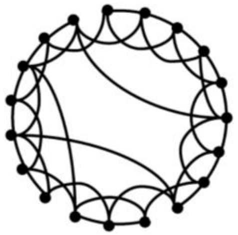




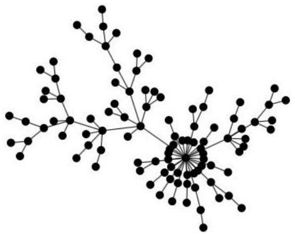


Generative models

1959



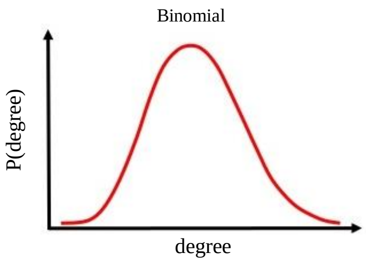
1998

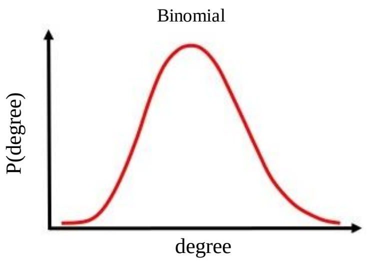


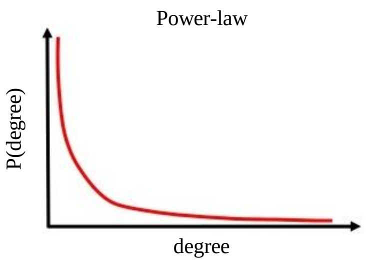
1999

Problem

	Erdős-Rényi	Watts-Strogatz	Barabási-Albert	????
Clustering	—	☑	—	☑
Small-world	☑	☑	☑	☑
Scale-free	—	—	☑	☑







OUTLINE of the talk

1. Introduction to network analysis and models

2. Network geometry, AI and applications

3. Tomorrow:
3.1 Network science for Sparse deep learning
3.2 Neuromorphic Computing

LETTER

2012

doi:10.1038/nature11459

Popularity versus similarity in growing networks

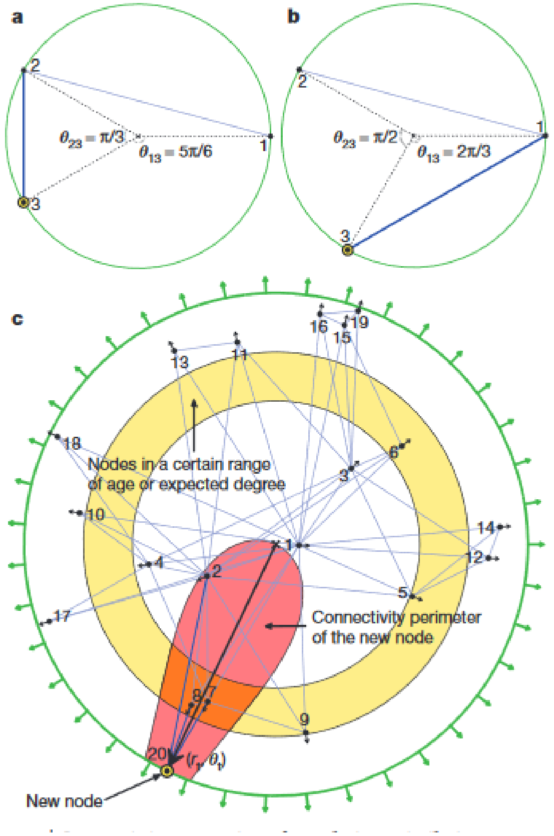
Fragkiskos Papadopoulos[1], Maksim Kitsak[2], M. Ángeles Serrano[3], Marián Boguñá[3] & Dmitri Krioukov[2]

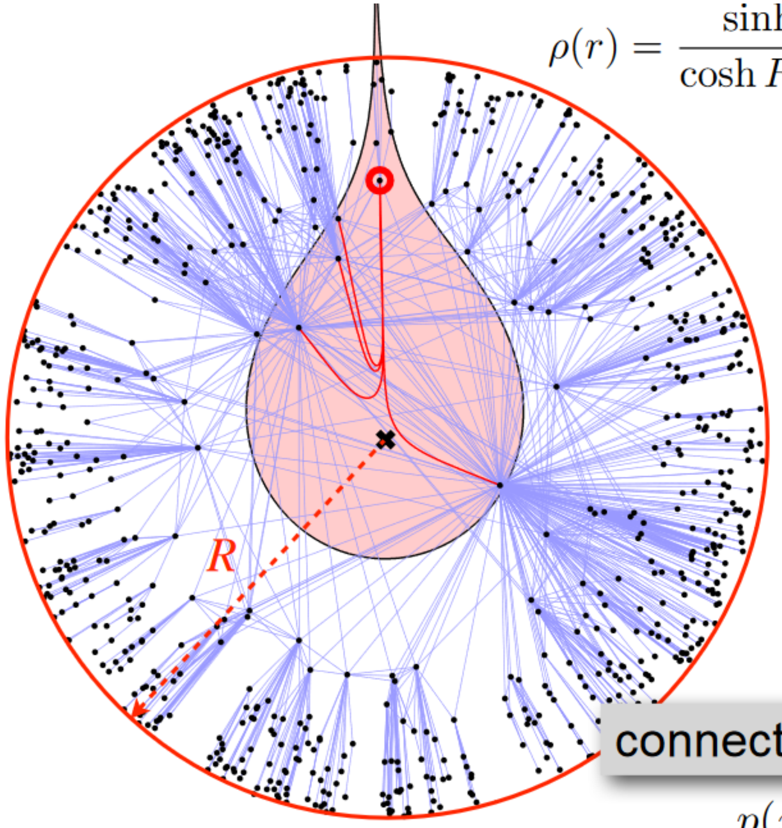
The principle[1] that 'popularity is attractive' underlies preferential attachment[2], which is a common explanation for the emergence of scaling in growing networks. If new connections are made preferentially to more popular nodes, then the resulting distribution of the number of connections possessed by nodes follows power laws[3,4], as observed in many real networks[5,6]. Preferential attachment has been directly validated for some real networks (including the Internet[7,8]), and can be a consequence of different underlying processes based on node fitness, ranking, optimization, random walks or duplication[9–16]. Here we show that popularity is just one dimension of attractiveness; another dimension is similarity[17–24]. We develop a framework in which new connections optimize certain trade-offs between popularity and similarity, instead of simply preferring popular nodes. The framework has a geometric interpretation in which popularity preference emerges from local optimization. As opposed to preferential attachment, our optimization framework accurately describes the large-scale evolution of technological (the Internet), social (trust relationships between people) and biological (*Escherichia coli* metabolic) networks, predicting the probability of new links with high precision. The framework that we have developed can thus be used for predicting new links in evolving networks, and provides a different perspective on preferential attachment as an emergent phenomenon.

Nodes that are similar have a higher chance of getting connected, even if they are not popular. This effect is known as homophily in social sciences[17,18], and it has been observed in many real networks[19–24]. In the web[23,24], for example, an individual creating her new homepage tends to link it not only to popular sites such as Google or Facebook, but also to not-so-popular sites that are close to her special interests—for example, sites devoted to the composer Tartini or to free solo climbing. These observations suggest the introduction of a measure of attractiveness that would somehow balance popularity and similarity.

The simplest proxy for popularity is the node birth time. All other things being equal, older nodes have more chances to become popular and attract connections[3,4]. If nodes join the network one by one, then the node birth time is simply the node number $t = 1, 2, \ldots$. To model similarity, we randomly place nodes on a circle that represents the simplest similarity space. That is, the angular distances between nodes model their similarity distances, such as the cosine similarity or any

connect simply to the closest m nodes on the plane, except that distances are not Euclidean but hyperbolic[25]. The hyperbolic distance between two nodes at polar coordinates (r_s, θ_s) and (r_t, θ_t) is approximately $x_{st} = r_s + r_t + \ln(\theta_{st}/2) = \ln(st\theta_{st}/2)$. Therefore the sets of nodes s minimizing x_{st} or $s\theta_{st}$ for each t are identical. The hyperbolic





$$\rho(r) = \frac{\sinh r}{\cosh R - 1} \approx e^{r-R} \sim e^r$$

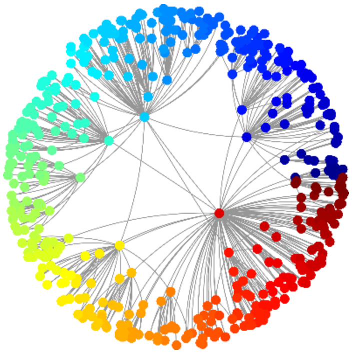
connection probability

$$p(x) = \Theta(R - x)$$

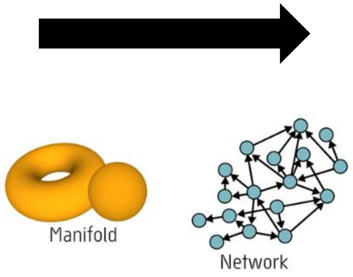
H2 in polar coordinates

Generative model for realistic complex networks

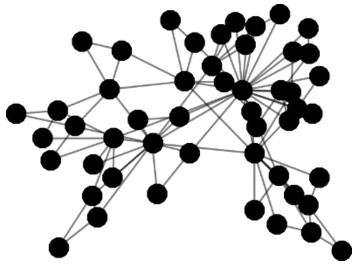
Geometry



PSO model



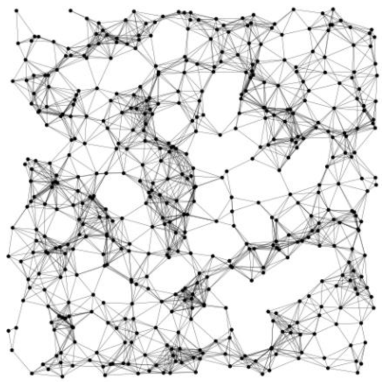
Topology



Synthetic network with:

- Clustering
- Small-word
- Scale-free

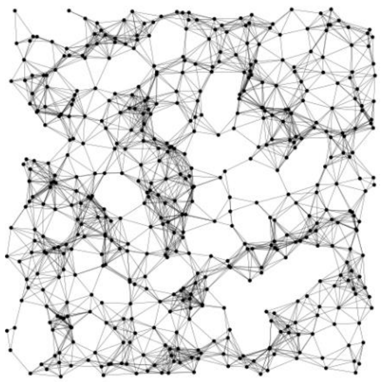
Generative models in geometric space
(soft random geometrical graph)

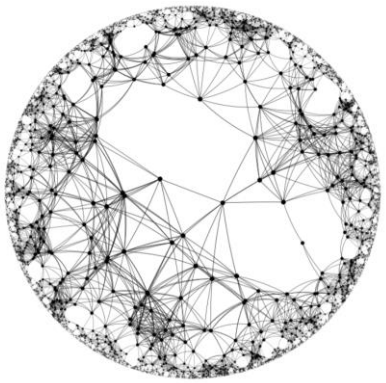


Problem

	Euclidean (soft)	??????
Clustering	☑	☑
Small-world	☑	☑
Scale-free	—	☑

Generative models in geometric space
(soft random geometrical graph)





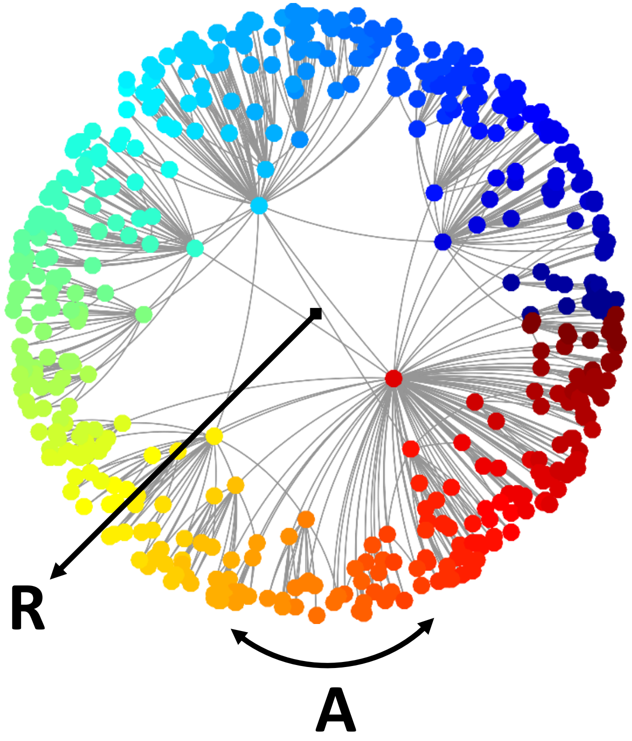
	Euclidean (soft)	Hyperbolic
Clustering	☑	☑
Small-world	☑	☑
Scale-free	—	☑

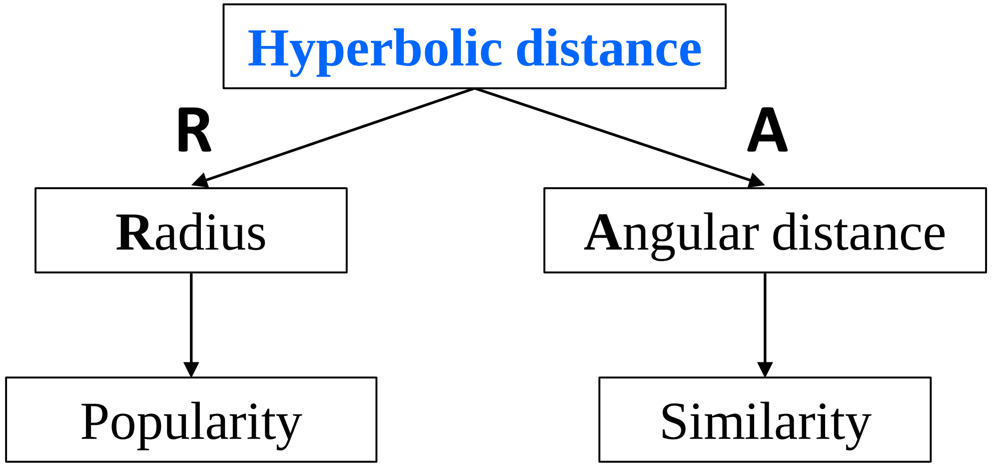
Hyperbolic distance

R → **R**adius

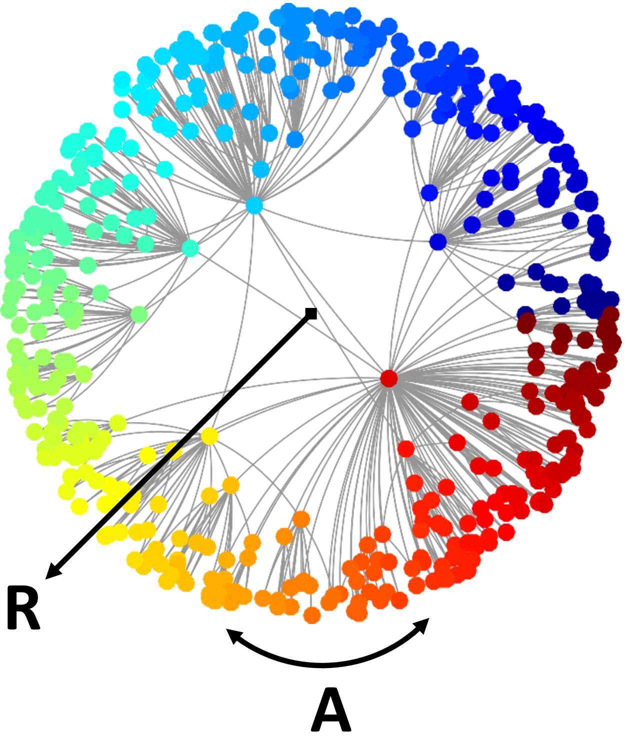
A → **A**ngular distance

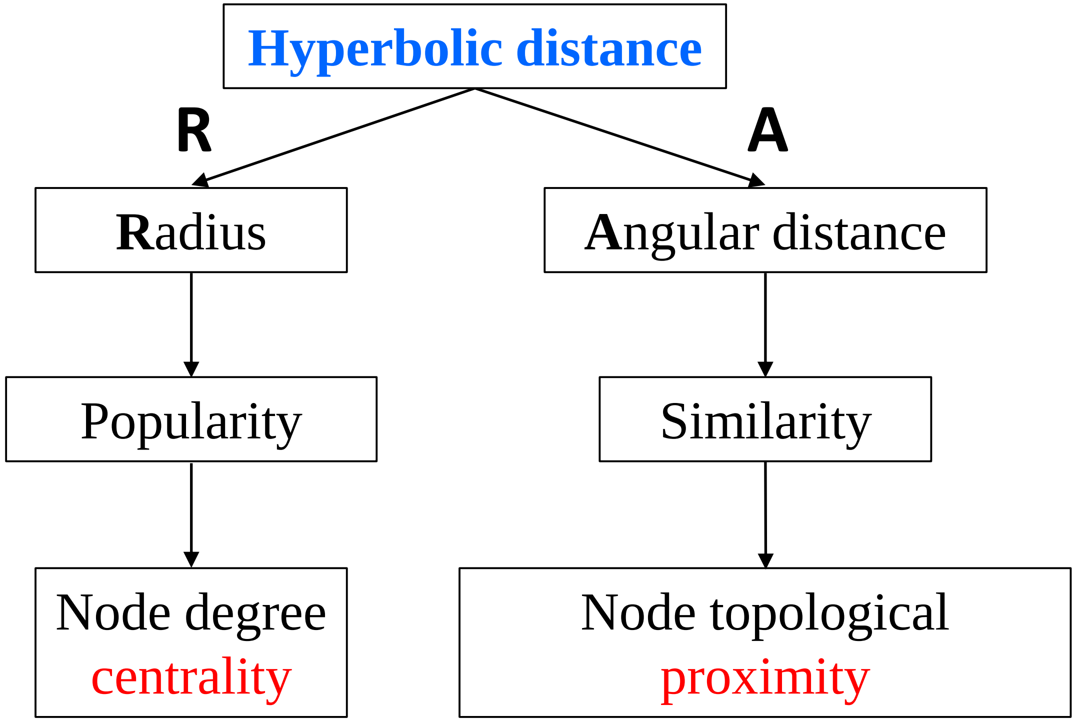
PSO model



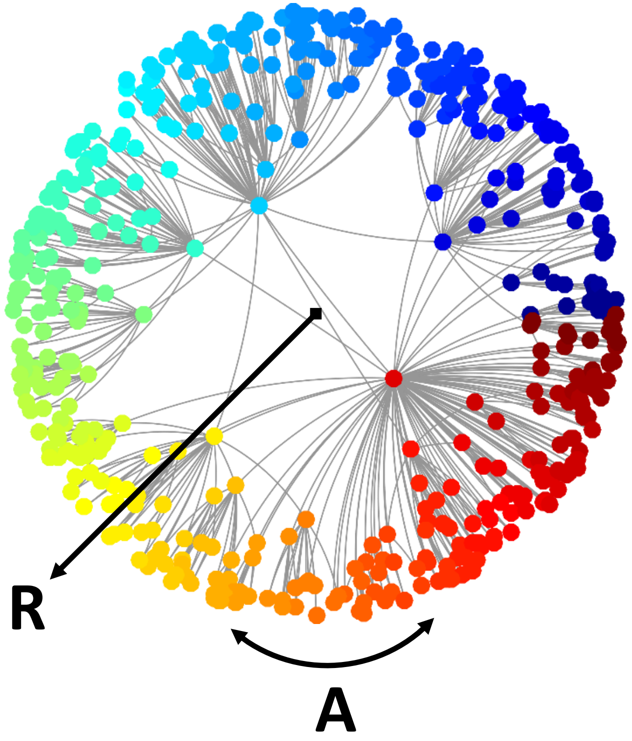

Hyperbolic distance
R
A
Radius
Angular distance
Popularity
Similarity

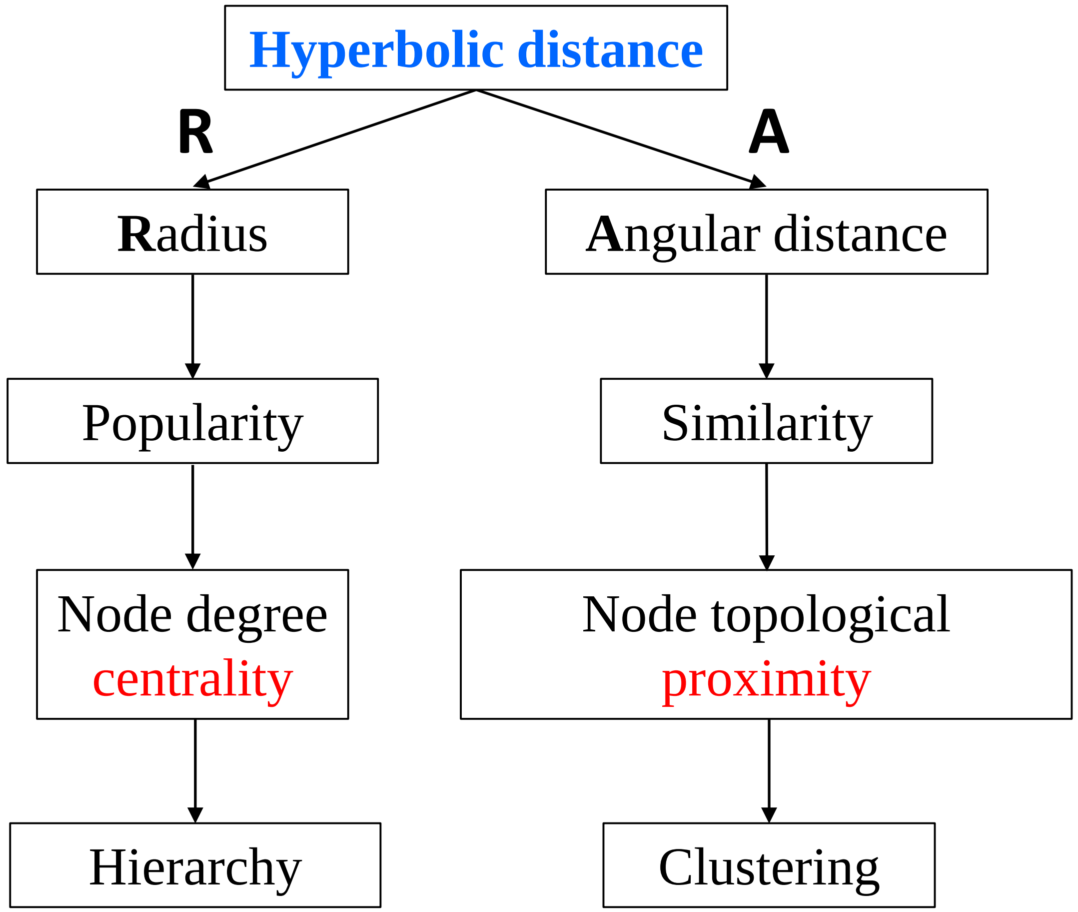
PSO model


R
A

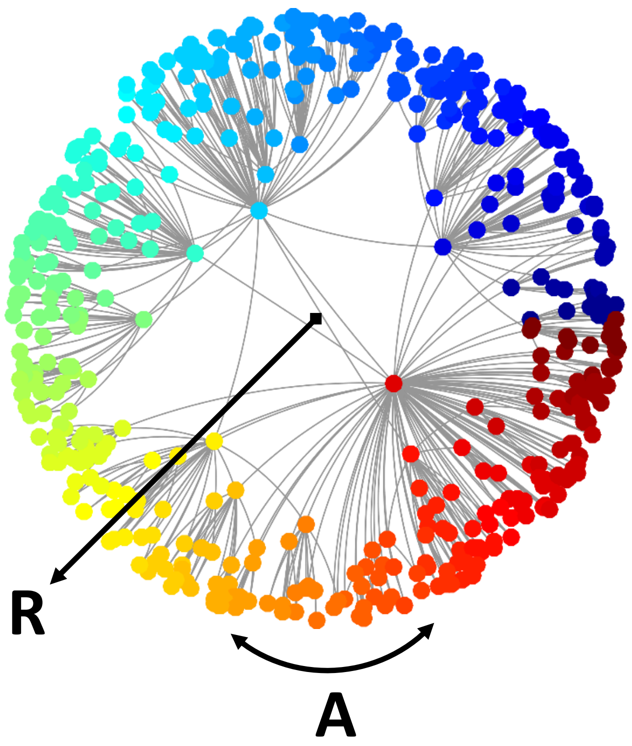

Hyperbolic distance
R
A
Radius
Angular distance
Popularity
Similarity
Node degree centrality
Node topological proximity

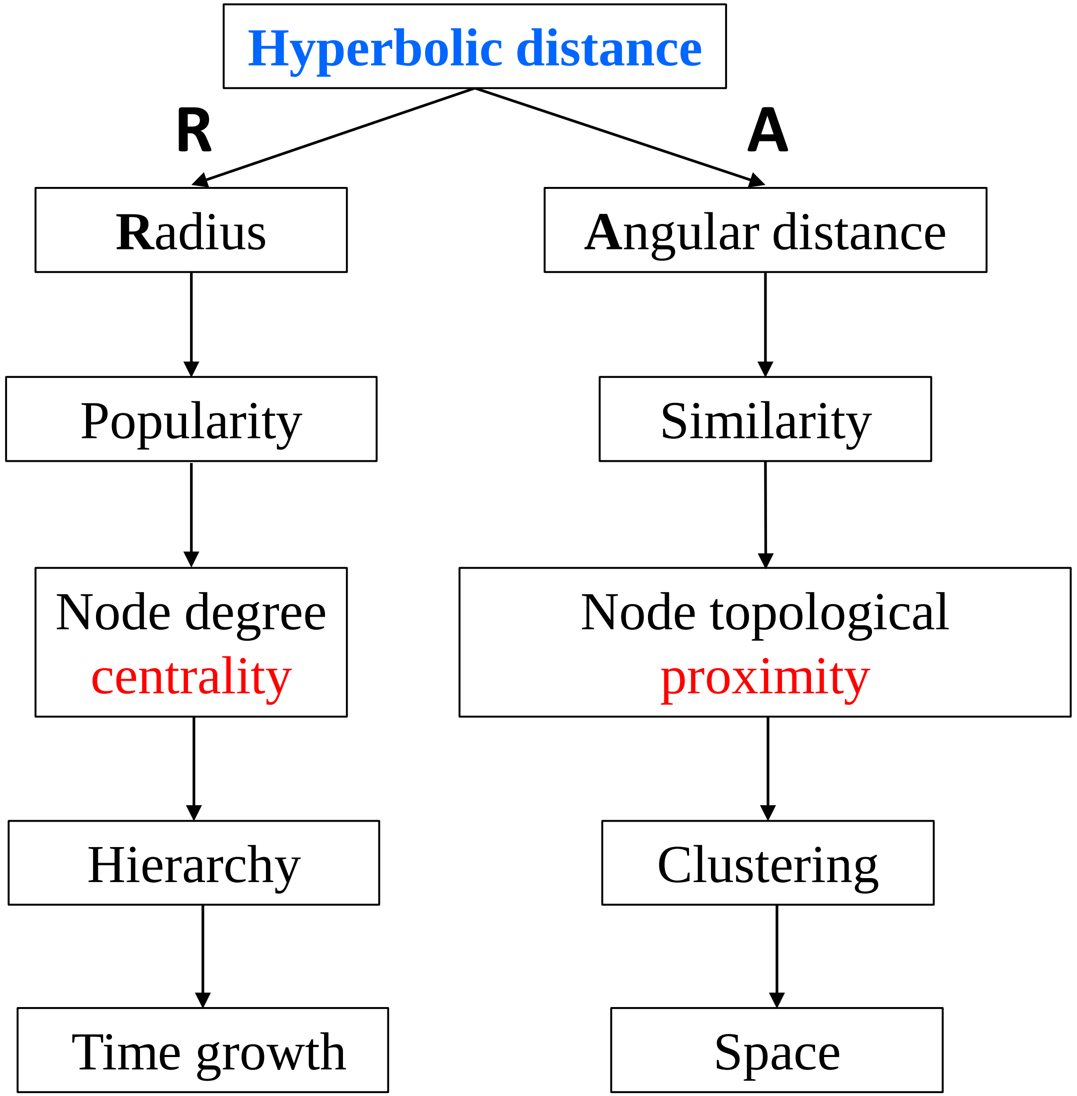
PSO model


R
A

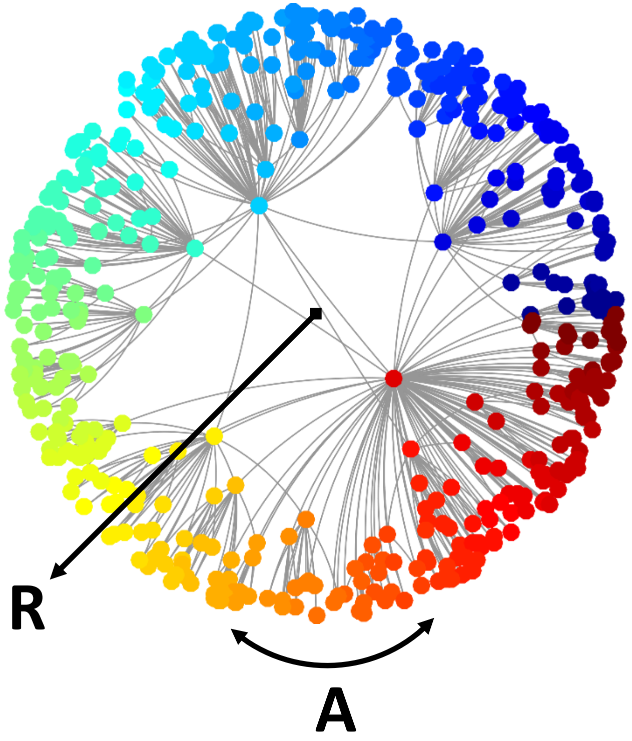

Hyperbolic distance
R
A
Radius
Angular distance
Popularity
Similarity
Node degree centrality
Node topological proximity
Hierarchy
Clustering

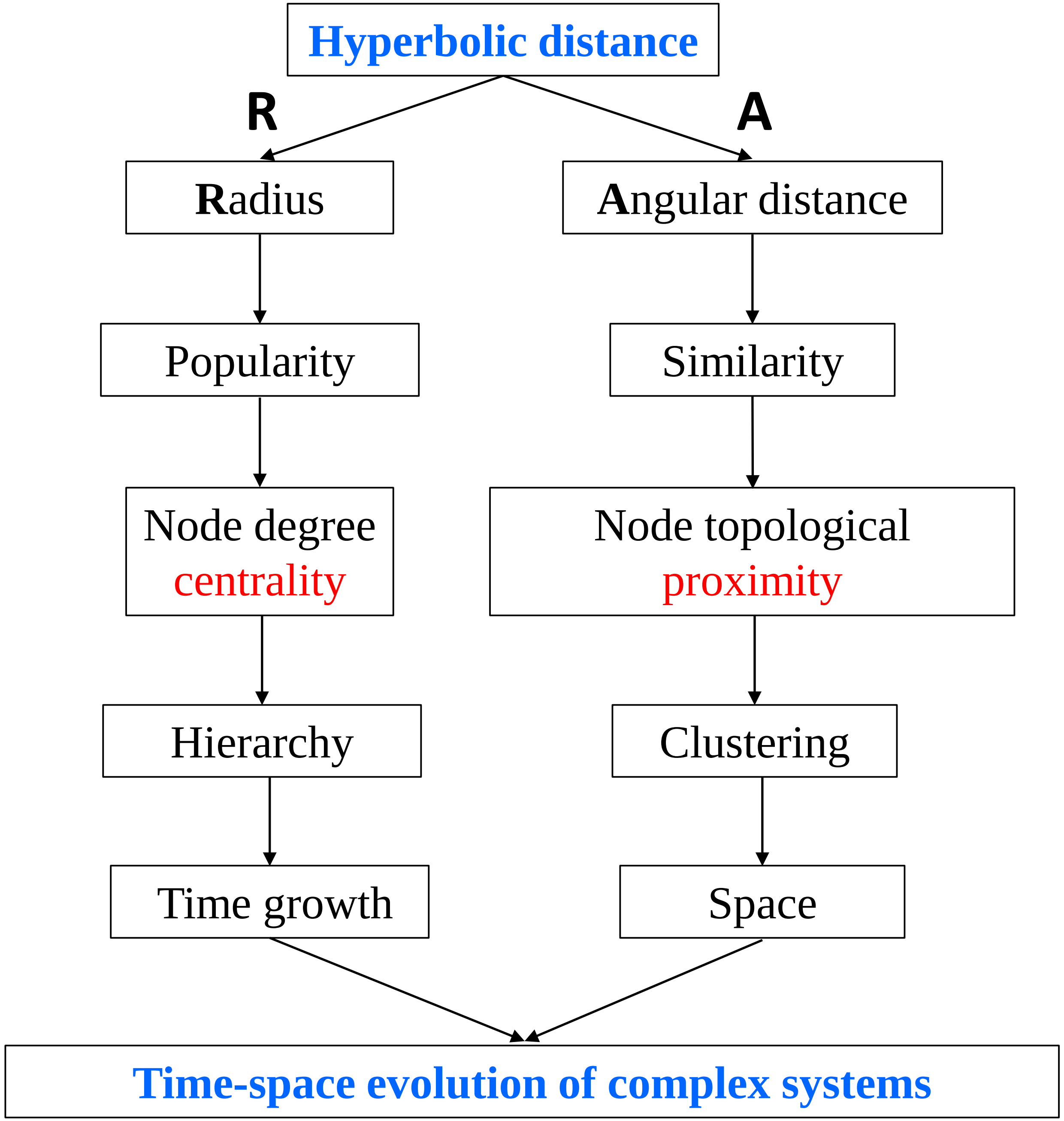
PSO model


R
A

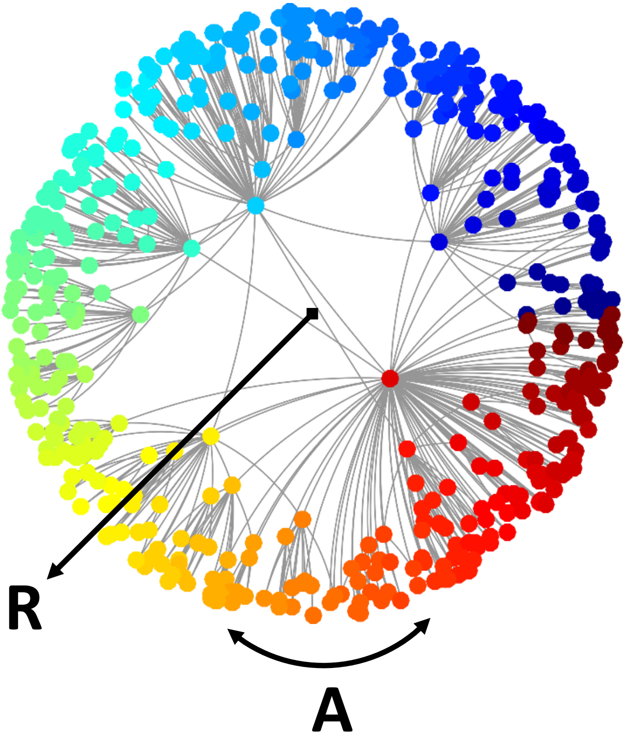

Hyperbolic distance
R
A
Radius
Angular distance
Popularity
Similarity
Node degree centrality
Node topological proximity
Hierarchy
Clustering
Time growth
Space

PSO model


R
A


Hyperbolic distance
R
A
Radius
Angular distance
Popularity
Similarity
Node degree centrality
Node topological proximity
Hierarchy
Clustering
Time growth
Space
Time-space evolution of complex systems

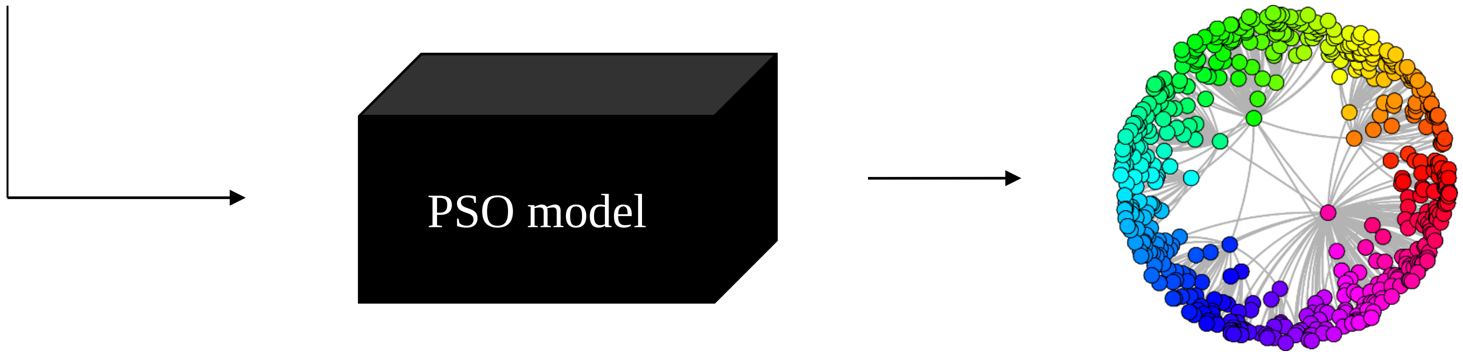
PSO model


R
A

Popularity-Similarity-Optimization (PSO) model

Input parameters:

$\boldsymbol{N}$ = number of nodes
$\boldsymbol{m}$ = half of the average node degree
$\boldsymbol{T}$ = temperature (inversely related to clustering)
$\boldsymbol{\gamma}$ = exponent of the power-law degree distribution



$$p(h_{ij}) = \frac{1}{1 + exp\left(\frac{h_{ij} - R_i}{T}\right)}$$

Papadopoulos et al., *Nature*, 2012.

The model has four input parameters:

Generation of synthetic networks by the PSO model

$\boldsymbol{m} > 0$, which defines the average node degree $\bar{k} = 2m$,

$\beta \in (0, 1]$ the exponent $\boldsymbol{\gamma} = 1 + 1/\beta$ of the power law degree distribution,

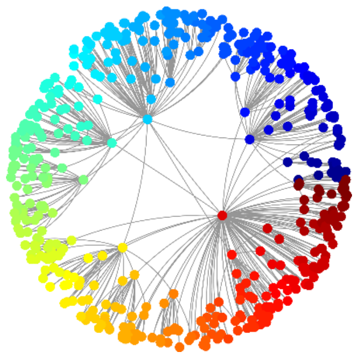
$\boldsymbol{T} \geq 0$, which controls the network clustering,

$\zeta = \sqrt{-K} > 0$, where $\boldsymbol{K}$ is the curvature of the hyperbolic plane. K generally fixed to -1

(1) all existing nodes $j < i$ **increase their radial coordinates according to** $\boldsymbol{r_j(i) = \beta r_j + (1 - \beta) r_i}$ **in order to simulate popularity fading**;

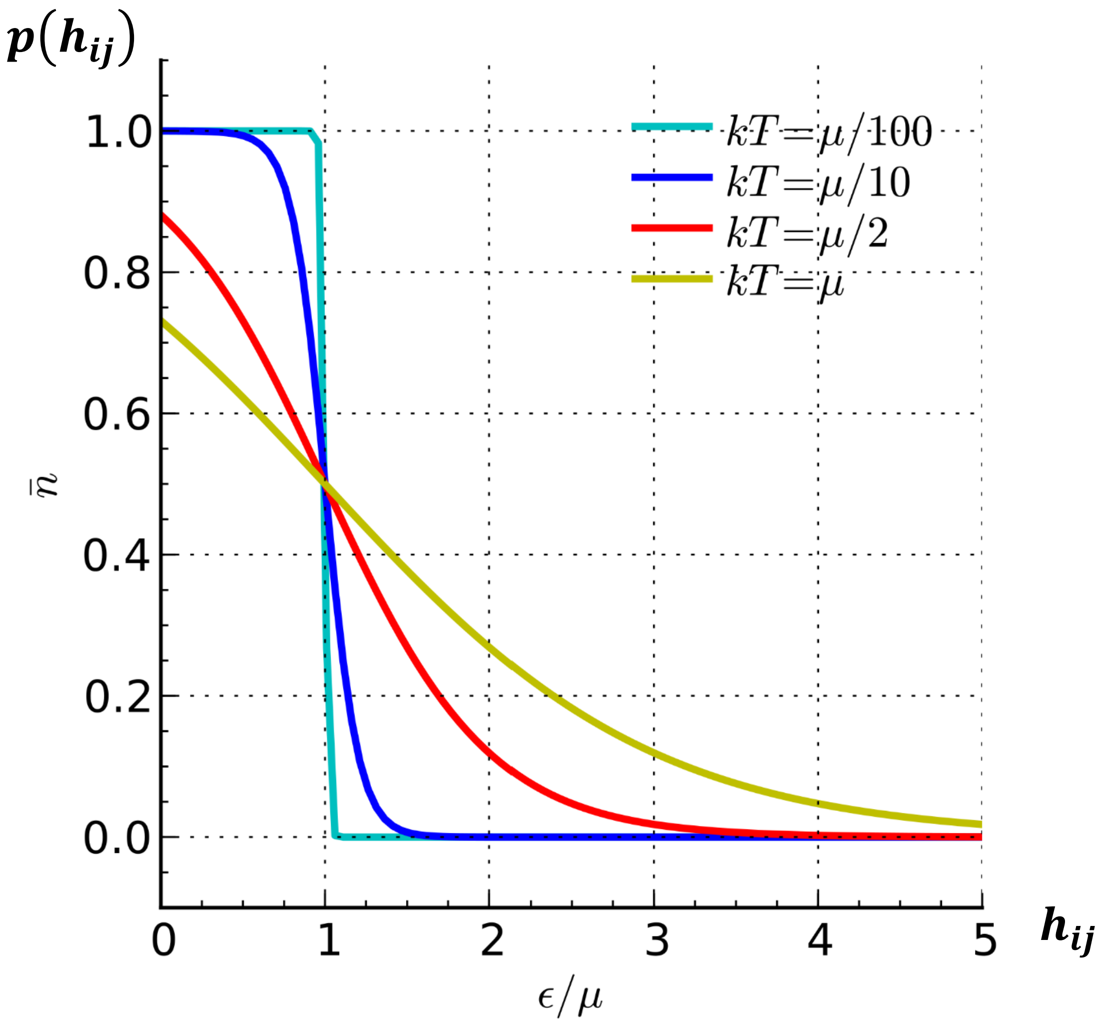
(2) The new node picks a randomly chosen existing node and connects to it with:

probability: $\boldsymbol{p(h_{ij}) = 1/(1 + exp((h_{ij} - R_i)/T))}$



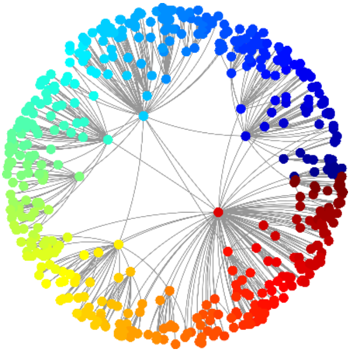
PSO model

Fermi-Dirac

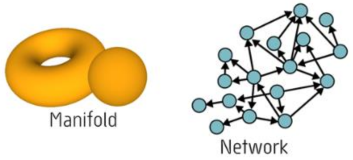

$p(h_{ij})$
1.0
0.8
0.6
0.4
0.2
0.0
$\bar{n}$
$kT = \mu/100$
$kT = \mu/10$
$kT = \mu/2$
$kT = \mu$
0
1
2
3
4
5
h_{ij}
ϵ/μ

Generative model for realistic complex networks

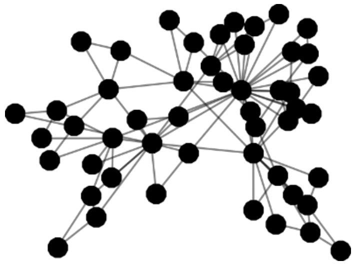
Geometry



PSO model

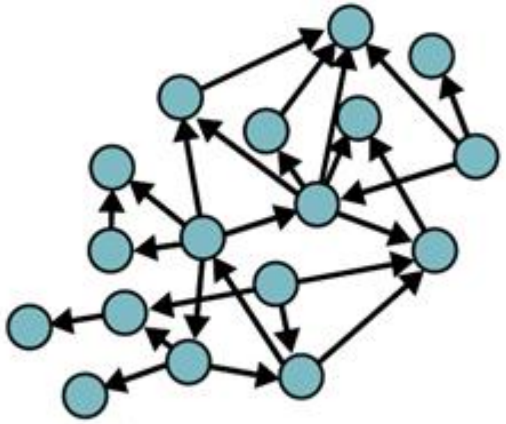


Topology



Synthetic network with:
- Clustering
- Small-word
- Scale-free

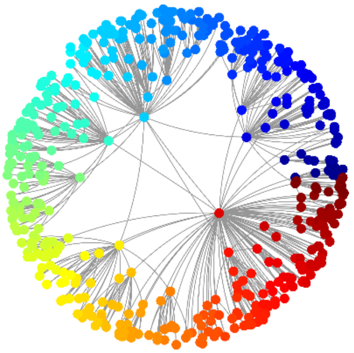
Problem 1
(inverse problem)

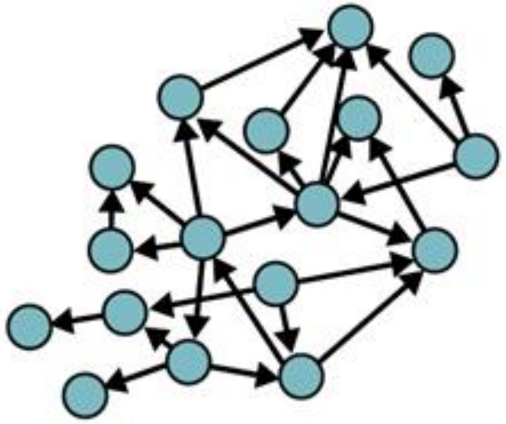


Given the network topology (just connectivity)
can we reverse the location of its nodes on the manifold?

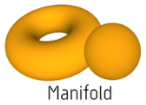
Geometry (Hyperbolic)

Network





inverse problem

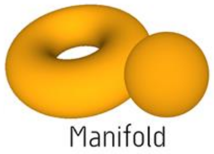


Given the network topology (just connectivity)
can we reverse the location of its nodes on the manifold?

Inverse problem in real networks

Geometry

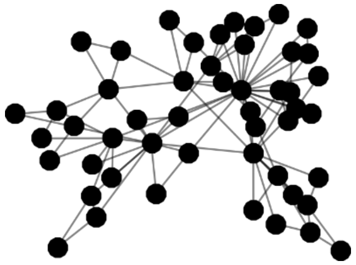
?



? Angular coordinates = node similarity ?

Radial coordinate ok!

Topology



real network

?

HyperMap (2015)

(Model-based)

- **Maximum likelihood estimation**

- Infer the coordinates maximizing the likelihood that the network has been generated by the PSO model (not community organization)

IEEE/ACM TRANSACTION ON NETWORKING

198 IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 23, NO. 1, FEBRUARY 2015

Network Mapping by Replaying Hyperbolic Growth

Fragkiskos Papadopoulos, Constantinos Psomas, and Dmitri Krioukov

Many limitations:

- **Time complexity $O(N^3)$ – $O(N^4)$**
- **Unweighted networks**
- **Only 2D-space**
- **Not community organization**

IEEE/ACM TRANSACTION ON NETWORKING

198 IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 23, NO. 1, FEBRUARY 2015

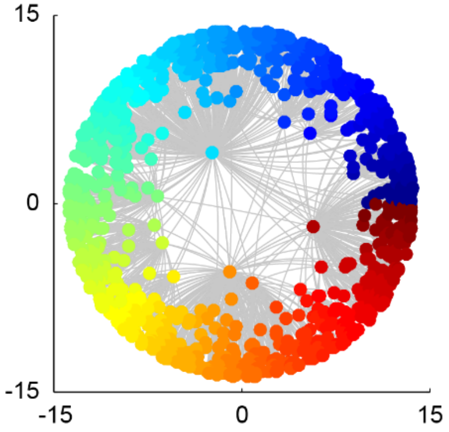
Network Mapping by Replaying Hyperbolic Growth

Fragkiskos Papadopoulos, Constantinos Psomas, and Dmitri Krioukov

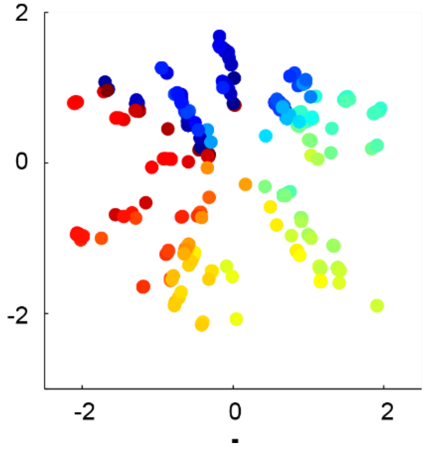
Intuition (2012)

using nonlinear dimension reduction unsupervised machine learning methods

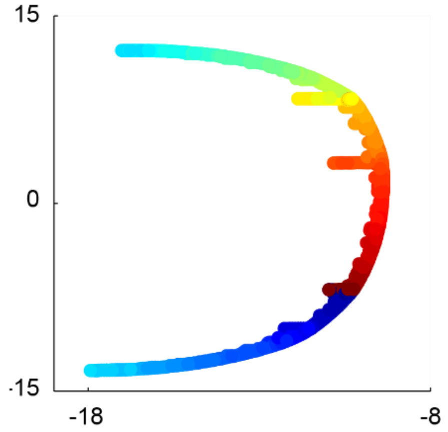
Original Network



Isomap
(manifold embedding)



MCE
(hierarchical embedding)



Why did I generate this intuition?

Modeling interactome: scale-free or geometric?

Natasa Pržulj, DG Corneil, I Jurisica, **Bioinformatics** 20 (18), **2004**

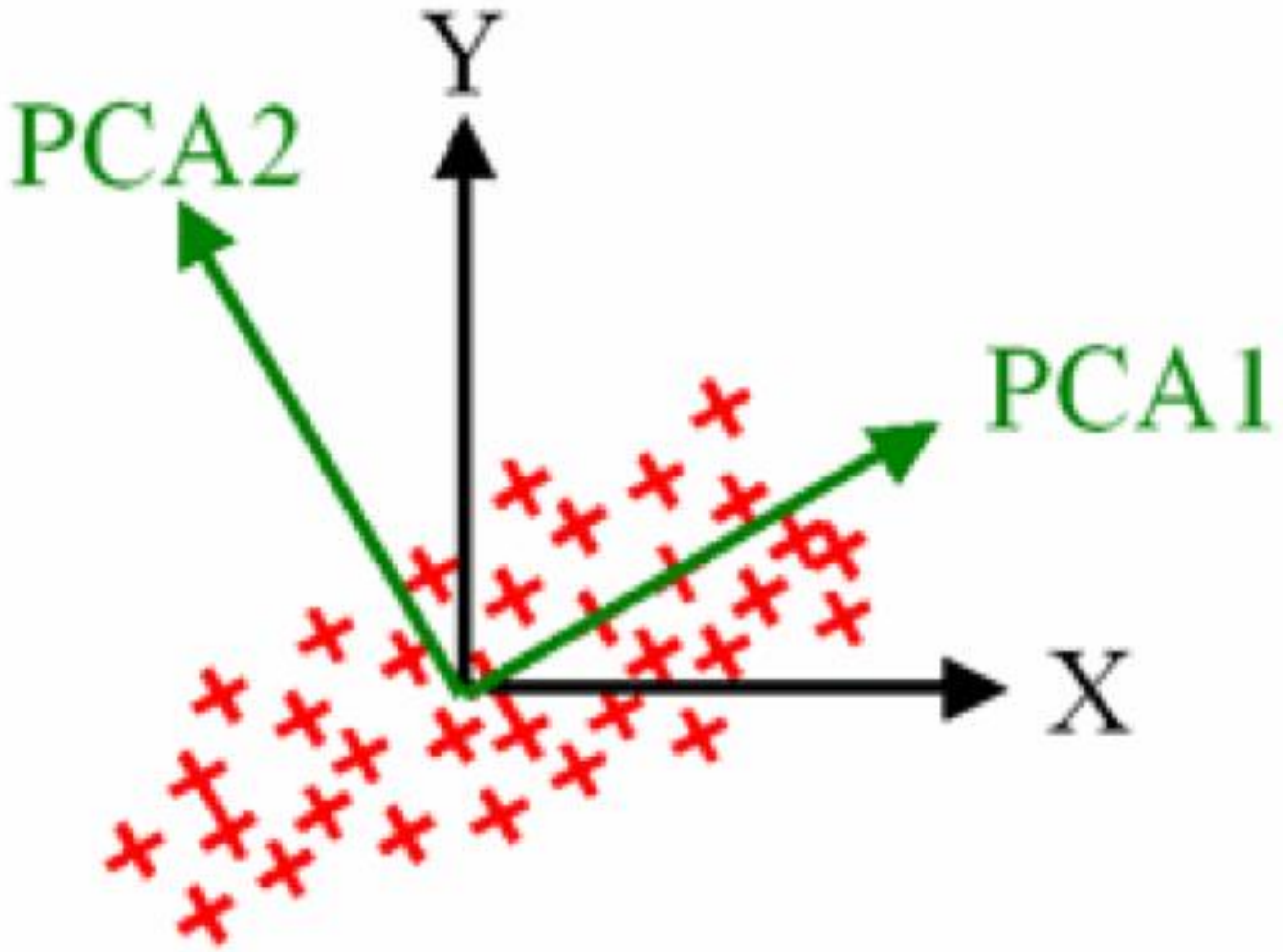
Fitting a geometric graph to a protein–protein interaction network

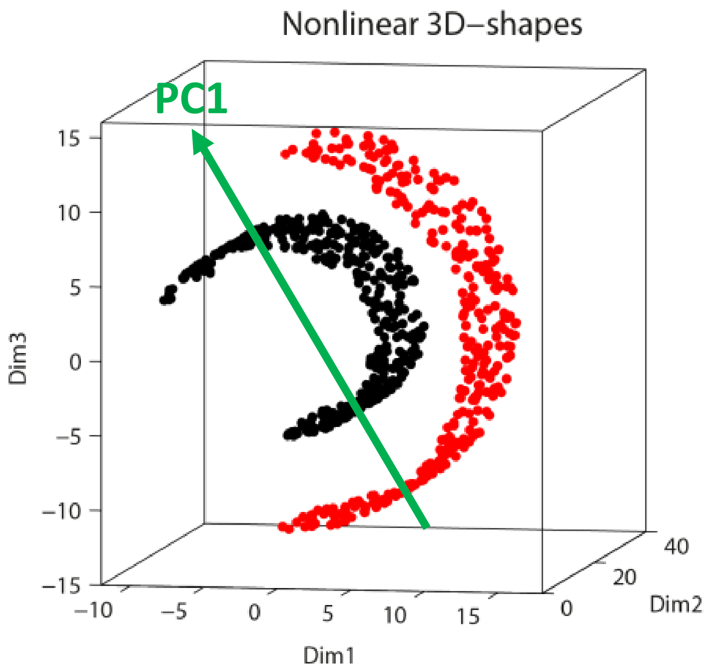
DJ Higham, M Rašajski, **Natasa Pržulj, Bioinformatics** 24 (8), **2008**

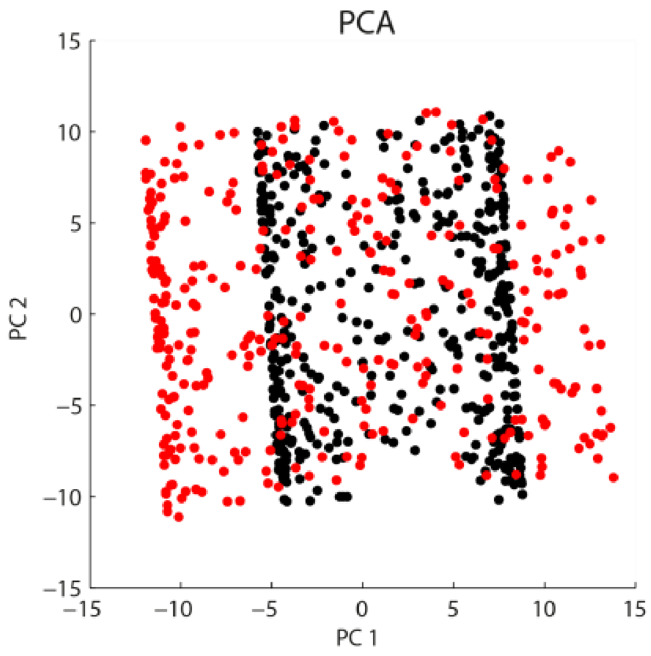
Network Biology was already discussing about this issue

How to address problems of data nonlinearity

Principal component analysis



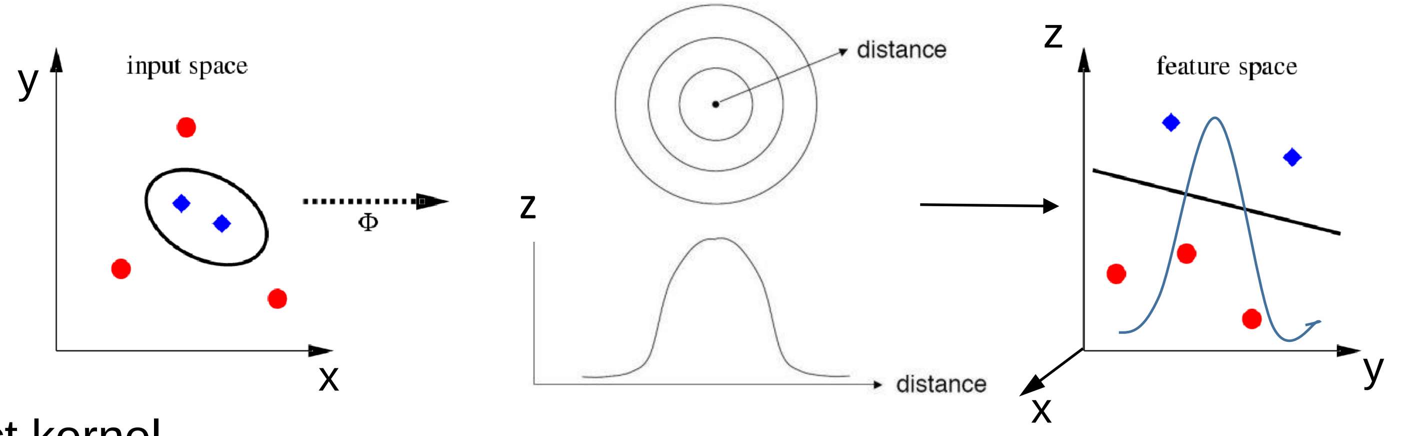




Nonlinear Dimension Reduction

- **Kernel** based

*{example: **Gaussian-PCA**}*

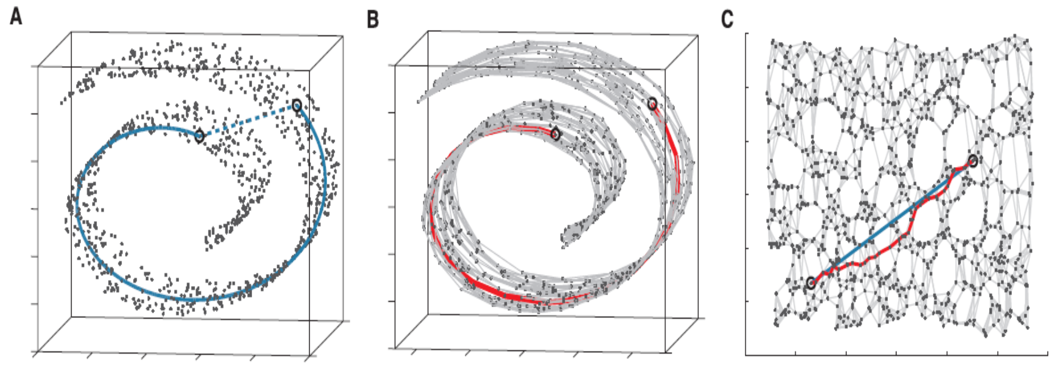


Issue 1: difficulty to know the correct kernel

Issue 2: presence of **free parameters to tune**!!!

Manifold based

*{example: **Isomap**}*



Tenenbaum et al. – Science, 2000

Issue 1: Hypothesis of **local continuity** of the manifold

Issue 2: presence of **free parameters to tune**!!!

General principles of organization of complex system

The inspiration (2008)

ARTICLES
PUBLISHED ONLINE: 16 NOVEMBER 2008 | DOI: 10.1038/NPHYS1130

nature physics

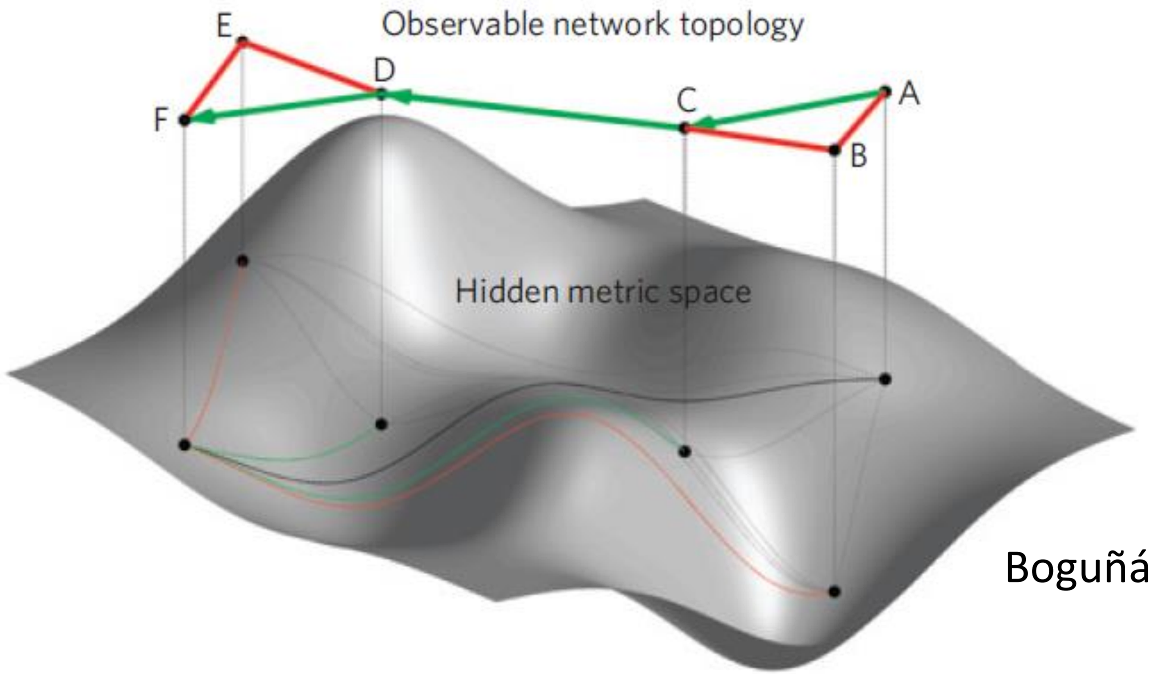
Navigability of complex networks

Marián Boguñá[1*], Dmitri Krioukov[2] and K. C. Claffy[2]

Routing information through networks is a universal phenomenon in both natural and man-made complex systems. When each node has full knowledge of the global network connectivity, finding short communication paths is merely a matter of distributed computation. However, in many real networks, nodes communicate efficiently even without such global intelligence. Here, we show that the peculiar structural characteristics of many complex networks support efficient communication without global knowledge. We also describe a general mechanism that explains this connection between network structure and function. This mechanism relies on the presence of a metric space hidden behind an observable network. Our findings suggest that real networks in nature have underlying metric spaces that remain undiscovered. Their discovery should have practical applications in a wide range of areas where networks are used to model complex systems.

The inspiration (greedy navigability)

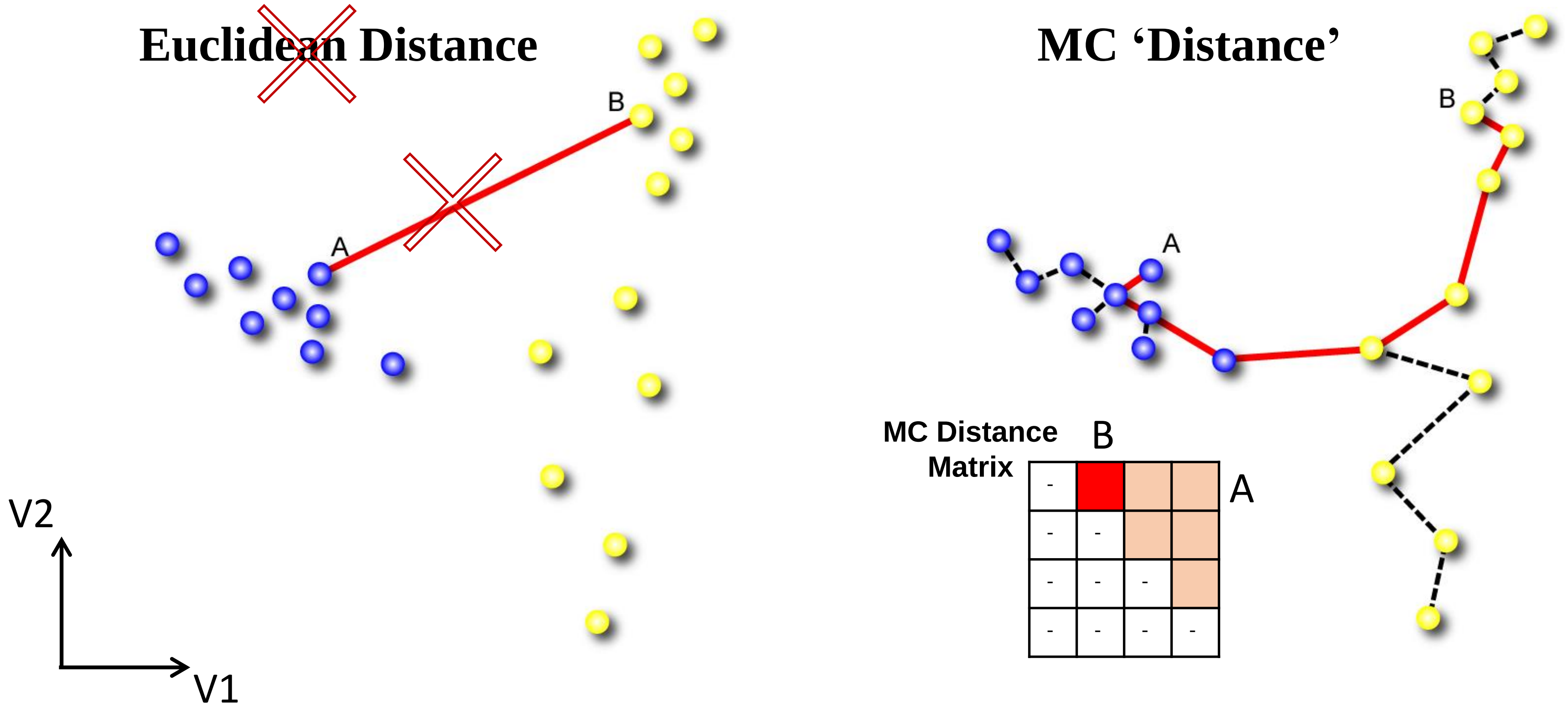
Important message
Navigating on the network
is like navigating in the valley
and hills of the hidden
geometrical space



Boguñá et al., *Nature Physics* 2008

The **observed topological** properties arise from
a **hidden geometric space** underlying the network

How MC works: *Navigating between the points with a greedy routing process: the minimum spanning tree (MST)!*

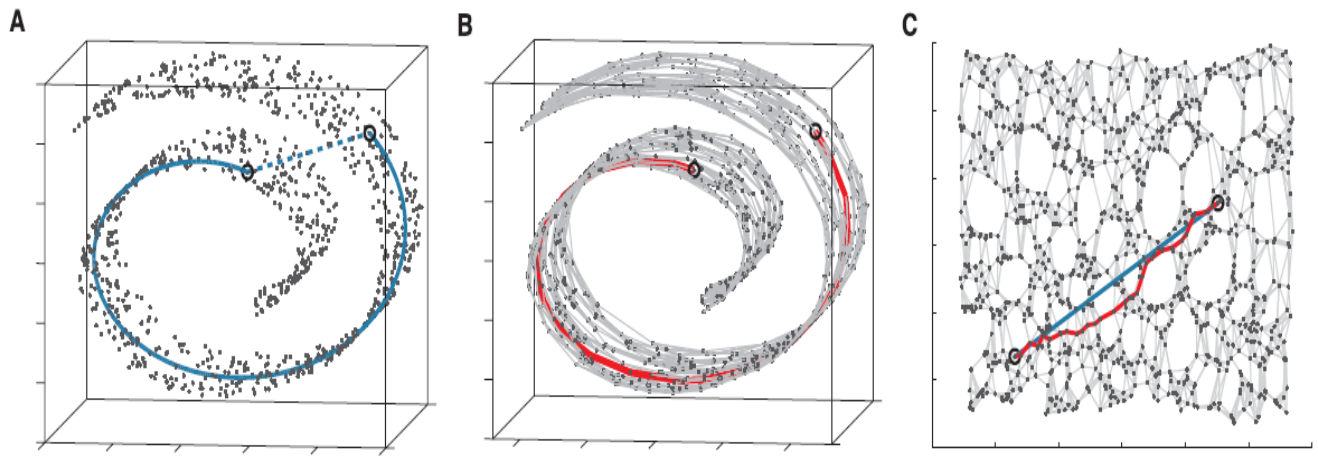


The greedy routing navigability is a way to map the hidden nonlinear topology

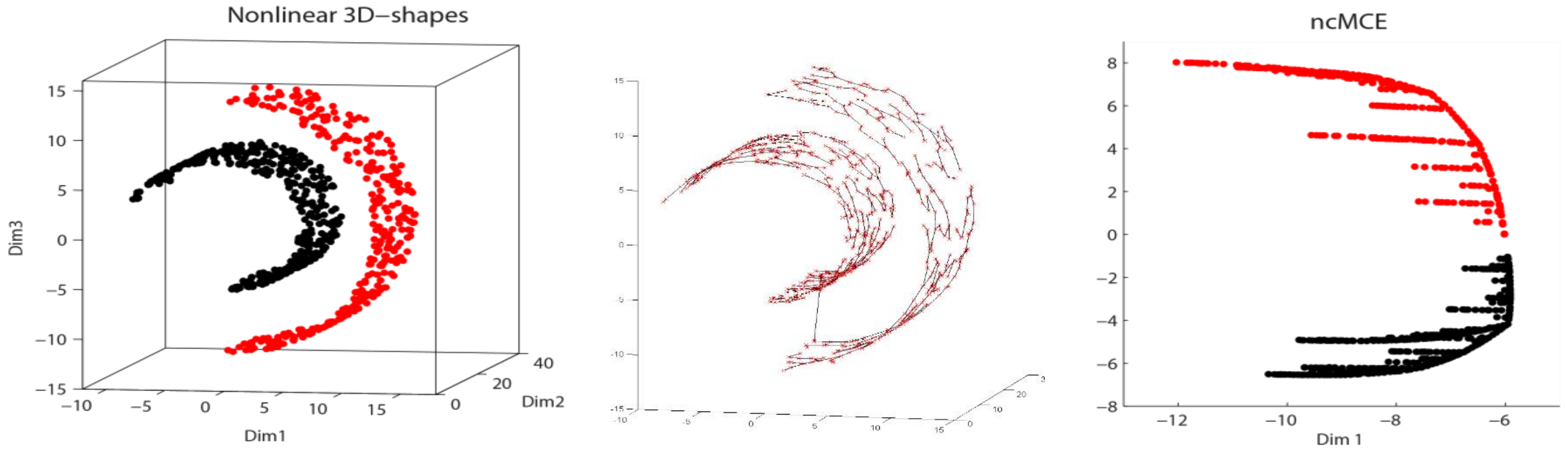
For MC: The *global mapping* **and the** *local fitting* **are reciprocally dependent**

MC *Minimize globally and fit locally!*

Nonlinear topological-based Dimension Reduction

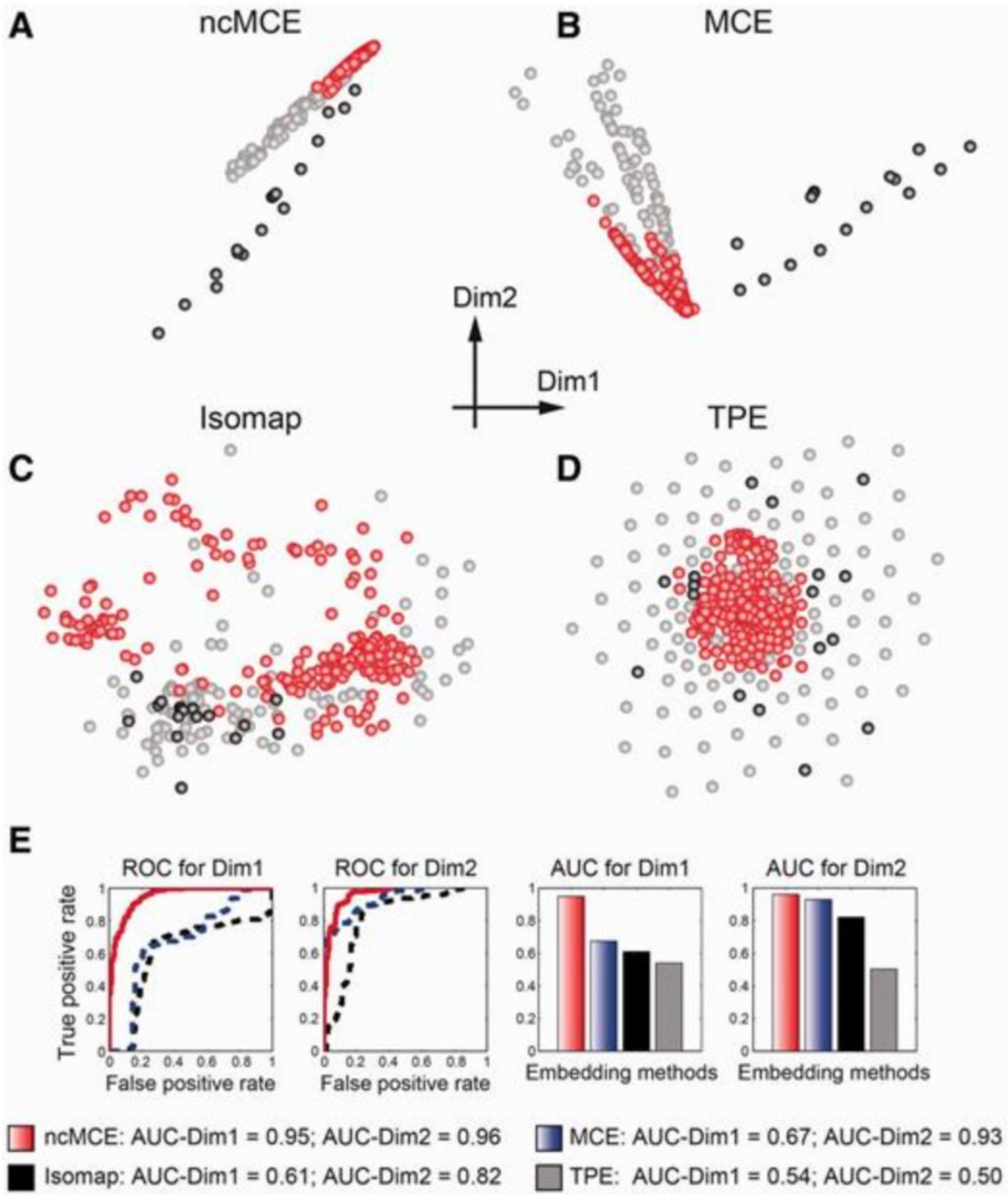


Isomap
(manifold embedding)



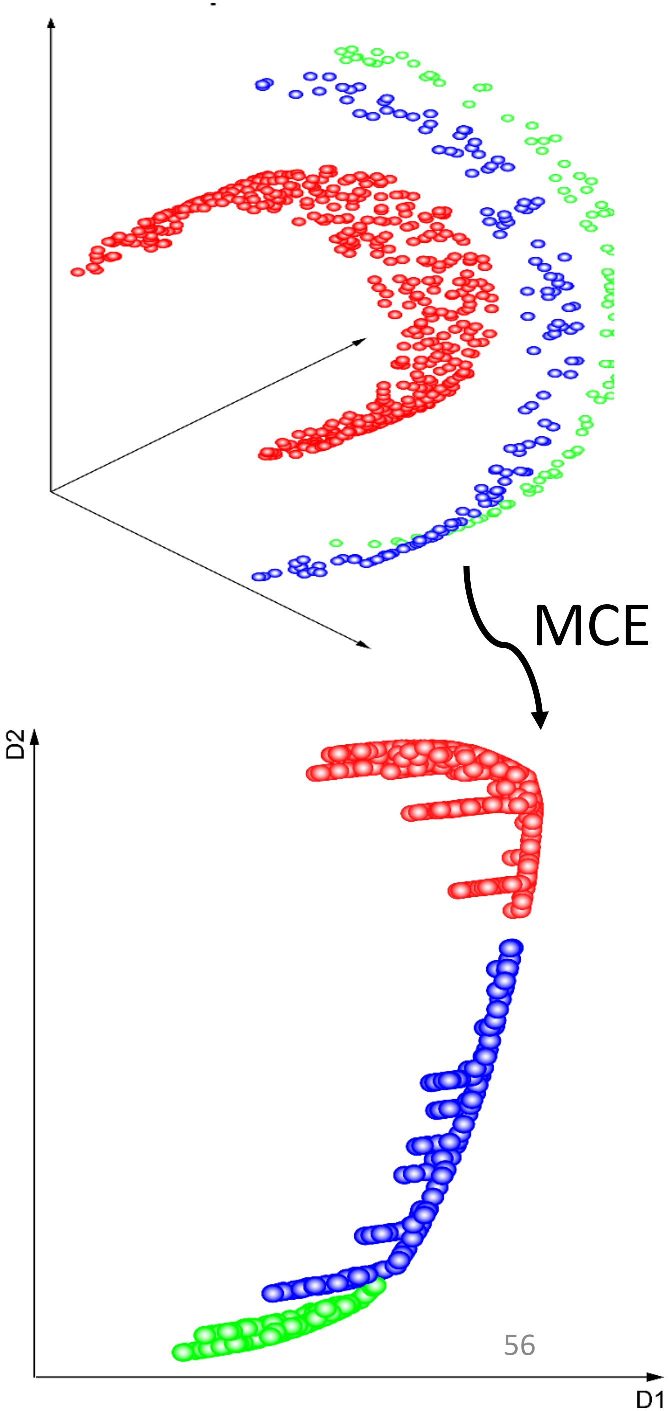
Minimum Curvilinear embedding (MCE)
(hierarchical embedding)

Radar Signal Dataset



The crescent-moon obsession

Similarity ordering of the samples !



Direction of maximum MC nonlinear similarity in the multidimensional space

ECCB 2010, Ghent, Belgium

BIOINFORMATICS

Vol. 26 ECCB 2010, pages i531–i539
doi:10.1093/bioinformatics/btq376

Nonlinear dimension reduction and clustering by Minimum Curvilinearity unfold neuropathic pain and tissue embryological classes

Carlo Vittorio Cannistraci[1,2,3,4,5,*], Timothy Ravasi[1,5], Franco Maria Montevecchi[3], Trey Ideker[5] and Massimo Alessio[2,*]

[1]Red Sea Integrative Systems Biology Lab, Computational Bioscience Research Center, Division of Chemical and Life Sciences and Engineering, King Abdullah University for Science and Technology (KAUST), Jeddah, Kingdom of Saudi Arabia, [2]Proteome Biochemistry, San Raffaele Scientific Institute, via Olgettina 58, 20132 Milan, [3]Department of Mechanics, [4]CMP Group, Microsoft Research, Politecnico di Torino, c/so Duca degli Abruzzi 24, 10129 Turin, Italy, [5]Department of Bioengineering and Department of Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093 USA

ISMB/ECCB 2013 , Berlin, Germany, July 2013

BIOINFORMATICS

Vol. 29 ISMB/ECCB 2013, pages i199–i209
doi:10.1093/bioinformatics/btt208

Minimum curvilinearity to enhance topological prediction of protein interactions by network embedding

Carlo Vittorio Cannistraci[1,2,*,†], Gregorio Alanis-Lobato[1,2,†] and Timothy Ravasi[1,2,*]

[1]Integrative Systems Biology Laboratory, Biological and Environmental Sciences and Engineering Division, Computer Electrical and Mathematical Sciences and Engineering Division, Computational Bioscience Research Center, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Kingdom of Saudi Arabia and [2]Division of Medical Genetics, Department of Medicine, University of California, San Diego, CA 92093-0688, USA

ABSTRACT

Motivation: Most functions within the cell emerge thanks to protein–protein interactions (PPIs), yet experimental determination of PPIs is both expensive and time-consuming. PPI networks present significant levels of noise and incompleteness. Predicting interactions using only PPI-network topology (topological prediction) is difficult but essential when prior biological knowledge is absent or unreliable.

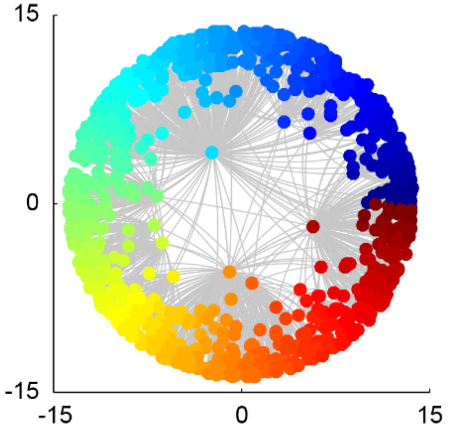
1 INTRODUCTION

Detection of new interactions between proteins is central to modern biology. Its application in protein function prediction, drug delivery control and disease diagnosis has developed alongside a deeper understanding of the processes that occur within the cell. One key task in systems biology is the experimental detection of new protein–protein interactions (PPIs). However, such experiments are time consuming and expensive. Because of

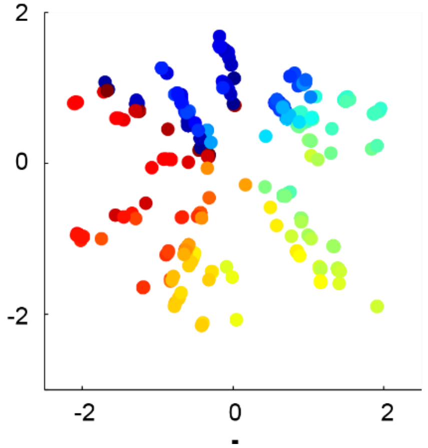
Intuition (2012)

using nonlinear dimension reduction unsupervised machine learning methods

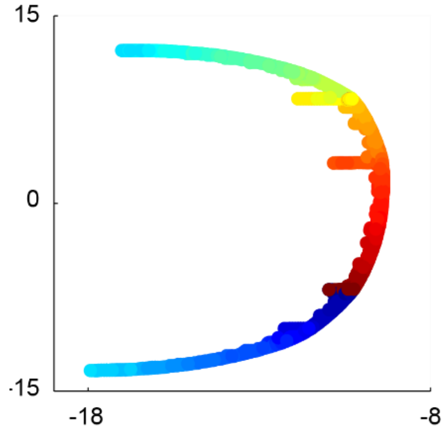
Original Network



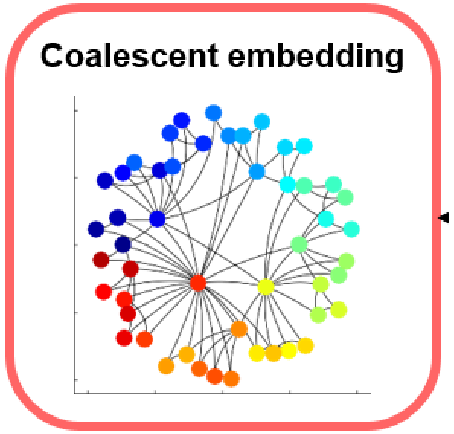
Isomap
(manifold embedding)



MCE
(hierarchical embedding)



Pre-weighting rules



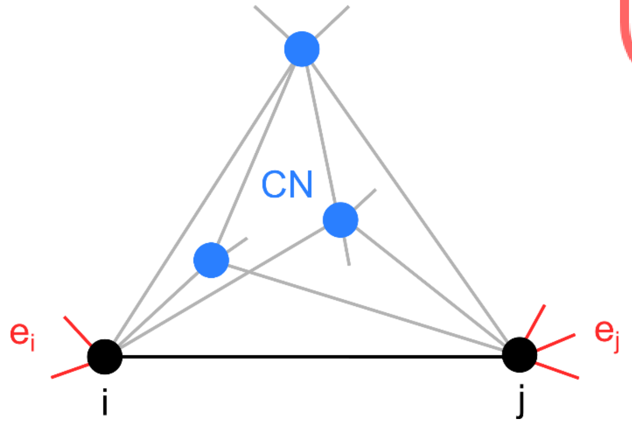
RA Distance Matrix

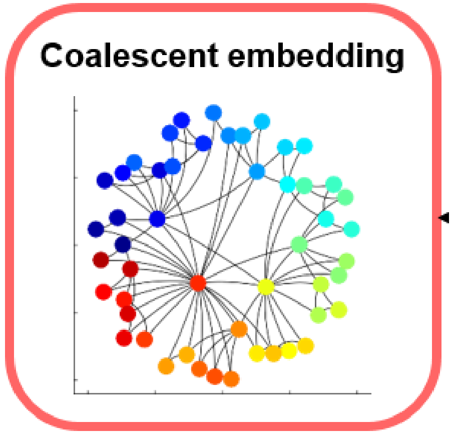
	B		
-			
-	-		
-	-	-	
-	-	-	-

A

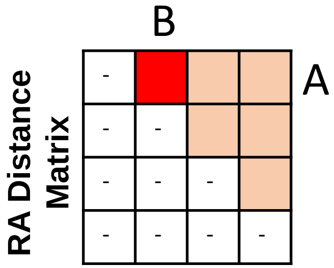
- Repulsion-Attraction (RA) - **local**

$$x_{ij}^{RA} = \frac{1 + e_i + e_j + e_i e_j}{1 + CN_{ij}}$$



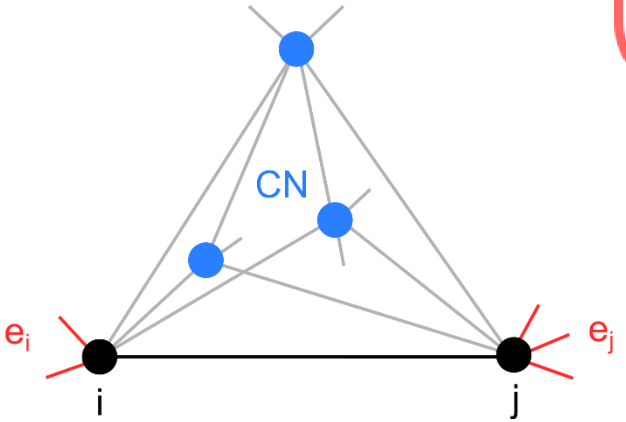


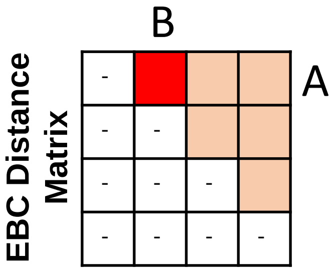
Pre-weighting rules



- Repulsion-Attraction (RA) - **local**

$$x_{ij}^{RA} = \frac{1 + e_i + e_j + e_i e_j}{1 + CN_{ij}}$$



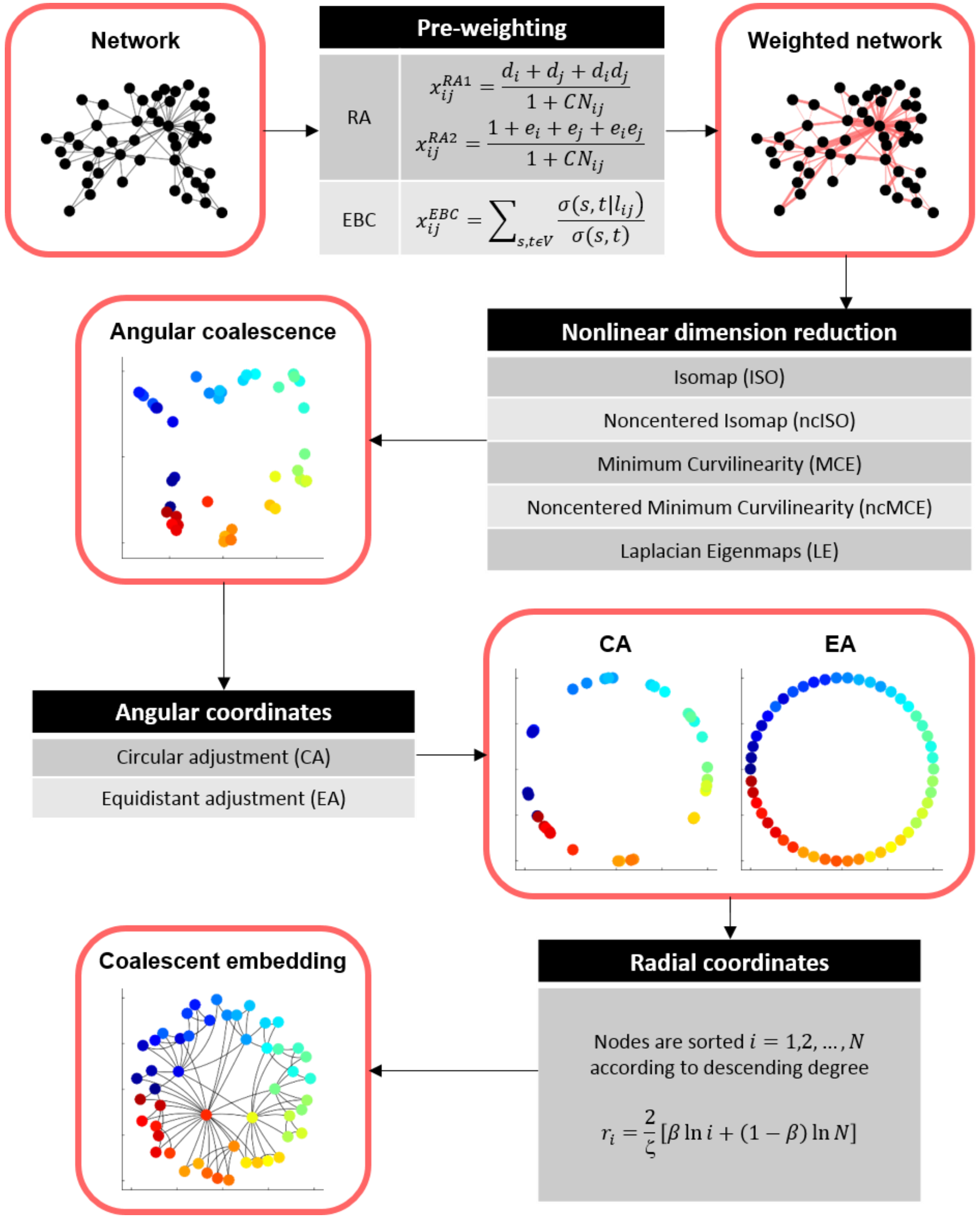


- Edge-Betweenness-Centrality (EBC) – **global**

$$x_{ij}^{EBC} = \sum_{s,t \epsilon V} \frac{\sigma(s,t|l_{ij})}{\sigma(s,t)}$$

Ginestra Bianconi

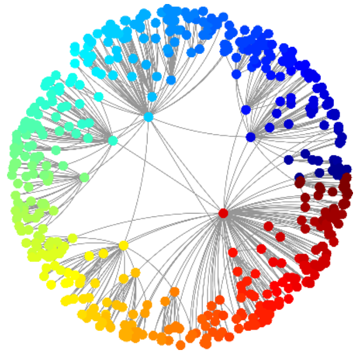
Coalescent Embedding (2017) - *(Model-free)*



Results

Testing on the benchmark

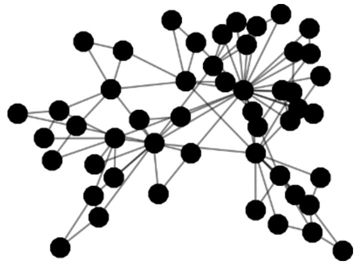
Geometry



PSO model

network generation →

Topology



Synthetic network

network embedding →

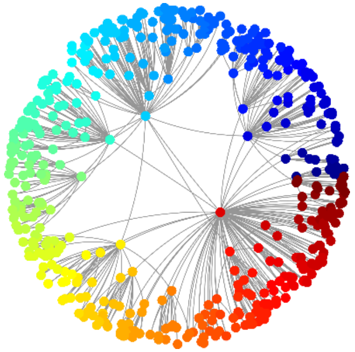
Geometry

?

Angular coordinates
Radial coordinate

Testing on the benchmark

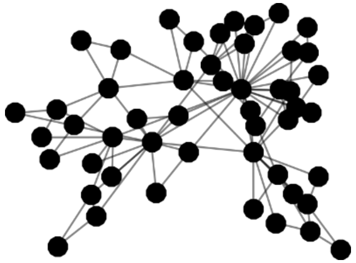
Geometry



PSO model

network generation

Topology



Synthetic network

network embedding

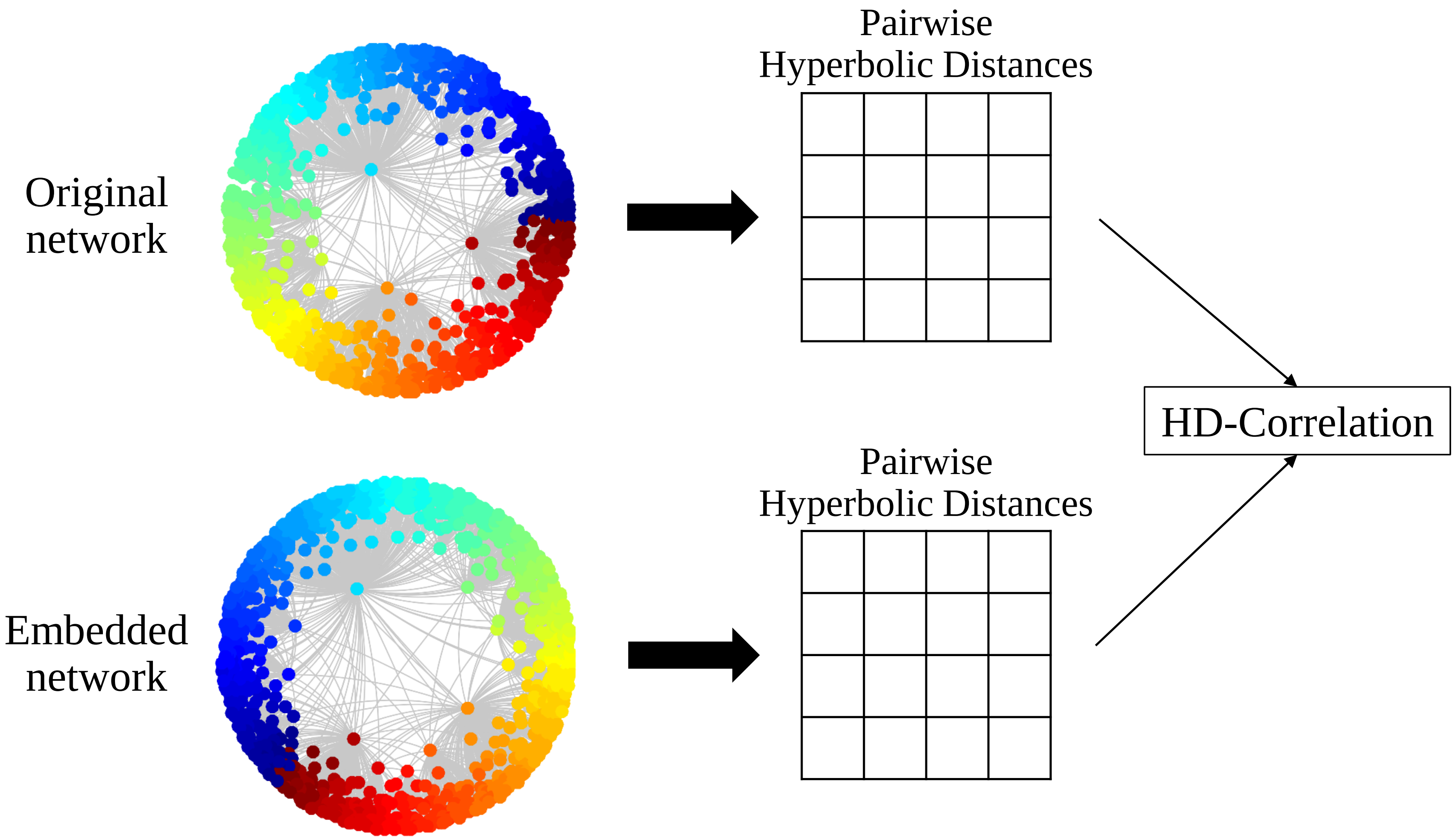
Geometry

?

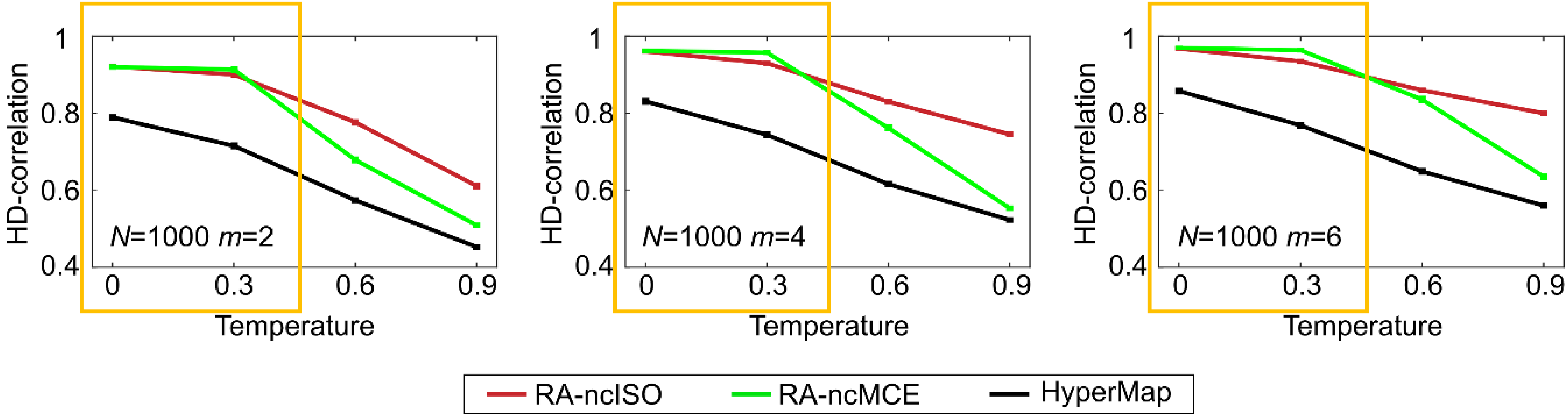
Angular coordinates
Radial coordinate

Compare to
the reference

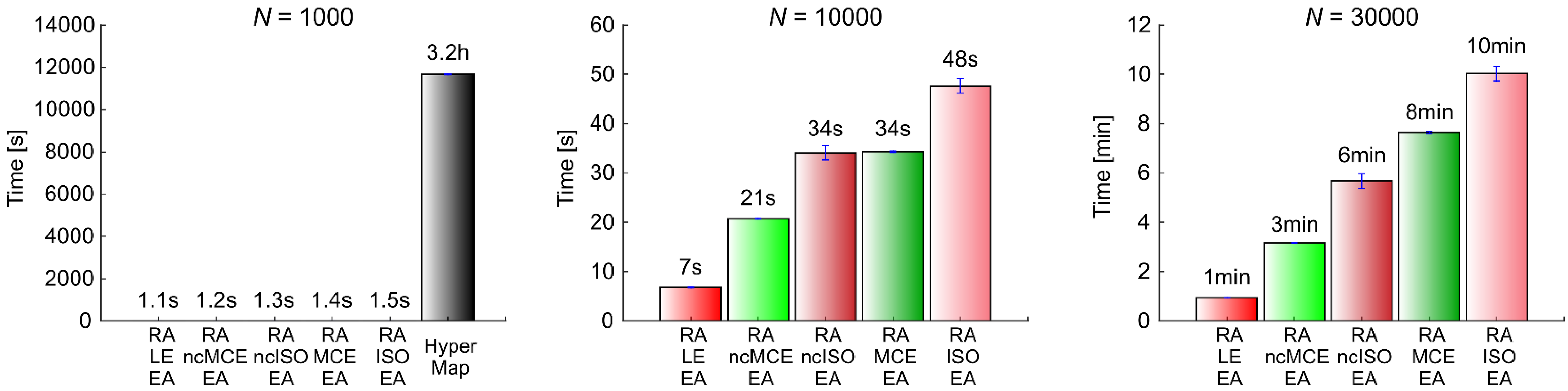
Evaluation on PSO networks



Evaluation on PSO networks: embedding quality

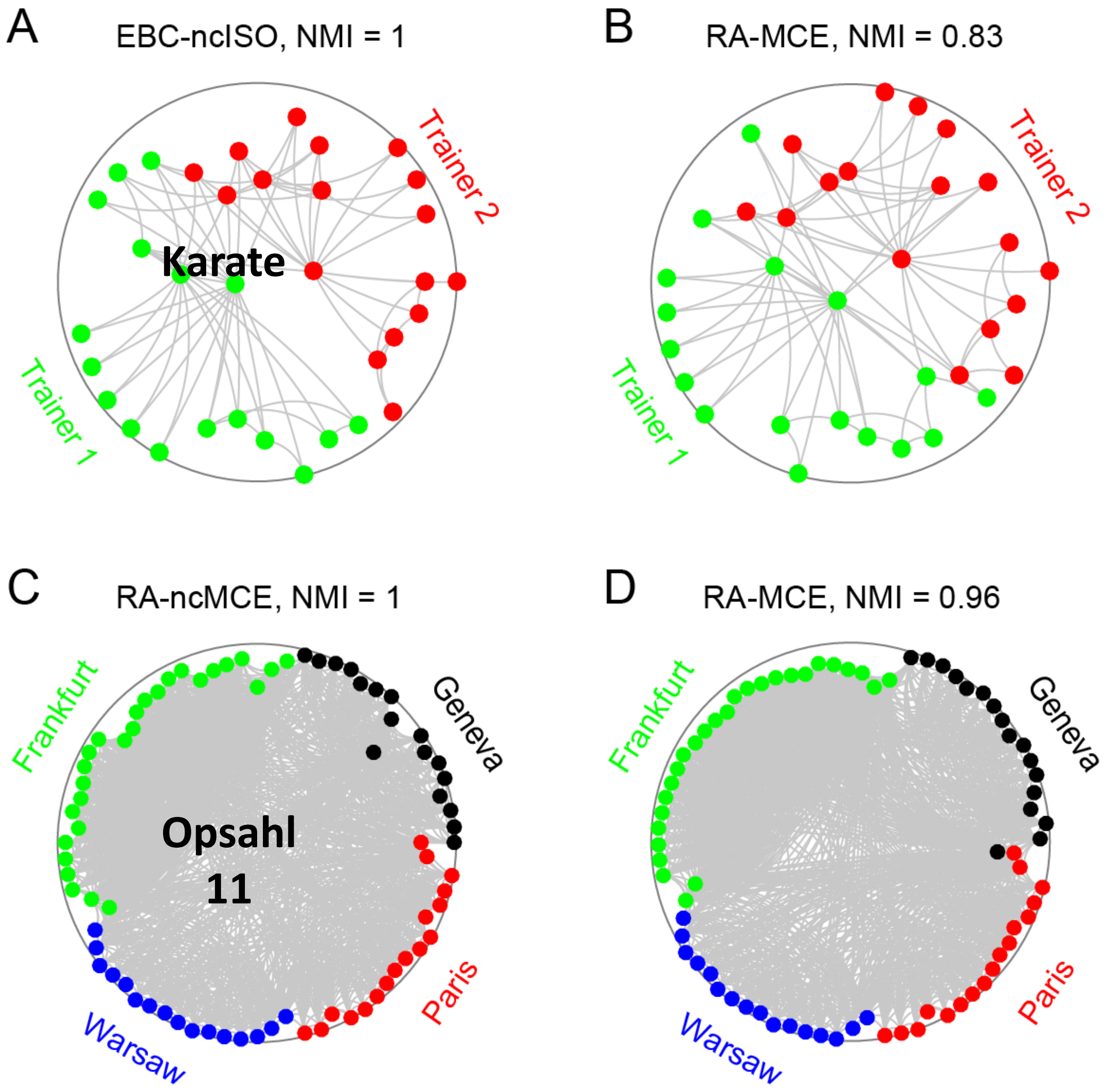


Evaluation on PSO networks: computational time

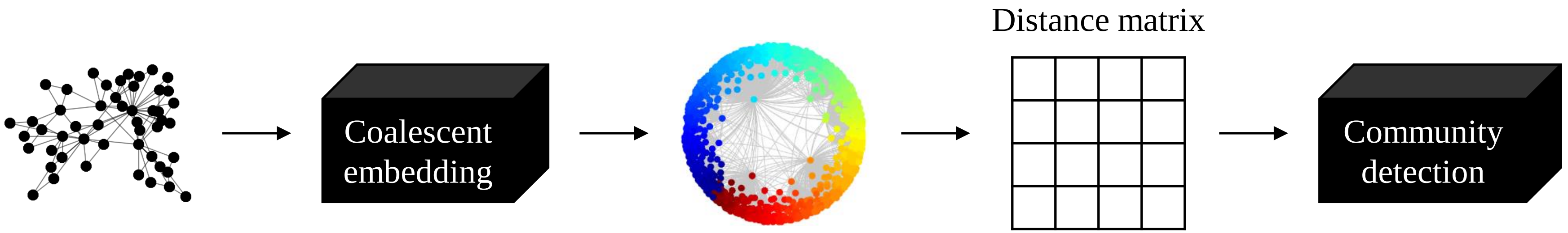


Time complexity ~ **$O(N^2)$**

Community detection

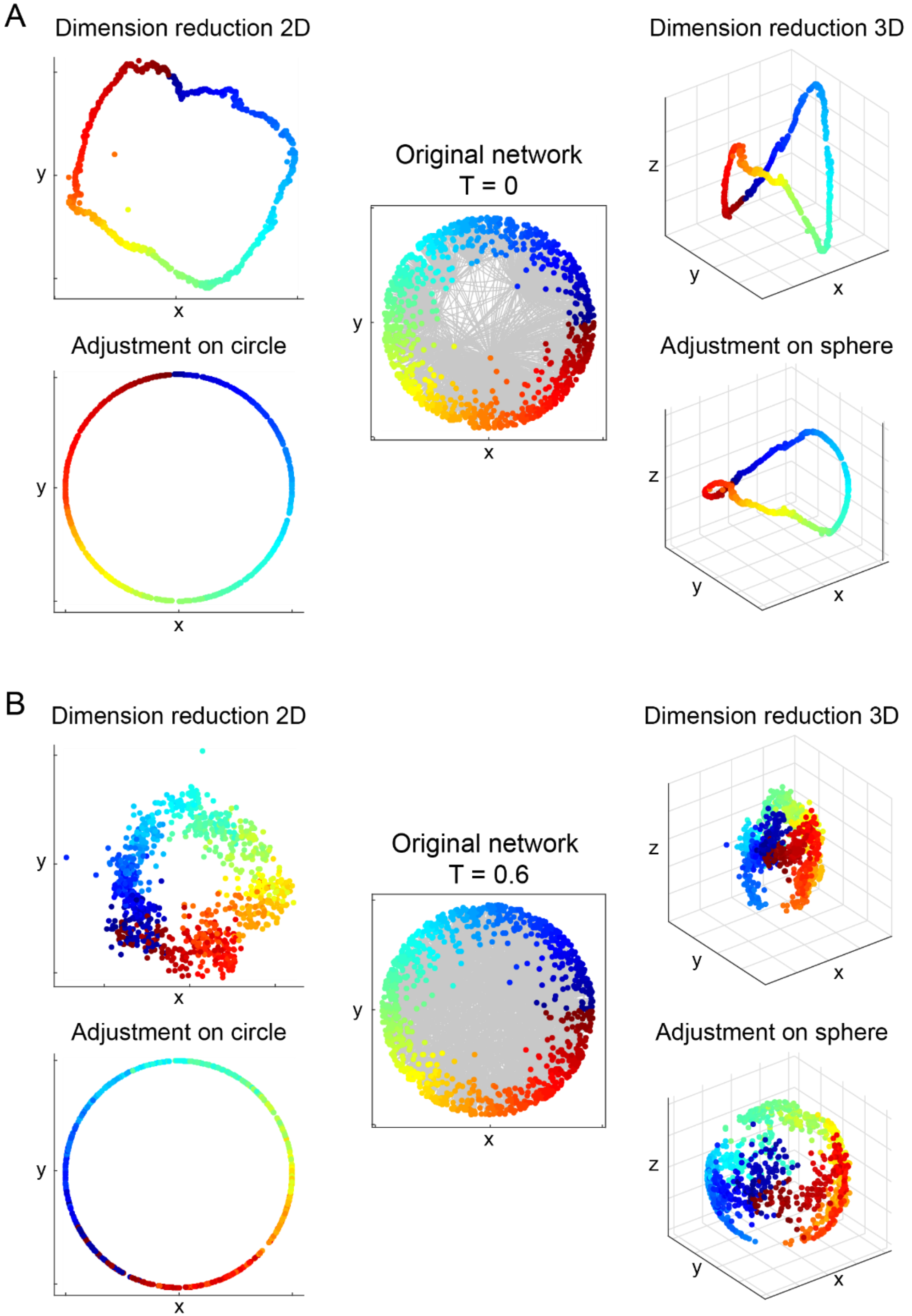


Community detection
(hyperbolic Louvain)



Method	Karate	Opsahl 8	Opsahl 9	Opsahl 10	Opsahl 11	Polbooks	Football	Polblogs	Mean NMI
Coalescent Embedding	1.00	0.57	0.47	1.00	0.93	0.59	0.90	0.68	**0.77**
Original algorithm	0.46	0.55	0.41	1.00	0.96	0.50	0.93	0.64	0.68
HyperMap mapping	0.56	0.60	0.28	0.92	0.85	0.50	0.83	0.69	0.65

Embedding in 3D



Conclusions on Coalescent Embedding

- ‘$O(n^3)$ or $O(n^4)$’ → **$O(n^2)$**
- Unweighted networks → **Weighted networks**
- Only 2D embedding → **also 3D or any dimensional space**
- **Community detection in the hyperbolic space algorithms**
- **Not hyperparameters to tune!**

November 2017

nature COMMUNICATIONS

ARTICLE

DOI: 10.1038/s41467-017-01825-5 OPEN

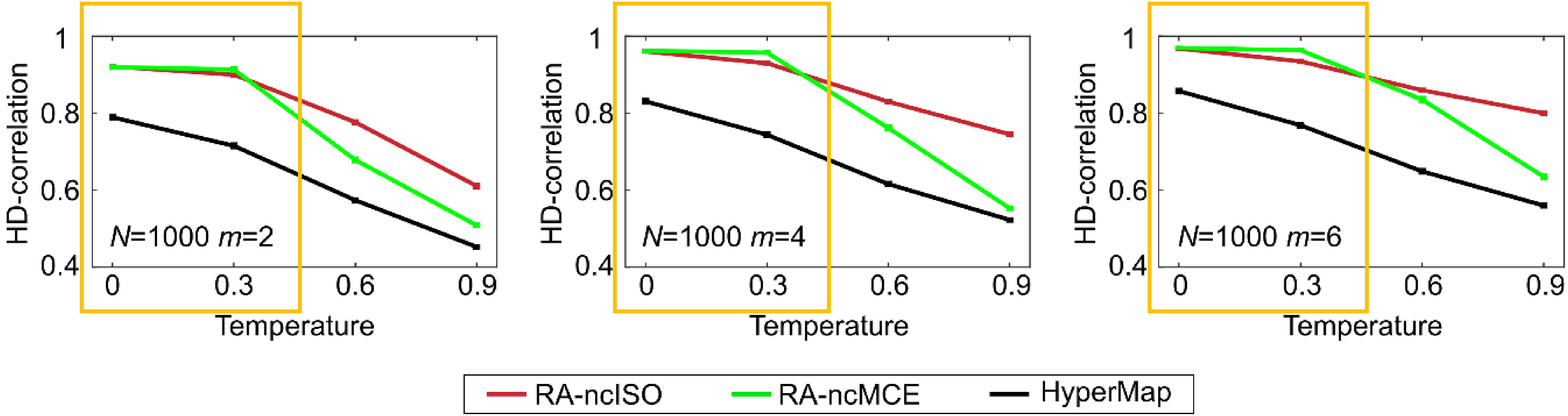
Machine learning meets complex networks via coalescent embedding in the hyperbolic space

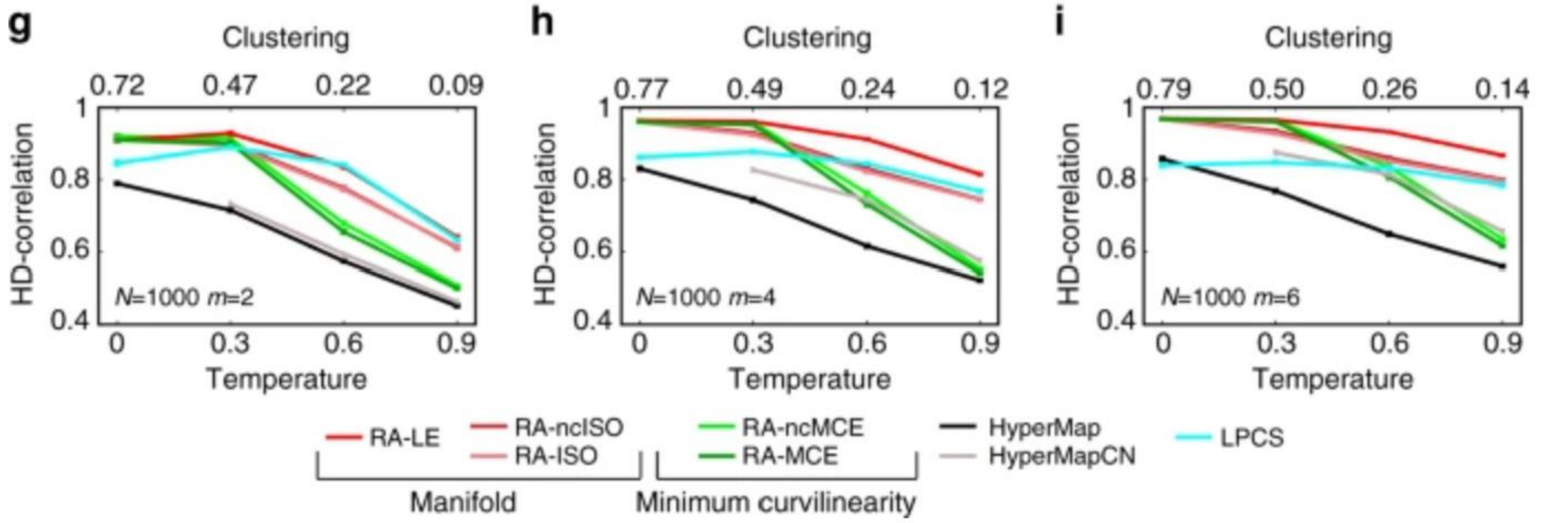
Alessandro Muscoloni[1], Josephine Maria Thomas[1], Sara Ciucci[1,2], Ginestra Bianconi[3] & Carlo Vittorio Cannistraci[1,4]

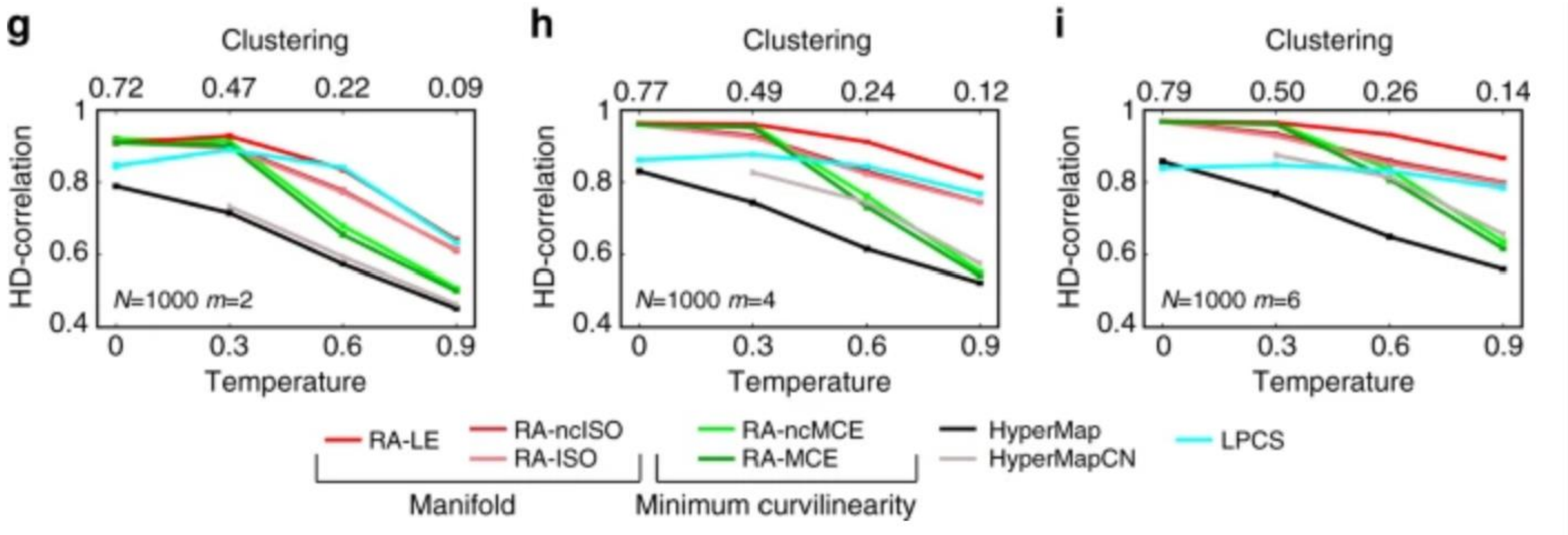
Physicists recently observed that realistic complex networks emerge as discrete samples from a continuous hyperbolic geometry enclosed in a circle: the radius represents the node centrality and the angular displacement between two nodes resembles their topological proximity. The hyperbolic circle aims to become a universal space of representation and analysis of many real networks. Yet, inferring the angular coordinates to map a real network back to its latent geometry remains a challenging inverse problem. Here, we show that intelligent machines for unsupervised recognition and visualization of similarities in big data can also infer the network angular coordinates of the hyperbolic model according to a geometrical organization that we term "angular coalescence." Based on this phenomenon, we propose a class of algorithms that offers fast and accurate "coalescent embedding" in the hyperbolic circle even for large networks. This computational solution to an inverse problem in physics of complex systems favors the application of network latent geometry techniques in disciplines dealing with big network data analysis including biology, medicine, and social science.

A question for you
?

Evaluation on PSO networks: embedding quality




g
h
i
Clustering
0.72 0.47 0.22 0.09
0.77 0.49 0.24 0.12
0.79 0.50 0.26 0.14
HD-correlation
N=1000 m=2
N=1000 m=4
N=1000 m=6
Temperature
RA-LE
RA-ncISO
RA-ISO
RA-ncMCE
RA-MCE
HyperMap
HyperMapCN
LPCS
Manifold
Minimum curvilinearity



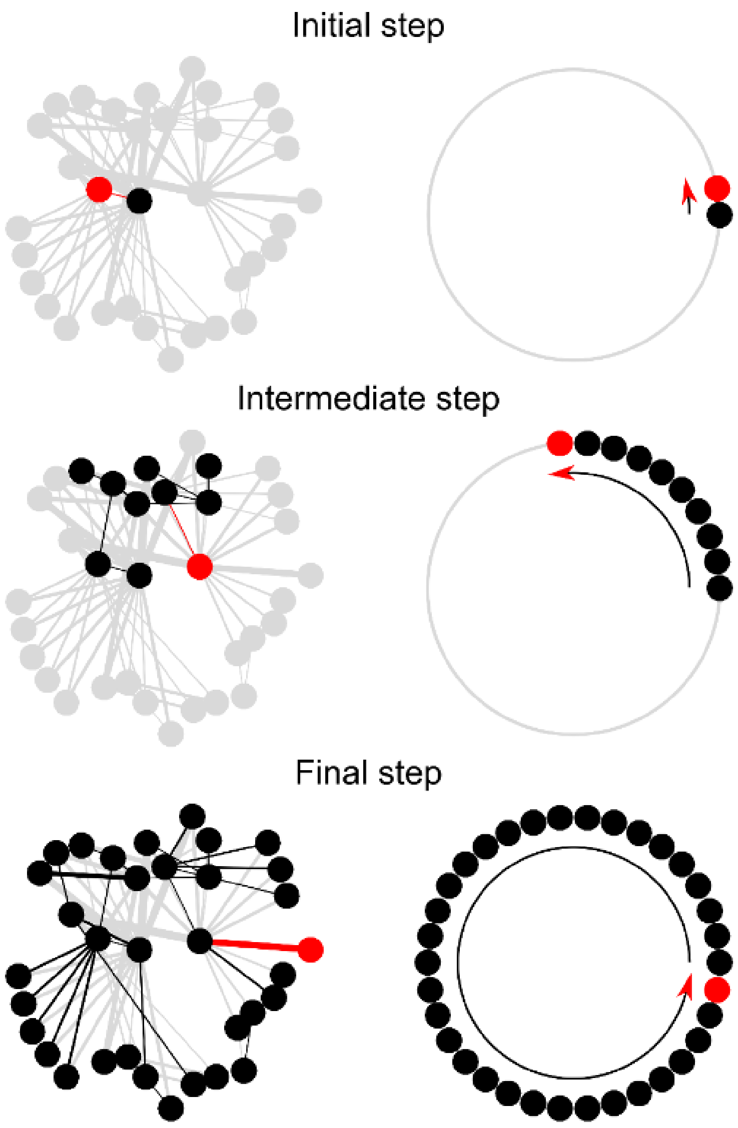
Article | Open access | Published: 06 August 2021

The inherent community structure of hyperbolic networks

Bianka Kovács & Gergely Palla

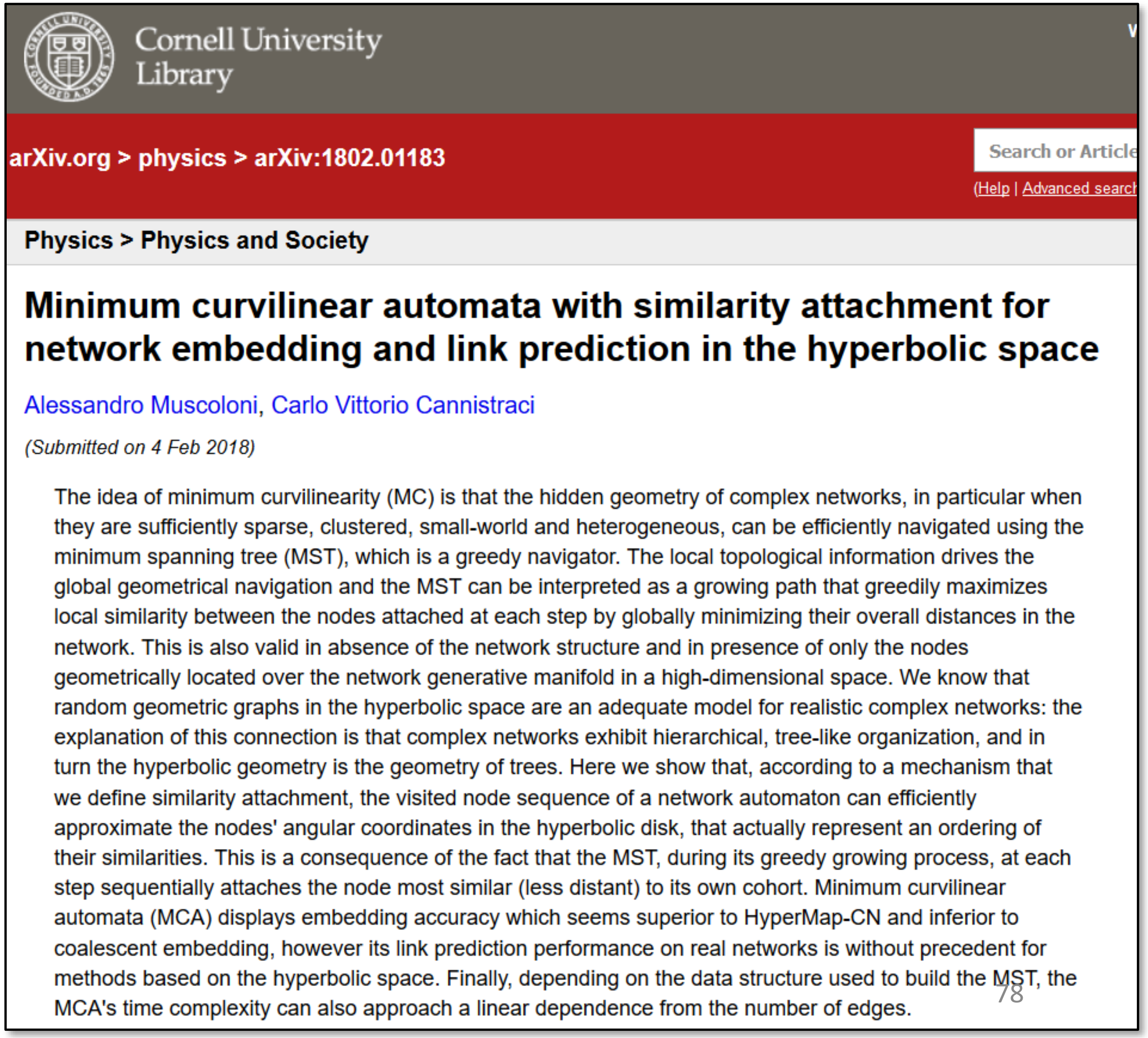
Scientific Reports **11**, Article number: 16050 (2021) | Cite this article

Mechanism of similarity attachment



Network automata solution

Time complexity ~ O(E)



Cornell University Library

arXiv.org > physics > arXiv:1802.01183

Search or Article

(Help | Advanced search

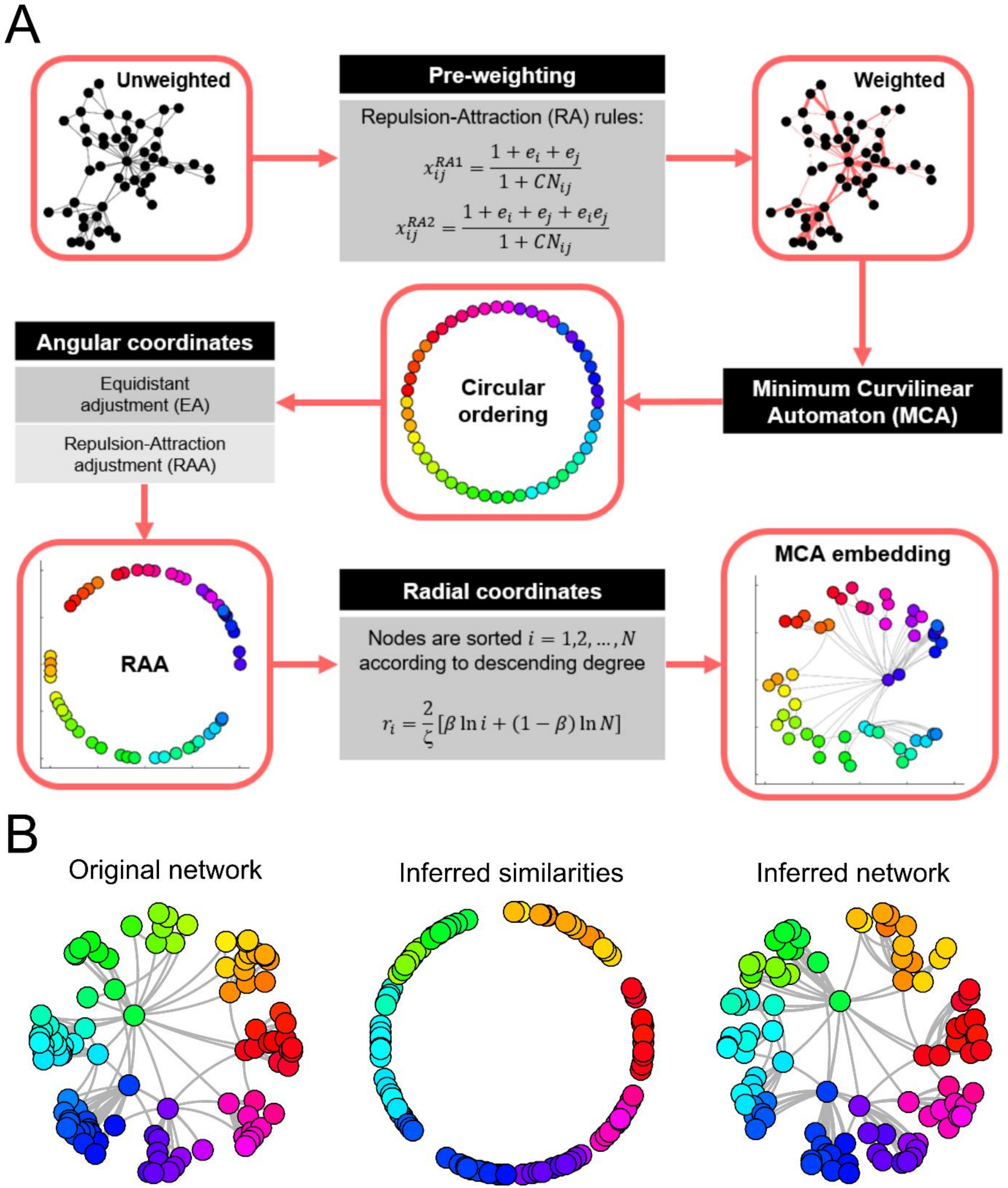
Physics > Physics and Society

Minimum curvilinear automata with similarity attachment for network embedding and link prediction in the hyperbolic space

Alessandro Muscoloni, Carlo Vittorio Cannistraci

(Submitted on 4 Feb 2018)

The idea of minimum curvilinearity (MC) is that the hidden geometry of complex networks, in particular when they are sufficiently sparse, clustered, small-world and heterogeneous, can be efficiently navigated using the minimum spanning tree (MST), which is a greedy navigator. The local topological information drives the global geometrical navigation and the MST can be interpreted as a growing path that greedily maximizes local similarity between the nodes attached at each step by globally minimizing their overall distances in the network. This is also valid in absence of the network structure and in presence of only the nodes geometrically located over the network generative manifold in a high-dimensional space. We know that random geometric graphs in the hyperbolic space are an adequate model for realistic complex networks: the explanation of this connection is that complex networks exhibit hierarchical, tree-like organization, and in turn the hyperbolic geometry is the geometry of trees. Here we show that, according to a mechanism that we define similarity attachment, the visited node sequence of a network automaton can efficiently approximate the nodes' angular coordinates in the hyperbolic disk, that actually represent an ordering of their similarities. This is a consequence of the fact that the MST, during its greedy growing process, at each step sequentially attaches the node most similar (less distant) to its own cohort. Minimum curvilinear automata (MCA) displays embedding accuracy which seems superior to HyperMap-CN and inferior to coalescent embedding, however its link prediction performance on real networks is without precedent for methods based on the hyperbolic space. Finally, depending on the data structure used to build the MST, the MCA's time complexity can also approach a linear dependence from the number of edges.


A
Unweighted
Pre-weighting
Repulsion-Attraction (RA) rules:
$x_{ij}^{RA1} = \frac{1 + e_i + e_j}{1 + CN_{ij}}$
$x_{ij}^{RA2} = \frac{1 + e_i + e_j + e_i e_j}{1 + CN_{ij}}$
Weighted
Angular coordinates
Equidistant adjustment (EA)
Repulsion-Attraction adjustment (RAA)
Circular ordering
Minimum Curvilinear Automaton (MCA)
RAA
Radial coordinates
Nodes are sorted $i = 1,2,\ldots,N$ according to descending degree
$r_i = \frac{2}{\zeta}[\beta \ln i + (1-\beta)\ln N]$
MCA embedding
B
Original network
Inferred similarities
Inferred network

Video of Minimum Curvilinear automata

The trilogy

Probabilistic-model based	Model-free	Mechanistic-model based
Maximum likelihood estimation	**Nonlinear dimension reduction**	**Network Automata**
- Time complexity **$O(N^3)$ – $O(N^4)$**	- Time complexity **$O(N^2)$**	- Time complexity ~ **$O(E)$**
- Unweighted networks	- Unweighted/weighted networks	- Unweighted/weighted networks
- Only 2D-space	- Any dimension	- Only 2D-space
- Higher GR accuracy	- Higher HD accuracy	- Higher LP accuracy
Bogugna et al. 2010, Nat. com. Papadopulos et al. 2015 IEEE/ACM	Muscoloni et al. 2018, Nat. com.	Muscoloni et al. 2018, ArXiv.

Problem 2
(absence of community structure)

Popularity-Similarity-Optimization (PSO) model (2012)

Real network properties:

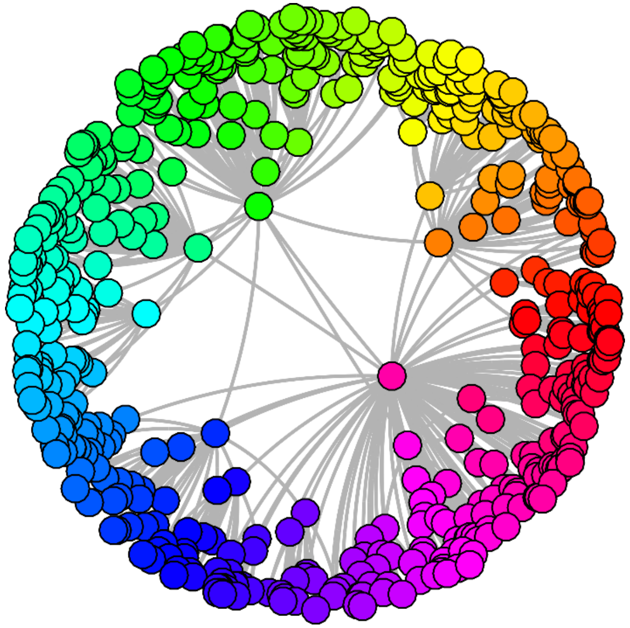
☑ Clustering

☑ Small-world

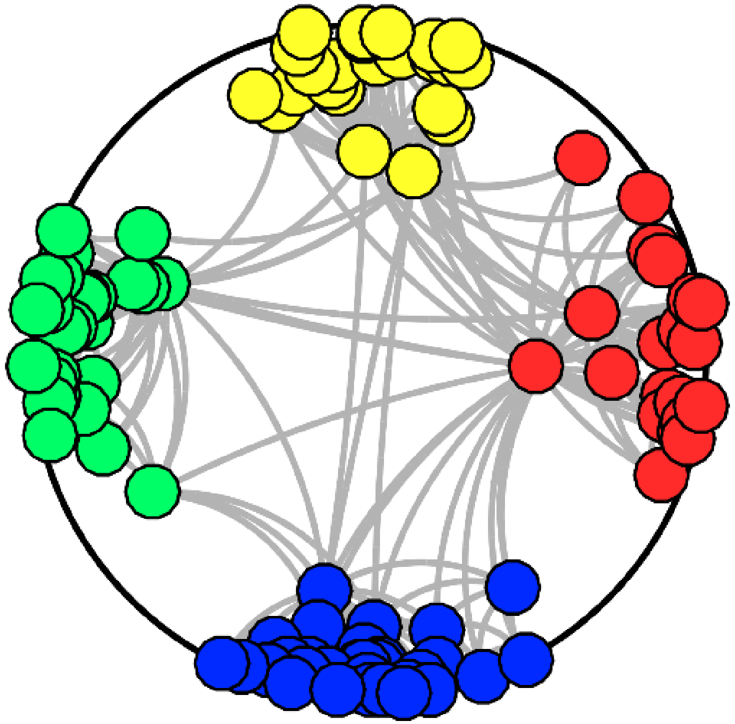
☑ Scale-free

– controllable ✗ Community structure

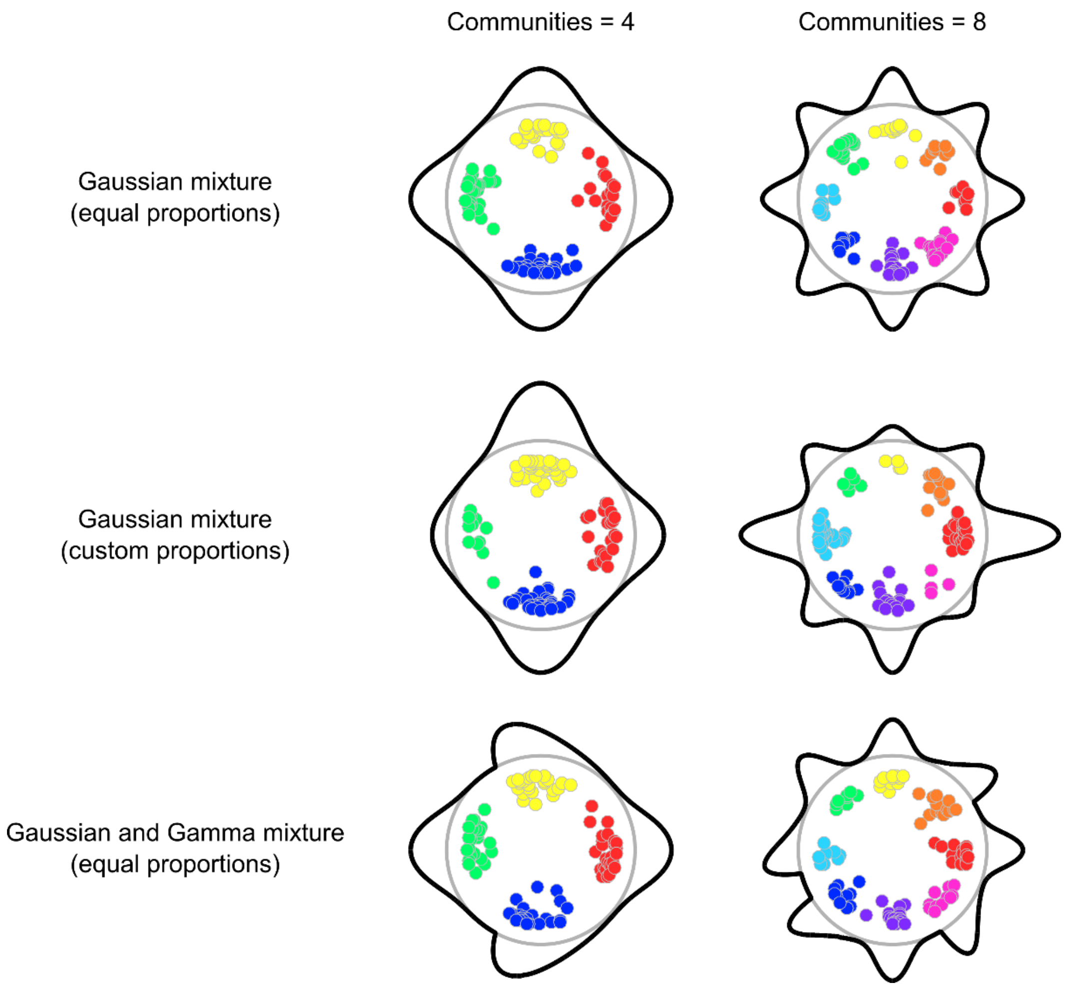
Uniform distribution of angular coordinates



Community structure

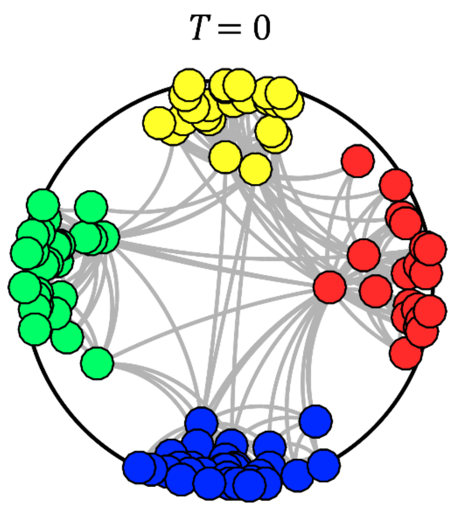


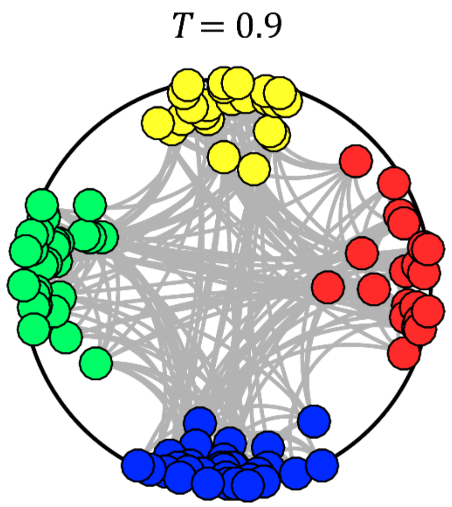
Community membership



- Node **label**: mixture component whose mean is at the lowest angular distance
- Scenarios with communities non-overlapping

Temperature

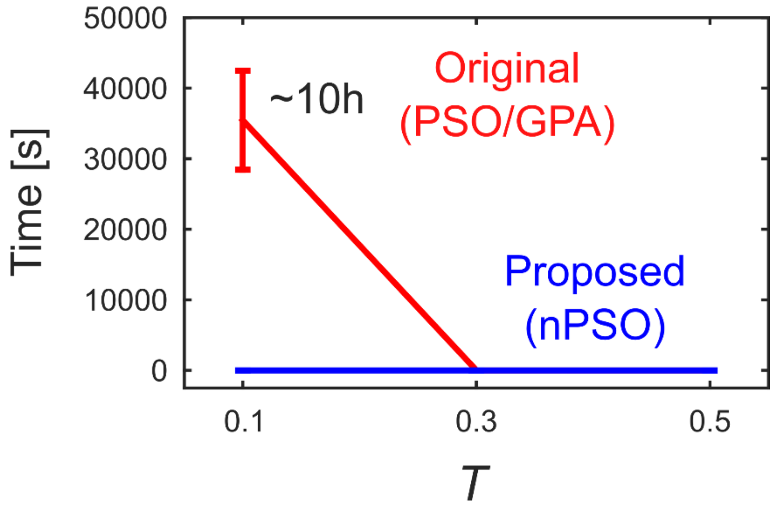




Temperature for tuning
clustering and
mixing between the communities

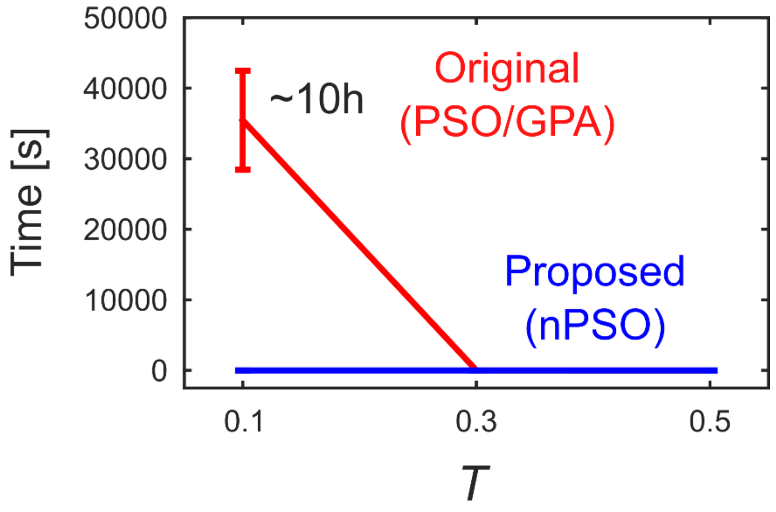
Implementation for links generation

Time for 1 network
(N = 1000)



Implementation for links generation

Time for 1 network
(N = 1000)



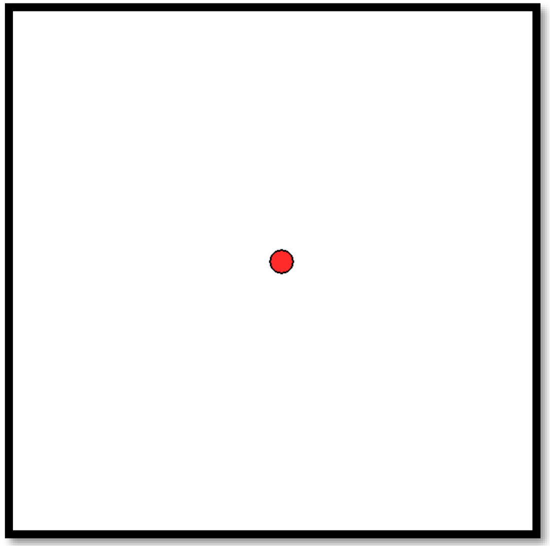
- **Time complexity O(*EN*)**
- ***N* = 10000 → 5 min**

Generative procedure (with community)

For each iteration $t = 1 \ldots \boldsymbol{N}$

1) Update radial coordinates of existing nodes $j < t$: $r_j = f(j, t, \boldsymbol{\gamma})$
2) Introduce new node
   - Radial coordinate: $f(t)$
   - Angular coordinate: sampled from **distribution**
3) Establish $\boldsymbol{m}$ links
   - Connection probability: $p_{tj} = f(h_{tj}, \boldsymbol{T}, \ldots)$

$N = 100$, $m = 4$, $T = 0$, $\boldsymbol{\gamma} = 3$
4 communities



Solution to problem 2 published in 2018

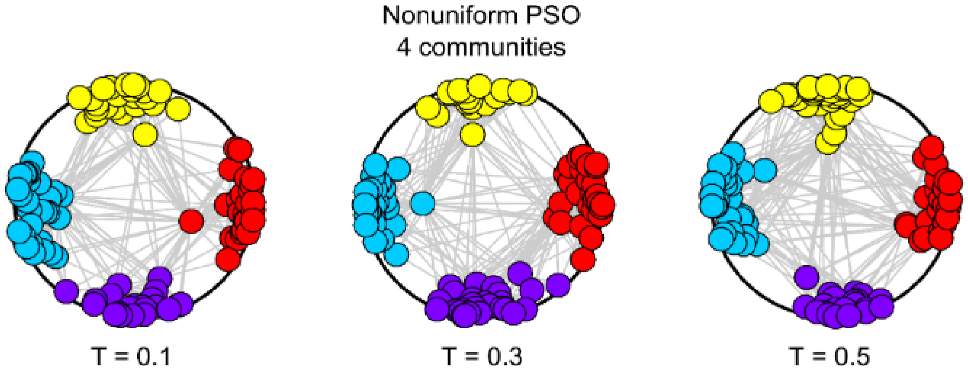
Alessandro Muscoloni

1. The first article on the **theoretical model**

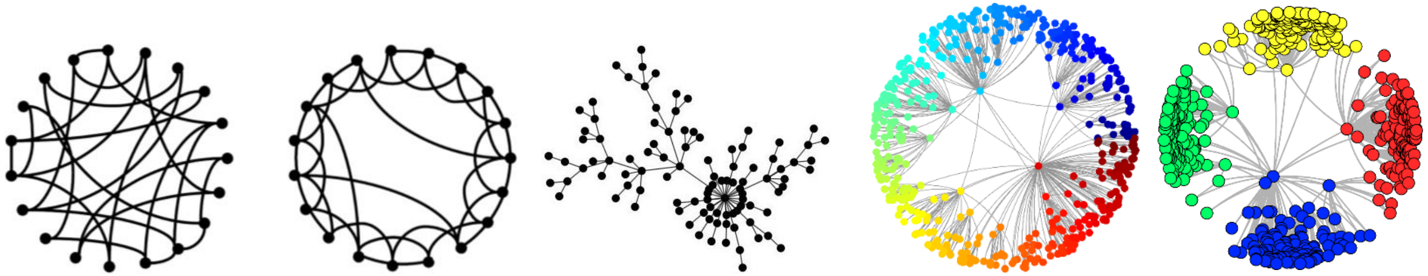
A nonuniform popularity-similarity optimization (nPSO) model to efficiently generate realistic complex networks with communities
Alessandro Muscolini and Carlo Vittorio Cannistraci
http://iopscience.iop.org/article/10.1088/1367-2630/aac06f/meta

2. The second article on **the application**

Leveraging the nonuniform PSO network model as a benchmark for performance evaluation in community detection and link prediction
Alessandro Muscolini and Carlo Vittorio Cannistraci
http://iopscience.iop.org/article/10.1088/1367-2630/aac6f9



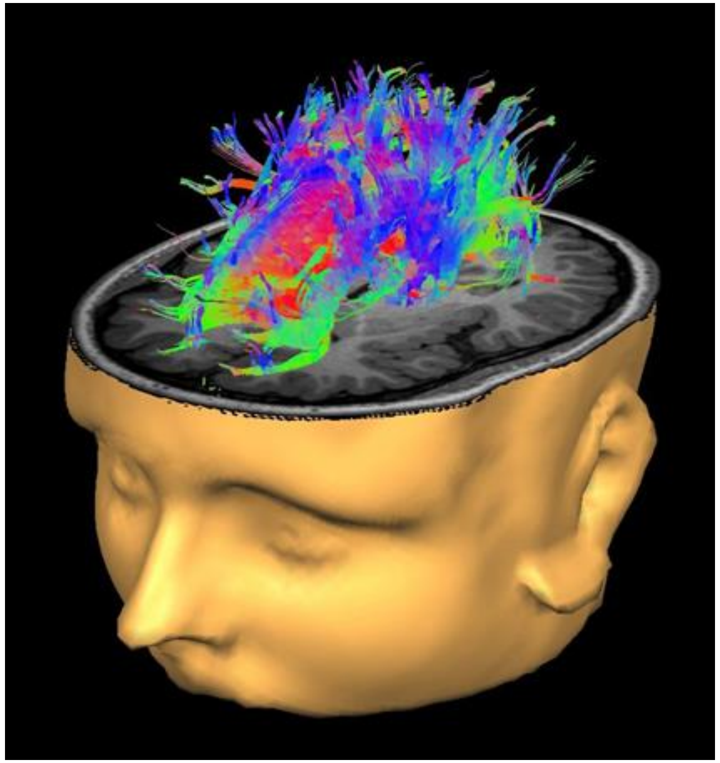
Generative models in Network Science

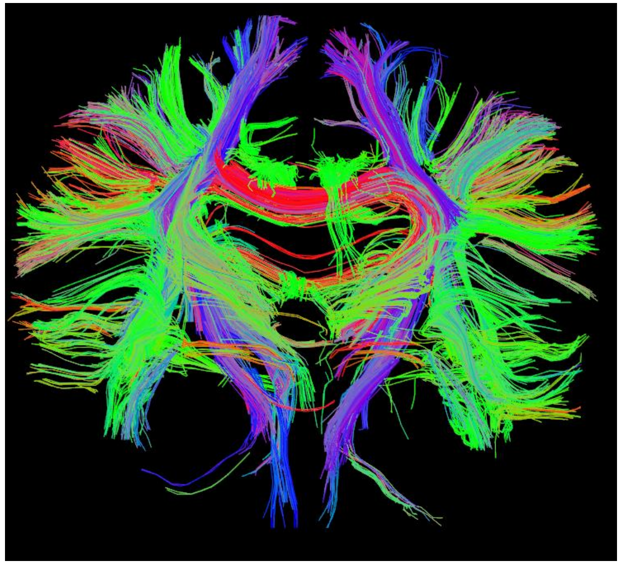


	Erdős-Rényi (1959)	Watts-Strogatz (1998)	Barabási-Albert (1999)	PSO (2012)	nPSO (2018)	????
Clustering	—	✓	—	✓	✓	✓
Small-world	✓	✓	✓	✓	✓	✓
Scale-free	—	—	✓	✓	✓	✓
Community	—	—	—	—	✓	✓
Assortativity	—	—	—	—	—	✓
Motifs	—	—	—	—	—	✓

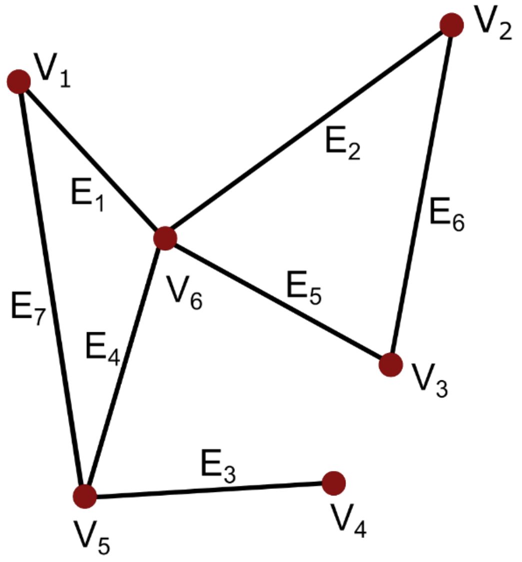
Problem 3
(Geometrical markers from brain diseases)

Human brain
diffusion weighted magnetic resonance imaging (DW-MRI)
structural connectomes

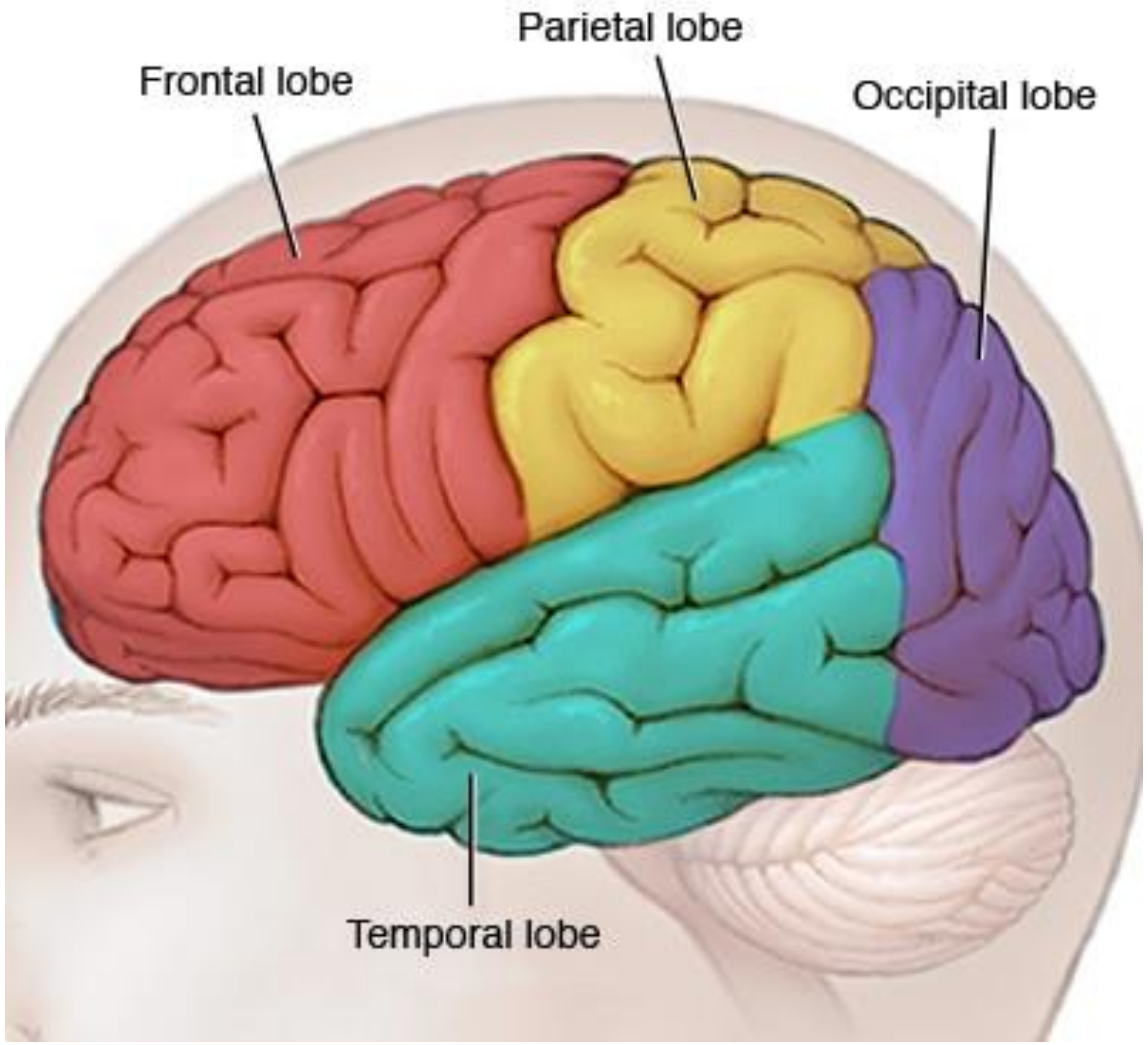


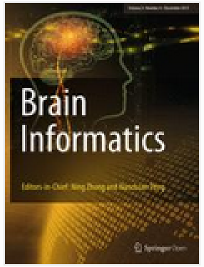


Nodes are brain areas



Anatomical Brain Lobe Characterization





Brain Informatics
December 2015, Volume 2, Issue 4, pp 197–210

2015

The intrinsic geometry of the human brain connectome

Authors | Authors and affiliations

Allen Q. Ye, Olusola A. Ajilore, Giorgio Conte, Johnson GadElkarim, Galen Thomas-Ramos, Liang Zhan, Shaolin Yang, Anand Kumar, Richard L. Magin, Angus G. Forbes, Alex D. Leow

Open Access | Article

First Online: 07 November 2015

DOI: 10.1007/s40708-015-0022-2

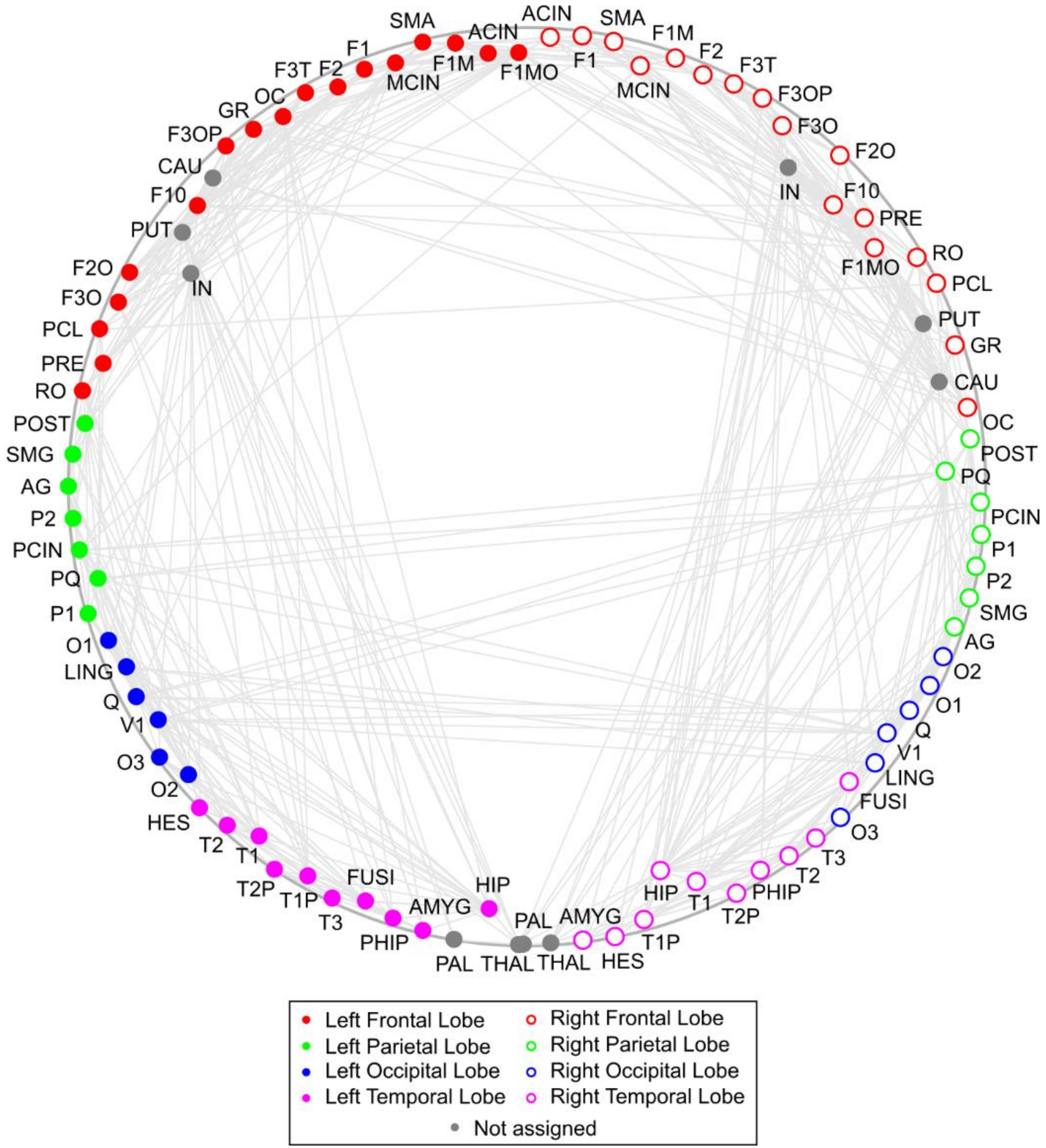
Cite this article as:
Ye, A.Q., Ajilore, O.A., Conte, G. et al. Brain Inf. (2015) 2: 197. doi:10.1007/s40708-015-0022-2

"Intrinsic brain network geometry only minimally relates to neuroanatomy!"

Open Problem to solve!

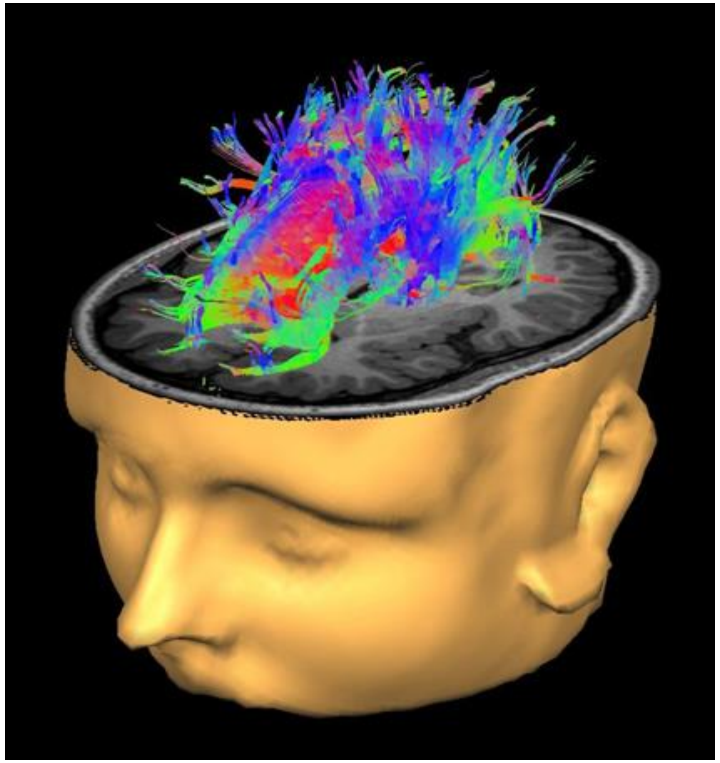
Coalescent Embedding

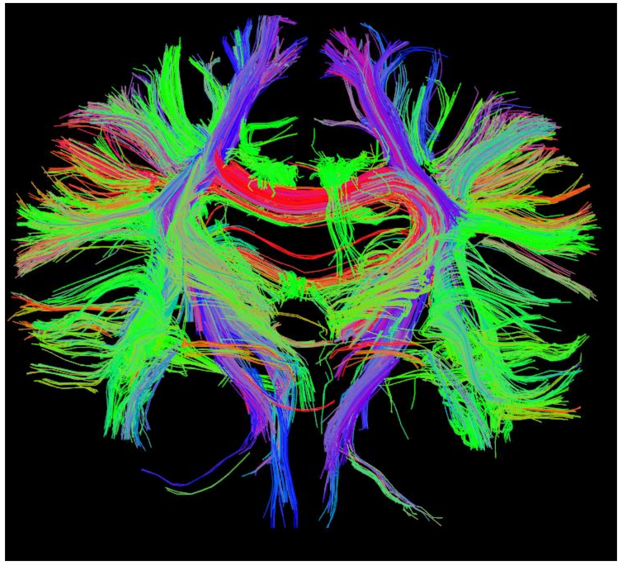
PRE	Precentral
F1	Frontal-Sup
F10	Frontal-Sup-Orb
F2	Frontal-Mid
F2O	Frontal-Mid-Orb
F3OP	Frontal-Inf-Oper
F3T	Frontal-Inf-Tri
F3O	Frontal-Inf-Orb
RO	Rolandic-Oper
SMA	Supp-Motor-Area
OC	Olfactory
F1M	Frontal-Sup-Medial
F1MO	Frontal-Med-Orb
GR	Rectus
IN	Insula
ACIN	Cingulum-Ant
MCIN	Cingulum-Mid
PCIN	Cingulum-Post
HIP	Hippocampus
PHIP	ParaHippocampal
AMYG	Amygdala
V1	Calcarine
Q	Cuneus
LING	Lingual
O1	Occipital-Sup
O2	Occipital-Mid
O3	Occipital-Inf
FUSI	Fusiform
POST	Postcentral
P1	Parietal-Sup
P2	Parietal-Inf
SMG	SupraMarginal
AG	Angular
PQ	Precuneus
PCL	Paracentralobule
CAU	Caudate
PUT	Putamen
PAL	Pallidum
THAL	Thalamus
HES	Heschl
T1	Temporal-Sup
T1P	Temporal-Pole-Sup
T2	Temporal-Mid
T2P	Temporal-Pole-Mid
T3	Temporal-Inf



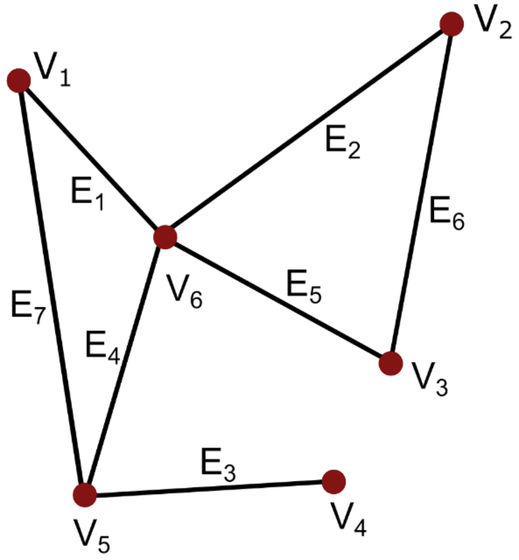
Applied example

Network-based markers for brain diseases: brain imaging quantification of disease state in psychiatric (depression) and neurodegenerative (Parkinson, Alzheimer, etc.) disorders.





Nodes are brain areas



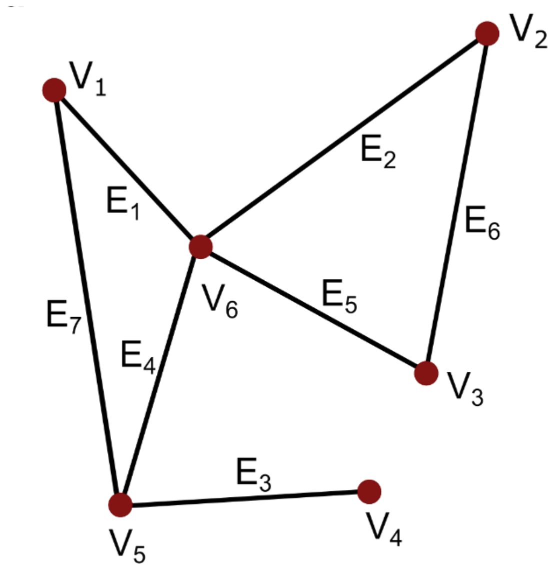
Geometrical modifications of the brain in diseases

de novo drug naïve Parkinson's Disease (PD) patients
compared with
Healthy Controls (HC)

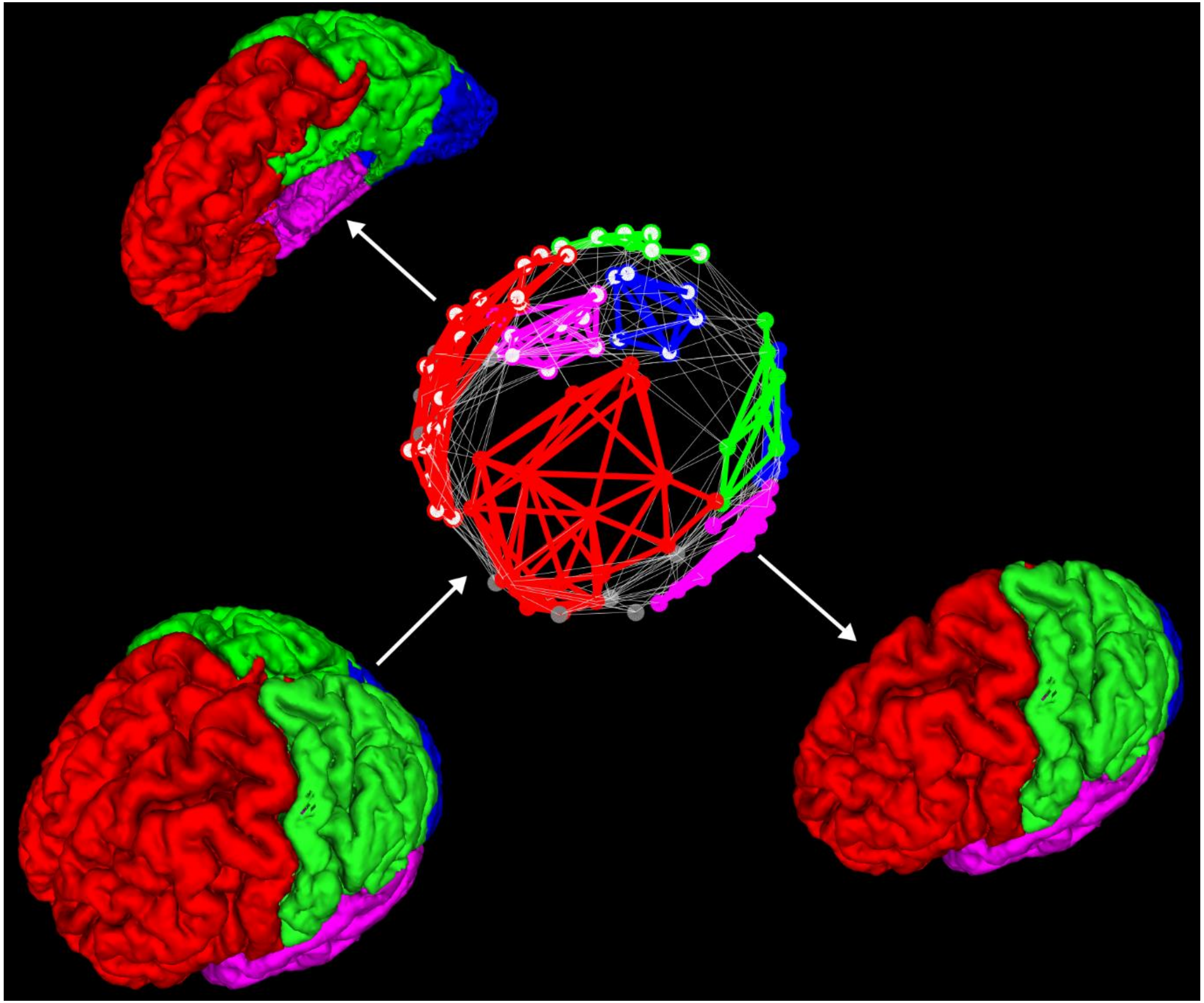
	mean marker (HC)	mean marker (PD)	MW p-value	AUC	AUPR
Coalescent embedding (MCE)	**14.7**	**15.7**	**0.006**	**0.87**	**0.82**
HyperMap	13.0	14.0	**0.026**	**0.80**	**0.76**
Original network	222.0	215.8	0.307	0.64	0.64

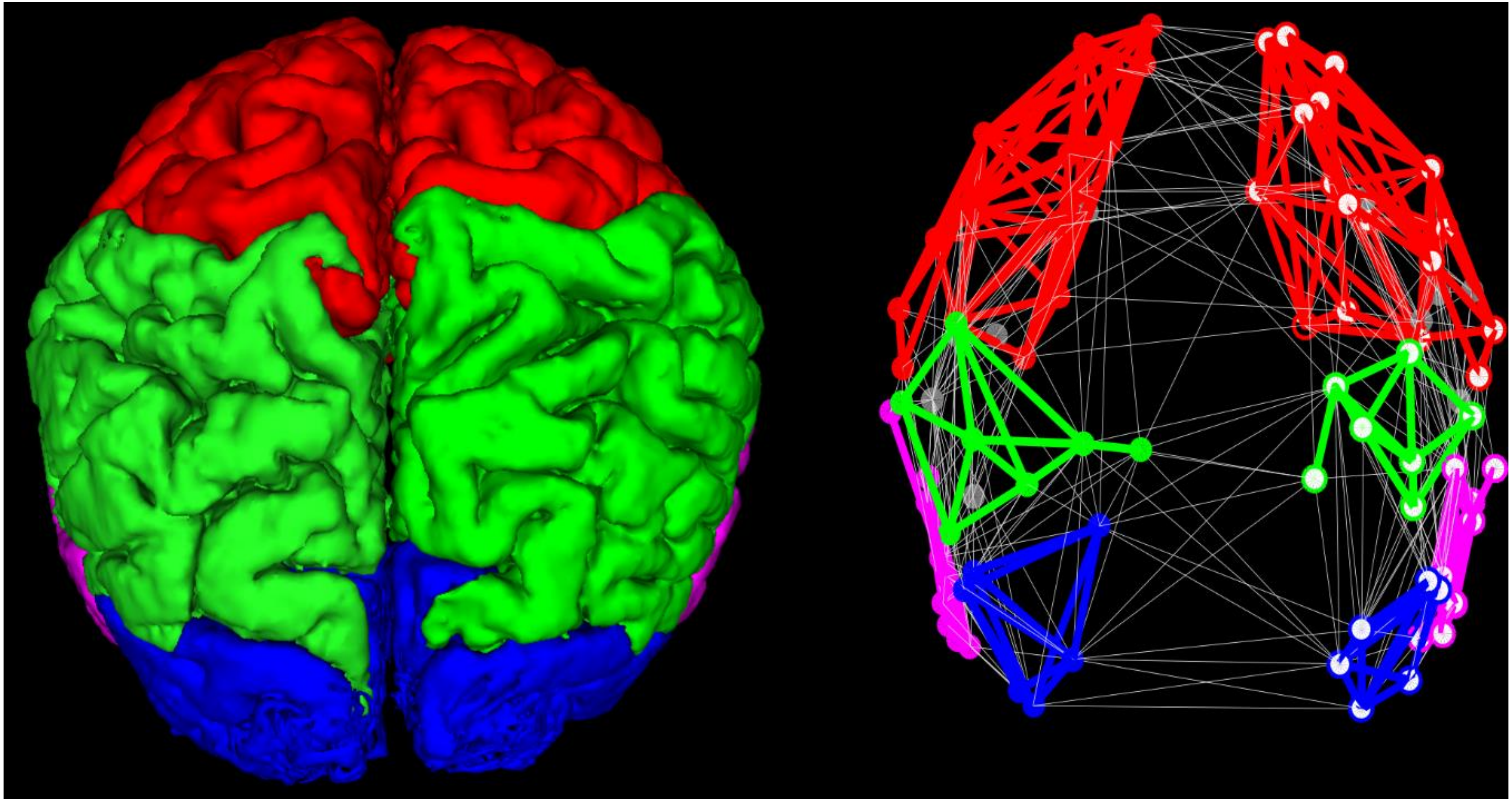
Coalescent Embedding in the 3D

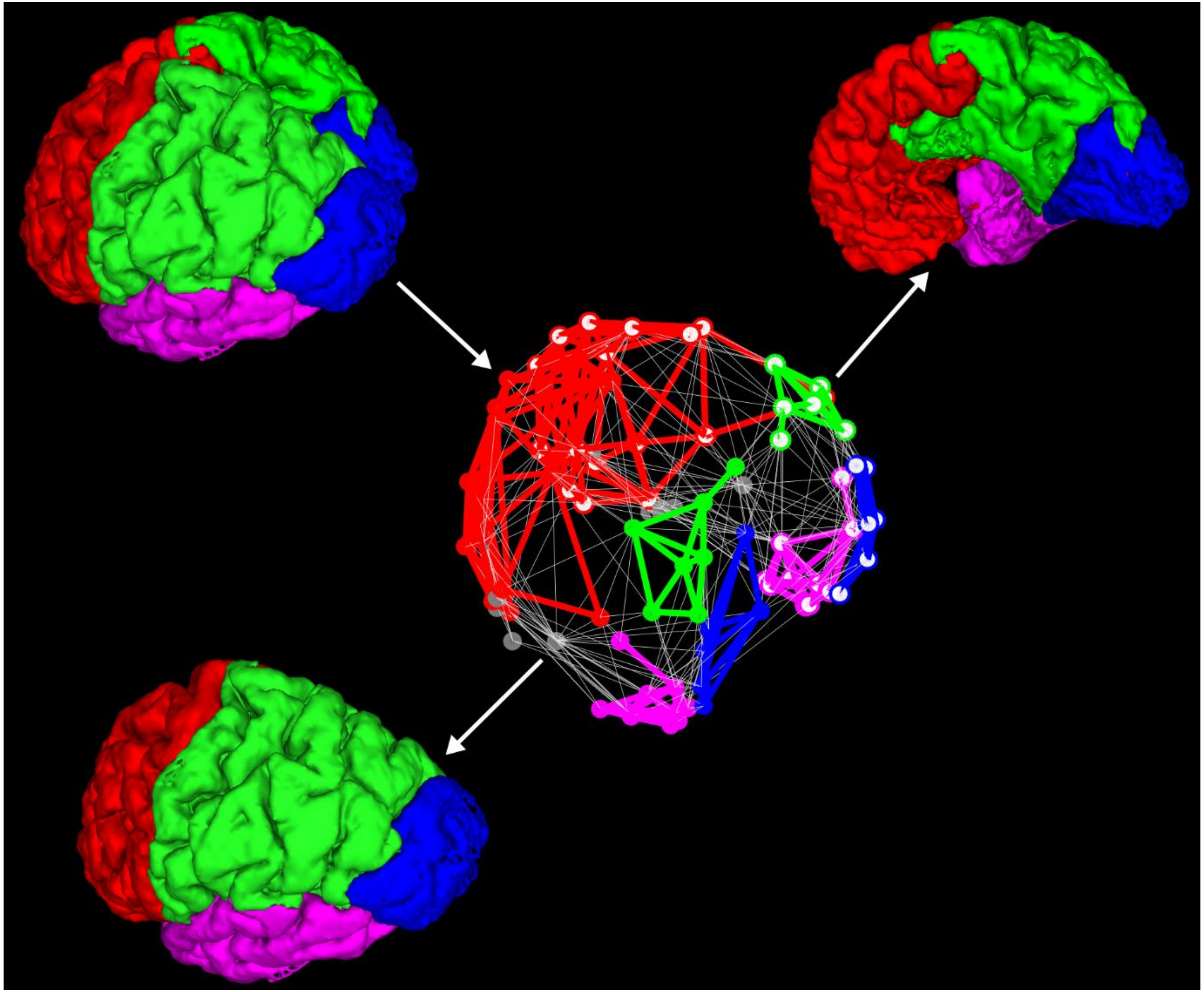
Coalescent Embedding in the 3D



Brain Network of an individual







Real application Network Neuroscience

Alberto Cacciola

Cornell University Library

arXiv.org > q-bio > arXiv:1705.04192

Quantitative Biology > Neurons and Cognition

Coalescent embedding in the hyperbolic space unsupervisedly discloses the hidden geometry of the brain

Alberto Cacciola, Alessandro Muscoloni, Vaibhav Narula, Alessandro Calamuneri, Salvatore Nigro, Emeran A. Mayer, Jennifer S. Labus, Giuseppe Anastasi, Aldo Quattrone, Angelo Quartarone, Demetrio Milardi, Carlo Vittorio Cannistraci

(Submitted on 10 May 2017)

The human brain displays a complex network topology, whose structural organization is widely studied using diffusion tensor imaging. The original geometry from which emerges the network topology is known, as well as the localization of the network nodes in respect to the brain morphology and anatomy. One of the most challenging problems of current network science is to infer the latent geometry from the mere topology of a complex network. The human brain structural connectome represents the perfect benchmark to test algorithms aimed to solve this problem. Coalescent embedding was recently designed to map a complex network in the hyperbolic space, inferring the node angular coordinates. Here we show that this methodology is able to unsupervisedly reconstruct the latent geometry of the brain with an incredible accuracy and that the intrinsic geometry of the brain networks strongly relates to the lobes organization known in neuroanatomy. Furthermore, coalescent embedding allowed the detection of geometrical pathological changes in the connectomes of Parkinson's Disease patients. The present study represents the first evidence of brain networks' angular coalescence in the hyperbolic space, opening a completely new perspective, possibly towards the realization of latent geometry network markers for evaluation of brain disorders and pathologies.

Problem 4
(Geometrical congruence of a network)

nature communications

Article https://doi.org/10.1038/s41467-022-34634-6

Geometrical congruence, greedy navigability and myopic transfer in complex networks and brain connectomes

Received: 6 July 2020

Accepted: 1 November 2022

Published online: 27 November 2022

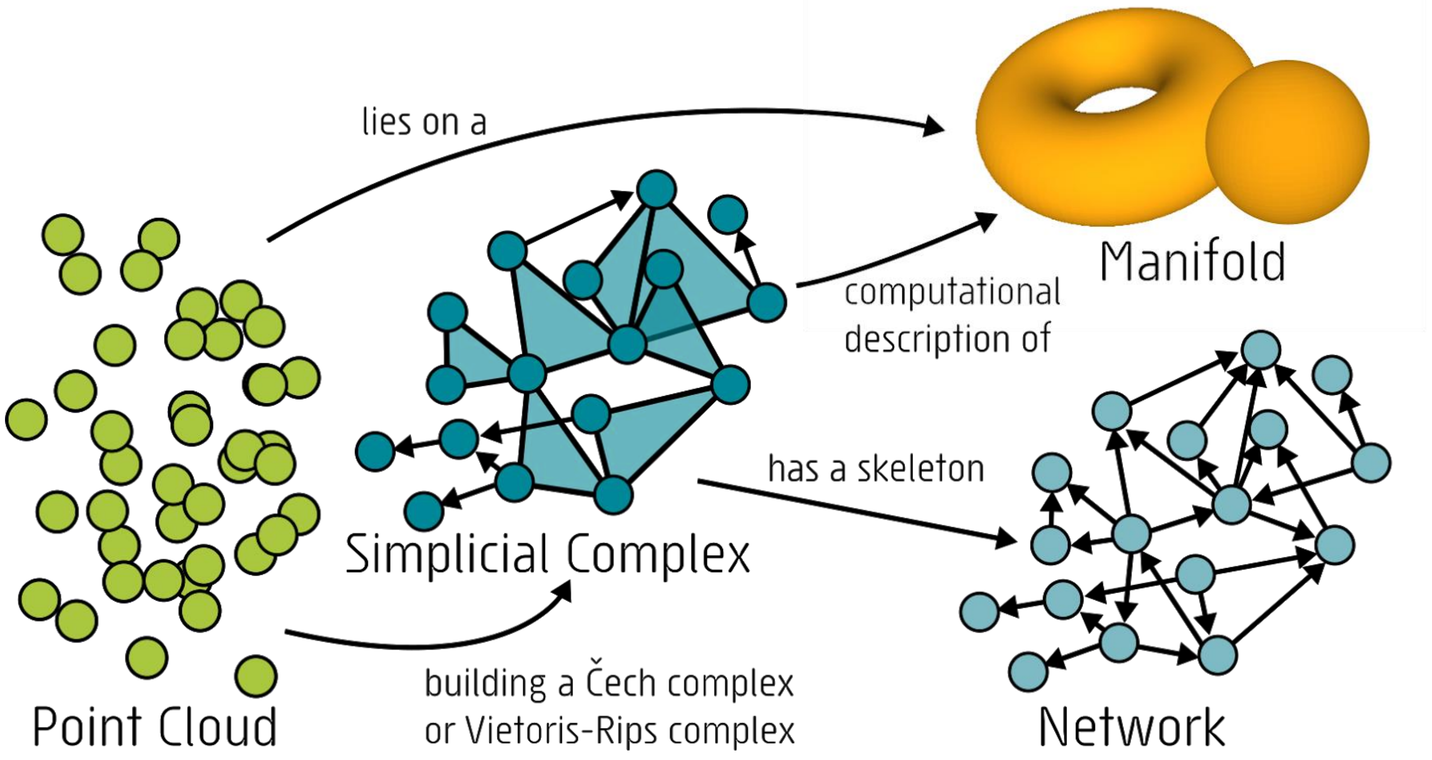
Check for updates

Carlo Vittorio Cannistraci[1,2,3,4,5,6] & Alessandro Muscoloni[1,4]

We introduce in network geometry a measure of *geometrical congruence (GC)* to evaluate the extent a network topology follows an underlying geometry. This requires finding all topological shortest-paths for each nonadjacent node pair in the network: a nontrivial computational task. Hence, we propose an optimized algorithm that reduces 26 years of worst scenario computation to one week parallel computing. Analysing artificial networks with patent geometry we discover that, different from current belief, hyperbolic networks do not show in general high GC and efficient greedy navigability (GN) with respect to the geodesics. The myopic transfer which rules GN works best only when degree-distribution power-law exponent is strictly close to two. Analysing real networks—whose geometry is often latent—GC overcomes GN as marker to differentiate phenotypical states in macroscale structural-MRI brain connectomes, suggesting connectomes might have a latent neurobiological geometry accounting for more information than the visible tridimensional Euclidean.

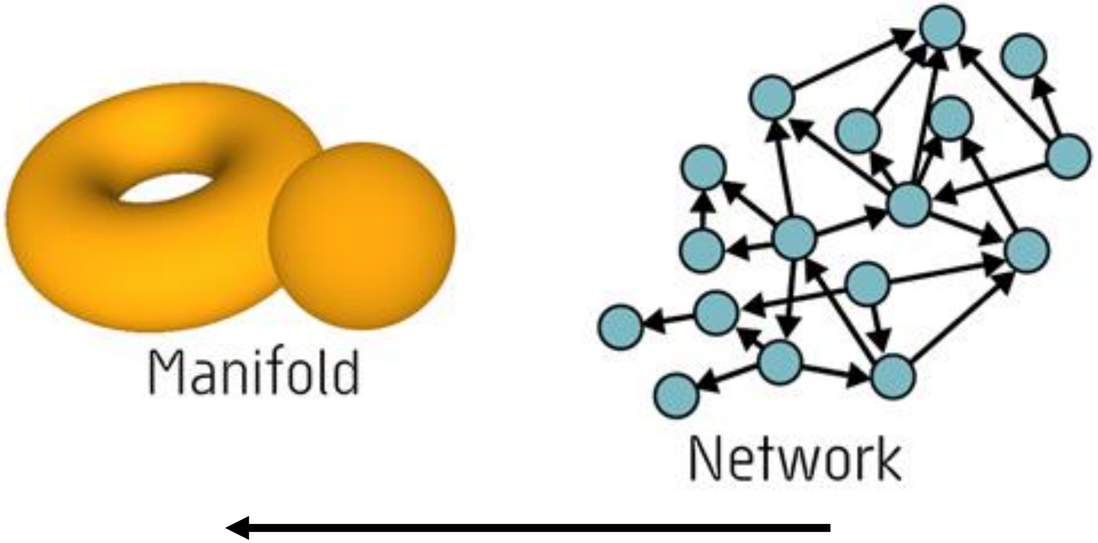
November 2022

Artistic representation of the topics of today



https://menchelab.com/higher-order-networks-and-the-topology-of-data

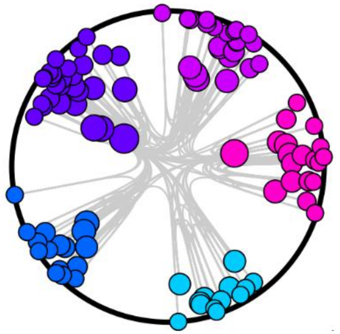
Computational measure of the soft congruence of a topology with a geometry



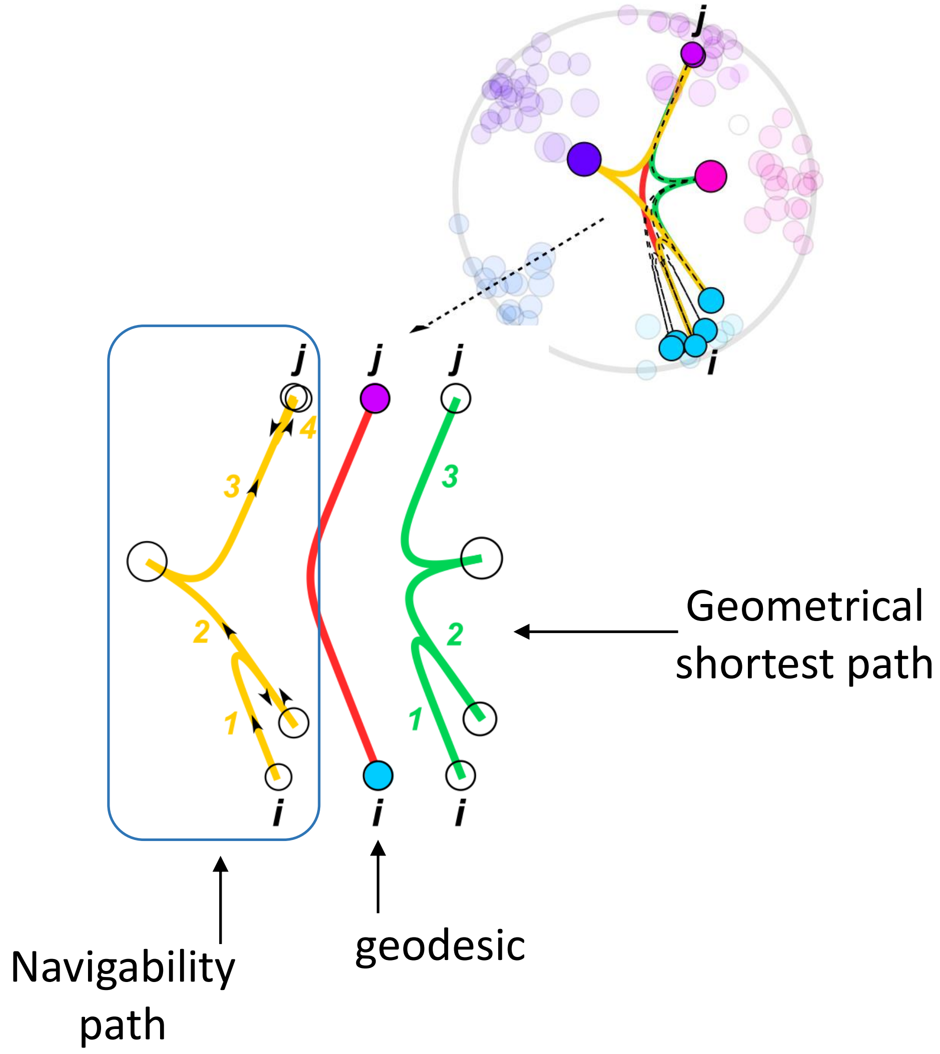
Soft congruence : Measuring to which extent
a topology of a network follows a generative geometry

Measuring how modifications of general network properties
modifies soft congruence

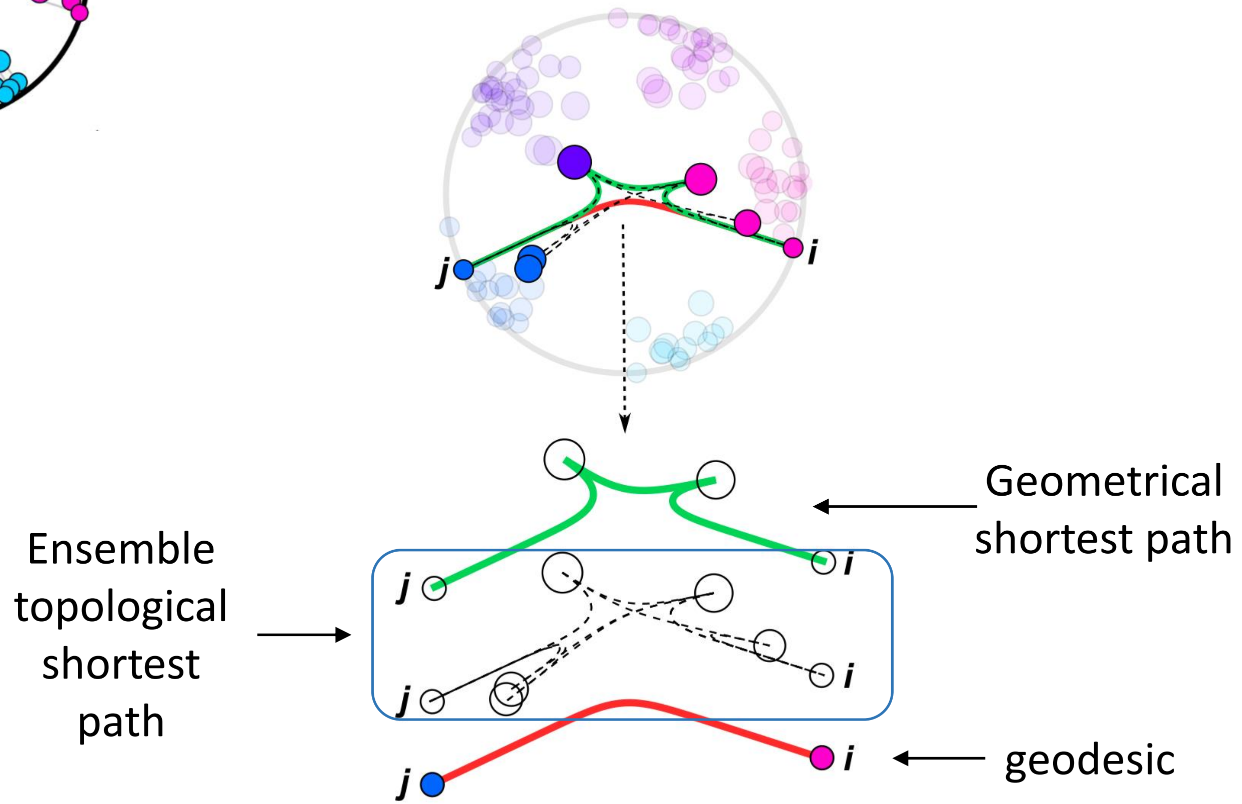
Network in a geometrical space



Navigability measure (previous)
Muscoloni,…, Cannistraci et al.
Nat. Com. 2017



Congruency measure (New)
Cannistraci et al. Nat. Com. 2022



Previous literature

Krioukov, D. et al. Hyperbolic geometry of complex networks. Phys Rev E (2010).

Boguñá, M. & Krioukov, D. Navigating ultrasmall worlds in ultrashort time. Phys Rev Lett 102, 058701 (2009).

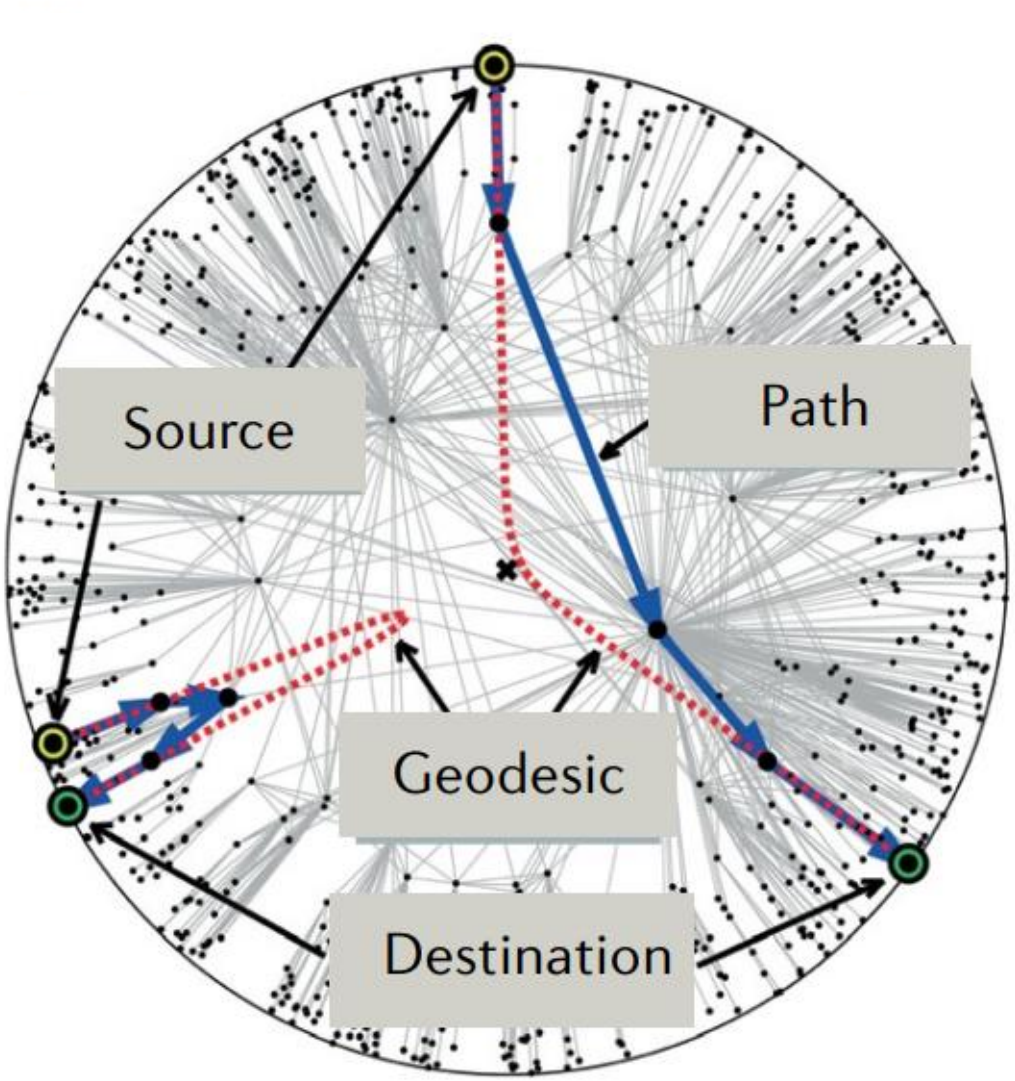
Boguñá, M. et al. Network geometry. Nature Reviews Physics 2021 3:2 3, 114–135 (2021).

- It was considering success ratio and stretch separately therefore there was not a unique measure of navigability efficiency

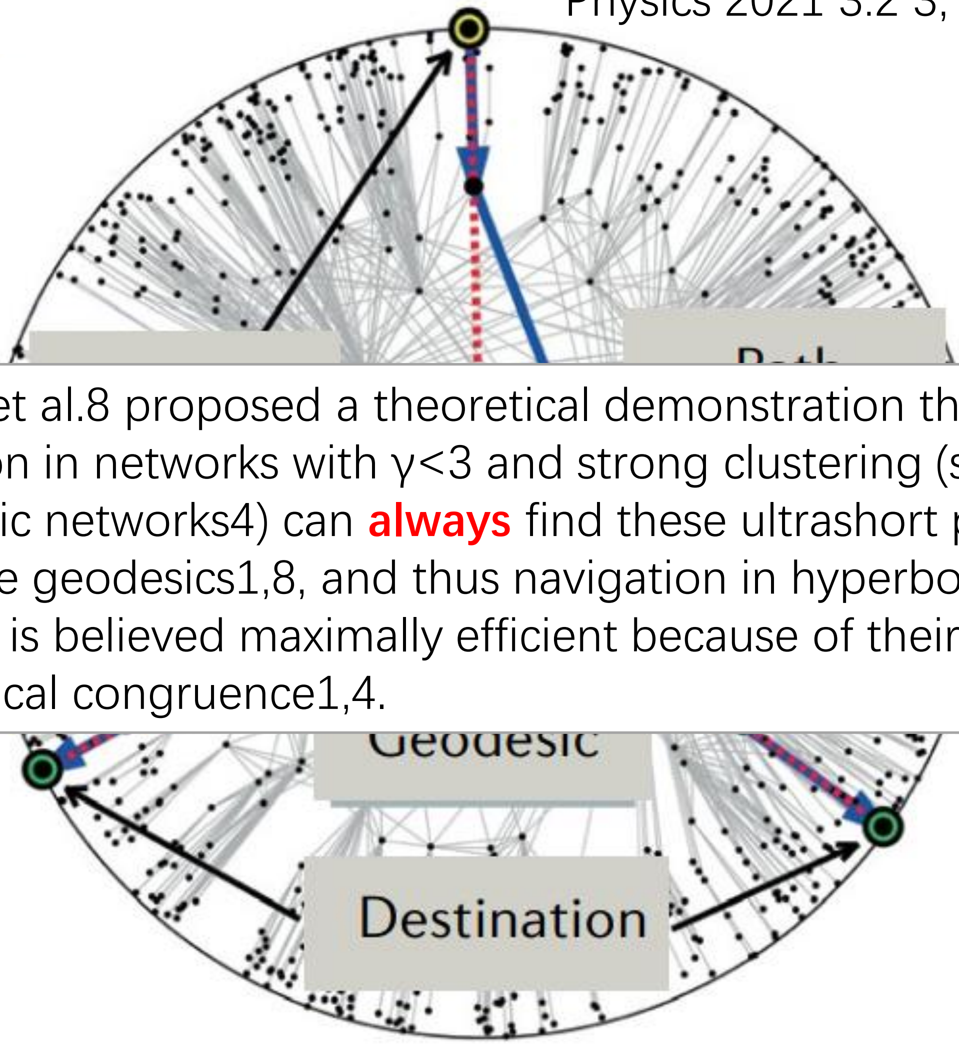
- It was considering the ‘congruency’ of successful greedy paths only supporting the believe that hyperbolic networks are maximally congruent with their underlying geometry

- It was not considering the ensemble of shortest path but only the geometrical shortest path.

Boguñá, M. et al. Network geometry. Nature Reviews Physics 2021 3:2 3, 114–135 (2021).



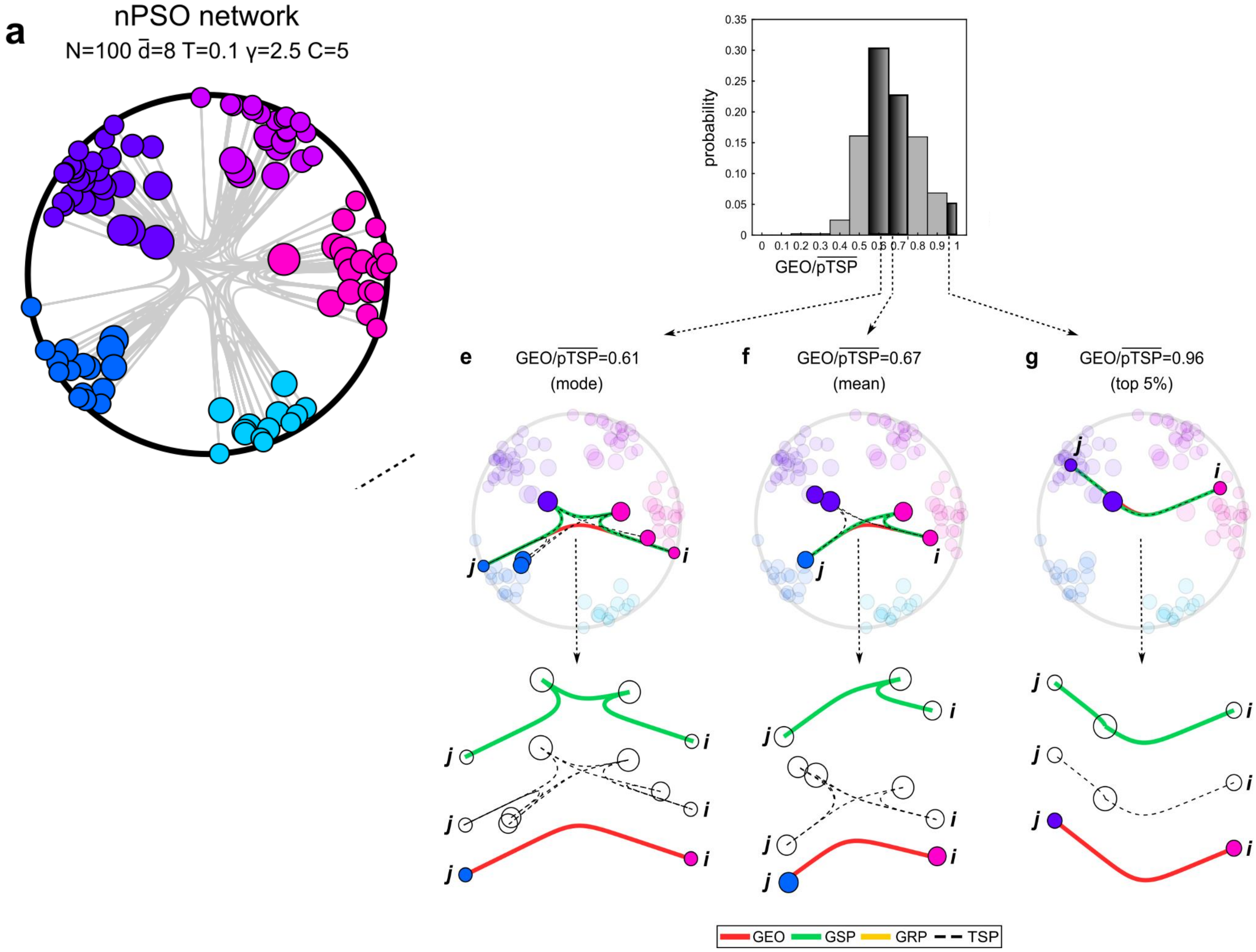
Boguñá, M. et al. Network geometry. Nature Reviews Physics 2021 3:2 3, 114–135 (2021).

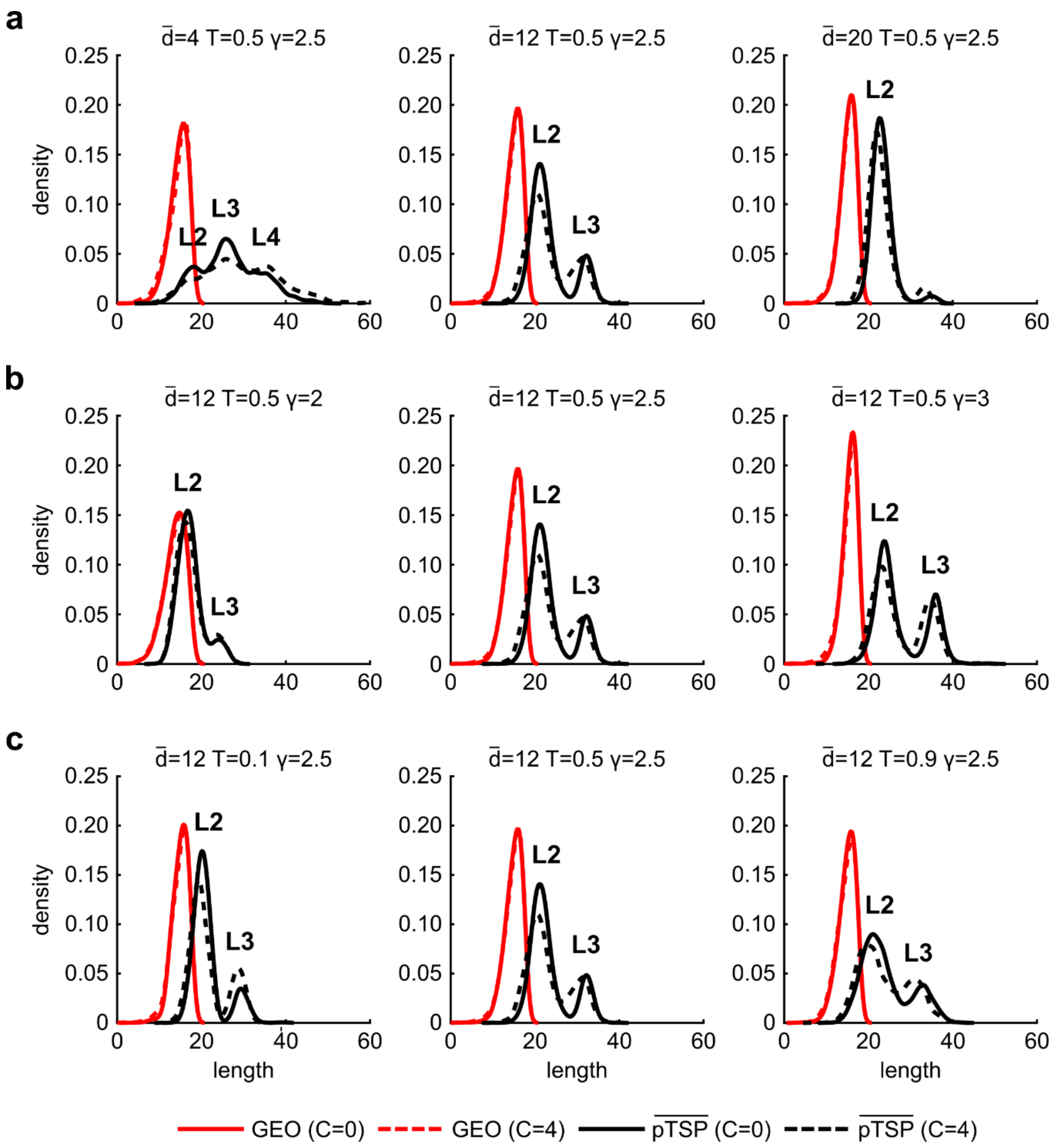


Boguñá et al.8 proposed a theoretical demonstration that greedy navigation in networks with γ<3 and strong clustering (such as hyperbolic networks4) can **always** find these ultrashort paths which follow the geodesics1,8, and thus navigation in hyperbolic networks with γ<3 is believed maximally efficient because of their supposed geometrical congruence1,4.

Krioukov, D. et al. Hyperbolic geometry of complex networks. Phys Rev E (2010).
Boguñá, M. & Krioukov, D. Navigating ultrasmall worlds in ultrashort time. Phys Rev Lett 102, 058701 (2009).

Results


a
nPSO network
N=100 d̄=8 T=0.1 γ=2.5 C=5
probability
GEO/pTSP
e
GEO/pTSP=0.61
(mode)
f
GEO/pTSP=0.67
(mean)
g
GEO/pTSP=0.96
(top 5%)
GEO
GSP
GRP
TSP


a
b
c
d̄=4 T=0.5 γ=2.5
d̄=12 T=0.5 γ=2.5
d̄=20 T=0.5 γ=2.5
d̄=12 T=0.5 γ=2
d̄=12 T=0.5 γ=2.5
d̄=12 T=0.5 γ=3
d̄=12 T=0.1 γ=2.5
d̄=12 T=0.5 γ=2.5
d̄=12 T=0.9 γ=2.5
density
length
L2
L3
L4
GEO (C=0)
GEO (C=4)
p̄TSP (C=0)
p̄TSP (C=4)

Navigability measure (previous)
Muscoloni,⋯, Cannistraci et al. Nat. Com. 2017

$$GRE(pGRP, RD) = \left(\frac{1}{n \cdot (n-1) - 2 \cdot e}\right) \cdot \sum \frac{RD(\mathrm{i},\mathrm{j})}{pGRP(\mathrm{i},\mathrm{j})} \;;\quad \text{with } (\mathrm{i},\mathrm{j}) \in \tilde{E}$$

Congruency measure (New)
Cannistraci et al. Nat. Com. 2022

$$GC(\overline{pTSP}, RD) = \left(\frac{2}{n \cdot (n-1) - 2 \cdot e}\right) \cdot \sum_{i<j} \frac{RD(\mathrm{i},\mathrm{j})}{\overline{pTSP}(\mathrm{i},\mathrm{j})} \;;\quad \text{with } (\mathrm{i},\mathrm{j}) \in \tilde{E}$$

Navigability measure (previous)
Muscoloni,⋯, Cannistraci et al. Nat. Com. 2017

$$GRE(pGRP, RD) = \left(\frac{1}{n \cdot (n-1) - 2 \cdot e}\right) \cdot \sum \frac{RD(\mathrm{i},\mathrm{j})}{pGRP(\mathrm{i},\mathrm{j})} \;;\quad \text{with } (\mathrm{i},\mathrm{j}) \in \tilde{E}$$

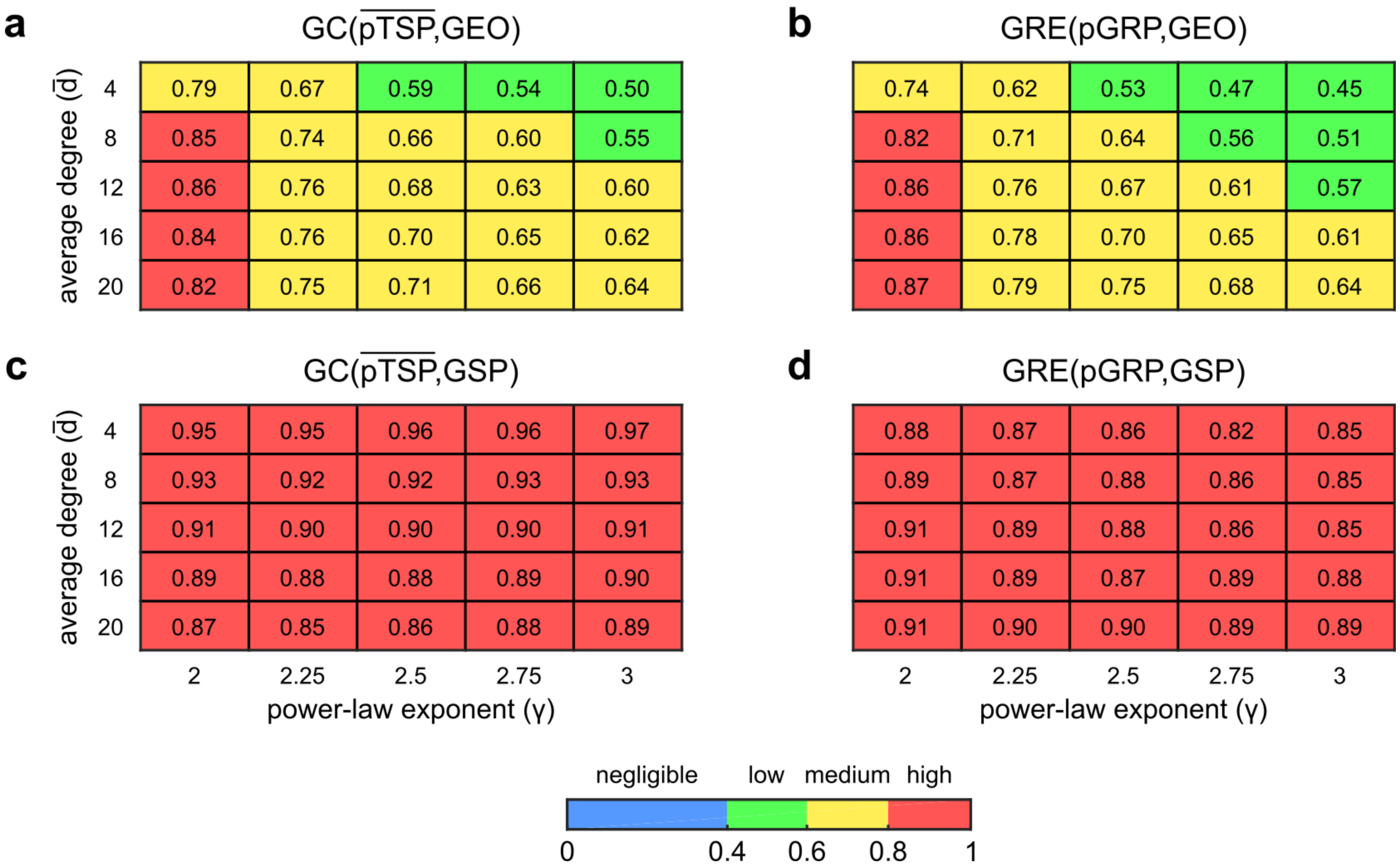
Greedy (approximated) measure

Congruency measure (New)
Cannistraci et al. Nat. Com. 2022

$$GC(\overline{pTSP}, RD) = \left(\frac{2}{n \cdot (n-1) - 2 \cdot e}\right) \cdot \sum_{i<j} \frac{RD(\mathrm{i},\mathrm{j})}{\overline{pTSP}(\mathrm{i},\mathrm{j})} \;;\quad \text{with } (\mathrm{i},\mathrm{j}) \in \tilde{E}$$

Exact measure

nPSO (C=4 T=0.1 N=100)



nPSO (C=4 N=10000)

T = 0.1

$\mathrm{GC}(\overline{pTSP}, GEO)$

average degree ($\bar{d}$)					
4	0.74	0.54	0.41	0.34	0.30
8	0.82	0.64	0.50	0.42	0.37
12	0.86	0.67	0.54	0.46	0.41
16	0.88	0.69	0.56	0.49	0.44
20	0.89	0.70	0.58	0.51	0.46

$\mathrm{GRE}(pGRP, GEO)$

0.68	0.47	0.34	0.28	0.24
0.75	0.56	0.43	0.37	0.32
0.80	0.60	0.48	0.40	0.36
0.82	0.62	0.49	0.43	0.39
0.83	0.63	0.51	0.45	0.40

$\mathrm{GC}(\overline{pTSP}, GSP)$

average degree ($\bar{d}$)					
4	0.94	0.96	0.97	0.98	0.98
8	0.92	0.95	0.96	0.97	0.97
12	0.93	0.93	0.95	0.96	0.96
16	0.93	0.92	0.94	0.95	0.96
20	0.93	0.92	0.94	0.95	0.95

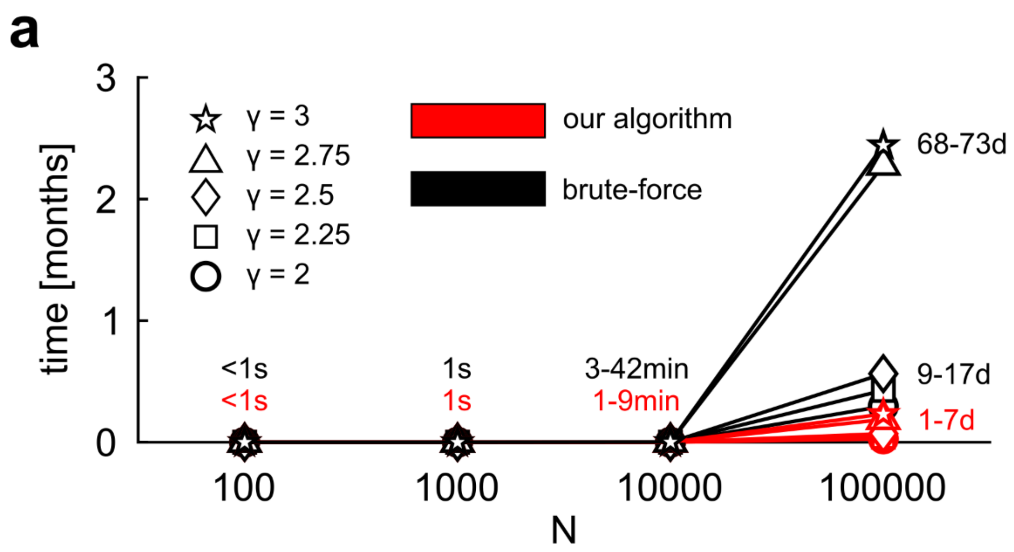
$\mathrm{GRE}(pGRP, GSP)$

0.86	0.82	0.80	0.81	0.80
0.84	0.83	0.82	0.83	0.82
0.85	0.82	0.82	0.82	0.83
0.86	0.82	0.82	0.83	0.83
0.87	0.82	0.82	0.82	0.82

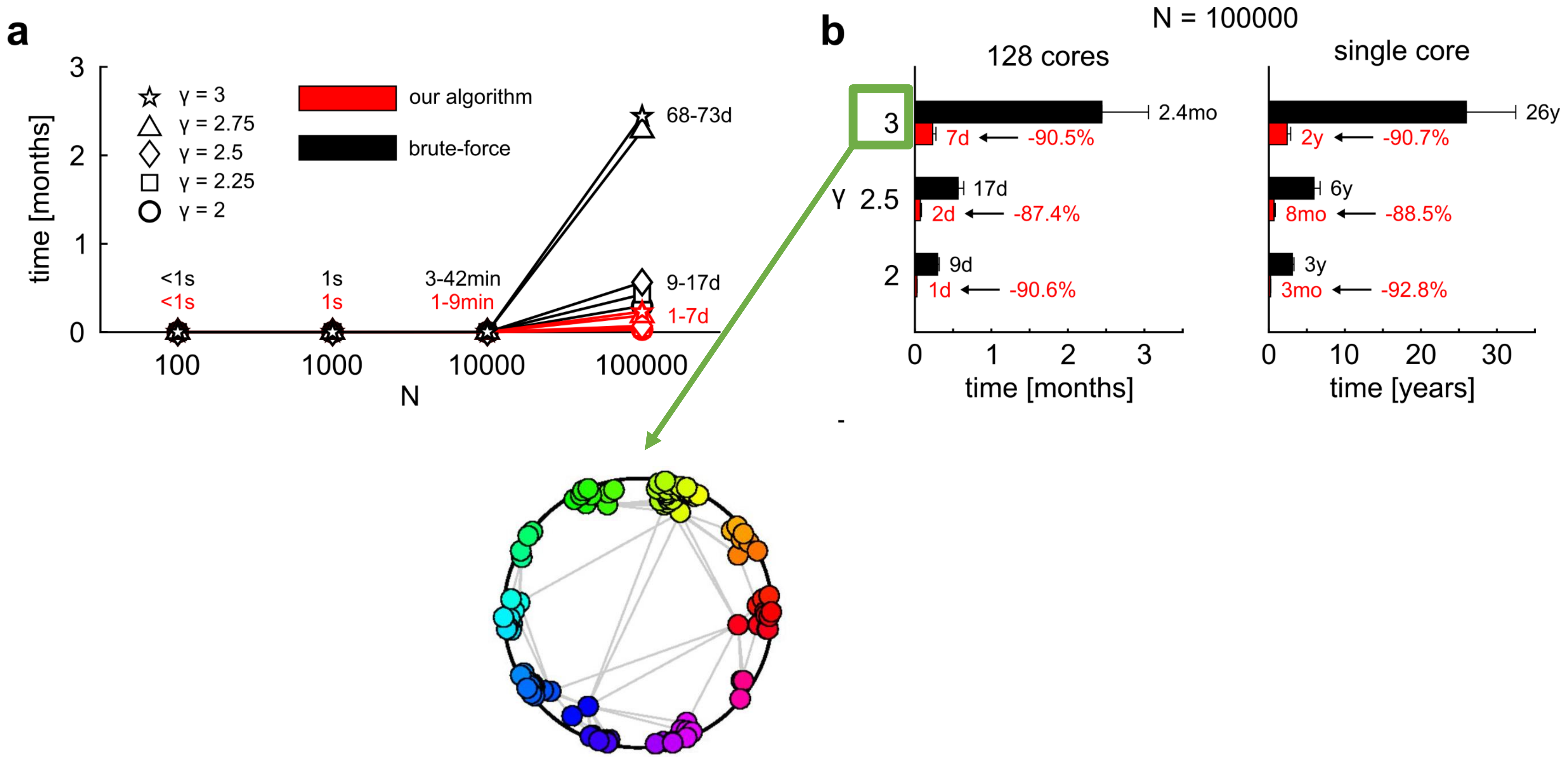
negligible low medium high

0 0.4 0.6 0.8 1

Optimized algorithm to find all shortest paths

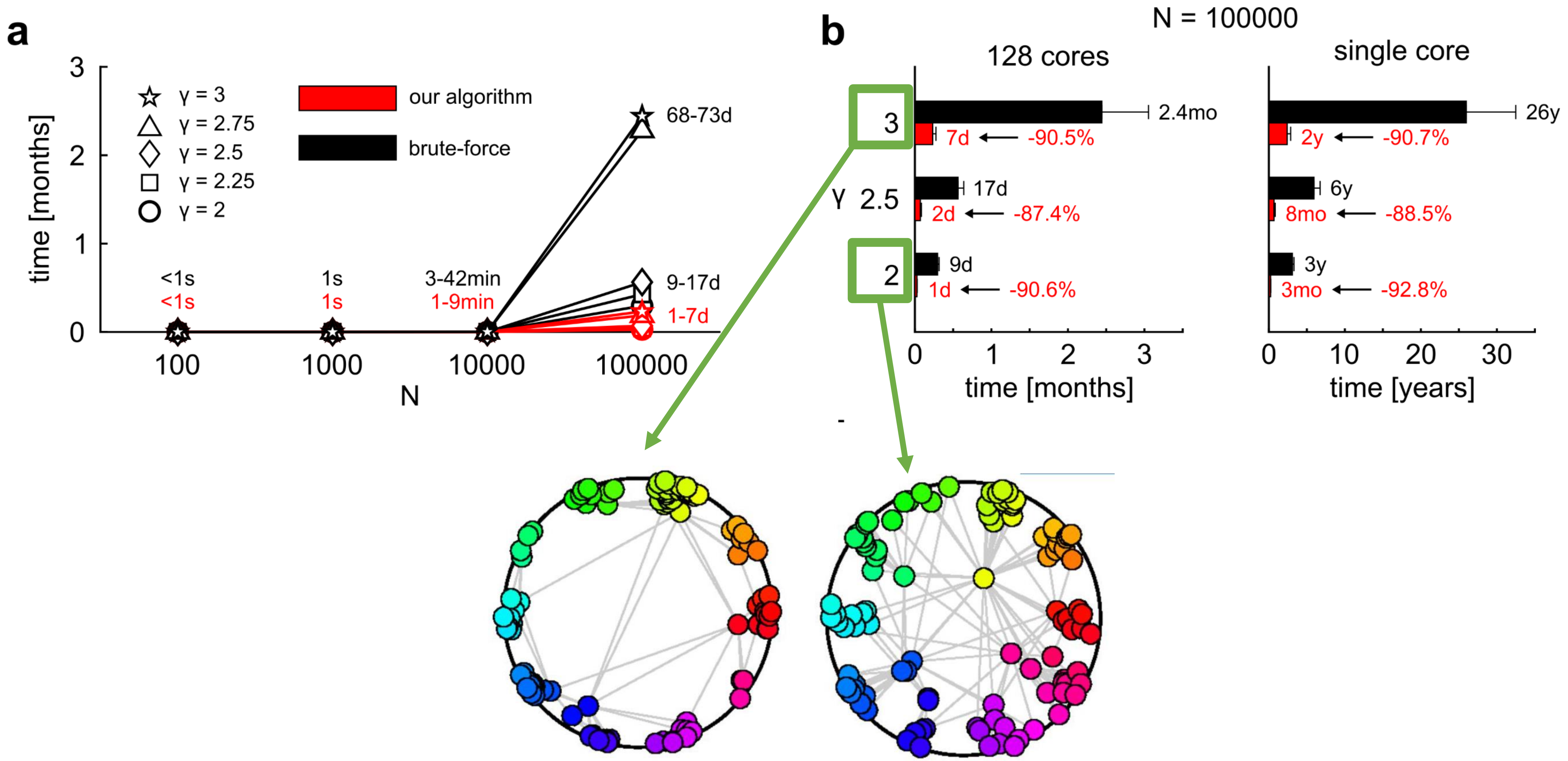


Optimized algorithm to find all shortest paths



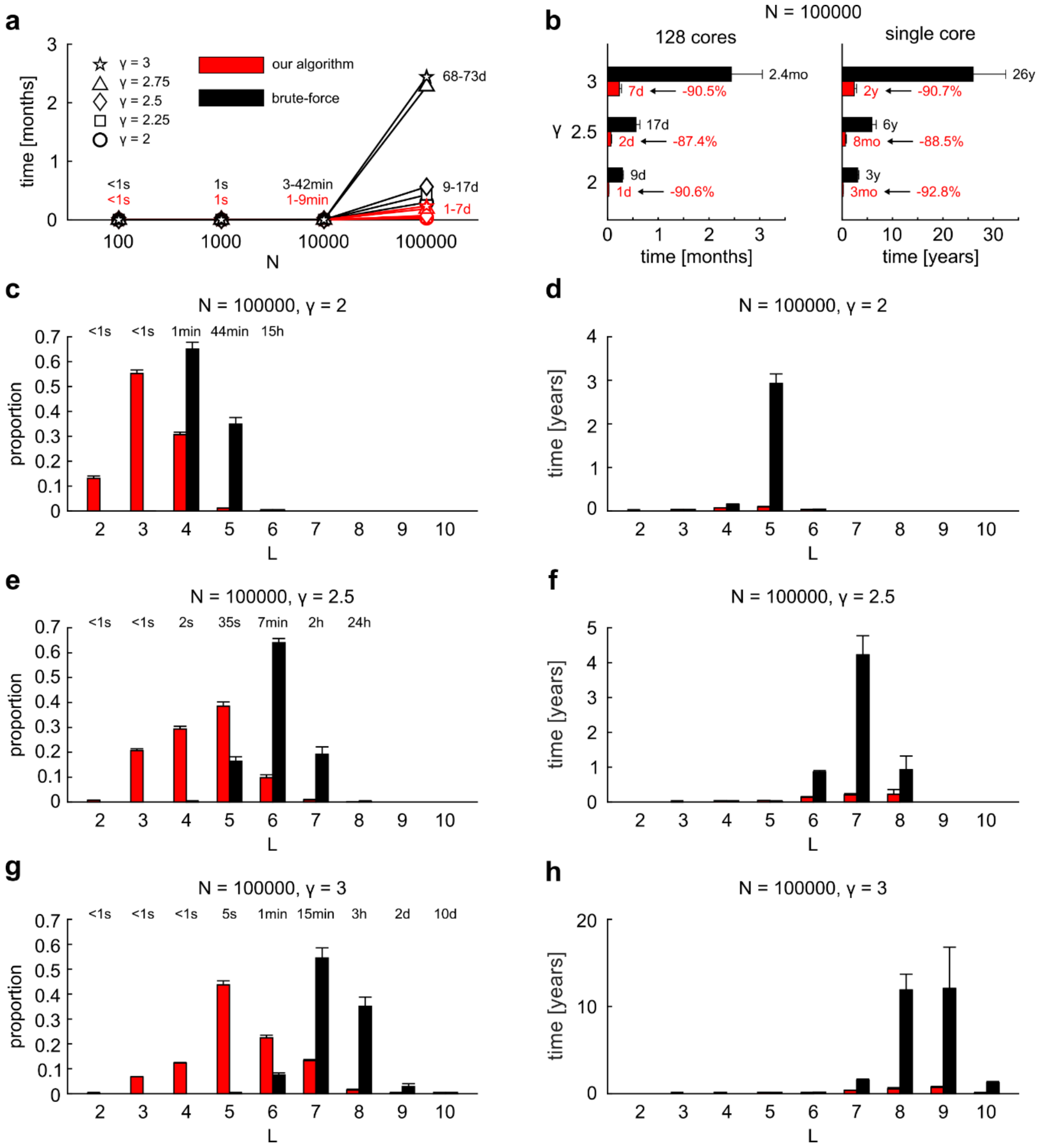
Gamma = 3

Optimized algorithm to find all shortest paths



Gamma = 3

Gamma = 2



a
time [months]
γ = 3
γ = 2.75
γ = 2.5
γ = 2.25
γ = 2
our algorithm
brute-force
<1s
<1s
1s
1s
3-42min
1-9min
68-73d
9-17d
1-7d
N
b
N = 100000
128 cores
single core
2.4mo
7d -90.5%
17d
2d -87.4%
9d
1d -90.6%
26y
2y -90.7%
6y
8mo -88.5%
3y
3mo -92.8%
γ
time [months]
time [years]
c
N = 100000, γ = 2
<1s <1s 1min 44min 15h
proportion
L
d
N = 100000, γ = 2
time [years]
L
e
N = 100000, γ = 2.5
<1s <1s 2s 35s 7min 2h 24h
proportion
L
f
N = 100000, γ = 2.5
time [years]
L
g
N = 100000, γ = 3
<1s <1s <1s 5s 1min 15min 3h 2d 10d
proportion
L
h
N = 100000, γ = 3
time [years]
L

Suppl. Note 2. Pseudocode to compute the $\overline{pTSP}$ between all node pairs.

INPUT

N – number of nodes

A – adjacency list, containing for each node the list of neighbours;
A[1] is the list of neighbours of node 1, A[2] the same for node 2, and so on for N nodes.

G – NxN matrix of geodesics between all node pairs.

T – NxN matrix of topological shortest paths between all node pairs.

OUTPUT

P = NxN matrix of $\overline{pTSP}$ between all node pairs.

```
function P = compute_pTSP(A, T, S, order)

# compute for each node the mean of the topological shortest paths to all other nodes
Tmean = numerical vector of N elements, initialized to zeros                O(N)
for s in [1…N]                                                              O(N)
    for t in [s+1…N]                                                        O(N)
        Tmean[s] = Tmean[s] + T[s,t]                                        O(1)
        Tmean[t] = Tmean[t] + T[s,t]                                        O(1)
    Tmean[s] = Tmean[s] / (N-1)                                             O(1)

# sort nodes by decreasing mean of the topological shortest paths
order = numerical vector of N elements, initialized to zeros                O(N)
order = get_sort_indexes(Tmean, ‘decreasing’)                               O(NlogN)
# the hypothetical function get_sort_indexes sorts the elements of Tmean
# by decreasing order and returns the indexes of the sorted elements

# compute L, which indicates for each node the maximum path length to evaluate
L = numerical vector of N elements, initialized to zeros                    O(N)
mask = logical vector of N elements, initialized to false                   O(N)
for i in [1…N]                                                              O(N)
    s = order[i]                                                            O(1)
    for t in [1…N]                                                          O(N)
        if (mask[t]==false) & (T[s,t]>L[s])                                 O(1)
            L[s] = T[s,t]                                                   O(1)
```

Suppl. Note 2. Pseudocode to compute the $\overline{pTSP}$ between all node pairs.

INPUT
N – number of nodes
A – adjacency list, containing for each node the list of neighbours;
   A[1] is the list of neighbours of node 1, A[2] the same for node 2, and so on for N nodes.
G – NxN matrix of geodesics between all node pairs.
T – NxN matrix of topological shortest paths between all node pairs.

OUTPUT
P = NxN matrix of $\overline{pTSP}$ between all node pairs.

```
function P = compute_pTSP(A, T, S, order)

```

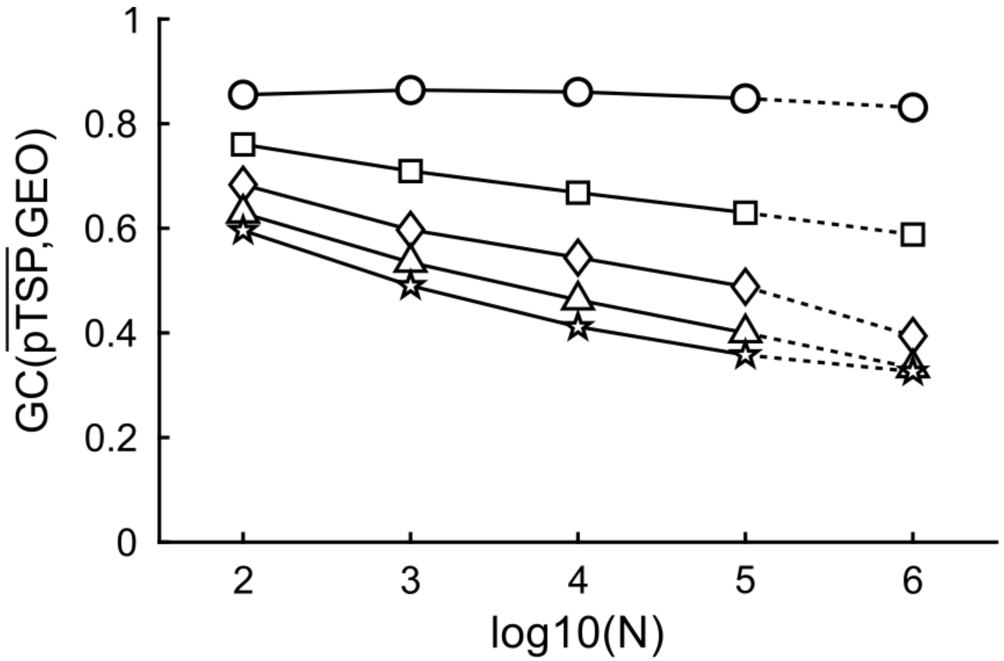
compute for each node the mean of the topological shortest paths to all other nodes

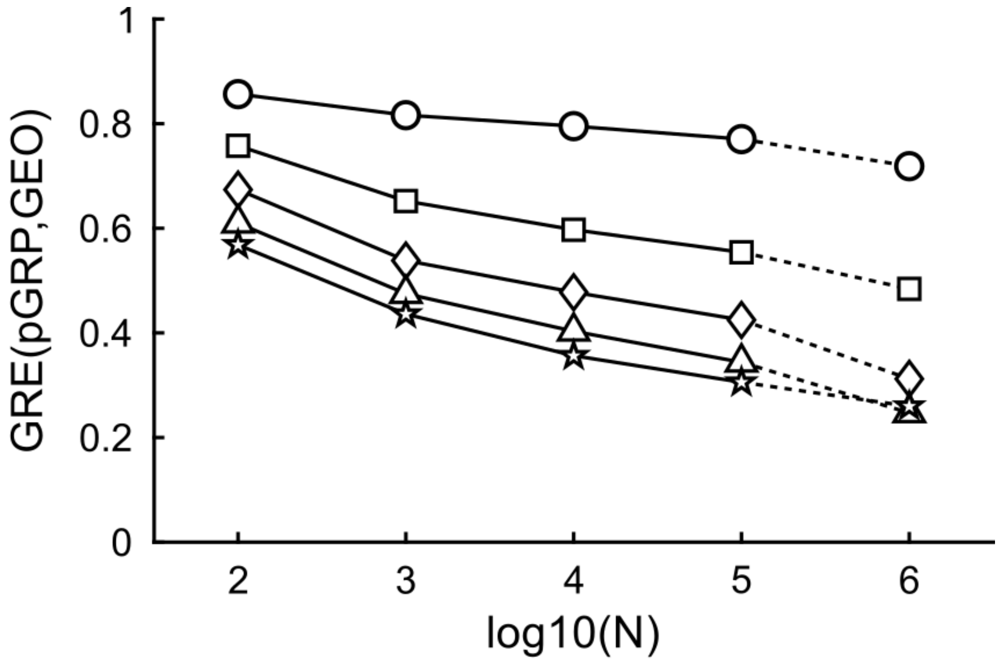
```
for s in [1...N]                                          O(N)
      for t in [s+1...N]                                  O(N)
            Tmean[s] = Tmean[s] + T[s,t]                  O(1)
            Tmean[t] = Tmean[t] + T[s,t]                  O(1)
      Tmean[s] = Tmean[s] / (N-1)                         O(1)
```

sort nodes by decreasing mean of the topological shortest paths

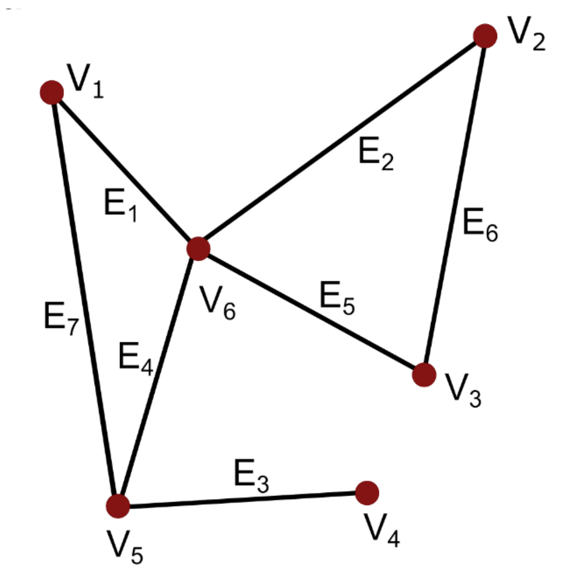
```
order = get_sort_indexes(Tmean, 'decreasing')             O(NlogN)
# the hypothetical function get_sort_indexes sorts the elements of Tmean
# by decreasing order and returns the indexes of the sorted elements

# compute L, which indicates for each node the maximum path length to evaluate
L = numerical vector of N elements, initialized to zeros  O(N)
mask = logical vector of N elements, initialized to false O(N)
for i in [1...N]                                          O(N)
      s = order[i]                                        O(1)
      for t in [1...N]                                    O(N)
            if (mask[t]==false) & (T[s,t]>L[s])           O(1)
                  L[s] = T[s,t]                           O(1)
```


GC(pTSP,GEO)
1
0.8
0.6
0.4
0.2
0
2
3
4
5
6
log10(N)


GRE(pGRP,GEO)
1
0.8
0.6
0.4
0.2
0
2
3
4
5
6
log10(N)

Applications



Brain Network of an individual

Navigability measure (previous)
Muscoloni,···, Cannistraci et al. Nat. Com. 2017

$$GRE(pGRP, RD) = \left(\frac{1}{n \cdot (n-1) - 2 \cdot e}\right) \cdot \sum \frac{RD(\mathrm{i},\mathrm{j})}{pGRP(\mathrm{i},\mathrm{j})} \; ; \quad \text{with } (\mathrm{i},\mathrm{j}) \in \tilde{E}$$

PNAS

ARTICLES FRONT MATTER AUTHORS TOPICS +

RESEARCH ARTICLE | BIOLOGICAL SCIENCES |

Navigation of brain networks

Caio Seguin, Martijn P. van den Heuvel, and Andrew Zalesky Authors Info & Affiliations

Edited by Edward T. Bullmore, University of Cambridge, Cambridge, United Kingdom, and accepted by Editorial Gazzaniga May 7, 2018 (received for review January 24, 2018)

May 30, 2018 | 115 (24) 6297-6302 | https://doi.org/10.1073/pnas.1801351115

Navigability measure (previous)
Muscoloni,⋯, Cannistraci et al. Nat. Com. 2017

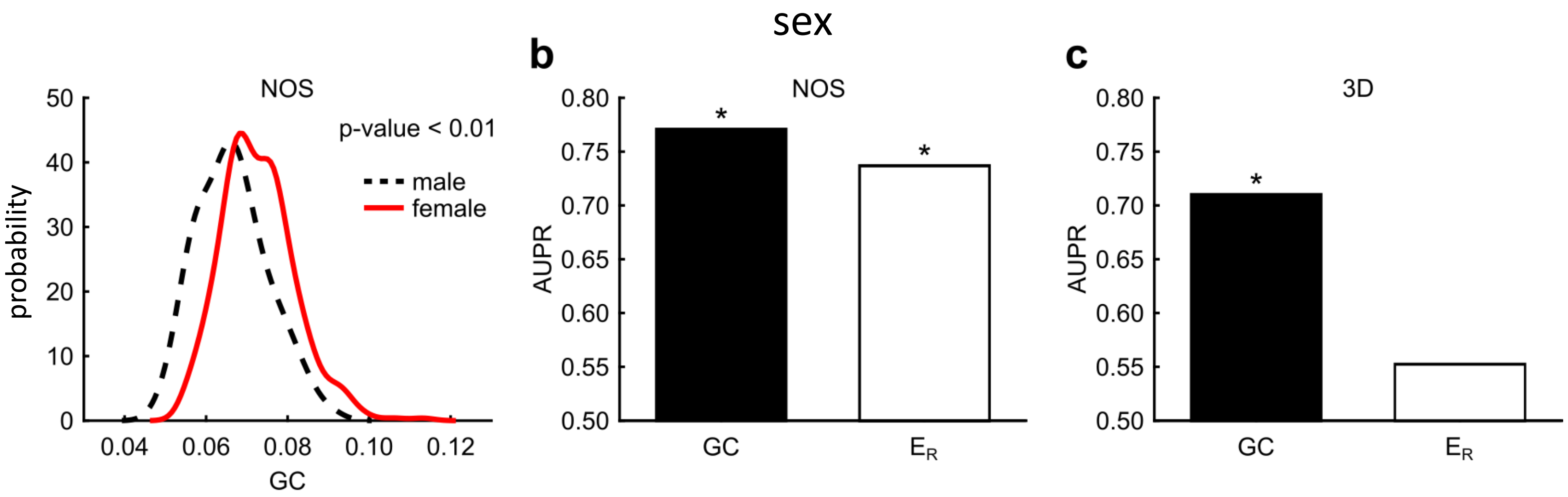
$$GRE(pGRP, RD) = \left(\frac{1}{n \cdot (n-1) - 2 \cdot e}\right) \cdot \sum \frac{RD(\mathrm{i},\mathrm{j})}{pGRP(\mathrm{i},\mathrm{j})} \; ; \quad \text{with } (\mathrm{i},\mathrm{j}) \in \tilde{E}$$

Congruency measure (New)
Cannistraci et al. Nat. Com. 2022

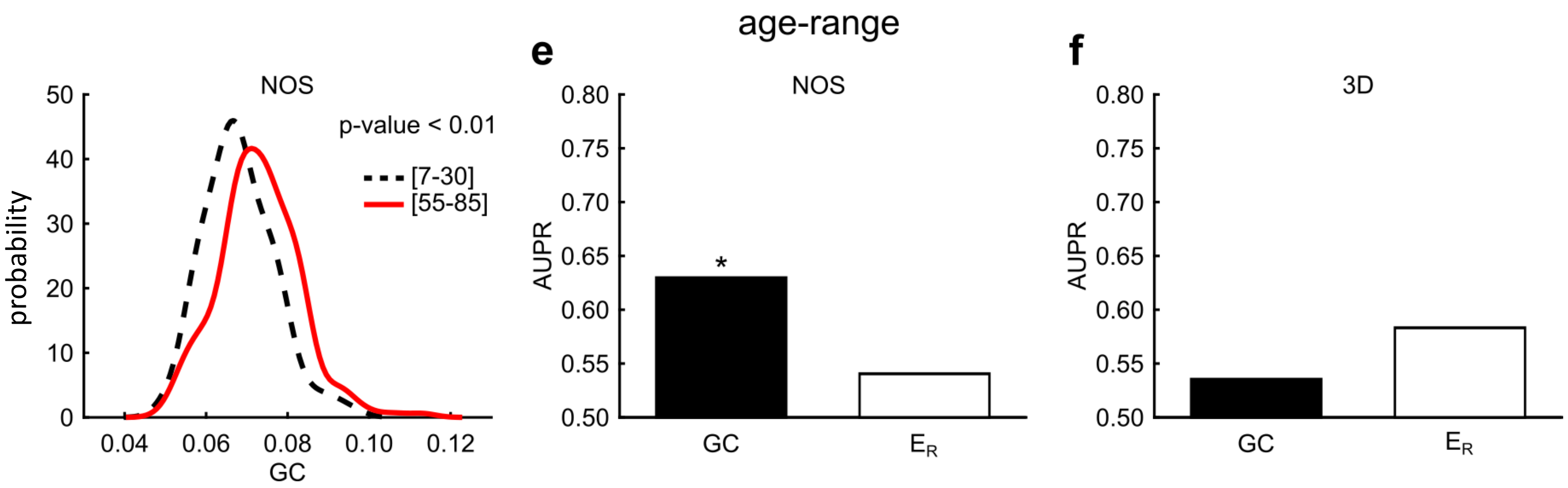
$$GC(\overline{pTSP}, RD) = \left(\frac{2}{n \cdot (n-1) - 2 \cdot e}\right) \cdot \sum_{i<j} \frac{RD(\mathrm{i},\mathrm{j})}{\overline{pTSP}(\mathrm{i},\mathrm{j})} \; ; \quad \text{with } (\mathrm{i},\mathrm{j}) \in \tilde{E}$$

Testing GC as a phenotypic marker on Human brain connectomes

230 males
Vs.
384 females



223 in [7 to 30]
Vs.
215 in [55 to 85])



nature communications

Article https://doi.org/10.1038/s41467-022-34634-6

Geometrical congruence, greedy navigability and myopic transfer in complex networks and brain connectomes

Received: 6 July 2020

Accepted: 1 November 2022

Published online: 27 November 2022

Check for updates

Carlo Vittorio Cannistraci [1,2,3,4,5,6] & Alessandro Muscoloni [1,4]

We introduce in network geometry a measure of *geometrical congruence (GC)* to evaluate the extent a network topology follows an underlying geometry. This requires finding all topological shortest-paths for each nonadjacent node pair in the network: a nontrivial computational task. Hence, we propose an optimized algorithm that reduces 26 years of worst scenario computation to one week parallel computing. Analysing artificial networks with patent geometry we discover that, different from current belief, hyperbolic networks do not show in general high GC and efficient greedy navigability (GN) with respect to the geodesics. The myopic transfer which rules GN works best only when degree-distribution power-law exponent is strictly close to two. Analysing real networks—whose geometry is often latent—GC overcomes GN as marker to differentiate phenotypical states in macroscale structural-MRI brain connectomes, suggesting connectomes might have a latent neurobiological geometry accounting for more information than the visible tridimensional Euclidean.

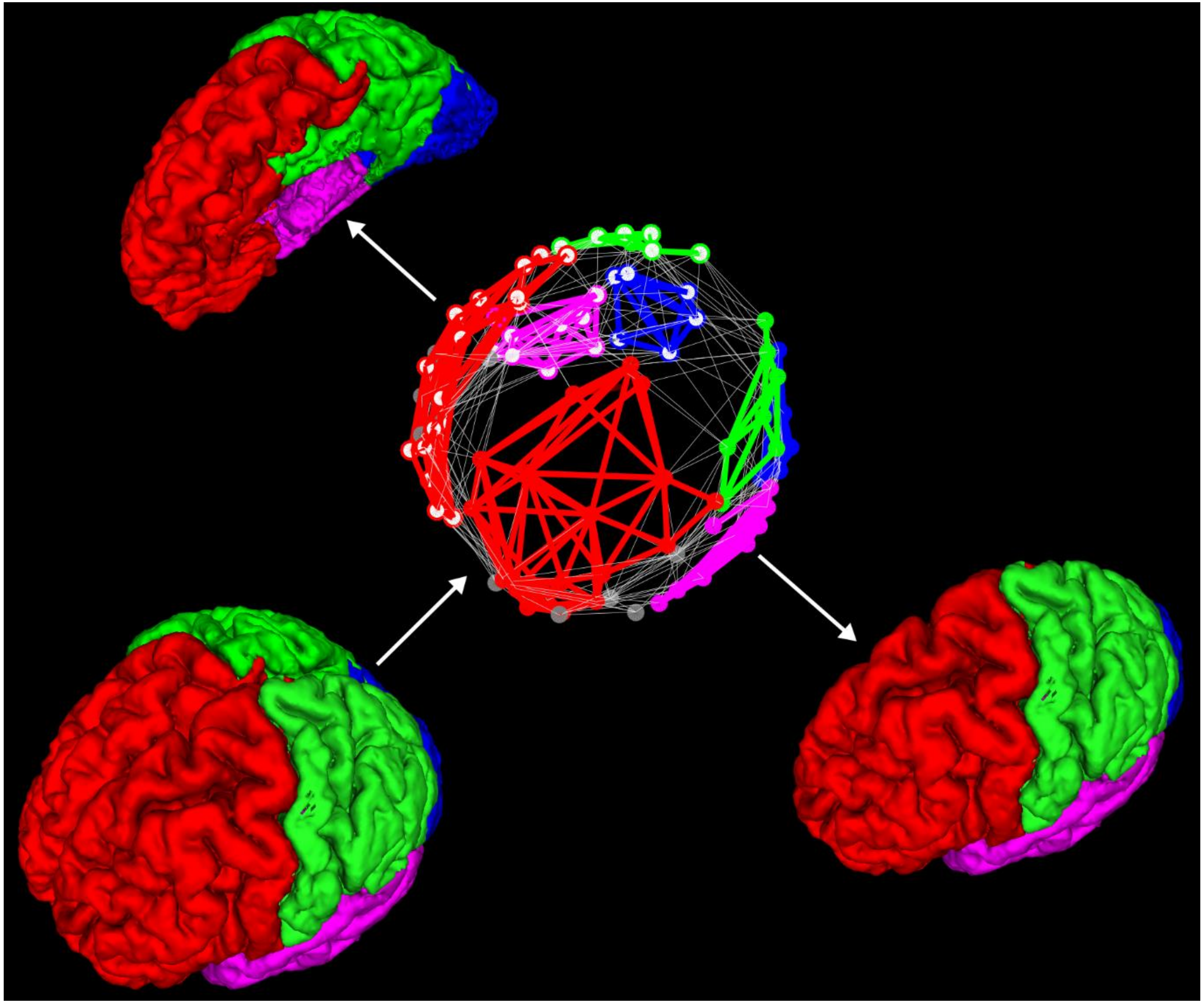
November 2022

Alessandro Muscoloni

Next direction

- Application of congruence to spatial network in general with any underlying geometry
- Application of congruence to design markers for brain diseases
- Application of navigability and congruency to urban science and human mobility networks.

Problem 5
(De Novo single cell spatial reconstruction by Coalescence Embedding)



Jing-Dong (Jackie) Han
Peking University

The idea of Jackie

<< To leverage the general concept of network embedding **angular coalescence** and the methodology of coalescence embedding for De novo (landmark and marker free) reverse-engineering the mesoscale spatial organization of single cell directly from their transcriptome. >>

ADVANCED SCIENCE
Open Access

Research Article | Open Access

Spatial Reconstruction of Oligo and Single Cells by De Novo Coalescent Embedding of Transcriptomic Networks

Yuxuan Zhao, Shiqiang Zhang, Jian Xu, Yangyang Yu, Guangdun Peng, Carlo Vittorio Cannistraci, Jing-Dong J. Han

First published: 15 June 2023 | https://doi.org/10.1002/advs.202206307

Theoretical Background and Method

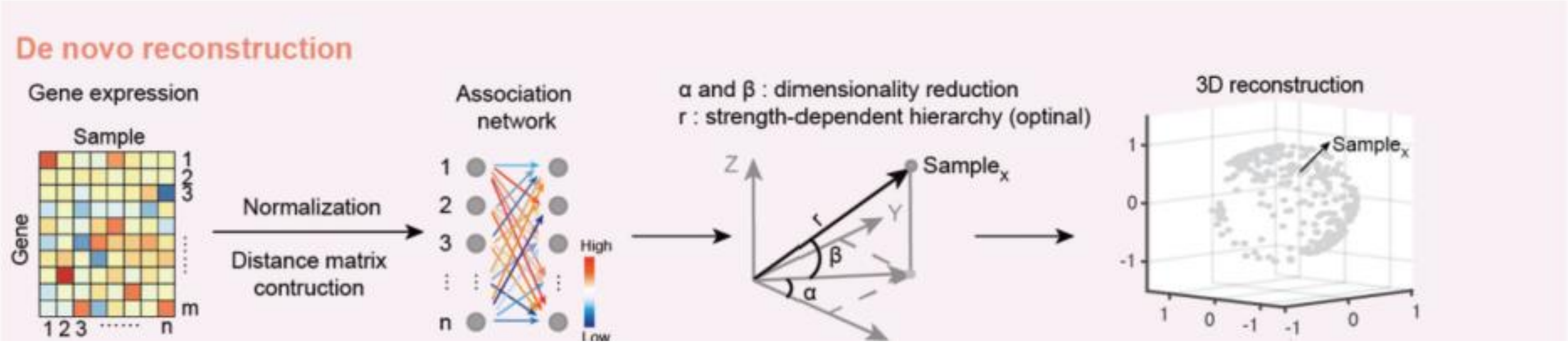
Coalescent embedding (CE)

This is a machine intelligence theory for nonlinear embedding of networks of complex interconnected systems in a geometrical space, which is called coalescent embedding (CE) because it relies on a phenomenon that in physics of complexity takes the name of angular coalescence. This phenomenon states that for a network that derives from a complex interconnected system, whose connections between its parts (nodes) emerge in a latent geometrical space, the network embedding in a 2D or 3D visualization space will display a typical pattern of node aggregation that respects the intrinsic geometry of the system in the latent geometrical space in terms of both congruence and navigability.

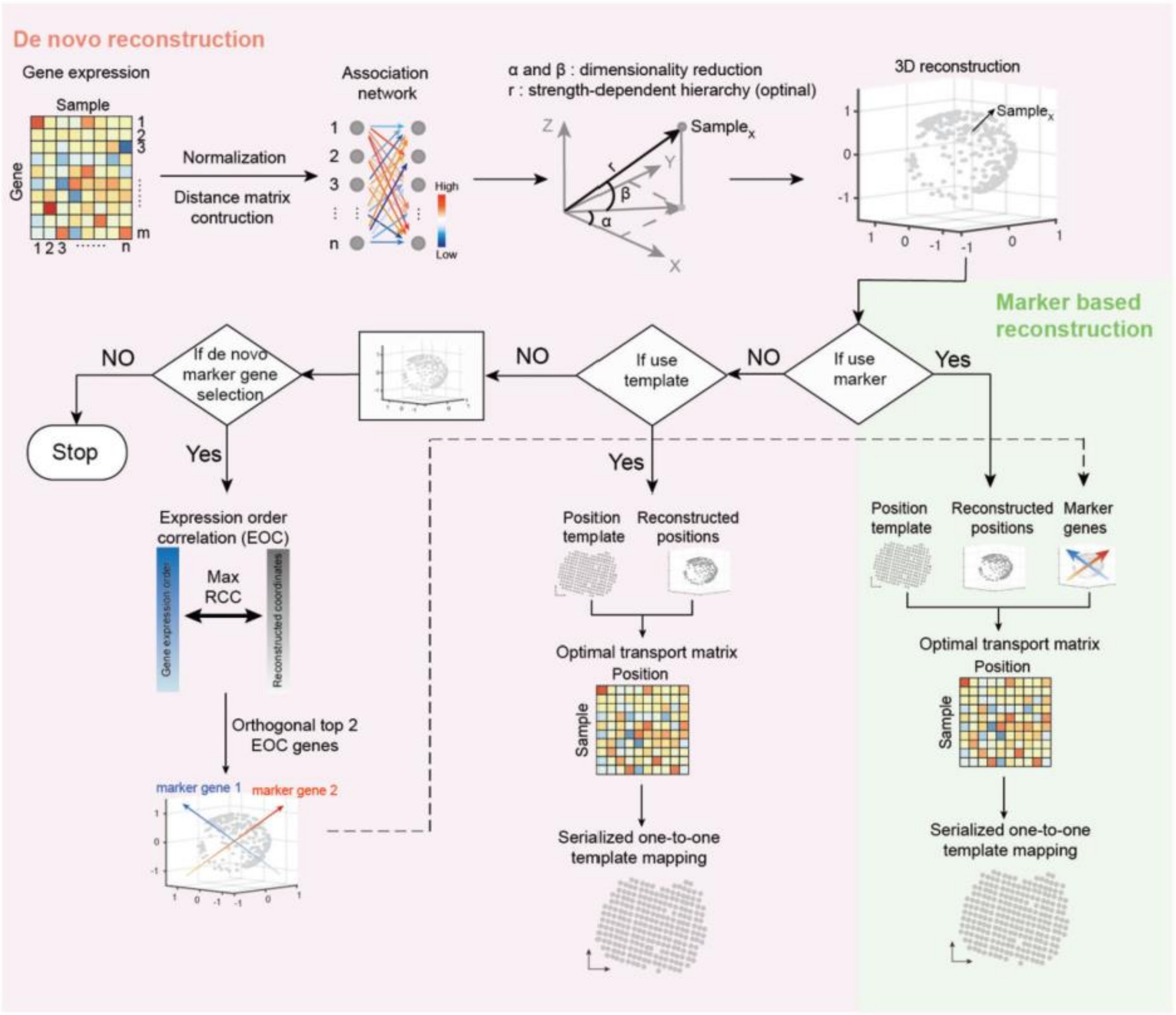
De Novo Coalescent Embedding (D-CE)

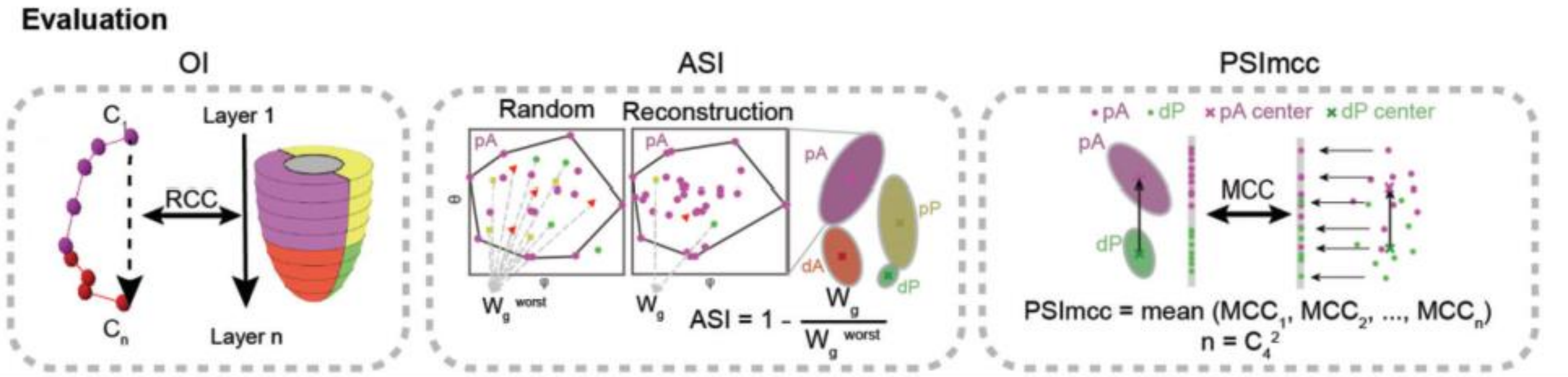
Building upon this theory, we developed a novel algorithm called De Novo Coalescent Embedding (D-CE) which unveils single-cell mesoscale spatial organization, where densely interacting network neighborhoods or communities are associated with spatial domains.

D-CE algorithm



D-CE algorithm





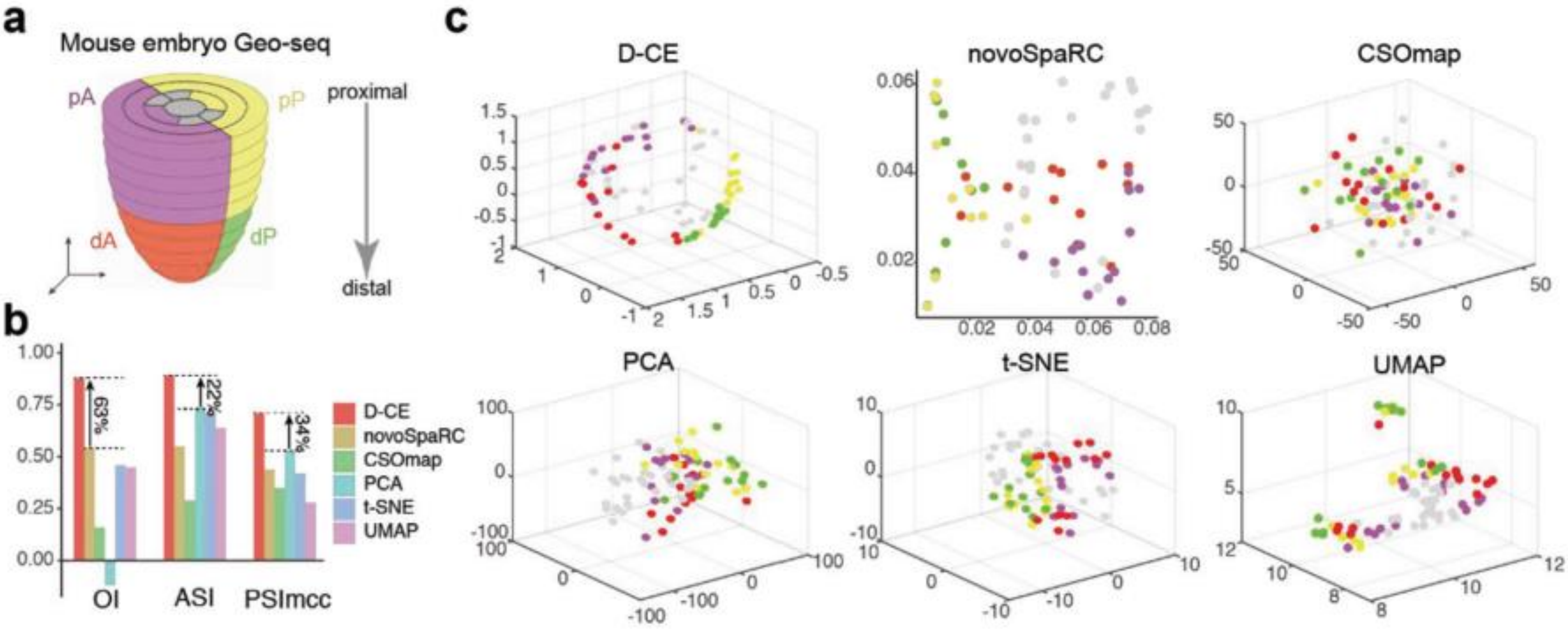
Ordering index (OI)
(ranging: -1 to 1)

Angular separation index (ASI)
(ranging: 0 to 1)

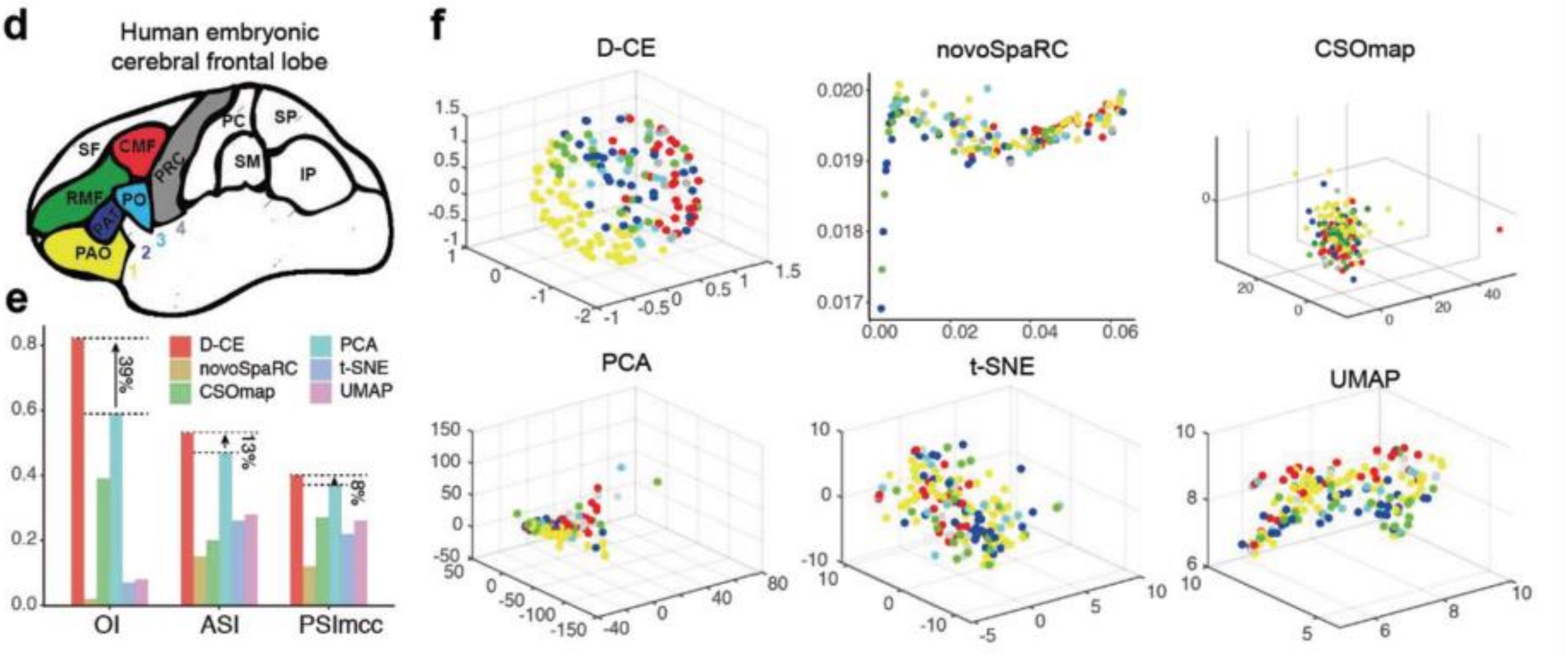
Projection separability Index
(ranging: -1 to 1)

Computational Results

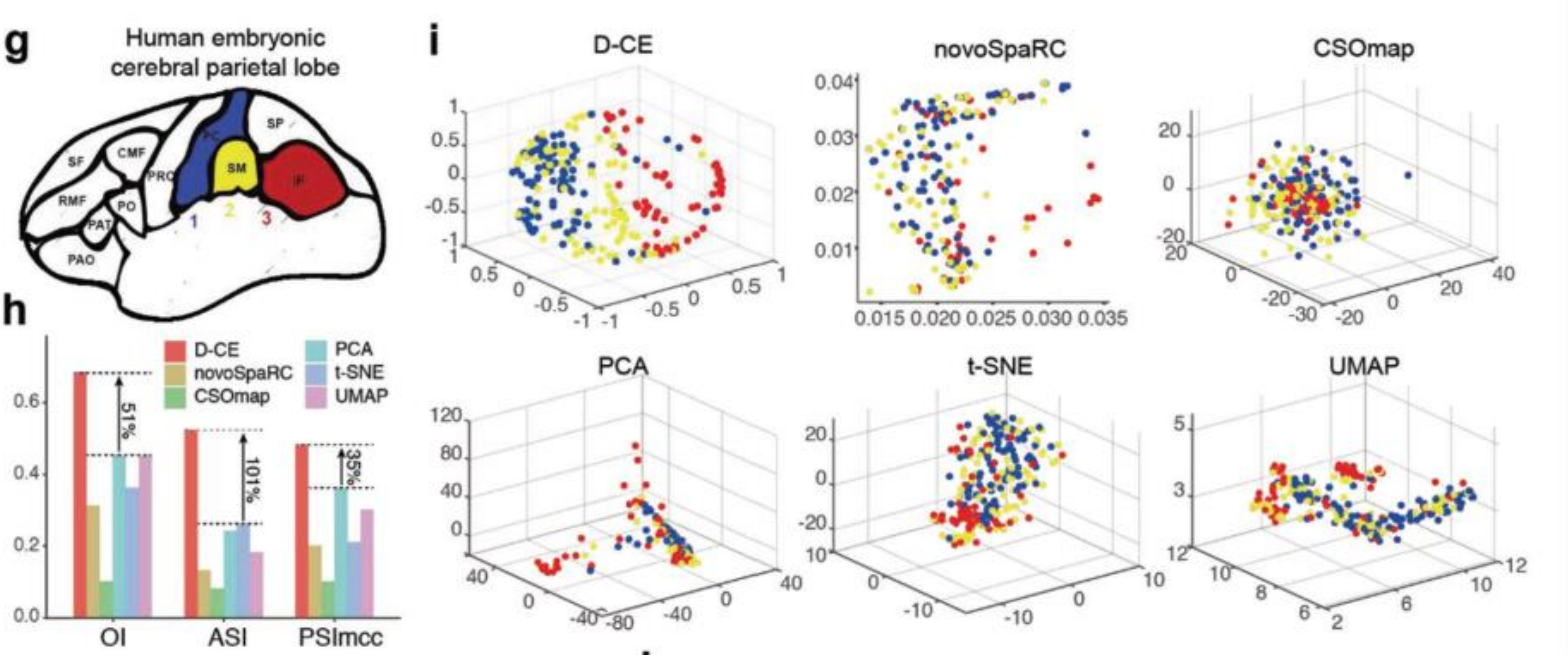
Reconstruction of spatial domain labels from oligo or single cell RNA-seq data



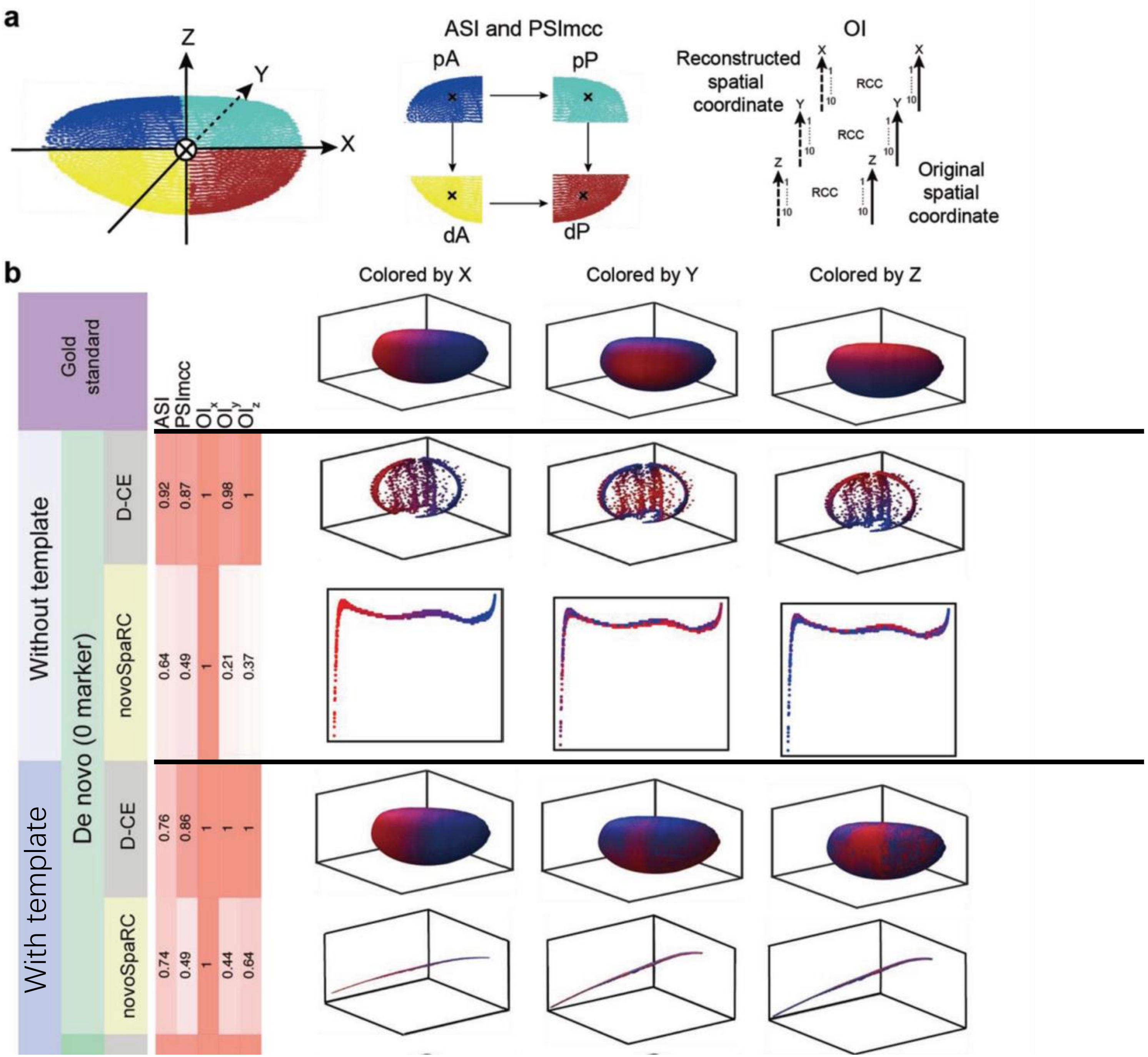
Reconstruction of spatial domain labels from oligo or single cell RNA-seq data



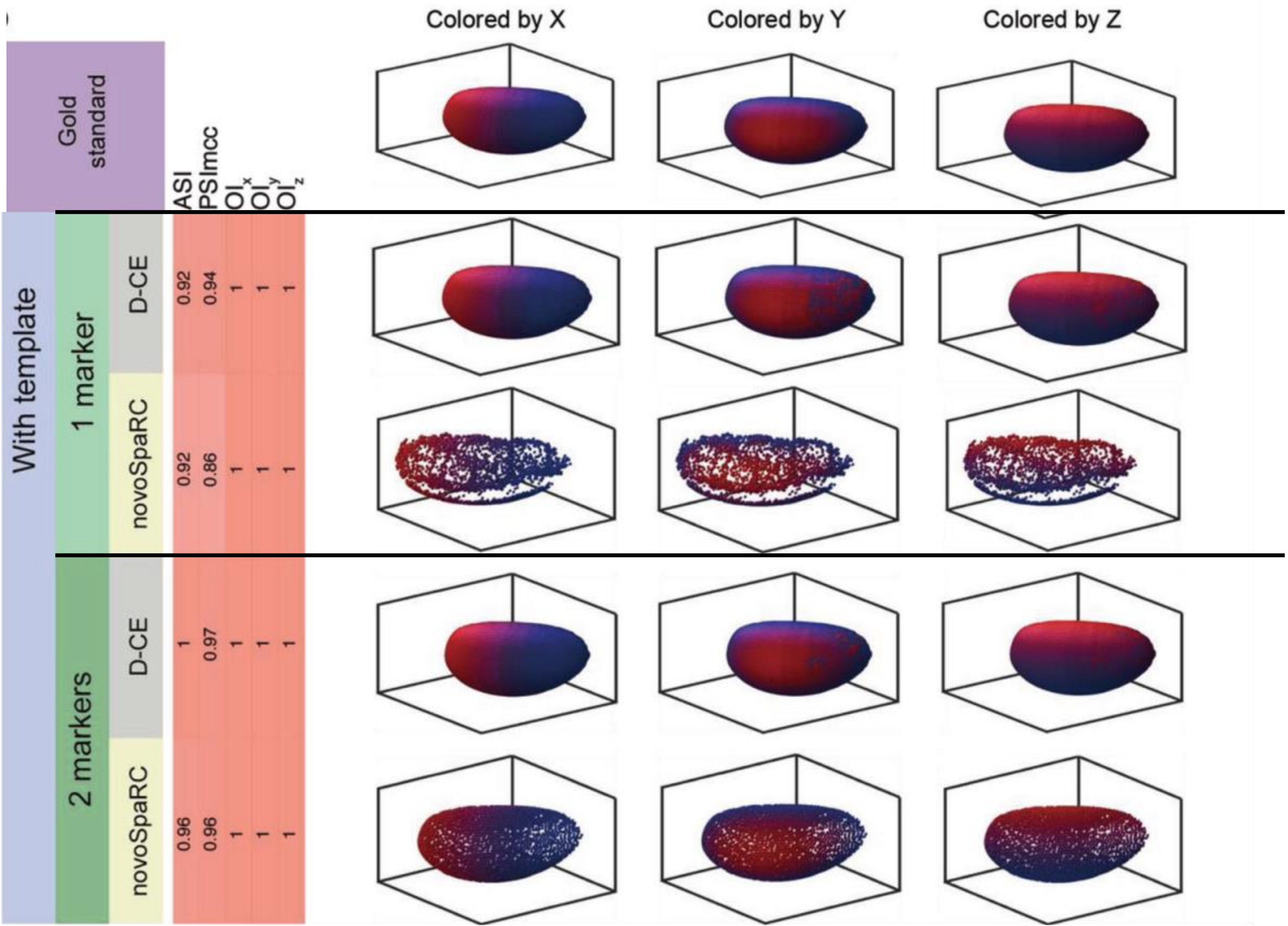
Reconstruction of spatial domain labels from oligo or single cell RNA-seq data



Drosophila embryo segmentation



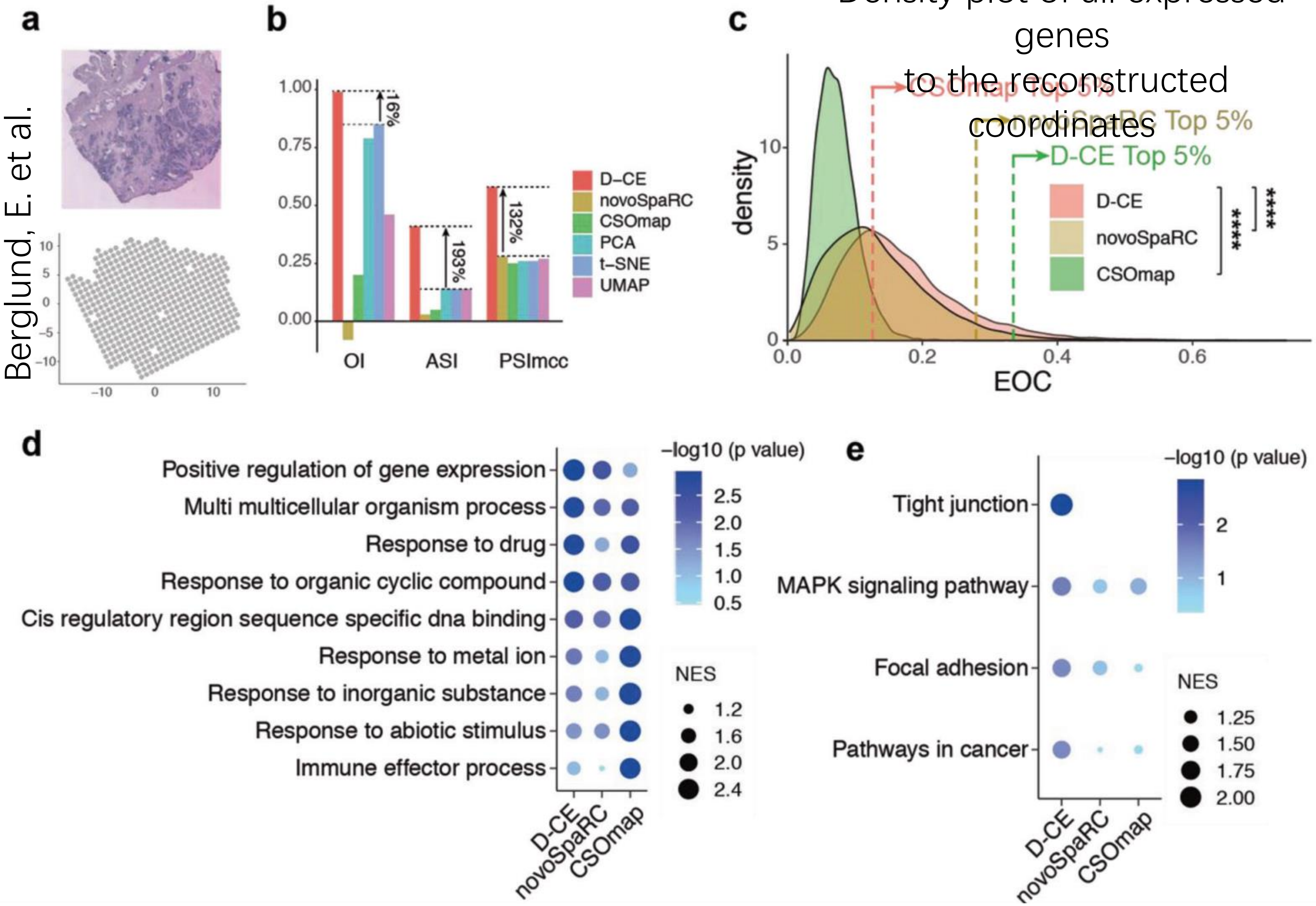
Drosophila embryo segmentation



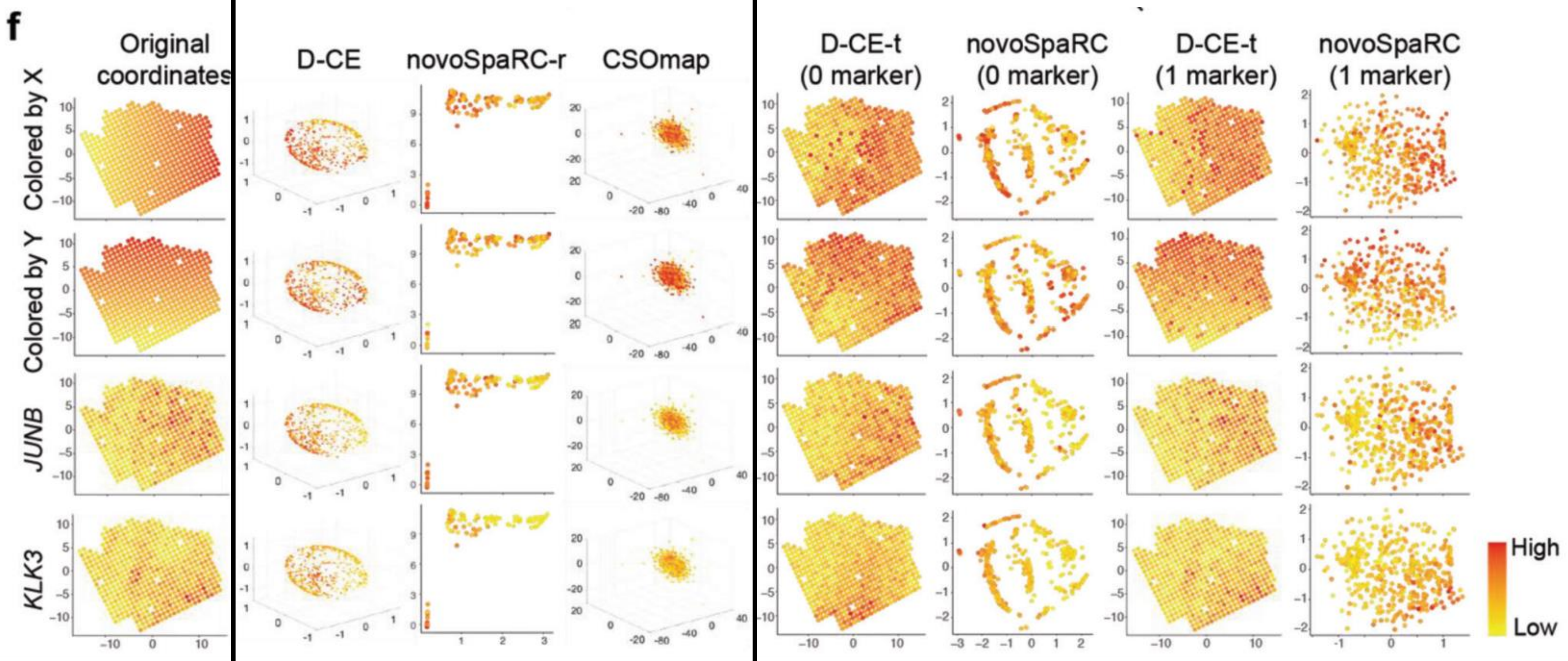
Spatial reconstruction and spatial marker gene detection of cancerous prostate spatial transcriptomic dataset

Density plot of all expressed genes
to the reconstructed coordinates

Berglund, E. et al.



Spatial reconstruction and spatial marker gene detection of cancerous prostate spatial transcriptomic dataset



Original

without marker and template fitting

marker and template fitting

1. We developed D-CE which is an effective **landmark free and model free** de novo 3D reconstruction method for oligo and single cell analysis.

2. The proposed algorithm for de novo coalescent embedding (D-CE) of oligo/single cell transcriptomic networks **is based on the physics principle of angular coalescence** and relies **only** on the spatial information encoded in the expression patterns of genes, without need of prior information.

3. We found that D-CE of cell–cell association transcriptomic networks, by preserving mesoscale network organization, **captures spatial domains,** identifies spatially expressed genes, reconstructs cell samples' 3D spatial distribution, and **uncovers spatial domains and markers necessary** for understanding the design principles on spatial organization and pattern formation.

4. Comparison to the novoSpaRC and CSOmap on **14 datasets and 497 reconstructions**, reveals a **significantly superior performance of D-CE**.

5. Angular coalescence: the same principle of organization of complex connected systems can be used to analyze systems at different scale from single-cell to brain to society.

Scientific Institutions

San Raffaele
Scientific Institute

Scuola Interpolitecnica

- Lipotype (Germany)
- Politecnico di Torino and Milano (Italy)
- Italian Interpolytechnic School of Doctorate (SIPD, Italy)
- San Raffaele Scientific Institute, Hospital and University (Italy)
- King Abdullah University of Science and Technology (KAUST, Saudi Arabia)
- University of California San Diego/ Ideker Lab (USA)
- ISMB/ECCB generously provided me (International)
- Italian National Research Council (CNR) / Bioengineering (It
- Technical University Dresden (Germany)
- Klaus Tschira Foundation (Germany)
- FANTOM Consortium and RIKEN institute (Japan)

biotec
Biotechnology Center TU Dresden

center for systems biology dresden

lipotype
Lipidomics for a better life

STVDIORVM VNIVERSITAS MESSANAE

CENTRO NEUROLESI
BONINO PULEJO
IRCCS MESSINA

Politecnico di Torino

清华大学脑与智能实验室
Tsinghua Laboratory of Brain and Intelligence

KAUST
King Abdullah University of Science and Technology

UC San Diego
SCHOOL OF MEDICINE

清華大學
Tsinghua University

TECHNISCHE UNIVERSITÄT DRESDEN

Carlo Vittorio Cannistraci

Center for Complex Network Intelligence

EMAIL: kalokagathos.agon@gmail.com

Thanks!

Tsinghua University