



# Updating Markov Chains

**Carl Meyer**  
**Amy Langville**

Department of Mathematics  
North Carolina State University  
Raleigh, NC

A. A. Markov Anniversary Meeting  
June 13, 2006



# Intro

## Assumptions

Very large irreducible chain

—  $m = O(10^9)$

number of states

—  $\mathbf{Q}_{m \times m}$

old transition matrix

—  $\phi^T = (\phi_1, \phi_2, \dots, \phi_m)$

old stationary distribution



# Intro

## Assumptions

Very large irreducible chain

- $m = O(10^9)$  number of states
- $\mathbf{Q}_{m \times m}$  old transition matrix
- $\phi^T = (\phi_1, \phi_2, \dots, \phi_m)$  old stationary distribution

## Updates

Change some transition probabilities

Add or delete some states

- $\mathbf{P}_{n \times n}$  new transition matrix (irreducible)
- $\pi^T = (\pi_1, \pi_2, \dots, \pi_n)$  new distribution (unknown)



# Intro

## Assumptions

Very large irreducible chain

- $m = O(10^9)$  number of states
- $\mathbf{Q}_{m \times m}$  old transition matrix
- $\phi^T = (\phi_1, \phi_2, \dots, \phi_m)$  old stationary distribution

## Updates

Change some transition probabilities

Add or delete some states

- $\mathbf{P}_{n \times n}$  new transition matrix (irreducible)
- $\pi^T = (\pi_1, \pi_2, \dots, \pi_n)$  new distribution (unknown)

## Aim

Use  $\phi^T$  to compute  $\pi^T$



# Exact (Theoretical) Updating

## Perturbation Formula

$$\mathbf{P} = \mathbf{Q} - \mathbf{E} \implies \boldsymbol{\pi}^T = \boldsymbol{\phi}^T - \boldsymbol{\epsilon}^T$$



# Exact (Theoretical) Updating

## Perturbation Formula

$$\mathbf{P} = \mathbf{Q} - \mathbf{E} \implies \boldsymbol{\pi}^T = \boldsymbol{\phi}^T - \boldsymbol{\epsilon}^T \quad \boldsymbol{\epsilon}^T = \boldsymbol{\phi}^T \mathbf{E} \mathbf{Z} (\mathbf{I} + \mathbf{E} \mathbf{Z})^{-1}$$

$$\mathbf{Z} = \begin{cases} \text{Fundamental Matrix} \\ (\mathbf{I} - \mathbf{Q})^\# \quad (\text{group inverse}) \end{cases}$$



# Exact (Theoretical) Updating

## Perturbation Formula

$$\mathbf{P} = \mathbf{Q} - \mathbf{E} \implies \boldsymbol{\pi}^T = \boldsymbol{\phi}^T - \boldsymbol{\epsilon}^T \quad \boldsymbol{\epsilon}^T = \boldsymbol{\phi}^T \mathbf{E} \mathbf{Z} (\mathbf{I} + \mathbf{E} \mathbf{Z})^{-1}$$

Requires  $m = n$  (No states added or deleted)

$$\mathbf{Z} = \begin{cases} \text{Fundamental Matrix} \\ (\mathbf{I} - \mathbf{Q})^\# \quad (\text{group inverse}) \end{cases}$$



# Exact (Theoretical) Updating

## Perturbation Formula

$$\mathbf{P} = \mathbf{Q} - \mathbf{E} \implies \boldsymbol{\pi}^T = \boldsymbol{\phi}^T - \boldsymbol{\epsilon}^T \quad \boldsymbol{\epsilon}^T = \boldsymbol{\phi}^T \mathbf{E} \mathbf{Z} (\mathbf{I} + \mathbf{E} \mathbf{Z})^{-1}$$

Requires  $m = n$  (No states added or deleted)

$$\mathbf{Z} = \begin{cases} \text{Fundamental Matrix} \\ (\mathbf{I} - \mathbf{Q})^\# \quad (\text{group inverse}) \end{cases}$$

## One Row At A Time (Sherman–Morrison)

$$\mathbf{p}_i^T = \mathbf{q}_i^T - \boldsymbol{\delta}_i^T \implies \boldsymbol{\pi}^T = \boldsymbol{\phi}^T - \boldsymbol{\epsilon}^T$$





# Exact (Theoretical) Updating

## Perturbation Formula

$$\mathbf{P} = \mathbf{Q} - \mathbf{E} \implies \boldsymbol{\pi}^T = \boldsymbol{\phi}^T - \boldsymbol{\epsilon}^T \quad \boldsymbol{\epsilon}^T = \boldsymbol{\phi}^T \mathbf{E} \mathbf{Z} (\mathbf{I} + \mathbf{E} \mathbf{Z})^{-1}$$

Requires  $m = n$  (No states added or deleted)

$$\mathbf{Z} = \begin{cases} \text{Fundamental Matrix} \\ (\mathbf{I} - \mathbf{Q})^\# \quad (\text{group inverse}) \end{cases}$$

## One Row At A Time (Sherman–Morrison)

$$\mathbf{p}_i^T = \mathbf{q}_i^T - \delta_i^T \implies \boldsymbol{\pi}^T = \boldsymbol{\phi}^T - \boldsymbol{\epsilon}^T$$

$$\boldsymbol{\epsilon}^T = \left( \frac{\phi_i}{1 + \delta_i^T \mathbf{A}_{*i}^\#} \right) \delta_i^T \mathbf{A}^\# \quad \mathbf{A}^\# = (\mathbf{I} - \mathbf{Q})^\#$$



# Exact (Theoretical) Updating

## Perturbation Formula

$$\mathbf{P} = \mathbf{Q} - \mathbf{E} \implies \boldsymbol{\pi}^T = \boldsymbol{\phi}^T - \boldsymbol{\epsilon}^T \quad \boldsymbol{\epsilon}^T = \boldsymbol{\phi}^T \mathbf{E} \mathbf{Z} (\mathbf{I} + \mathbf{E} \mathbf{Z})^{-1}$$

Requires  $m = n$  (No states added or deleted)

$$\mathbf{Z} = \begin{cases} \text{Fundamental Matrix} \\ (\mathbf{I} - \mathbf{Q})^\# \quad (\text{group inverse}) \end{cases}$$

## One Row At A Time (Sherman–Morrison)

$$\mathbf{p}_i^T = \mathbf{q}_i^T - \delta_i^T \implies \boldsymbol{\pi}^T = \boldsymbol{\phi}^T - \boldsymbol{\epsilon}^T$$

$$\boldsymbol{\epsilon}^T = \left( \frac{\phi_i}{1 + \delta_i^T \mathbf{A}_{*i}^\#} \right) \delta_i^T \mathbf{A}^\# \quad \mathbf{A}^\# = (\mathbf{I} - \mathbf{Q})^\#$$

$$(\mathbf{I} - \mathbf{P})^\# = \mathbf{A}^\# + \mathbf{e} \boldsymbol{\epsilon}^T [\mathbf{A}^\# - \gamma \mathbf{I}] - \frac{\mathbf{A}_{*i}^\# \boldsymbol{\epsilon}^T}{\phi_i} \quad \gamma = \frac{\boldsymbol{\epsilon}^T \mathbf{A}_{*i}^\#}{\phi_i}$$



# Exact (Theoretical) Updating

## Perturbation Formula

$$\mathbf{P} = \mathbf{Q} - \mathbf{E} \implies \boldsymbol{\pi}^T = \boldsymbol{\phi}^T - \boldsymbol{\epsilon}^T \quad \boldsymbol{\epsilon}^T = \boldsymbol{\phi}^T \mathbf{E} \mathbf{Z} (\mathbf{I} + \mathbf{E} \mathbf{Z})^{-1}$$

Requires  $m = n$  (No states added or deleted)

$$\mathbf{Z} = \begin{cases} \text{Fundamental Matrix} \\ (\mathbf{I} - \mathbf{Q})^\# \quad (\text{group inverse}) \end{cases}$$

## One Row At A Time (Sherman–Morrison)

$$\mathbf{p}_i^T = \mathbf{q}_i^T - \delta_i^T \implies \boldsymbol{\pi}^T = \boldsymbol{\phi}^T - \boldsymbol{\epsilon}^T$$

$$\boldsymbol{\epsilon}^T = \left( \frac{\phi_i}{1 + \delta_i^T \mathbf{A}_{*i}^\#} \right) \delta_i^T \mathbf{A}^\# \quad \mathbf{A}^\# = (\mathbf{I} - \mathbf{Q})^\#$$

$$(\mathbf{I} - \mathbf{P})^\# = \mathbf{A}^\# + \mathbf{e} \boldsymbol{\epsilon}^T [\mathbf{A}^\# - \gamma \mathbf{I}] - \frac{\mathbf{A}_{*i}^\# \boldsymbol{\epsilon}^T}{\phi_i} \quad \gamma = \frac{\boldsymbol{\epsilon}^T \mathbf{A}_{*i}^\#}{\phi_i}$$

**Not Practical For Large Problems**



# Restarted Power Method

$$\mathbf{x}_{j+1}^T = \mathbf{x}_j^T \mathbf{P} \quad \text{with} \quad \mathbf{x}_0^T = \phi^T$$

(assume aperiodic)



# Restarted Power Method

$$\mathbf{x}_{j+1}^T = \mathbf{x}_j^T \mathbf{P} \quad \text{with} \quad \mathbf{x}_0^T = \phi^T \quad (\text{assume aperiodic})$$

— Requires  $m = n$  (no states added or deleted)



# Restarted Power Method

$$\mathbf{x}_{j+1}^T = \mathbf{x}_j^T \mathbf{P} \quad \text{with} \quad \mathbf{x}_0^T = \phi^T \quad (\text{assume aperiodic})$$

- Requires  $m = n$  (no states added or deleted)
- Requires  $\phi^T \approx \pi^T$  (generally means  $\mathbf{P} \approx \mathbf{Q}$ )



# Restarted Power Method

$$\mathbf{x}_{j+1}^T = \mathbf{x}_j^T \mathbf{P} \quad \text{with} \quad \mathbf{x}_0^T = \phi^T \quad (\text{assume aperiodic})$$

- Requires  $m = n$  (no states added or deleted)
- Requires  $\phi^T \approx \pi^T$  (generally means  $\mathbf{P} \approx \mathbf{Q}$ )
- Asymptotic rate of convergence:  $R = -\log_{10} |\lambda_2|$

Need about  $1/R$  iterations to eventually gain one additional significant digit of accuracy



# Restarted Power Method

$$\mathbf{x}_{j+1}^T = \mathbf{x}_j^T \mathbf{P} \quad \text{with} \quad \mathbf{x}_0^T = \phi^T \quad (\text{assume aperiodic})$$

- Requires  $m = n$  (no states added or deleted)
- Requires  $\phi^T \approx \pi^T$  (generally means  $\mathbf{P} \approx \mathbf{Q}$ )
- Asymptotic rate of convergence:  $R = -\log_{10} |\lambda_2|$

Need about  $1/R$  iterations to eventually gain one additional significant digit of accuracy

## Example:

- ▷ Suppose  $digit_1(\phi^T) = digit_1(\pi^T)$
- ▷ Want 12 digits of accuracy





# Restarted Power Method

$$\mathbf{x}_{j+1}^T = \mathbf{x}_j^T \mathbf{P} \quad \text{with} \quad \mathbf{x}_0^T = \phi^T \quad (\text{assume aperiodic})$$

- Requires  $m = n$  (no states added or deleted)
- Requires  $\phi^T \approx \pi^T$  (generally means  $\mathbf{P} \approx \mathbf{Q}$ )
- Asymptotic rate of convergence:  $R = -\log_{10} |\lambda_2|$   
Need about  $1/R$  iterations to eventually gain one additional significant digit of accuracy

## Example:

- ▷ Suppose  $digit_1(\phi^T) = digit_1(\pi^T)$
- ▷ Want 12 digits of accuracy
- ▷ RPM — about  $11/R$  iterations
- ▷ Start from scratch — about  $12/R$  iterations



# Restarted Power Method

$$\mathbf{x}_{j+1}^T = \mathbf{x}_j^T \mathbf{P} \quad \text{with} \quad \mathbf{x}_0^T = \phi^T \quad (\text{assume aperiodic})$$

- Requires  $m = n$  (no states added or deleted)
- Requires  $\phi^T \approx \pi^T$  (generally means  $\mathbf{P} \approx \mathbf{Q}$ )
- Asymptotic rate of convergence:  $R = -\log_{10} |\lambda_2|$

Need about  $1/R$  iterations to eventually gain one additional significant digit of accuracy

## Example:

- ▷ Suppose  $digit_1(\phi^T) = digit_1(\pi^T)$
- ▷ Want 12 digits of accuracy
- ▷ RPM — about  $11/R$  iterations
- ▷ Start from scratch — about  $12/R$  iterations
- ▷ Only 8.3% reduction



# Restarted Power Method

$$\mathbf{x}_{j+1}^T = \mathbf{x}_j^T \mathbf{P} \quad \text{with} \quad \mathbf{x}_0^T = \phi^T \quad (\text{assume aperiodic})$$

- Requires  $m = n$  (no states added or deleted)
- Requires  $\phi^T \approx \pi^T$  (generally means  $\mathbf{P} \approx \mathbf{Q}$ )
- Asymptotic rate of convergence:  $R = -\log_{10} |\lambda_2|$

Need about  $1/R$  iterations to eventually gain one additional significant digit of accuracy

Example:

- ▷ Suppose  $digit_1(\phi^T) = digit_1(\pi^T)$
- ▷ Want 12 digits of accuracy
- ▷ RPM — about  $11/R$  iterations
- ▷ Start from scratch — about  $12/R$  iterations
- ▷ Only 8.3% reduction

**A Little Better, But Not Great**



# Censoring

Partition (not necessarily NCD)

$$\mathbf{P}_{n \times n} = \begin{matrix} & G_1 & G_2 & \cdots & G_k \\ G_1 & \mathbf{P}_{11} & \mathbf{P}_{12} & \cdots & \mathbf{P}_{1k} \\ G_2 & \mathbf{P}_{21} & \mathbf{P}_{22} & \cdots & \mathbf{P}_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ G_k & \mathbf{P}_{k1} & \mathbf{P}_{k2} & \cdots & \mathbf{P}_{kk} \end{matrix}$$



# Censoring

Partition (not necessarily NCD)

$$\mathbf{P}_{n \times n} = \begin{matrix} & G_1 & G_2 & \cdots & G_k \\ \begin{matrix} G_1 \\ G_2 \\ \vdots \\ G_k \end{matrix} & \begin{pmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} & \cdots & \mathbf{P}_{1k} \\ \mathbf{P}_{21} & \mathbf{P}_{22} & \cdots & \mathbf{P}_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{P}_{k1} & \mathbf{P}_{k2} & \cdots & \mathbf{P}_{kk} \end{pmatrix} \end{matrix}$$

## Censored Chains

Records state of process only when chain visits states in  $G_i$ .

Visits to states outside of  $G_i$  are ignored.



# Censoring

Partition (not necessarily NCD)

$$\mathbf{P}_{n \times n} = \begin{matrix} & G_1 & G_2 & \cdots & G_k \\ \begin{matrix} G_1 \\ G_2 \\ \vdots \\ G_k \end{matrix} & \begin{pmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} & \cdots & \mathbf{P}_{1k} \\ \mathbf{P}_{21} & \mathbf{P}_{22} & \cdots & \mathbf{P}_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{P}_{k1} & \mathbf{P}_{k2} & \cdots & \mathbf{P}_{kk} \end{pmatrix} \end{matrix}$$

## Censored Chains

Records state of process only when chain visits states in  $G_i$ .

Visits to states outside of  $G_i$  are ignored.

## Censored Transition Matrices

$$\mathbf{C}_i = \mathbf{P}_{ii} + \mathbf{P}_{i\star}(\mathbf{I} - \mathbf{P}_i^{\star})^{-1}\mathbf{P}_{\star i}$$

Stochastic Complements



# Aggregation

## Censored Distributions

$$\mathbf{s}_i^T \mathbf{C}_i = \mathbf{s}_i^T$$



# Aggregation

## Censored Distributions

$$\mathbf{s}_i^T \mathbf{C}_i = \mathbf{s}_i^T$$

## Aggregation Matrix

$$\mathcal{A}_{k \times k} = \begin{bmatrix} \mathbf{s}_1^T \mathbf{P}_{11} \mathbf{e} & \cdots & \mathbf{s}_1^T \mathbf{P}_{1k} \mathbf{e} \\ \vdots & \ddots & \vdots \\ \mathbf{s}_k^T \mathbf{P}_{k1} \mathbf{e} & \cdots & \mathbf{s}_k^T \mathbf{P}_{kk} \mathbf{e} \end{bmatrix}$$

$$\mathbf{e} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$





# Aggregation

## Censored Distributions

$$\mathbf{s}_i^T \mathbf{C}_i = \mathbf{s}_i^T$$

## Aggregation Matrix

$$\mathcal{A}_{k \times k} = \begin{bmatrix} \mathbf{s}_1^T \mathbf{P}_{11} \mathbf{e} & \cdots & \mathbf{s}_1^T \mathbf{P}_{1k} \mathbf{e} \\ \vdots & \ddots & \vdots \\ \mathbf{s}_k^T \mathbf{P}_{k1} \mathbf{e} & \cdots & \mathbf{s}_k^T \mathbf{P}_{kk} \mathbf{e} \end{bmatrix} \quad \mathbf{e} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

## Aggregated Distribution

$$\boldsymbol{\alpha}^T \mathcal{A} = \boldsymbol{\alpha}^T \quad \boldsymbol{\alpha}^T = (\alpha_1, \alpha_2, \dots, \alpha_k)$$



# Aggregation

## Censored Distributions

$$\mathbf{s}_i^T \mathbf{C}_i = \mathbf{s}_i^T$$

## Aggregation Matrix

$$\mathcal{A}_{k \times k} = \begin{bmatrix} \mathbf{s}_1^T \mathbf{P}_{11} \mathbf{e} & \cdots & \mathbf{s}_1^T \mathbf{P}_{1k} \mathbf{e} \\ \vdots & \ddots & \vdots \\ \mathbf{s}_k^T \mathbf{P}_{k1} \mathbf{e} & \cdots & \mathbf{s}_k^T \mathbf{P}_{kk} \mathbf{e} \end{bmatrix} \quad \mathbf{e} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

## Aggregated Distribution

$$\alpha^T \mathcal{A} = \alpha^T \quad \alpha^T = (\alpha_1, \alpha_2, \dots, \alpha_k)$$

## Aggregation Theorem

$$\boldsymbol{\pi}^T = (\boldsymbol{\pi}_1^T \mid \boldsymbol{\pi}_2^T \mid \cdots \mid \boldsymbol{\pi}_k^T) = (\alpha_1 \mathbf{s}_1^T \mid \alpha_2 \mathbf{s}_2^T \mid \cdots \mid \alpha_k \mathbf{s}_k^T)$$



# Partitioning

## Intuition

Update relatively small number of states in large sparse chain

- Effects are primarily local
- Most stationary probabilities not significantly affected.



# Partitioning

## Intuition

Update relatively small number of states in large sparse chain

- Effects are primarily local
- Most stationary probabilities not significantly affected.

## State Space Partition

$$\mathcal{S} = G \cup \bar{G} \quad G = \begin{cases} \text{states likely to be most affected} \\ \text{newly added states} \end{cases}$$



# Partitioning

## Intuition

Update relatively small number of states in large sparse chain

- Effects are primarily local
- Most stationary probabilities not significantly affected.

## State Space Partition

$$\mathcal{S} = G \cup \overline{G} \quad G = \begin{cases} \text{states likely to be most affected} \\ \text{newly added states} \end{cases}$$

$$— \quad g = |G| \ll |\overline{G}|$$



# Partitioning

## Intuition

Update relatively small number of states in large sparse chain

- Effects are primarily local
- Most stationary probabilities not significantly affected.

## State Space Partition

$$\mathcal{S} = G \cup \overline{G} \quad G = \begin{cases} \text{states likely to be most affected} \\ \text{newly added states} \end{cases}$$

$$— g = |G| \ll |\overline{G}|$$

## Induced Matrix Partition

$$\mathbf{P}_{n \times n} = \begin{matrix} & G & \overline{G} \\ \begin{matrix} G \\ \overline{G} \end{matrix} & \begin{pmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{pmatrix} \end{matrix} = \begin{bmatrix} p_{11} & \cdots & p_{1g} & \mathbf{P}_{1\star} \\ \vdots & & \vdots & \vdots \\ p_{g1} & \cdots & p_{gg} & \mathbf{P}_{g\star} \\ \mathbf{P}_{\star 1} & \cdots & \mathbf{P}_{\star g} & \mathbf{P}_{22} \end{bmatrix}$$



# Specialized Aggregation

## Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1 \times 1} \quad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$



# Specialized Aggregation

## Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1 \times 1} \quad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

## Censored Distributions

$$\mathbf{s}_1^T = \cdots = \mathbf{s}_g^T = 1 \quad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T \mathbf{C}_{g+1}$$





# Specialized Aggregation

## Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1 \times 1} \quad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

## Censored Distributions

$$\mathbf{s}_1^T = \cdots = \mathbf{s}_g^T = 1 \quad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T \mathbf{C}_{g+1}$$

## Aggregation Matrix

$$\mathcal{A} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}_{g+1}^T \mathbf{P}_{21} & 1 - \mathbf{s}_{g+1}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1) \times (g+1)}$$



# Specialized Aggregation

## Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1 \times 1} \quad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

## Censored Distributions

$$\mathbf{s}_1^T = \cdots = \mathbf{s}_g^T = \mathbf{1} \quad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T \mathbf{C}_{g+1}$$

## Aggregation Matrix

$$\mathcal{A} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}_{g+1}^T \mathbf{P}_{21} & \mathbf{1} - \mathbf{s}_{g+1}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1) \times (g+1)}$$

## Aggregated Distribution

$$\boldsymbol{\alpha}^T = (\alpha_1, \dots, \alpha_g, \alpha_{g+1})$$

## Aggregation Theorem

$$\boldsymbol{\pi}^T = (\pi_1, \dots, \pi_g \mid \bar{\boldsymbol{\pi}}^T) = (\alpha_1, \dots, \alpha_g, \mid \alpha_{g+1} \mathbf{s}_{g+1}^T)$$



# Specialized Aggregation

## Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1 \times 1} \quad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

## Censored Distributions

$$\mathbf{s}_1^T = \cdots = \mathbf{s}_g^T = 1 \quad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T \mathbf{C}_{g+1}$$

## Aggregation Matrix

$$\mathcal{A} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}_{g+1}^T \mathbf{P}_{21} & 1 - \mathbf{s}_{g+1}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1) \times (g+1)}$$

## Aggregated Distribution

$$\boldsymbol{\alpha}^T = (\alpha_1, \dots, \alpha_g, \alpha_{g+1})$$

## Aggregation Theorem

$$\boldsymbol{\pi}^T = (\pi_1, \dots, \pi_g \mid \bar{\boldsymbol{\pi}}^T) = (\alpha_1, \dots, \alpha_g, \mid \alpha_{g+1} \mathbf{s}_{g+1}^T)$$

## Old Distribution (reordered)

$$\boldsymbol{\phi}^T = (\phi_1, \phi_2, \dots \mid \bar{\boldsymbol{\phi}}^T)$$



# Specialized Aggregation

## Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1 \times 1} \quad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

## Censored Distributions

$$\mathbf{s}_1^T = \cdots = \mathbf{s}_g^T = 1 \quad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T \mathbf{C}_{g+1}$$

## Aggregation Matrix

$$\mathcal{A} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}_{g+1}^T \mathbf{P}_{21} & 1 - \mathbf{s}_{g+1}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1) \times (g+1)}$$

## Aggregated Distribution

$$\alpha^T = (\alpha_1, \dots, \alpha_g, \alpha_{g+1})$$

## Aggregation Theorem

$$\pi^T = (\pi_1, \dots, \pi_g \mid \bar{\pi}^T) = (\alpha_1, \dots, \alpha_g, \mid \alpha_{g+1} \mathbf{s}_{g+1}^T)$$

## Old Distribution (reordered)

$$\phi^T = (\phi_1, \phi_2, \dots \mid \bar{\phi}^T)$$

## The Assumption

$$\bar{\phi}^T \approx \bar{\pi}^T$$



# Specialized Aggregation

## Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1 \times 1} \quad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

## Censored Distributions

$$\mathbf{s}_1^T = \cdots = \mathbf{s}_g^T = 1 \quad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T \mathbf{C}_{g+1}$$

## Aggregation Matrix

$$\mathcal{A} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}_{g+1}^T \mathbf{P}_{21} & 1 - \mathbf{s}_{g+1}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1) \times (g+1)}$$

## Aggregated Distribution

$$\alpha^T = (\alpha_1, \dots, \alpha_g, \alpha_{g+1})$$

## Aggregation Theorem

$$\pi^T = (\pi_1, \dots, \pi_g \mid \bar{\pi}^T) = (\alpha_1, \dots, \alpha_g, \mid \alpha_{g+1} \mathbf{s}_{g+1}^T)$$

## Old Distribution (reordered)

$$\phi^T = (\phi_1, \phi_2, \dots \mid \bar{\phi}^T)$$

## The Assumption

$$\bar{\phi}^T \approx \bar{\pi}^T \implies \bar{\phi}^T \approx \alpha_{g+1} \mathbf{s}_{g+1}^T$$



# Specialized Aggregation

## Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1 \times 1} \quad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

## Censored Distributions

$$\mathbf{s}_1^T = \cdots = \mathbf{s}_g^T = 1 \quad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T \mathbf{C}_{g+1}$$

## Aggregation Matrix

$$\mathcal{A} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}_{g+1}^T \mathbf{P}_{21} & 1 - \mathbf{s}_{g+1}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1) \times (g+1)}$$

## Aggregated Distribution

$$\alpha^T = (\alpha_1, \dots, \alpha_g, \alpha_{g+1})$$

## Aggregation Theorem

$$\pi^T = (\pi_1, \dots, \pi_g \mid \bar{\pi}^T) = (\alpha_1, \dots, \alpha_g, \mid \alpha_{g+1} \mathbf{s}_{g+1}^T)$$

## Old Distribution (reordered)

$$\phi^T = (\phi_1, \phi_2, \dots \mid \bar{\phi}^T)$$

## The Assumption

$$\bar{\phi}^T \approx \bar{\pi}^T \implies \bar{\phi}^T \approx \alpha_{g+1} \mathbf{s}_{g+1}^T \implies \mathbf{s}_{g+1}^T \approx \frac{\bar{\phi}^T}{\bar{\phi}^T \mathbf{e}}$$



# Specialized Aggregation

## Censored Transition Matrices

$$\mathbf{C}_1 = \dots = \mathbf{C}_g = [1]_{1 \times 1} \quad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

## Censored Distributions

$$\mathbf{s}_1^T = \dots = \mathbf{s}_g^T = 1 \quad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T \mathbf{C}_{g+1}$$

## Aggregation Matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}_{g+1}^T \mathbf{P}_{21} & 1 - \mathbf{s}_{g+1}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1) \times (g+1)}$$

## Aggregated Distribution

$$\alpha^T = (\alpha_1, \dots, \alpha_g, \alpha_{g+1})$$

## Aggregation Theorem

$$\pi^T = (\pi_1, \dots, \pi_g \mid \bar{\pi}^T) = (\alpha_1, \dots, \alpha_g, \mid \alpha_{g+1} \mathbf{s}_{g+1}^T)$$

## Old Distribution (reordered)

$$\phi^T = (\phi_1, \phi_2, \dots \mid \bar{\phi}^T)$$

## The Assumption

$$\bar{\phi}^T \approx \bar{\pi}^T \implies \bar{\phi}^T \approx \alpha_{g+1} \mathbf{s}_{g+1}^T \implies \mathbf{s}_{g+1}^T \approx \frac{\bar{\phi}^T}{\phi^T \mathbf{e}}$$



# Specialized Aggregation

## Censored Transition Matrices

$$\mathbf{C}_1 = \dots = \mathbf{C}_g = [1]_{1 \times 1} \quad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

## Censored Distributions

$$\mathbf{s}_1^T = \dots = \mathbf{s}_g^T = \mathbf{1} \quad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T \mathbf{C}_{g+1}$$

## Aggregation Matrix

$$\mathcal{A} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}_{g+1}^T \mathbf{P}_{21} & \mathbf{1} - \mathbf{s}_{g+1}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1) \times (g+1)}$$

## Aggregated Distribution

$$\longrightarrow \boldsymbol{\alpha}^T = (\alpha_1, \dots, \alpha_g, \alpha_{g+1})$$

## Aggregation Theorem

$$\boldsymbol{\pi}^T = (\pi_1, \dots, \pi_g \mid \bar{\boldsymbol{\pi}}^T) = (\alpha_1, \dots, \alpha_g \mid \alpha_{g+1} \mathbf{s}_{g+1}^T)$$

## Old Distribution (reordered)

$$\boldsymbol{\phi}^T = (\phi_1, \phi_2, \dots \mid \bar{\boldsymbol{\phi}}^T)$$

## The Assumption

$$\bar{\boldsymbol{\phi}}^T \approx \bar{\boldsymbol{\pi}}^T \implies \bar{\boldsymbol{\phi}}^T \approx \alpha_{g+1} \mathbf{s}_{g+1}^T \implies \mathbf{s}_{g+1}^T \approx \frac{\bar{\boldsymbol{\phi}}^T}{\boldsymbol{\phi}^T \mathbf{e}}$$





# Specialized Aggregation

## Censored Transition Matrices

$$\mathbf{C}_1 = \dots = \mathbf{C}_g = [1]_{1 \times 1} \quad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

## Censored Distributions

$$\mathbf{s}_1^T = \dots = \mathbf{s}_g^T = 1 \quad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T \mathbf{C}_{g+1}$$

## Aggregation Matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}_{g+1}^T \mathbf{P}_{21} & 1 - \mathbf{s}_{g+1}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1) \times (g+1)}$$

## Aggregated Distribution

$$\boldsymbol{\alpha}^T = (\alpha_1, \dots, \alpha_g, \alpha_{g+1})$$

## Aggregation Theorem

$$\boldsymbol{\pi}^T = (\alpha_1, \dots, \alpha_g, | \alpha_{g+1} \mathbf{s}_{g+1}^T)$$

## Old Distribution (reordered)

$$\boldsymbol{\phi}^T = (\phi_1, \phi_2, \dots | \bar{\boldsymbol{\phi}}^T)$$

## Updated Distribution

## The Assumption

$$\bar{\boldsymbol{\phi}}^T \approx \boldsymbol{\pi}^T \implies \bar{\boldsymbol{\phi}}^T \approx \alpha_{g+1} \mathbf{s}_{g+1}^T \implies \mathbf{s}_{g+1}^T \approx \frac{\bar{\boldsymbol{\phi}}^T}{\boldsymbol{\phi}^T \mathbf{e}}$$



# Summary

## Reorder & Partition Updated State Space

$$\mathcal{S} = G \cup \overline{G}$$

$$G = \{\text{New} + \text{Most affected}\} \quad \overline{G} = \{\text{Less affected}\}$$



# Summary

## Reorder & Partition Updated State Space

$$\mathcal{S} = G \cup \bar{G} \quad G = \{\text{New} + \text{Most affected}\} \quad \bar{G} = \{\text{Less affected}\}$$

## Use Less Affected Components From Old Distribution

$\bar{\phi}^T \leftarrow$  components from  $\phi^T$  corresponding to states in  $\bar{G}$



# Summary

## Reorder & Partition Updated State Space

$$\mathcal{S} = G \cup \bar{G} \quad G = \{\text{New} + \text{Most affected}\} \quad \bar{G} = \{\text{Less affected}\}$$

## Use Less Affected Components From Old Distribution

$$\bar{\phi}^T \leftarrow \text{components from } \phi^T \text{ corresponding to states in } \bar{G}$$

## Approximate Censored Distribution

$$\mathbf{s}^T \leftarrow \bar{\phi}^T / (\bar{\phi}^T \mathbf{e})$$



# Summary

## Reorder & Partition Updated State Space

$$\mathcal{S} = G \cup \bar{G} \quad G = \{\text{New} + \text{Most affected}\} \quad \bar{G} = \{\text{Less affected}\}$$

## Use Less Affected Components From Old Distribution

$$\bar{\phi}^T \leftarrow \text{components from } \phi^T \text{ corresponding to states in } \bar{G}$$

## Approximate Censored Distribution

$$\mathbf{s}^T \leftarrow \bar{\phi}^T / (\bar{\phi}^T \mathbf{e})$$

## Form Approximate Aggregation Matrix

$$\mathcal{A} \leftarrow \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}^T \mathbf{P}_{21} & \mathbf{1} - \mathbf{s}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1) \times (g+1)}$$



# Summary

## Reorder & Partition Updated State Space

$$\mathcal{S} = G \cup \bar{G} \quad G = \{\text{New} + \text{Most affected}\} \quad \bar{G} = \{\text{Less affected}\}$$

## Use Less Affected Components From Old Distribution

$$\bar{\phi}^T \leftarrow \text{components from } \phi^T \text{ corresponding to states in } \bar{G}$$

## Approximate Censored Distribution

$$\mathbf{s}^T \leftarrow \bar{\phi}^T / (\bar{\phi}^T \mathbf{e})$$

## Form Approximate Aggregation Matrix

$$\mathcal{A} \leftarrow \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}^T \mathbf{P}_{21} & 1 - \mathbf{s}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1) \times (g+1)}$$

## Compute Approximate Aggregated Distribution

$$\alpha^T \leftarrow (\alpha_1, \dots, \alpha_g, \alpha_{g+1})$$



# Summary

## Reorder & Partition Updated State Space

$$\mathcal{S} = G \cup \bar{G} \quad G = \{\text{New} + \text{Most affected}\} \quad \bar{G} = \{\text{Less affected}\}$$

## Use Less Affected Components From Old Distribution

$$\bar{\phi}^T \leftarrow \text{components from } \phi^T \text{ corresponding to states in } \bar{G}$$

## Approximate Censored Distribution

$$\mathbf{s}^T \leftarrow \bar{\phi}^T / (\bar{\phi}^T \mathbf{e})$$

## Form Approximate Aggregation Matrix

$$\mathcal{A} \leftarrow \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}^T \mathbf{P}_{21} & 1 - \mathbf{s}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1) \times (g+1)}$$

## Compute Approximate Aggregated Distribution

$$\alpha^T \leftarrow (\alpha_1, \dots, \alpha_g, \alpha_{g+1})$$

## Approximate Updated Distribution

$$\pi^T \leftarrow (\alpha_1, \dots, \alpha_g, | \alpha_{g+1} \mathbf{s}^T)$$



# Summary

## Reorder & Partition Updated State Space

$$\mathcal{S} = G \cup \bar{G} \quad G = \{\text{New + Most affected}\} \quad \bar{G} = \{\text{Less affected}\}$$

## Use Less Affected Components From Old Distribution

$\bar{\phi}^T \leftarrow$  components from  $\phi^T$  corresponding to states in  $\bar{G}$

## Approximate Censored Distribution

$$\mathbf{s}^T \leftarrow \bar{\phi}^T / (\bar{\phi}^T \mathbf{e})$$

## Form Approximate Aggregation Matrix

$$\mathcal{A} \leftarrow \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}^T \mathbf{P}_{21} & 1 - \mathbf{s}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1) \times (g+1)}$$

## Compute Approximate Aggregated Distribution

$$\alpha^T \leftarrow (\alpha_1, \dots, \alpha_g, \alpha_{g+1})$$

## Approximate Updated Distribution

$$\pi^T \leftarrow (\alpha_1, \dots, \alpha_g, \alpha_{g+1} \mathbf{s}^T)$$

**Iterate ?**





# Summary

## Reorder & Partition Updated State Space

$$\mathcal{S} = G \cup \bar{G} \quad G = \{\text{New + Most affected}\} \quad \bar{G} = \{\text{Less affected}\}$$

## Use Less Affected Components From Old Distribution

$$\bar{\phi}^T \leftarrow \text{components from } \phi^T \text{ corresponding to states in } \bar{G}$$

## Approximate Censored Distribution

$$\mathbf{s}^T \leftarrow \bar{\phi}^T / (\bar{\phi}^T \mathbf{e})$$

## Form Approximate Aggregation Matrix

$$\mathcal{A} \leftarrow \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}^T \mathbf{P}_{21} & 1 - \mathbf{s}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1) \times (g+1)}$$

## Compute Approximate Aggregated Distribution

$$\alpha^T \leftarrow (\alpha_1, \dots, \alpha_g, \alpha_{g+1})$$

## Approximate Updated Distribution

$$\pi^T \leftarrow (\alpha_1, \dots, \alpha_g, \alpha_{g+1} \mathbf{s}^T)$$

Iterate ?

No! — At A Fixed Point



# Iterative Aggregation

## Move Off Of Fixed Point With Power Step

$$\mathbf{s}^T \leftarrow \overline{\phi}^T / (\overline{\phi}^T \mathbf{e})$$

$$\mathcal{A} \leftarrow \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}^T \mathbf{P}_{21} & 1 - \mathbf{s}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1) \times (g+1)}$$

$$\boldsymbol{\alpha}^T \leftarrow (\alpha_1, \dots, \alpha_g, \alpha_{g+1})$$

$$\boldsymbol{\pi}^T \leftarrow (\alpha_1, \dots, \alpha_g, | \alpha_{g+1} \mathbf{s}^T)$$

$$\boldsymbol{\psi}^T \leftarrow \boldsymbol{\pi}^T \mathbf{P}$$

(Also makes progress toward convergence when aperiodic)

If  $\|\boldsymbol{\psi}^T - \boldsymbol{\chi}^T\| < \tau$  then quit — else

$$\mathbf{s}^T \leftarrow \overline{\boldsymbol{\psi}^T} / (\overline{\boldsymbol{\psi}^T} \mathbf{e})$$



# Iterative Aggregation

Move Off Of Fixed Point With Power Step

$$\mathbf{s}^T \leftarrow \overline{\phi}^T / (\overline{\phi}^T \mathbf{e})$$

$$\mathcal{A} \leftarrow \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}^T \mathbf{P}_{21} & 1 - \mathbf{s}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1) \times (g+1)}$$

$$\boldsymbol{\alpha}^T \leftarrow (\alpha_1, \dots, \alpha_g, \alpha_{g+1})$$

$$\boldsymbol{\pi}^T \leftarrow (\alpha_1, \dots, \alpha_g, | \alpha_{g+1} \mathbf{s}^T)$$

$$\boldsymbol{\psi}^T \leftarrow \boldsymbol{\pi}^T \mathbf{P}$$

(Also makes progress toward convergence when aperiodic)

If  $\|\boldsymbol{\psi}^T - \boldsymbol{\chi}^T\| < \tau$  then quit — else

$$\mathbf{s}^T \leftarrow \overline{\boldsymbol{\psi}^T} / (\overline{\boldsymbol{\psi}^T} \mathbf{e})$$

## Theorem

If  $\mathbf{C} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$  is aperiodic, then convergent for all partitions  $\mathcal{S} = G \cup \overline{G}$



# Iterative Aggregation

Move Off Of Fixed Point With Power Step

$$\mathbf{s}^T \leftarrow \overline{\phi}^T / (\overline{\phi}^T \mathbf{e})$$

$$\mathcal{A} \leftarrow \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}^T \mathbf{P}_{21} & 1 - \mathbf{s}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1) \times (g+1)}$$

$$\boldsymbol{\alpha}^T \leftarrow (\alpha_1, \dots, \alpha_g, \alpha_{g+1})$$

$$\boldsymbol{\pi}^T \leftarrow (\alpha_1, \dots, \alpha_g, | \alpha_{g+1} \mathbf{s}^T)$$

$$\boldsymbol{\psi}^T \leftarrow \boldsymbol{\pi}^T \mathbf{P}$$

(Also makes progress toward convergence when aperiodic)

If  $\|\boldsymbol{\psi}^T - \boldsymbol{\chi}^T\| < \tau$  then quit — else

$$\mathbf{s}^T \leftarrow \overline{\boldsymbol{\psi}^T} / (\overline{\boldsymbol{\psi}^T} \mathbf{e})$$

## Theorem

If  $\mathbf{C} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$  is aperiodic, then convergent for all partitions  $\mathcal{S} = G \cup \overline{G}$

—  $|\lambda_2(\mathbf{C})|$  determines rate of convergence



# Google's PageRank

## Random Walk On WWW Link Structure

$$\mathbf{H}_{ij} = \begin{cases} 1/(\text{total \# outlinks from page } \mathcal{P}_i) & \text{if } \mathcal{P}_i \rightarrow \mathcal{P}_j, \\ 0 & \text{otherwise} \end{cases}$$



# Google's PageRank

## Random Walk On WWW Link Structure

$$\mathbf{H}_{ij} = \begin{cases} 1/(\text{total \# outlinks from page } \mathcal{P}_i) & \text{if } \mathcal{P}_i \rightarrow \mathcal{P}_j, \\ 0 & \text{otherwise} \end{cases}$$

## Google Matrix

$$\mathbf{P} = \alpha(\mathbf{H} + \mathbf{E}) + (1 - \alpha)\mathbf{F}$$

- $(\mathbf{H} + \mathbf{E})$  &  $\mathbf{F}$  are stochastic  $\text{rank}(\mathbf{E}) = \text{rank}(\mathbf{F}) = 1$
- $0 < \alpha < 1$



# Google's PageRank

## Random Walk On WWW Link Structure

$$\mathbf{H}_{ij} = \begin{cases} 1/(\text{total \# outlinks from page } \mathcal{P}_i) & \text{if } \mathcal{P}_i \rightarrow \mathcal{P}_j, \\ 0 & \text{otherwise} \end{cases}$$

## Google Matrix

$$\mathbf{P} = \alpha(\mathbf{H} + \mathbf{E}) + (1 - \alpha)\mathbf{F}$$

—  $(\mathbf{H} + \mathbf{E})$  &  $\mathbf{F}$  are stochastic  $\text{rank}(\mathbf{E}) = \text{rank}(\mathbf{F}) = 1$

—  $0 < \alpha < 1$

— PageRank =  $\pi^T$



# Google's PageRank

## Random Walk On WWW Link Structure

$$\mathbf{H}_{ij} = \begin{cases} 1/(\text{total \# outlinks from page } \mathcal{P}_i) & \text{if } \mathcal{P}_i \rightarrow \mathcal{P}_j, \\ 0 & \text{otherwise} \end{cases}$$

## Google Matrix

$$\mathbf{P} = \alpha(\mathbf{H} + \mathbf{E}) + (1 - \alpha)\mathbf{F}$$

- $(\mathbf{H} + \mathbf{E})$  &  $\mathbf{F}$  are stochastic  $\text{rank}(\mathbf{E}) = \text{rank}(\mathbf{F}) = 1$
- $0 < \alpha < 1$
- PageRank =  $\pi^T$

## Power Law Distribution

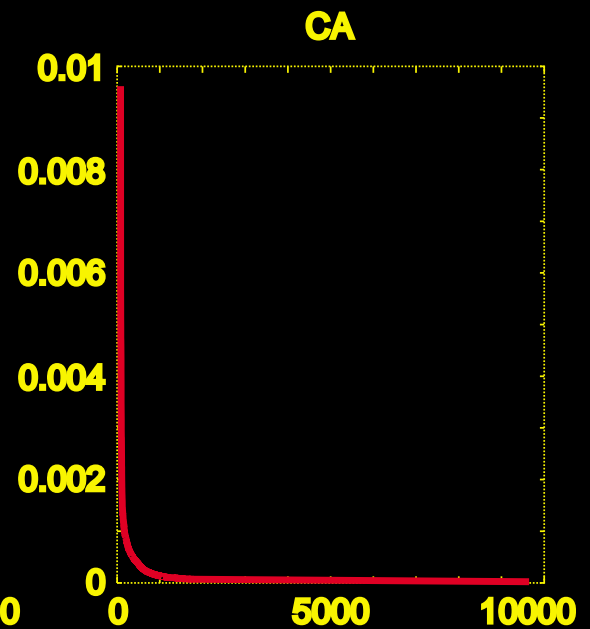
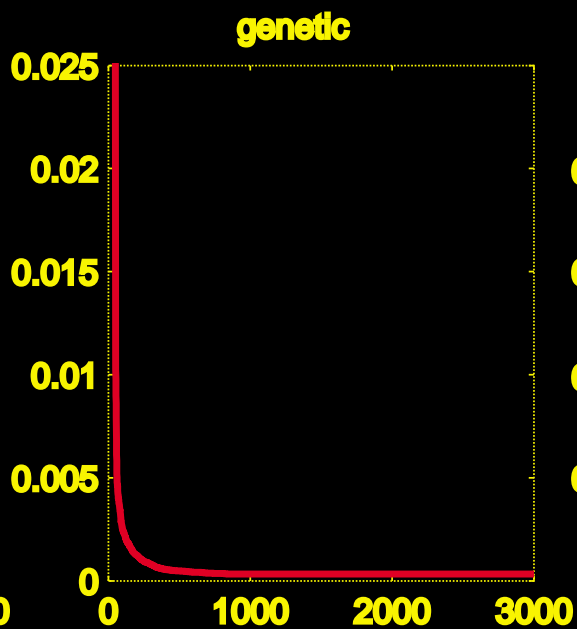
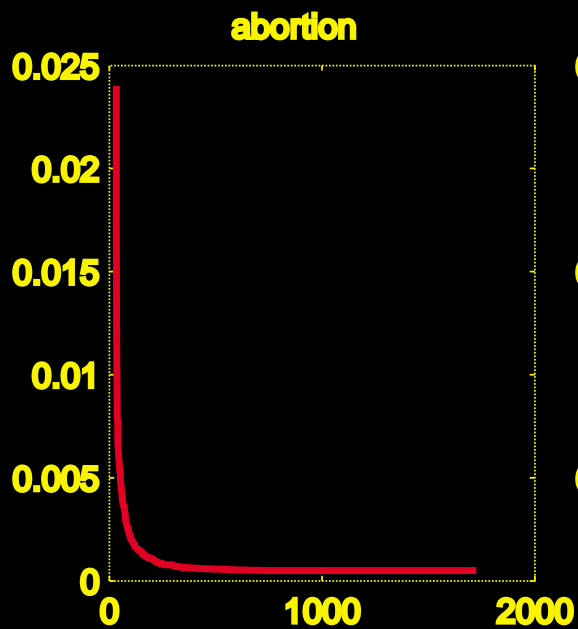
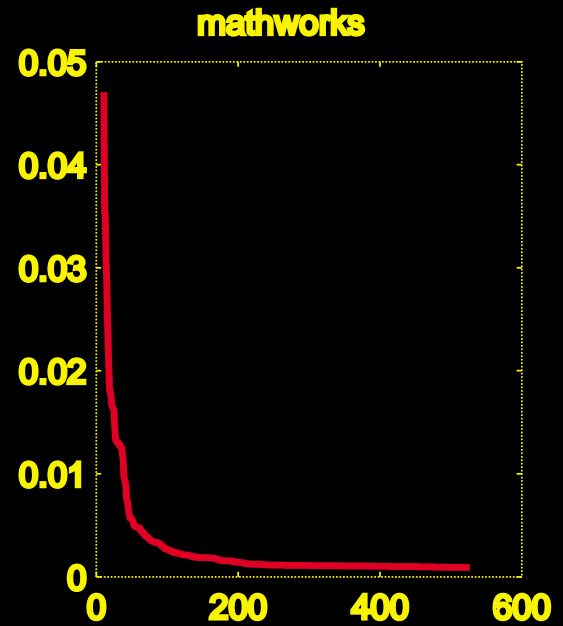
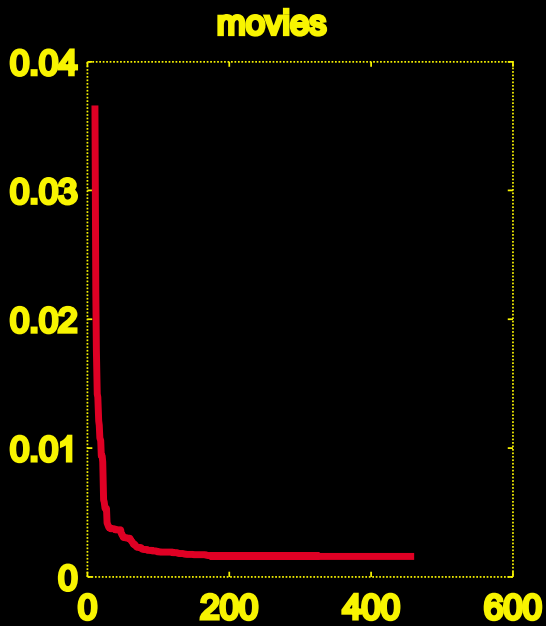
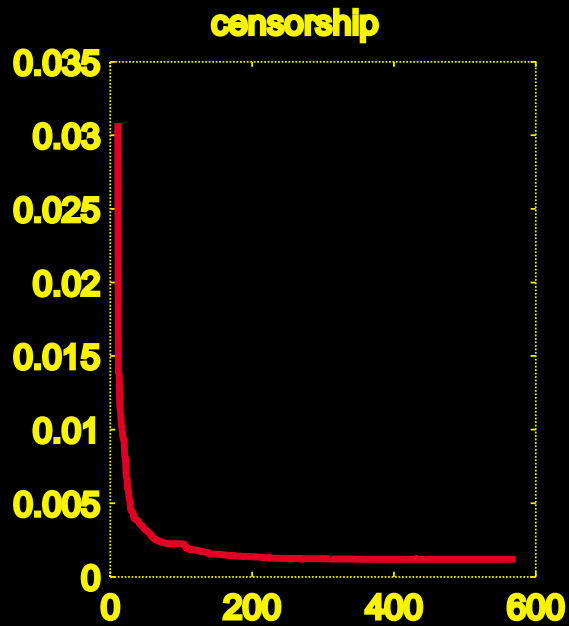
If ordered by magnitude  $\pi(1) \geq \pi(2) \geq \dots \geq \pi(n)$ , then

- $\pi(i) \approx \alpha i^{-k}$  for  $k \approx 2.109$  [Donato, Laura, Leonardi, 2002]  
[Pandurangan, Raghavan, & Upfal, 2004]
- Relatively few large states (i.e., important sites)





# “L” Curves





# Experiments

## The Updates

# Nodes Added = 3

# Nodes Removed = 50

# Links Added = 10

(Different values have little effect on results)

# Links Removed = 20

## Stopping Criterion

1-norm of residual  $< 10^{-10}$



# Movies

## Power Method

<u>Iterations</u>	<u>Time</u>
17	.40

## Iterative Aggregation

<u><math> G </math></u>	<u>Iterations</u>	<u>Time</u>
5	12	.39
10	12	.37
15	11	.36
20	11	.35
<b>25</b>	<b>11</b>	<b>.31</b>
<b>50</b>	<b>9</b>	<b>.31</b>
100	9	.33
200	8	.35
300	7	.39
400	6	.47

*nodes = 451 links = 713*



# Censorship

## Power Method

Iterations	Time
38	1.40

## Iterative Aggregation

$ G $	Iterations	Time
5	38	1.68
10	38	1.66
15	38	1.56
20	20	1.06
25	20	1.05
50	10	.69
100	8	.55
<b>200</b>	<b>6</b>	<b>.53</b>
300	6	.65
400	5	.70

*nodes = 562 links = 736*



# MathWorks

## Power Method

<u>Iterations</u>	<u>Time</u>
54	1.25

## Iterative Aggregation

<u><math> G </math></u>	<u>Iterations</u>	<u>Time</u>
5	53	1.18
10	52	1.29
15	52	1.23
20	42	1.05
25	20	1.13
<b>50</b>	<b>18</b>	<b>.70</b>
<b>100</b>	<b>16</b>	<b>.70</b>
<b>200</b>	<b>13</b>	<b>.70</b>
300	11	.83
400	10	1.01

*nodes = 517    links = 13,531*



# Abortion

## Power Method

<u>Iterations</u>	<u>Time</u>
106	37.08

## Iterative Aggregation

<u><math> G </math></u>	<u>Iterations</u>	<u>Time</u>
5	109	38.56
10	105	36.02
15	107	38.05
20	107	38.45
25	97	34.81
50	53	18.80
<b>100</b>	<b>13</b>	<b>5.18</b>
250	12	5.62
500	6	5.21
750	5	10.22
1000	5	14.61

*nodes = 1,693    links = 4,325*



# Genetics

## Power Method

<u>Iterations</u>	<u>Time</u>
92	91.78

## Iterative Aggregation

<u><math> G </math></u>	<u>Iterations</u>	<u>Time</u>
5	91	88.22
10	92	92.12
20	71	72.53
50	25	25.42
100	19	20.72
250	13	14.97
<b>500</b>	<b>7</b>	<b>11.14</b>
1000	5	17.76
1500	5	31.84

*nodes = 2,952    links = 6,485*



# California

## Power Method

<u>Iterations</u>	<u>Time</u>
176	5.85

## Iterative Aggregation

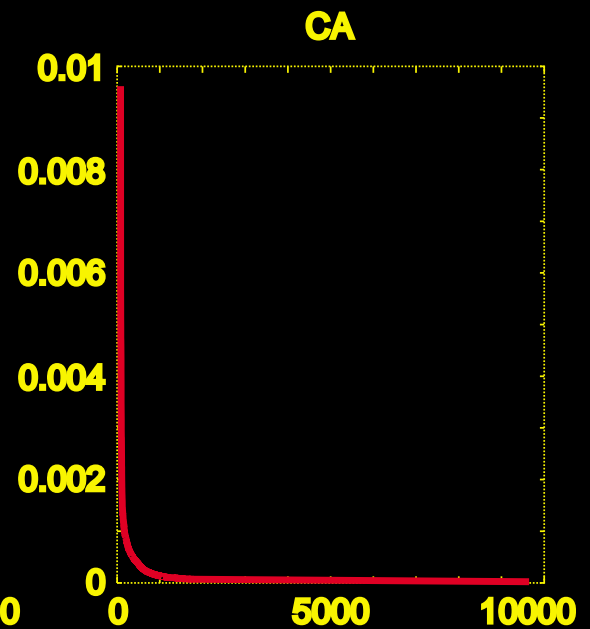
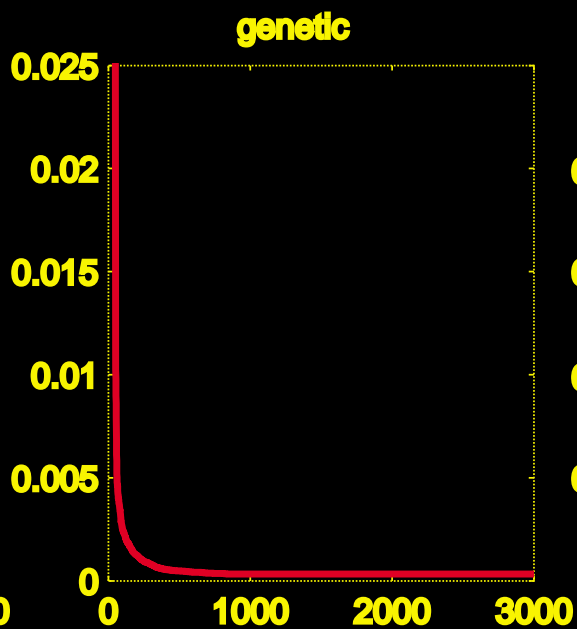
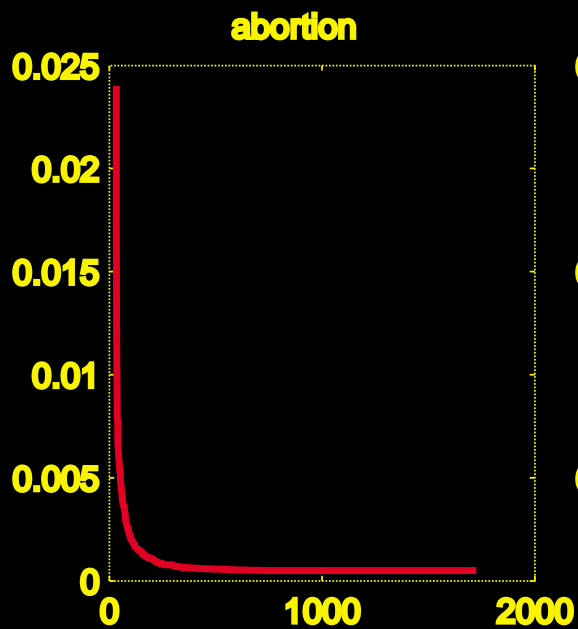
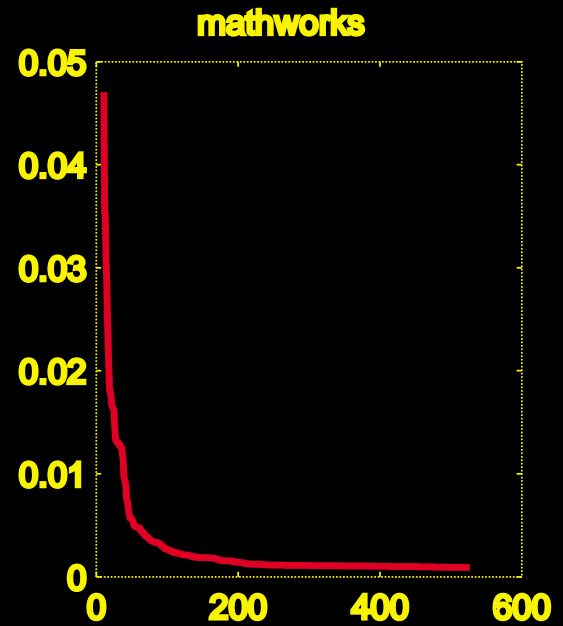
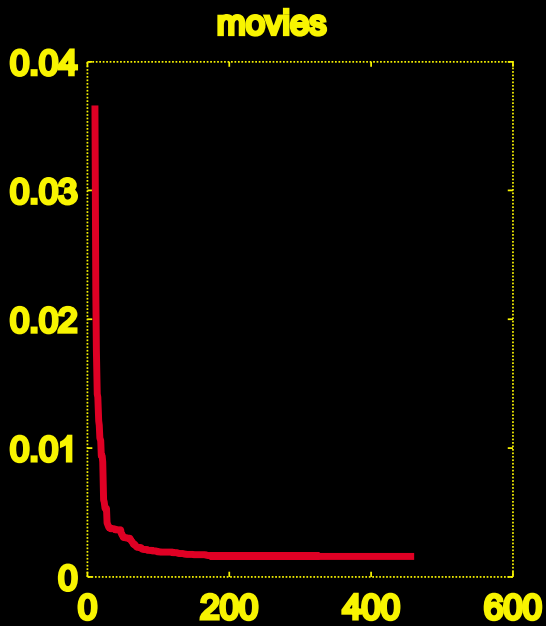
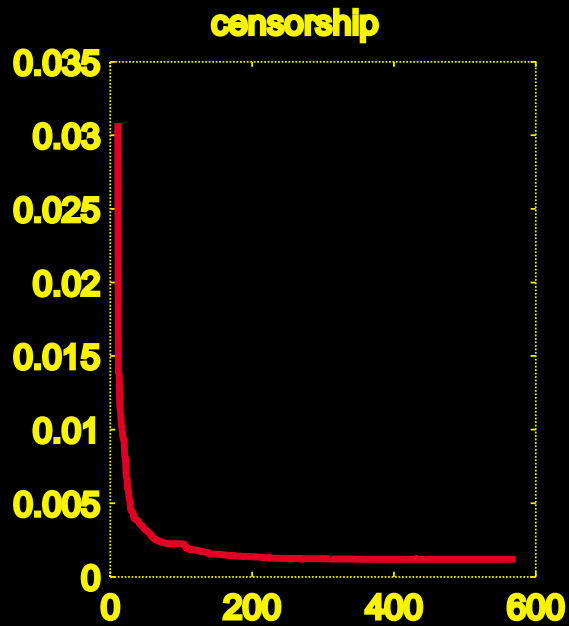
<u><math> G </math></u>	<u>Iterations</u>	<u>Time</u>
500	19	1.12
1000	15	.92
1250	20	1.04
<b>1500</b>	<b>14</b>	<b>.90</b>
2000	13	1.17
5000	6	1.25

*nodes = 9,664 links = 16,150*



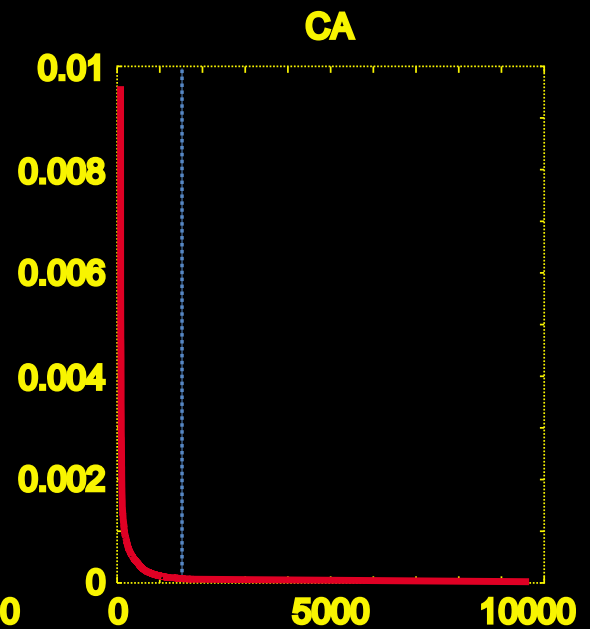
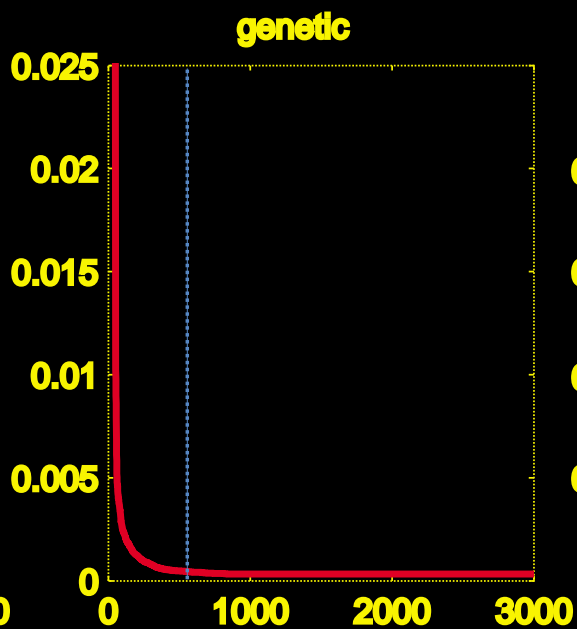
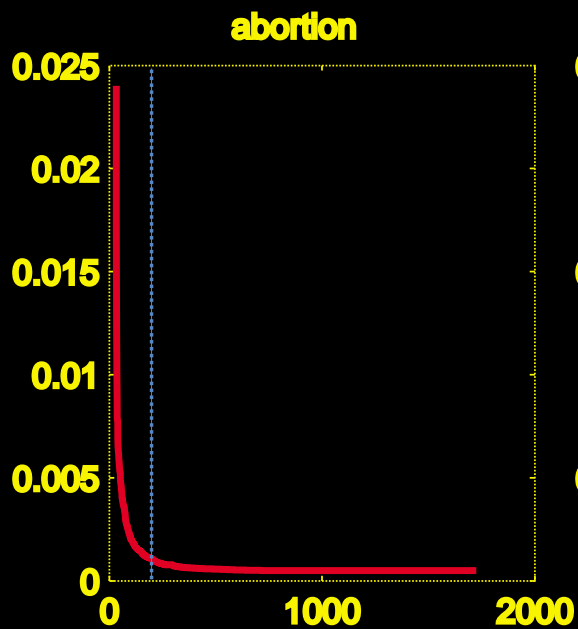
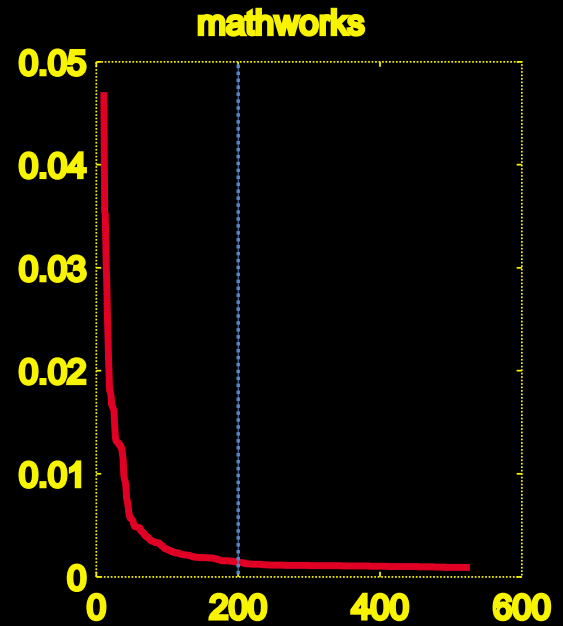
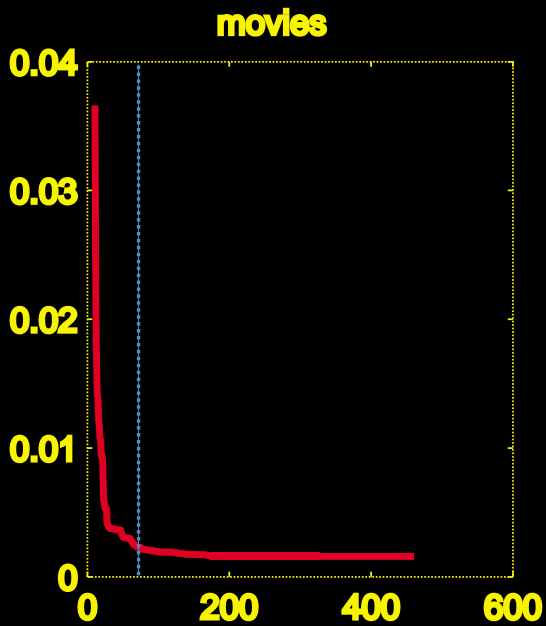
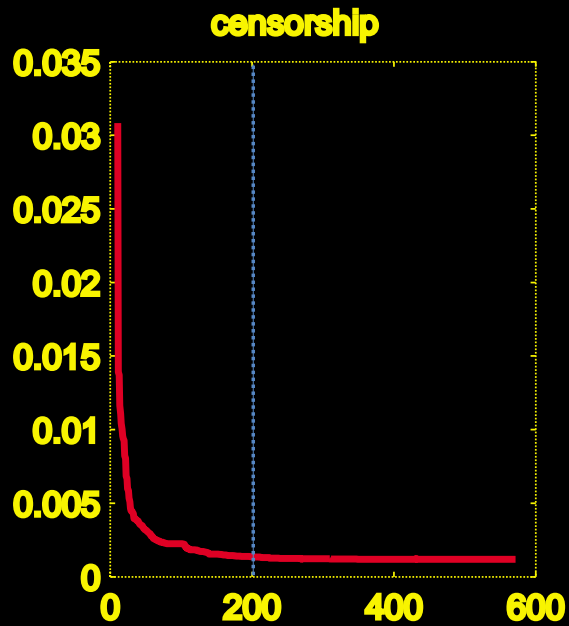


# “L” Curves





# “L” Curves











# Timings

	<b>Iterations</b>	<b>Time (sec)</b>	<b> G </b>
<b>Power</b>	162	9.69	
<b>Power+Quad</b>	81	5.93	
<b>IAD</b>	21	2.22	2000
<b>IAD+Quad</b>	16	1.85	2000

*nodes = 10,000    links = 101,118*



# Conclusion

- ✦ **Iterative aggregation shows  
promise for updating  
Markov chains**
- ✦ **Especially for those having  
power law distributions**



# Leveling Off Point

$$\pi(i) \approx \alpha i^{-k}$$

$$\left| \frac{d\pi(i)}{di} \right| \approx \epsilon \quad \text{for some user-defined tolerance } \epsilon$$

$$i_{level} \approx \left( \frac{k\alpha}{\epsilon} \right)^{1/k+1}$$

Perhaps better: 
$$i_{level} \approx f(n) \left( \frac{k\alpha}{\epsilon} \right)^{1/k+1}$$

For WWW: 
$$g_{opt} \approx f(n) \left[ \frac{2.109\alpha}{\epsilon} \right]^{1/3.109}$$