Communities in statistically-validated networks

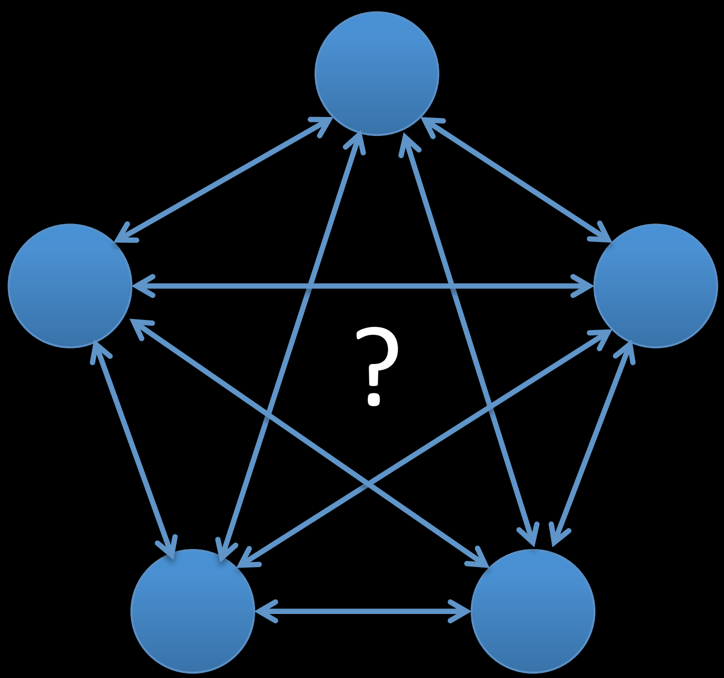
Chester Curme

In collaboration with:

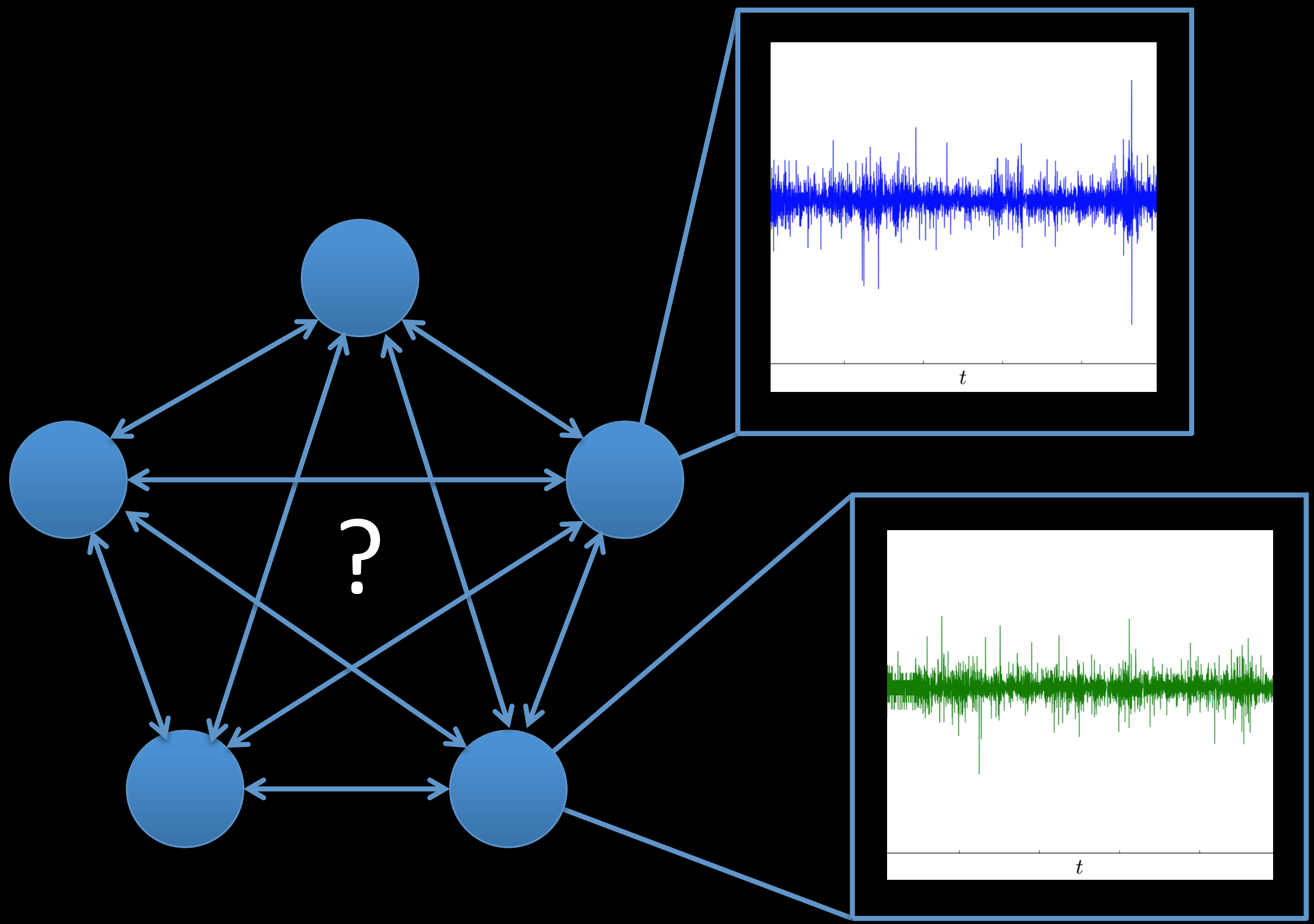
Irena Vodenska

H. Eugene Stanley

Big picture

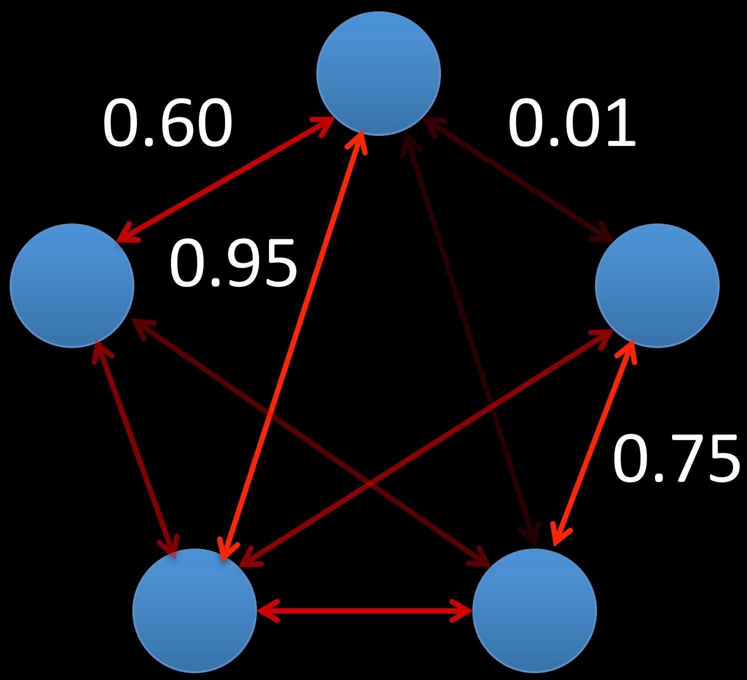


Big picture

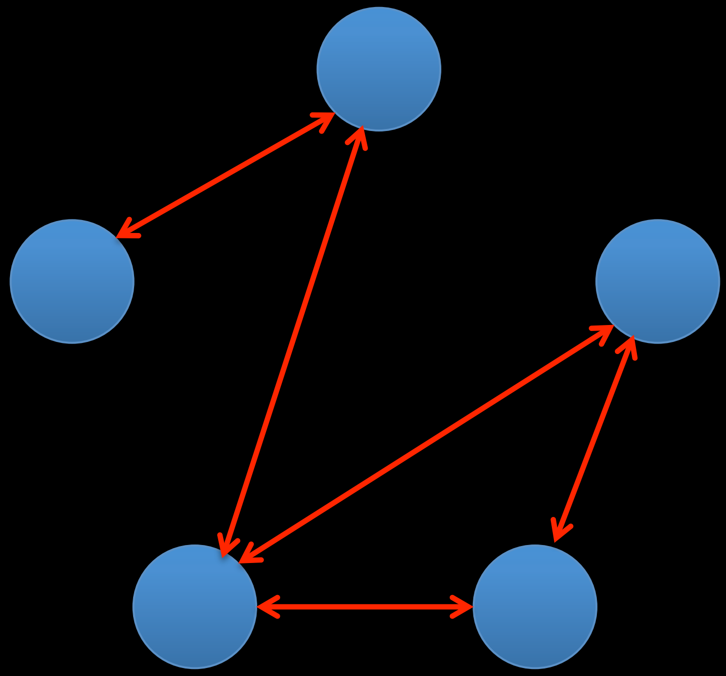


Big picture

$$A = \begin{bmatrix} 1 & 0.60 & 0.01 & 0.95 & \ldots \\ 0.60 & 1 & \ldots & \ldots & \ldots \\ 0.01 & \ldots & 1 & \ldots & 0.75 \\ 0.95 & \ldots & \ldots & 1 & \ldots \\ \ldots & \ldots & 0.75 & \ldots & 1 \end{bmatrix}$$



Big picture



$$A = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

Thesis overview

Focus on one experimental setting:

- Associate time series to nodes

Thesis overview

Focus on one experimental setting:

- Associate time series to nodes
- Interactions measured by lagged correlations
  - A is asymmetric

Thesis overview

Focus on one experimental setting:

- Associate time series to nodes
- Interactions measured by lagged correlations
  - A is asymmetric
- Low signal-to-noise
  - Statistical uncertainty is important

Thesis overview

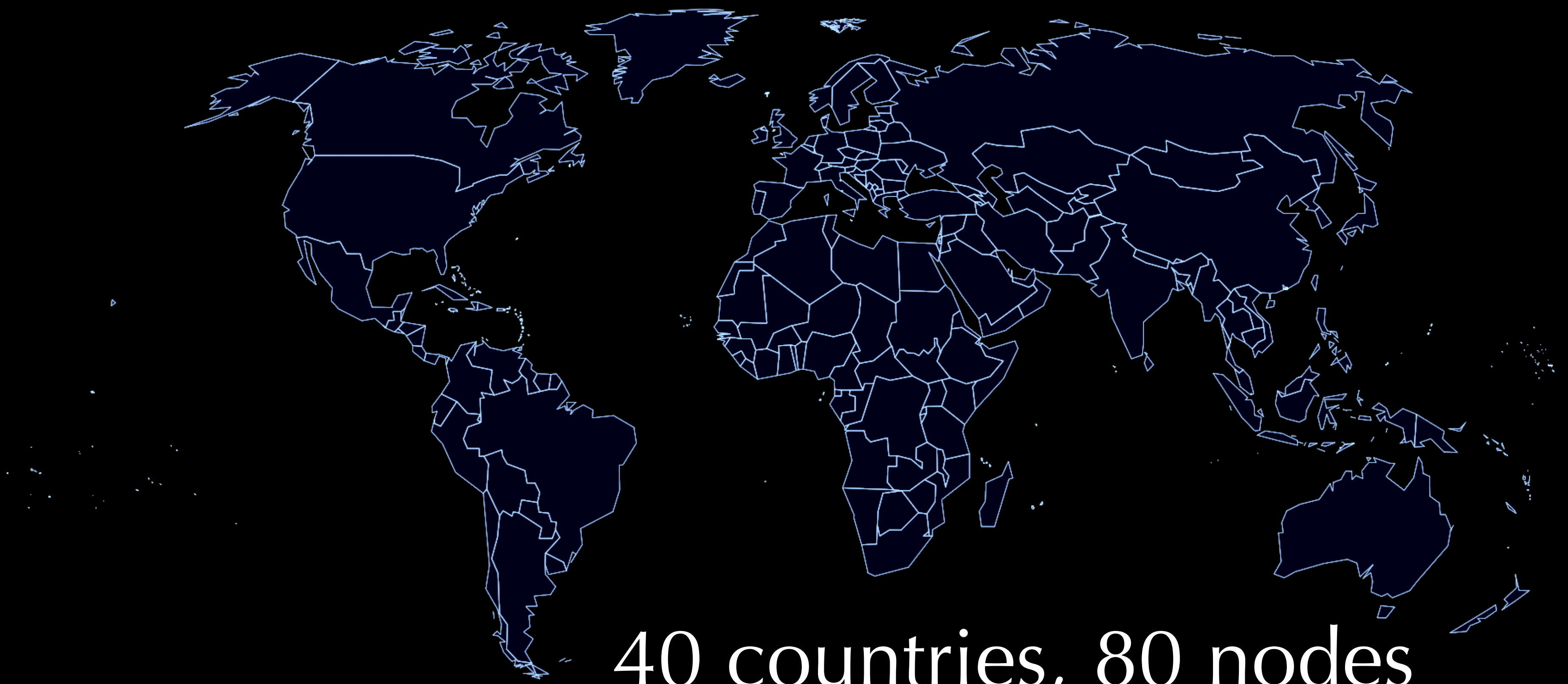
1) Methodology and applications
   - Curme, Tumminello, Mantegna, Stanley, Kenett. *Quantitative Finance* (2014).
2) Extend to seasonal time series and use to explain a phenomenon in financial markets
   - Curme, Tumminello, Mantegna, Stanley, Kenett (in preparation).
3) Relate community structures to statistical model performance
   - Curme, Vodenska, Stanley (submitted).
4) Using topic models to explain market movements
   - Curme, Zhuo, Moat, Preis (in preparation).
   - Curme, Preis, Stanley, Moat. *PNAS* (2014).
   - Moat, Curme, Avakian, Kenett, Stanley, Preis. *Scientific Reports* (2013).
   - Moat, Curme, Stanley, Preis. Book chapter in *Nonlinear phenomena in complex systems: from nano to macro scale* (2014).

Thesis overview

1) Methodology and applications
   - Curme, Tumminello, Mantegna, Stanley, Kenett. *Quantitative Finance* (2014).
2) Extend to seasonal time series and use to explain a phenomenon in financial markets
   - Curme, Tumminello, Mantegna, Stanley, Kenett (submitted).
3) Relate community structures to statistical model performance
   - Curme, Vodenska, Stanley (submitted).
4) Using topic models to explain market movements
   - Curme, Zhuo, Moat, Preis (submitted).
   - Curme, Preis, Stanley, Moat. *PNAS* (2014).
   - Moat, Curme, Avakian, Kenett, Stanley, Preis. *Scientific Reports* (2013).
   - Moat, Curme, Stanley, Preis. Book chapter in *Nonlinear phenomena in complex systems: from nano to macro scale* (2014).

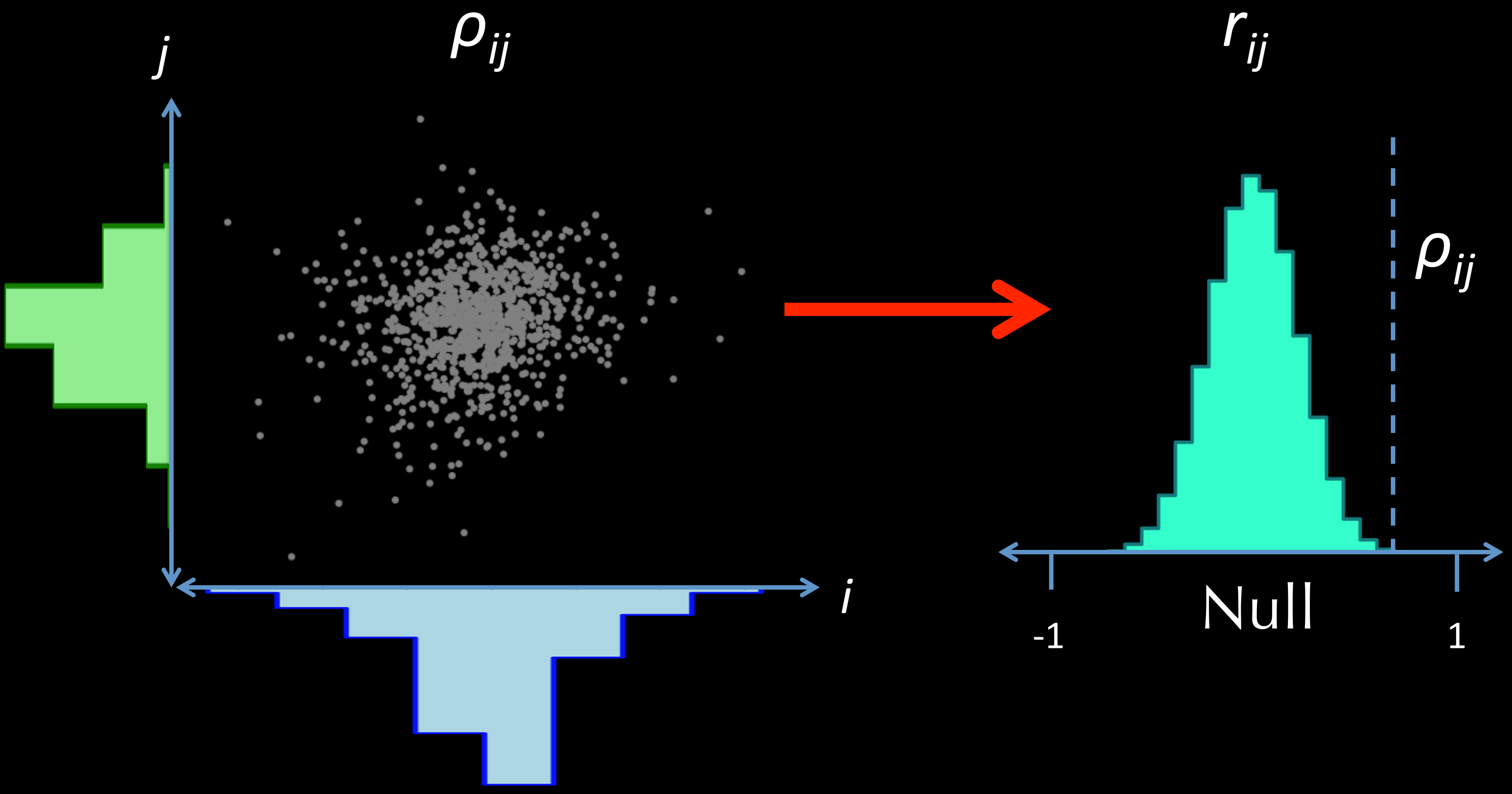
Data

Market returns
News sentiments



40 countries, 80 nodes

Statistical validation



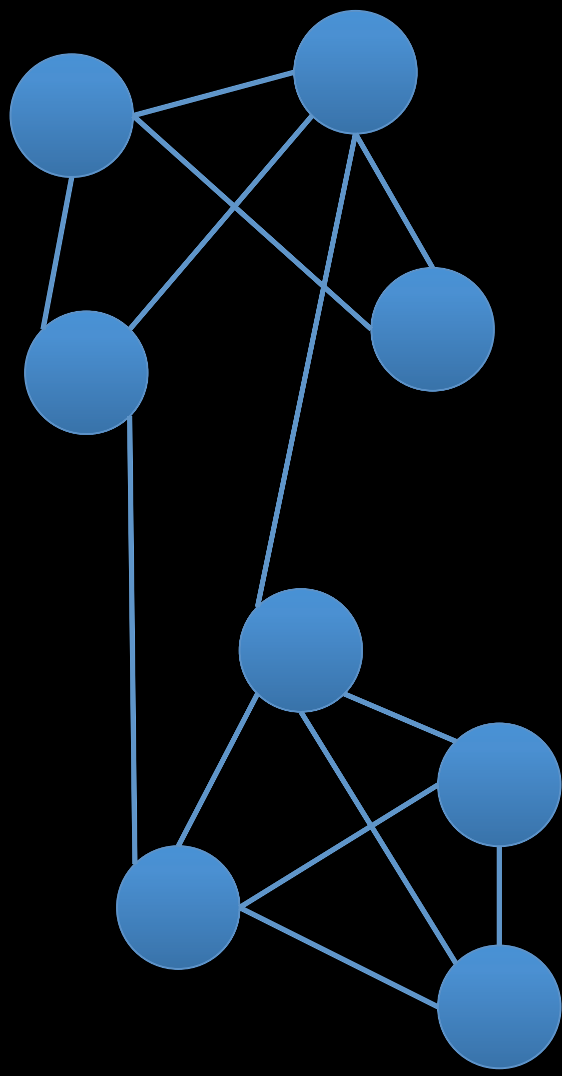
Statistically-validated network

Statistically-validated network

News → Markets: 12 links

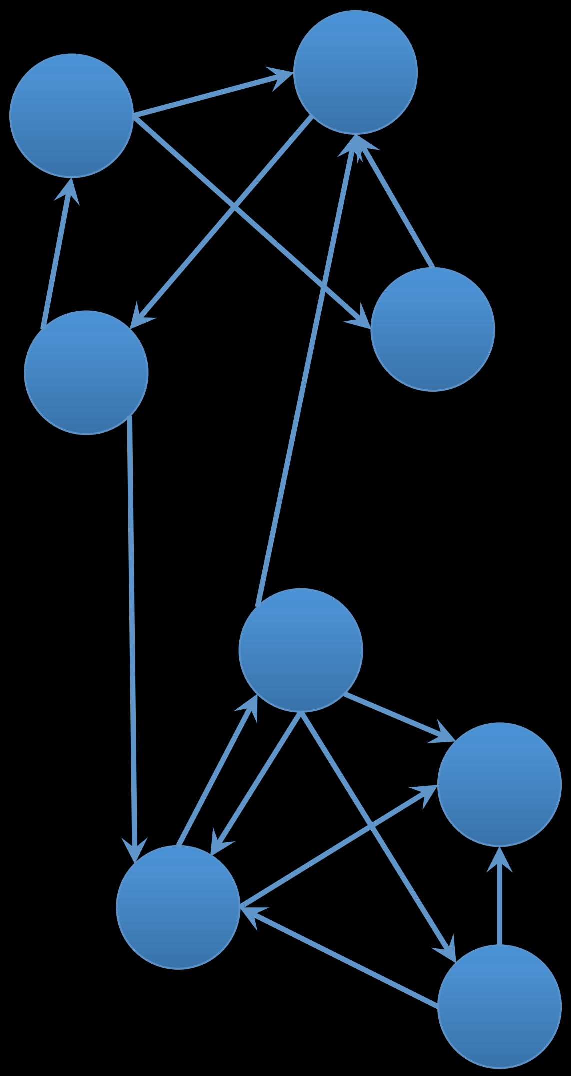
Markets → News: 174 links

Community Structures



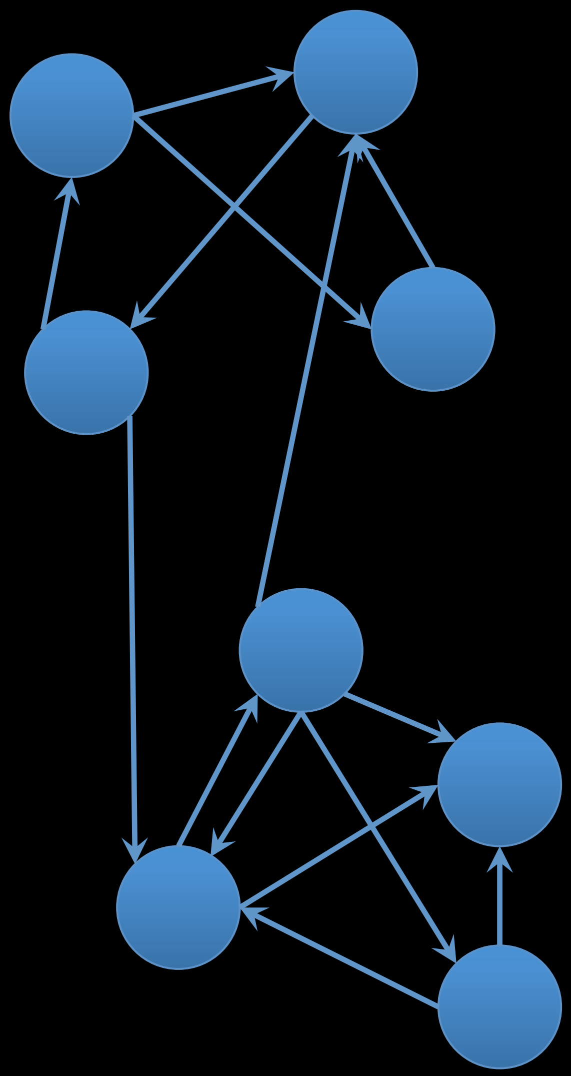
In undirected networks...

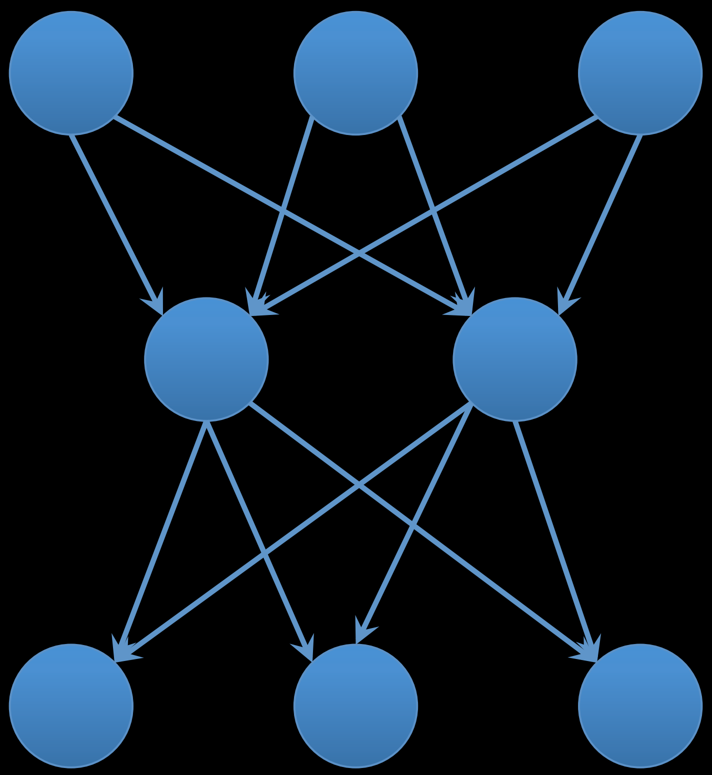
Community Structures



In directed networks…

Community Structures



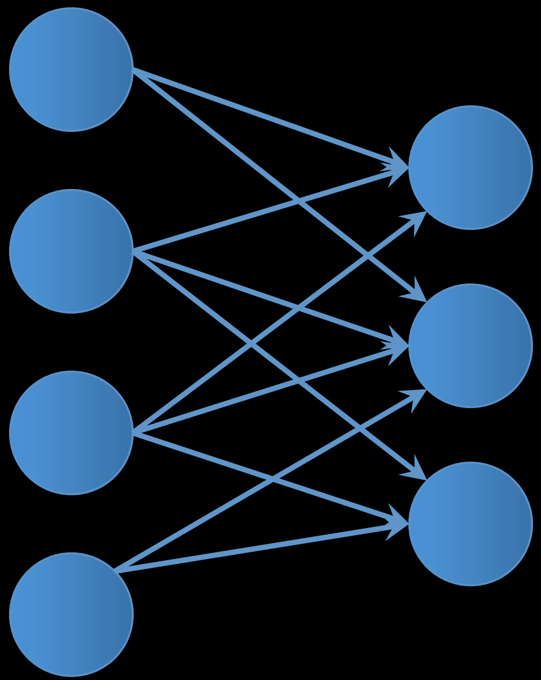


In directed networks...

Finding bipartite substructure

Singular value decomposition:

$$A = U\,\Sigma\,V^T$$



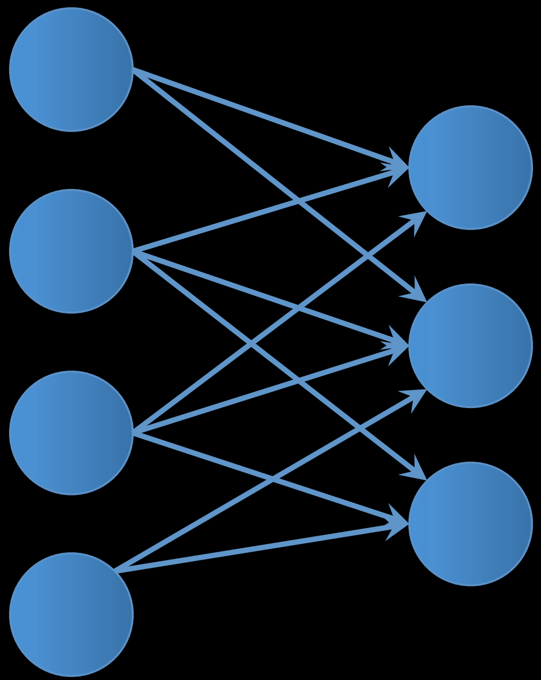
Finding bipartite substructure

Singular value decomposition:

$$A = U \, \Sigma \, V^T$$

$$\begin{bmatrix} \text{Eigenvectors} \\ \text{of } AA^T \\ \\ \end{bmatrix}$$

"Left singular vectors"



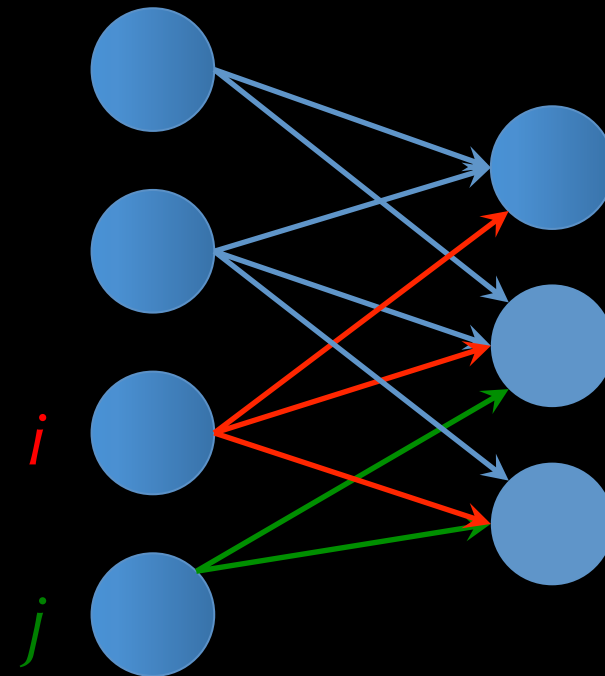
Finding bipartite substructure

Singular value decomposition:

$$A = U\,\Sigma\,V^T$$

$$\begin{bmatrix} \text{Eigenvectors} \\ \text{of } AA^T \end{bmatrix}$$

"Left singular vectors"

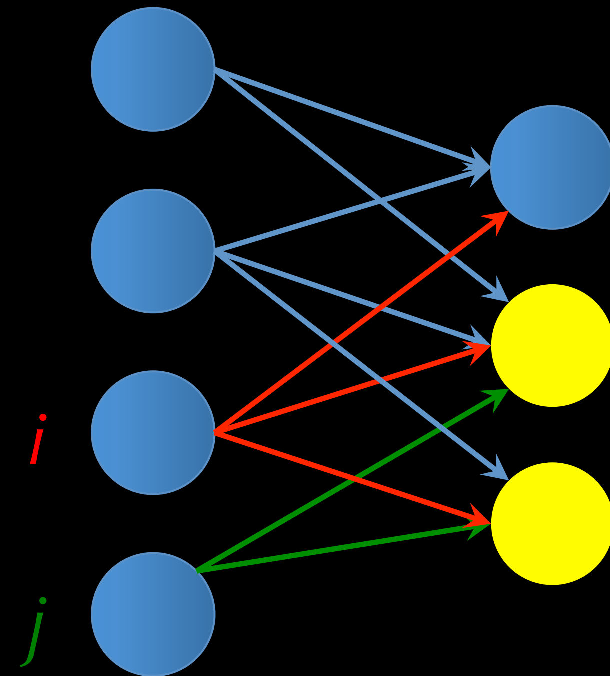


$$AA^T = \begin{array}{c} i \\ \\ j \end{array}\begin{bmatrix} \ldots\ 1,\ 1,\ 1,\ \ldots \\ \\ \ldots\ 0,\ 1,\ 1,\ \ldots \end{bmatrix}\begin{array}{c} \begin{array}{cc} i & \quad j \end{array} \\ \begin{bmatrix} \vdots & \quad \vdots \\ 1, & \quad 0, \\ 1, & \quad 1, \\ 1, & \quad 1, \\ \vdots & \quad \vdots \end{bmatrix} \end{array}$$

Finding bipartite substructure

Singular value decomposition:

$$A = U \, \Sigma \, V^T$$



$$\begin{bmatrix} \text{Eigenvectors} \\ \text{of } AA^T \end{bmatrix}$$

"Left singular vectors"

$$AA^T = \begin{array}{cc} & \begin{matrix} i & \quad j \end{matrix} \\ \begin{matrix} i \\ \\ j \end{matrix} & \begin{bmatrix} & \ldots, 2, \ldots \\ & \\ \ldots, 2, \ldots & \end{bmatrix} \end{array}$$

Finding bipartite substructure

Singular value decomposition:

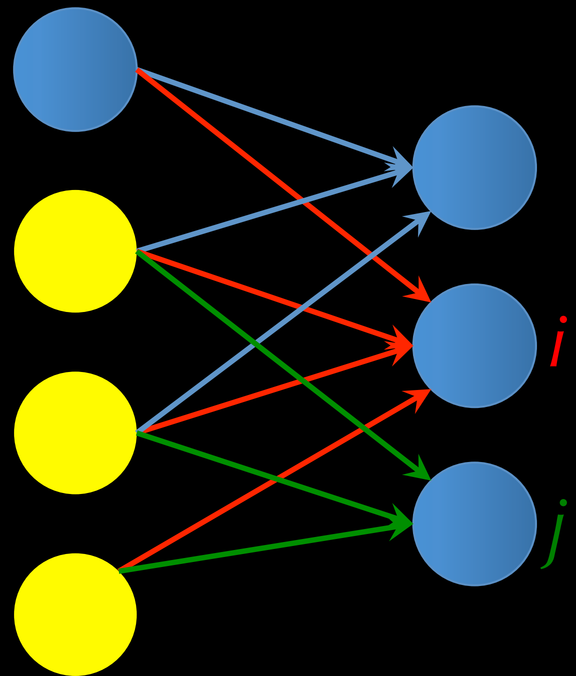
$$A = U \, \Sigma \, V^T$$

U: $\begin{bmatrix} \text{Eigenvectors} \\ \text{of } AA^T \\ \\ \end{bmatrix}$

"Left singular vectors"

V: $\begin{bmatrix} \text{Eigenvectors} \\ \text{of } A^TA \\ \\ \end{bmatrix}$

"Right singular vectors"



Finding bipartite substructure

Singular value decomposition:

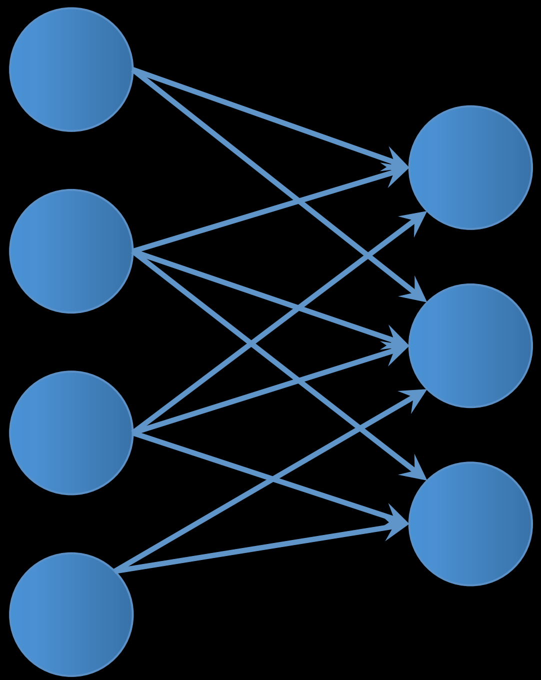
$$A = U \, \Sigma \, V^T$$

$$\begin{bmatrix} \text{Eigenvectors} \\ \text{of } AA^T \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ldots \end{bmatrix} \begin{bmatrix} \text{Eigenvectors} \\ \text{of } A^TA \end{bmatrix}$$

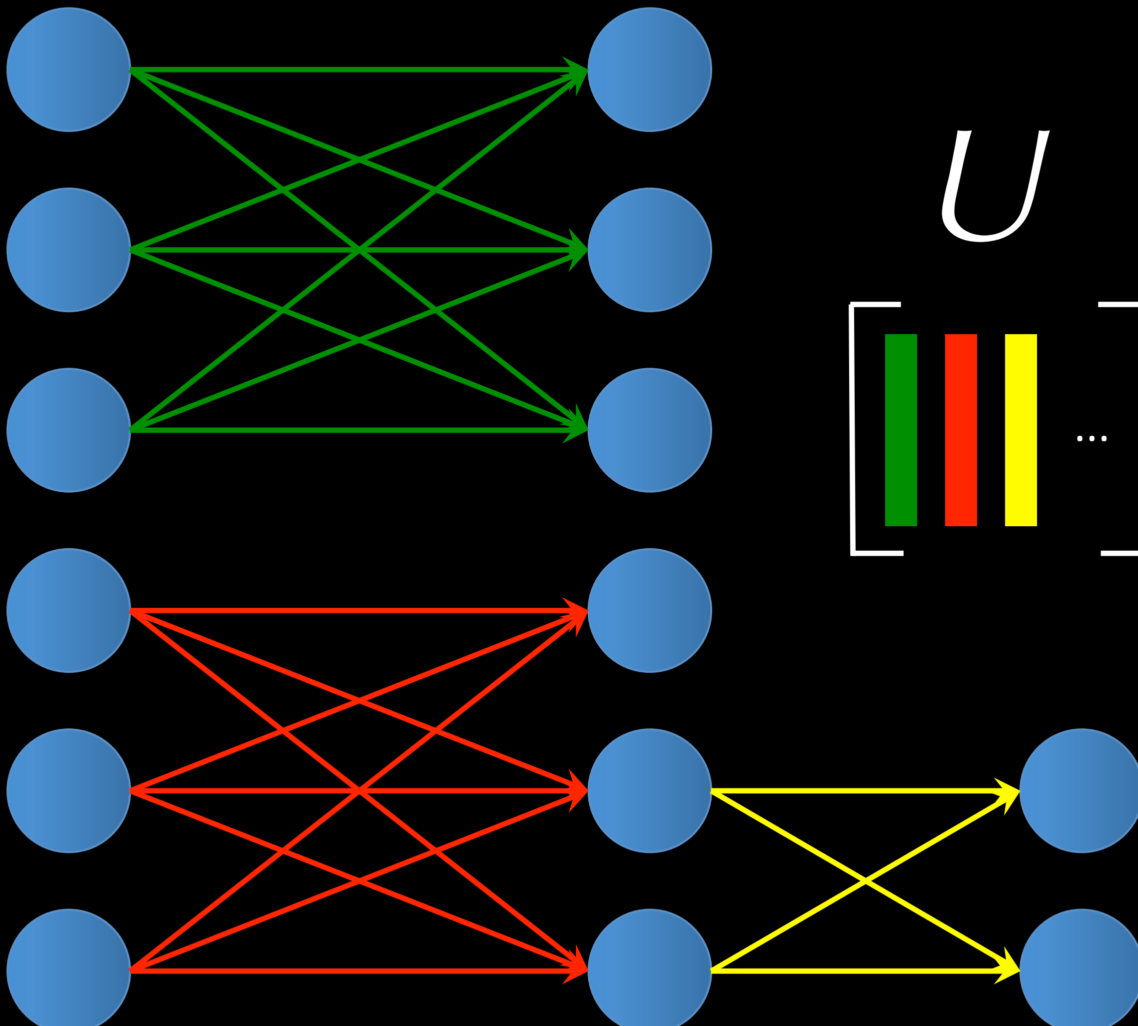
“Left singular vectors”

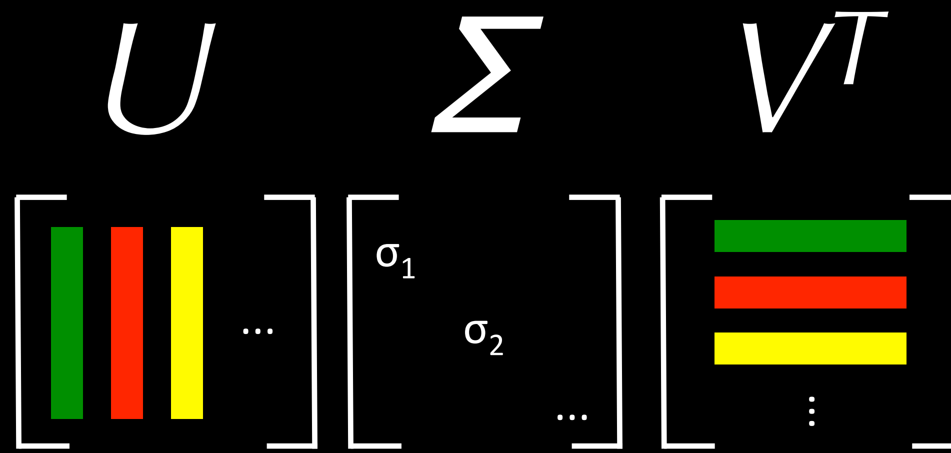
“Singular values”

“Right singular vectors”



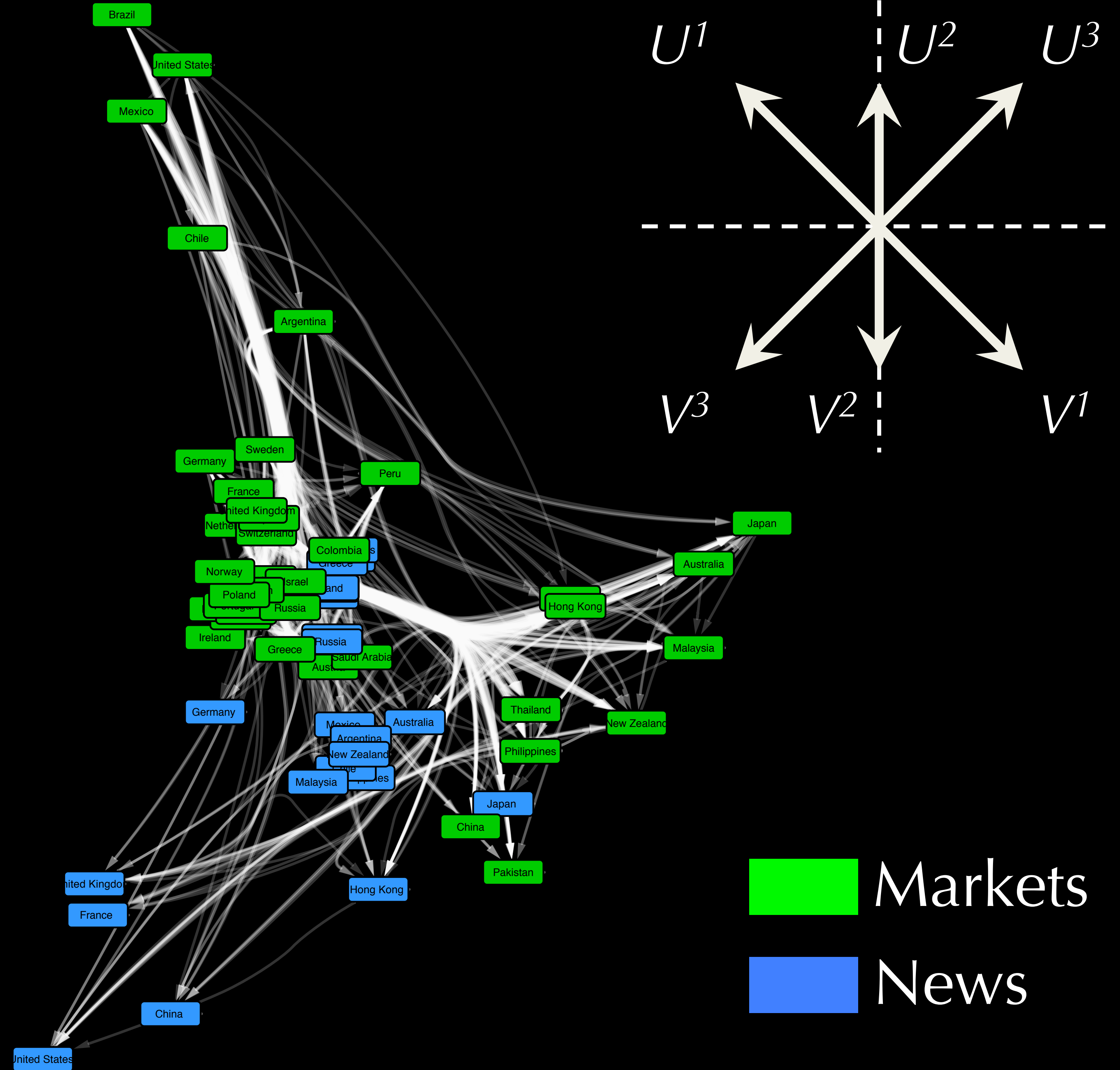
Finding bipartite substructure





Finding bipartite substructure

U^1		V^1
United States		New Zealand
Mexico		Philippines
Brazil	→	Australia
Chile		Japan
Argentina		Malaysia
U^2		V^2
United States		France
Mexico		United Kingdom
Brazil	→	Sweden
Chile		Finland
Saudi Arabia		Belgium
U^3		V^3
Japan		China News
Australia		United States News
Philippines	→	United Kingdom News
Hong Kong		Hong Kong News
Malaysia		Japan News



U^1
U^2
U^3
V^3
V^2
V^1
Markets
News
Brazil
Mexico
Chile
Argentina
Sweden
Germany
Peru
France
Switzerland
Japan
Colombia
Australia
Norway
Israel
Poland
Russia
Hong Kong
Ireland
Greece
Saudi Arabia
Malaysia
Germany
Thailand
Australia
New Zealand
Mexico
Argentina
New Zealand
Philippines
Malaysia
Japan
China
Pakistan
Hong Kong
France
China

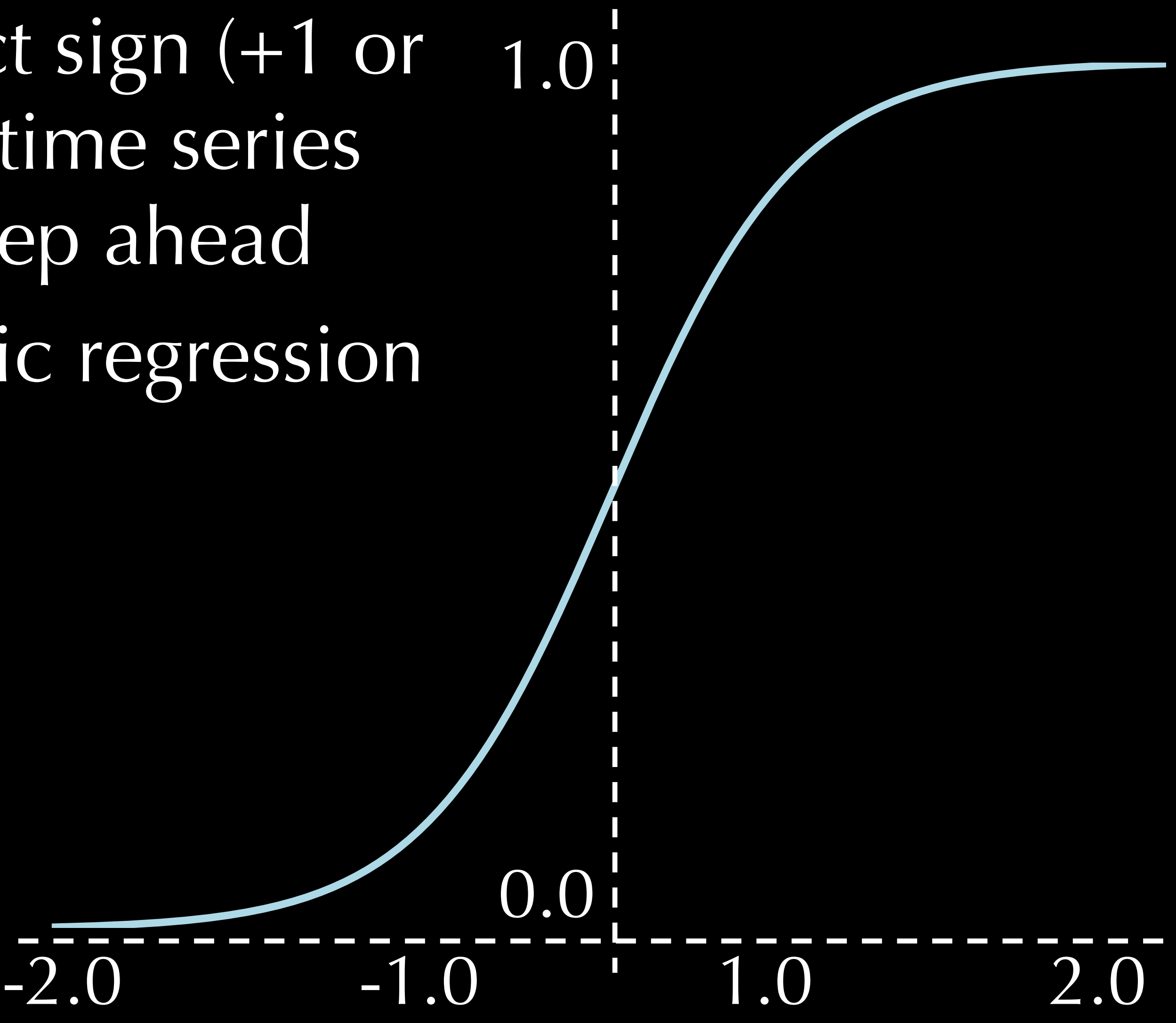
Making predictions

Making predictions

- Predict sign (+1 or -1) of time series one step ahead

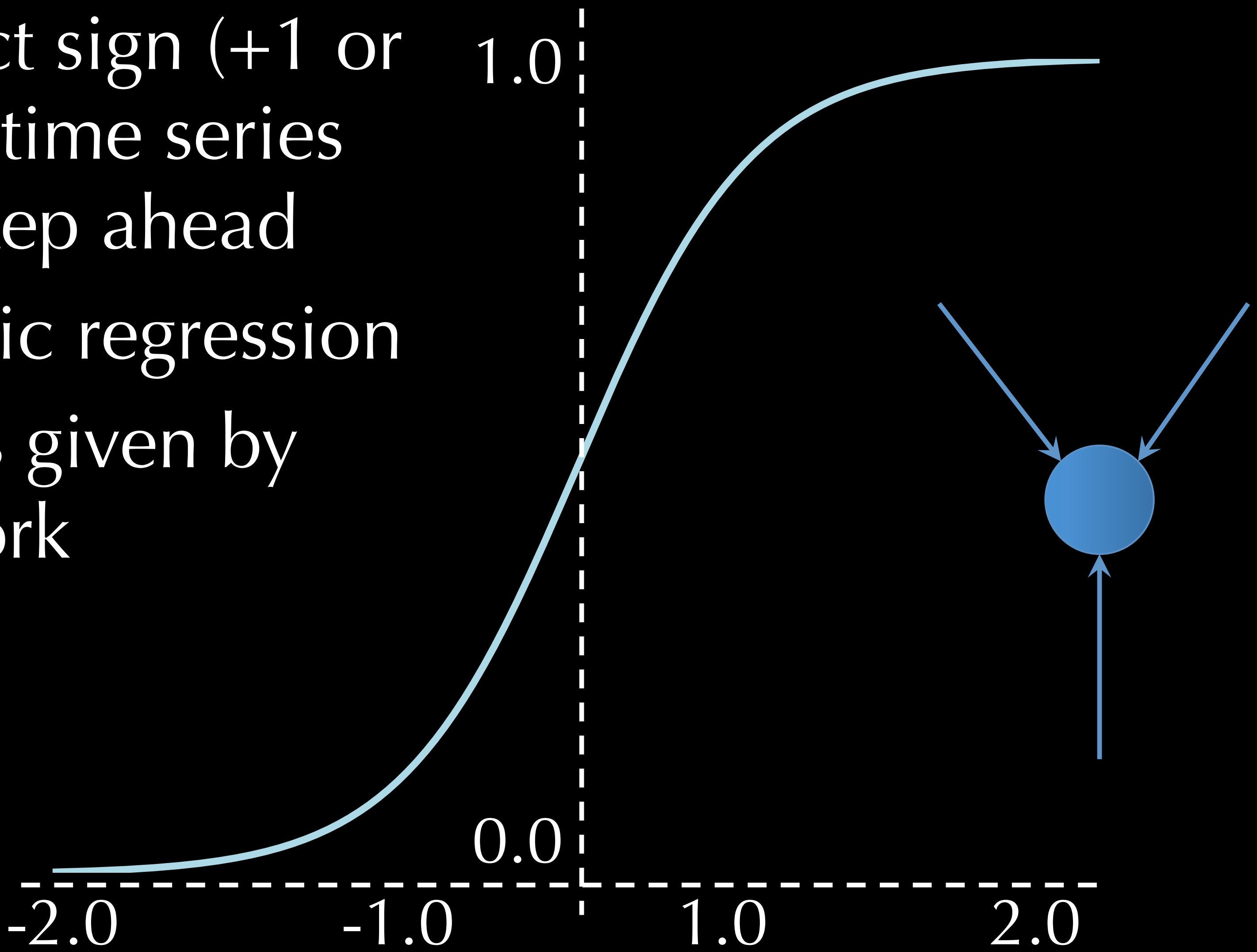
Making predictions

- Predict sign (+1 or -1) of time series one step ahead
- Logistic regression



Making predictions

- Predict sign (+1 or -1) of time series one step ahead
- Logistic regression
- Inputs given by network



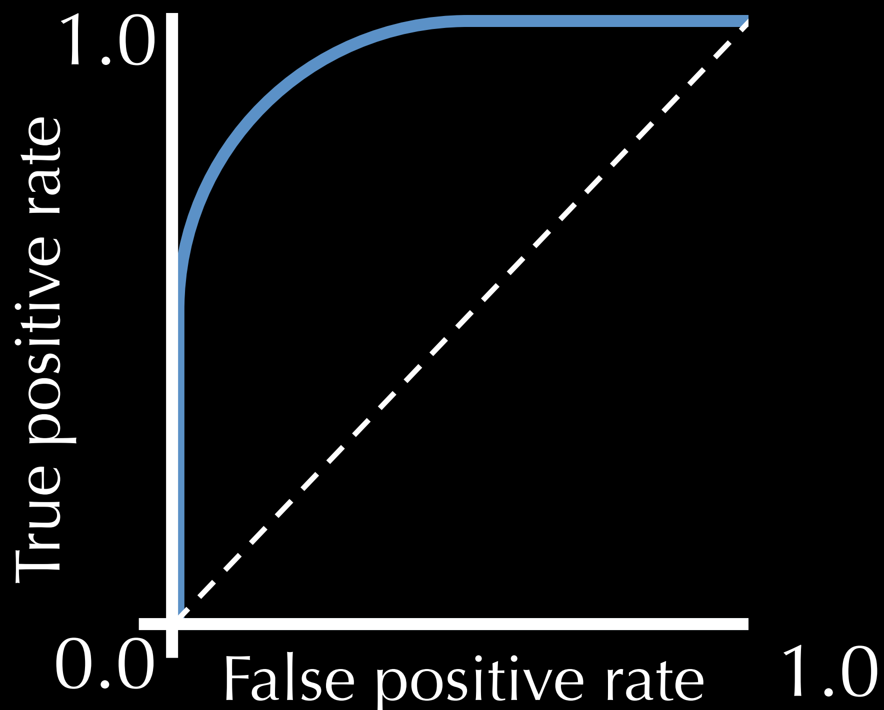
Evaluation

- Divide data into training and test sets.
- Construct network and train logistic regressions using training set.
- Evaluate accuracy on test set.

Evaluation

- Divide data into training and test sets.
- Construct network and train logistic regressions using training set.
- Evaluate accuracy on test set.

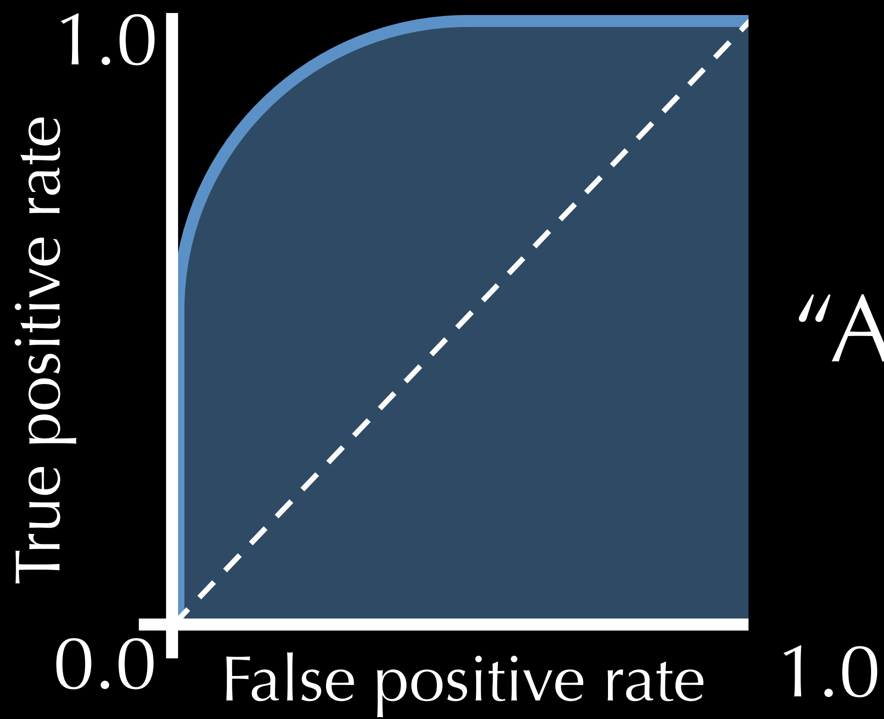
"ROC Curve"



Evaluation

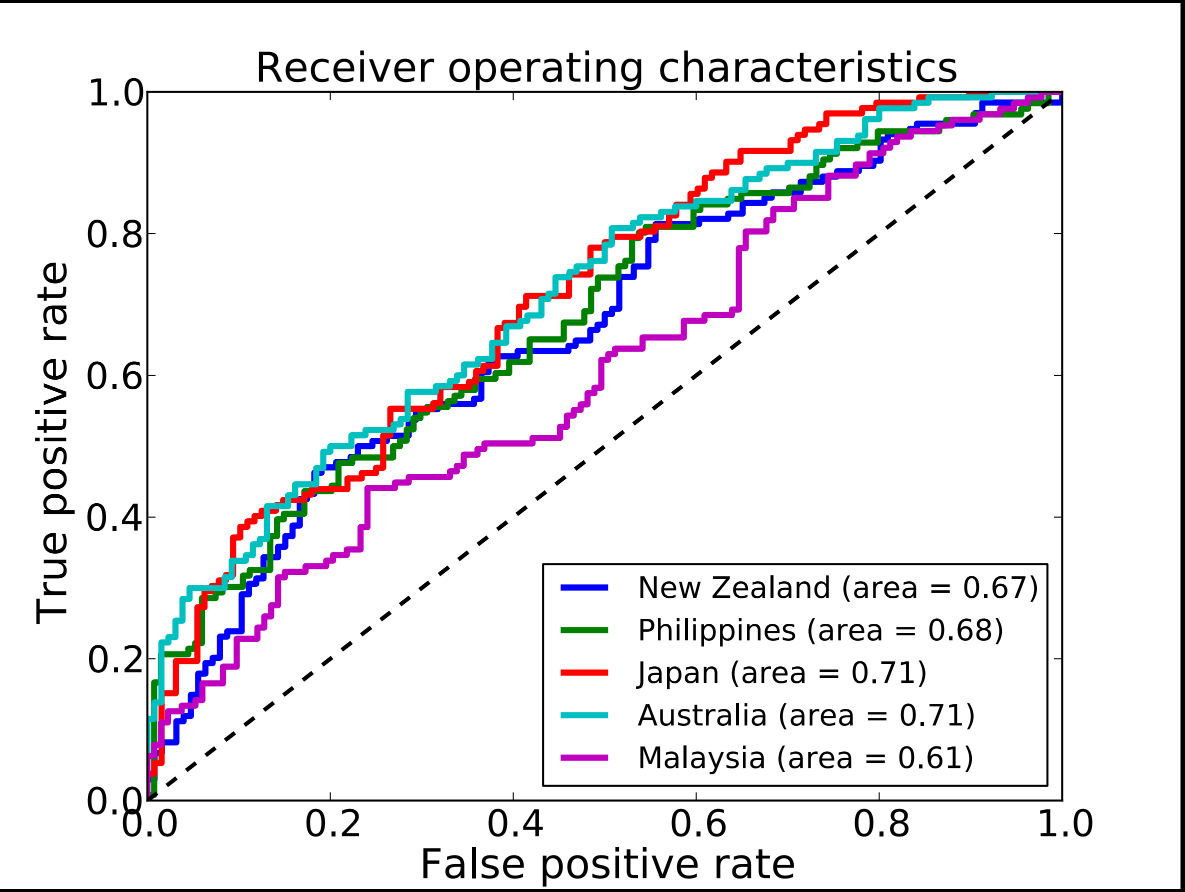
- Divide data into training and test sets.
- Construct network and train logistic regressions using training set.
- Evaluate accuracy on test set.

“ROC Curve”

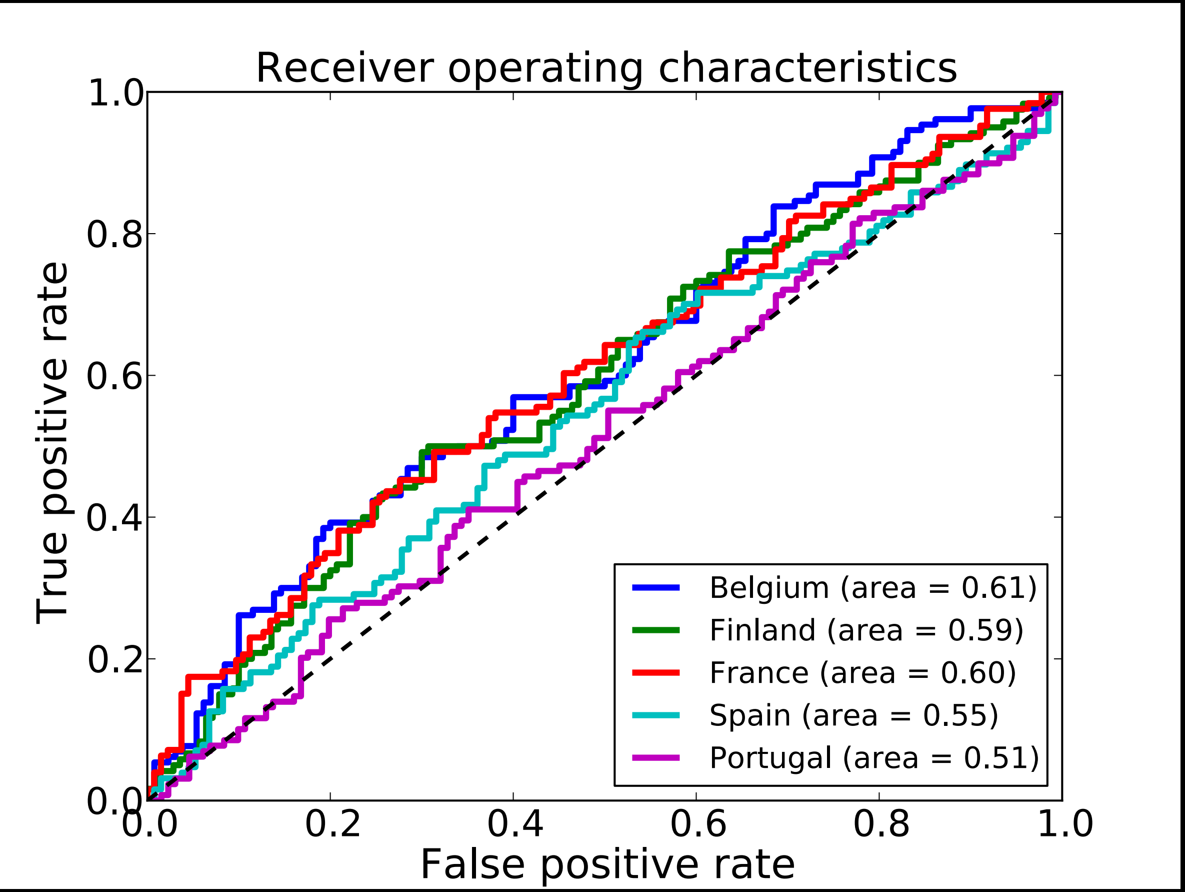


“AUC”

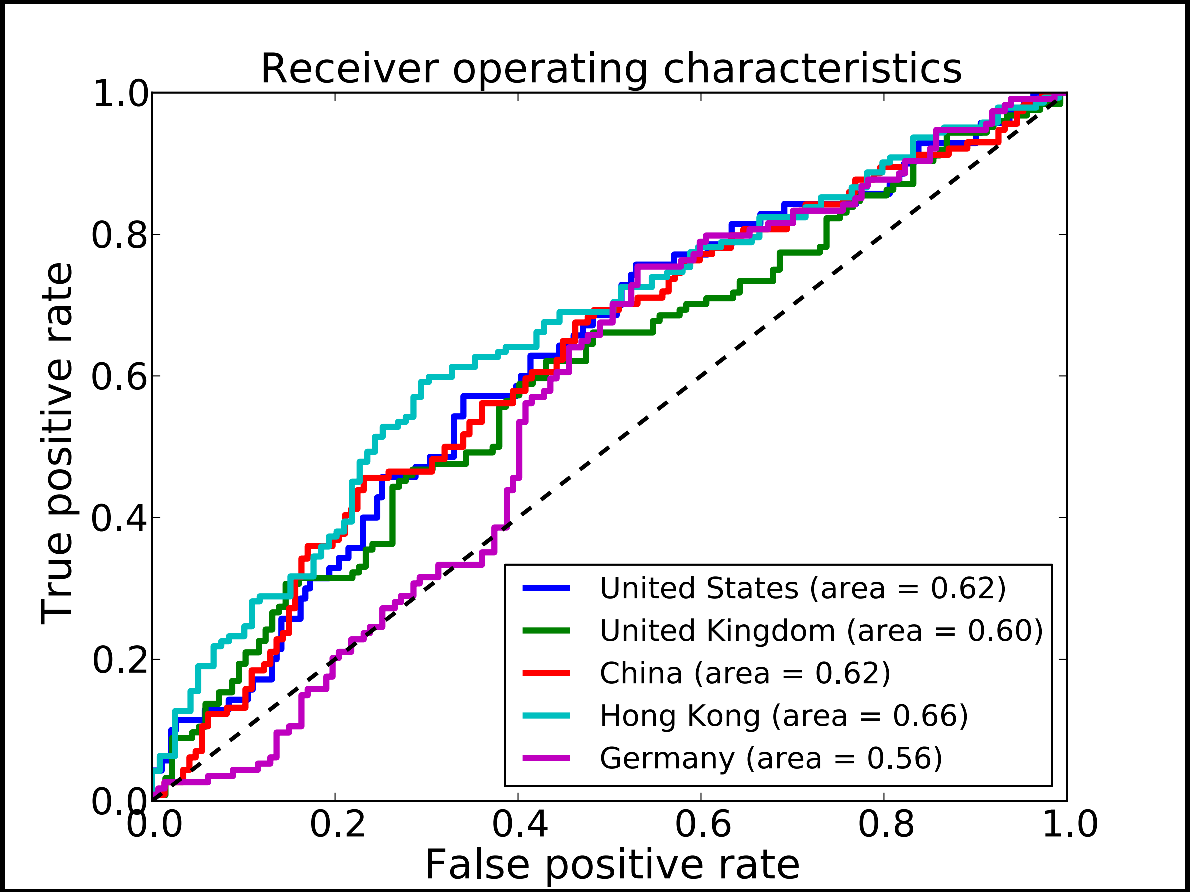
Network captures predictive relationships



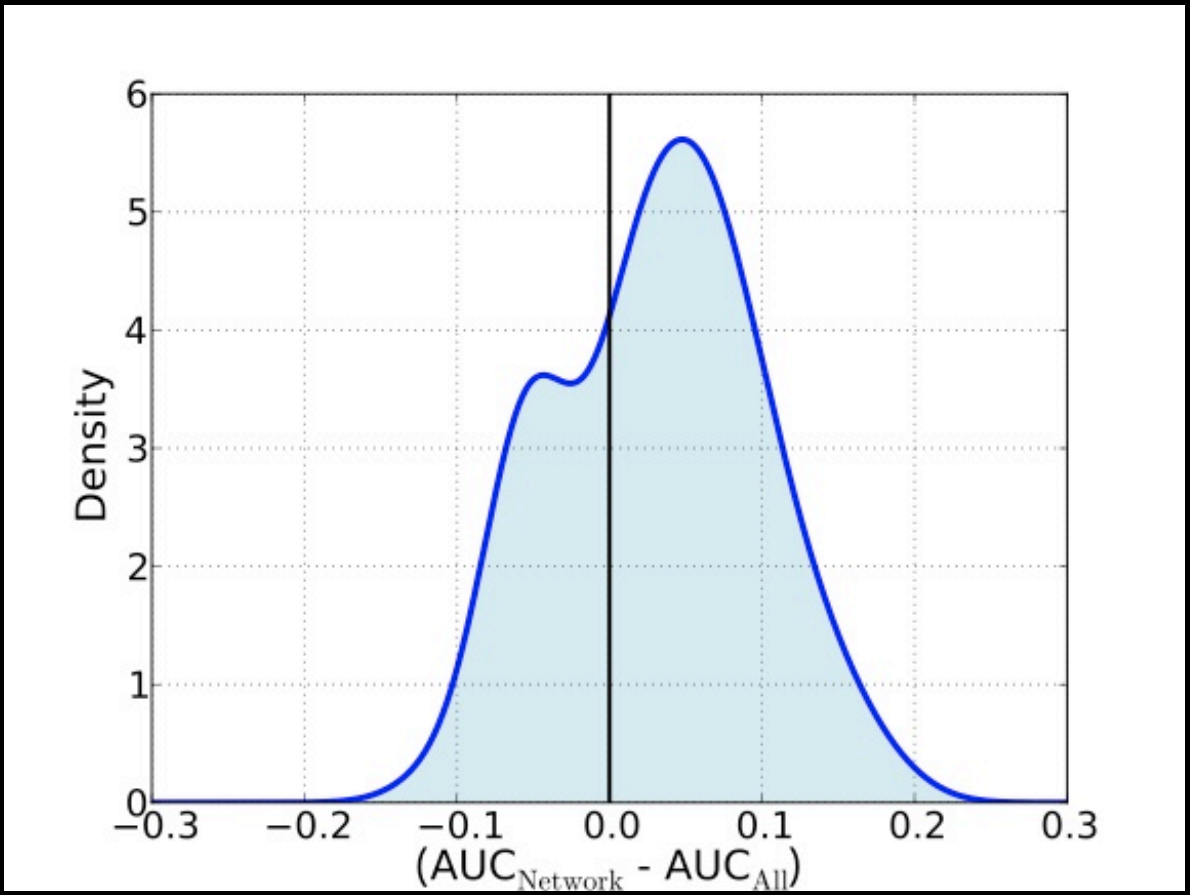
Network captures predictive relationships



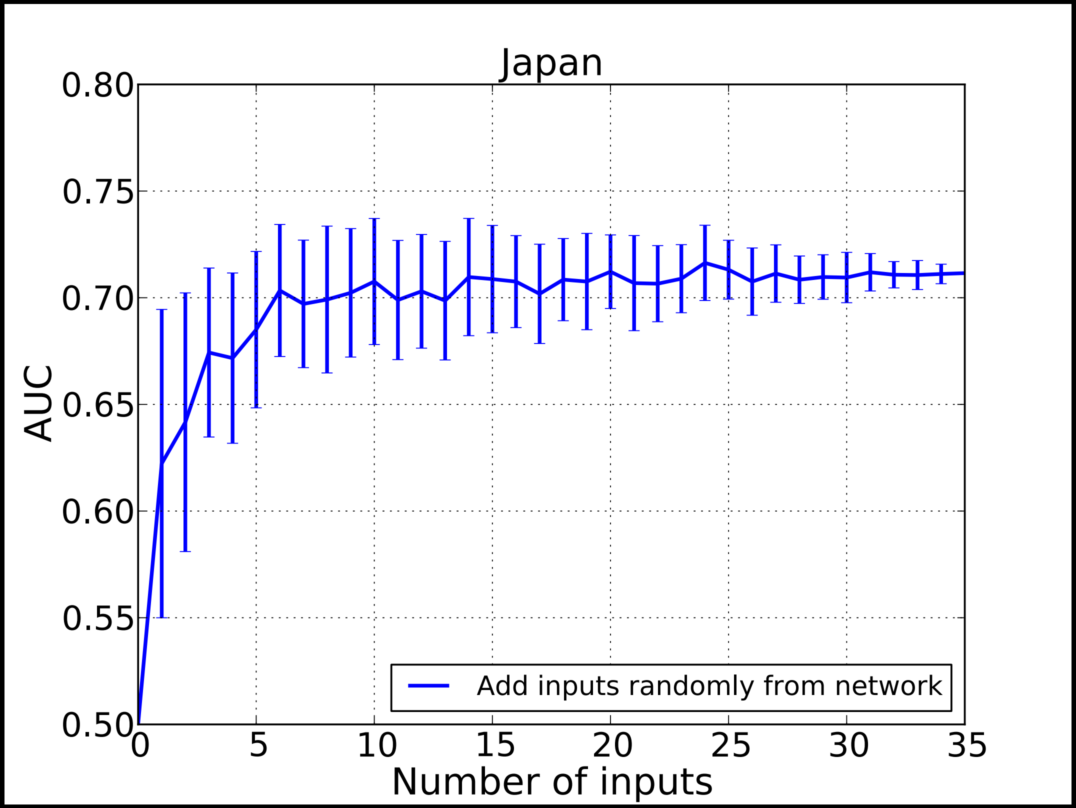
Network captures predictive relationships

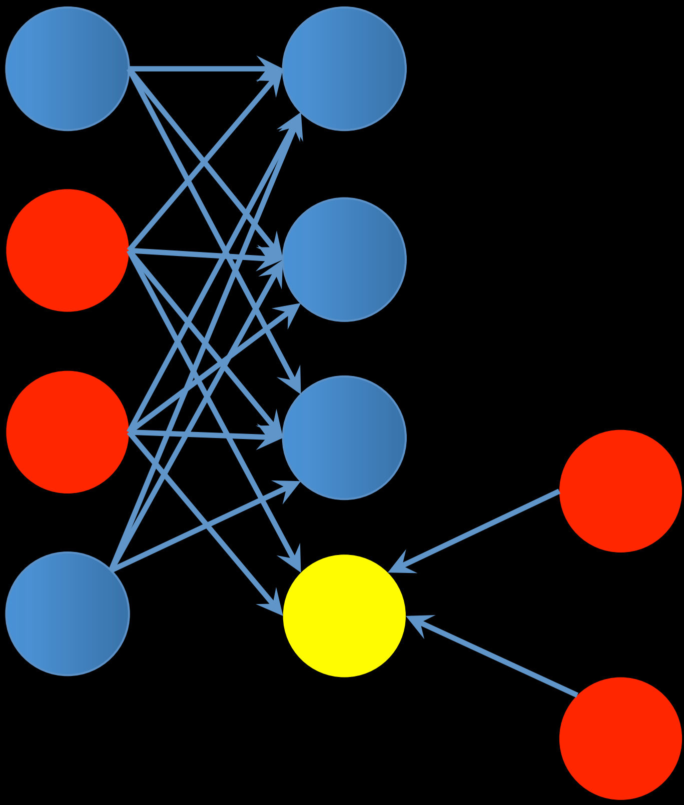


Restricting to network inputs boosts accuracy



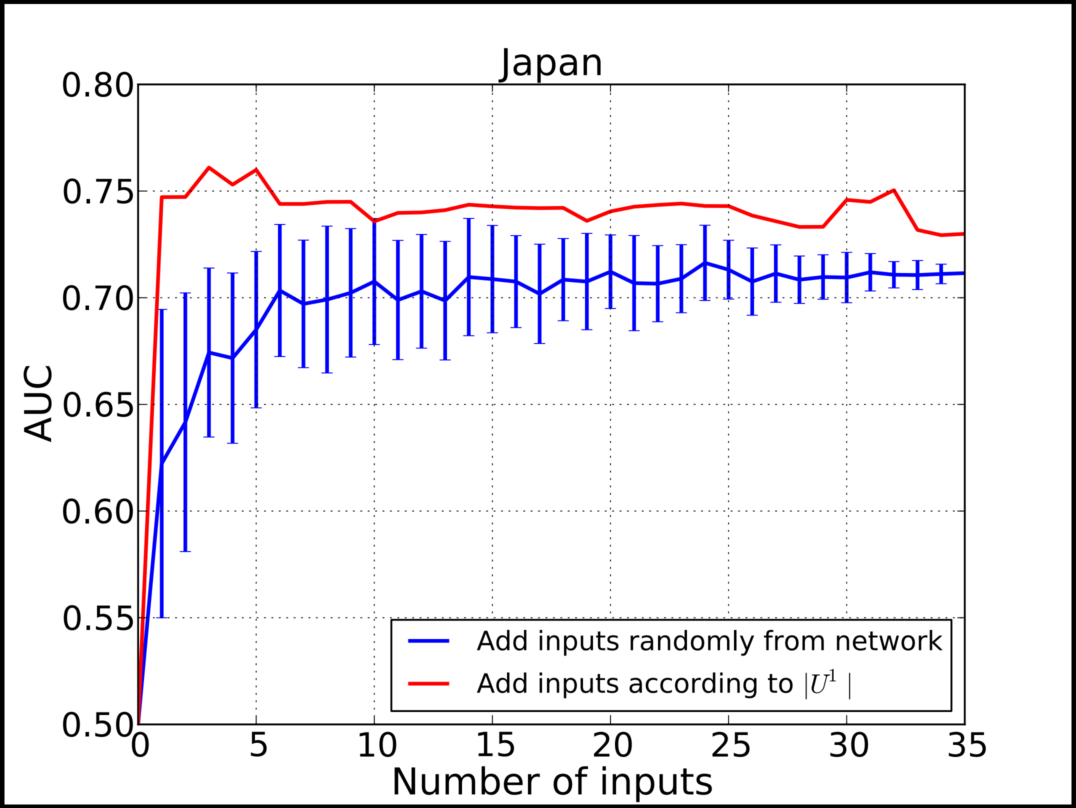
Using communities to recommend inputs

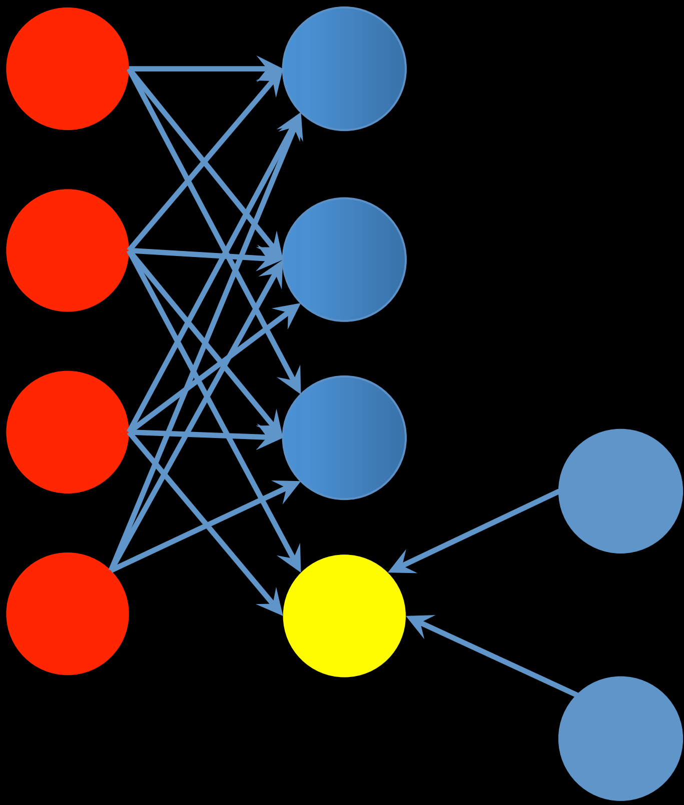




Element of V^1

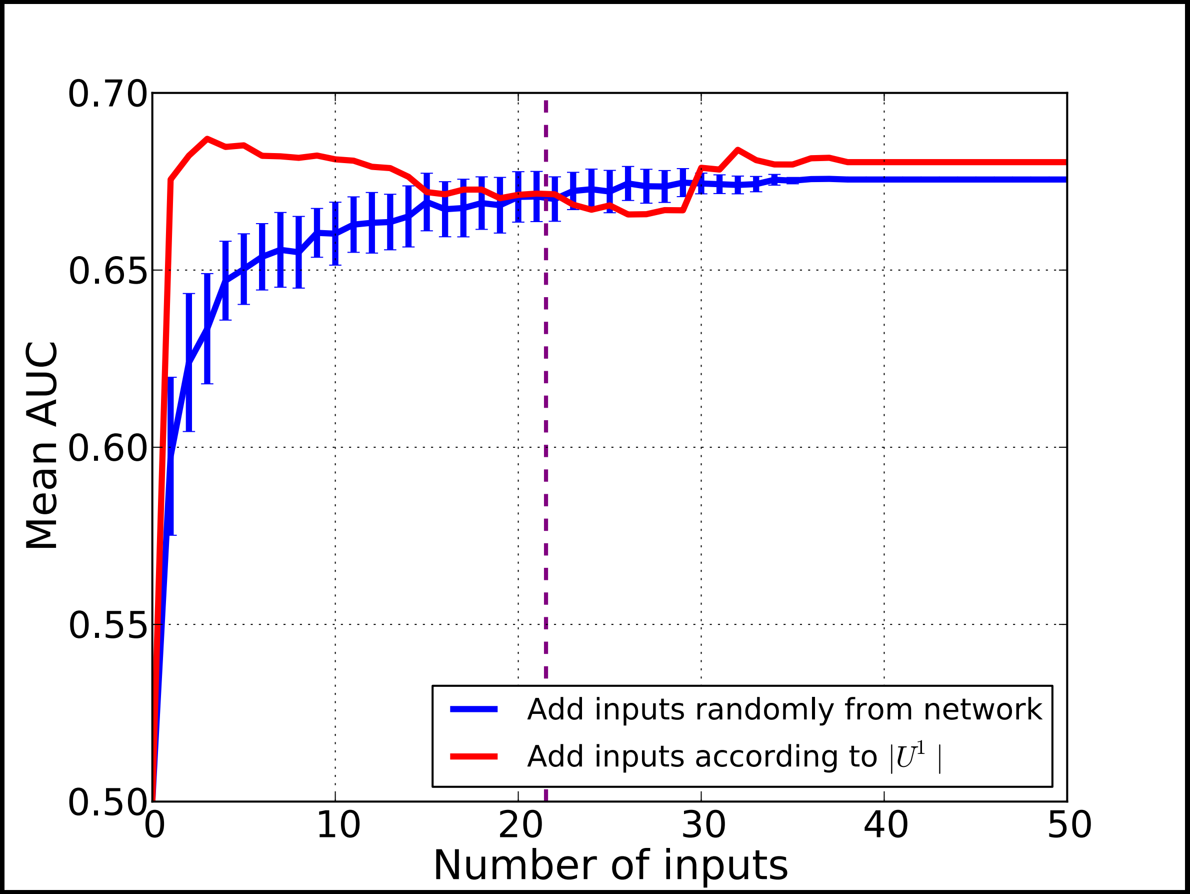
Using communities to recommend inputs





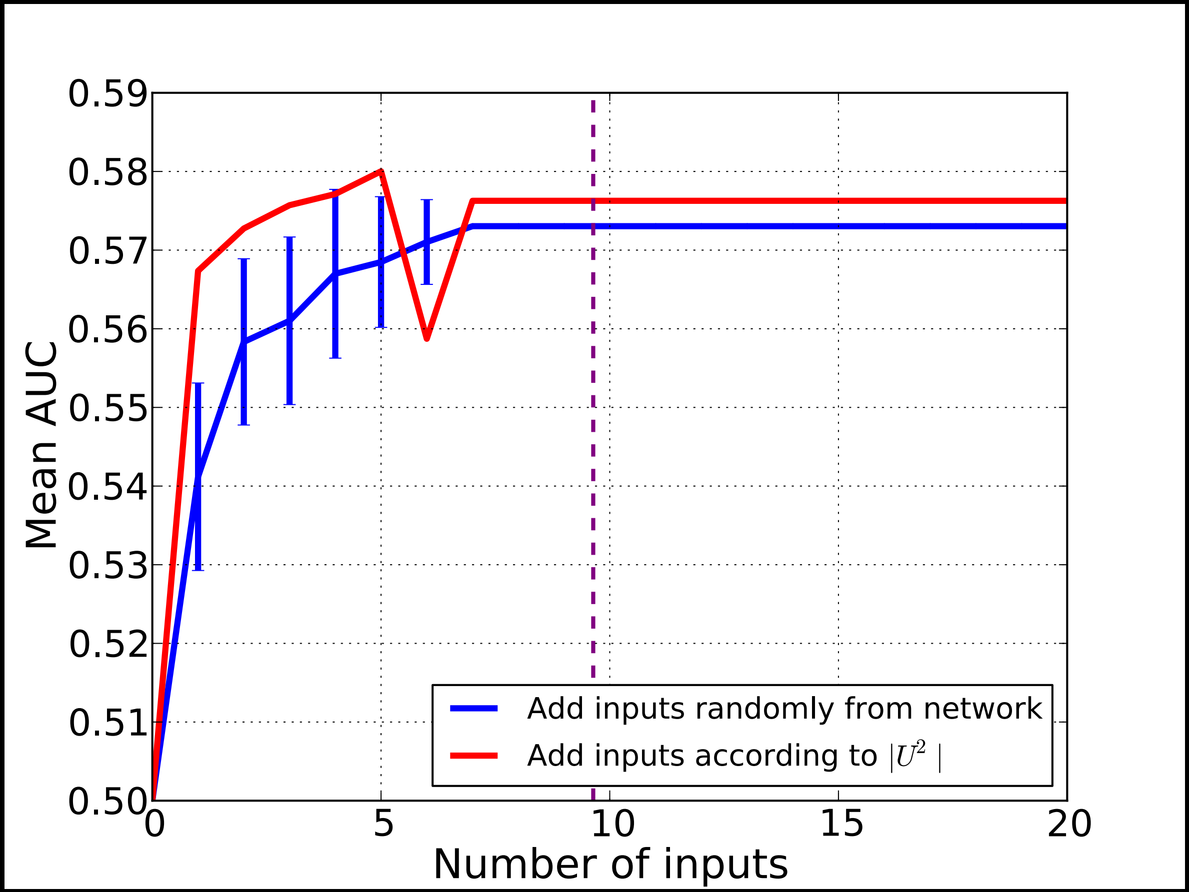
Element of V^1

Using communities to recommend inputs



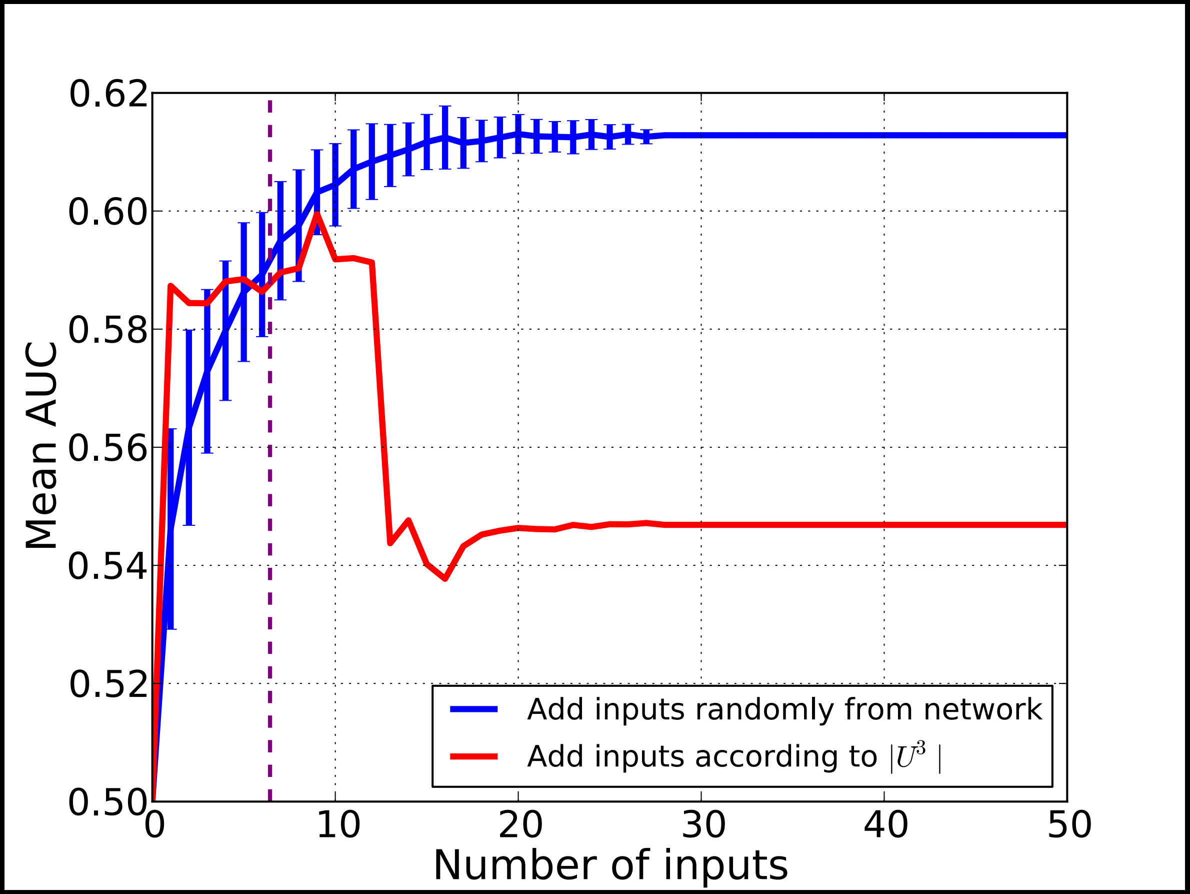
Elements of V^1

Using communities to recommend inputs



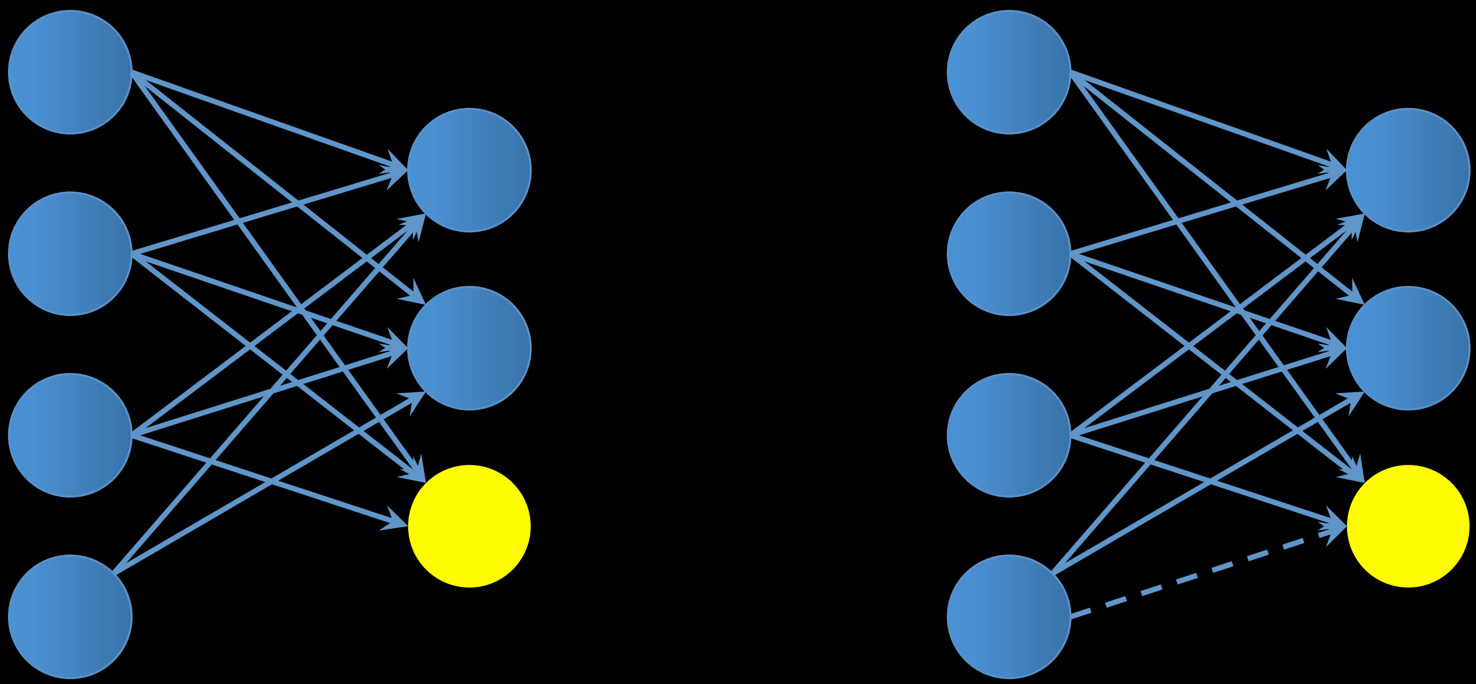
Elements of V^2

Using communities to recommend inputs



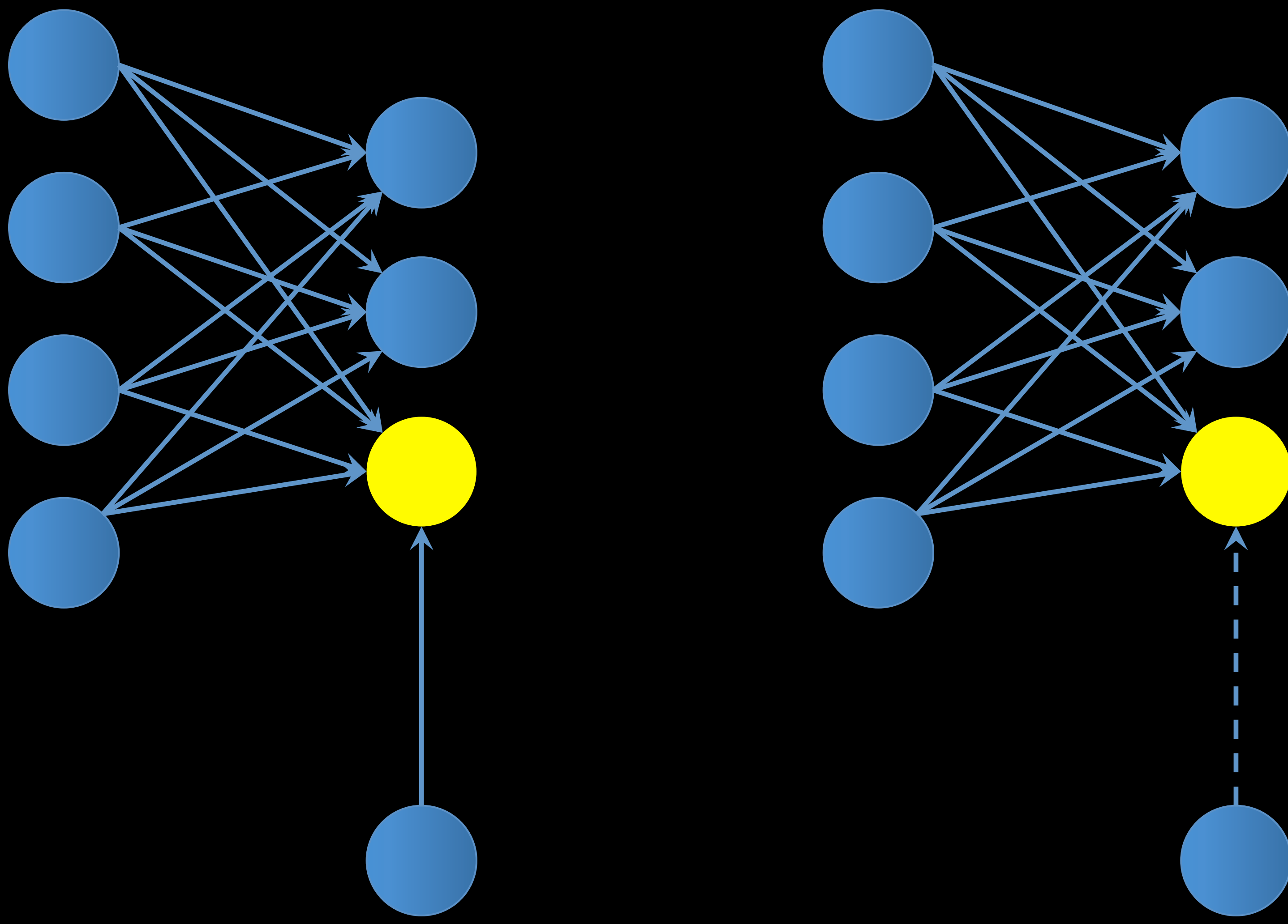
Elements of V^3

Recommender system interpretation



Bolster against missing links

Recommender system interpretation



...and spurious links.

Summary

- SVN methodology reveals global network of interactions among market movements and financial news sentiment signals.
  - News responds to market movements.

Summary

- SVN methodology reveals global network of interactions among market movements and financial news sentiment signals.
  - News responds to market movements.
- Community structures show collective interactions among groups of countries.

Summary

- SVN methodology reveals global network of interactions among market movements and financial news sentiment signals.
  - News responds to market movements.
- Community structures show collective interactions among groups of countries.
- In this setting, community structures simultaneously form the basis of a “recommender system” for model inputs.

Thank you!

Questions?