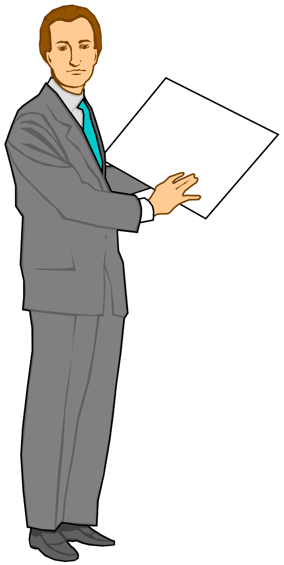
Graph Spectra & the N-intertwined Mean-field SIS approximation on Networks

Piet Van Mieghem

work in collaboration with Eric Cator, Huijuan Wang, Rob Kooij

Outline



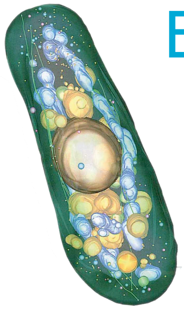
Introduction & Definitions

Exact SIS model

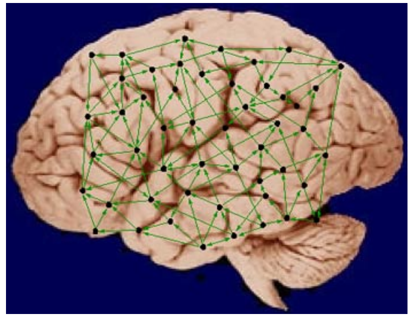
The N-intertwined MF approximation

Extensions

Summary



Effect of Network on Function



Biology

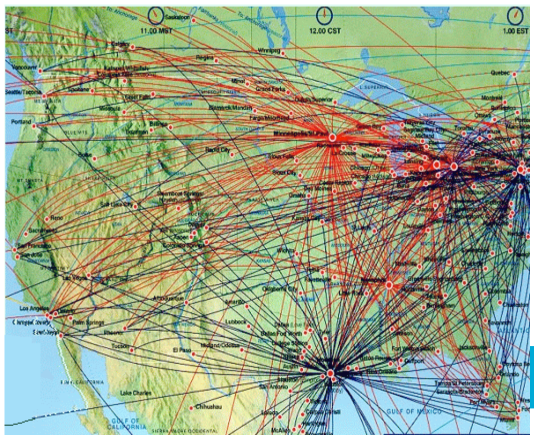
Brain

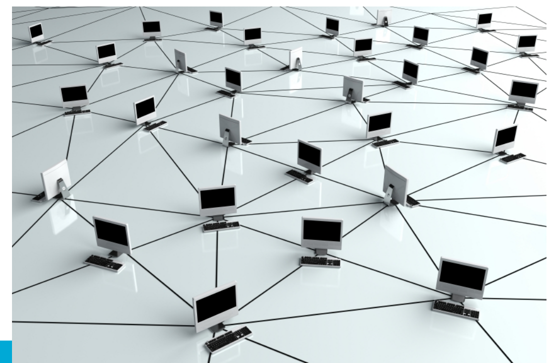
Economy

Network Science

Social

networking





Man-made Infrastructures
Internet, power-grid

Brain/Bio-inspired Networking
to design superior man-made
robust networks/systems

TU Delft

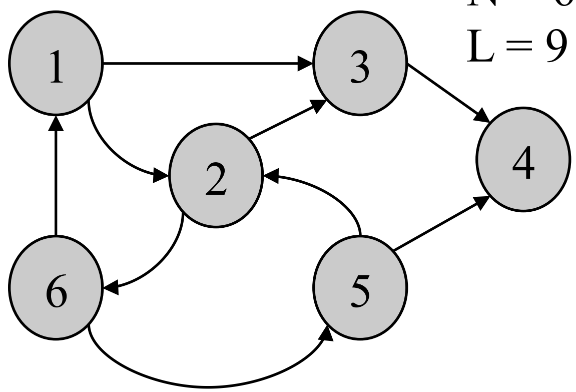
Motivation for virus spread in networks

- **Digital world:**
  - Information spread in on-line social networks
  - security threat to Internet (Code Red worm: several billion $ $ in damage)
- **Real world: Biological epidemics** (e.g. Mexican flue)

- **Why do we care**?
  - Understanding the spread of a virus is the first step in prevention
  - How fast do we need to disinfect nodes so that the virus dies quickly? Which nodes?

Algebraic graph theory

Any graph G can be represented by an adjacency matrix A and an incidence matrix B, and a Laplacian Q



N = 6
L = 9

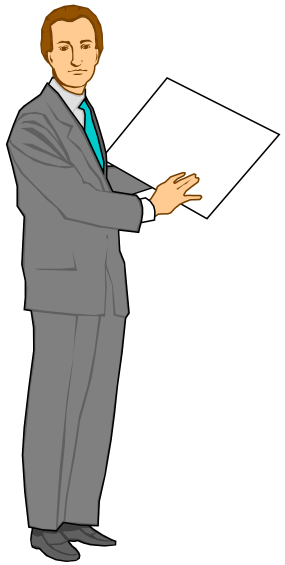
$$A_{N\times N} = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \end{bmatrix} = A^T$$

$$B_{N\times L} = \begin{bmatrix} 1 & 1 & -1 & \cdots & 0 \\ -1 & 0 & 0 & & 0 \\ 0 & -1 & 0 & & 0 \\ 0 & 0 & 0 & & 0 \\ 0 & 0 & 0 & & -1 \\ 0 & 0 & 1 & & 1 \end{bmatrix}$$

$$Q = BB^T = \Delta - A$$

$$\Delta = diag(d_1 \quad d_2 \quad \ldots \quad d_N)$$

Outline



Introduction & Definitions

Exact SIS model

The N-intertwined MF approximation

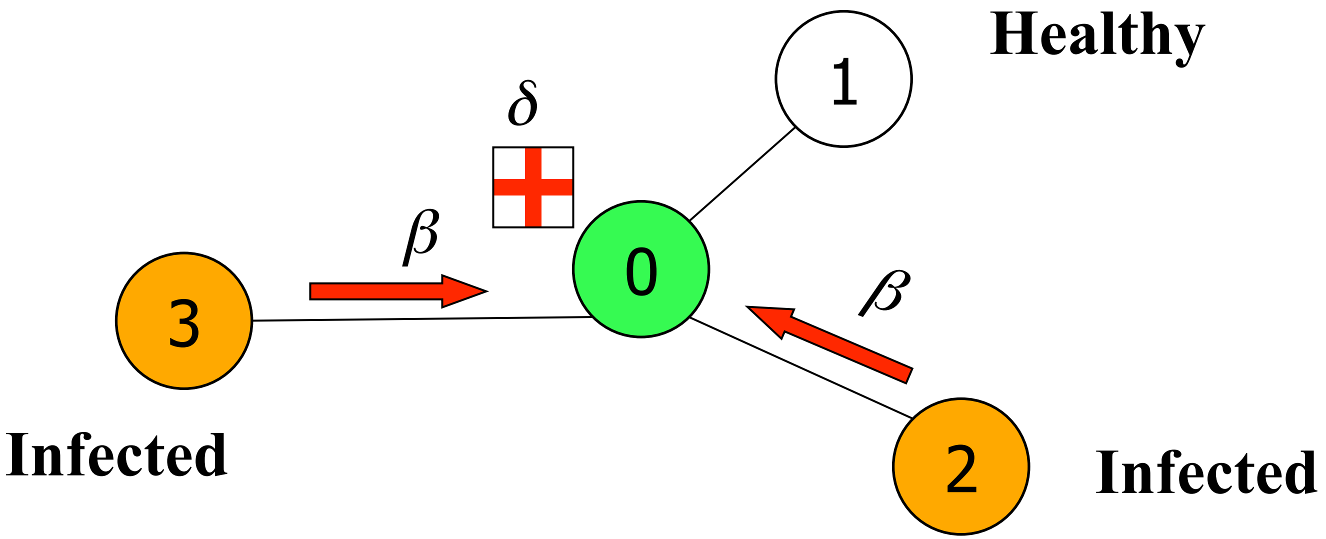
Extensions

Summary

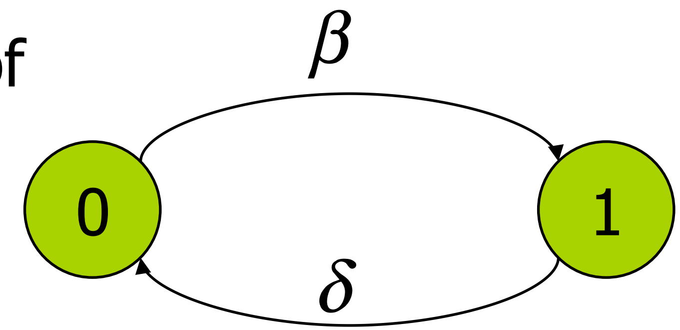
Simple SIS model

- Homogeneous birth (infection) rate β on all edges between infected and susceptible nodes
- Homogeneous death (curing) rate δ for infected nodes

$\tau = \beta / \delta$: effective spreading rate



Definition of the states in SIS



- Each node *j* can be in either of the two states:
  - "0": healthy
  - "1": infected
- **Markov continuous time**:
  - infection rate β
  - curing rate δ
- Mathematically:
  - X_j is the state of node j
  - infinitesimal generator $Q_j(t) = \begin{vmatrix} -q_{1j} & q_{1j} \\ q_{2j} & -q_{2j} \end{vmatrix} = \begin{bmatrix} -q_{1j} & q_{1j} \\ \delta & -\delta \end{bmatrix}$

Governing SIS equation for node *j*

$$\frac{dE[X_j]}{dt} = E\left[-\delta X_j + (1 - X_j)\beta\sum_{k=1}^{N} a_{kj}X_k\right]$$

time-change of $E[X_j] = \Pr[X_j = 1]$, probability that node *j* is infected

if *infected*: probability of curing per unit time

if *not infected (healthy)*: probability of infection per unit time

$$\frac{dE[X_j]}{dt} = -\delta E\left[X_j\right] + \beta\sum_{k=1}^{N} a_{kj}E\left[X_k\right] - \beta\sum_{k=1}^{N} a_{kj}E\left[X_jX_k\right]$$

Joint probabilities

$$\frac{dE\left[X_iX_j\right]}{dt} = E\left[-2\delta X_iX_j + X_j\beta(1-X_i)\sum_{k=1}^{N} a_{ik}X_k + X_i\beta(1-X_j)\sum_{k=1}^{N} a_{jk}X_k\right]$$

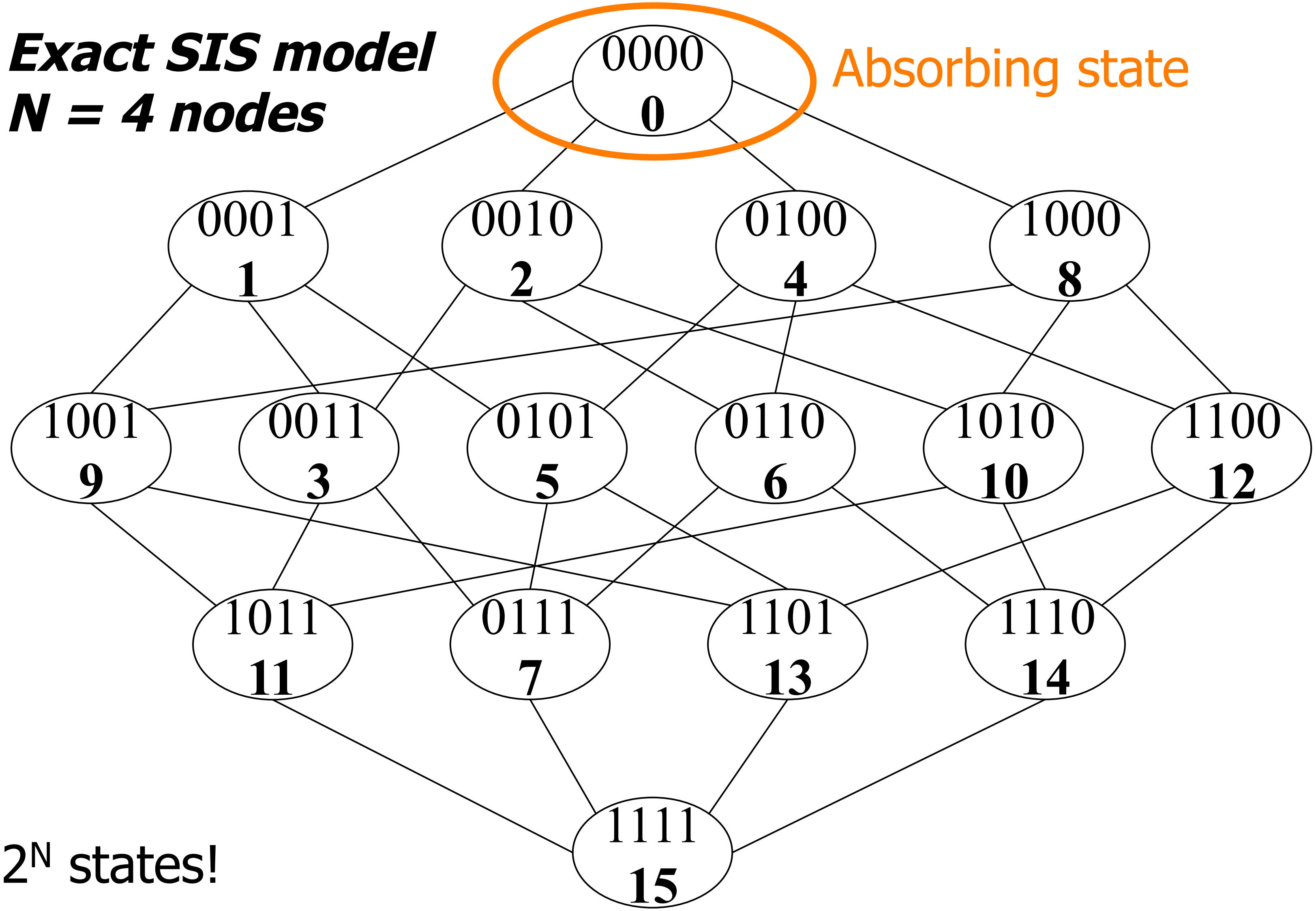
$$= -2\delta E\left[X_iX_j\right] + \beta\sum_{k=1}^{N} a_{ik}E\left[X_jX_k\right] + \beta\sum_{k=1}^{N} a_{jk}E\left[X_iX_k\right] - \beta\sum_{k=1}^{N}\left(a_{jk}+a_{ik}\right)E\left[X_iX_jX_k\right]$$

Next, we need the $\binom{N}{3}$ differential equations for $E[X_iX_jX_k]$...

In total, the SIS process is defined by $2^N = \sum_{k=1}^{N}\binom{N}{k} + 1$ linear equations

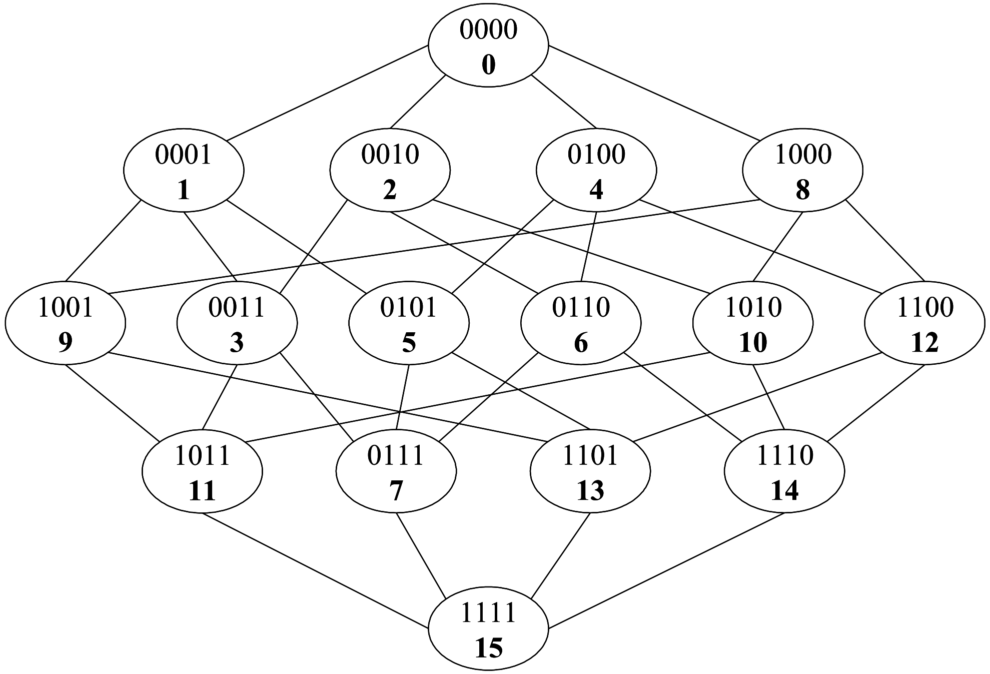
E. Cator and P. Van Mieghem, 2012, "Second-order mean-field susceptible -infected-susceptible epidemic threshold", Physical Review E, vol. 85, No. 5, May, p. 056111.

Exact SIS model
N = 4 nodes

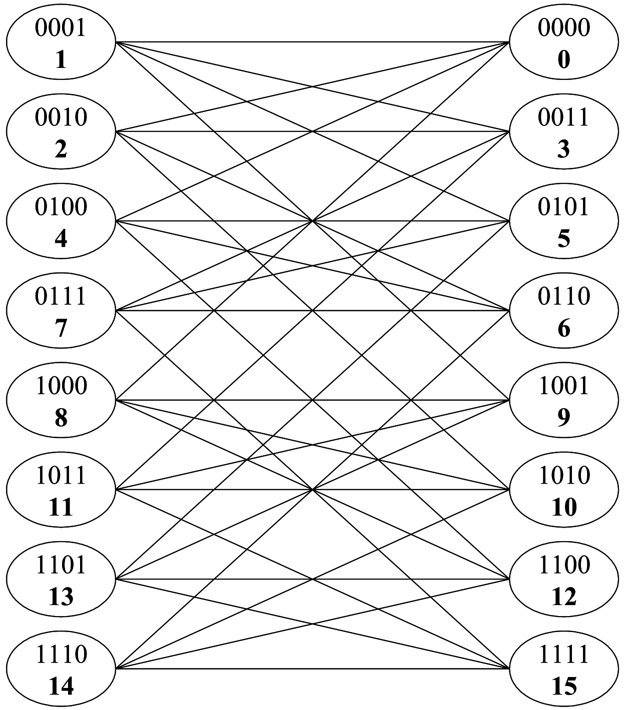


2^N states!

Markov theory



bi-partite Markov graph



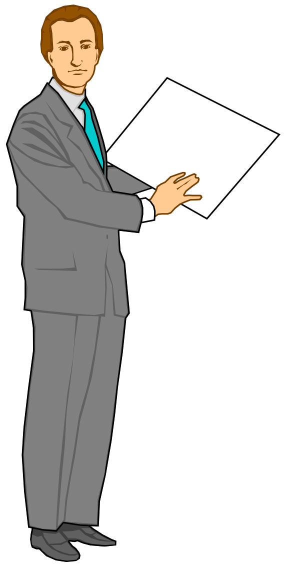
Recursive structure of infinitesimal general Q_N

Van Mieghem, P. and E. Cator, ε-SIS epidemics and the epidemic threshold,
Physical Review E, to appear 2012

Markov Theory

- SIS model is exactly described as a continuous-time Markov process on 2^N states, with infinitesimal generator Q_N.
- **Drawbacks**:
  - no easy structure in Q_N
  - computationally intractable for N>20
  - steady-state is the absorbing state (reached after unrealistically long time)
  - very few exact results...
- The mathematical community (e.g. Liggett, Durrett,...) uses:
  - duality principle & coupling & asymptotics
  - graphical representation of the Poisson infection and recovery events

Outline



Introduction & Definitions

Exact SIS model

The N-intertwined MF approximation

Extensions

Summary

TUDelft

N-intertwined mean-field approximation

$$\frac{dE[X_j]}{dt} = -\delta E\left[X_j\right] + \beta \sum_{k=1}^{N} a_{kj} E\left[X_k\right] - \beta \sum_{k=1}^{N} a_{kj} E\left[X_j X_k\right]$$

$E\left[X_j X_k\right] = \Pr\left[X_j = 1, X_k = 1\right] = \Pr\left[X_j = 1 \middle| X_k = 1\right] \Pr\left[X_k = 1\right]$ and $\Pr\left[X_j = 1 \middle| X_k = 1\right] \geq \Pr\left[X_j = 1\right]$

$$\Rightarrow \quad E\left[X_i X_k\right] \geq \Pr\left[X_i = 1\right] \Pr\left[X_k = 1\right] = E\left[X_i\right] E\left[X_k\right]$$

$$\boxed{\frac{dE[X_j]}{dt} \leq -\delta E\left[X_j\right] + \beta \sum_{k=1}^{N} a_{kj} E\left[X_k\right] - \beta E\left[X_j\right] \sum_{k=1}^{N} a_{kj} E\left[X_k\right]}$$

N-intertwined mean-field approx. (= equality above)
upper bounds the probability of infection

E. Cator and P. Van Mieghem, 2012, "Second-order mean-field susceptible -infected-susceptible epidemic threshold", Physical Review E, vol. 85, No. 5, May, p. 056111.

N-intertwined non-linear equations

$$\begin{cases} \dfrac{dv_1}{dt} = (1-v_1)\beta\displaystyle\sum_{k=1}^{N} a_{1k}v_k - \delta v_1 \\ \dfrac{dv_2}{dt} = (1-v_2)\beta\displaystyle\sum_{k=1}^{N} a_{2k}v_k - \delta v_2 \\ \vdots \\ \dfrac{dv_N}{dt} = (1-v_N)\beta\displaystyle\sum_{k=1}^{N} a_{Nk}v_k - \delta v_N \end{cases}$$

where the viral probability of infection is

$$v_k(t) = E[X_k(t)] = \Pr[X_k(t) = 1]$$

In matrix form:

$$\boxed{\frac{dV(t)}{dt} = \beta A.V(t) - diag(v_i(t))(\beta A.V(t) + \delta u)}$$

where the vector $u^T = [1\ 1\ \dots\ 1]$ and $V^T = [v_1\ v_2\ \dots\ v_N]$

P. Van Mieghem, J. Omic, R. E. Kooij, "Virus Spread in Networks",
IEEE/ACM Transaction on Networking, Vol. 17, No. 1, pp. 1-14, (2009).

Lower bound for the epidemic threshold

$$\frac{dv_j(t)}{dt} = -\delta v_j + \beta \sum_{k=1}^{N} a_{kj} v_k - \beta \sum_{k=1}^{N} a_{kj} E\left[X_i X_k\right] \qquad v_k(t) = E[X_k(t)]$$

Is the point V = 0 stable?
For a very few infected nodes, we can ignore the quadratic terms

$$\frac{dV(t)}{dt} = \left(-\delta I + \beta A\right) V(t)$$

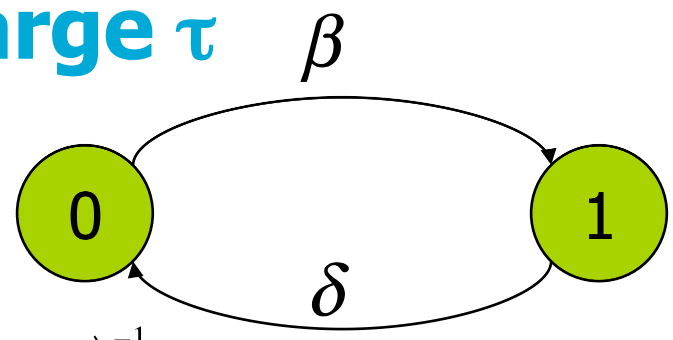
The origin V=0 is stable attractor if all eigenvalues of $\beta A - \delta I$ are negative (v_j tends exponentially fast to zero with *t*). Hence, if

$$\beta \lambda_1(A) - \delta < 0 \quad \Longrightarrow \quad \tau = \frac{\beta}{\delta} < \frac{1}{\lambda_1(A)} < \tau_c$$

The N-intertwined mean-field epidemic threshold is precisely

$$\boxed{\tau_c^{(1)} = \frac{1}{\lambda_1(A)} < \tau_c} \qquad \boxed{\tau_c^{(1)} = \frac{1}{\lambda_1(A)} < \tau_c^{(2)} = \frac{1}{\lambda_1(H)} < \tau_c}$$

Exact in steady-state for large τ



Almost all neighbors of node *j* are infected: independence

$$\Pr\left[X_j = 1\right] \cong \frac{\beta d_j}{\delta + \beta d_j} = \left(1 + \frac{1}{\tau d_j}\right)^{-1} = \left(1 + \frac{s}{d_j}\right)^{-1}$$

Exact steady-state fraction of infected nodes:

$$y_\infty(s) \cong \frac{1}{N}\sum_{j=1}^{N} \Pr\left[X_j = 1\right] = \frac{1}{N}\sum_{j=1}^{N}\left(1 + \frac{s}{d_j}\right)^{-1}$$

Slope: $$\left.\frac{dy_\infty(s)}{ds}\right|_{s=0} = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{d_j} = E\left[\frac{1}{D}\right]$$

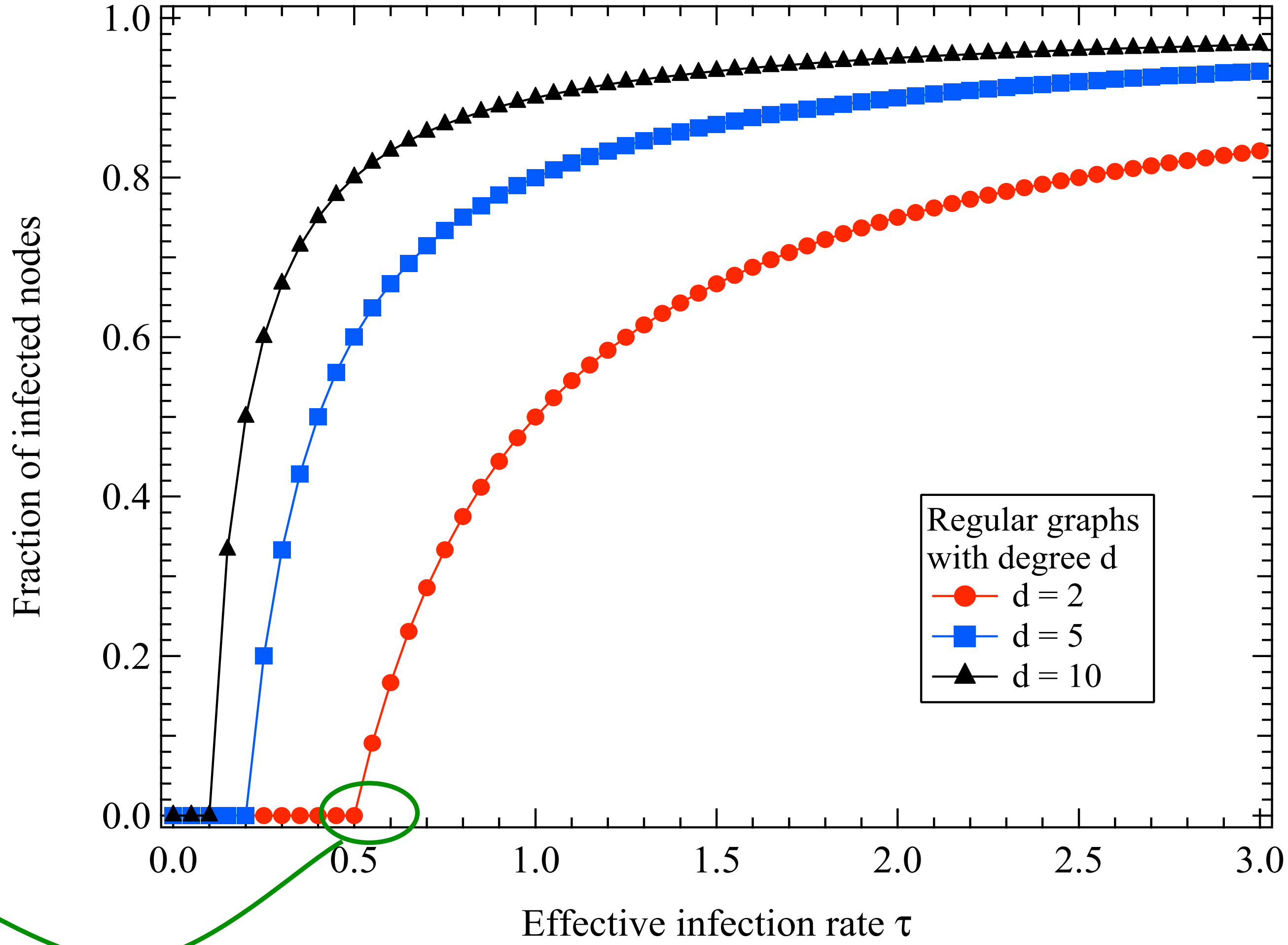
P. Van Mieghem, 2012, “The Viral Conductance in Networks”, Computer Communications, Vol 35, No. 12, pp. 1494-1509.

What is so interesting about epidemics?

β : infection rate per link

δ : curing rate per node

τ= β/ δ : effective spreading rate



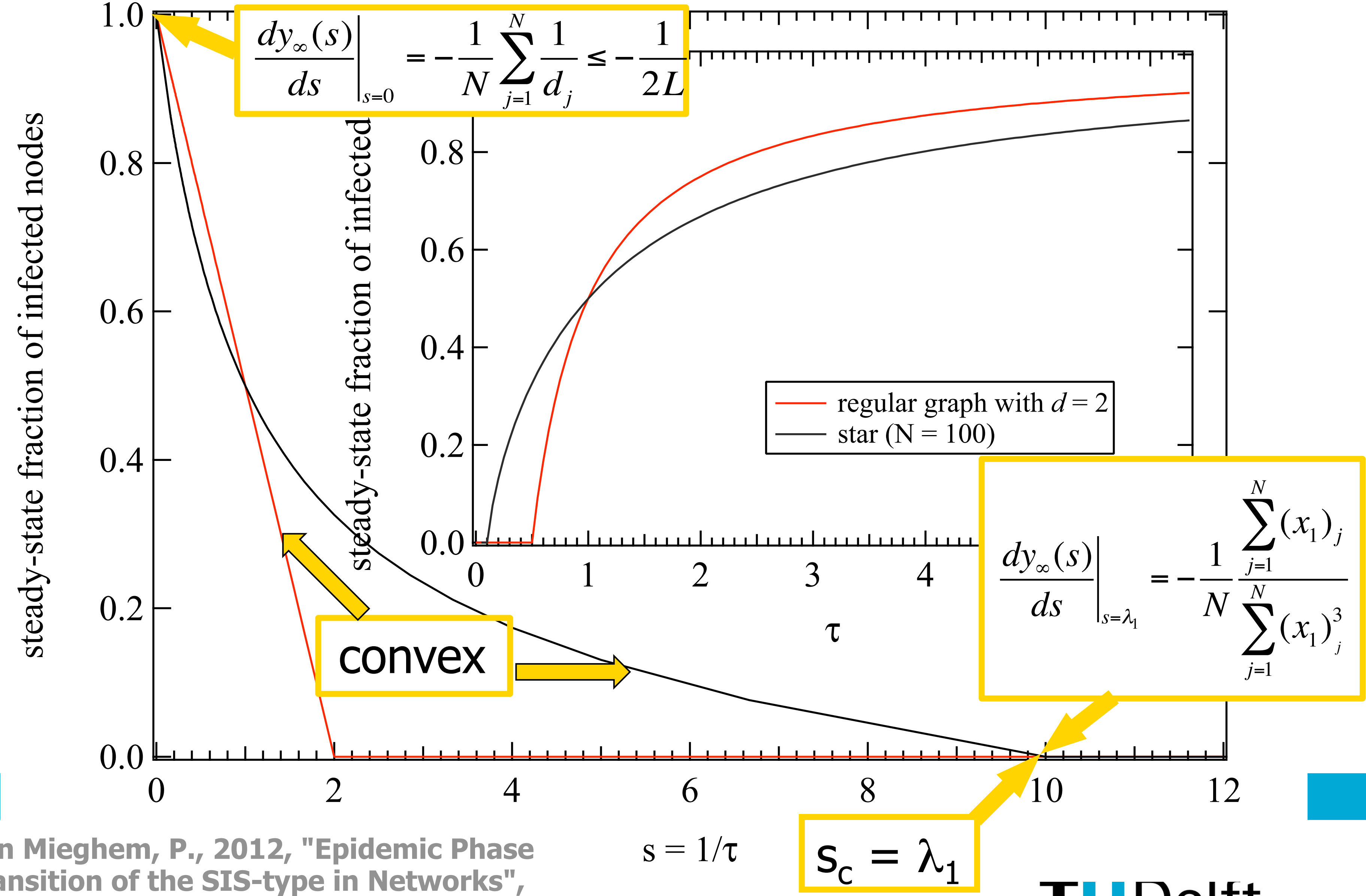
- **Final epidemic state**
- **Rate of propagation**
- **Epidemic threshold**

$$\tau_c = \frac{1}{\lambda_1(A)}$$

$$\max\left(E[D]\sqrt{1+\frac{Var[D]}{(E[D])^2}}, \sqrt{d_{\max}} \right) \le \lambda_1(A) \le d_{\max}$$

TU Delft

Transformation $s = \dfrac{1}{\tau}$ & principal eigenvector



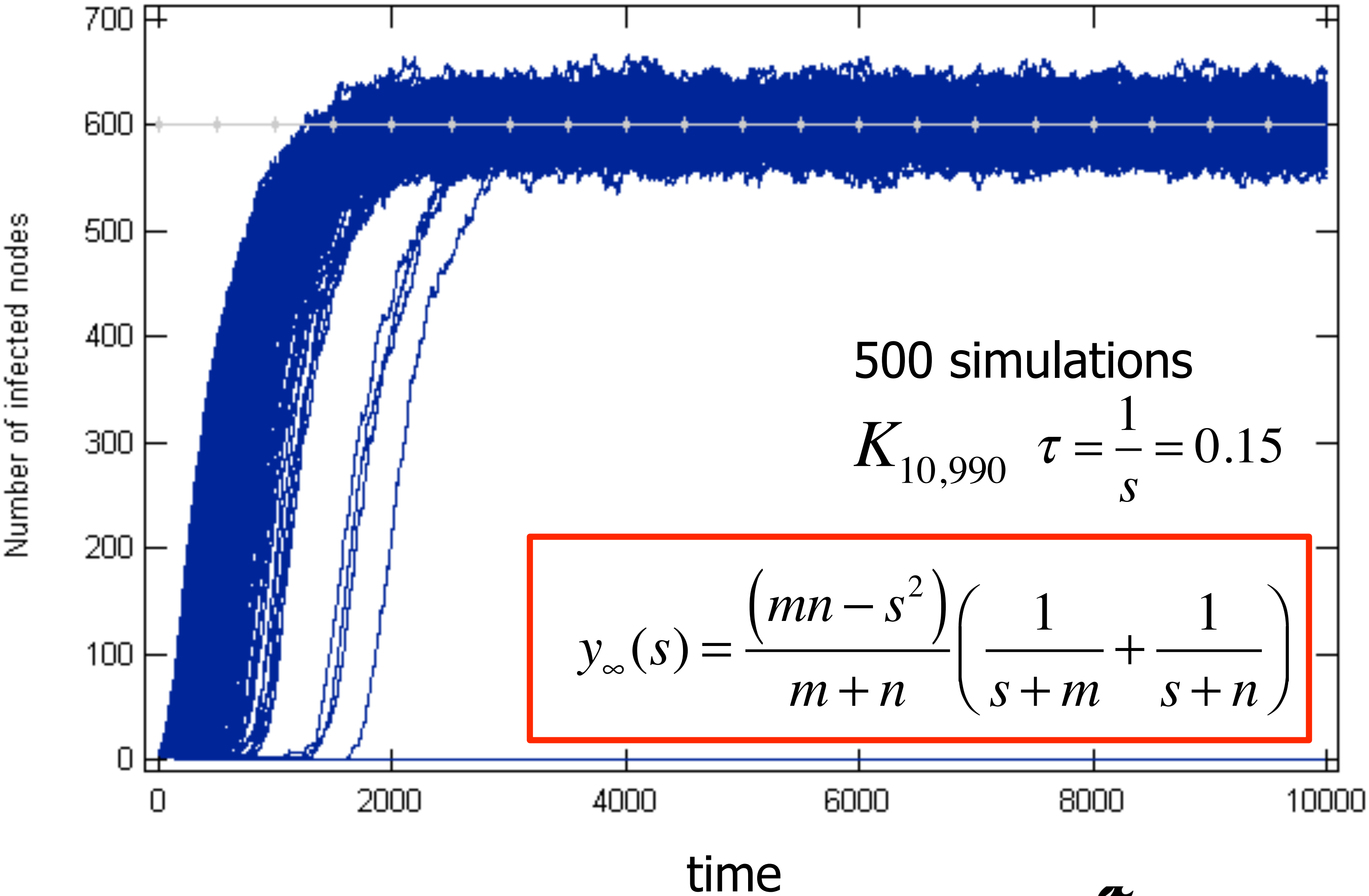
$$\left.\frac{dy_\infty(s)}{ds}\right|_{s=0} = -\frac{1}{N}\sum_{j=1}^{N}\frac{1}{d_j} \le -\frac{1}{2L}$$

$$\left.\frac{dy_\infty(s)}{ds}\right|_{s=\lambda_1} = -\frac{1}{N}\frac{\sum_{j=1}^{N}(x_1)_j}{\sum_{j=1}^{N}(x_1)_j^3}$$

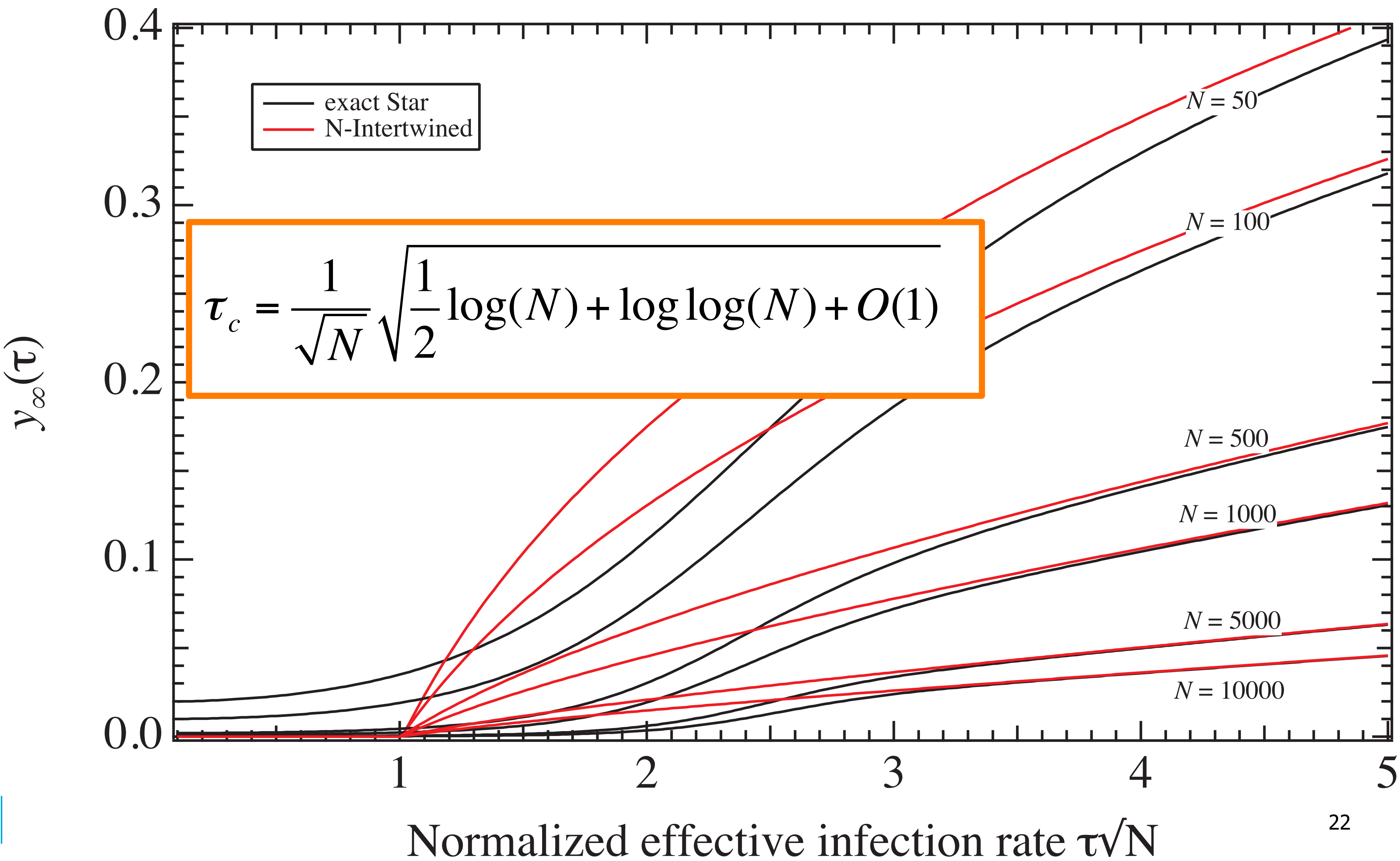
Van Mieghem, P., 2012, "Epidemic Phase Transition of the SIS-type in Networks", Europhysics Letters (EPL), Vol. 97, Februari, p. 48004.

TU Delft

Simulations



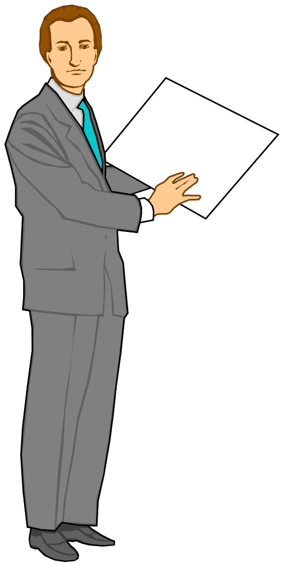
Star graph



$$\tau_c = \frac{1}{\sqrt{N}}\sqrt{\frac{1}{2}\log(N) + \log\log(N) + O(1)}$$

E. Cator and P. Van Mieghem, 2012, **"SIS epidemics on the complete graph and the star graph: exact analysis"**, unpublished.

Outline



Introduction & Definitions

Exact SIS model

The N-intertwined MF approximation

Extensions

Summary

Heterogeneous virus spread

- The N-intertwined model is extended to a hetergeneous setting:

$$\boxed{\frac{dV(t)}{dt} = A\,diag(\beta_i)V(t) - diag(v_i(t))\left(A\,diag(\beta_i)V(t) + C\right)}$$

where the curing rate vector $C^T = [\delta_1\ \delta_2 \ldots \delta_N]$

- *Results:*
  - Extended multi-dim. threshold for virus spread
  - Generalized Laplacian that extends the classical Laplacian of a graph: $\boxed{Q(q_k) = diag(q_k) - A}$
  - Strong convexity v_k with respect to δ_k, concave with respect to others δ_j (j different from k).
  - Choose C vector in network protection via game theory

- J. Omic, P. Van Mieghem, and A. Orda, "Game Theory and Computer Viruses", *IEEE Infocom09.*
- P. Van Mieghem and J. Omic, "In-homogeneous Virus Spread in Networks", TUDelft report (see my website)

Extensions of the N-intertwined model

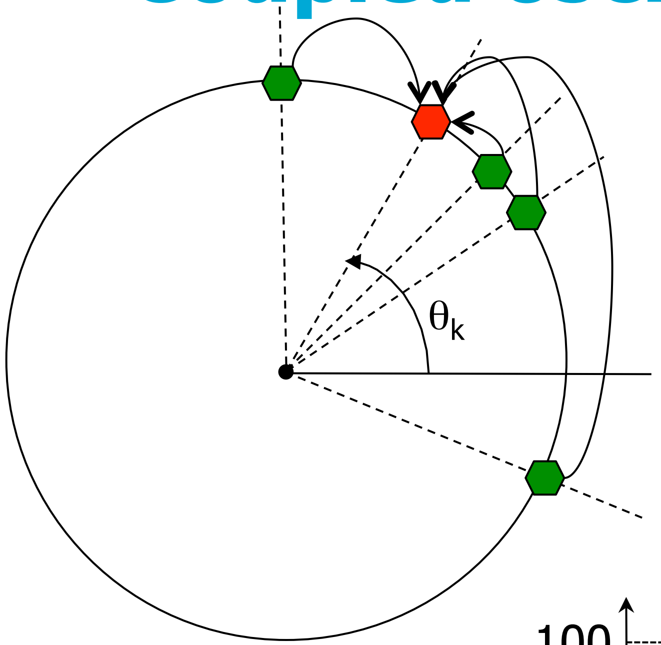
- SAIS instead of SIS:
  - From 2 states (Infected and Susceptible) to a 3-states (Infected, Susceptible, Alert)
  - "Epidemic Spread in Human Networks", F. Darabi Sahneh and C. Scoglio, 50th IEEE Conf. Decision and Contol, Orlando, Florida (2011)
- SIR instead of SIS:
  - "An individual-based approach to SIR epidemics in contact networks", M. Youssef and C. Scoglio, Journal of Theoretical Biology 283, pp. 136-144, (2011).
- Very general extension: m compartments (includes both SIS, SAIS, SIR,...):
  - "Generalized Epidemic Mean-Field Model for Spreading Processes over Multi-Layer Complex Networks", F. Darabi Sahneh, C. Scoglio, P. Van Mieghem, submitted IEEE/ACM Transactions on Networking

Affecting the epidemic threshold

- Degree-preserving rewiring (assortativity of the graph)
  - Van Mieghem, P., H. Wang, X. Ge, S. Tang and F. A. Kuipers, 2010, "Influence of Assortativity and Degree-preserving Rewiring on the Spectra of Networks", The European Physical Journal B, vol. 76, No. 4, pp. 643-652.
  - Van Mieghem, P., X. Ge, P. Schumm, S. Trajanovski and H. Wang, 2010, "Spectral Graph Analysis of Modularity and Assortativity", Physical Review E, Vol. 82, November, p. 056113.
  - Li, C., H. Wang and P. Van Mieghem, 2012, "Degree and Principal Eigenvectors in Complex Networks", IFIP Networking 2012, May 21-25, Prague, Czech Republic.
- Removing links/nodes (optimal way is NP-complete)
  - Van Mieghem, P., D. Stevanovic, F. A. Kuipers, C. Li, R. van de Bovenkamp, D. Liu and H. Wang, 2011,"Decreasing the spectral radius of a graph by link removals", Physical Review E, Vol. 84, No. 1, July, p. 016101.
- Quarantining and network protection
  - Omic, J., J. Martin Hernandez and P. Van Mieghem, 2010, "Network protection against worms and cascading failures using modularity partitioning", 22nd International Teletraffic Congress (ITC 22), September 7-9, Amsterdam, Netherlands.
  - Gourdin, E., J. Omic and P. Van Mieghem, 2011, "Optimization of network protection against virus spread", 8th International Workshop on Design of Reliable Communication Networks (DRCN 2011), October 10-12, Krakow, Poland.

E.R. van Dam, R.E. Kooij, The minimal spectral radius of graphs with a given diameter, *Linear Algebra and its Applications,* 423, 2007, pp. 408-419.

Coupled oscillators (Kuramoto model)

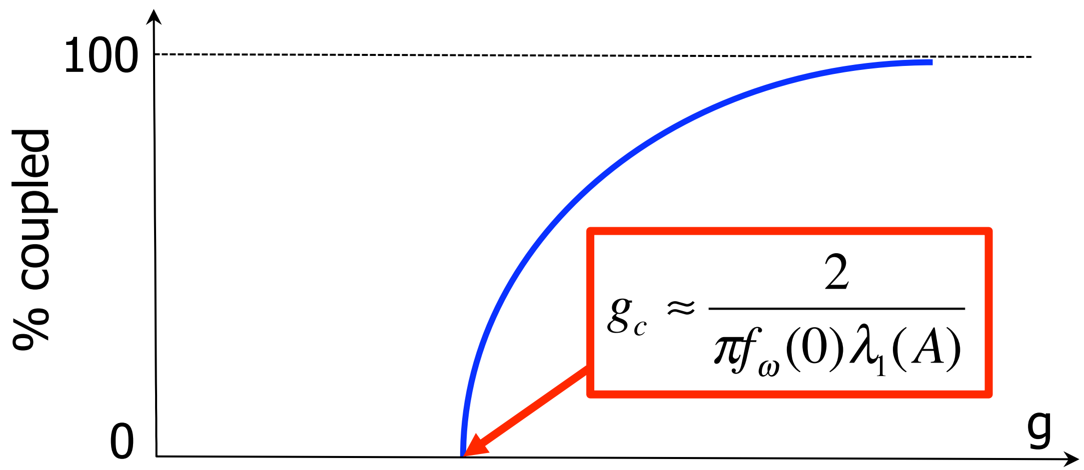


Interaction equals sums of sinus of phase difference of each neighbor:

$$\dot{\theta}_k = \omega_k + g\sum_{j=1}^{N} a_{kj} \sin\left(\theta_j - \theta_k\right)$$

coupling strength

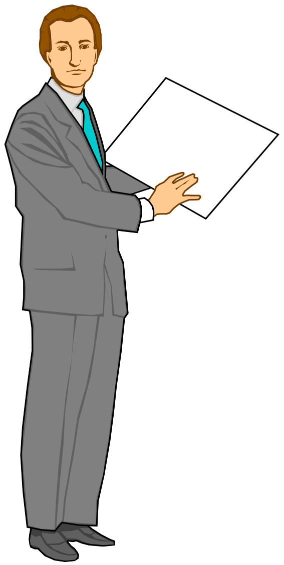
natural frequency



J. G. Restrepo, E. Ott, and B. R. Hunt. Onset of synchronization in large networks of coupled oscillators, Phys. Rev. E, vol. 71, 036151, 2005

TU Delft

Outline



Introduction & Definitions

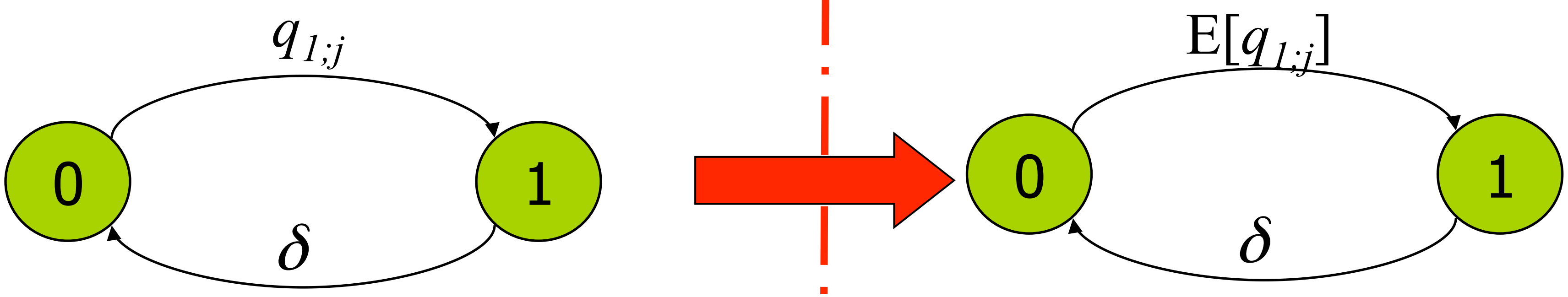
Exact SIS model

The N-intertwined MF approximation

Extensions

$\boxed{\text{Summary}}$

Mean-field approximation



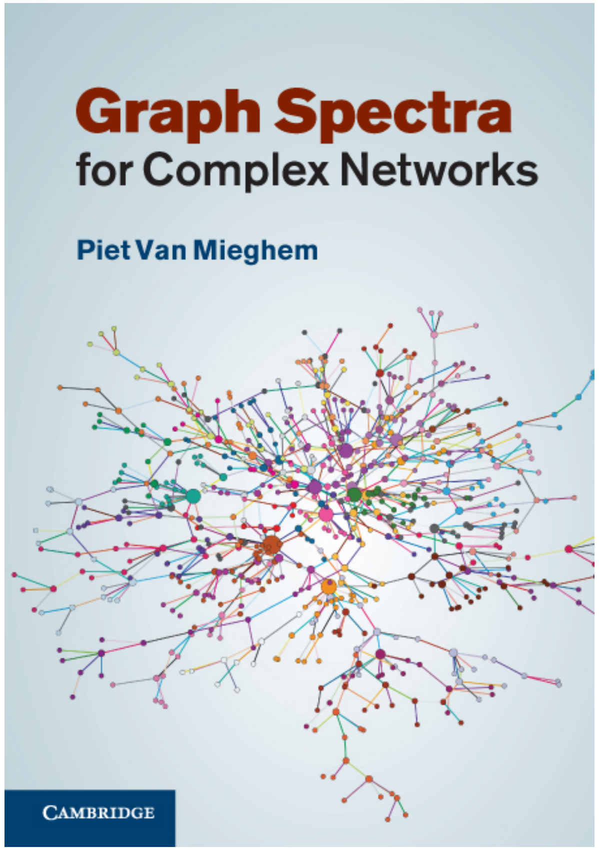
- 2^N linear equations
- Steady-state
  - absorbing (healthy) state
  - reached after unrealistically long time
- difficult to analyze

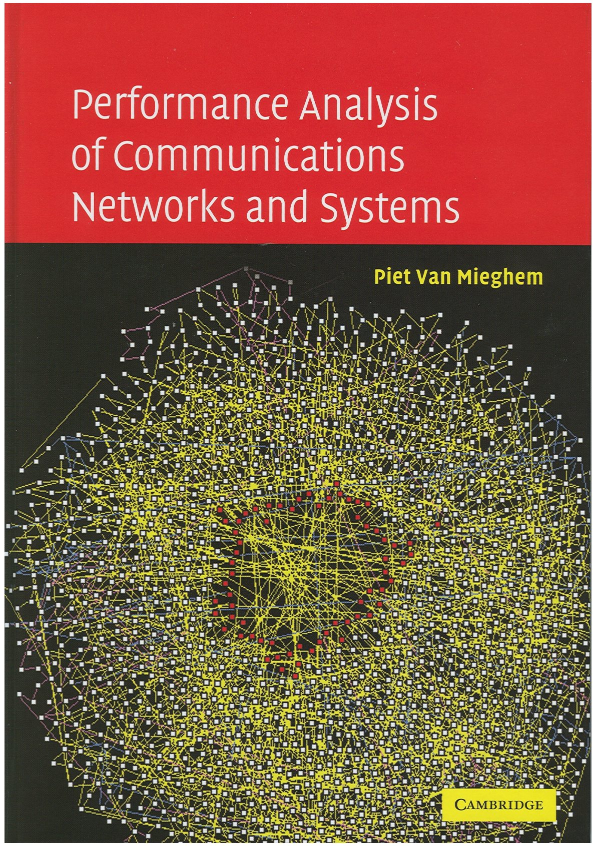
- *N* non-linear equations
- Meta-stable state:
  - phase-transition
  - epidemic threshold
  - realistic
- analytically tractable

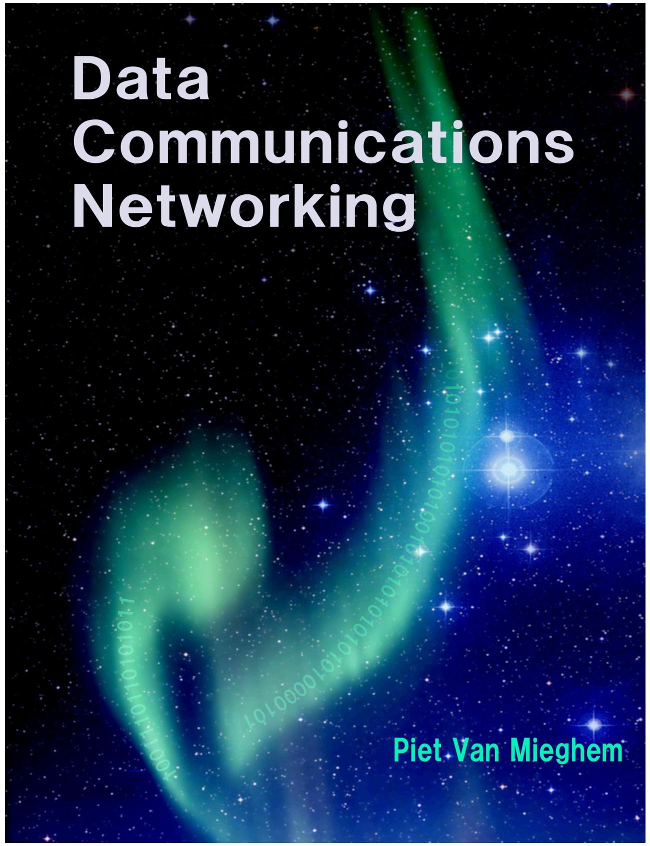
Agenda for future research

- Accuracy of *N*-intertwined mean-field approximation
- Exact computations for graphs beside the complete graph and the star
- Coupling of the virus spread process and the underlying topology (adaptive networks)
- Multiple, simultaneous viruses on a network
- Eigenvectors and eigenvalues of a graph: what do they "physically" mean?

Books







Articles: **http://www.nas.ewi.tudelft.nl**

Thank You

Piet Van Mieghem
NAS, TUDelft
P.F.A.VanMieghem@tudelft.nl

TUDelft