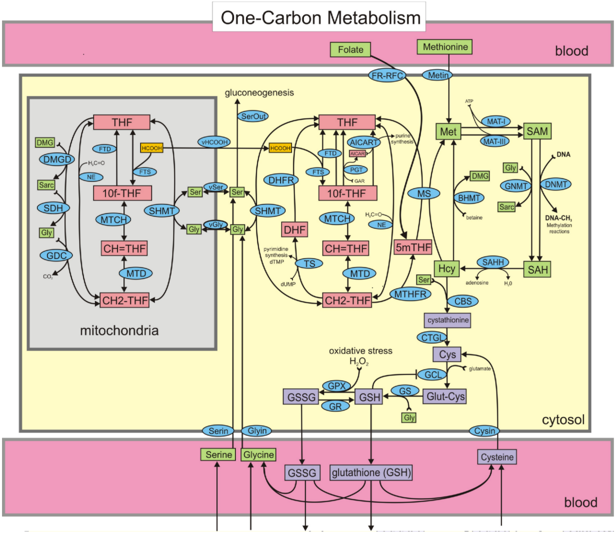
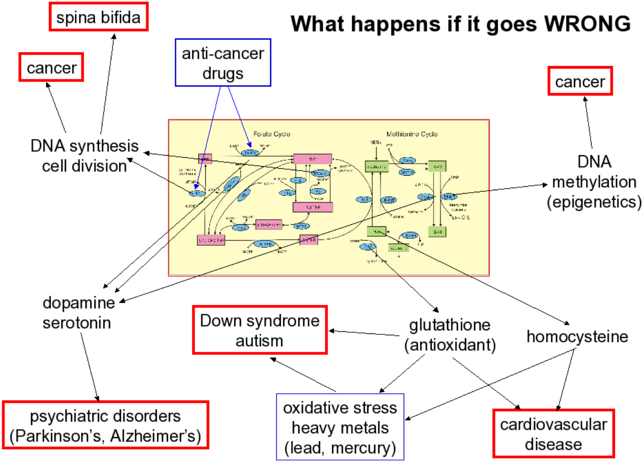
# Random Input to Networks

Rachel Thomas

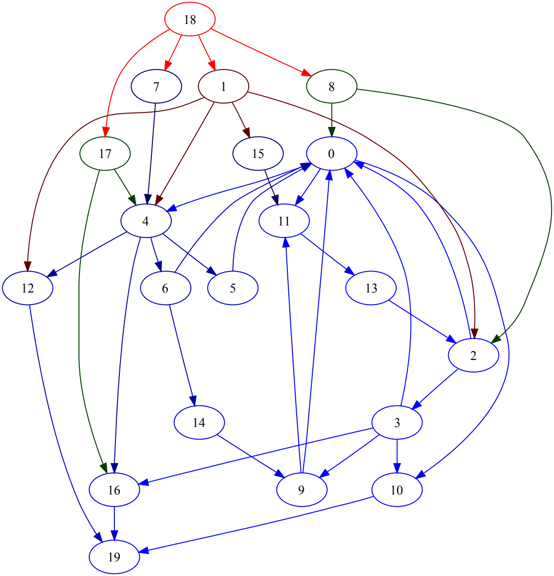
March 28, 2008

# A Biology Problem




What happens if it goes WRONG
spina bifida
cancer
anti-cancer drugs
cancer
DNA synthesis cell division
Folate Cycle
Methionine Cycle
DNA methylation (epigenetics)
dopamine serotonin
Down syndrome autism
glutathione (antioxidant)
homocysteine
psychiatric disorders (Parkinson's, Alzheimer's)
oxidative stress heavy metals (lead, mercury)
cardiovascular disease

# A Math Problem



# Questions

- How do changes in input propagate through the network?
- What network properties affect variance?
- Can we decompose a complicated network into simpler components?

# A Simple Example

$$
\begin{tikzcd}
{} \arrow[r, "I+\xi"] & X_1 \arrow[r, "k_3"] \arrow[d, shift right, "k_1"'] & X_3 \arrow[r, "k_4"] & {} \\
& X_2 \arrow[u, shift right, "k_2"'] & &
\end{tikzcd}
$$

$$
\begin{aligned}
\dot{x}(t) &= Ax(t) + I + \xi(t) \\
x(0) &= x_0
\end{aligned}
$$

$$
A = \begin{pmatrix} -(k_1 + k_3) & k_2 & 0 \\ k_1 & -k_2 & 0 \\ k_3 & 0 & -k_4 \end{pmatrix}
$$

# SSC Networks

- Weakly reversible
- Mass action kinetics (linear ODEs)
- Single linkage class
- Each complex is a single substrate

# The Equations

$$\begin{cases} \dot{x}(t) &= Ax(t) + I + \xi(t) \\ x(0) &= x_0 \end{cases}$$

$$x^*(t,\xi) = \int_{-\infty}^{t} e^{A(t-s)} I ds + \int_{-\infty}^{t} e^{A(t-s)} \xi(s) ds$$

$$m_i = \mathbf{E}(x_i) = I \int_{-\infty}^{t} e^{A(t-s)} e_1 \cdot e_i ds$$

# Simple Chains

Variance decreases down a chain (Anderson)

$$\xrightarrow{I+\xi(t)} X_1 \xrightarrow{k_1} \cdots \xrightarrow{k_{m-1}} X_m \xrightarrow{k_m}$$

For a non-reversible chain and stationary stochastic process $\xi(t)$ with finite variance, mean zero, and $\xi(t) \geq -I$,

$$\begin{aligned} Var(k_i x_i^*) &< Var(\xi) \\ Var(k_{i+1} x_{i+1}^*) &< Var(k_i x_i^*) \end{aligned}$$

# Variance Lowers

**Theorem:** (Anderson) Let $x^*(t)$ be the solution of an SSC system with one input $I + \xi(t)$, where $\xi(t)$ is a stationary stochastic process with finite variance, mean zero, and $\xi(t) \geq -I$. Let $m_i$ be the mean of species $x_i$. Then

$$var(x_i^*) < \left(\frac{m_i}{I}\right)^2 Var(\xi)$$

Proof:

$$
\begin{aligned}
Var(x_i^*(t)) &= \mathbf{E}\left(\int_{-\infty}^{t} \xi(s) e^{A(t-s)} e_1 \cdot e_i ds\right)^2 \\
&= \mathbf{E}\left(\int_{-\infty}^{t} \xi(s) (e^{A(t-s)} e_1 \cdot e_i)^{1/2} (e^{A(t-s)} e_1 \cdot e_i)^{1/2} ds\right)^2 \\
&< \mathbf{E}\left(\int_{-\infty}^{t} \xi(s)^2 e^{A(t-s)} e_1 \cdot e_i ds\right)\left(\int_{-\infty}^{t} e^{A(t-s)} e_1 \cdot e_i ds\right) \\
&= Var(\xi)\left(\int_{-\infty}^{t} e^{A(t-s)} e_1 \cdot e_i ds\right)^2 \\
&= \left(\frac{m_i}{I}\right)^2 Var(\xi)
\end{aligned}
$$

# Back to simple chains

$$\xrightarrow{I+\xi(t)} X_1 \xrightarrow{k_1} \cdots \xrightarrow{k_{m-1}} X_m \xrightarrow{k_m}$$

Consider input to $k_{i+1}x^*_{i+1}$ to be

$$I + (k_i x^*_i - I)$$

Then $\psi = (k_i x^*_i - I)$ is a finite variance, mean zero, stationary stochastic process

By the above theorem, $Var(k_{i+1}x^*_{i+1}) < Var(k_i x^*_i)$

# Reversible Chains

(Anderson) Variance of flux decreases down a reversible chain.

$$0 \underset{b_0}{\overset{J+\psi}{\rightleftharpoons}} X_1 \underset{b_1}{\overset{k_1}{\rightleftharpoons}} X_2 \underset{b_2}{\overset{k_2}{\rightleftharpoons}} \cdots \quad \underset{b_{n-1}}{\overset{k_{n-1}}{\rightleftharpoons}} X_n \xrightarrow{k_n}$$

$$0 \xrightarrow{I+\xi} X_1 \underset{b_1}{\overset{k_1}{\rightleftharpoons}} X_2 \underset{b_2}{\overset{k_2}{\rightleftharpoons}} \cdots \quad \underset{b_{n-1}}{\overset{k_{n-1}}{\rightleftharpoons}} X_n \xrightarrow{k_n}$$

$\xi$ is a stationary, mean zero, finite variance stochastic process

$$\begin{aligned} I &= J + h \qquad \text{where } h = \mathbf{E}[\psi - b_0 x_1] \\ \xi(t) &= \psi(t) - h \end{aligned}$$

# Reversible chains proof

Variance of flux decreases down reversible chain

Pf: Define fluxes of the chain

$$\begin{aligned} y_1 &= k_1 x_1 - b_1 x_2 \\ &\vdots \\ y_{n-1} &= k_{n-1} x_{n-1} - b_{n-1} x_n \\ y_n &= k_n x_n \end{aligned}$$

# Reversible chains proof ctd...

$$
\begin{aligned}
\dot{y}_i &= k_i y_{i-1} - (k_i + b_i) y_i + b_i y_{i+1} \qquad \text{for } 1 < i < n \\
\dot{y}_1 &= k_1 (I + \xi - y_1) - b_1 (y_1 - y_2) \\
\dot{y}_n &= k_n (y_{n-1} - y_n)
\end{aligned}
$$

$$
y(t) = k_1 I \int_{-\infty}^{t} e^{A(t-s)} e_1 ds + k_1 \int_{-\infty}^{t} e^{A(t-s)} \xi(s) ds
$$

Same argument as before to show that $Var(y_i) < Var(\xi)$

# Side Reactions

$$\begin{tikzcd} {} \arrow[r, "I+\xi(t)"] & X_1 \arrow[r, "k_1"] \arrow[d, shift left, "k_3"] & {} \\ & \boxed{\text{Side System}} \arrow[u, shift left, "k_2"] & \end{tikzcd}$$

$$\xrightarrow{I+\xi(t)} \tilde{X}_1 \xrightarrow{k_1}$$

Assume: for $s < t$, $\mathbf{E}\xi(t)\xi(s) > 0$ and $\mathbf{E}\xi(t)\xi(s)$ is increasing in $s$. (Anderson) Then a side reaction system lowers variance:

$$Var(k_1 x_1^*) < Var(k_1 \tilde{x}_1^*)$$

# Feedback Loops

Feedback loops lower variance (Anderson)

$$\begin{tikzcd}
{} \arrow[r, "I+\xi(t)"] & X_1 \arrow[r, "k_1"] & \cdots \arrow[r, "k_{n-1}"] & X_n \arrow[r, "k_n"] \arrow[dl, "c"] & {} \\
 & & \boxed{\text{Subsystem}} \arrow[ul, "f_1(t)"] & &
\end{tikzcd}$$

$$\xrightarrow{I+\xi(t)} \tilde{X}_1 \xrightarrow{k_1} \cdots \xrightarrow{k_{n-1}} \tilde{X}_n \xrightarrow{k_n}$$

With the same hypotheses on $\xi(t)$ and $\mathbf{E}\xi(t)\xi(s)$ as before, then

$$Var(k_n x_n^*) < Var(k_n \tilde{x}_n^*)$$

## Feedback Loops

Proof:

$$
\begin{tikzcd}
{} \arrow[r, "I+\xi(t)"] & X_1 \arrow[r, "k_1"] & \cdots \arrow[r, "k_{n-1}"] & X_n \arrow[r, "k_n"] \arrow[dl, "c"] & {} \\
& & \boxed{\text{Subsystem}} \arrow[ul, "f_1(t)"] & &
\end{tikzcd}
$$

$$
\begin{tikzcd}
{} \arrow[r, "I+\xi(t)"] & Z_1 \arrow[r, "k_1"] & \cdots \arrow[r, "k_{n-1}"] & Z_n \arrow[r, "k_n"] \arrow[dll, "c"'] & {} & \\
& \boxed{\text{Subsystem}} \arrow[rr, "f_1(t)"'] & & Y_1 \arrow[r, "k_1"'] & \cdots & Y_{n-1} \arrow[ull, "k_{n-1}"']
\end{tikzcd}
$$

$$
\begin{aligned}
X_i &= Y_i + Z_i \text{ for } 1 \le i \le n-1 \\
X_n &= Z_n
\end{aligned}
$$

# Feedforward Loops

$$
\begin{tikzcd}
{} \arrow[r, "I+\xi(t)"] & X_1 \arrow[r, "k_1"] \arrow[dr, "h"'] & \cdots \arrow[r, "k_{n-1}"] & X_n \arrow[r, "k_n"] & {} \\
 & & \boxed{\text{Subsystem}} \arrow[ur, "f(t)"'] & &
\end{tikzcd}
$$

# Feedforward Loops

$$\begin{tikzcd} {} \arrow[r, "I+\xi(t)"] & X_1 \arrow[r, "k_1"] \arrow[rr, bend right, "h"'] & \cdots \arrow[r, "k_{n-1}"] & X_n \arrow[r, "k_n"] & {} \end{tikzcd}$$

$$Var(k_n x_n^*) < \left(\frac{h + k_1}{k_1}\right)^2 Var(k_{n-1} x_{n-1}^*)$$

# Splitting

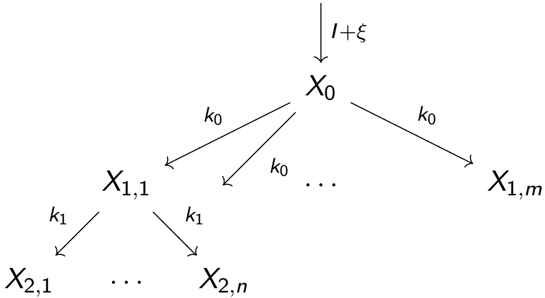
$$
\begin{tikzcd}
 & & X_{1,2} \arrow[r, "k_2"] & \cdots \arrow[r, "k_{n-2}"] & X_{1,n-1} \arrow[dr, "k_{n-1}"] & & \\
{} \arrow[r, "I+\xi"] & X_1 \arrow[ur, "\frac{k_1}{m}"] \arrow[dr, "\frac{k_1}{m}"'] & \vdots & & & X_n \arrow[r, "k_n"] & {} \\
 & & X_{m,2} \arrow[r, "k_2"'] & \cdots \arrow[r, "k_{n-2}"'] & X_{m,n-1} \arrow[ur, "k_{n-1}"'] & &
\end{tikzcd}
$$

$$
\xrightarrow{I+\xi} \tilde{X}_1 \xrightarrow{k_1} \tilde{X}_2 \xrightarrow{k_2} \cdots \xrightarrow{k_{n-2}} \tilde{X}_{n-1} \xrightarrow{k_{n-1}} \tilde{X}_n \xrightarrow{k_n}
$$

# Trees

Two key variables: node degree and level in tree

$$Var(k_1 X_{1,1}) < \frac{1}{n^2} Var(k_0 X_0) < \frac{1}{m^2 n^2} Var(\xi)$$



# Large networks

How can larger networks be studied?
What are different ways to randomly generate networks?

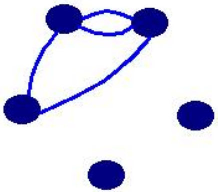
# Erdös-Rényi

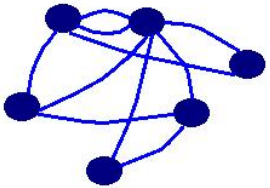
Given $n$ vertices, $\frac{n(n-1)}{2}$ edges are possible.
Pick each edge with equal probability.

This is an exponential network $P(k) \propto e^{-k}$ for large $k$

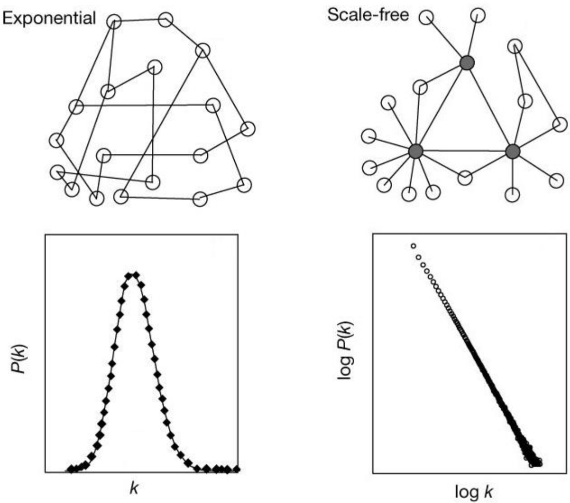
# Albert-Barabási

At each step, add a new vertex with $m$ edges.
Probability $p_i$ a new edge connects to $v_i$ is $\propto k_i$ $\quad p_i = \frac{k_i}{\sum_j k_j}$





Degree distribution is a power law $P(k) \propto k^{-\gamma}$
Scale-free networks are dominated by a few highly connected nodes.




from *Nature*, Oct 2000

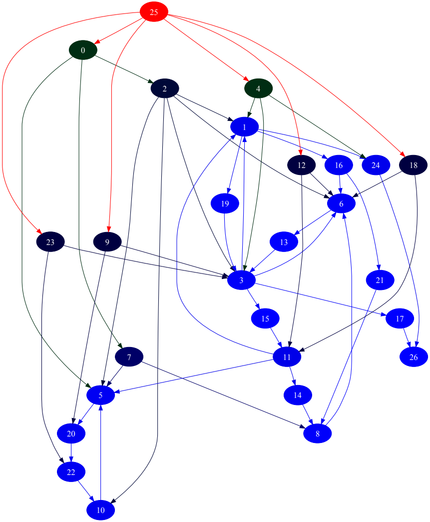
# Question:

How can any network with one input and one output be reduced to a chain with:

- Side reaction systems
- Feedback loops
- Feed-forward loops

# Computer Simulation

1. Generate random network
2. Assign edge directions
3. Add input node
4. Add output node
5. Provide noisy input and simulate
6. Analyze time series of node concentrations



# High dimensional data set

Ex: Record the concentrations at 50 nodes over 20,000 time steps

$$\left[\quad \text{species x time} \quad\right]$$

Can think of this as 20,000 points in 50-dimensional space.
The set may have a lower intrinsic dimensionality.

# Singular Value Decomposition

Any $m \times n$ matrix $A$ can be factored

$$A = U\Sigma V^* \qquad \begin{array}{ll} U & m \times m \text{ orthonormal} \\ V & n \times n \text{ orthonormal} \\ \Sigma & m \times n \text{ diagonal} \end{array}$$

The $\sigma_i$ are the singular values of $A$. They are in descending order:

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_{m \wedge n}$$

# Singular Value Decomposition

SVD is a generalization of eigenvector decomposition

$$\begin{aligned} A &= U \Sigma V^* \\ AA^* &= U(\Sigma\Sigma^*)U^* \end{aligned}$$

The squares of the singular values of $A$ are the eigenvalues of $AA^*$

$$\begin{aligned} AV &= U\Sigma \\ Av &= \sigma u \end{aligned}$$

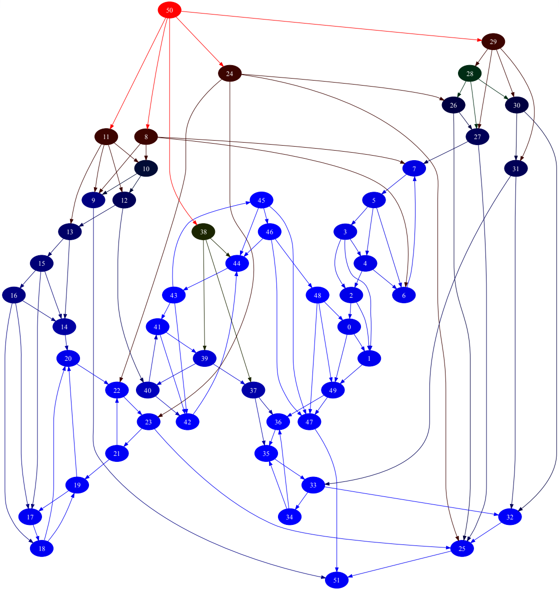
$U$ is an orthonormal basis for the range of $A$
Vectors in $V$ corresponding to $\sigma_i = 0$ are a basis for $Null(A)$

$U$ is an orthonormal basis for the range of $A$.

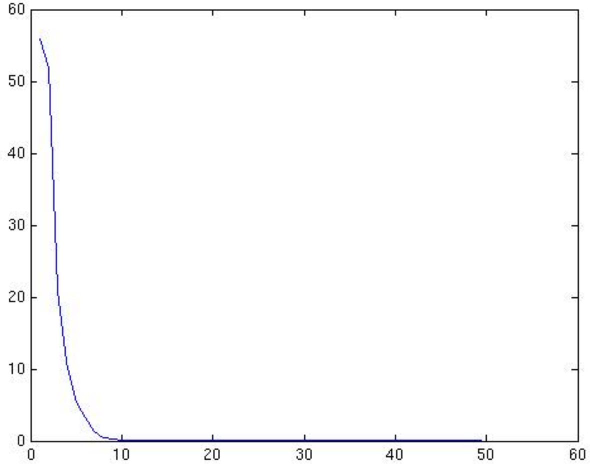
$$\begin{aligned} AV &= U\Sigma \\ Av &= \sigma u \end{aligned}$$

We can pick off the most significant singular values to reduce a higher dimensional data set to lower dimensional set.

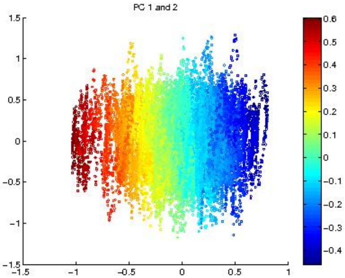
$$\begin{bmatrix} \text{species x time} \end{bmatrix} \begin{bmatrix} & \text{V} & \end{bmatrix} = \begin{bmatrix} \text{U} \end{bmatrix} \begin{bmatrix} \sigma_1 & \dots & \end{bmatrix}$$


50
29
24
28
26
30
11
8
27
10
7
31
9
12
45
5
13
38
46
3
15
44
4
16
43
48
2
6
14
41
0
20
39
1
22
40
37
49
23
42
36
47
21
35
19
33
17
34
32
18
25
51

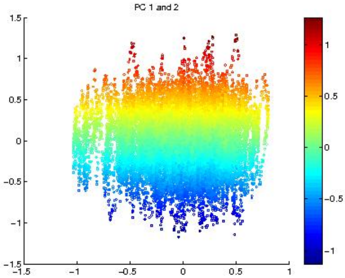
# Singular Values



# Principal Components 1,2

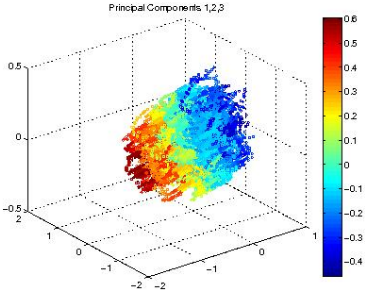



colored by concentration of output node

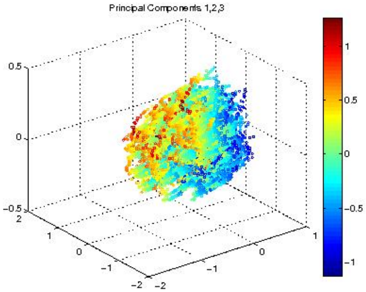



colored by concentration of input node

# Principal Components 1,2,3



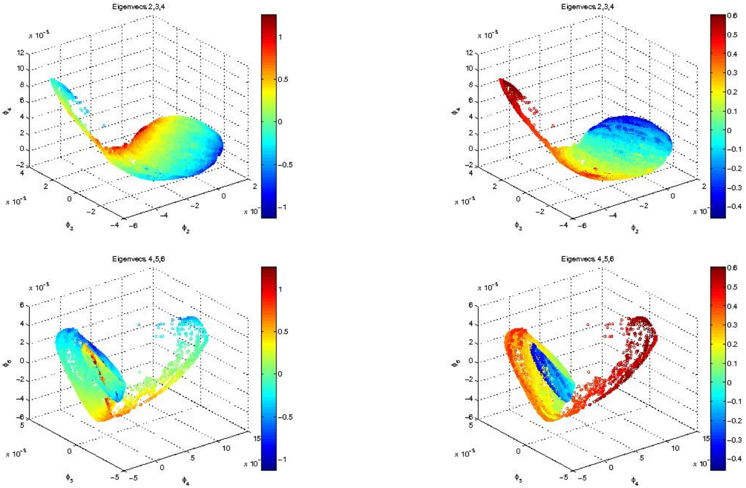

colored by concentration
of output node




colored by concetration
of input node

# Eigenvectors

Eigenvectors over time (in 8,000 dim)




colored by conc. of input node

colored by conc. of output node

## Conclusion

Ways to explore networks:

- Analytic estimates on particular structures
- Simulations on large networks
- Decomposing large networks and putting them back together

# Acknowledgements

Many thanks to Mike Reed, Jonathan Mattingly, Mauro Maggioni, and Dave Anderson.

And thank you for coming today.