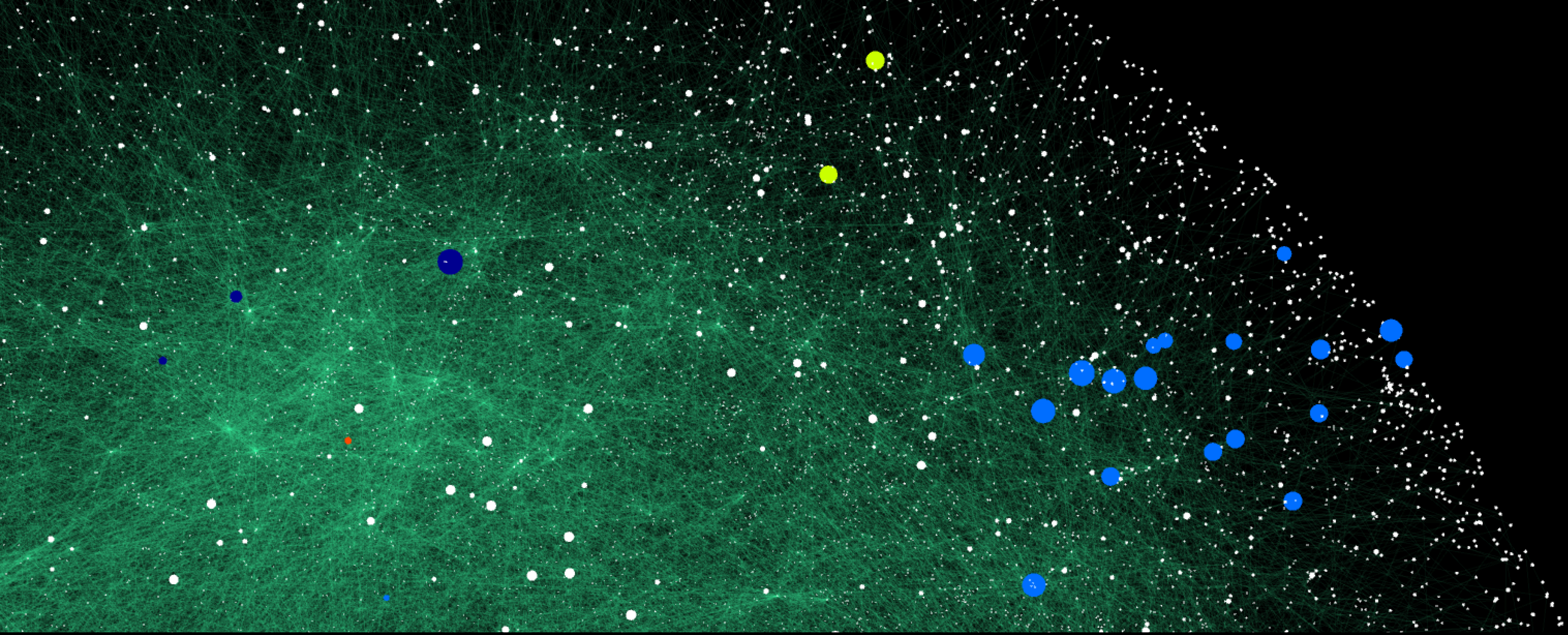


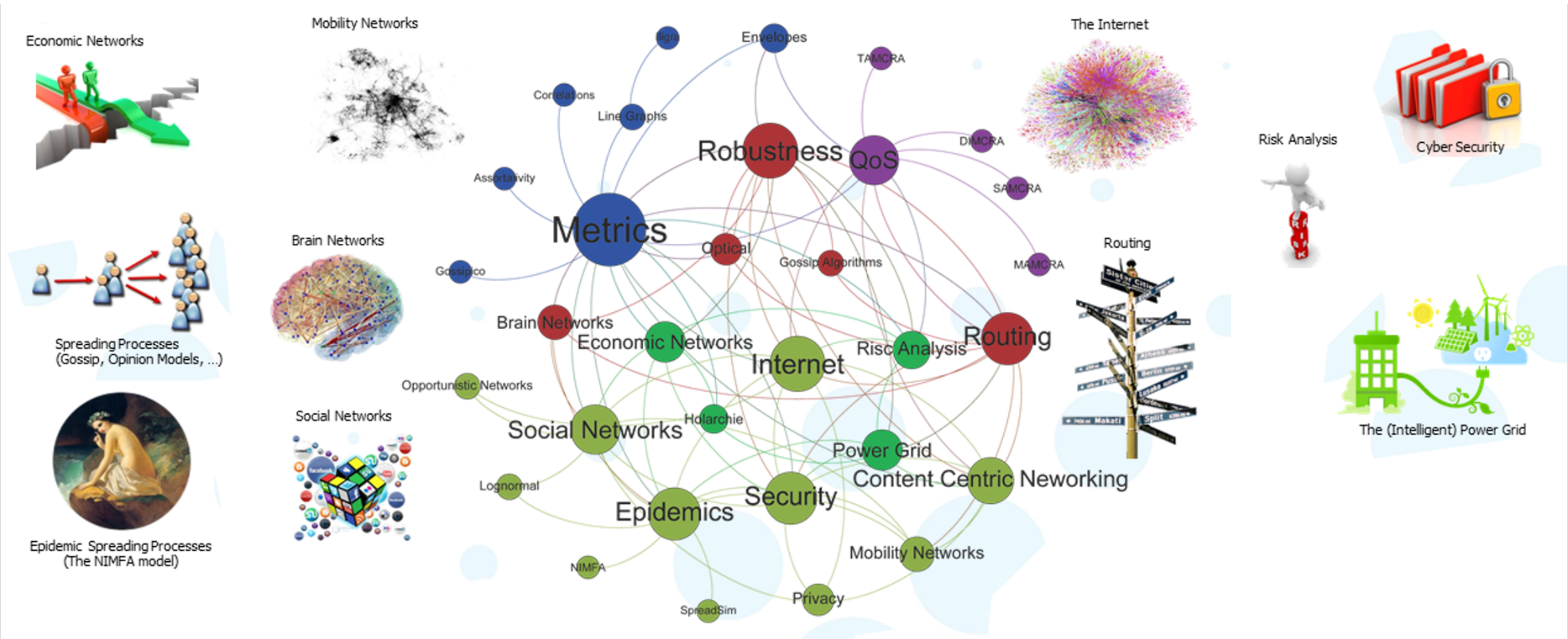
Geometric Representations of Complementarity-Driven Networks

Maksim Kitsak, Ph.D.

Electrical Engineering, Mathematics and Computer Science
Delft University of Technology

GRAPH&DATA Seminar, TU Delft, October 5, 2023

Network Architecture & Services @ EWI/TU Delft

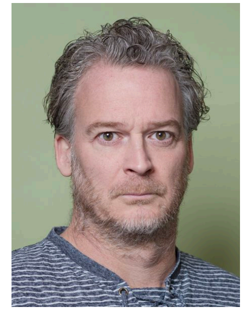


We are interested in networks, broadly defined

Network Architecture & Services @ EWI/TU Delft

Staff members

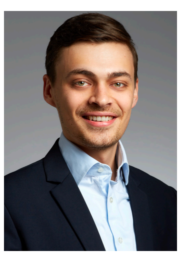


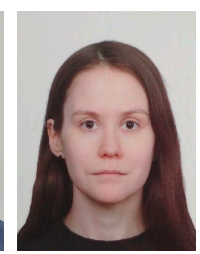
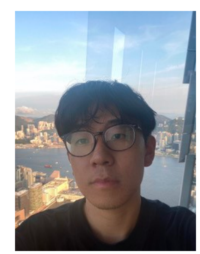
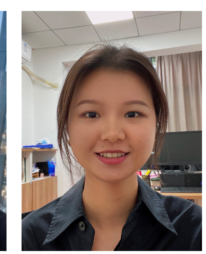
Senior researchers



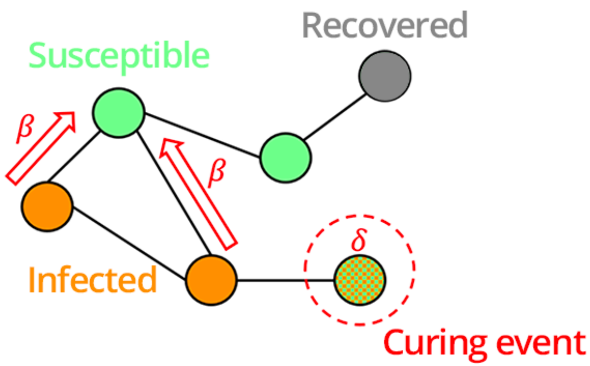
Junior researchers

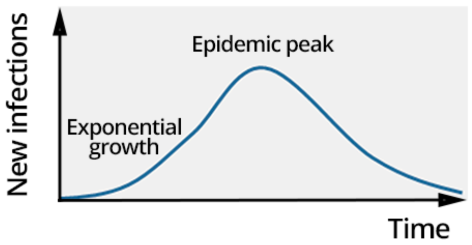


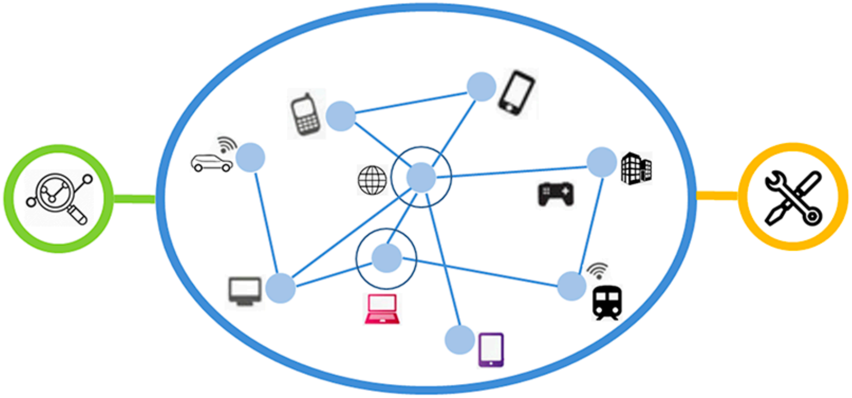

NAS Section: Research Interests



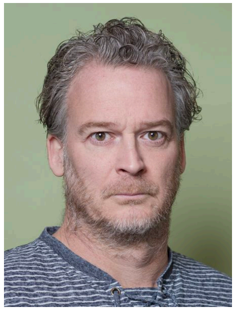
Epidemics on Networks



NAS Section: Research Interests

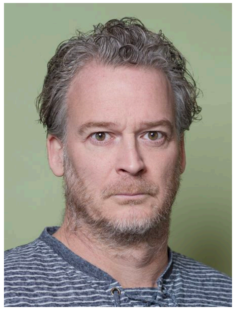


Control of Networks

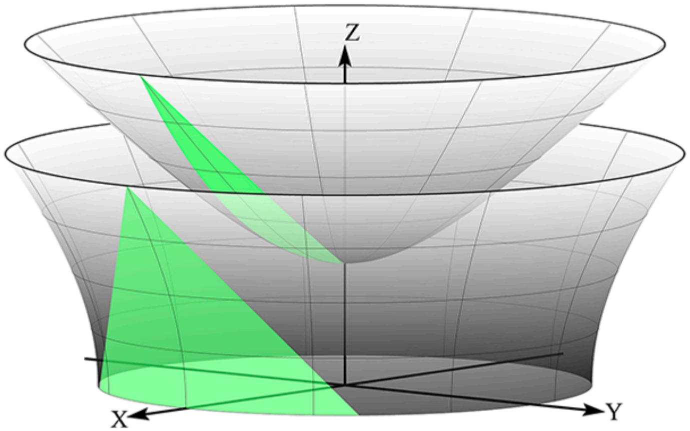


NAS Section: Research Interests

Next generation (5/6G) wireless




Network Geometry



Network Geometry aka Network Embeddings

Network <u>Node</u> Embedding:

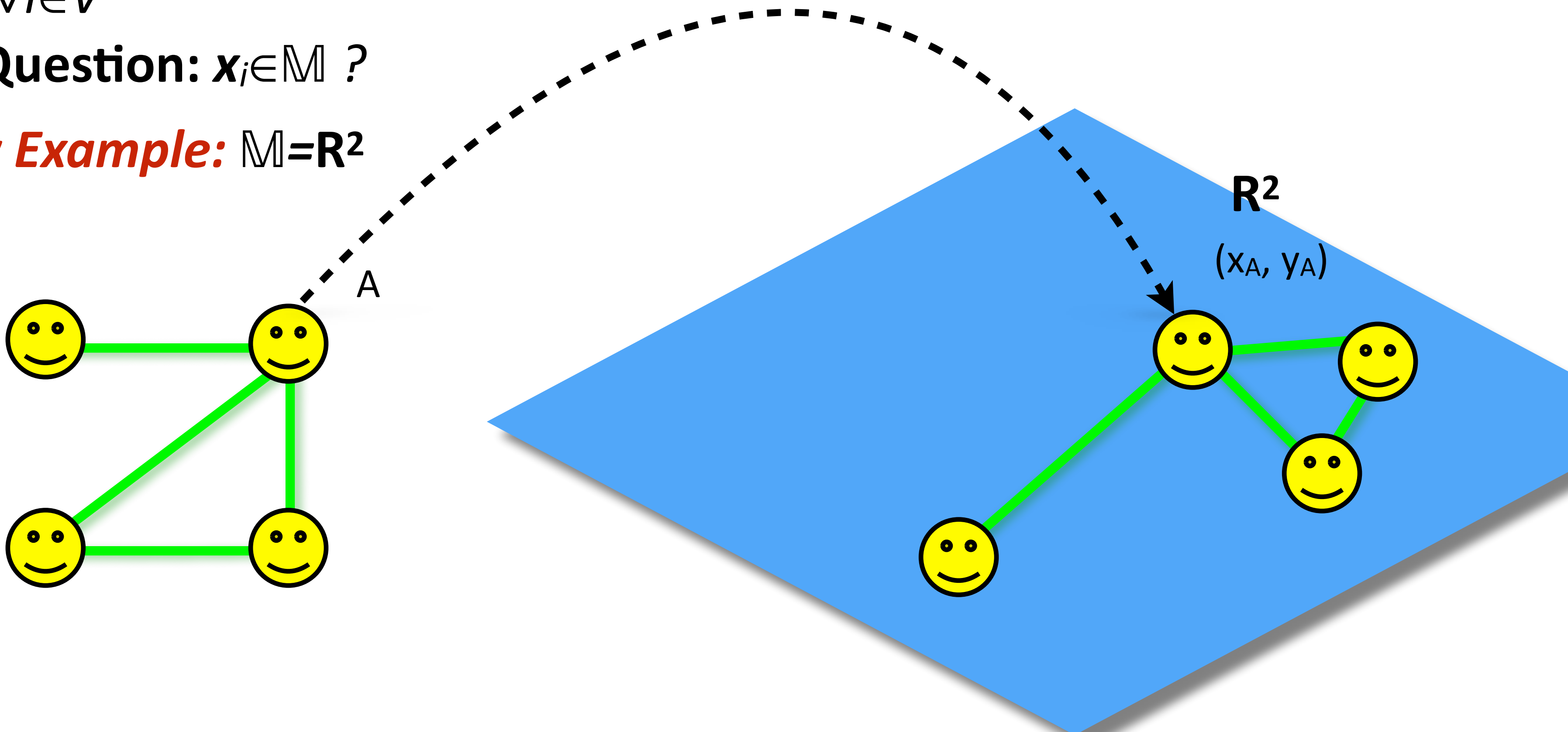
Mapping of a network to a (relatively) low-dimensional (metric) space.

Given: network *G(V,E)* and *space* $\mathbb{M}$.

for $\forall i \in V$

Question: $\boldsymbol{x}_i \in \mathbb{M}$?

Toy Example: $\mathbb{M}=\mathbf{R}^2$



Interpretation: $\boldsymbol{d}_{ij}$ reflects similarity between *i* and *j*:

smaller $\boldsymbol{d}_{ij}$ implies higher similarity.

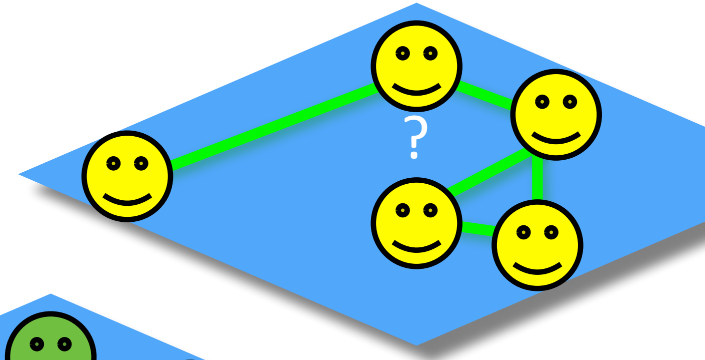
Network Embeddings: Applications

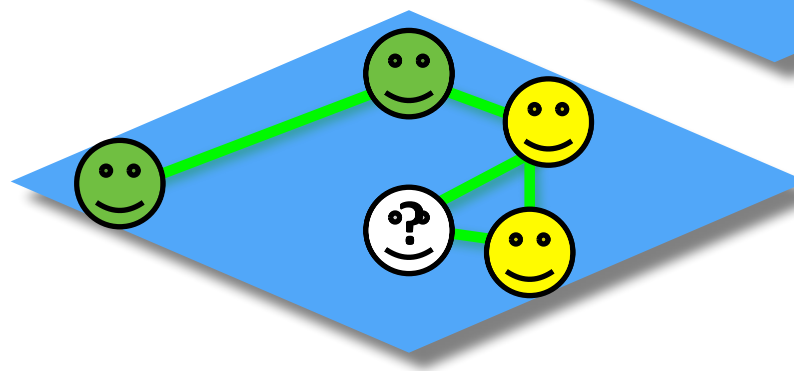
Interpretation: d_{ij} reflects similarity between i and j:

smaller d_{ij} implies higher similarity.

Application 1 (<u>Network reconstruction</u>):

e.g., unconnected nodes at small distances are missing links





Application 2 (<u>Classification</u>):

e.g., learn missing node attributes

Advanced Applications:

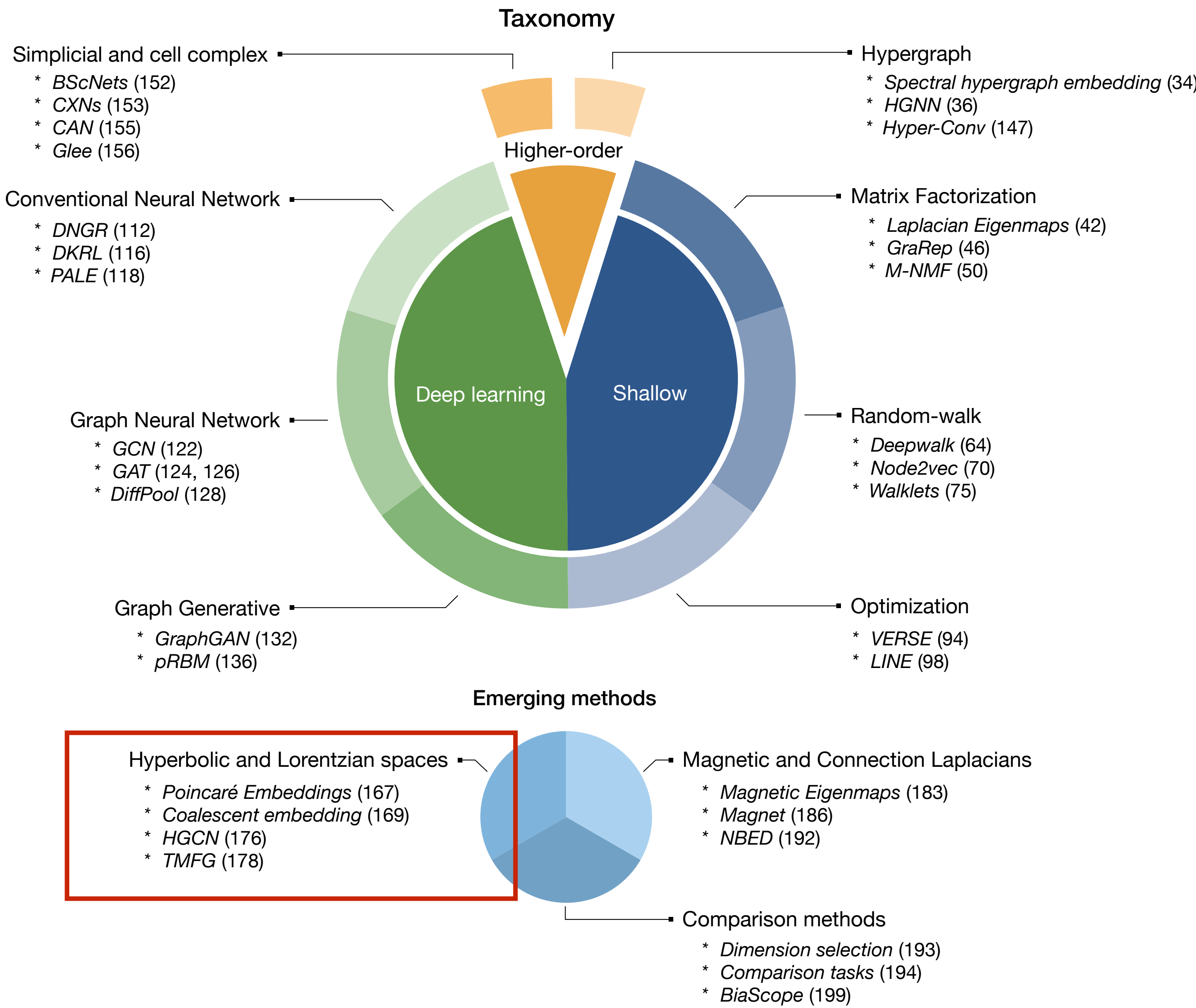
<u>Search</u>, e.g., S. Ratnasamy et al, *ACM SIGCOMM CCR* (2001).

<u>Routing,</u> e.g., M. Boguñá et al, *Nature Communications* (2010).

<u>Shortest Path Finding in Incomplete Networks</u>

M. Kitsak et al, *Nature Communications* (2023).

Taxonomy of Network Embeddings

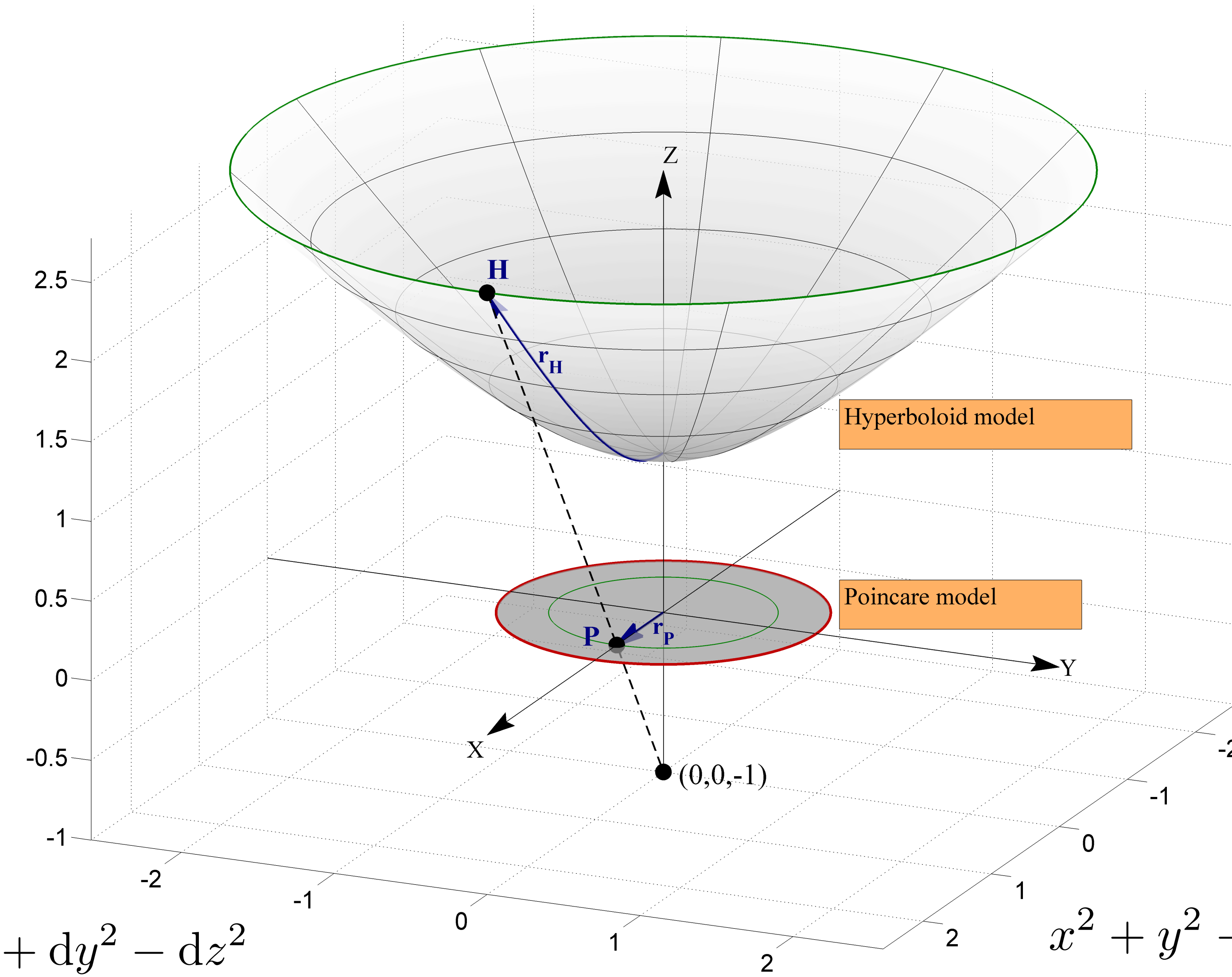


A. Baptista et al, arXiv 2305.03474 (2023).

Hyperbolic Network Embeddings

Embedding space is a low-dimensional hyperbolic disk, $\mathbb{M}=H^D$

Hyperbolic = Curvature K < 0



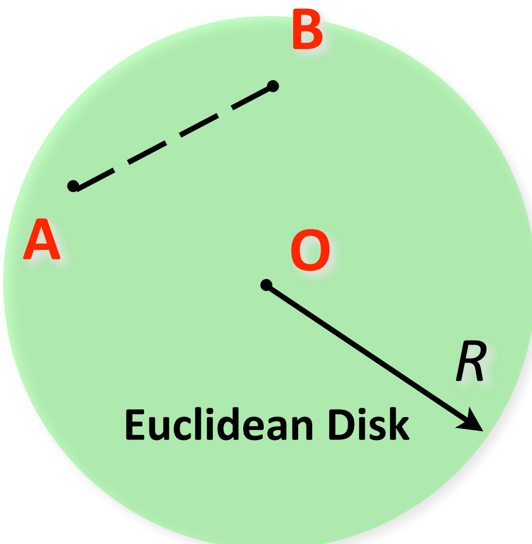
$$ds^2 = dx^2 + dy^2 - dz^2$$

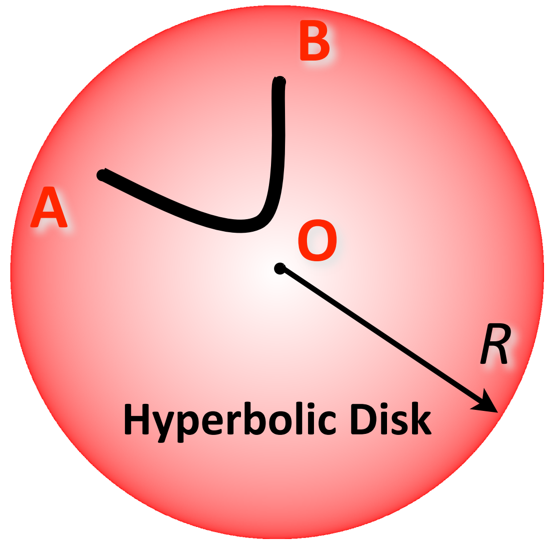
$$x^2 + y^2 - z^2 = -1$$

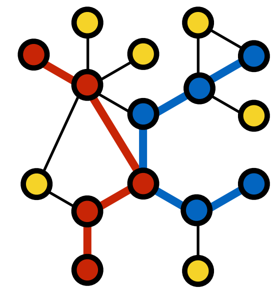
Why Hyperbolic Space?

Volume: $V \propto R^2$ $\quad$ $V \propto Cosh(R) \sim Exp(R)$ $\quad$ $d \sim Ln(N)$

Geodesic: ***Straight line*** $\quad$ ***"Biased" towards center*** $\quad$ ***Shortest paths tend to pass large degree nodes***







Network

Hyperbolic disk expands exponentially!
Hyperbolic geodesics are "biased" towards disk center!

D. Krioukov et al, *Phys.Rev. E.* (2010)
M. Kitsak et al, *Phys. Rev. Research* (2020)
M. Boguñá et al, *Nat. Rev. Physics* (2021)
M. Kitsak et al, *Nat. Comm.* (2023)

Are All Networks Embeddable (Effectively Geometric)? Not at all!

Counter Example: *G(n,p)* Connect each node pair with probability *p*

Open Hard Questions that are Underrated:

Q1: Is the network of interest effectively geometric?

D. Krioukov, "Clustering implies geometry in networks", Phys. Rev. Lett. (2016).

M. Boguñá et al, "Small worlds and clustering in spatial networks", Phys. Rev. Res. (2020).

Q2: What are the properties of the latent space $\mathbb{M}$?

E.g., curvature? dimensionality? volume?

Y. Ollivier, "Ricci curvature of metric spaces", *C. R. Math. Acad. Sci. Paris* (2007).

P. van der Hoorn et al, "Ollivier Curvature of Random Geometric Graphs Converges to Ricci Curvature of Their Riemannian Manifolds", *Discrete & Computational Geometry* (2023).

P. A. Blanco et al, "Detecting the ultra low dimensionality of real networks", *Nat. Comm.* (2022).

Why Geometric Methods Dominate NetSci/ML/DataSci?

<u>Similarity Principle</u>: *Birds of a Feather Flock Together*

<u>In Network Science</u>: *Connections are Likely Between Similar Nodes*

Why?

Network science has been fueled by *social sciences* in early 2000s because social network data is relatively easy to connect

Why Geometric Methods Dominate NetSci/ML/DataSci?

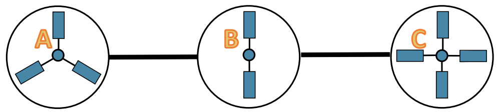
<u>Similarity Principle</u>: *Birds of a Feather Flock Together*

<u>In Network Science</u>: *Connections are Likely Between Similar Nodes*

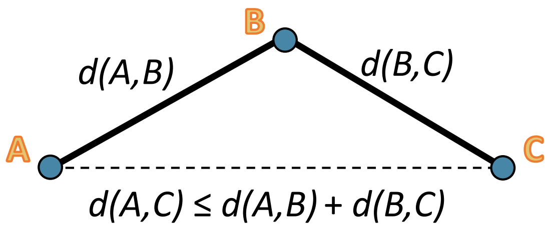
Why? Similarity is very *intuitive* owing to its **transitivity**:

A similar to B, B similar to C -> A similar to C

a



So is our 3D(+1) *Euclidean* world: Euclidean space is metric: d(A,C) ≤ d(A,B) + d(B,C)



It easy to think of networks using physics intuition:

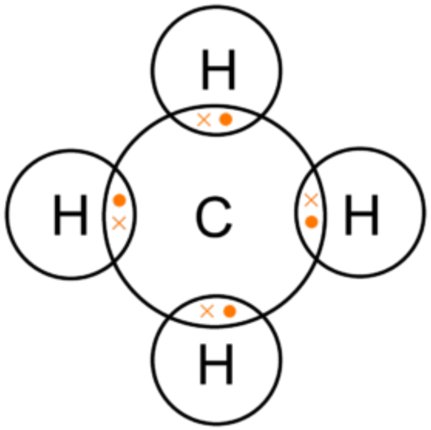
Nodes are points, links appear between close nodes.

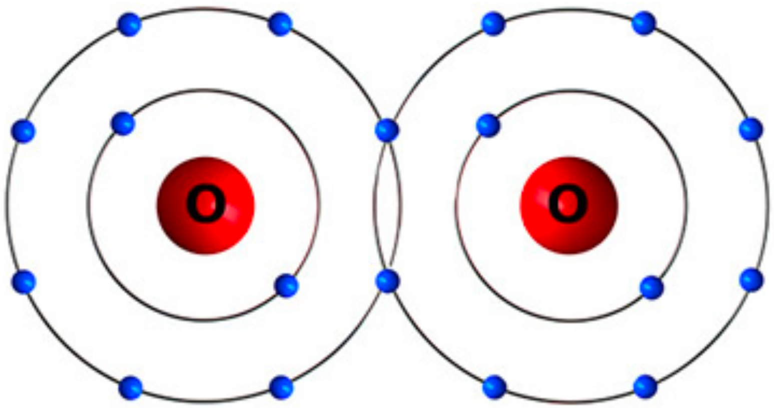
<u>Paths</u> are nothing else but trajectories.

<u>Communities</u> are dense regions of space where a lot of nodes are.

A lot of Networks that are Complementarity-Based

Example 1 (Chemistry 101): Covalent Bonds: sharing electrons between atoms.





Proteins perform a vast array of functions within organisms

myoglobin (source: Wikipedia)

A. J. McCoy, V. Chandana Epa, and P. M. Colman, *Electrostatic complementarity at protein/protein interfaces 1 1Edited by B. Honig,* J. Mol. Biol. **268**, 570 (1997).

Q. Zhang, M. Sanner, and A. J. Olson, *Shape complementarity of protein-protein complexes at multiple resolutions,* Proteins Struct. Funct. Bioinforma. **75**, 453 (2009).

Y. Li, X. Zhang, and D. Cao, *The Role of Shape Complementarity in the Protein-Protein Interactions,* Sci. Rep. **3**, 3271 (2013).

A lot of Networks that are Complementarity-Based

Example 2: Interdisciplinary collaborations:

Scientists with complementary skills can benefit in collaborations:

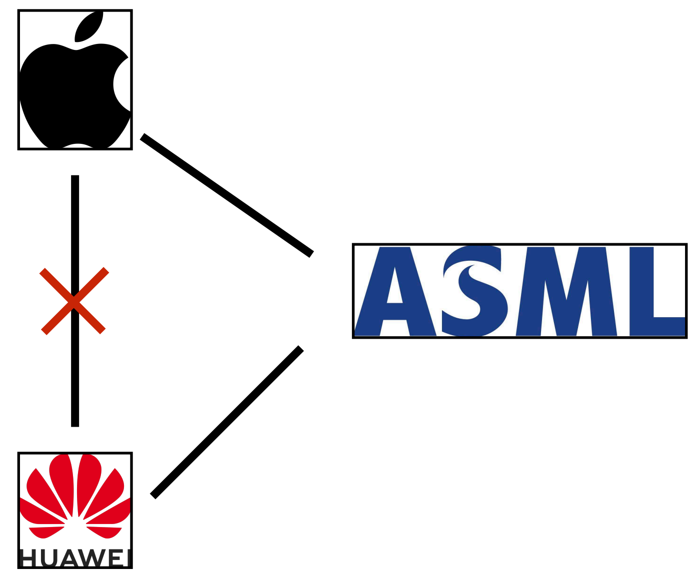
BioInformatics

Quantum Computing

Medical Physics

Political Data Science

Example 3: Business (trading) networks



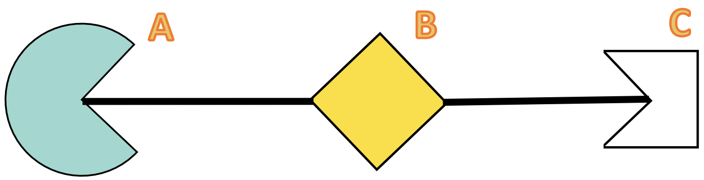
Example 4: Text

Words appear in text to complement each other towards certain message.

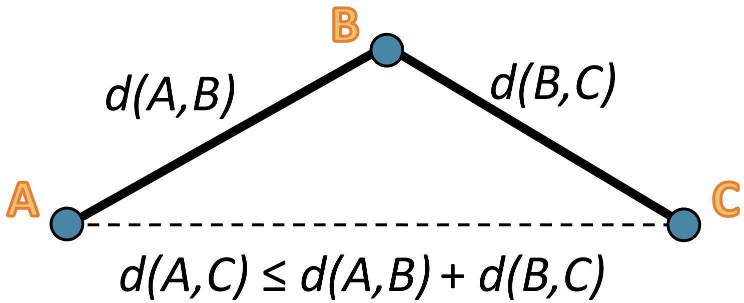
He was an old man who fished alone in a skiff in the Gulf Stream and he had gone eighty-four days now without taking a fish.

Off-the-Shelf Geometric Methods Fail for Complementarity Networks!

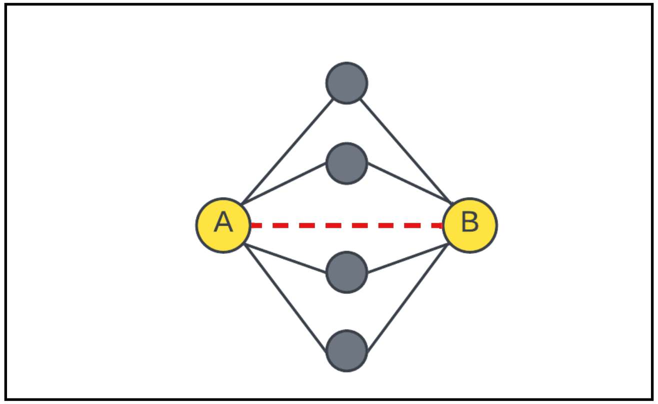
Complementarity is not transitive! A compl B, B compl C, **do not imply** A compl C



Applying similarity philosophy leads to inference errors!



Common neighbors do not imply links!



I. Kovacs et al, "Network-based prediction of protein interactions", *Nat. Comm.* (2023).

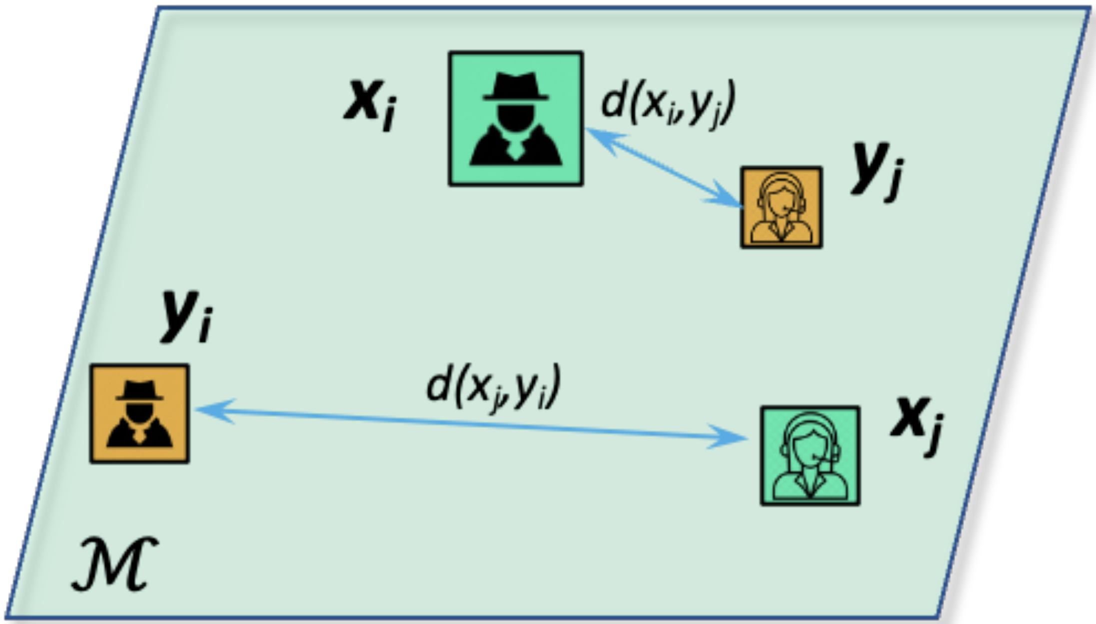
Can we rescue network embedding methods?

Quick Solution:

Nodes Represented by Several Points in $\mathbb{M}$.

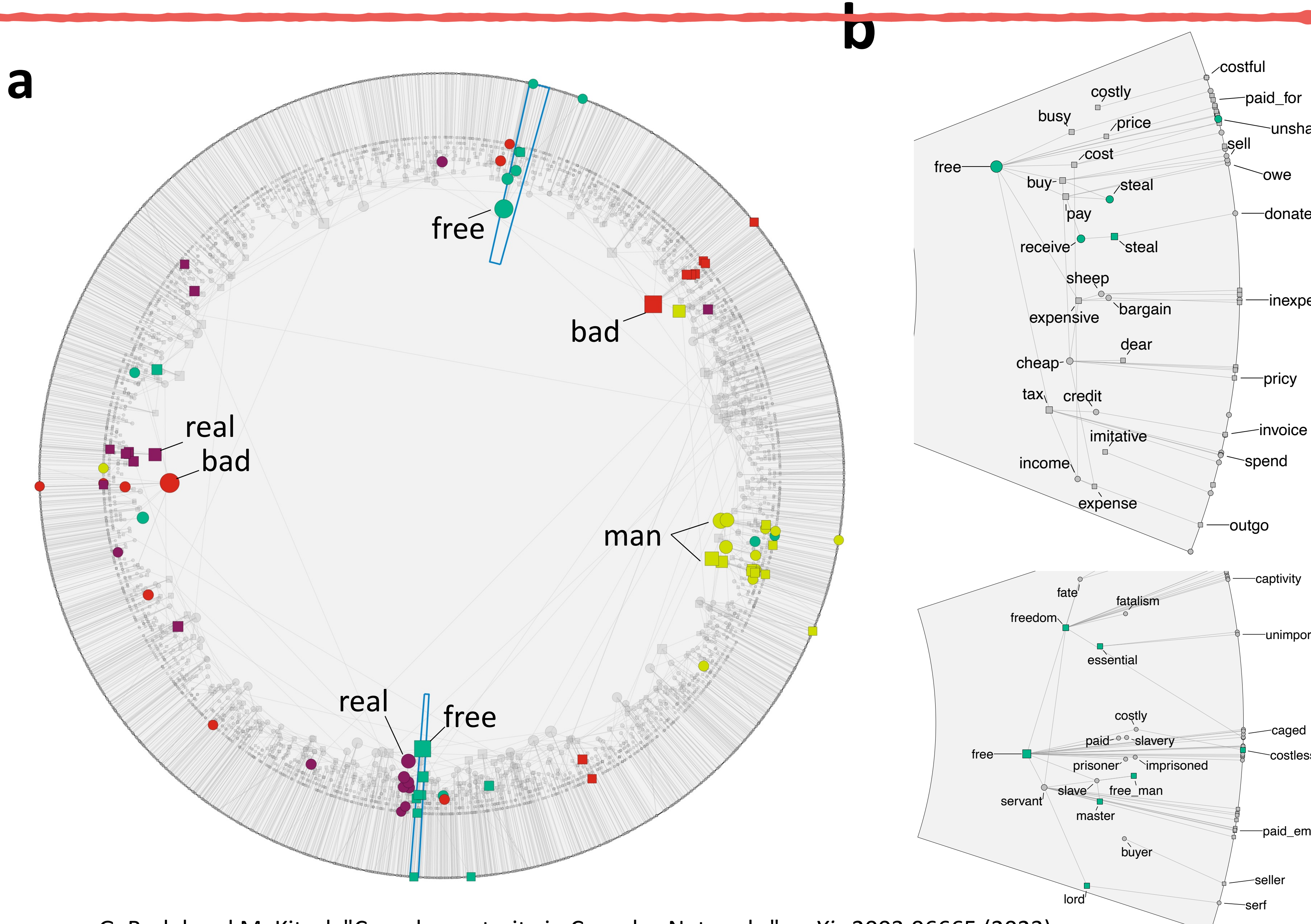
Each point can be viewed as a distinct property (or skill, or interface).

Distances between points of different type quantify complementarity.



G. Budel and M. Kitsak "Complementarity in Complex Networks", *arXiv* 2003.06665 (2023).

Complementarity Representation of the Antonym Network



G. Budel and M. Kitsak "Complementarity in Complex Networks", *arXiv* 2003.06665 (2023).

Can we rescue network embedding methods?

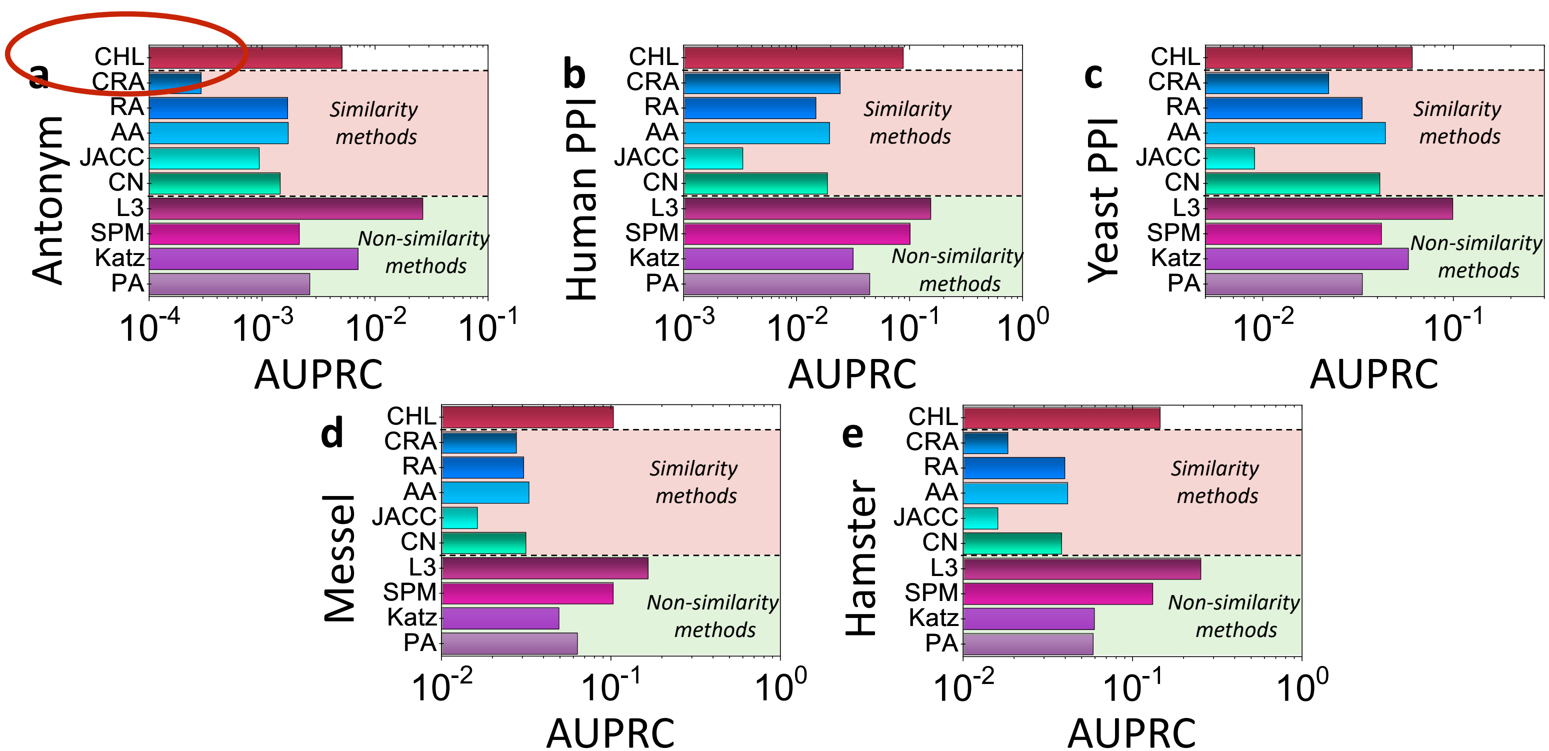
Quick Solution:

Nodes Represented by Several Points in $\mathbb{M}$.

Each point can be viewed as a distinct property (or skill, or interface).

Distances between points of different type quantify complementarity.

Prediction of missing links in five complementarity-based networks



G. Budel and M. Kitsak, "Complementarity in Complex Networks", *arXiv* 2003.06665 (2023).

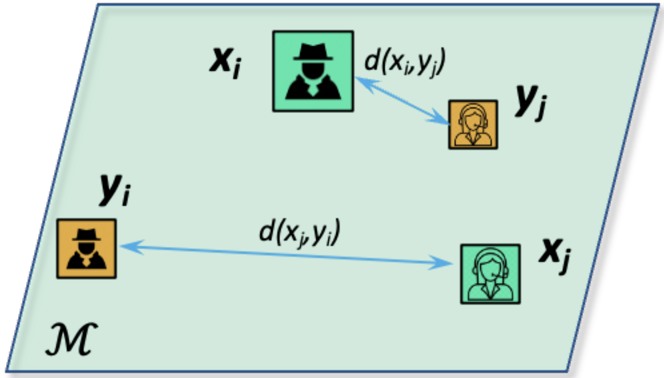
Can we rescue network embedding methods?

Quick Solution:

Nodes Represented by Several Points in $\mathbb{M}$.

Each point can be viewed as a distinct property (or skill, or interface).

Distances between points of different type quantify complementarity.



Open Questions:

What if there are more skills per node?

Why different skills belong to the *<u>same</u> space?*

How can we even compare different skills?

Why distance is a measure of complementarity?

G. Budel and M. Kitsak "Complementarity in Complex Networks", *arXiv* 2003.06665 (2023).

Towards the Principled Complementarity Framework

"Two people or things that are complementary are different but together form a useful or attractive combination of skills, qualities or physical features.”

-Oxford Dictionary

G. Budel and M. Kitsak "Complementarity in Complex Networks", *arXiv* 2003.06665 (2023).

Towards the Principled Complementarity Framework

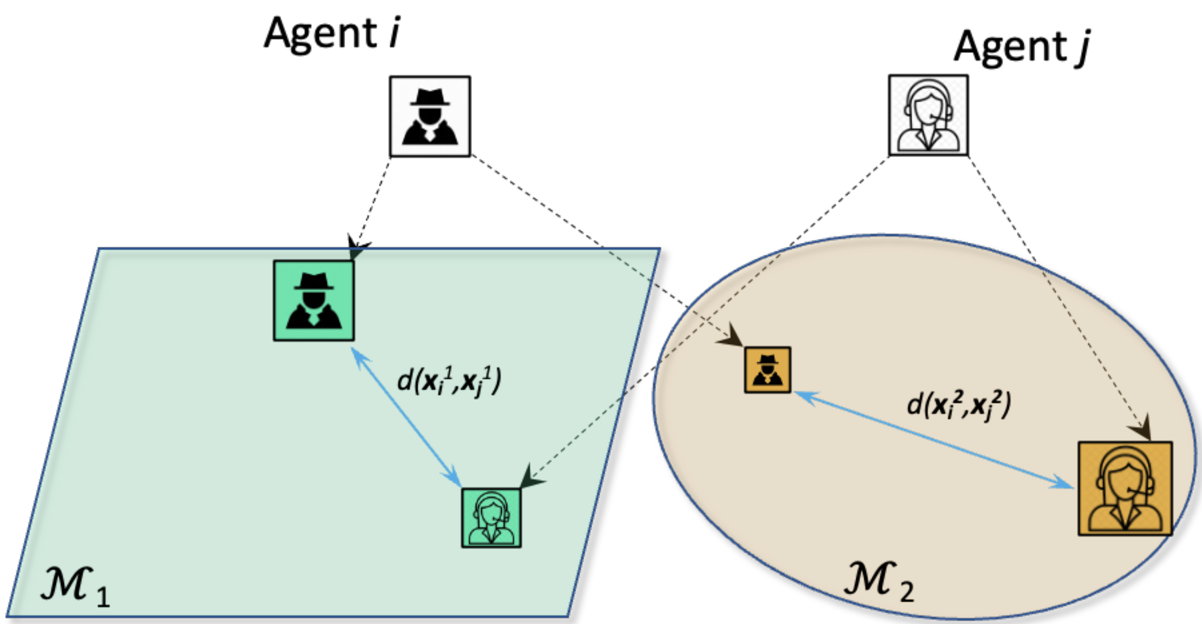
NWO enw-M1 proposal, M.Kitsak (PI), 2022

"Two people or things that are complementary are different but together form a useful or attractive **combination of skills**, qualities or physical features."

-Oxford Dictionary

There must be at least two different skills at play!

How do we describe skills? Points in a metric space. We need at least two metric spaces! Every node has a corresponding point in each space!



Distances quantify similarities. Distances are defined within each space, **not between spaces!**

Towards the Principled Complementarity Framework

NWO enw-M1 proposal, M.Kitsak (PI), 2022

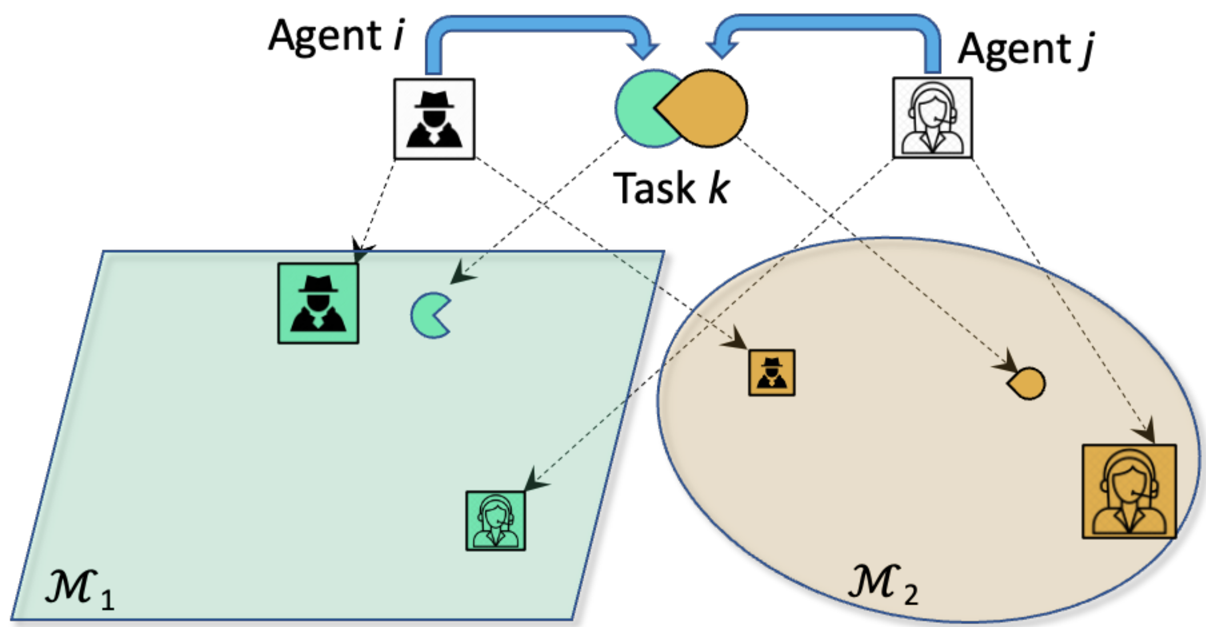
"Two people or things that are complementary are different but together **form a useful or attractive combination** of skills, qualities or physical features."

-Oxford Dictionary

One way to quantify "usefulness" is by introducing tasks!

Agents (nodes) complement each other in executing joint tasks.

Every task has a corresponding point in each space.



Towards the Principled Complementarity Framework

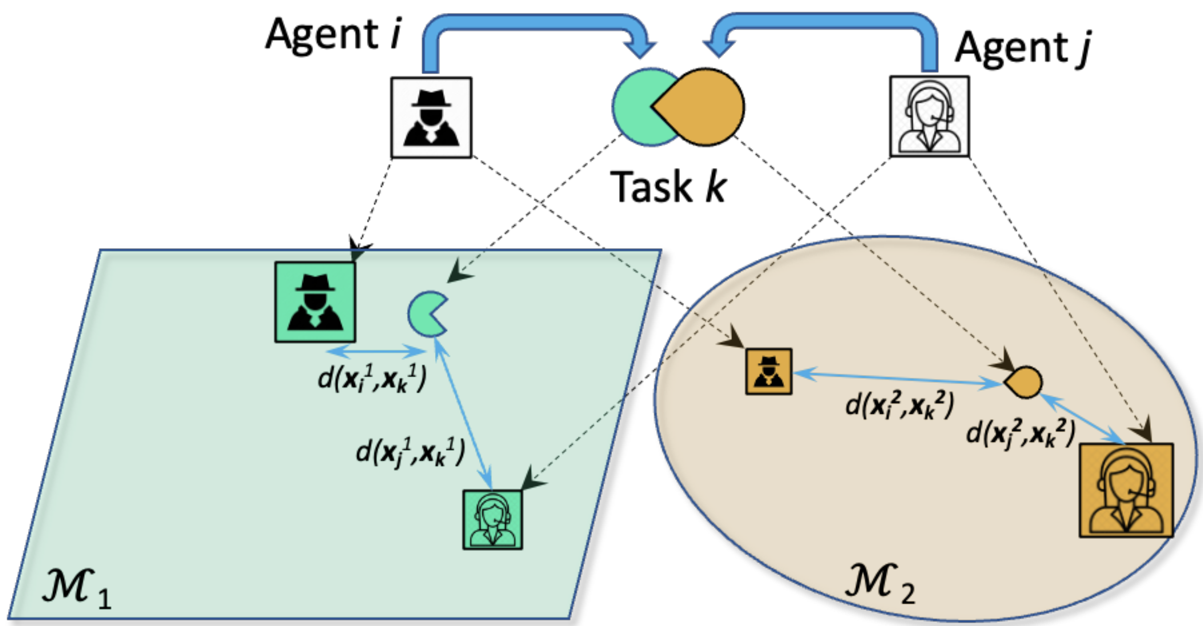
NWO enw-M1 proposal, M.Kitsak (PI), 2022

"Two people or things that are complementary are different but together **form a useful or attractive combination** of skills, qualities or physical features."

-Oxford Dictionary

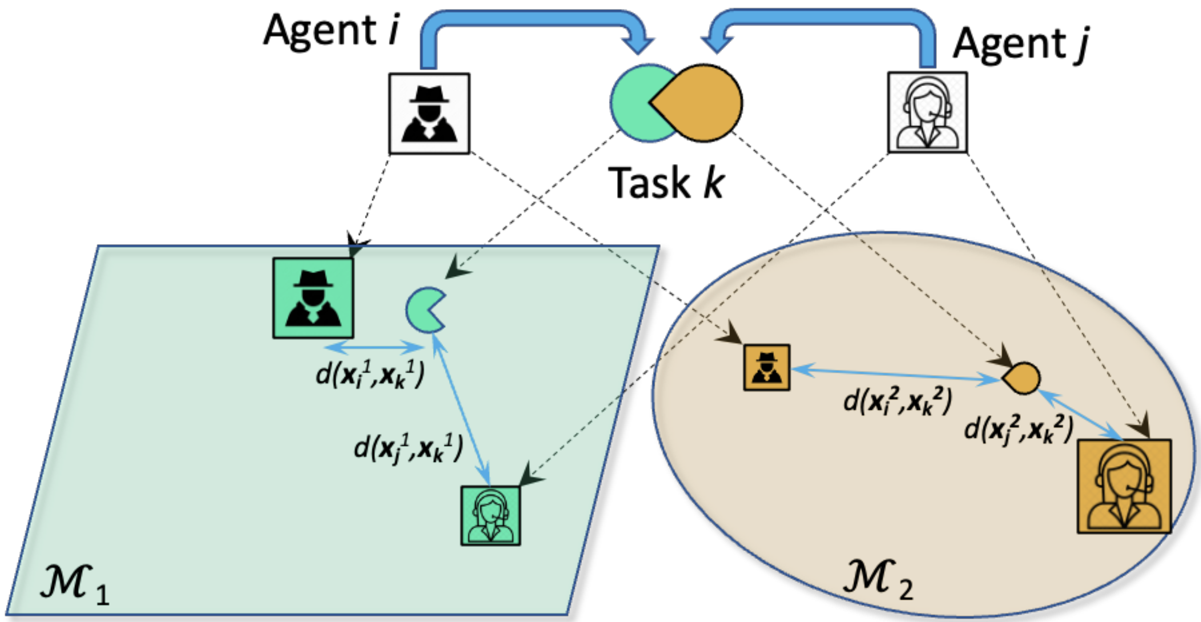
Complementarity is the ability to co-execute tasks:

each agent *independently* solves their part of the task



The closer you are to the given part of the task, the more likely you are to solve it.

Towards the Principled Complementarity Framework



Probability that i solves part 1 and j solves part 2: $r_1\left(d\left[\mathbf{x}_i^1, \mathbf{x}_k^1\right]\right) r_2\left(d\left[\mathbf{x}_j^2, \mathbf{x}_k^2\right]\right)$

$r_{\{1,2\}}(d)$ are decreasing functions of distances d

Probability that j solves part 1 and i solves part 2: $r_1\left(d\left[\mathbf{x}_j^1, \mathbf{x}_k^1\right]\right) r_2\left(d\left[\mathbf{x}_i^2, \mathbf{x}_k^2\right]\right)$

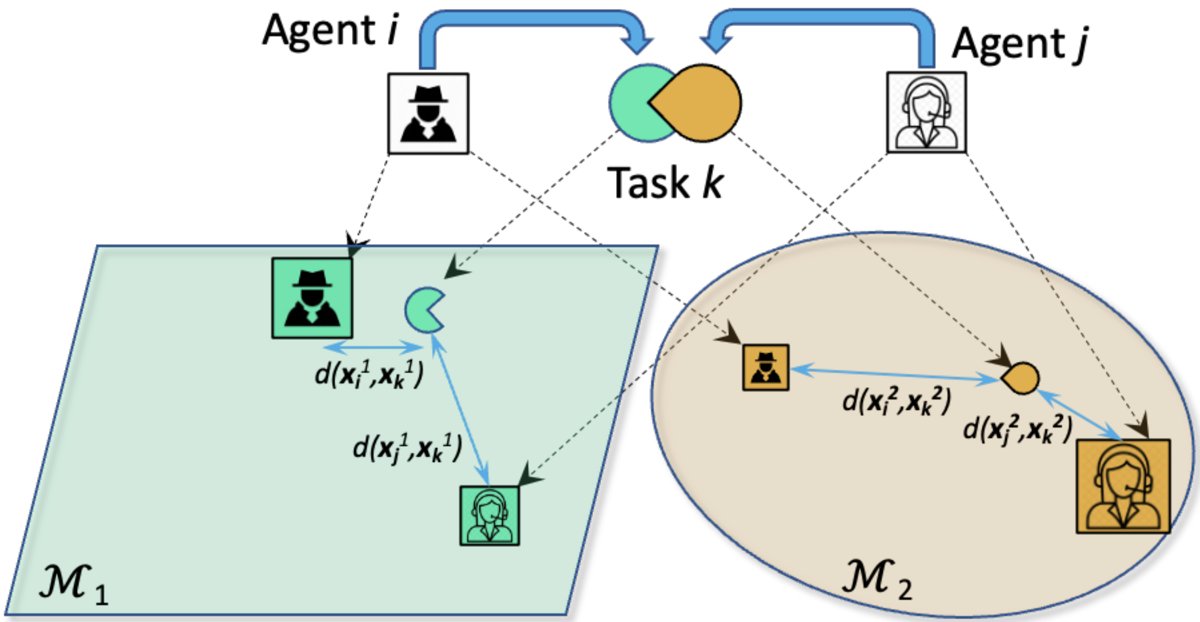
Probability i and j solve task is the union of two probabilities:

$$p(i,j|k) = 1 - \left[1 - r_1\left(d\left[\mathbf{x}_i^1, \mathbf{x}_k^1\right]\right) r_2\left(d\left[\mathbf{x}_j^2, \mathbf{x}_k^2\right]\right)\right]\left[1 - r_1\left(d\left[\mathbf{x}_j^1, \mathbf{x}_k^1\right]\right) r_2\left(d\left[\mathbf{x}_i^2, \mathbf{x}_k^2\right]\right)\right]$$

Complementarity = Pr-ty *i* and *j* solve at least one task *k:*

$$p_{ij} = 1 - \prod_k (1 - p(i,j|k))$$

Towards the Principled Complementarity Framework



Highlights:

The principled complementarity framework contains the minimal framework as a special case.

Complementarity contains similarity as a special case.

(When two points corresponding to a node are the same.)

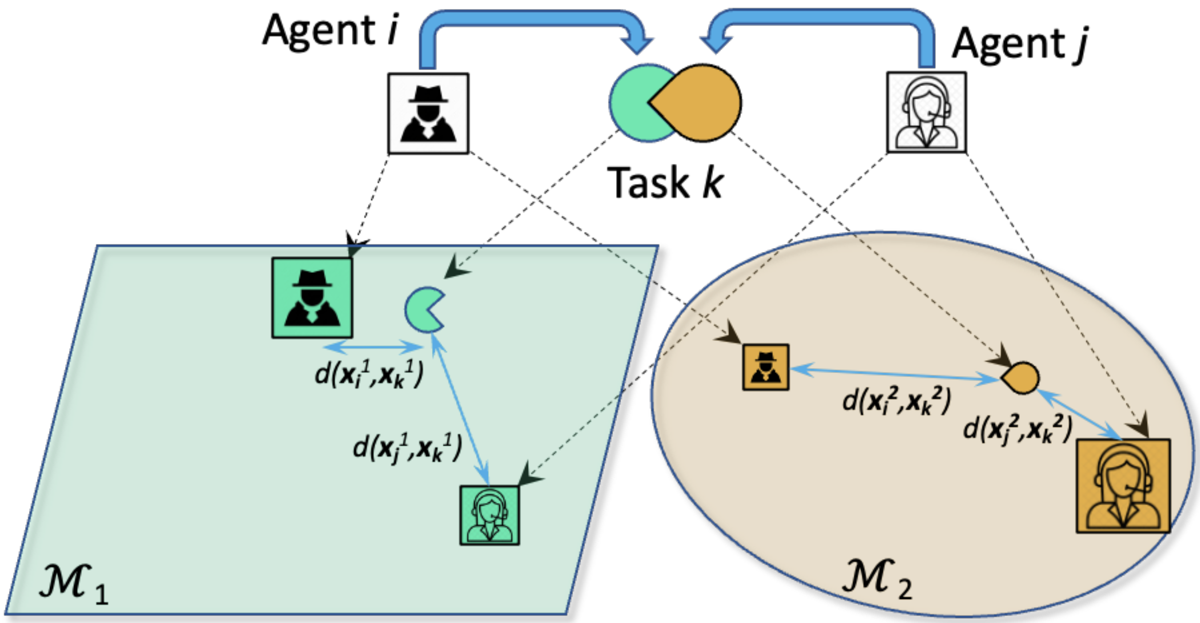
Open Questions:

Is the Principled Complementarity Framework Learnable in its General Form?

Simplified Versions of the Principled Complementarity Framework for Efficient Learning?

G. Budel and M. Kitsak "Complementarity in Complex Networks", *arXiv* 2003.06665 (2023).

Towards the Principled Complementarity Framework



Highlights:

The principled complementarity framework contains the minimal framework as a special case.

Complementarity contains similarity as a special case.

(When two points corresponding to a node are the same.)

Open Questions:

Is the Principled Complementarity Framework Learnable in its General Form?

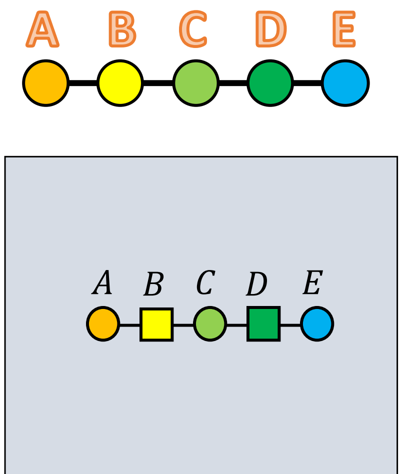
Simplified Versions of the Principled Complementarity Framework for Efficient Learning?

G. Budel and M. Kitsak "Complementarity in Complex Networks", *arXiv* 2003.06665 (2023).

Back to High Level: Implications for Shortest Paths

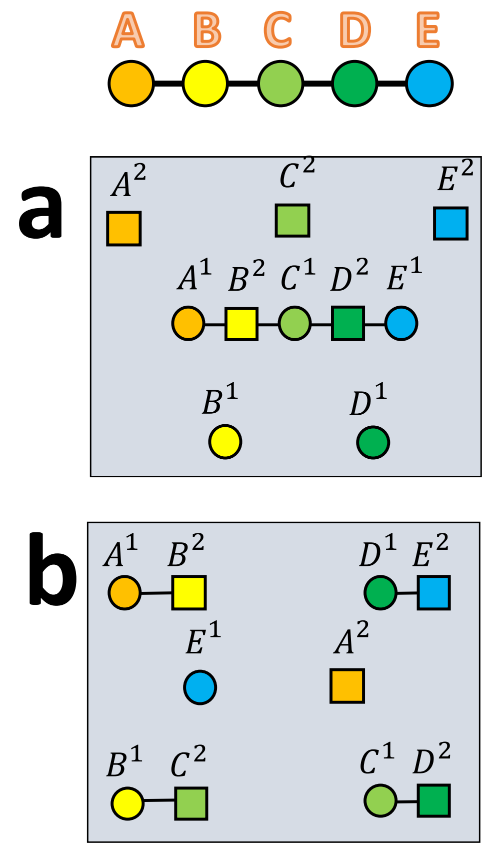
Similarity-based intuition:
shortest path is a ***trajectory***.

By traveling along a path you hop from one node to another, and neighboring nodes are relatively close to each other.



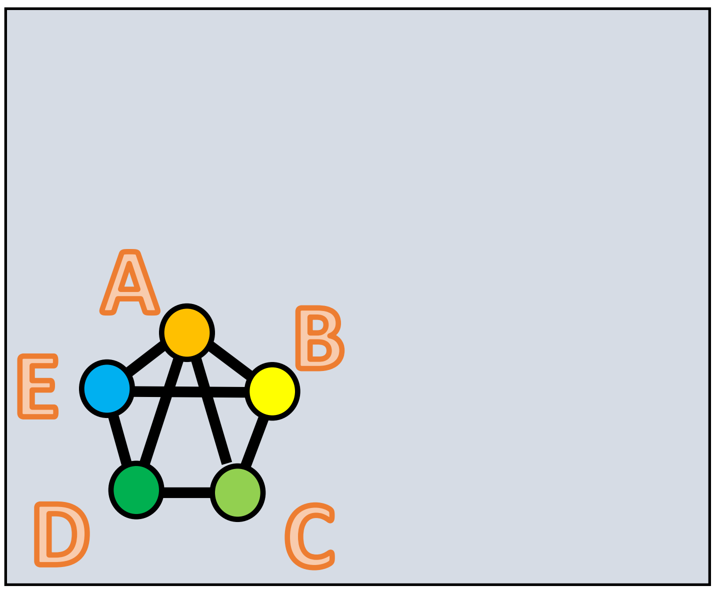
Complementarity-based intuition:
shortest path ***is not a trajectory***.

You may have a trajectory, it is much more likely that the trajectory is broken.



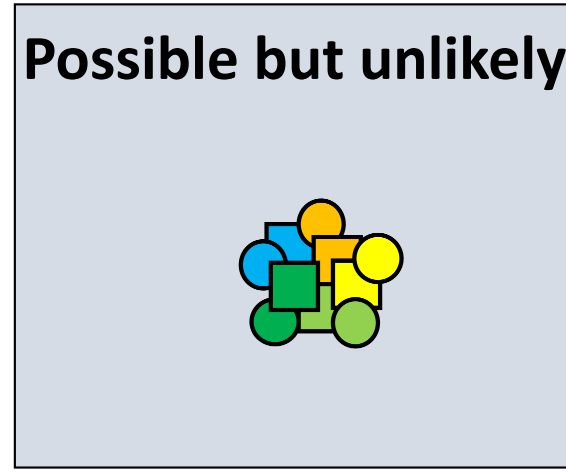
Back to High Level: Implications for Communities/Cliques/Modules

Similarity-based intuition: community is a collection of nodes that is **localized** in a region of space.

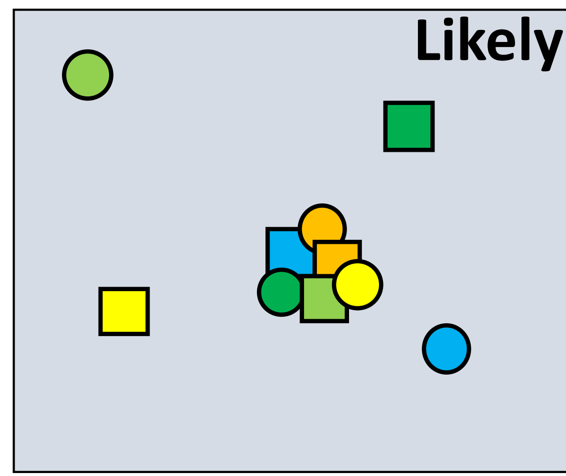


Complementarity-based intuition: communities are **rarely localized**.

a



b



Implications for (Representations) of Real Networks

Biomedicine:

Missing PPI interactions.

(Better) functional classification of proteins.

Uncovering relationships between cellular pathways.

Interpreting existing drug-target relationships and identifying novel drug targets.

……….

Science of Science:

Success in interdisciplinary collaborations.

Scientific trajectories.

Relations between subfields of related disciplines.

Uncovering relationships between scientific problems.

……….

Your feedback, and intuition, and collaboration is very welcome!

G. Budel and M. Kitsak "Complementarity in Complex Networks", *arXiv* 2003.06665 (2023).

Take-Home Message

1.Geometric representations of complementarity-driven nets possible.
But one needs more than 1 point per node.

2. A minimal *null* model with 2 points per node shows significant improvement at predicting missing links in biological and ecological networks.

3. Important implications for *Biomedicine* and *Science of Science.*

Our intuition in network science comes almost exclusively from social networks, which are governed by the similarity rule.

We need to rethink/adjust our approaches for complementarity-based networks: biological, collaboration, ecological etc.

Interpretable Complementarity Framework is coming soon!
Similarity is the <u>*special case*</u> of Complementarity!

G. Budel and M. Kitsak "Complementarity in Complex Networks", *arXiv* 2003.06665 (2023).