Market Connectivity

Anup Rao
Paul G. Allen School of Computer Science
University of Washington

Who am I?

- Not an economist
- A theoretical computer scientist (essentially ~ mathematician)
- I study *communication complexity.*
- A dad

Mika — 7 months

Market Connectivity

- New useful statistics for economies
- Paper - March 2017 (available on arxiv and my website)
- With Alan Griffith:
  analyze data about Uganda

High Level Objectives

- Show how to meaningfully mine data about transactions in economy.

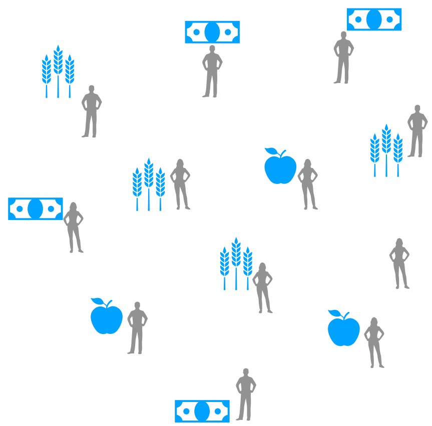
- Give **you** some tools to use on data
  and/or obtain data **from you** to analyze.
  All code (matlab) available for download on my website:
  https://homes.cs.washington.edu/~anuprao/

- Discover new ways to use these ideas

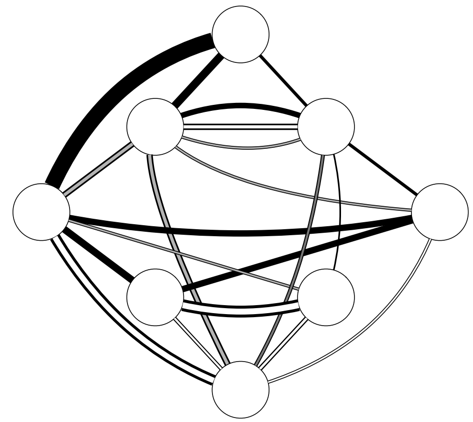
User’s guide

- *Please*:
  Interrupt!
  Ask questions!
  Make clarifying comments!
  Help me understand what is confusing!

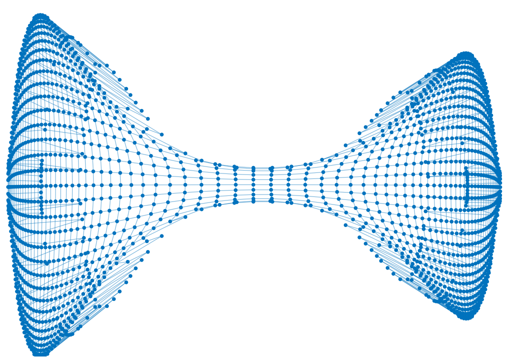
Outline



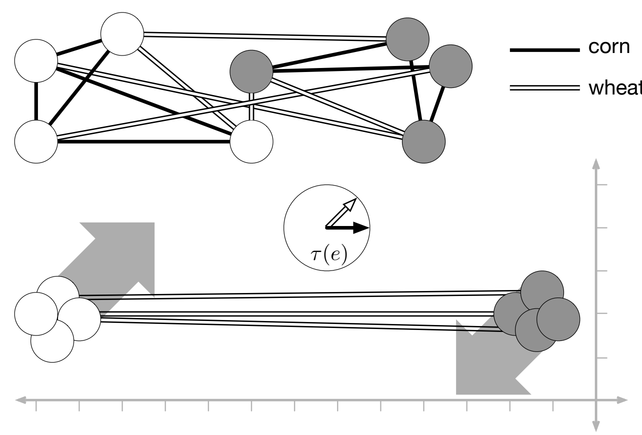
1. CS and Economics



2. Trade Networks



3. Graph Laplacians

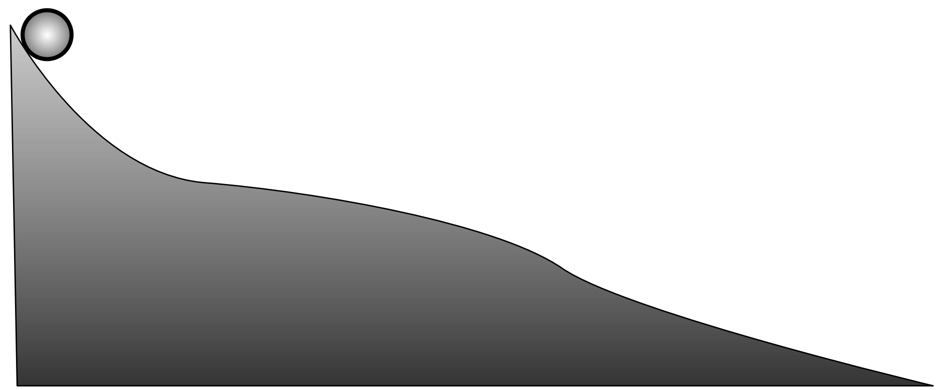


4. Modifications for Econ

A new theoretical concept

- … as in physics—concept to allow math to explain object of study
- Goal: useful, measurable statistics about economy.
- No assumptions made about underlying economy

Newtonian mechanics (1697)



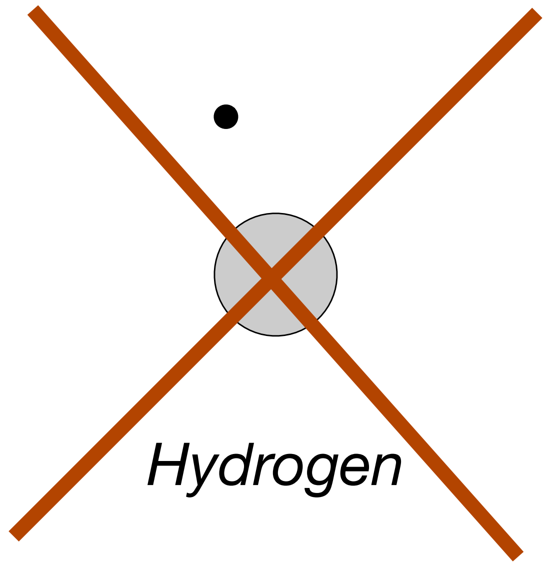
How long before boulder reaches bottom?

Focus
Position, velocity, mass, forces

Goal: new statistics for Econ

Ignored
Friction, air resistance,
exact shape/density of the boulder

Photoelectric effect (1902)



Focus
~~*Position, velocity*~~, mass, forces

Notes:
1. Newton still made progress!
2. Relevance of concept established using real-world data.
3. Data gives circumstantial evidence, at best.

CS model: Distributed System

Example: internet

Processors: computers on the internet

Inputs: information—user requests,
webpages, airline schedules…

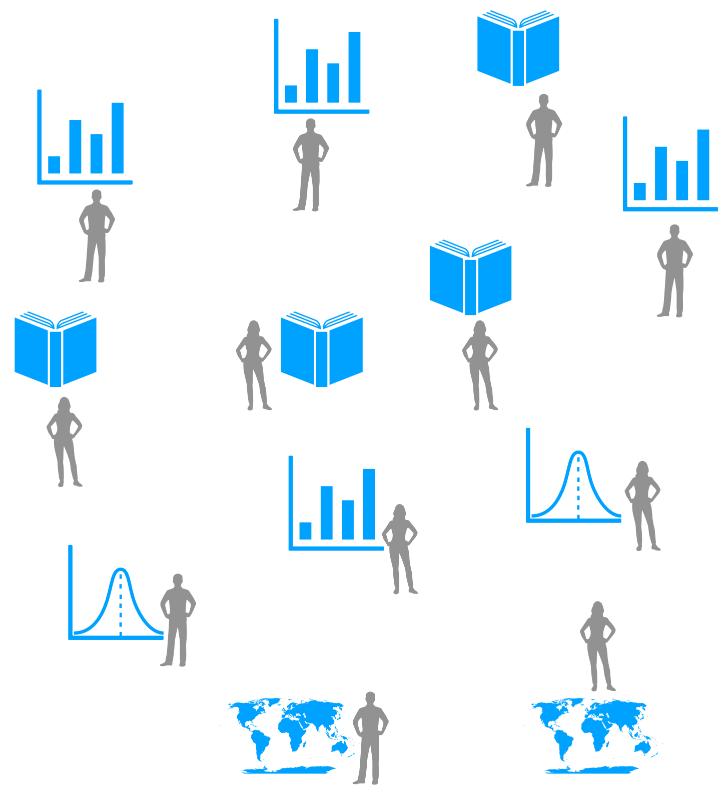
CS model:
Distributed System

Example: internet

Processors: computers on the internet

Inputs: information—user requests,
webpages, airline schedules…

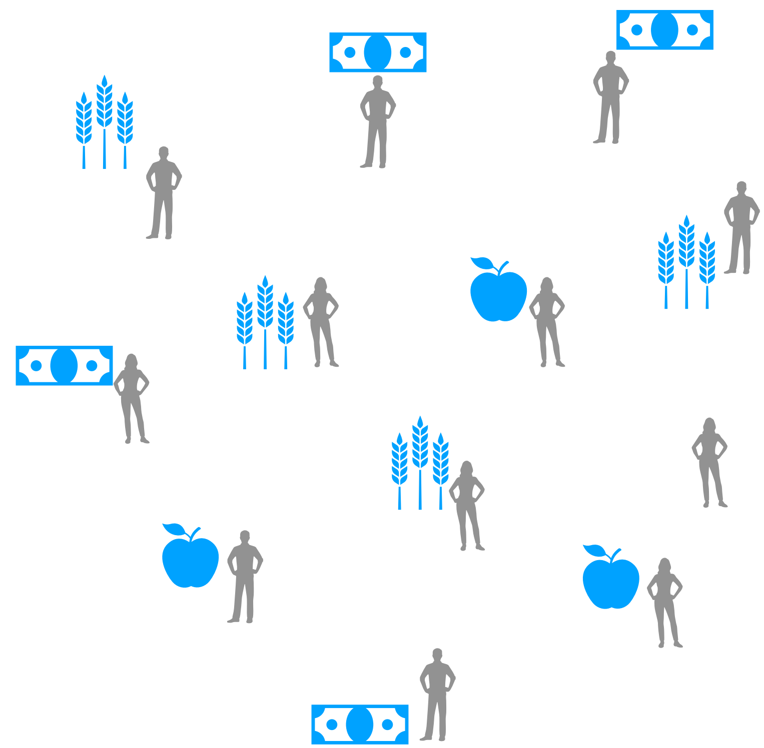
Output: relevant information



How to make this fast?
How to make this reliable?
How to connect computers?
How to direct traffic?
…

New perspective
Economy = Distributed System

Processors: agents
Inputs: allocation of resources

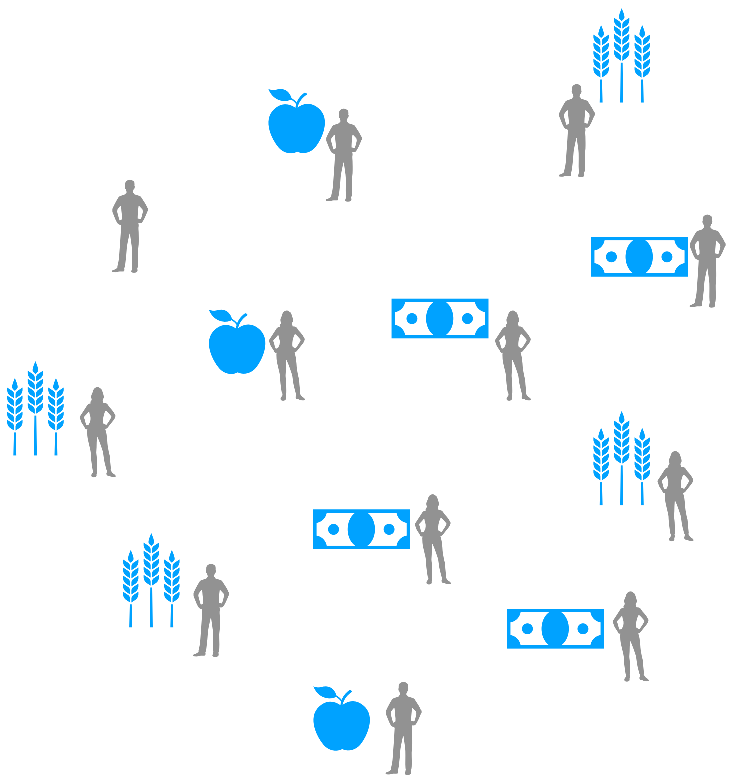


New perspective
Economy = Distributed System

Processors: agents
Inputs: allocation of resources
Output: new allocation of resources

This view opens up economics to CS style reasoning;
computations have associated costs

eg: How to measure *efficiency of communication*?



New perspective
Economy = Distributed System

- Analogy more accurate than one might think!

- **Internet**
  messy, organic, decentralized

- Challenges for Distributed Systems:
  Unreliable computers
  Malicious participants
  Inability to program all computers in the same way
  Enabling fast communication with limited connectivity
  …

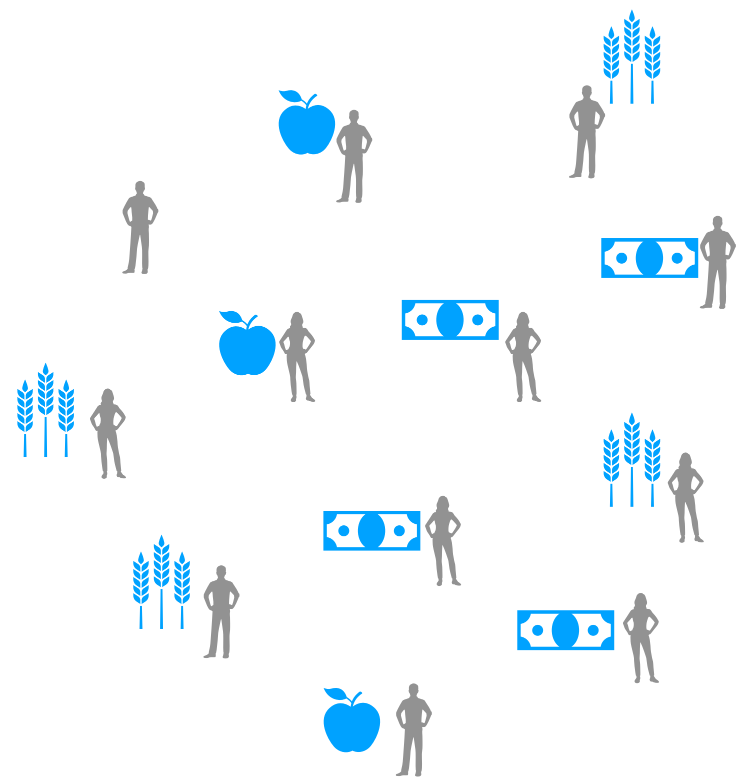
What is novel?

- price of milk @U-district ~ price of milk @cap-hill

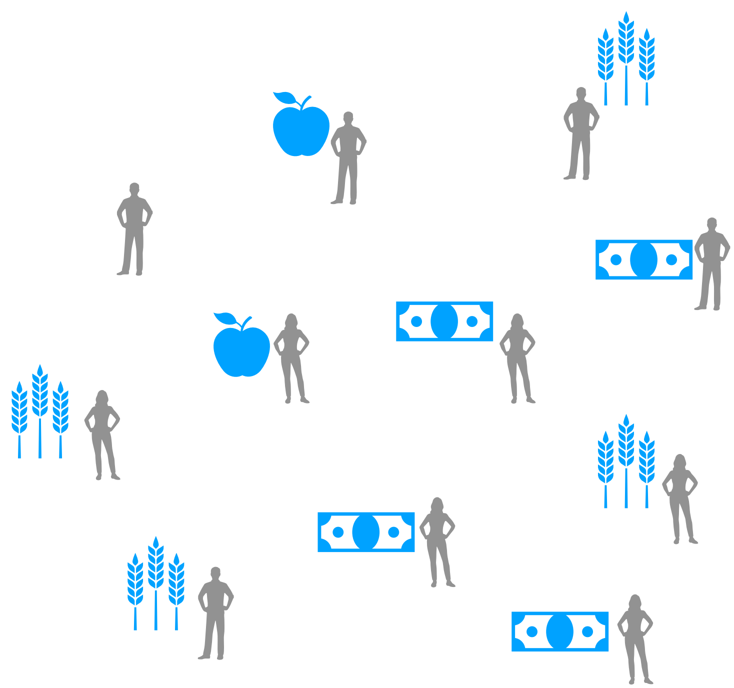
- From a CS perspective:
  this is a *computation,* so evaluate
  speed/accuracy

- A *trace* of computation is visible!

- Benefits to economics:
  Measure impact of policy
  design better systems/regulations



Focus of my work

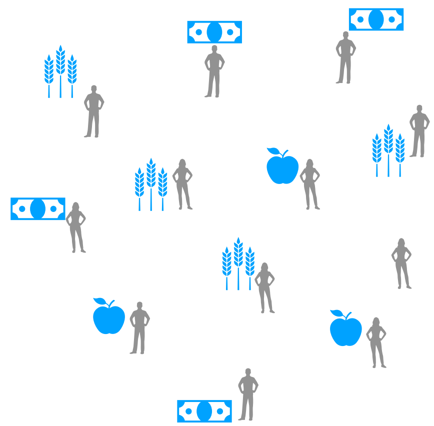


How well does information flow?

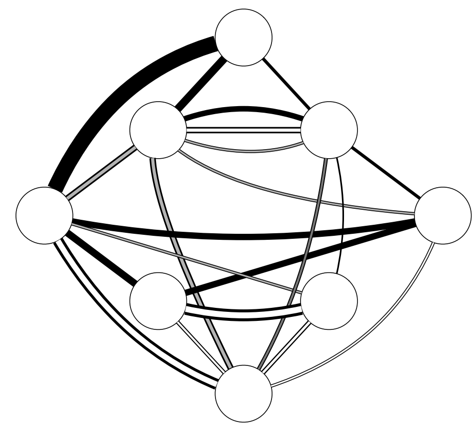
Focus
Statistics about
network of transactions

Ignored
Behavior of agents, rationality of agents, transaction costs…

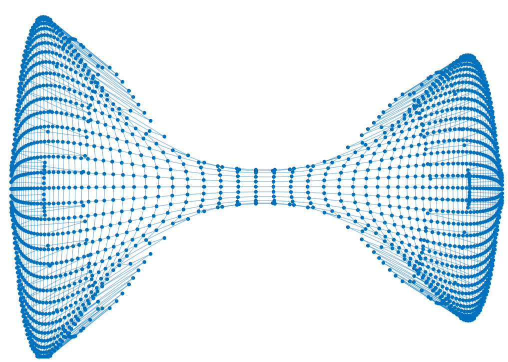
Outline



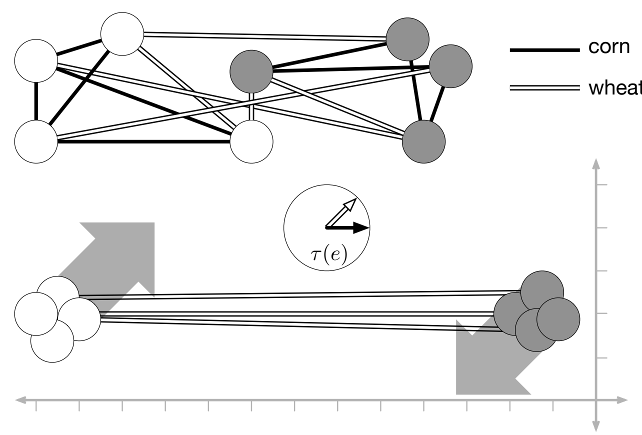
1. CS and Economics



2. Trade Networks



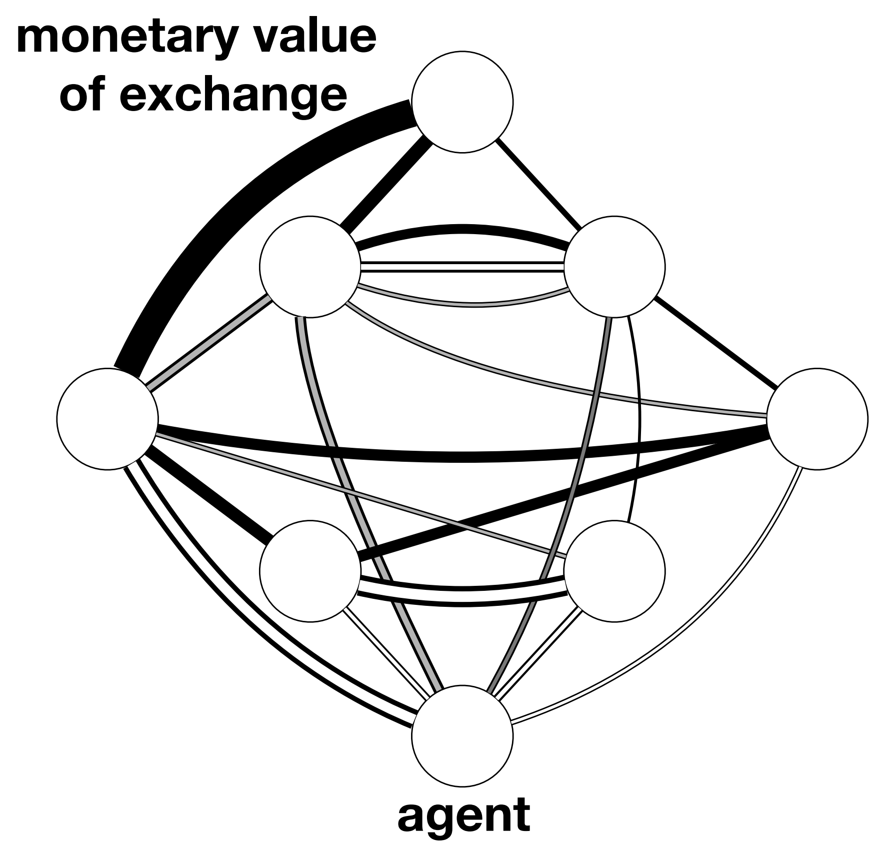
3. Graph Laplacians

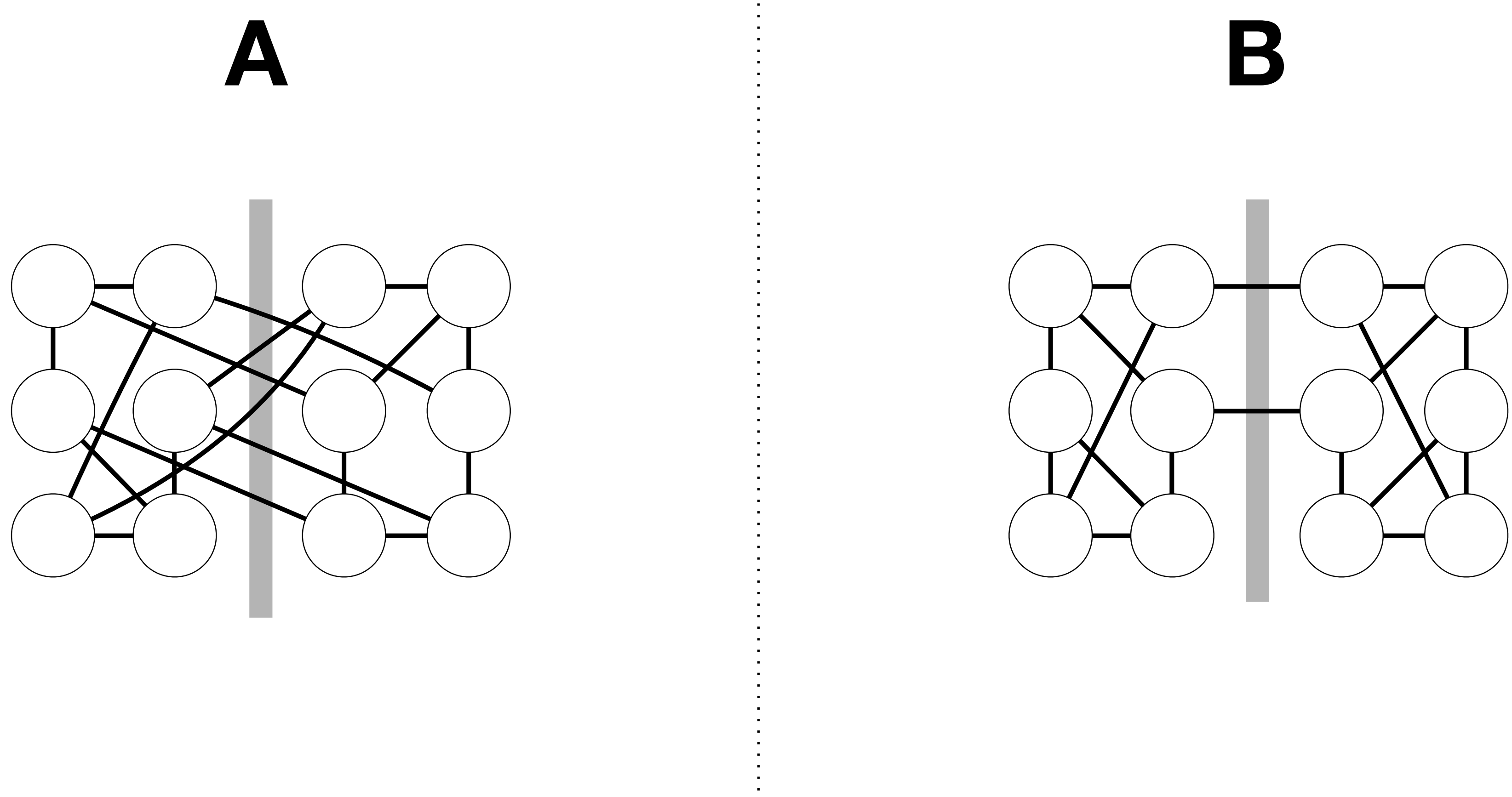


4. Modifications for Econ

Trade Network

- Fixed time interval
  say [Jan 2017- Jan 2018]
- **Nodes** = agents
- **Edges/links** = transactions
  (weights = monetary value)
- Can be **measured!**
  (data sets available)
- Encodes many useful statistics
  about economy
- GDP = sum of edge weights





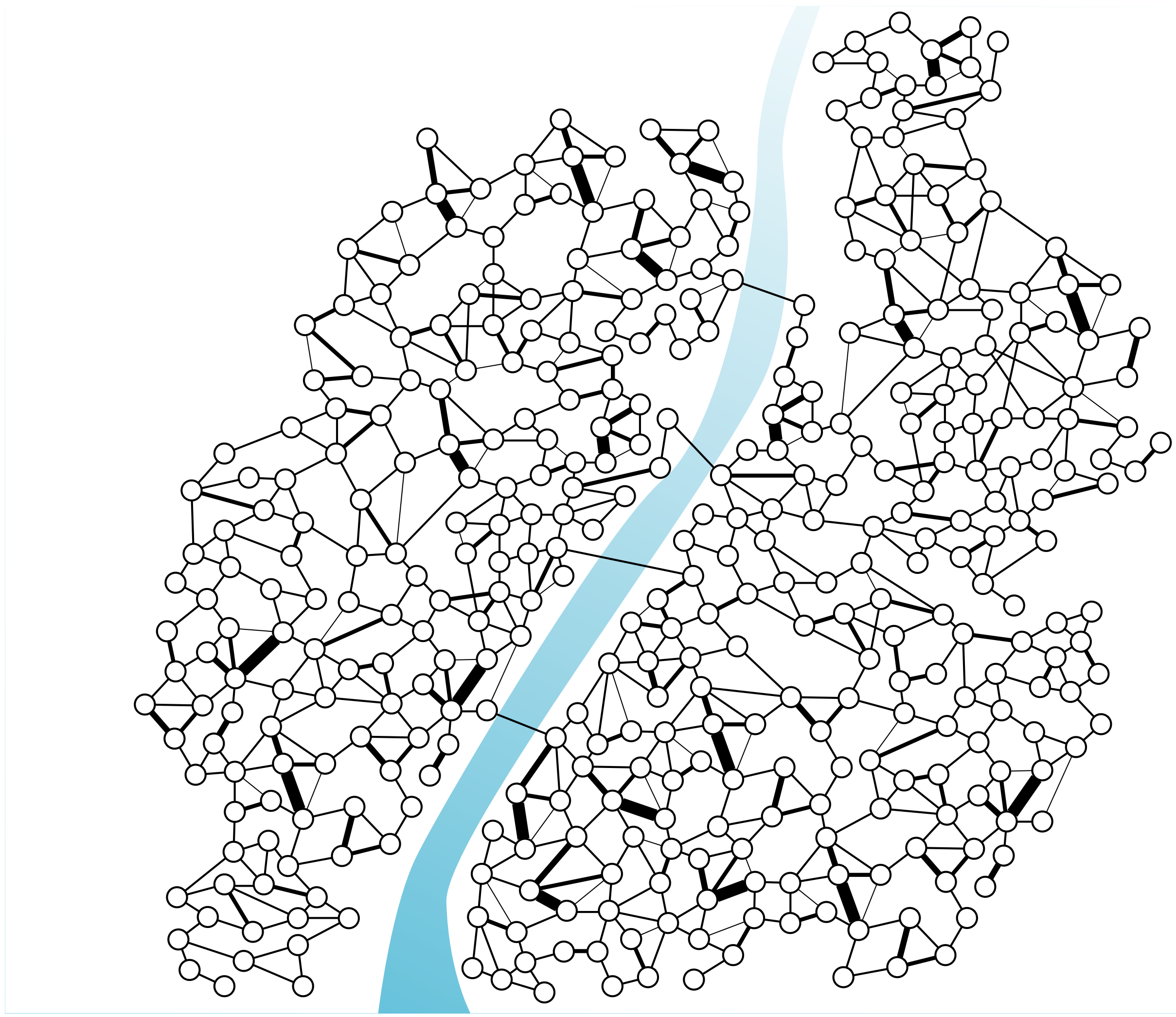
Total **number of links is equal,**
but A is better *connected***:**

1. Shorter paths between any two nodes
2. Any balanced partition of nodes *cuts* more links

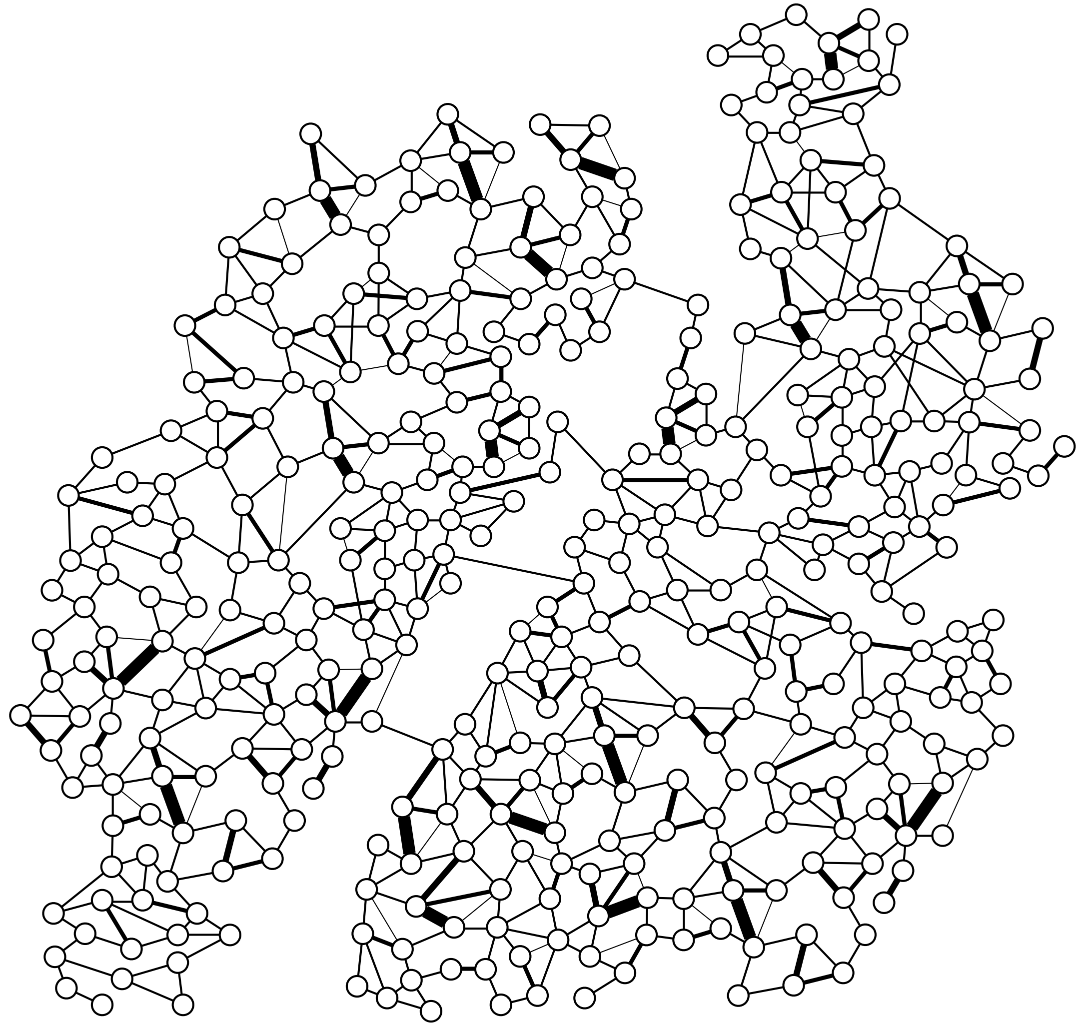
In general: network encodes *degree of connectivity*, not just volume of trade!

2000 BC

2000 BC

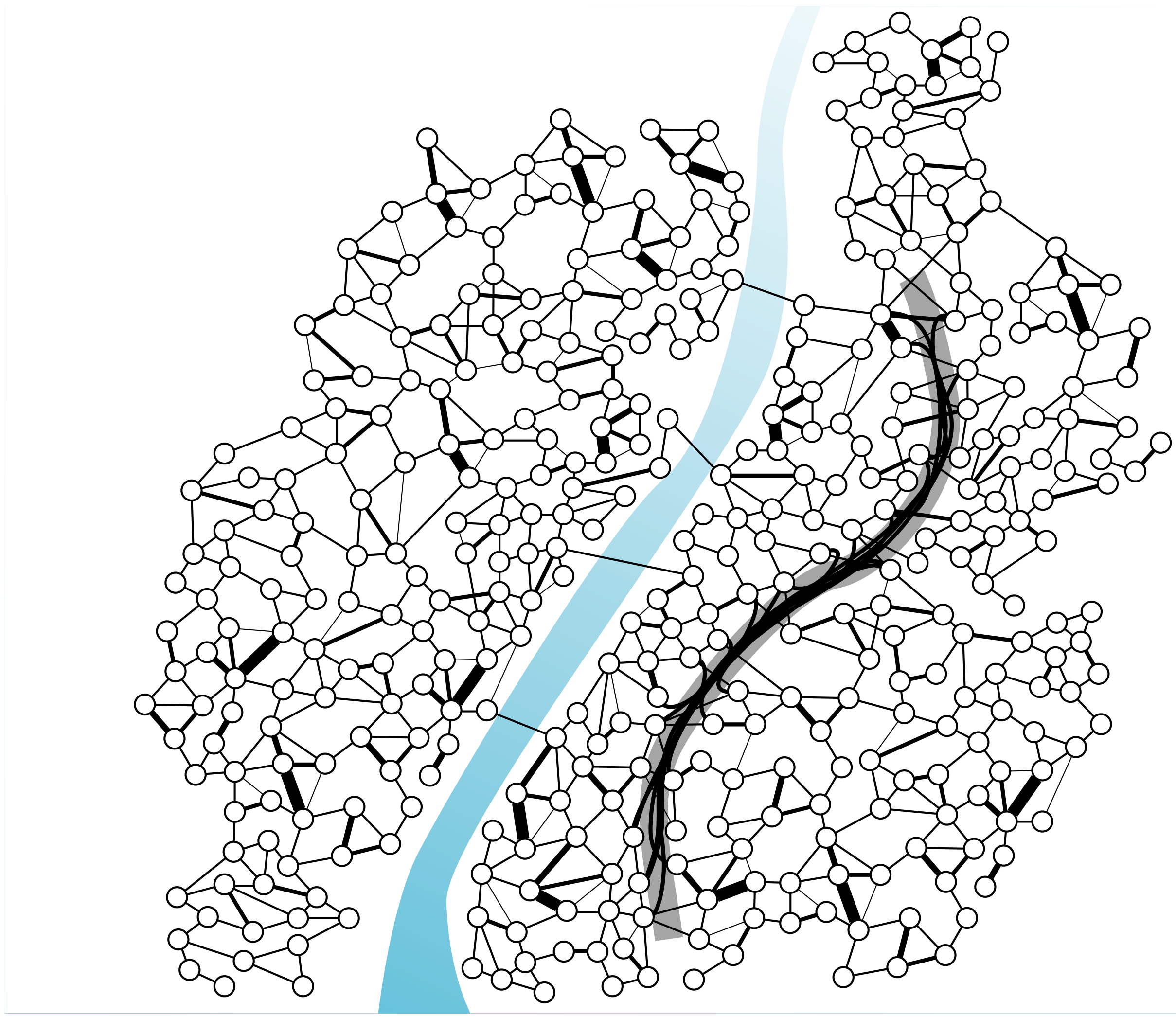


2000 BC

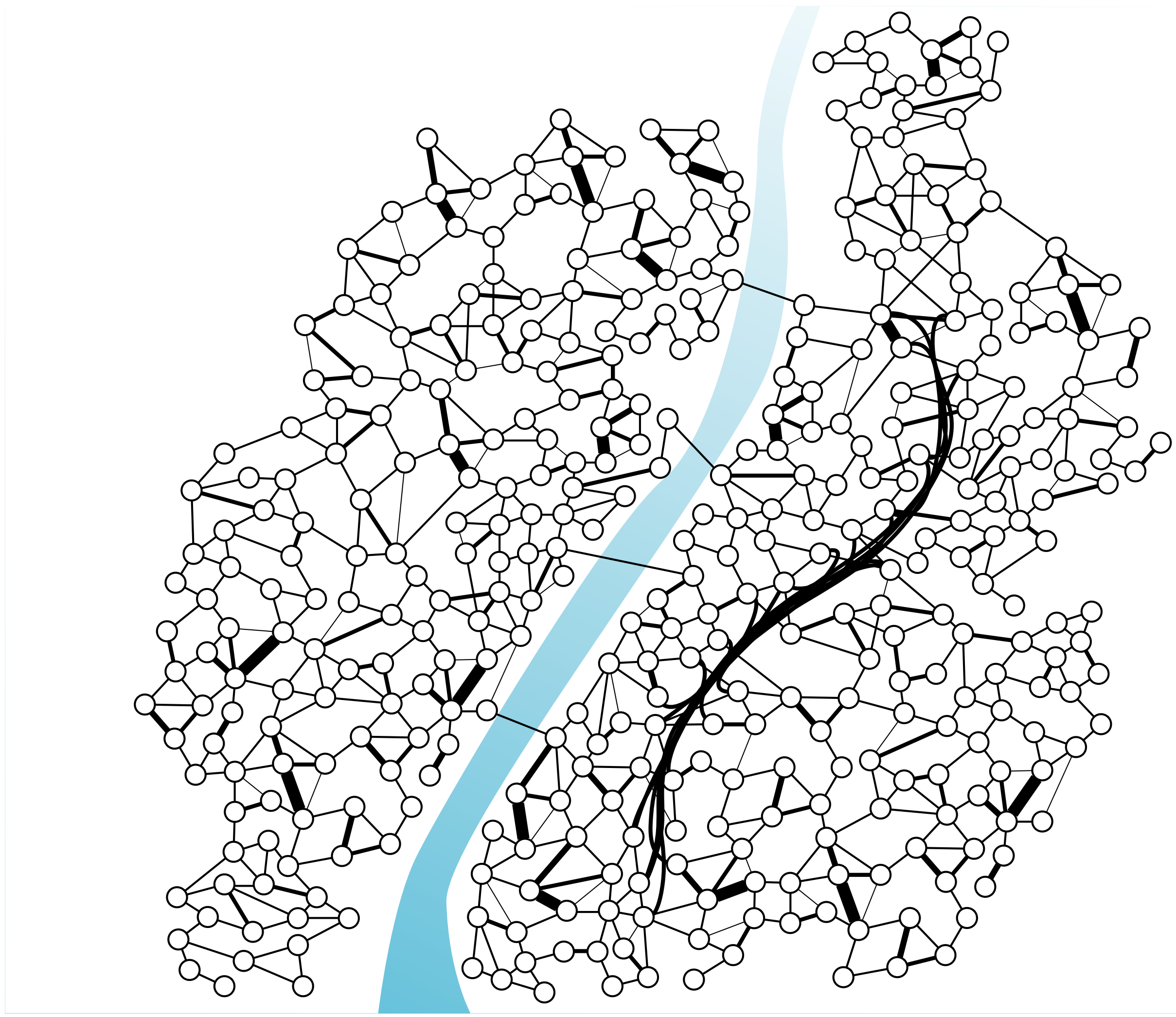


1500 BC

1500 BC

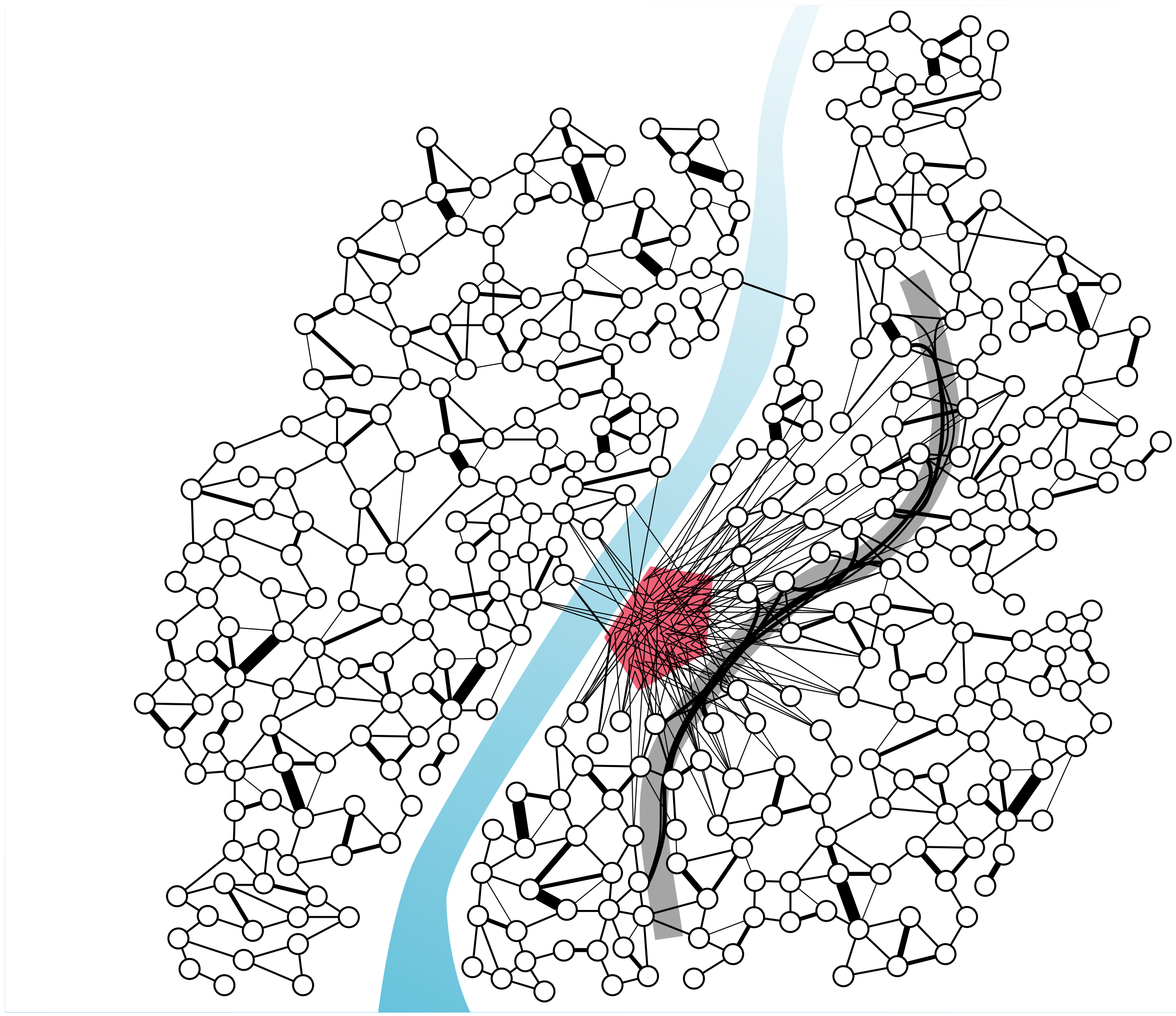


1500 BC

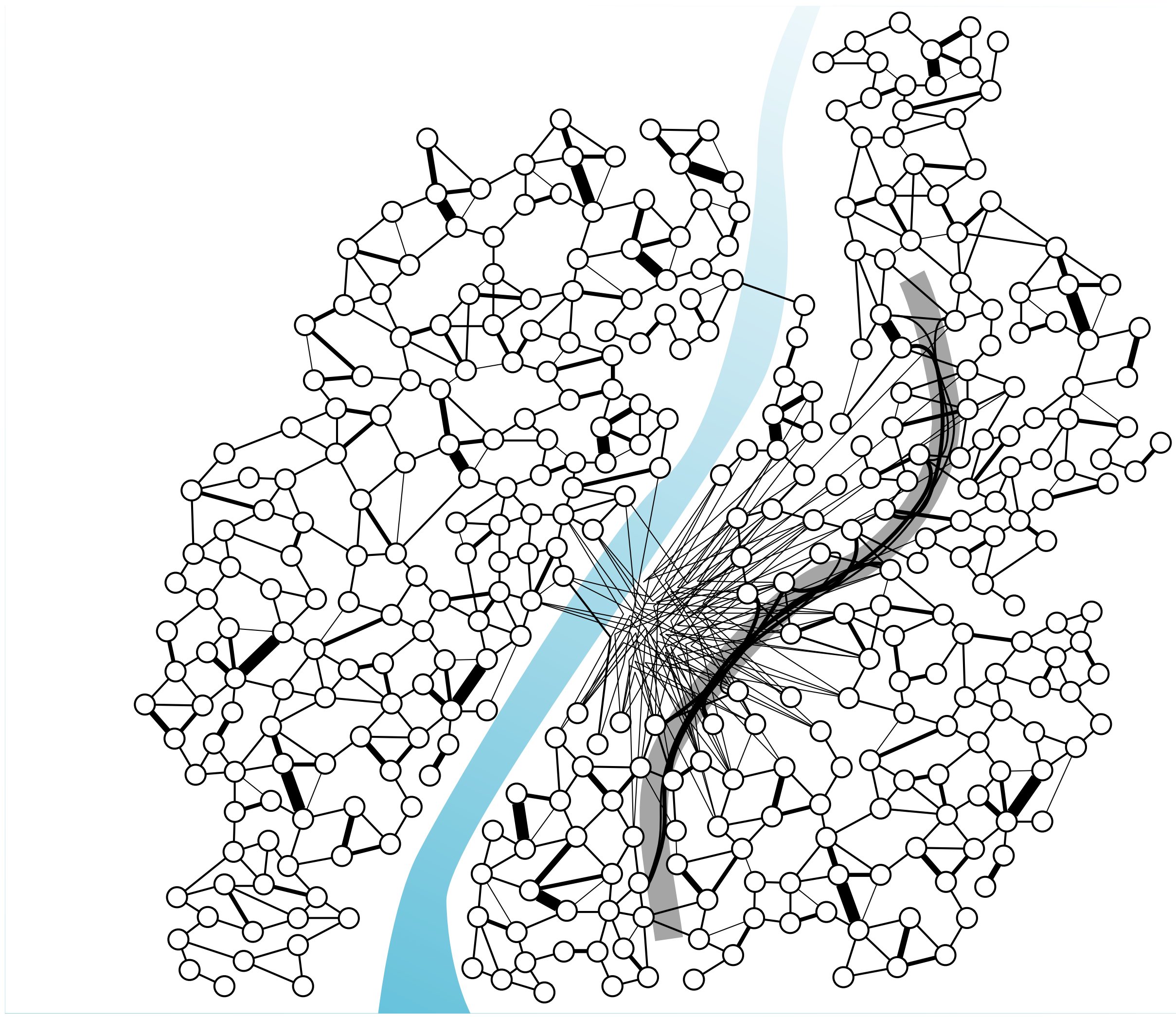


1000 BC

1000 BC

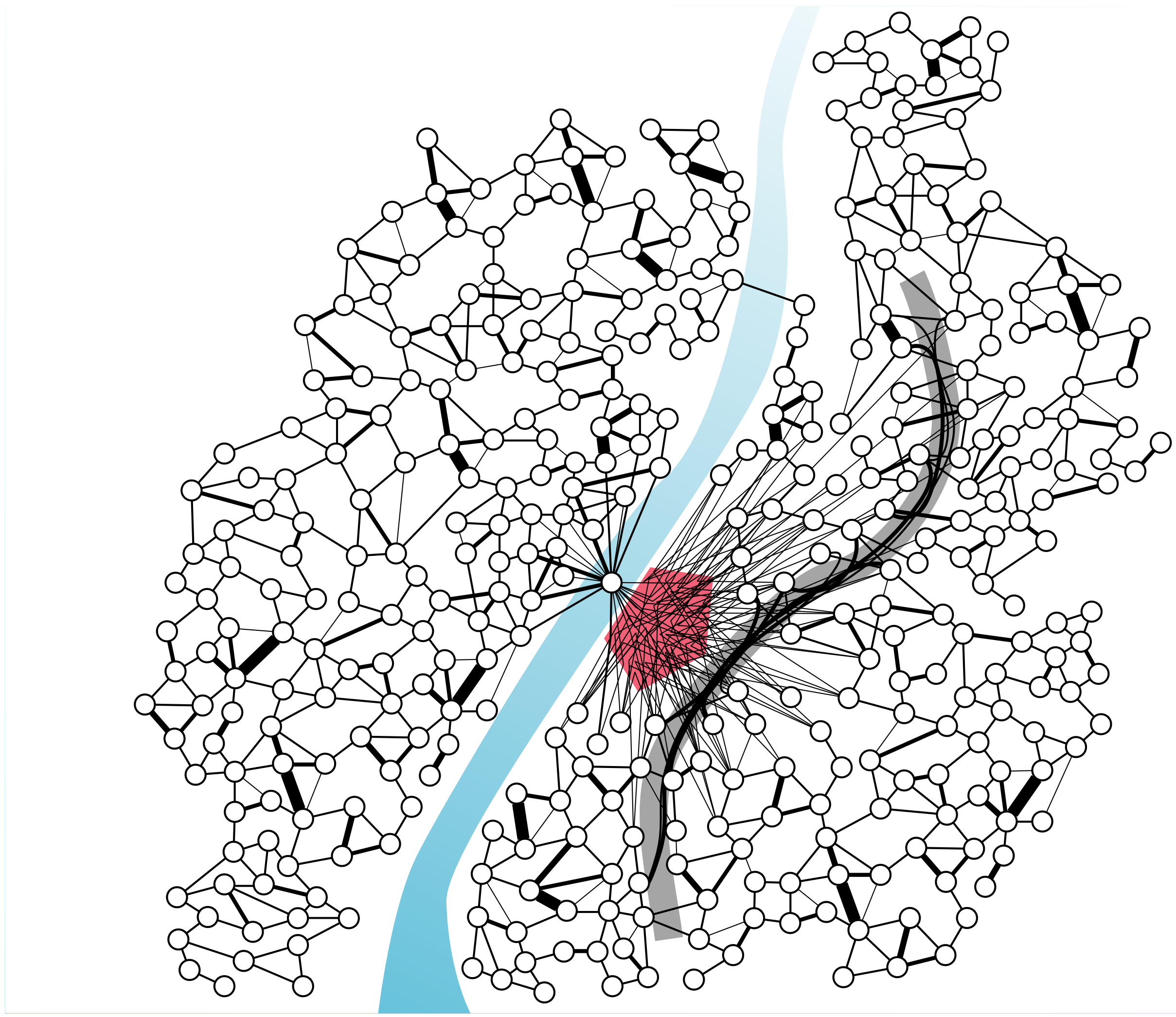


1000 BC

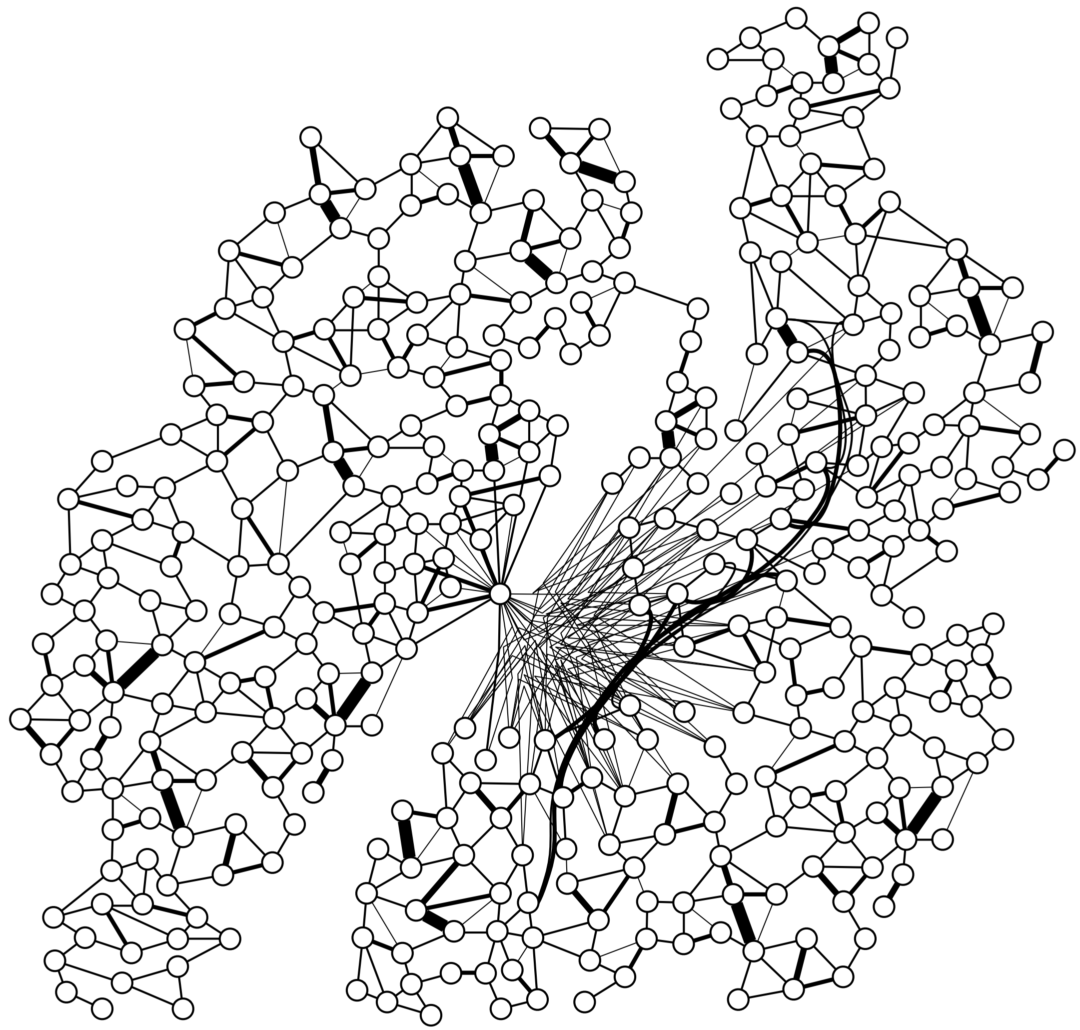


500 BC

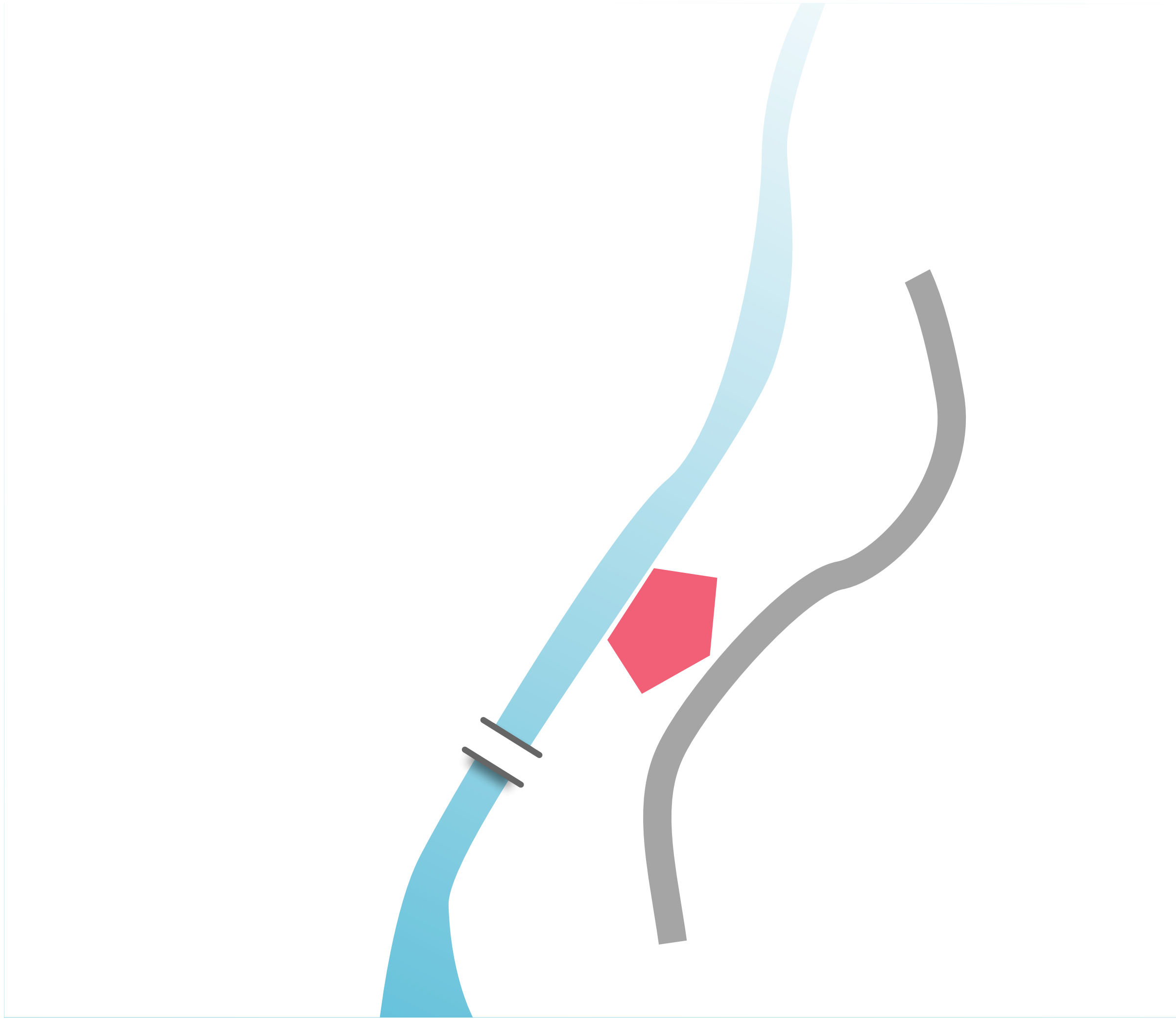
500 BC



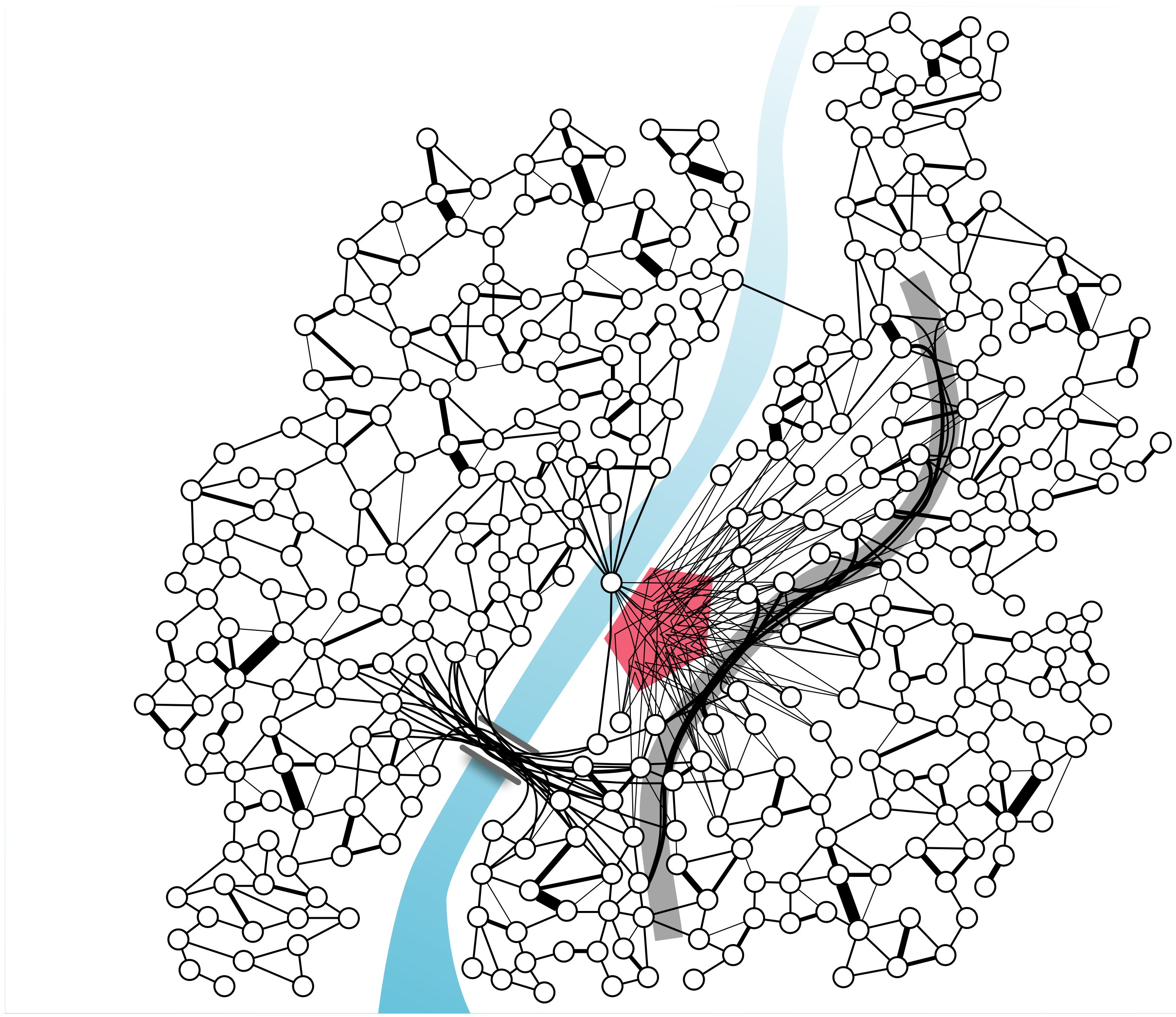
500 BC



100 AD

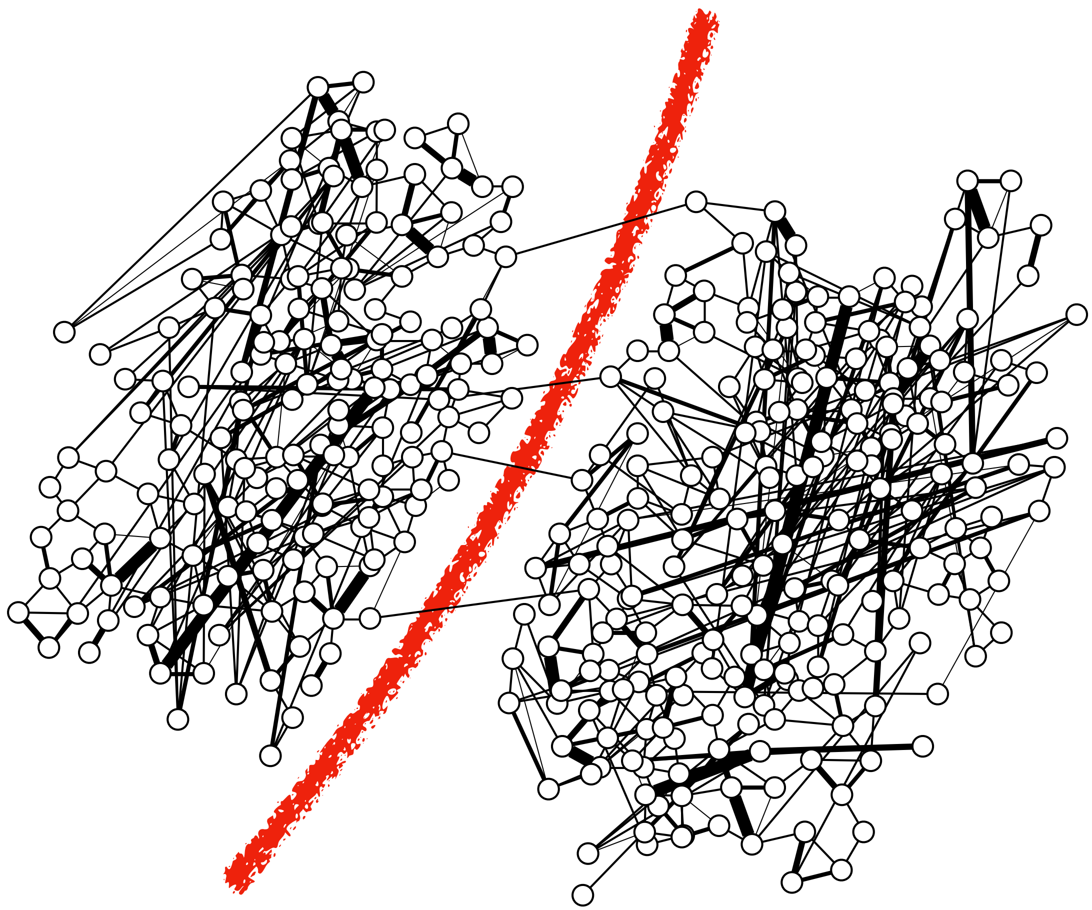


100 AD

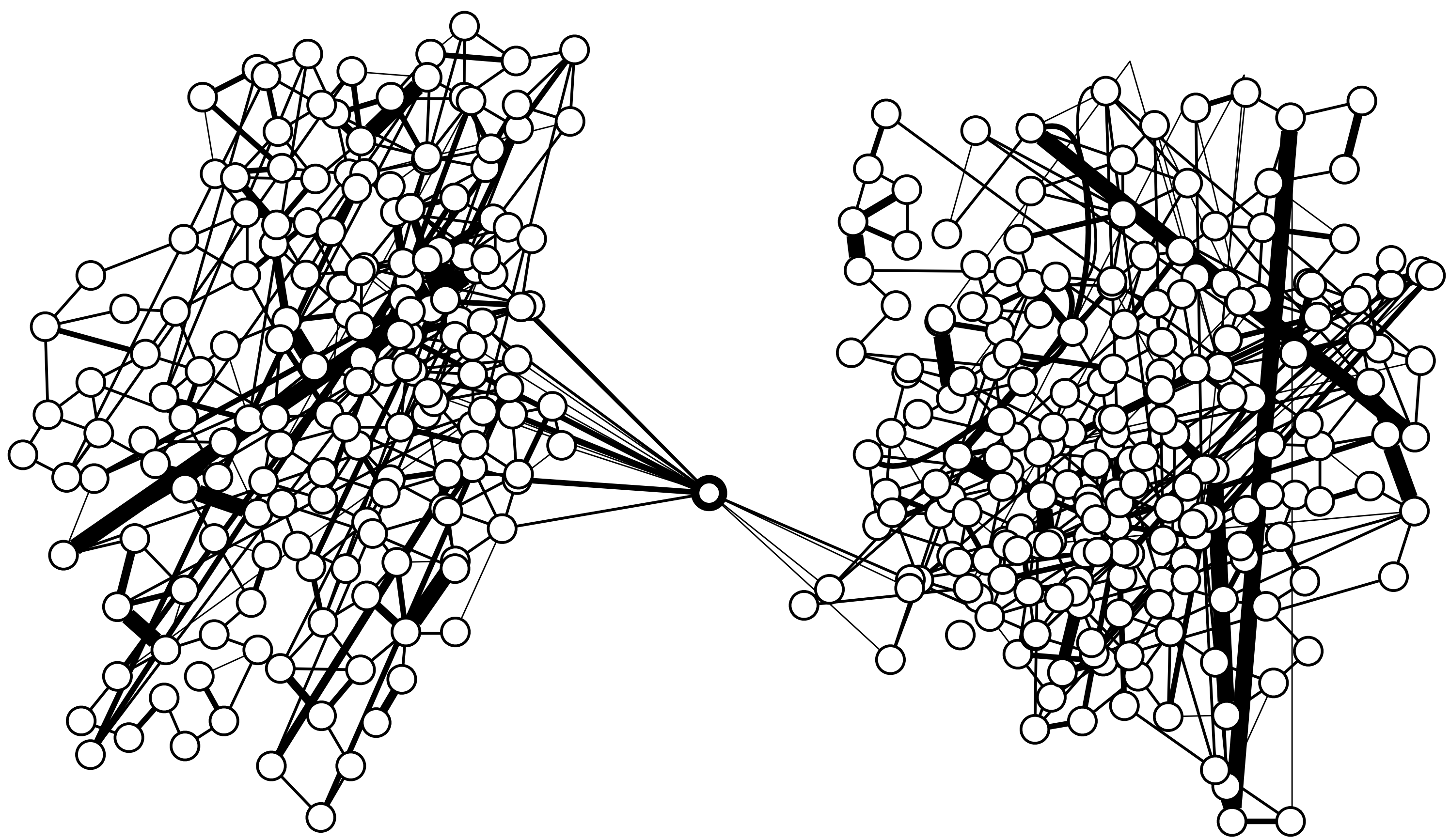


Goal: given trade network, find interesting features of economy

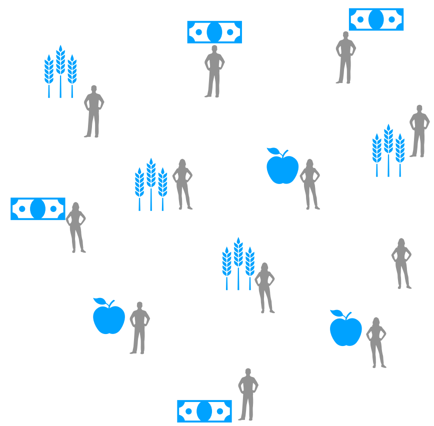
**Given this divide,
what is going on in this economy?**



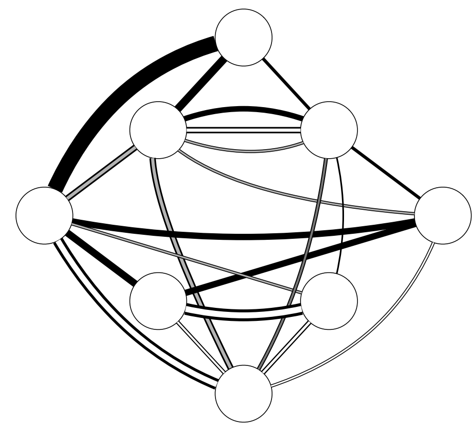
Monopoly: A node like this?



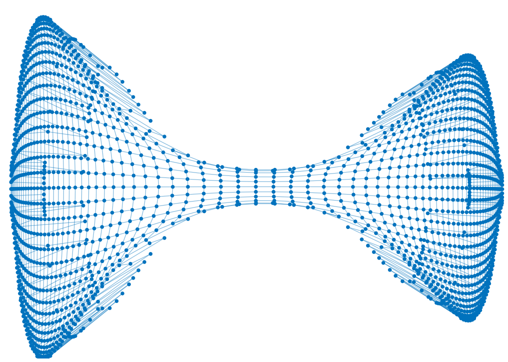
Outline



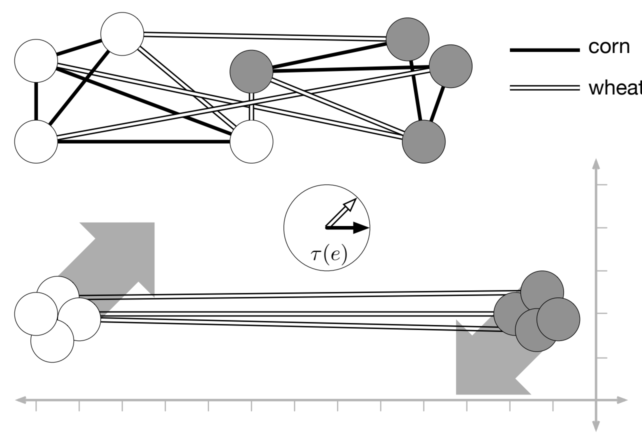
1. CS and Economics



2. Trade Networks



3. Graph Laplacians



4. Modifications for Econ

Graph Laplacian—
tool to interpret trade network

- Goal: mine data in trade network to find interesting features of economy
- Laplacian—matrix created from network
- Use eigenvalues and eigenvectors of matrix to determine *large scale shape* of the network.
- Use shape to discover interesting features of economy!

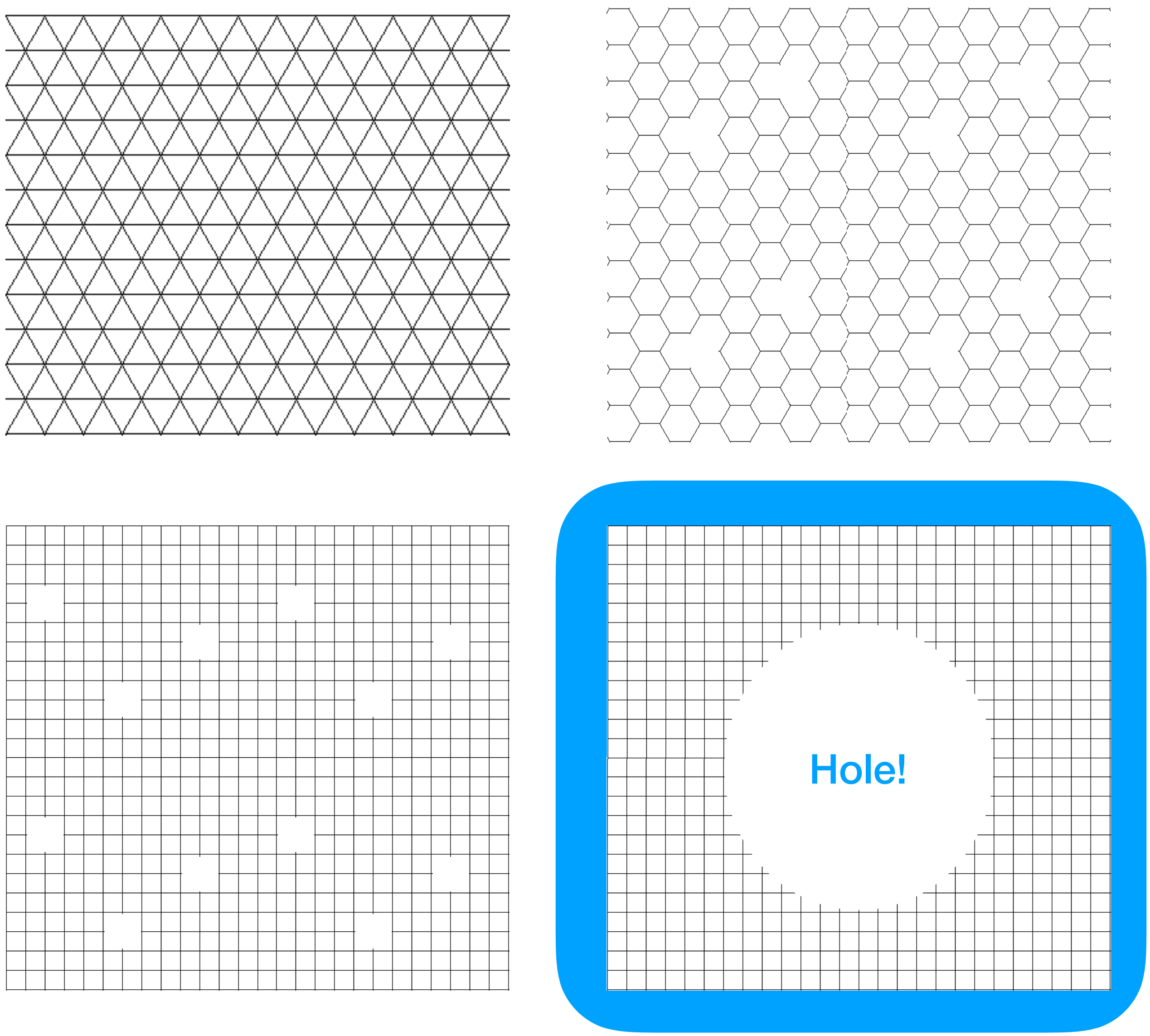
Table or Chair?

Table or Chair?

Table or Chair?

Intuition: at scale network encodes *shape* of economy

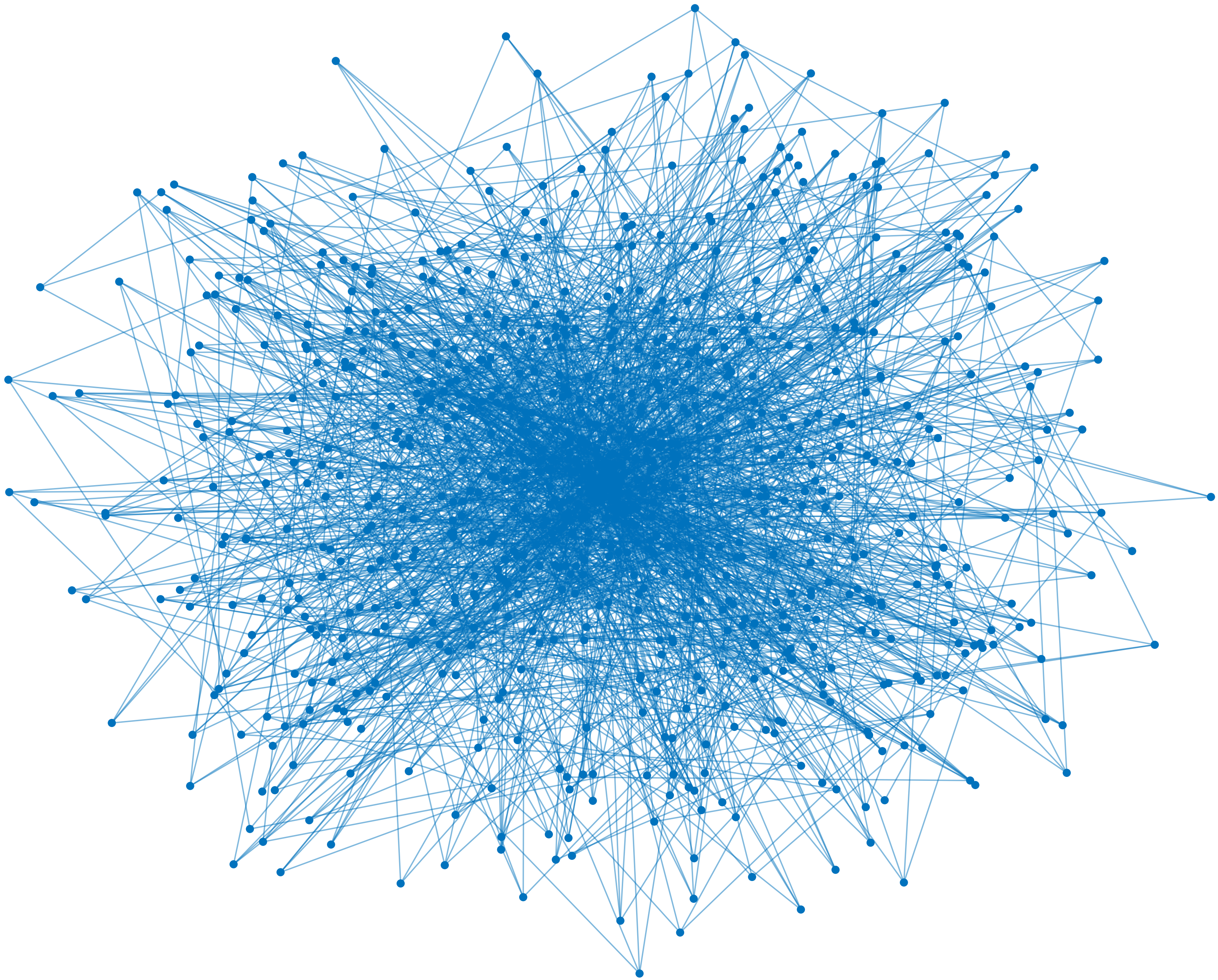
- Can use algorithms to determine *large scale* properties of network while ignoring small scale properties
- Robust to small changes!
- (right kind of) approximations to network are enough!



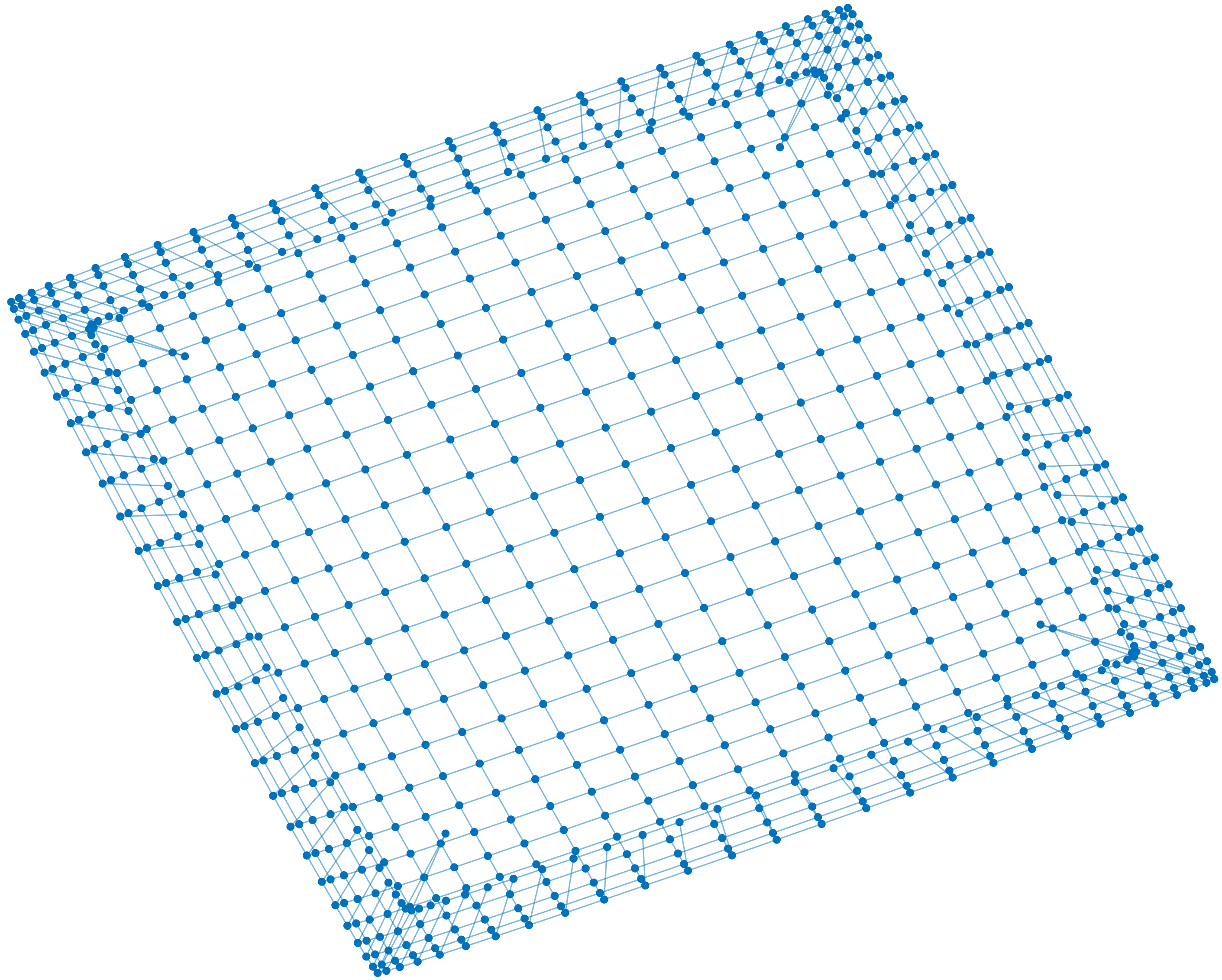
one of these is not like others

Show and Tell

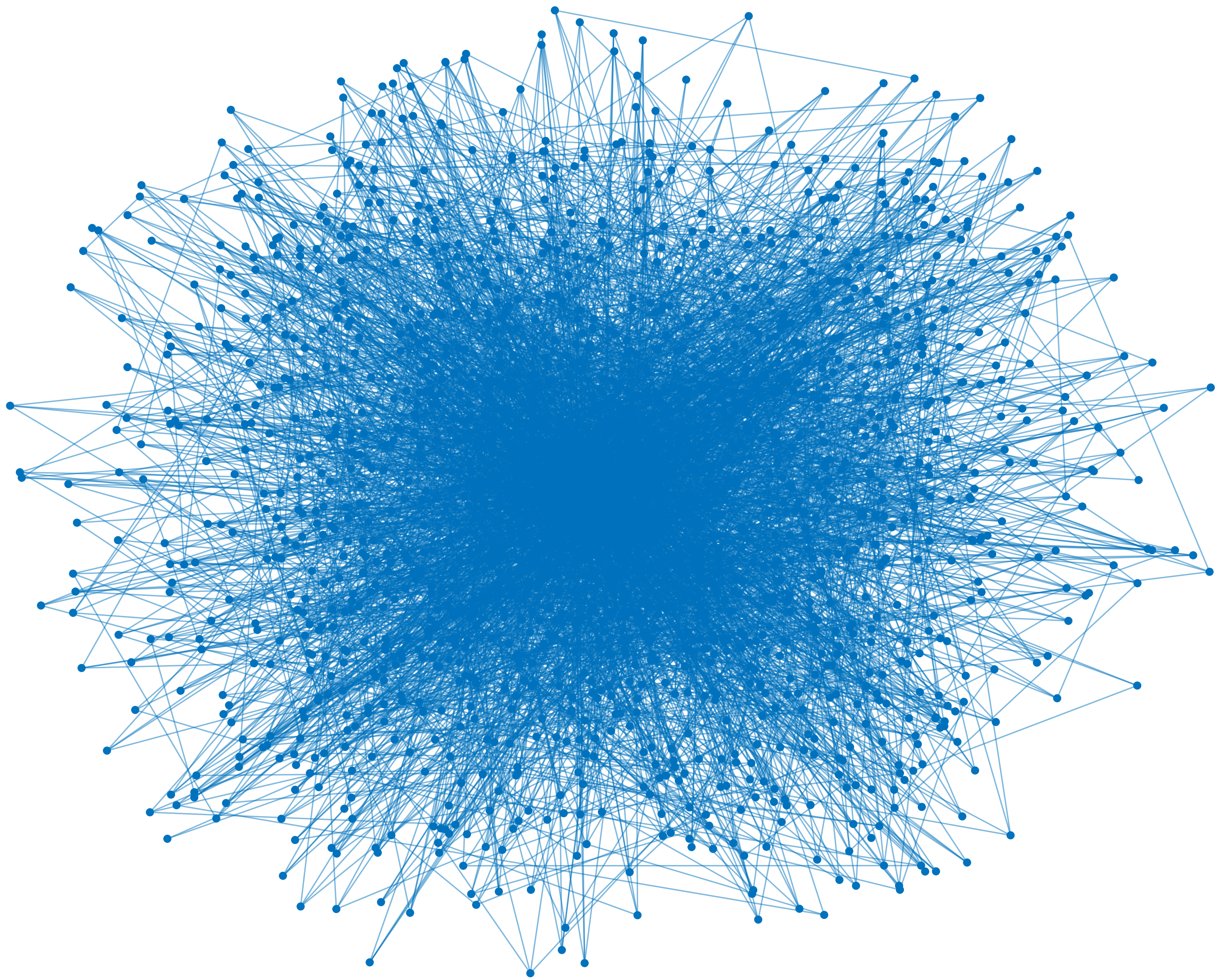
- Algorithm gets trade network as input
- **No other data** is shown to algorithm
- Algorithm untangles the mess and draws picture
- <u>Intuition for algorithm:</u>
  links are **strong** rubber bands—attraction
  nodes repel each other **weakly** —repulsion
  nodes are released and left to settle
- Intuition is only approximate, and sometimes completely wrong



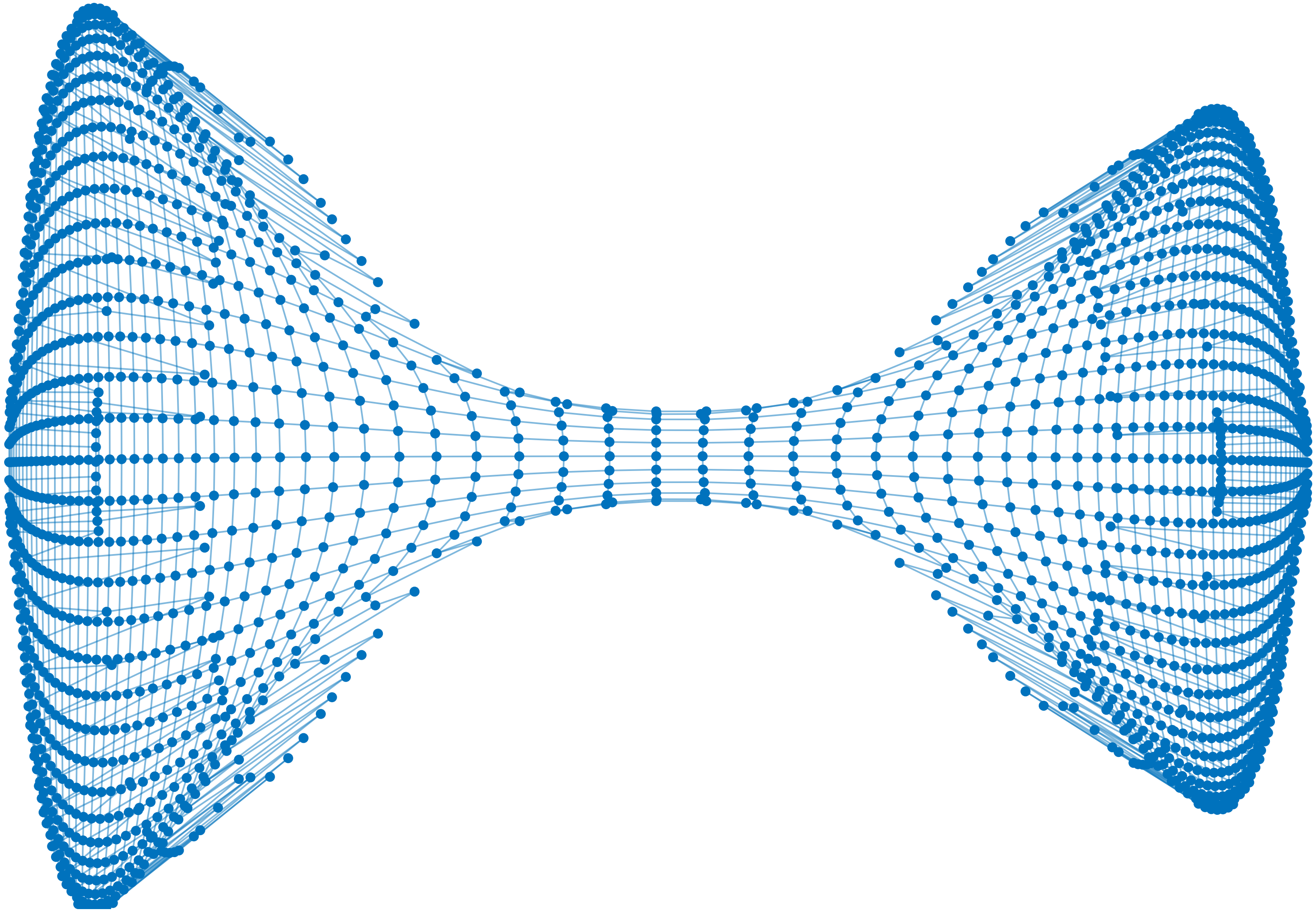
30 by 30 grid



30 by 30 grid
(unscrambled by algorithm)



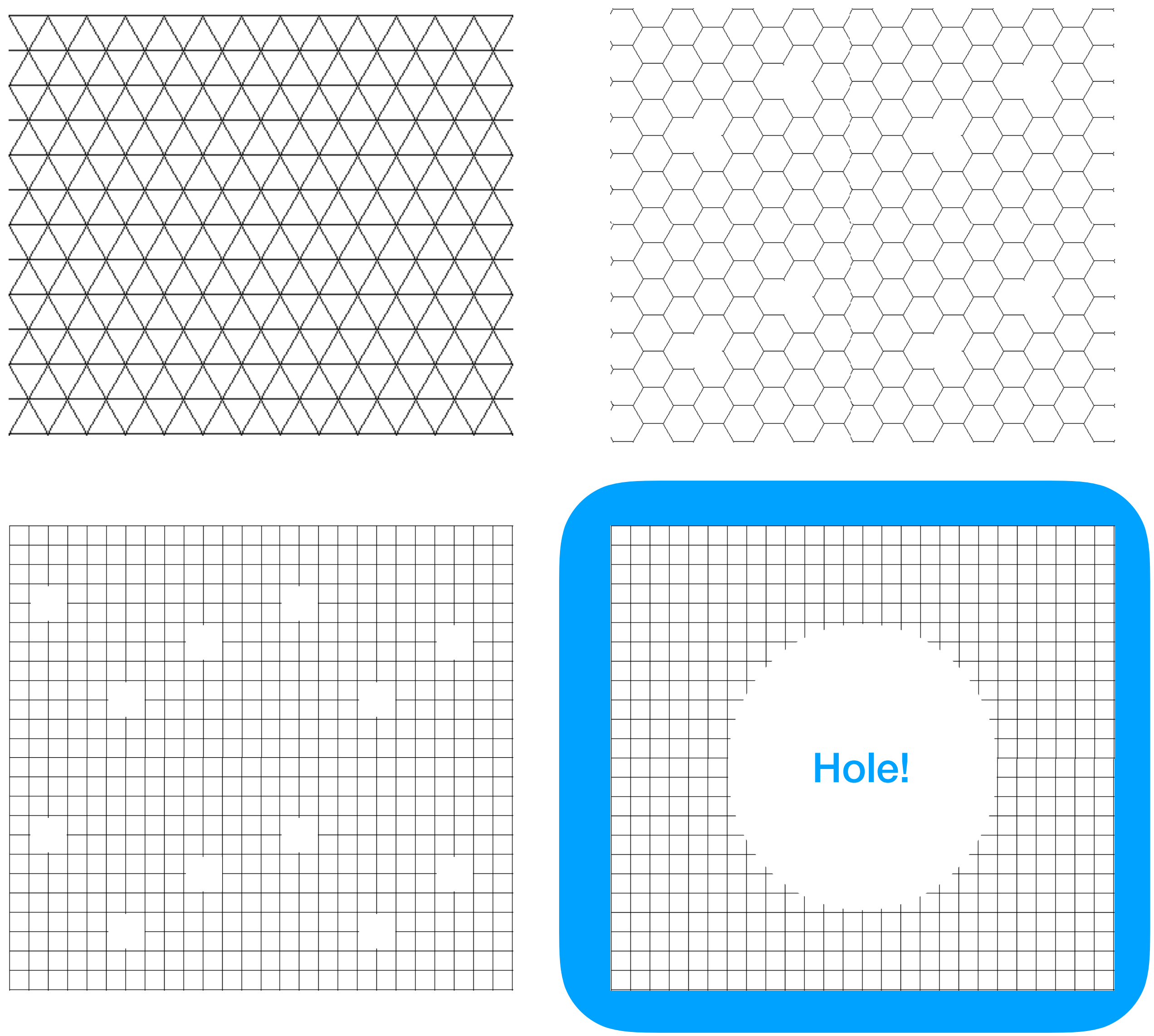
Dumbell network



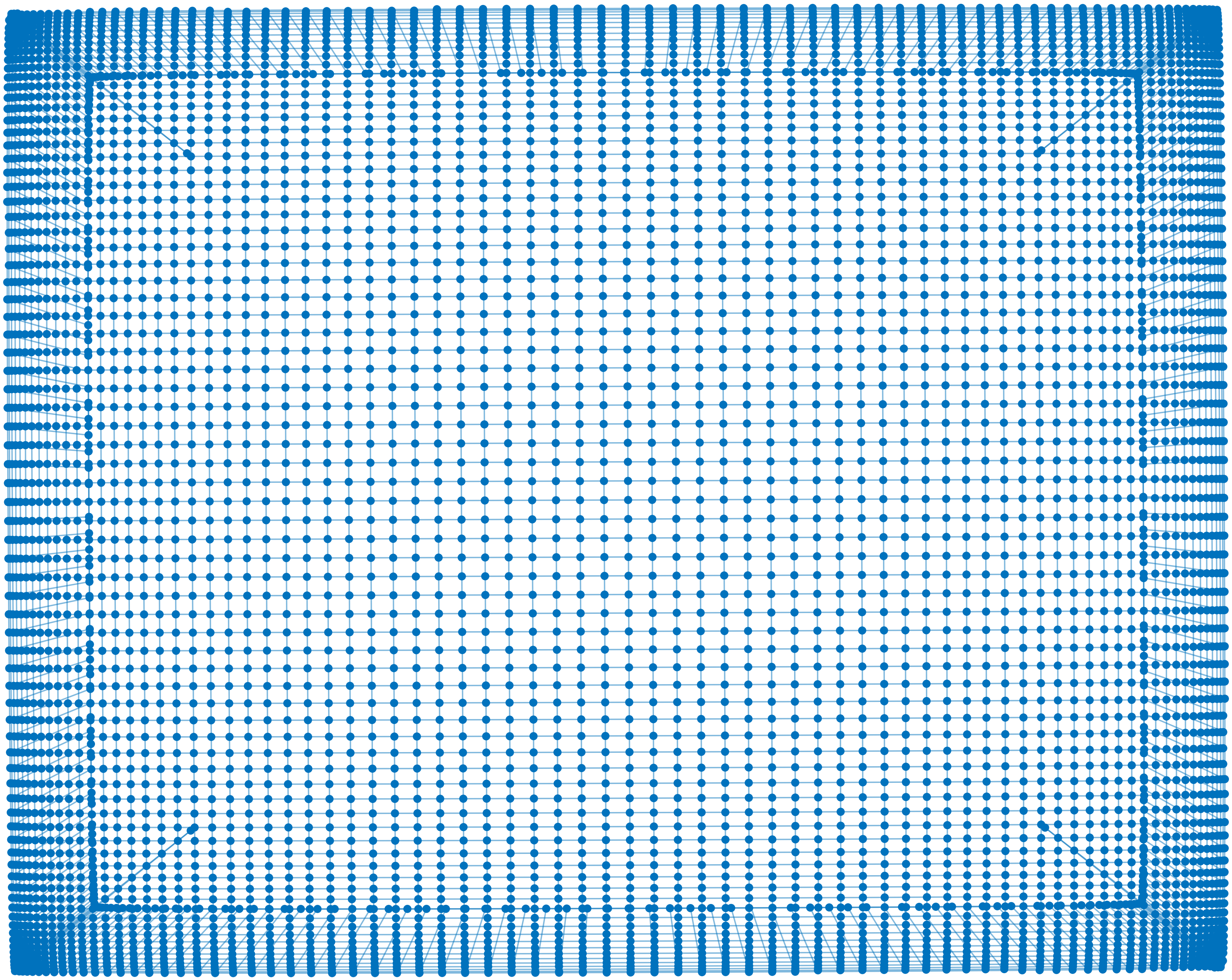
Dumbell network
(unscrambled by algorithm)

Can answer things like: what is the center of the network?
(Eigenvalue centrality would give a different kind of answer!)

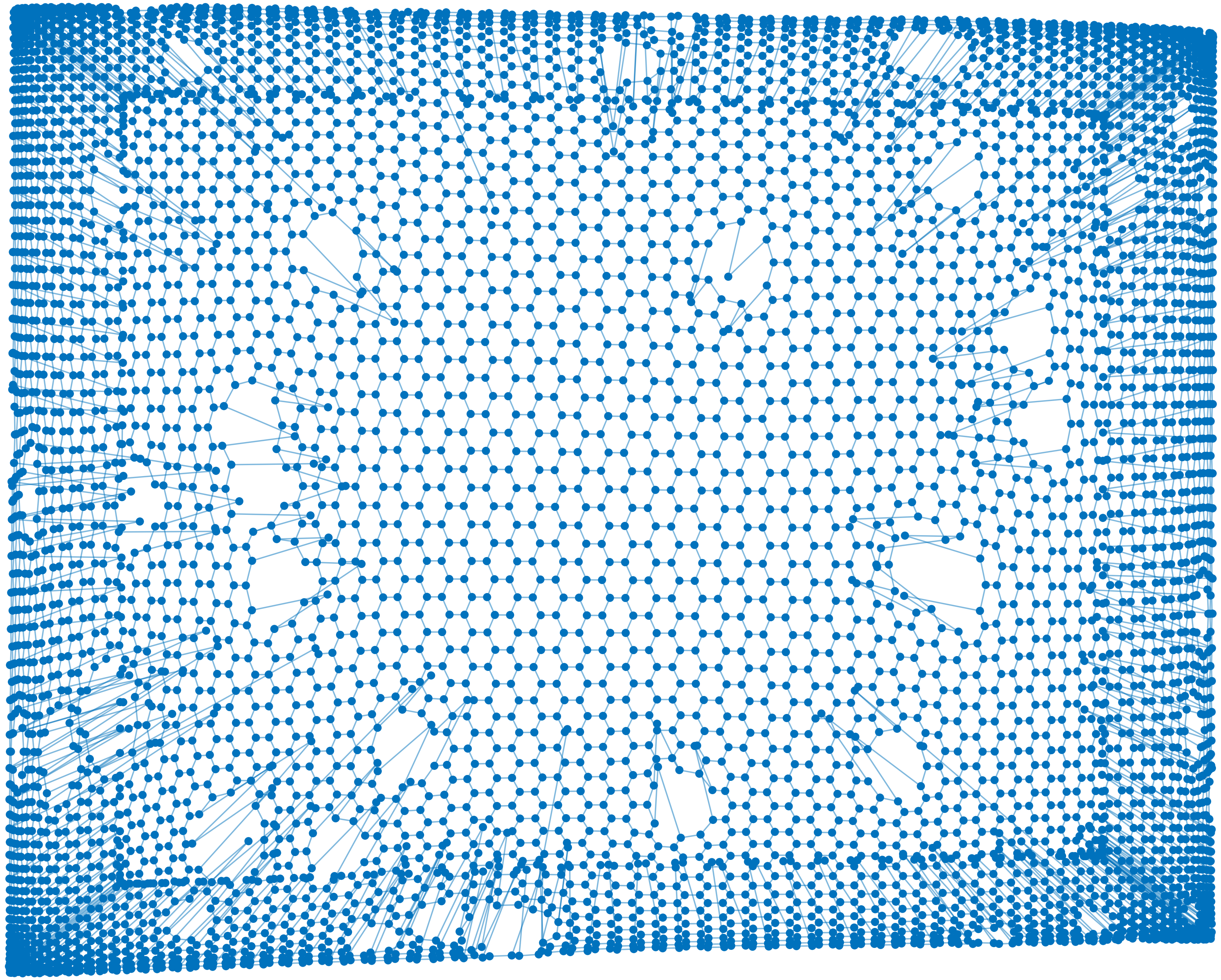
Intuition: at scale network encodes *shape* of economy



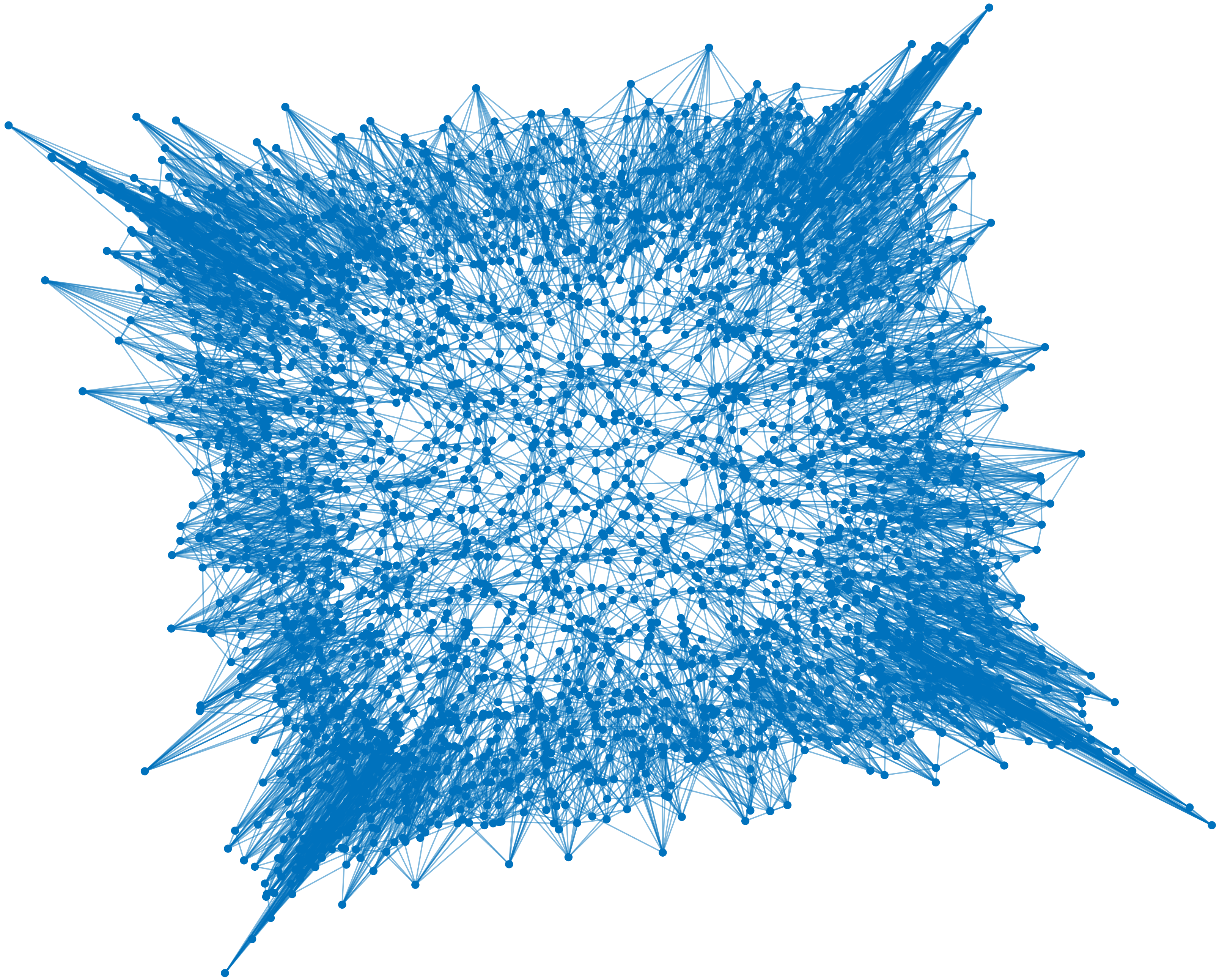
one of these is not like others



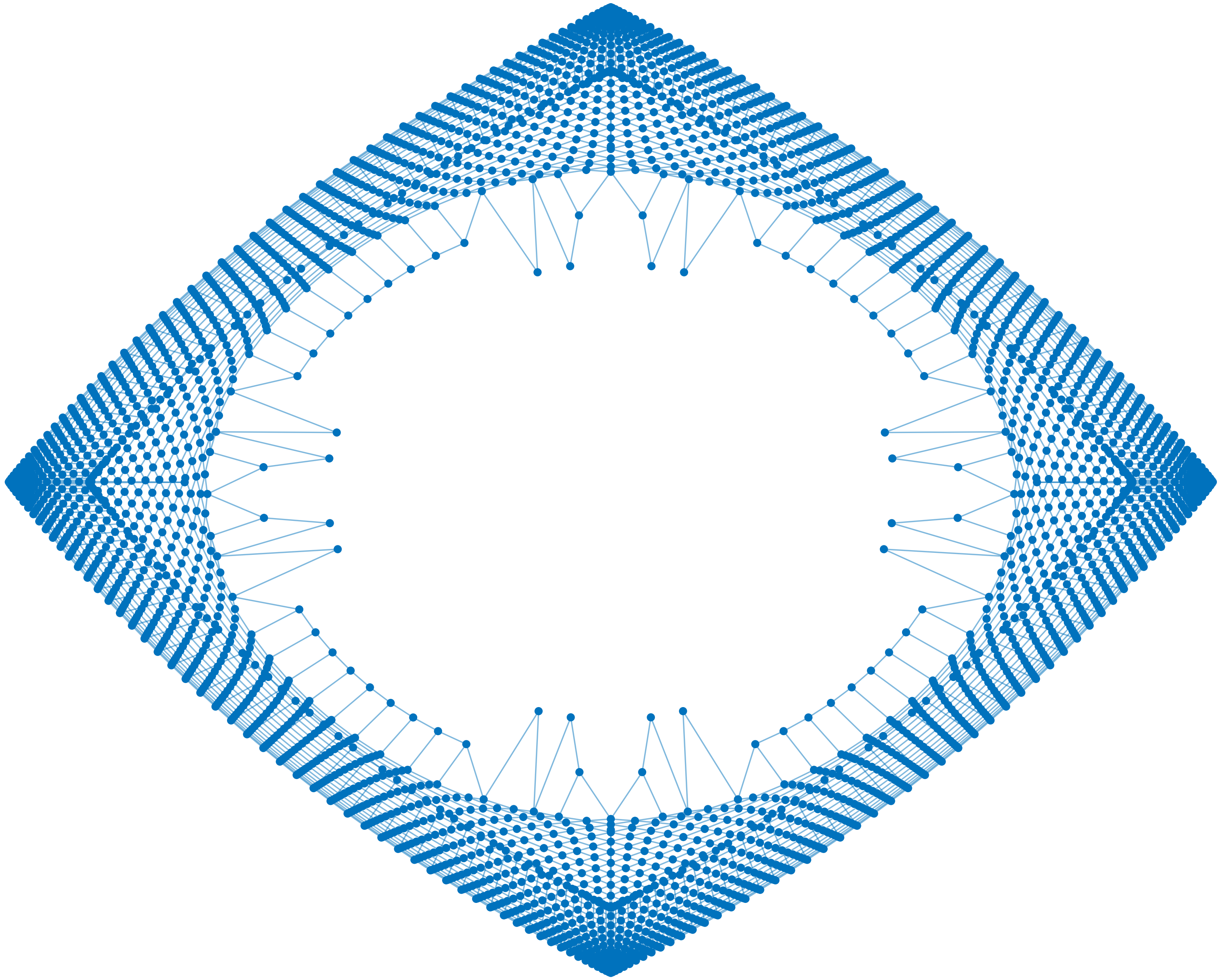
80 by 80 grid



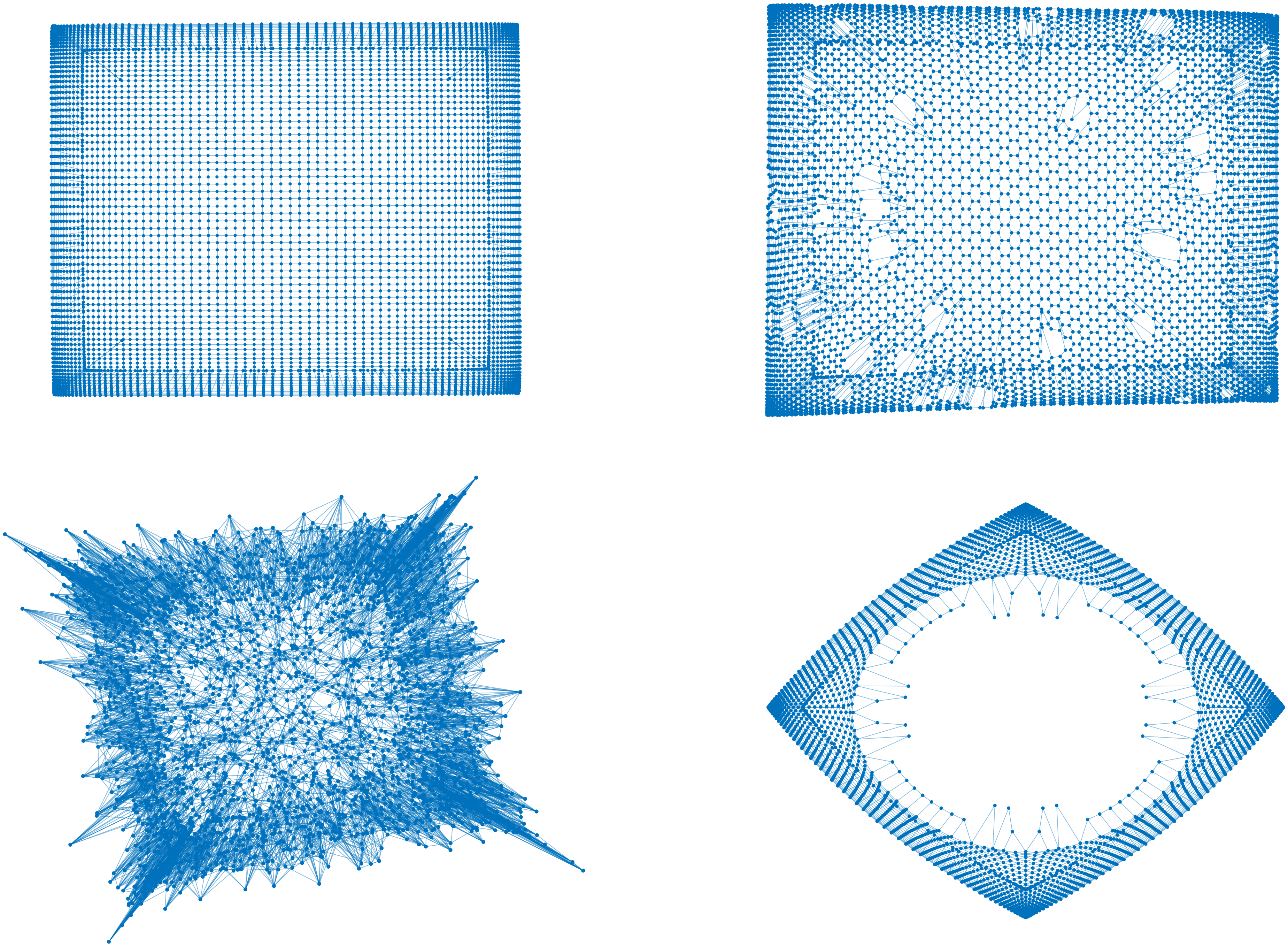
Hexagonal grid with 50 random holes



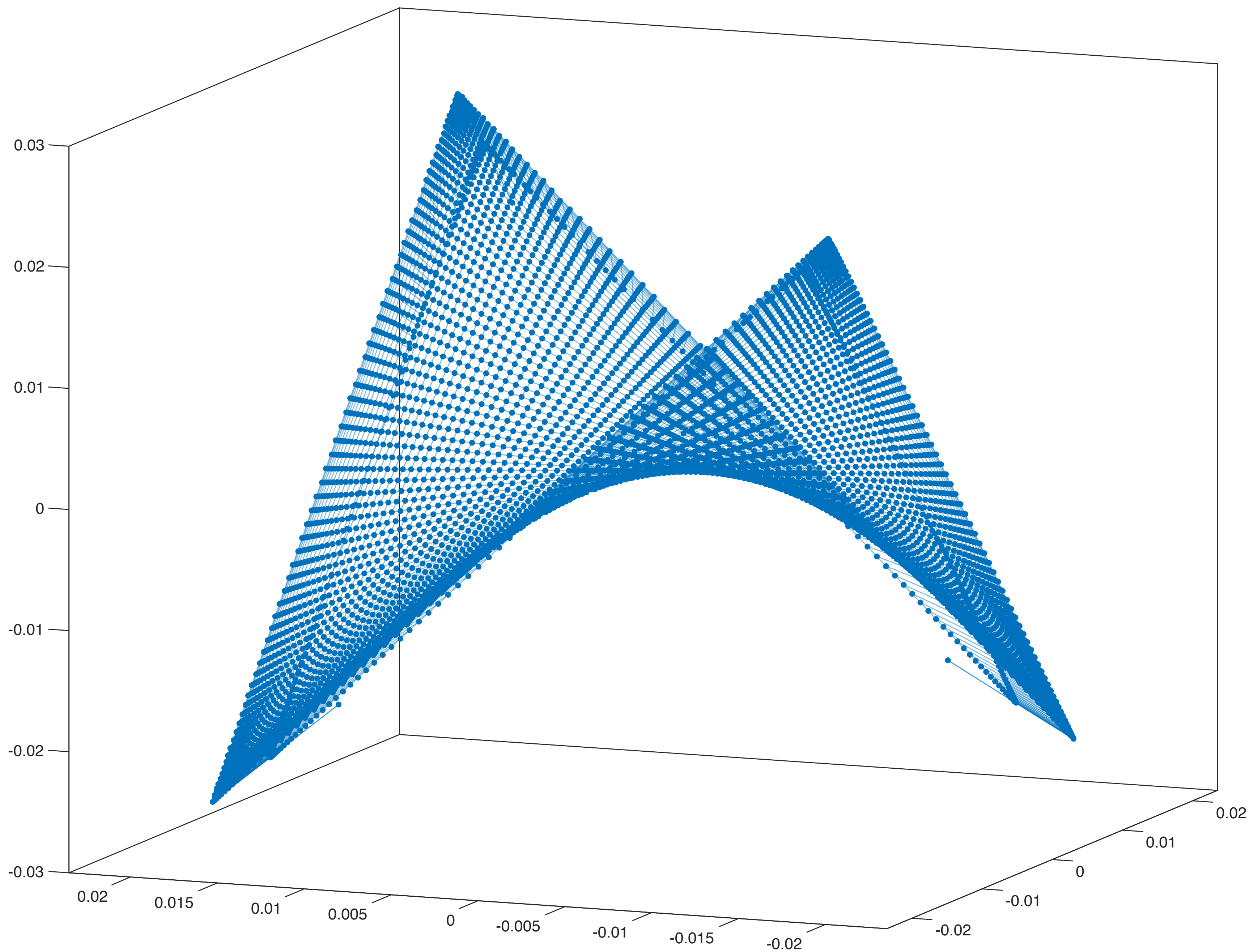
Random links to nearby points on grid



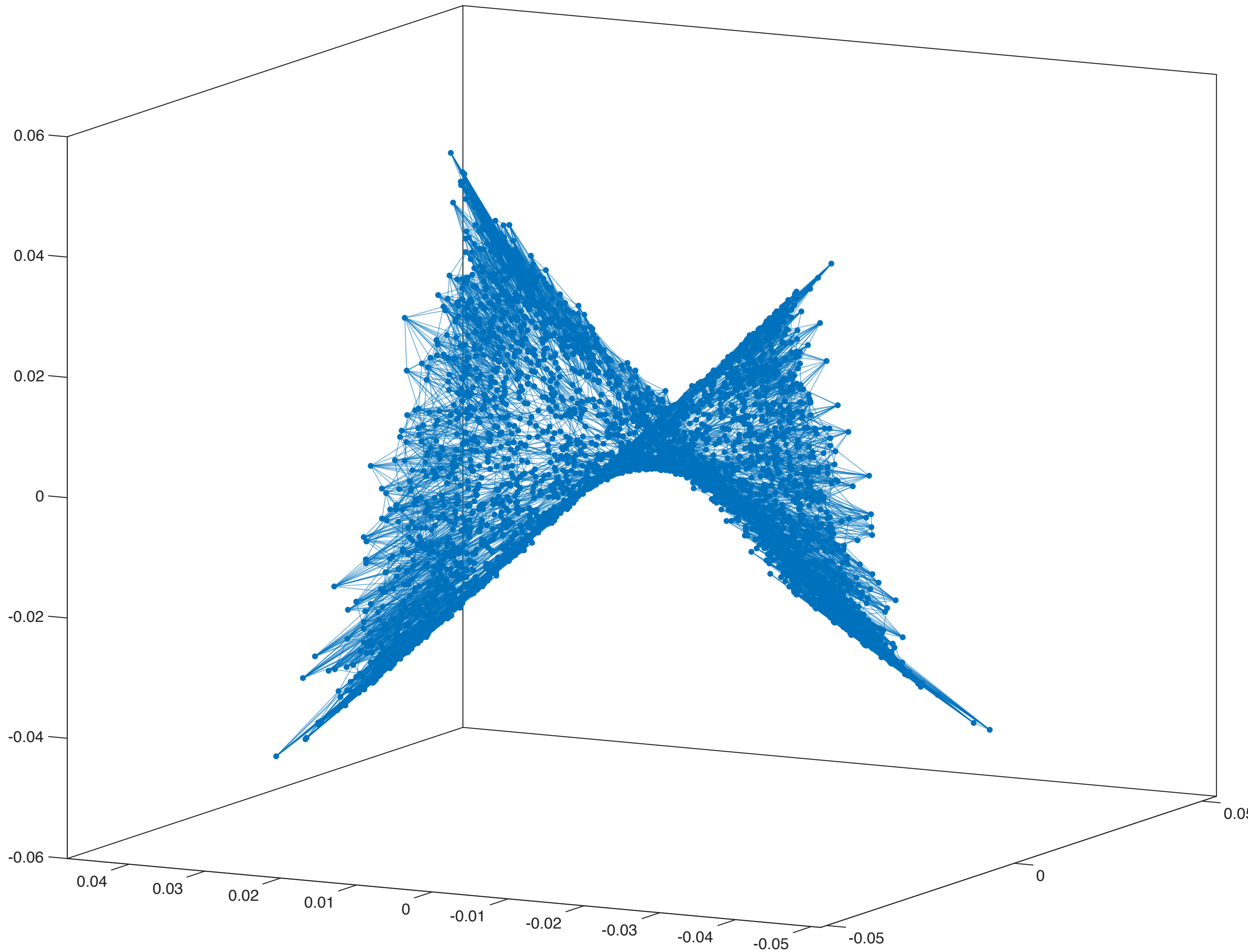
Grid with hole



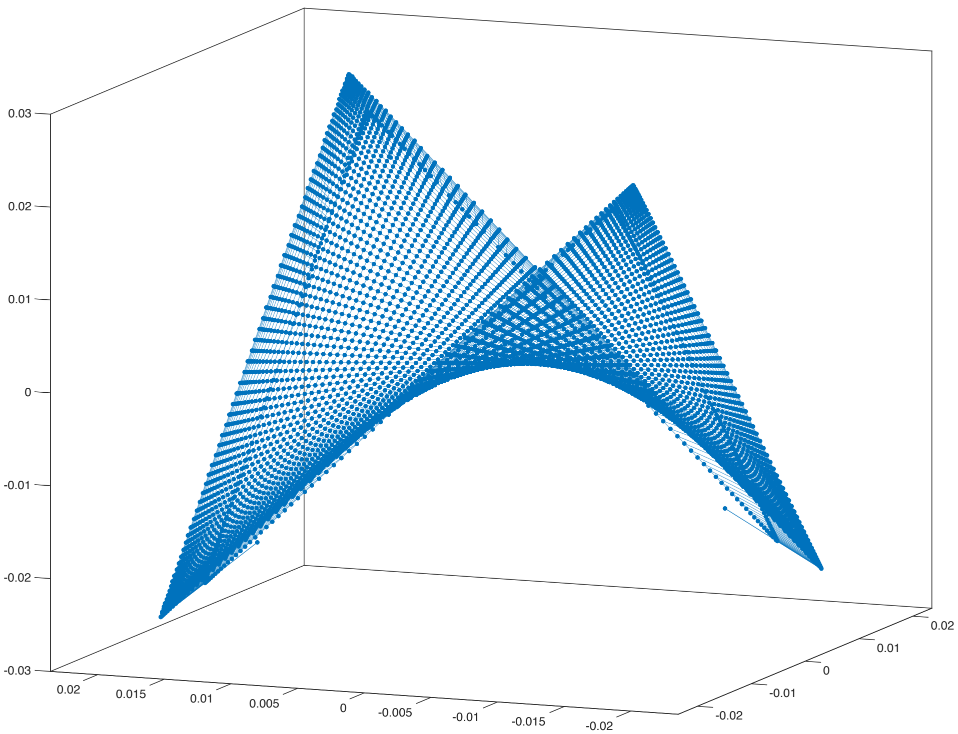
Even computer can distinguish hole from no hole

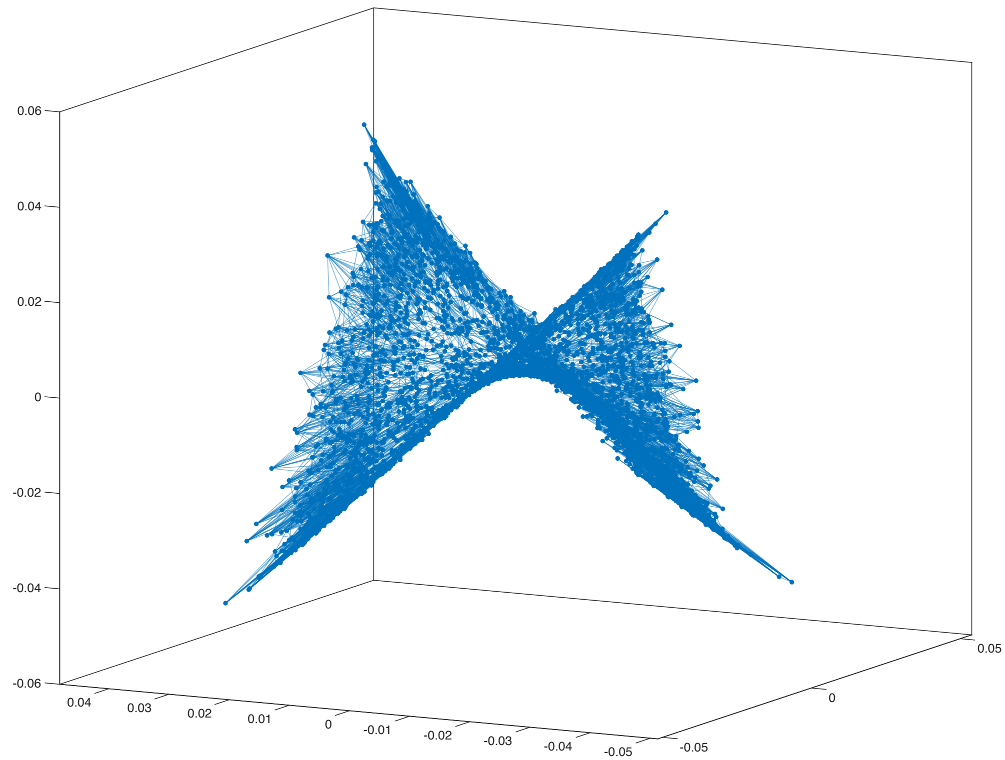


80 by 80 grid
(alternate view)



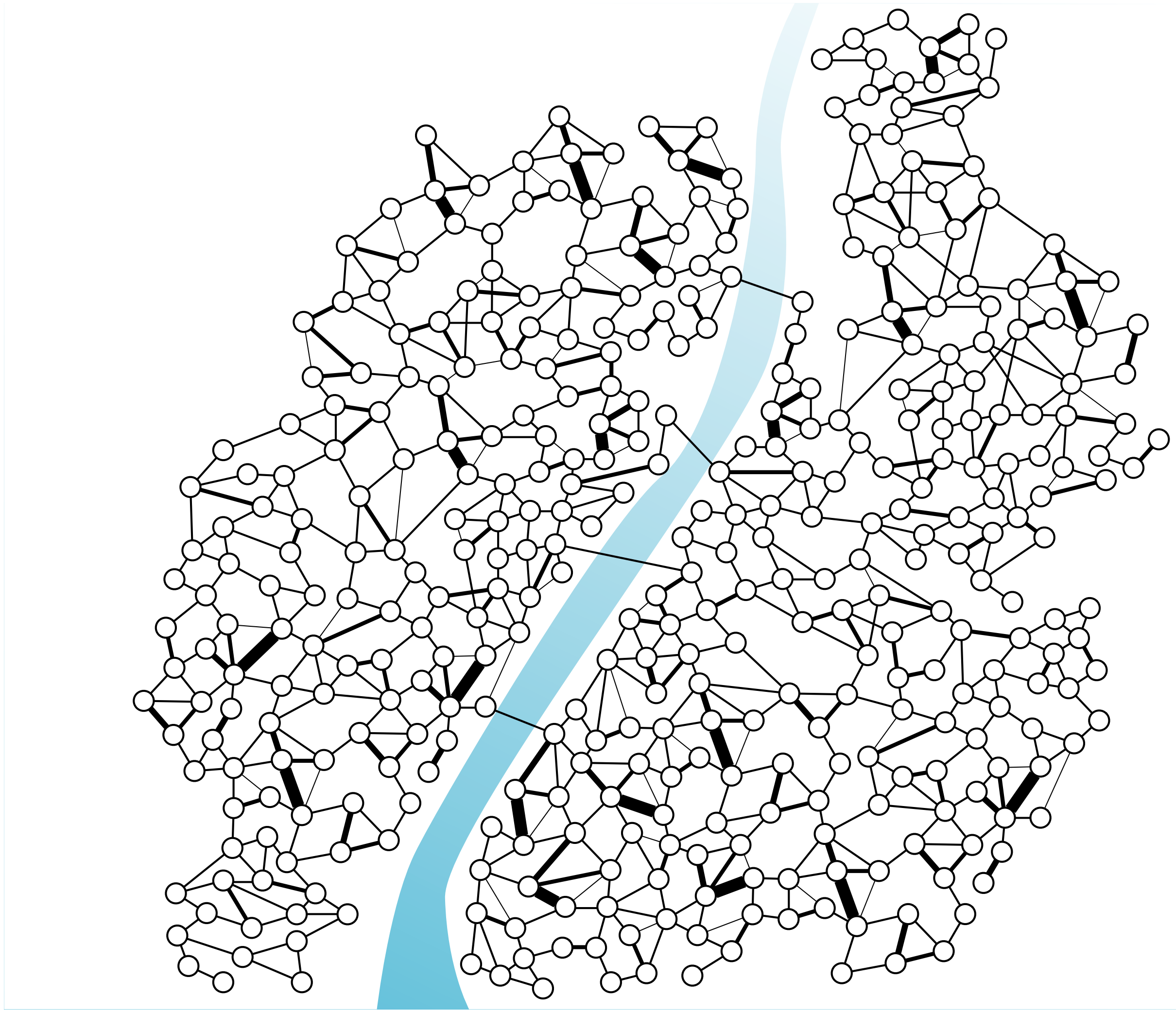
80 by 80 grid with random links
(alternate view)



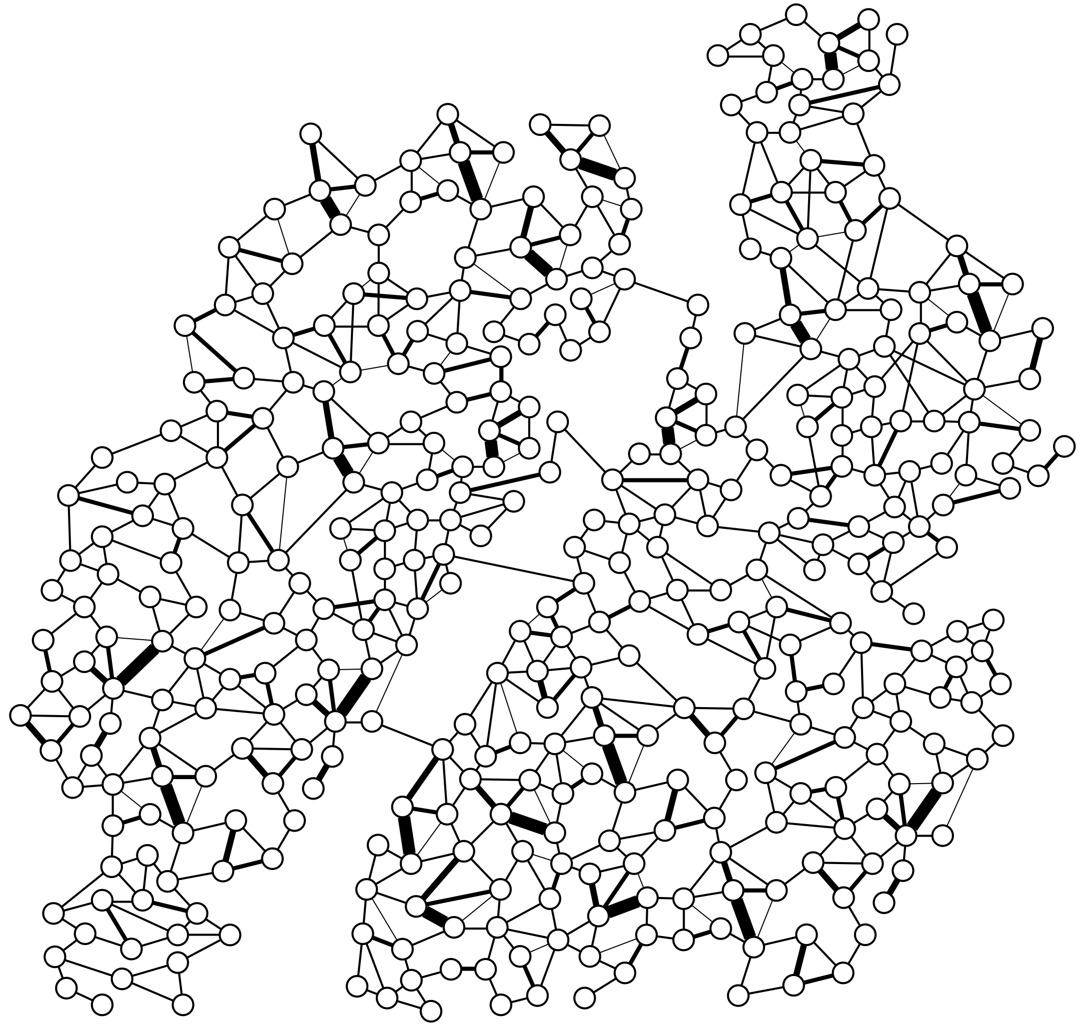


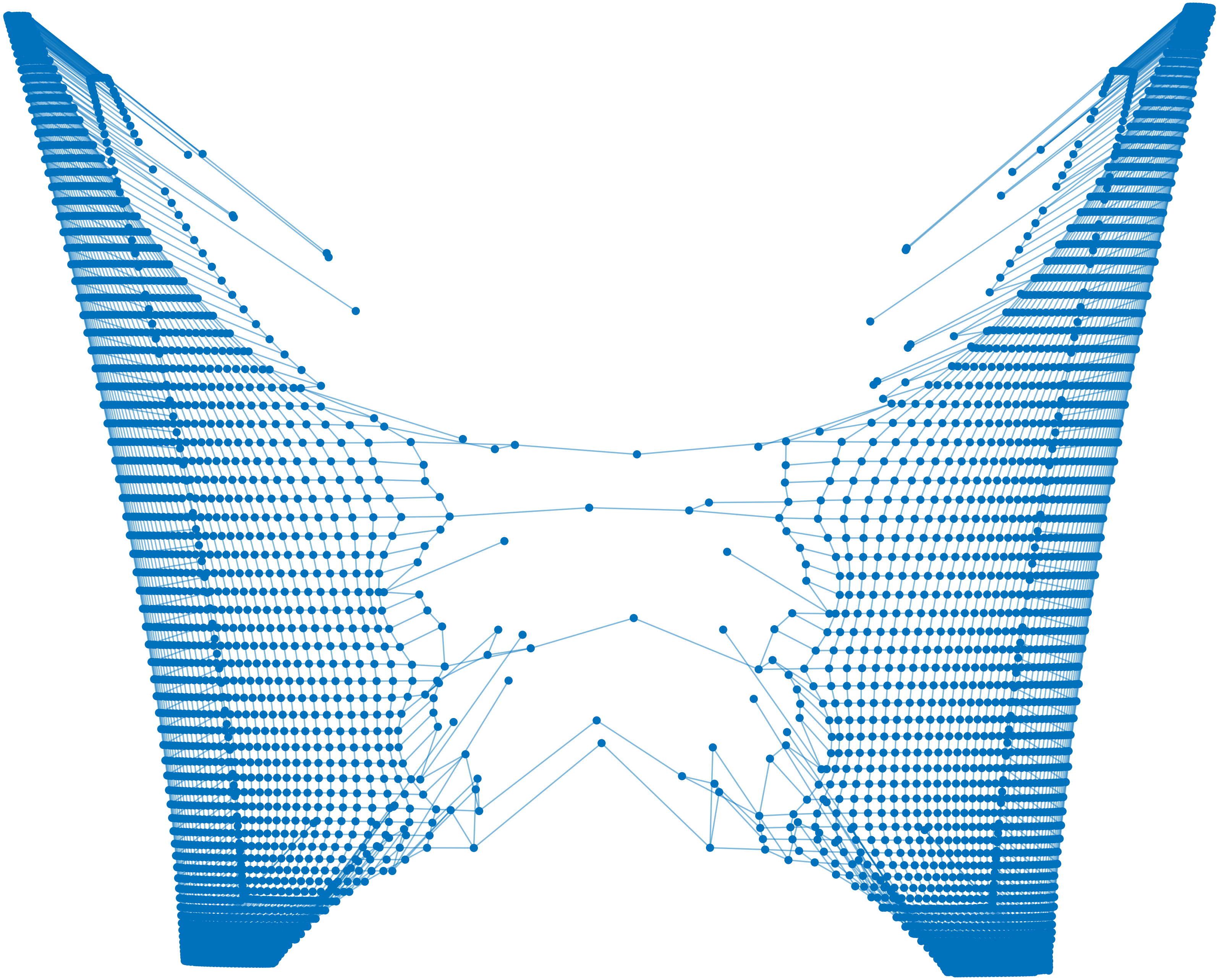
This is the same shape!
(quantifiably so)

2000 BC



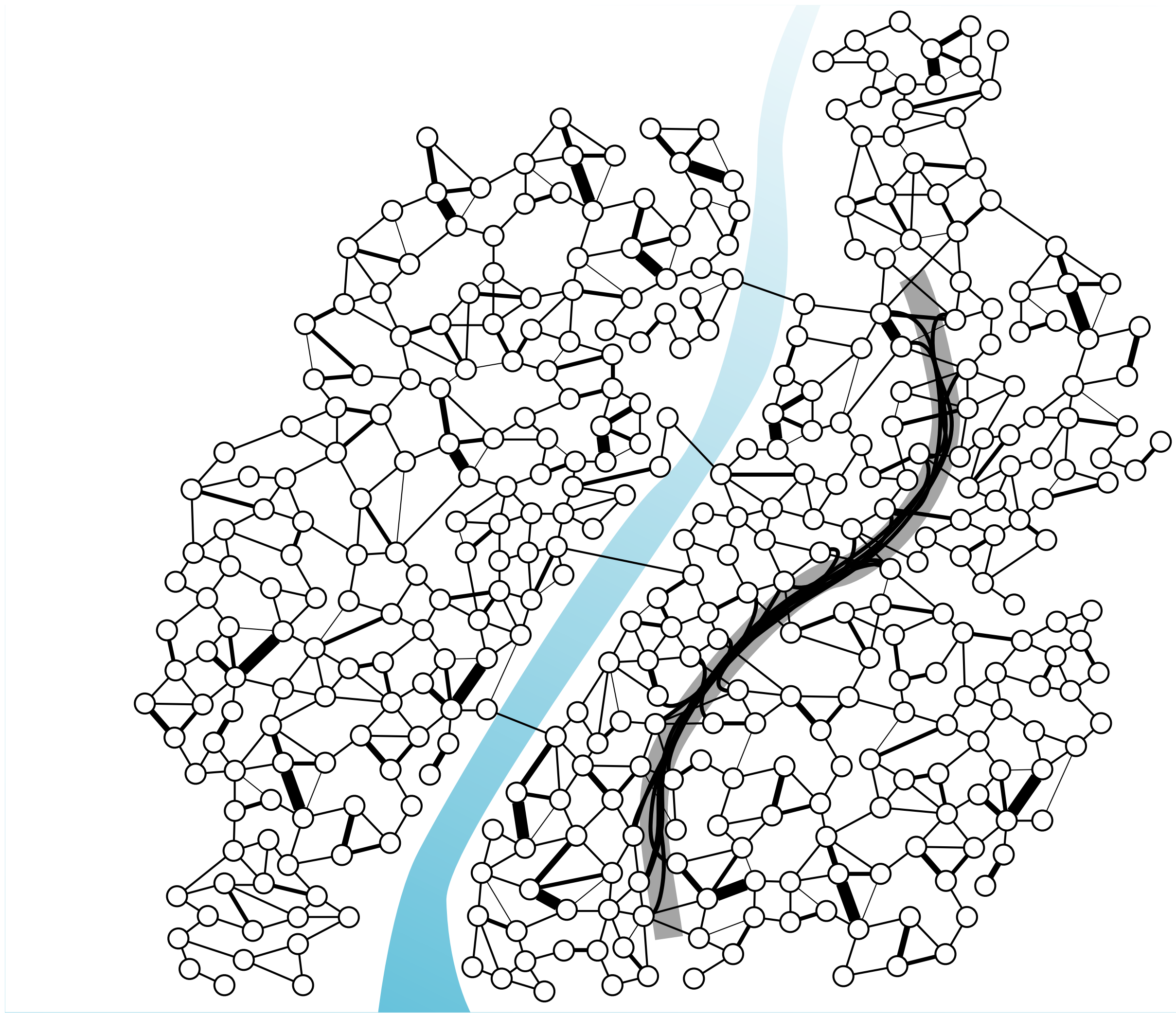
2000 BC

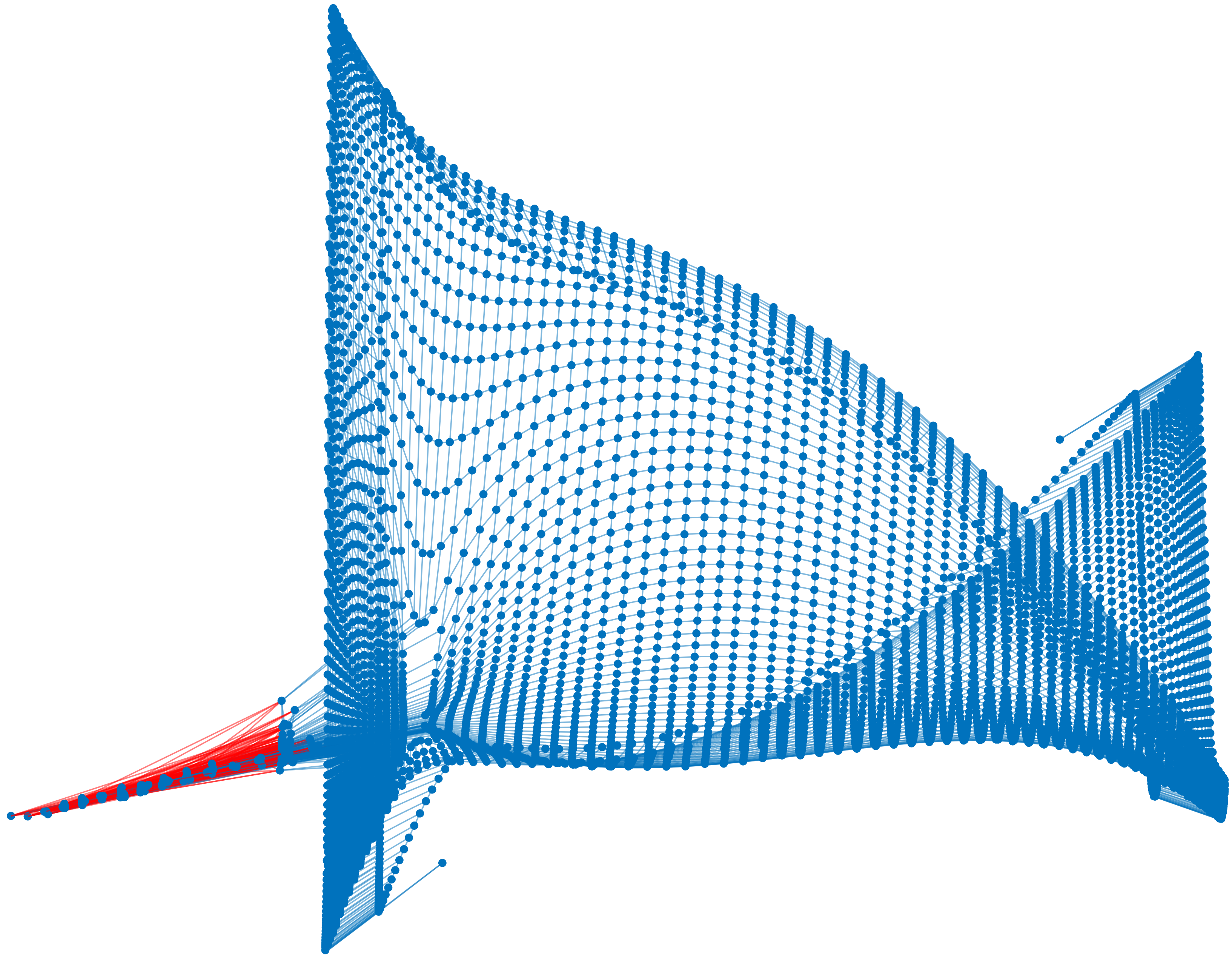




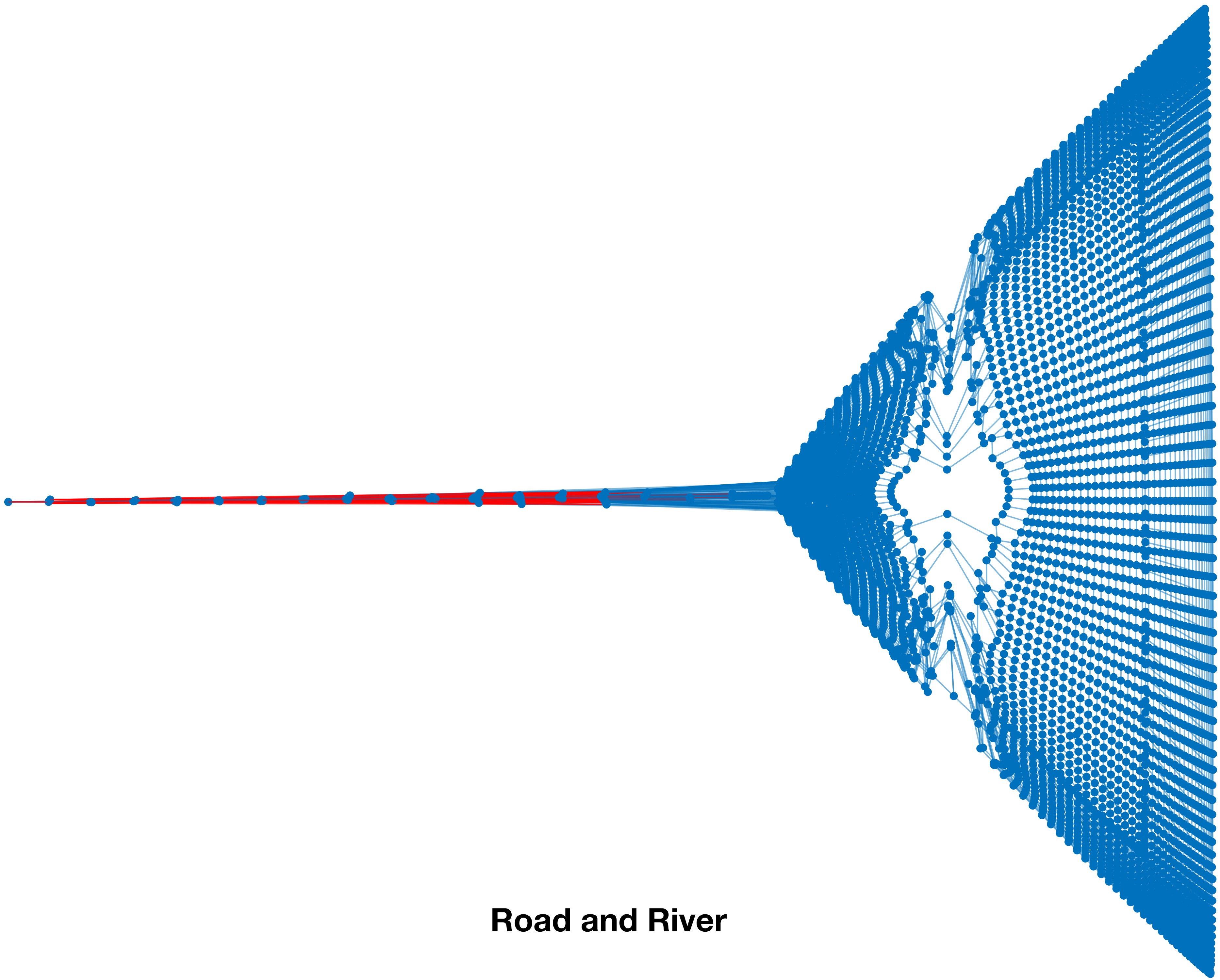
River: nodes in middle 2 columns randomly deleted

1500 BC

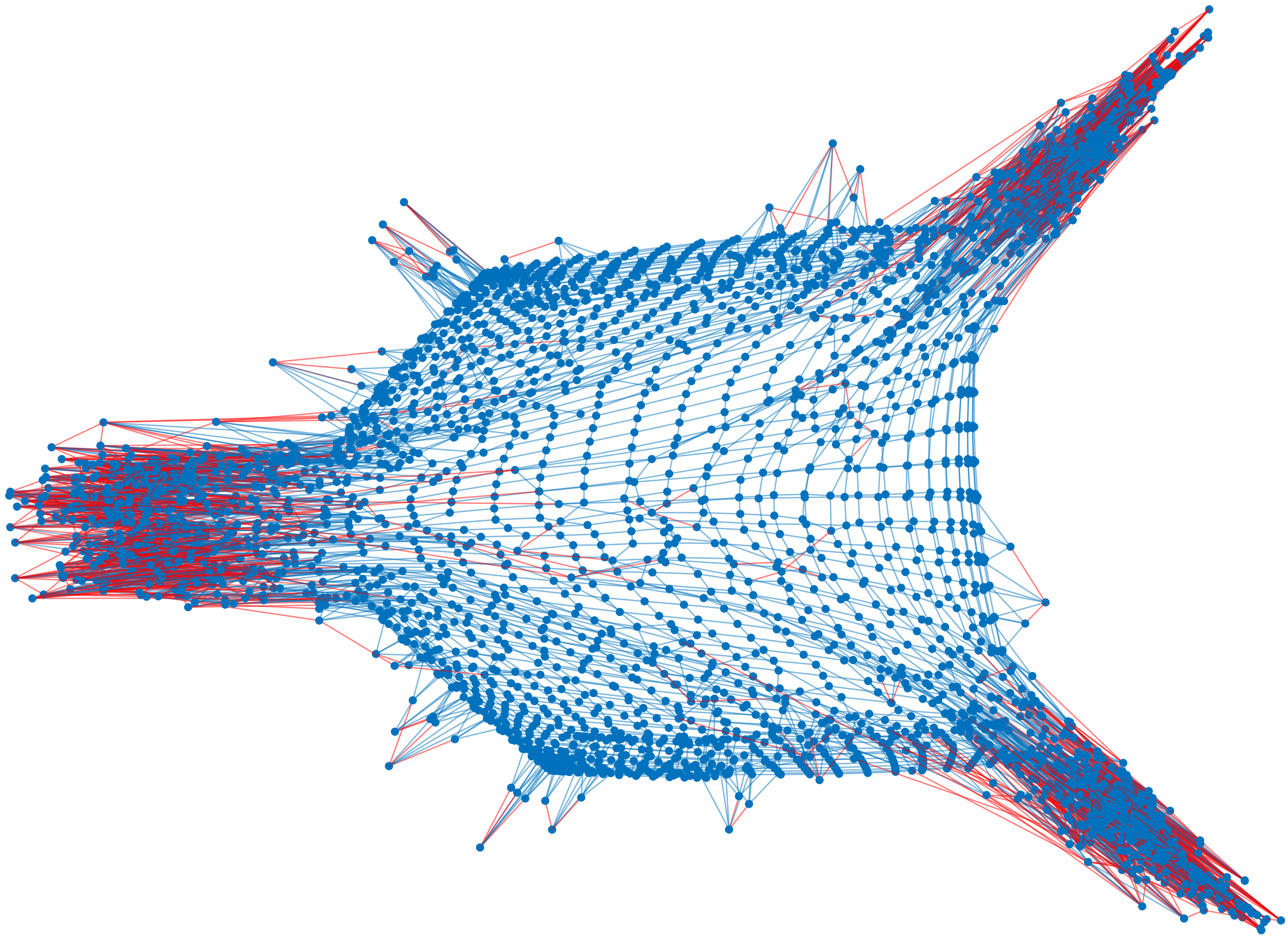




Road: Nodes along small strip randomly trade with each other
red links were added to grid



Road and River

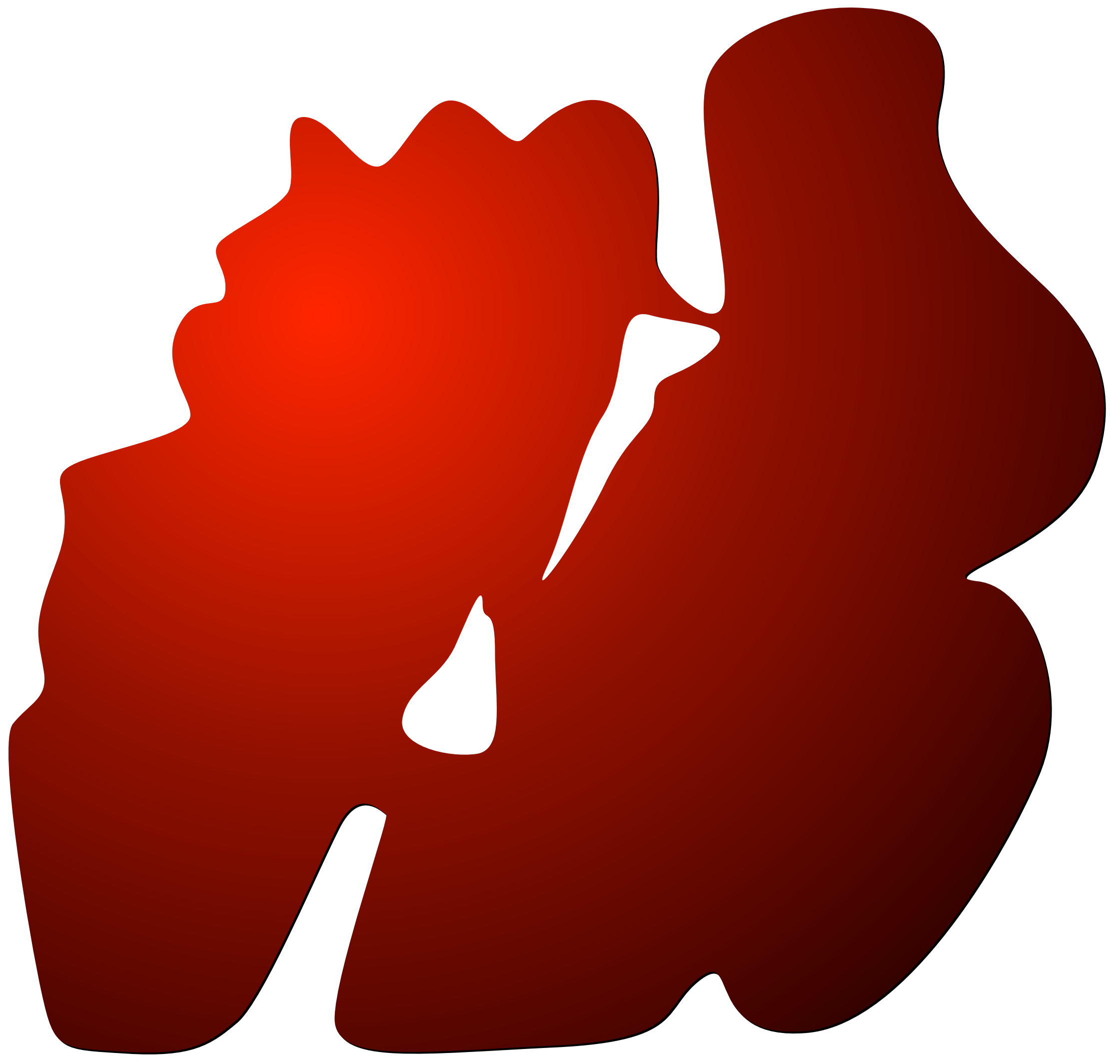


3 regions of increased activity added
many isolated nodes with increased activity randomly added
red links were added to grid

How does the algorithm actually work?

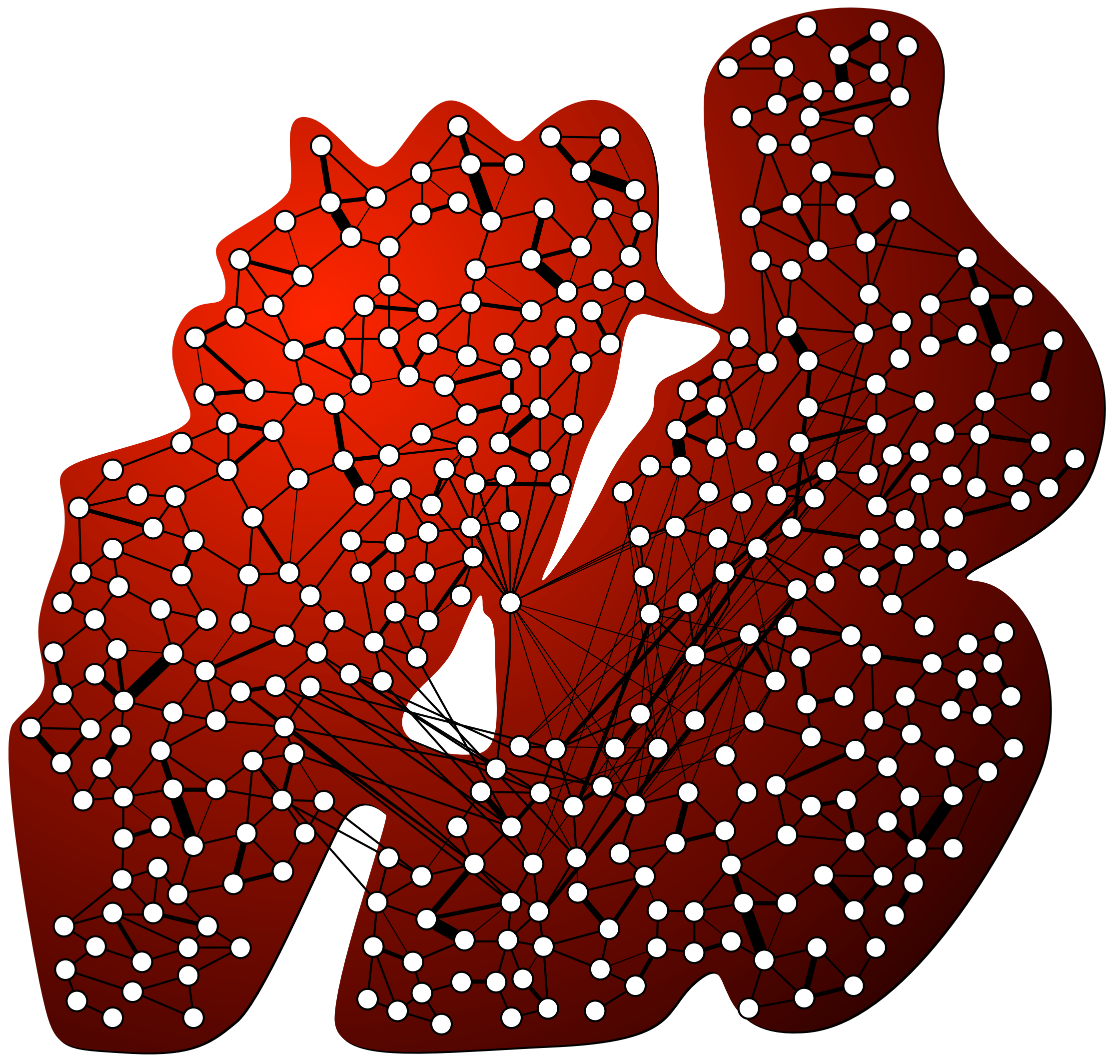
Network ~ piece of metal

Network ~ piece of metal



How well does it conduct heat?

Network ~ piece of metal

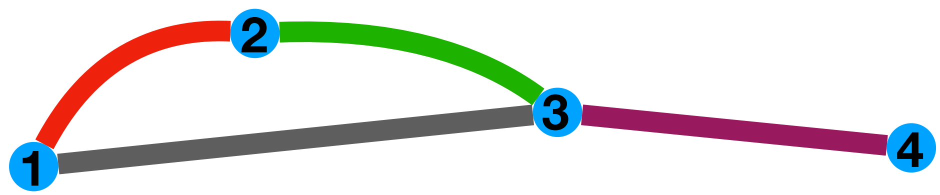


How well does it conduct information?

~

How well does it conduct heat?

Laplacian



$$L =$$

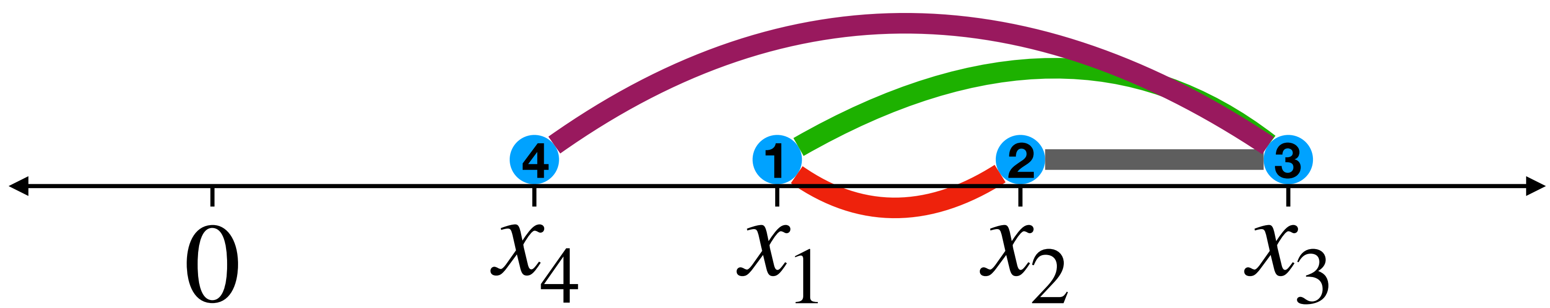
$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 2 & -1 & 0 \\ -1 & -1 & 3 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

Edge (i,j): $B_{i,i} = B_{j,j} = 1$

$B_{i,j} = B_{j,i} = -1$

Laplacian



$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

$$x^\intercal L x = (x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_1 - x_3)^2 + (x_3 - x_4)^2$$

Intuition: good drawings should *minimize* this!

Laplacian

$$x^\intercal L x = (x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_1 - x_3)^2 + (x_3 - x_4)^2 = 0$$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

Intuition: good drawings should *minimize* this!
Problem: the best solution is boring!
Solution: Find the minimum among all interesting solutions

Properties of Laplacian

Eigenvectors and Eigenvalues

Vector $v \neq 0$

Scalar λ

$$Lv = \lambda v \longrightarrow$$

v is eigenvector

λ is eigenvalue

Eigenbasis

Orthogonal Unit Eigenvectors $v_1, v_2, \ldots, v_n$

$$0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$$

Properties of Laplacian

Eigenbasis

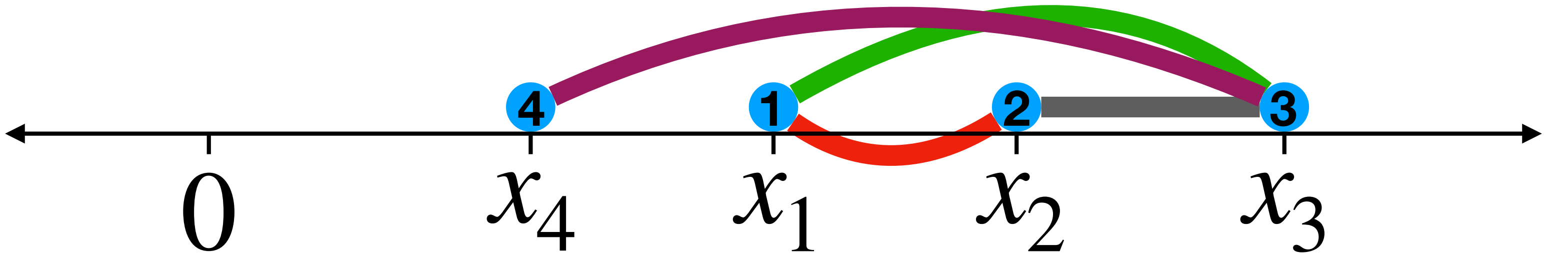
Orthogonal Unit Eigenvectors $v_1, v_2, \ldots, v_n$

$$0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$$

For every $x = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$

$$x^\intercal L x = \alpha_1^2 \lambda_1 + \alpha_2^2 \lambda_2 + \ldots + \alpha_n^2 \lambda_n$$

Laplacian



$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

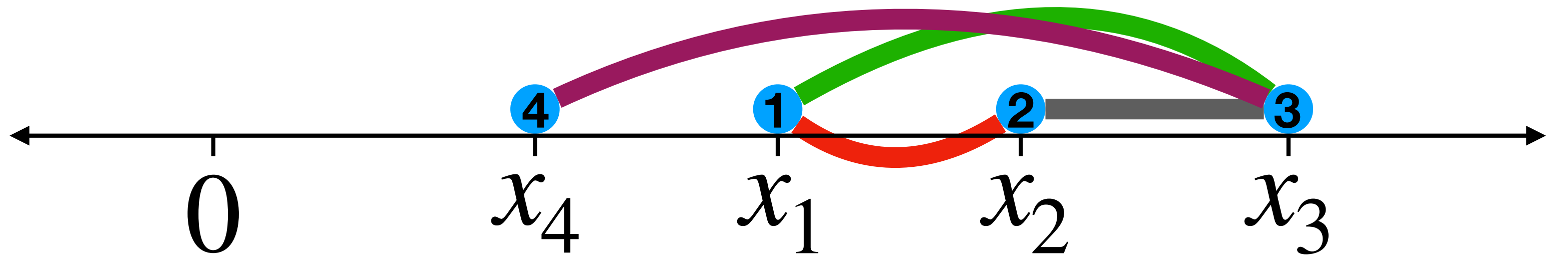
$$x = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$$

$$x^\intercal L x = (x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_1 - x_3)^2 + (x_3 - x_4)^2$$

$$= \alpha_1^2 \lambda_1 + \alpha_2^2 \lambda_2 + \ldots + \alpha_n^2 \lambda_n$$

Boring solution: $x = v_1$

Interesting solutions: $x = v_2, x = v_3, \ldots$



$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

$$x = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$$

$$x^\intercal L x = (x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_1 - x_3)^2 + (x_3 - x_4)^2$$

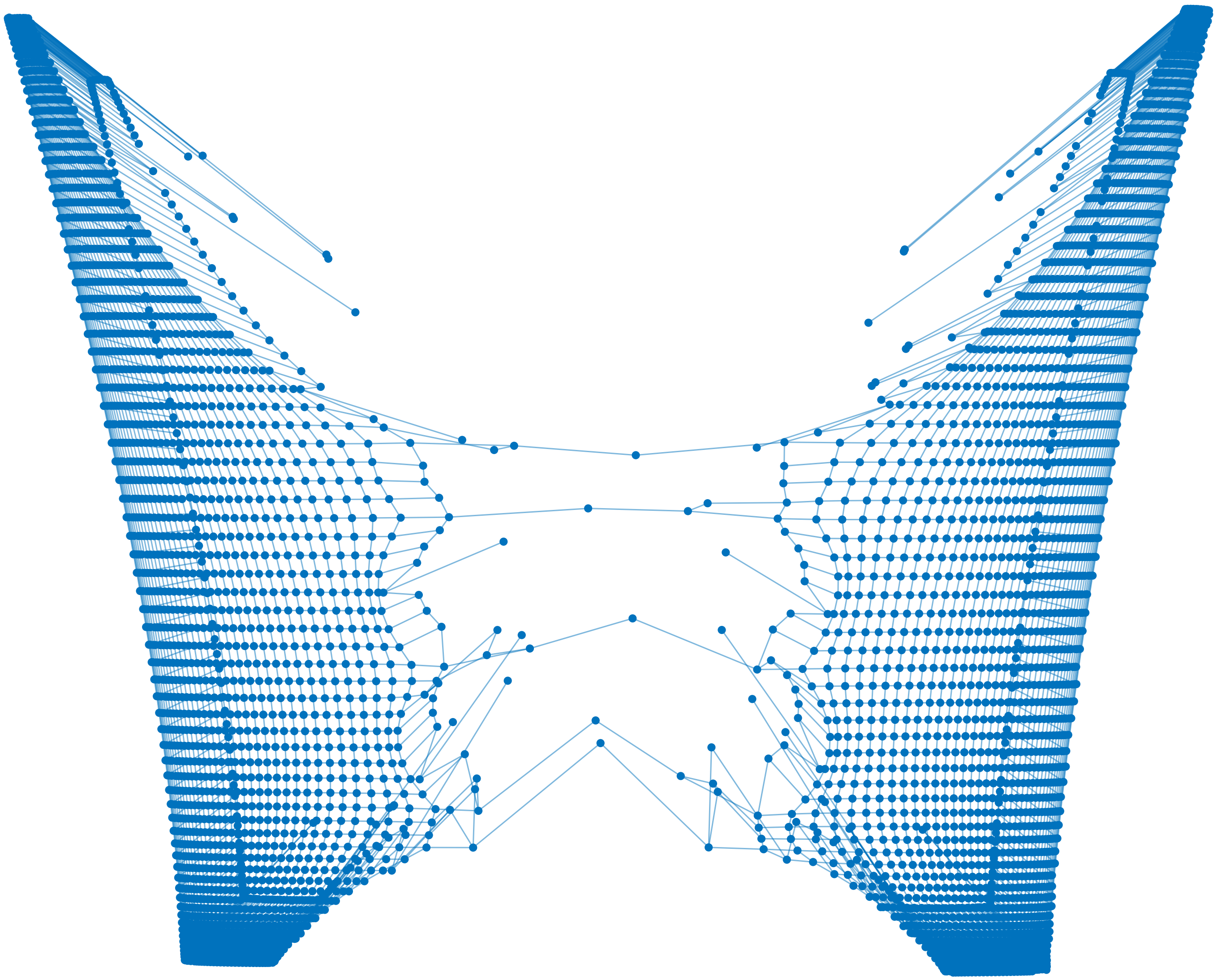
$$= \alpha_1^2 \lambda_1 + \alpha_2^2 \lambda_2 + \ldots + \alpha_n^2 \lambda_n$$

Boring solution: $x = v_1$

Interesting solutions: $x = v_2, x = v_3, \ldots$

$v_2 =$ choice of x minimizing $x^\intercal L x$ among all choices orthogonal to v_1

$v_3 =$ choice of x minimizing $x^\intercal L x$ among all choices orthogonal to v_1, v_2



v_2, v_3

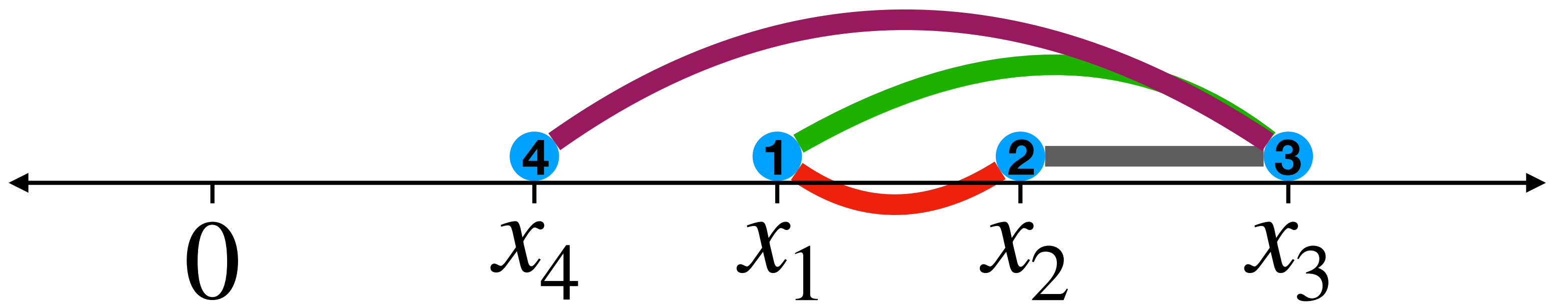
eigenvectors of laplacian

Place node i at location

$(v_{2,i}, v_{3,i})$

—this is not
how the picture was drawn—

Normalized Laplacian



$$
\begin{aligned}
x^\intercal \mathscr{L} x = {} & (x_1/\sqrt{2} - x_2/\sqrt{2})^2 + \\
& (x_1/\sqrt{2} - x_3/\sqrt{3})^2 + \\
& (x_2/\sqrt{2} - x_3/\sqrt{2})^2 + \\
& (x_3/\sqrt{3} - x_4)^2
\end{aligned}
$$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

Intuition: pay more attention to nodes with many links

Properties of Normalized Laplacian

Eigenvectors and Eigenvalues

Vector $v \neq 0$

Scalar λ

$$Lv = \lambda v \longrightarrow$$

v is eigenvector

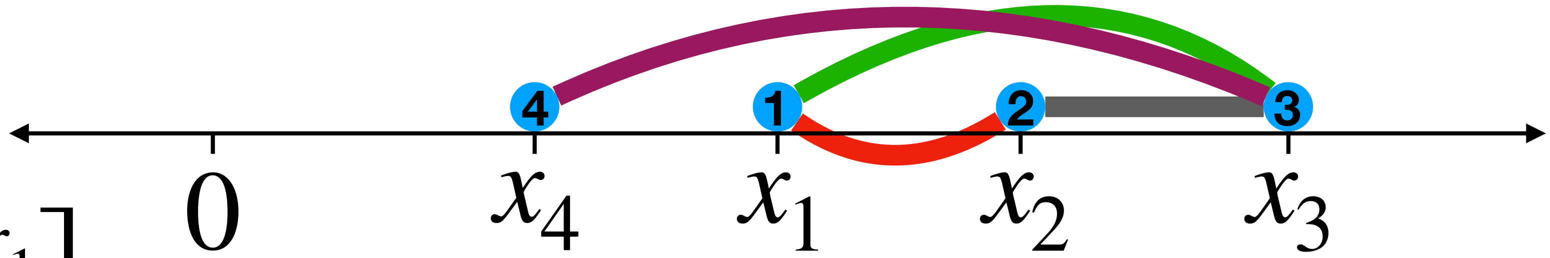
λ is eigenvalue

Eigenbasis

Orthogonal Unit Eigenvectors $v_1, v_2, \ldots, v_n$

$$0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n \leq 2$$

Normalized Laplacian



$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

$$x = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$$

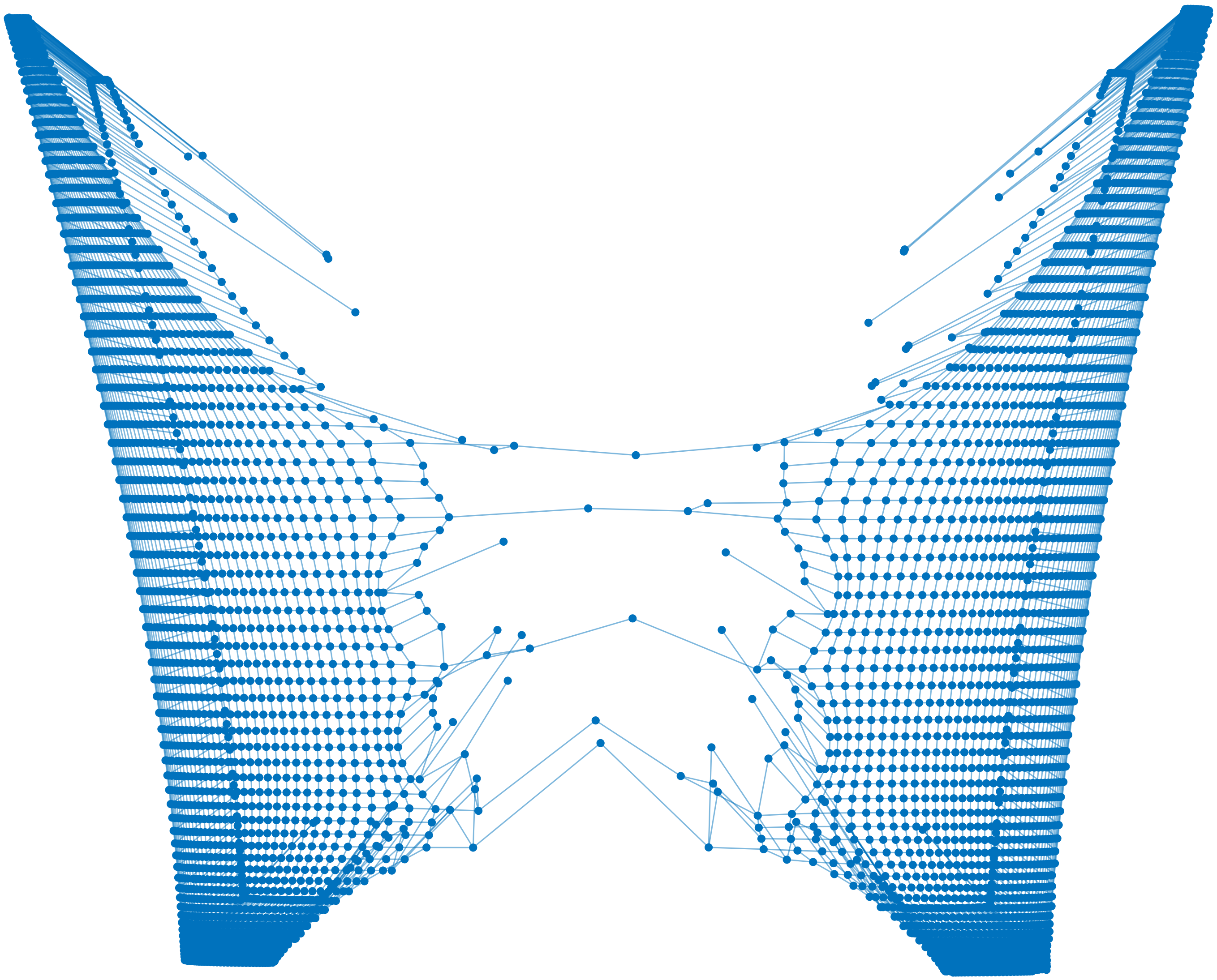
$$x^\intercal \mathscr{L} x = \alpha_1^2 \lambda_1 + \alpha_2^2 \lambda_2 + \ldots + \alpha_n^2 \lambda_n$$

Boring solution: $x = v_1$

$v_2 =$ choice of x minimizing $x^\intercal L x$ among all choices skew orthogonal to v_1

$v_3 =$ choice of x minimizing $x^\intercal L x$ among all choices skew orthogonal to v_1, v_2

λ_2 is a measure of the degree of connectivity of the whole network



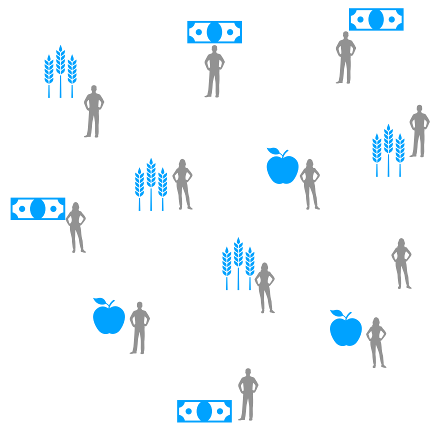
v_2, v_3

eigenvectors of normalized laplacian

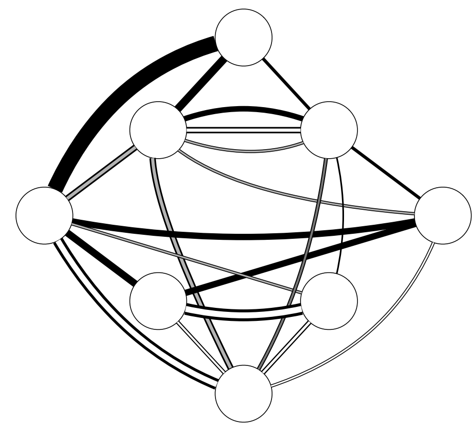
Place node i at location $(v_{2,i}, v_{3,i})$

actual drawing

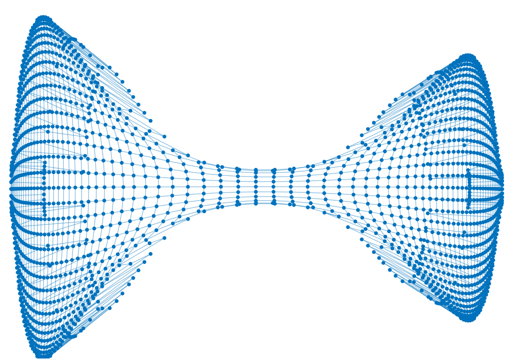
Outline



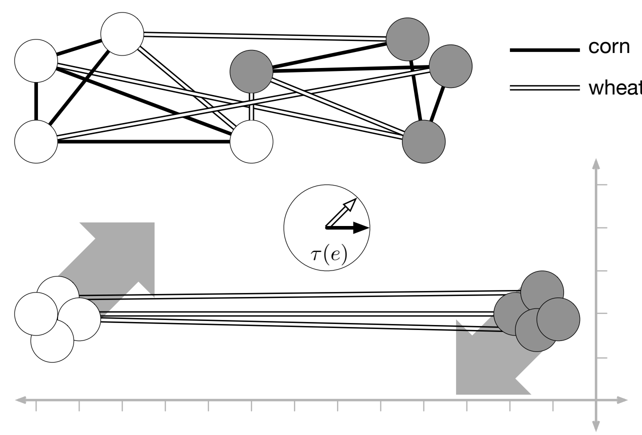
1. CS and Economics



2. Trade Networks



3. Graph Laplacians



4. Modifications for Econ

Questions

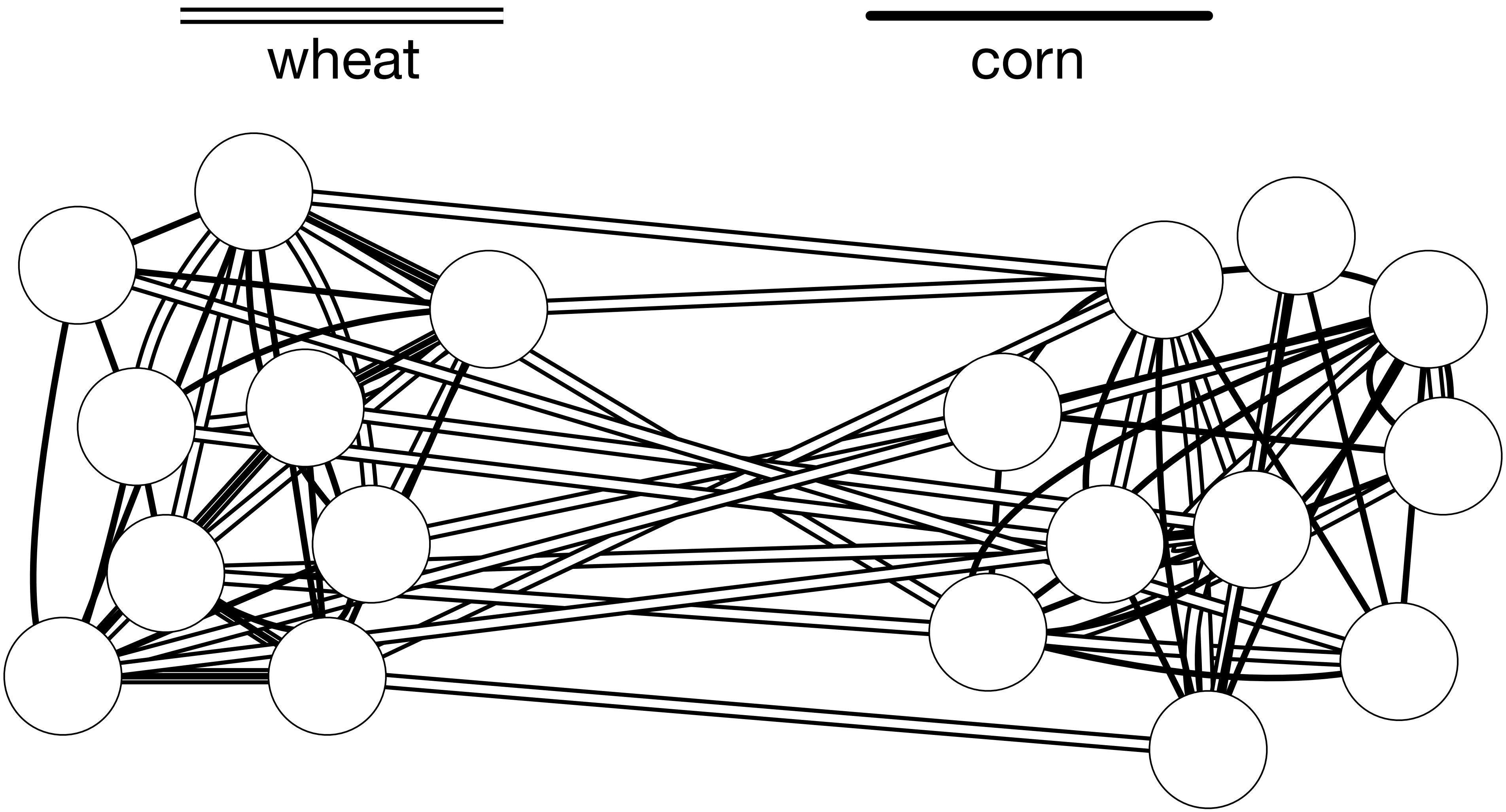
- Laplacian? Normalized Laplacian? Something else? What is the right notion to use to determine shape of Trade Network?

- What is the formulation that gives the most meaningful solutions from the perspective of economics?

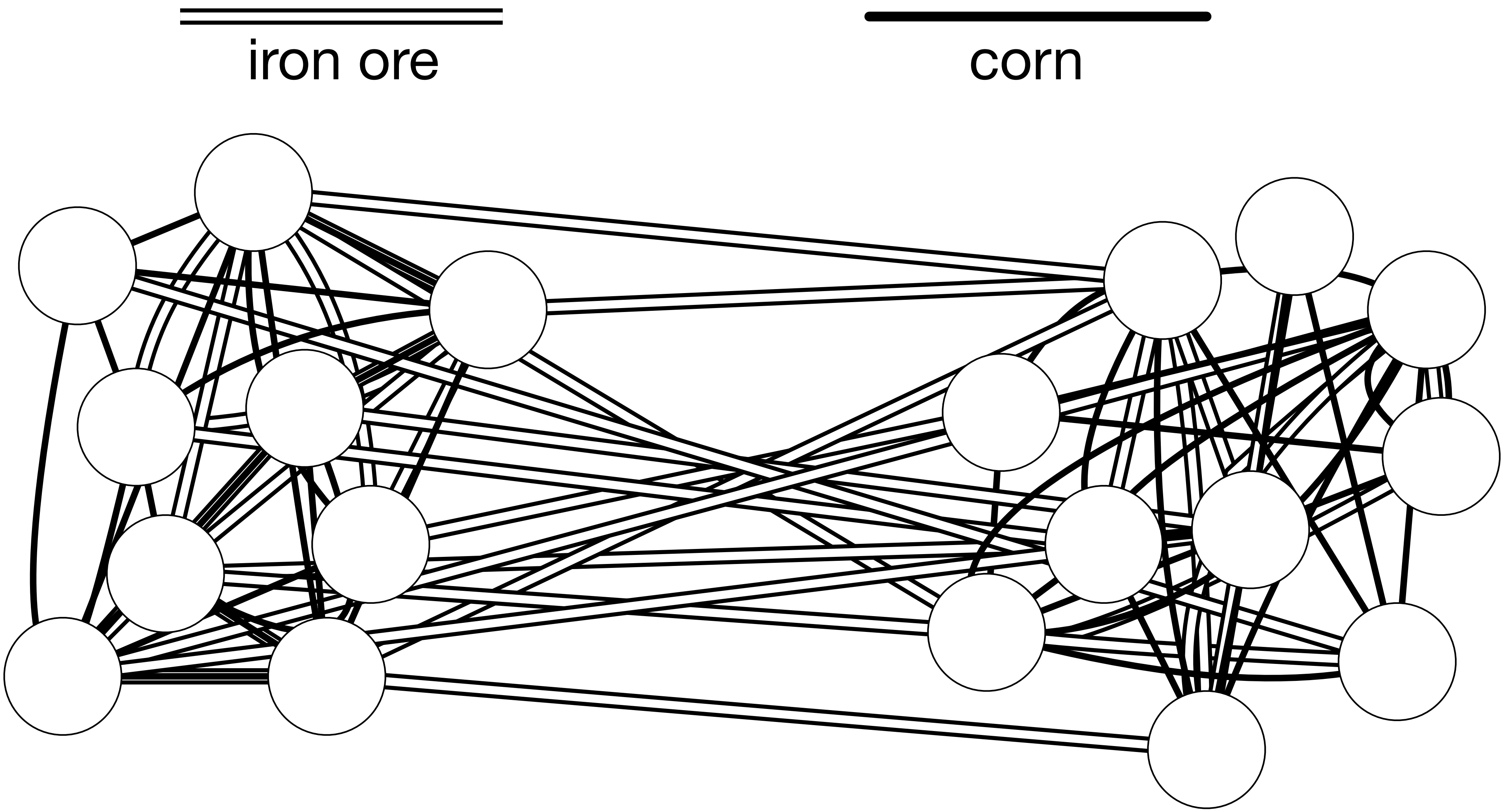
- For eg: so far the model does not distinguish type of item being bought/sold

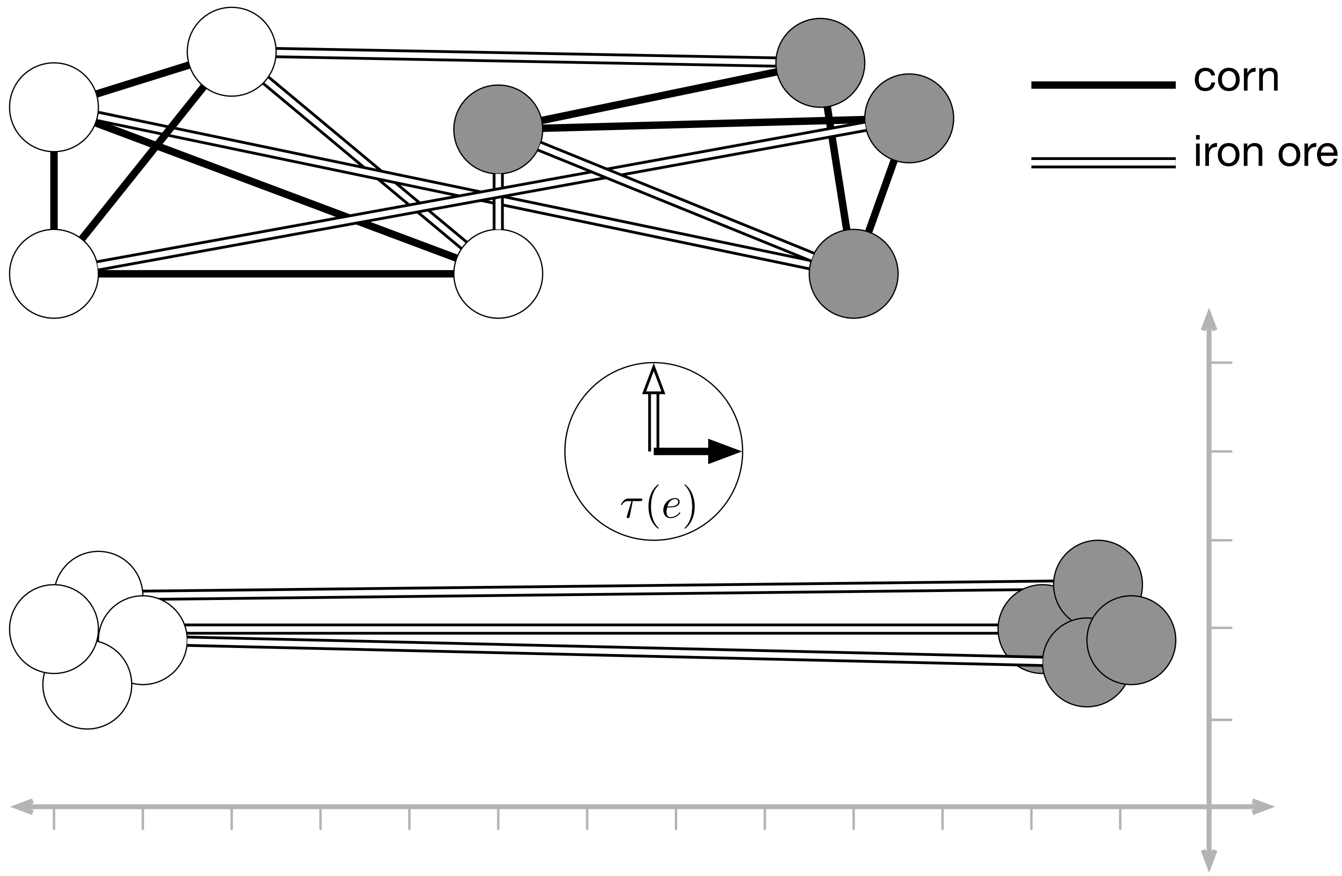
- **My paper: find definitions leading to most meaningful space of solutions for economics**

Should these two clusters be far apart or close together?

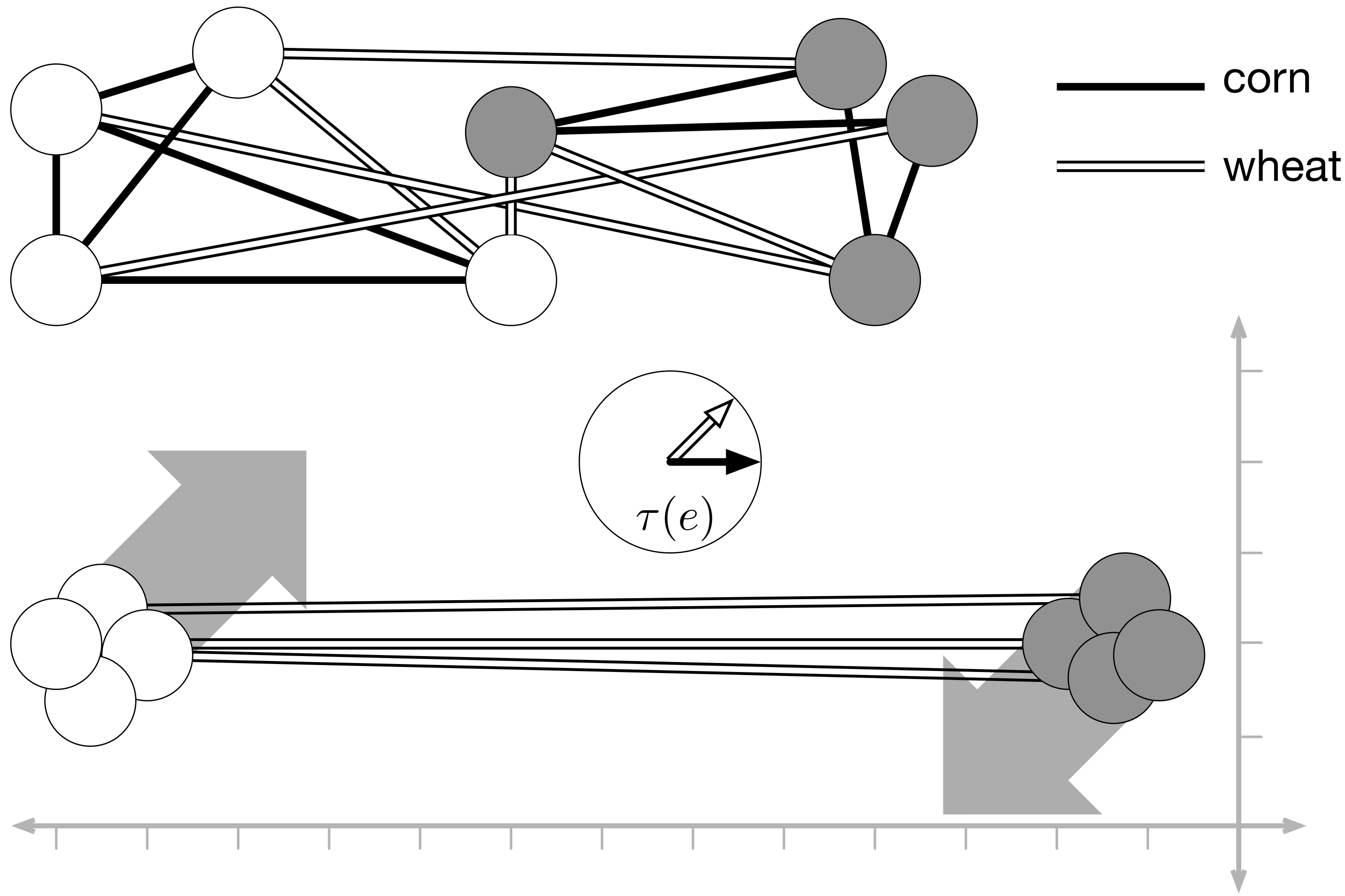


Should these two clusters be far apart or close together?



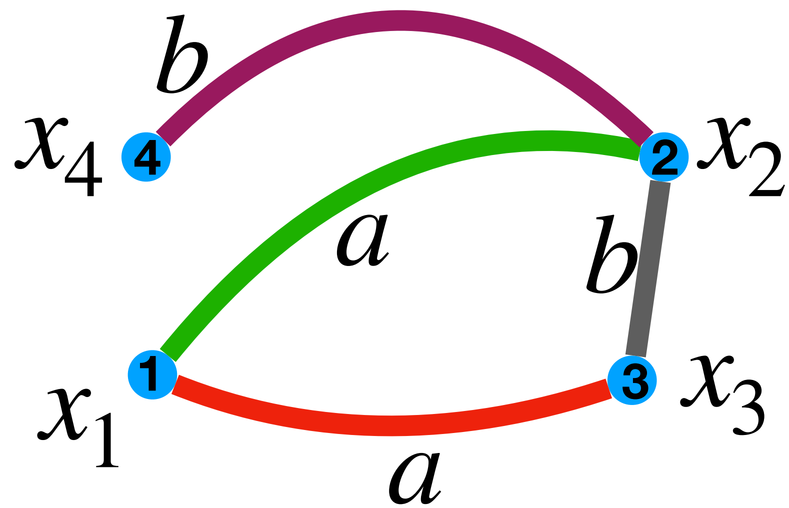


- Nodes are placed on plane instead of the line
- Each item is associated with a direction
- Each link pulls along the corresponding direction



- Nodes are placed on plane instead of the line
- Each item is associated with a direction
- Each link pulls along the corresponding direction

Typed Laplacians



$$
\begin{aligned}
x^\intercal \tilde{L} x = \; & ((x_1 - x_2) \cdot \tau_a)^2 + \\
& ((x_2 - x_3) \cdot \tau_b)^2 + \\
& ((x_1 - x_3) \cdot \tau_a)^2 + \\
& ((x_3 - x_4) \cdot \tau_b)^2
\end{aligned}
$$

$$
x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}
$$

Intuition: important large scale features come from minimizing this!

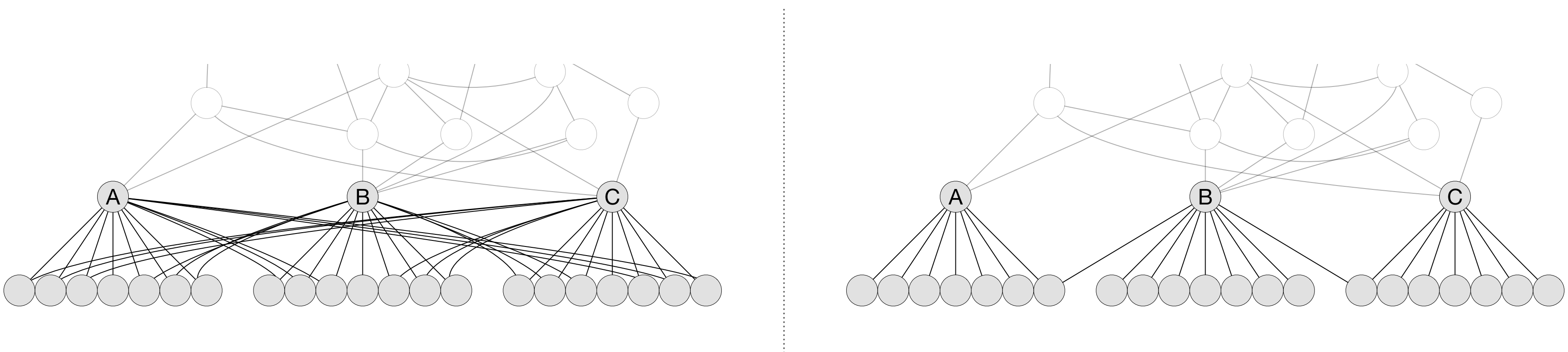
Questions I answered

- What are the boring eigenvectors/eigenvalues?
- How to compute the type vectors to get good results?
  My guess: Use data from the network itself! Items traded by similar nodes have similar type.

Using Trade Networks to pick out features relevant to economics

- New understanding of well-known phenomena!
- Examples: Monopolies, exchange traded funds, high switching costs, vertical integration…

Does market-share determine Monopoly?



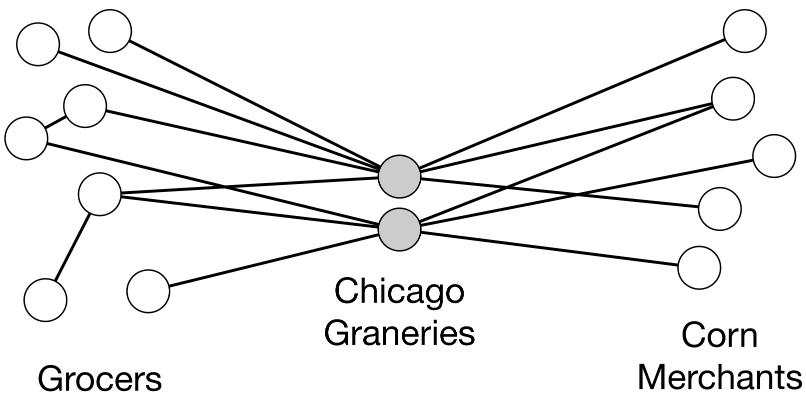
2 networks of cell phone providers and customers.

- What kind of network do we want to see?
- How can we change regulations to make a better network?

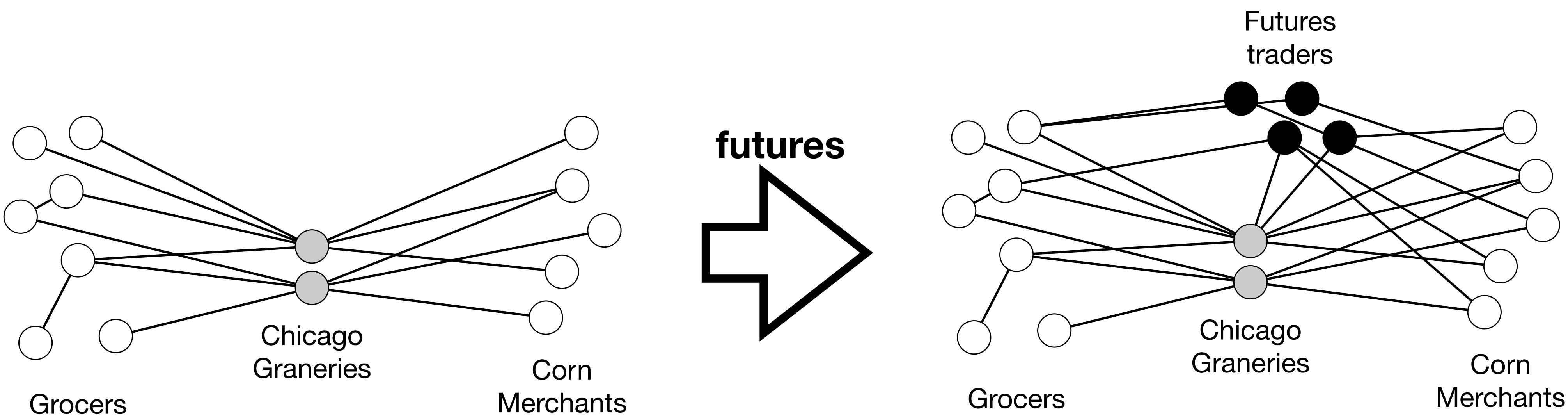
How futures break monopolies [Heironymous]

- Chicago in the 1850s
- Corn merchants would buy corn from farmers,
  sell to granaries,
  who sell to grocers
- Granaries are monopolies
- Corn merchants need capital.
  Futures contracts—contract for delivery of corn in future
- Other residents of Chicago buy and sell futures contracts
- Eliminated ability of granaries to fix prices, even though flow of corn is the same!

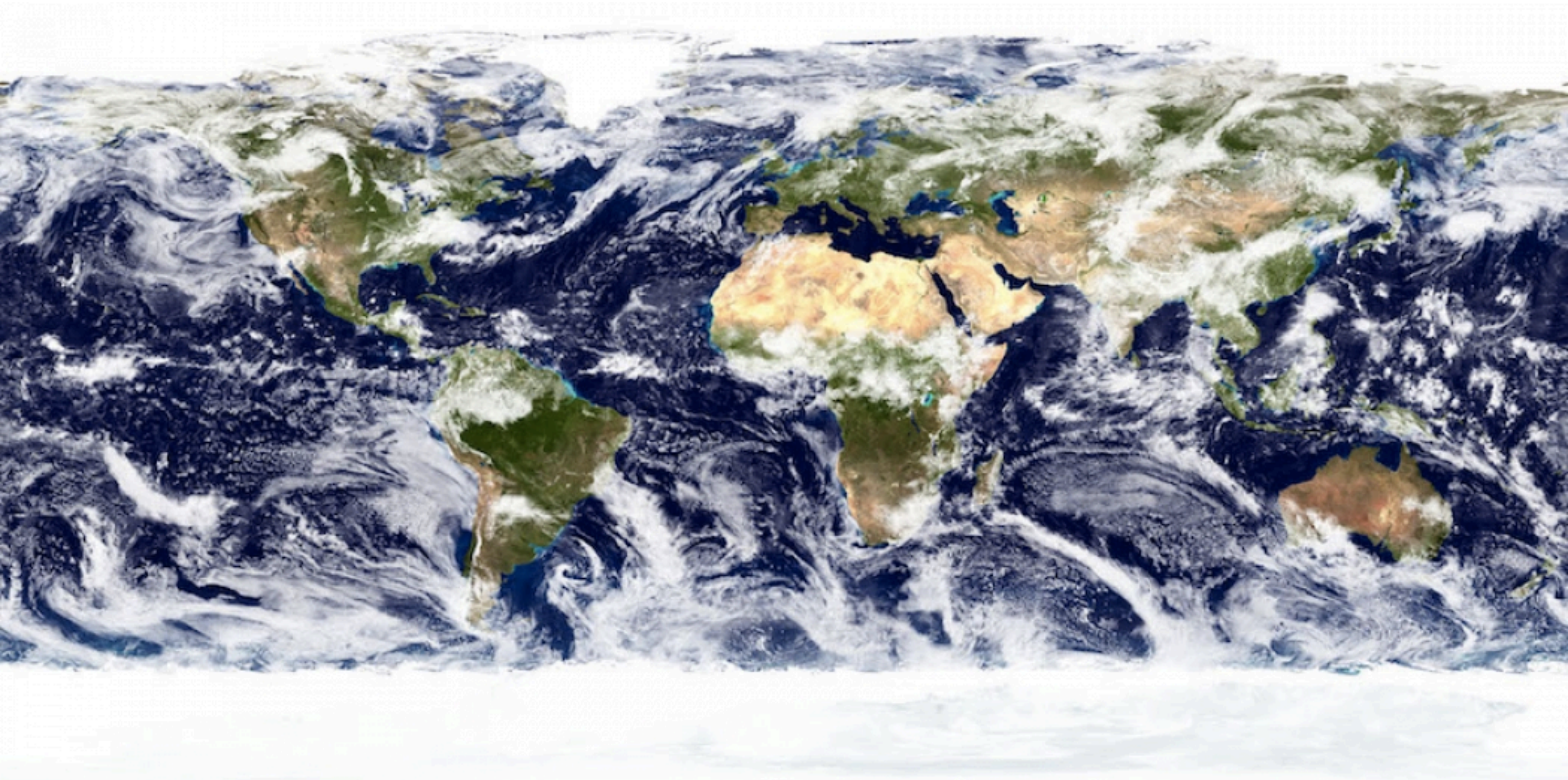
Futures markets and monopolies



Futures markets and monopolies



- Futures market broke monopoly, even though flow of goods is the same!



“Clouds are frozen over the great plains, but the whole world is not hidden.”

–The Blue Cliff Record