Mixing time of random walks on networks

Keith Briggs Keith.Briggs@bt.com

Joint work with Min Chen

research.btexact.com/teralab/keithbriggs.html

PICT seminar 2004 December 16 1500

typeset 2005 January 7 10:45 in pdfLaTeX on a linux system

Outline

- What is a random walk?
- How do we quantify a random walk?
- Fastest mixing problem
- Computational results
- Applications

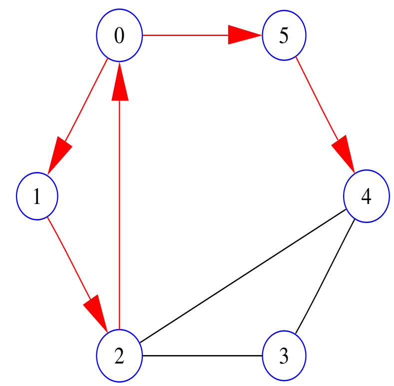
mixing time $\longrightarrow$ convergence speed

fast mixing $\longrightarrow$ fast convergence
of dynamical process on network

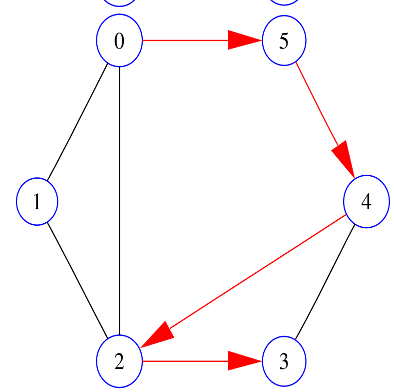
Random walk

- $\Gamma(v)$: neighbors of v in G
- $x_0 \xrightarrow{\text{random}} x_1 \in \Gamma(x_0) \xrightarrow{\text{random}} x_2 \in \Gamma(x_1) \xrightarrow{\text{random}} \cdots \xrightarrow{\text{random}} x_k \in \Gamma(x_{k-1})$
- Example:

$\{v_0 v_1 v_2 v_0 v_5 v_4\}$ is one instance of a random walk



$\{v_0 v_5 v_4 v_2 v_3\}$ is another instance of a random walk



Random walk and Markov chain

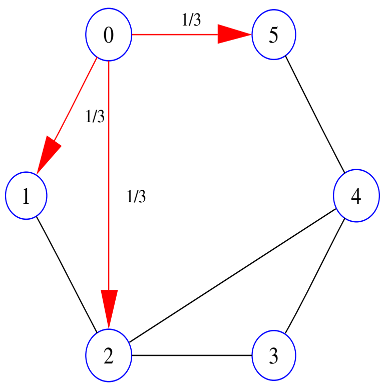
- Transition probability matrix P

$$P_{ij} = \begin{cases} \text{Pr}\left[x_{t+1} = j | x_t = i\right], & \text{if } ij \in E \\ 0, & \text{otherwise} \end{cases}$$

- We can choose each node uniformly

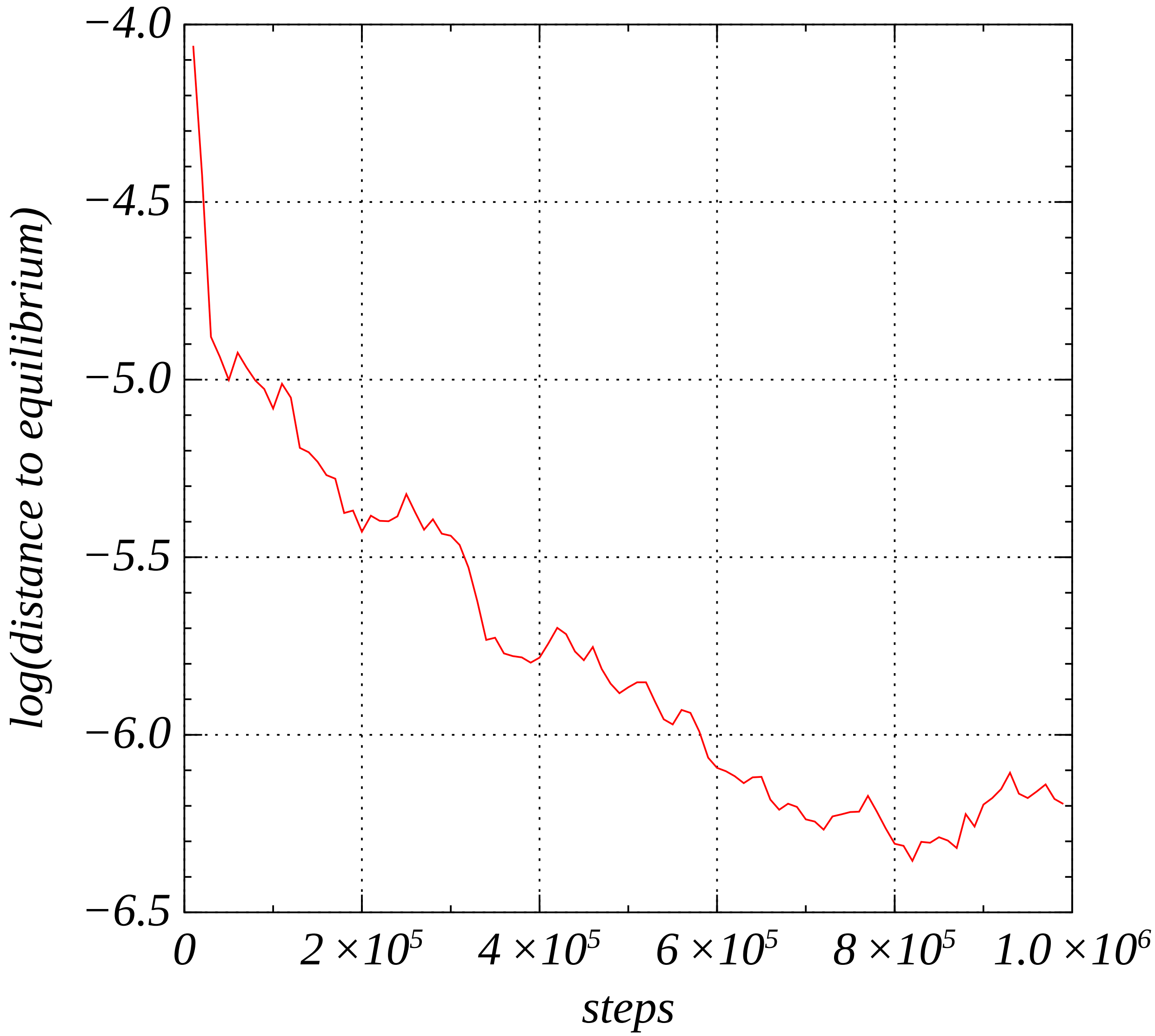
$$P_{ij} = \begin{cases} 1/d(i), & \text{if } ij \in E \\ 0, & \text{otherwise} \end{cases}$$

where $d(i)$ is the degree of i



- distribution at t: $\pi(t) = \pi(0)P^t$
- equilibrium distribution: $\pi = \pi P$

The approach to equilibrium



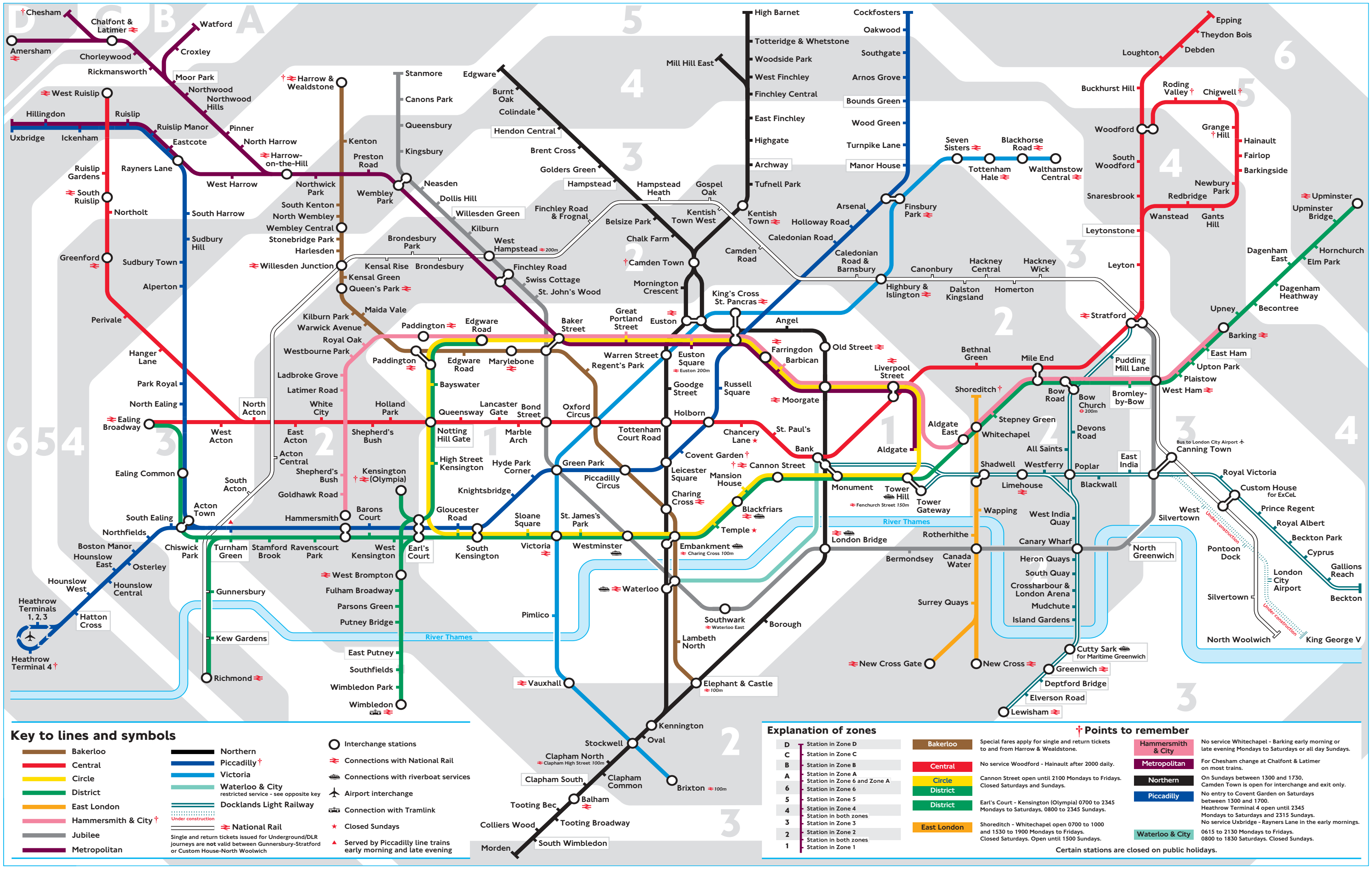
start a walk

plot $\log ||\pi(t) - \pi(\infty)||$

vs. t

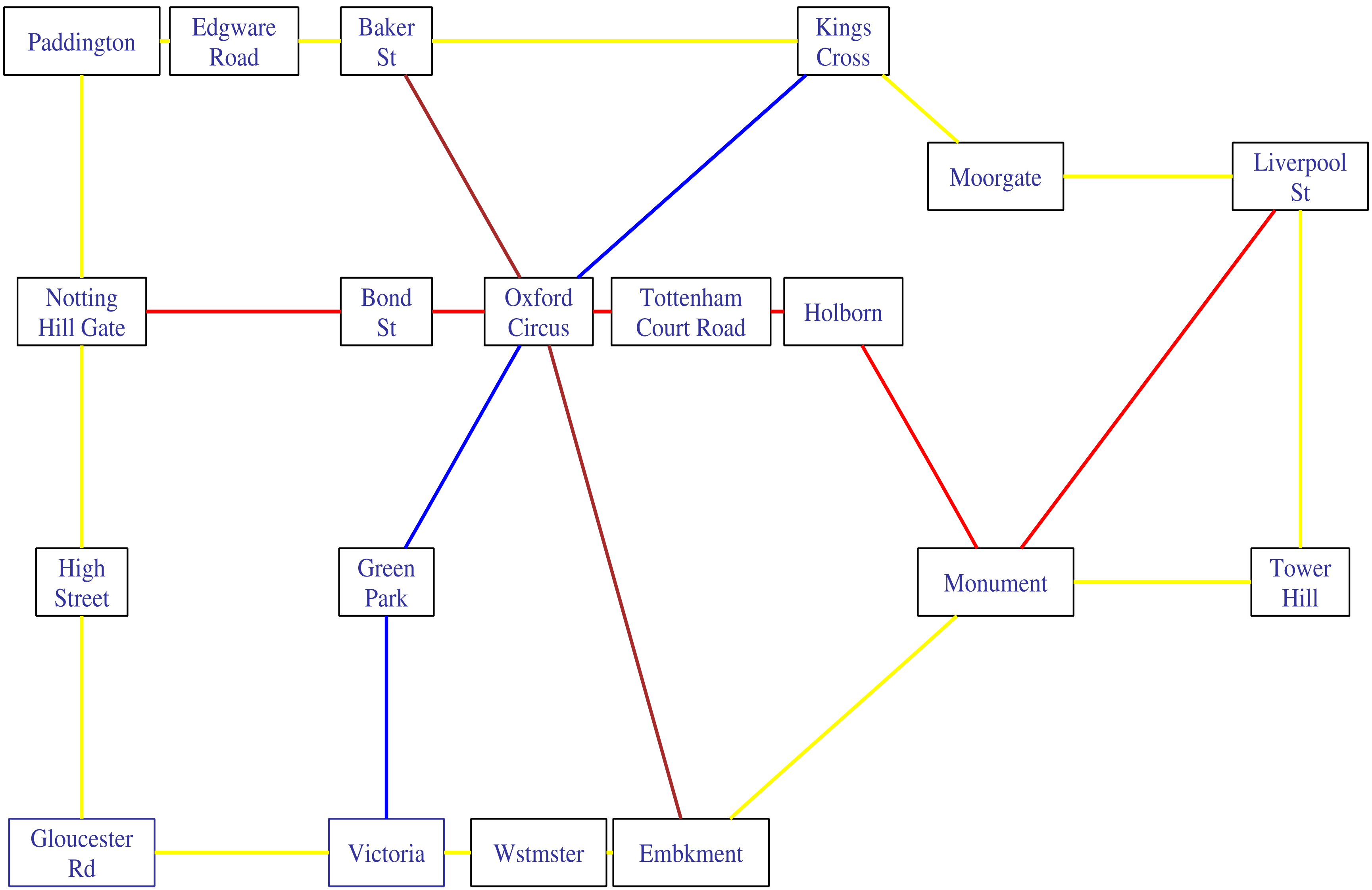
Statistical properties of random walks

- Example: London underground graph



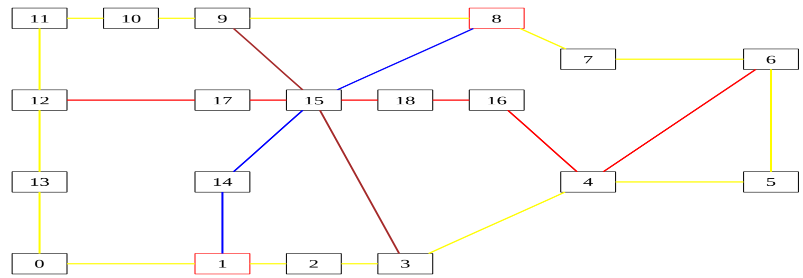
Transport for London

Example: London underground graph



Example: London underground graph - hitting time

- King's Cross $\underbrace{\overset{\text{random walk}}{- - - - \longrightarrow}}_{\text{? steps}}$ Victoria



- Hitting time [LL00]:

$$H_{ij} = 2m \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left[\frac{v_{ki}^2}{d(i)} - \frac{v_{ki}v_{kj}}{\sqrt{d(i)d(j)}} \right]$$

$\{\lambda, v\}$ is the eigensystem of $N = D^{1/2}AD^{1/2}$, A is the adjacency matrix, $D = \text{diag}(1/d(i))$, m is the number of edges in the graph

- the mean number of steps between King's Cross and Victoria is

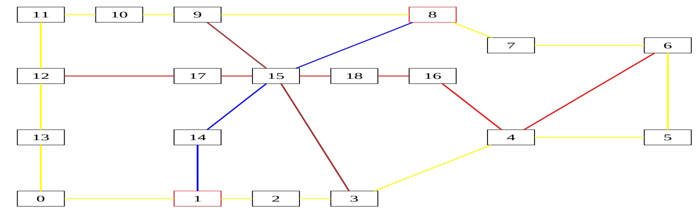
$$H_{kc-v} = H_{81} = 38.5$$

London underground graph - Hitting time matrix

0.0	12.7	30.1	24.9	36.1	58.8	47.5	49.7	31.2	30.0	46.5	45.0	25.4	20.9	25.9	13.9	50.8	31.3	43.4
28.1	0.0	21.3	20.1	32.6	55.4	44.4	47.4	29.6	29.8	49.6	51.4	35.0	39.8	18.5	11.8	47.7	35.1	40.8
40.6	16.4	0.0	11.0	27.0	50.5	40.0	44.7	28.7	29.9	51.3	54.5	39.7	48.3	26.6	11.5	43.9	37.3	38.8
51.1	30.9	26.7	0.0	19.5	43.5	33.7	40.1	25.8	28.1	50.9	55.7	42.3	54.9	32.7	9.3	38.1	37.5	34.7
58.4	39.4	38.7	15.5	0.0	26.4	18.9	32.5	25.4	30.3	54.0	59.6	47.1	61.0	39.0	13.3	26.5	41.9	30.9
60.3	41.5	41.4	18.8	5.7	0.0	10.5	27.8	24.4	30.8	54.8	60.7	48.5	62.6	40.8	14.8	30.8	43.3	33.8
60.2	41.7	42.2	20.1	9.4	21.6	0.0	21.1	21.5	29.2	53.5	59.7	47.9	62.2	40.6	14.3	33.1	42.7	34.7
58.8	41.1	43.3	23.0	19.4	35.4	17.6	0.0	11.8	23.5	48.8	55.9	45.0	60.1	39.1	11.8	38.9	40.1	36.4
55.5	38.5	42.5	23.9	27.4	47.2	33.2	26.9	0.0	15.9	42.0	50.1	40.1	56.0	35.5	7.3	42.7	35.4	36.0
53.3	37.7	42.7	25.2	31.4	52.6	39.9	37.7	14.9	0.0	29.8	41.6	35.3	52.5	35.0	7.0	45.3	32.8	37.2
50.1	37.8	44.4	28.3	35.4	56.9	44.5	43.3	21.4	10.1	0.0	21.8	25.5	46.0	36.9	10.7	49.2	29.8	41.0
45.0	36.0	44.0	29.4	37.4	59.1	47.1	46.8	25.8	18.2	18.2	0.0	13.8	37.6	36.8	12.5	51.1	24.8	42.8
37.8	32.1	41.6	28.6	37.3	59.4	47.7	48.3	28.3	24.4	34.3	26.2	0.0	27.1	34.8	12.2	51.0	17.8	42.6
19.9	23.4	36.8	27.7	37.7	60.1	48.6	50.0	30.7	28.2	41.4	36.6	13.7	0.0	31.4	14.1	51.9	25.5	44.0
40.7	17.9	30.8	21.3	31.5	54.0	42.7	44.7	26.0	26.4	48.0	51.6	37.1	47.1	0.0	6.9	45.3	33.7	37.1
51.3	33.7	38.4	20.4	28.4	50.7	39.0	40.0	20.3	21.0	44.5	49.8	37.1	52.4	29.5	0.0	41.0	30.2	31.5
58.0	39.5	40.6	19.2	11.5	36.5	27.6	37.0	25.7	29.2	52.8	58.3	45.8	60.1	37.8	10.9	0.0	40.0	16.5
45.6	33.9	41.0	25.5	33.9	56.0	44.3	45.2	25.3	23.7	40.4	39.0	19.6	40.8	33.1	7.1	47.0	0.0	38.1
55.7	37.6	40.5	20.8	21.0	44.6	34.3	39.5	24.0	26.1	49.7	55.1	42.5	57.3	34.7	6.4	21.5	36.1	0.0

London underground graph - commute time

- King's Cross $\underbrace{\xrightarrow{\text{rw}} \text{Victoria} \xrightarrow{\text{rw}}}_{?\ \text{steps}}$ King's Cross



- commute time [LL00]

$$k_{ij} = H_{ij} + H_{ji}$$

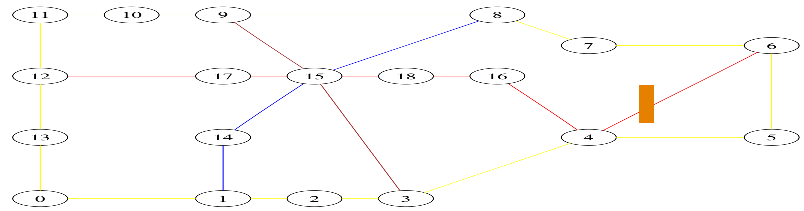
- So we know that the commute time from King's Cross to Victoria then back to King's Cross is

$$k_{k-v} = k_{81} = H_{81} + H_{18} = 38.5 + 29.6 = 68.1$$

- Symmetry
  - $H_{ij} \neq H_{ji}$ unless i and j are vertex-transitive.
  - $k_{ij} = k_{ji}$ for all i, j

London underground graph - mixing time

- $\underbrace{\pi(0) \to \pi(1) \to \cdots \to \pi}_{\text{? steps}}$



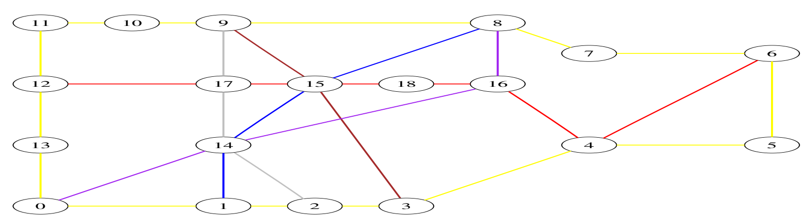
- $\mathit{mixing\ rate} = \log(1/\mu(P))$

  where $\mu(P) = \max_{i=2,\dots,n} |\lambda_i(P)| = \max\{\lambda_2(P), -\lambda_n(P)\}$

- mixing time: $\tau = 1/(\text{mixing rate}) = 1/\log(1/\mu)$
- $\mathit{mixing\ rate} = 0.101568$

  $\mathit{mixing\ time} = 9.845655$

- build new lines



$\mathit{mixing\ rate} = 0.168313$

$\mathit{mixing\ time} = 5.941294$

Definition of random walk properties

hitting time (access time)	H_{ij} *is the expected number of steps in a random walk starting from node* i *and before node* j.	$2m\sum_{k=2}^{n}\frac{1}{1-\lambda_k}\left[\frac{v_{ki}^2}{d(i)}-\frac{v_{ki}v_{kj}}{\sqrt{d(i)d(j)}}\right]$
commute time	k_{ij} *is the expected number of steps in a random walk starting at* i, *the first time return to* i *via* j.	$H_{ij}+H_{ji}$
mixing rate	*measure of how fast the random walk converges to its stationary distribution.*	$\rho=-\log(\mu(P))$
mixing time	*the time scale (in steps) for reaching the stationary distribution.*	$\tau=-\frac{1}{\log(\mu)}$

Fastest mixing graph problem

- fastest mixing (minimum mixing time) by changing topology

change the topology of the graph

$$\downarrow$$

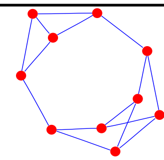
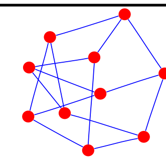
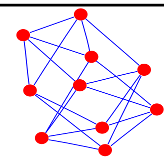
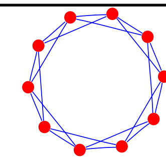
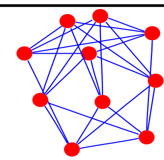
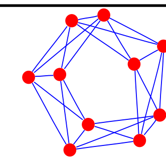
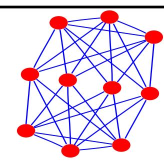
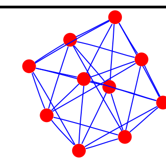
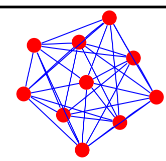
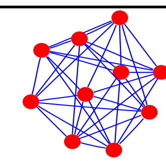
fastest mixing chain

- Optimization description

$$\begin{array}{ll} \min_G & \mu(P(G)) \\ \text{s.t.} & P_{\cdot} \geqslant 0 \\ & P\mathbf{1} = \mathbf{1} \\ & P_{ij} = \begin{cases} 1/d(i), & \text{if } ij \in E \\ 0, & \text{otherwise} \end{cases} \end{array}$$

Computational results - small regular graphs

- the min/max mixing time for 10 nodes regular graphs

n	deg	num	maxtime	graph	mintime	graph	avertime
10	3	17	15.5896		2.4663		6.8216
10	4	58	7.7220		1.7195		3.4093
10	5	59	7.4542		1.2427		2.2145
10	6	21	2.4663		0.9102		1.5722
10	7	5	1.1802		1.0168		1.1475

The fastest mixing problem

- fastest mixing (minimum mixing time) by adjusting weights

fix the topology, changing the weights

$\downarrow$ *(reversible chain)*

fastest mixing chain

- Optimization description

$$\begin{array}{ll} \min_P & \mu(P) \\ \textbf{s.t.} & P_{\bullet} \geqslant 0 \\ & P\mathbb{1} = \mathbb{1} \\ & \Pi P = P^T \Pi \\ & P_{ij} = 0, \quad i,j \notin E \end{array}$$

Fastest mixing problem - combined problem

- Fastest mixing chain both on topology and weights

all topologies

↓ *change the weights*

fastest mixing weights on each topology

↓

fastest mixing chain

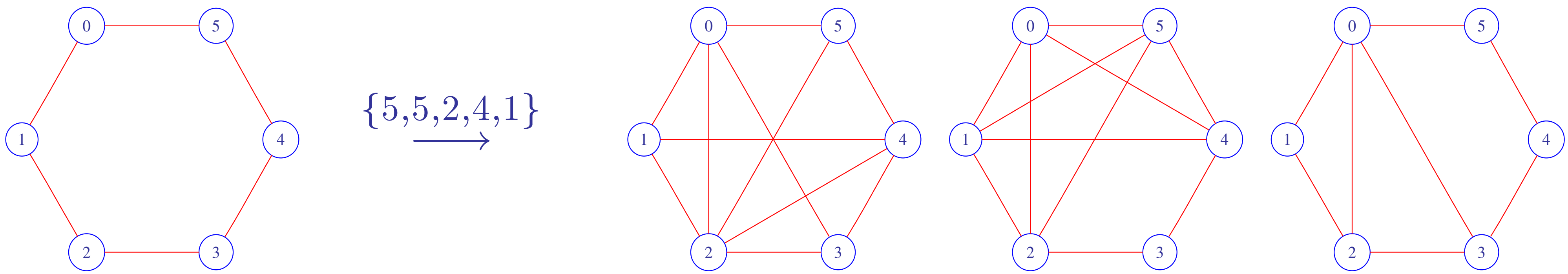
- Optimization description

$$
\begin{array}{ll}
\min_G \min_{P(G)} & \mu(P) \\
\text{s.t.} & P_{\cdot} \geqslant 0 \\
 & P1 = 1 \\
 & \Pi P = P^T \Pi \\
 & P_{ij} = 0, \quad i, j \notin E
\end{array}
$$

Small-world model $SW\{n,p\}$

- $SW\{n,p\}$ model:

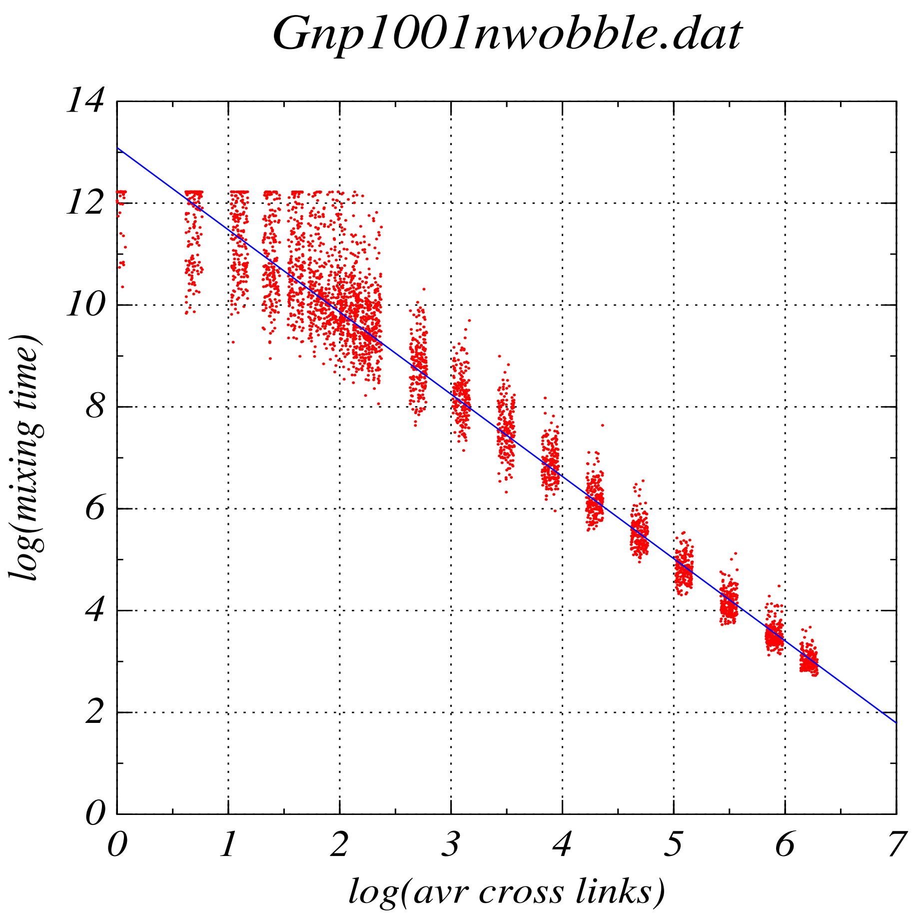
  n-node cycle $\Rightarrow$ $\{m_0, m_1, \dots, m_k\} \in$ Poisson distribution with mean $\bar{m} = \binom{n}{2}p$ $\Rightarrow$ $SW(n,m_0), \dots, SW(n,m_k)$



  $SW\{6, 1/5\}, \bar{m} = 15 \times 0.2 = 3 \dots$

- What is the relation between $\bar{m}$ and mixing time?

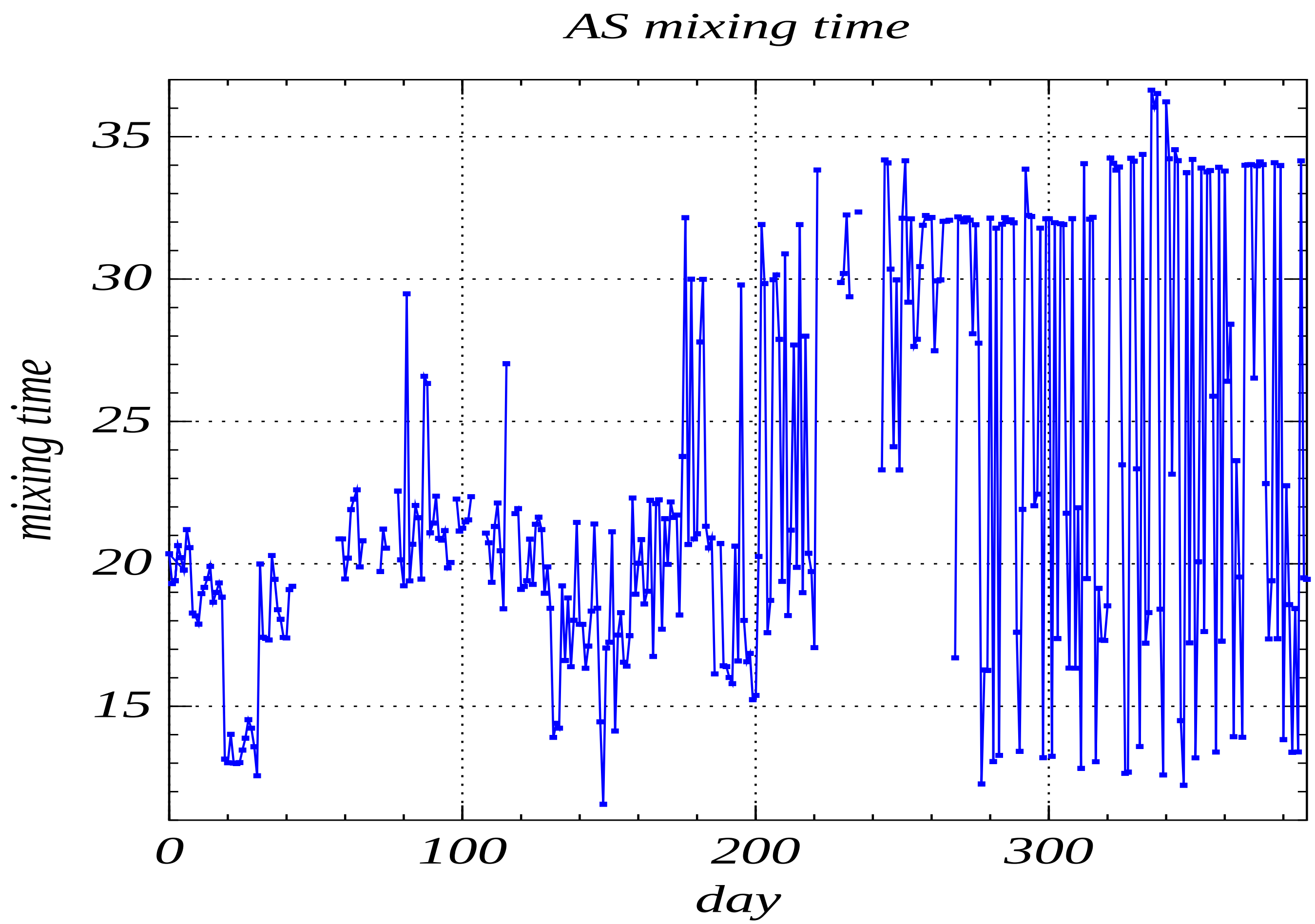
Small-world $SW\{n,p\}$ - mixing time vs. links



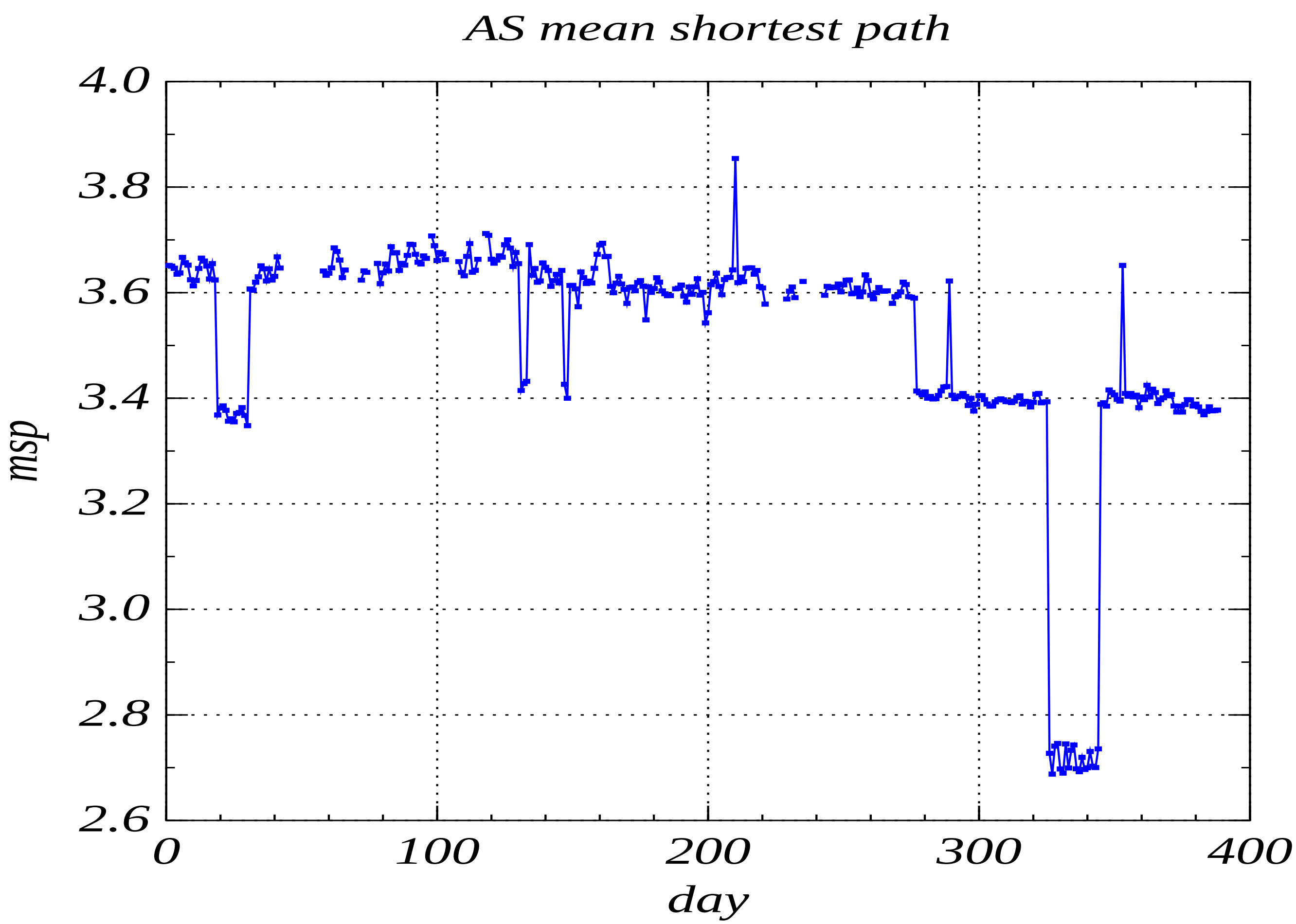
Here is the mixing time vs. the mean number of cross links

Mixing time of the internet AS graph

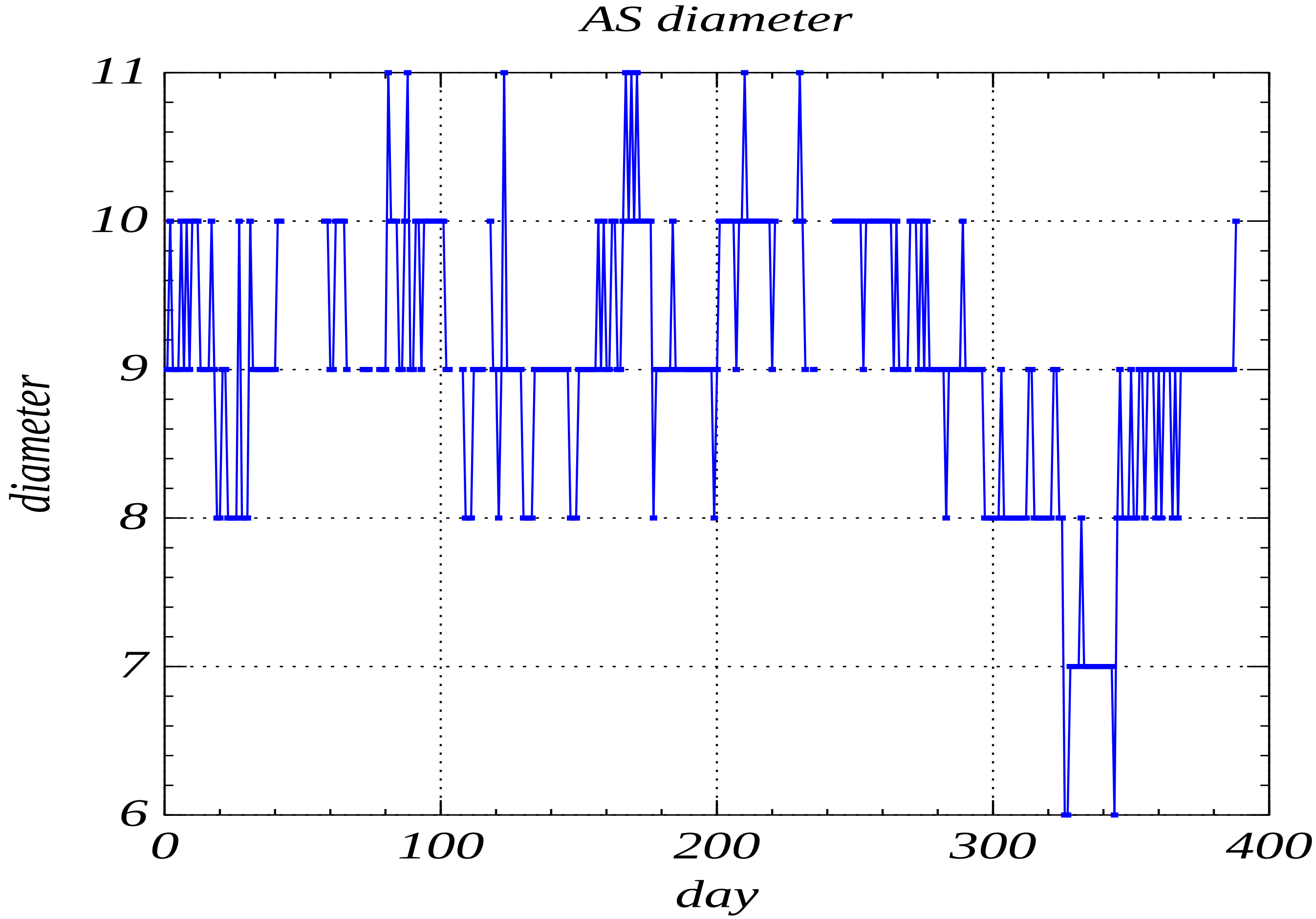
- Internet AS graph [CAIDA] has about 10,000 nodes and changes topology constantly



Mean shortest path of the AS graph



Diameter of the internet AS graph



Applications

- Information diffusion

  --- How fast does information/virus/spam spread on the internet?

- Search engine (Google)

  --- How fast can google rank the pages?

- Sampling problems

  --- e.g. Markov chain Monte Carlo, distributed averaging

Diffusion

- Eigenvalues λ_i and eigenfunctions ϕ_i
- heat kernel

$$H_t(x,y) = \sum_i \exp(-\lambda_i t)\ \phi_i(x)\ \phi_i(y)$$

- satisfies

$$\frac{d}{dt} f = -\mathcal{L}_S\ f$$

- $H_t = \exp(-t\mathcal{L}_S)$
- this solves the diffusion problem

References

CAIDA data from sk-aslinks.caida.org/data/

R. Diestel, *Graph Theory*, Springer 2000

B. Bollobás, *Random Graphs*, CUP 2001

S. Boyd, P. Diaconis & L. Xiao, *Fastest Mixing Markov Chain on A Graph*

L. Lovász, *Random walks on graphs: a survey (2001)*

Fan Chung & S-T Yau 2000 `Discrete Green's functions`

N. Biggs, *Algebraic graph theory*, CUP 1993

B. Bollobás, *Modern graph theory*, Springer-Verlag 2002