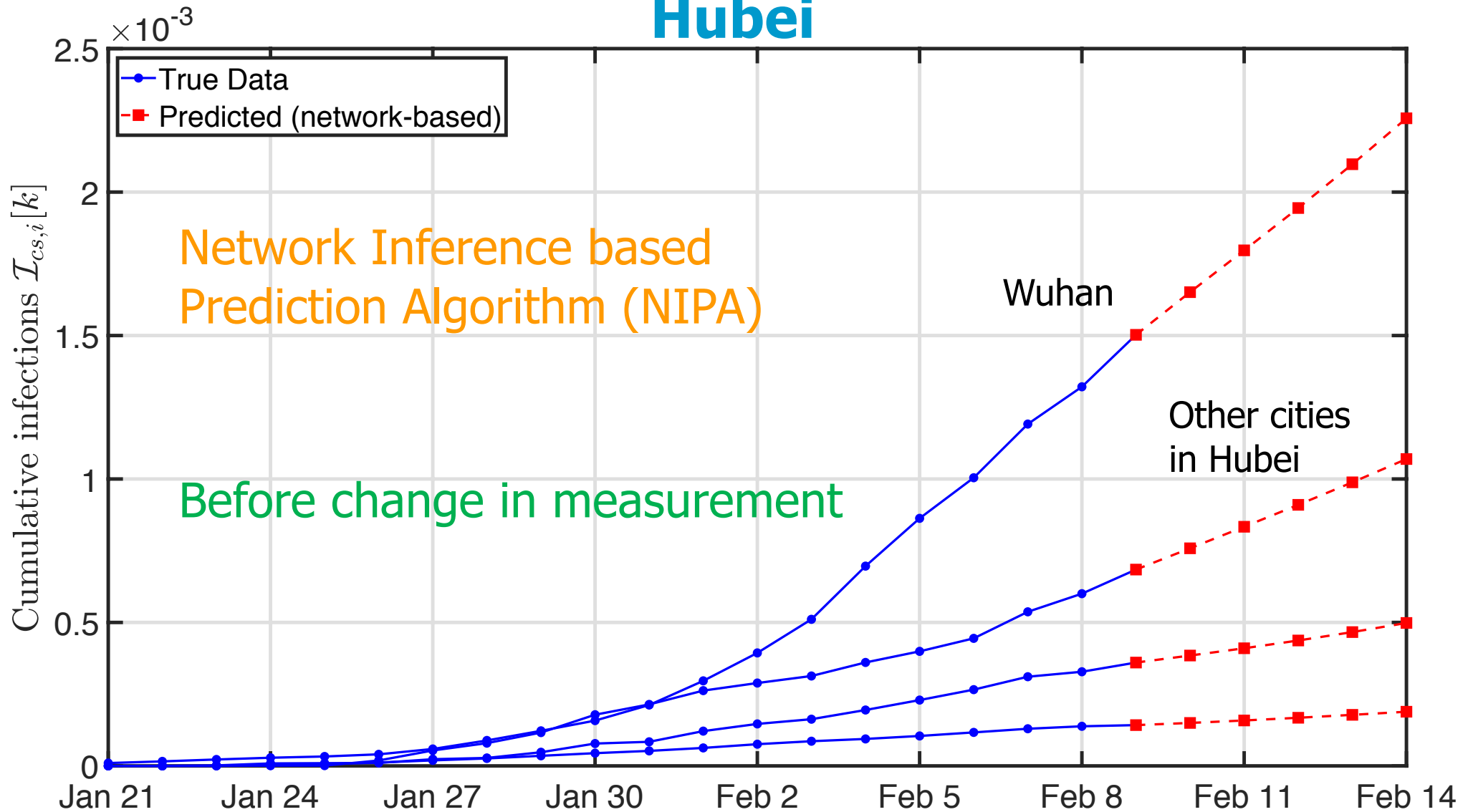


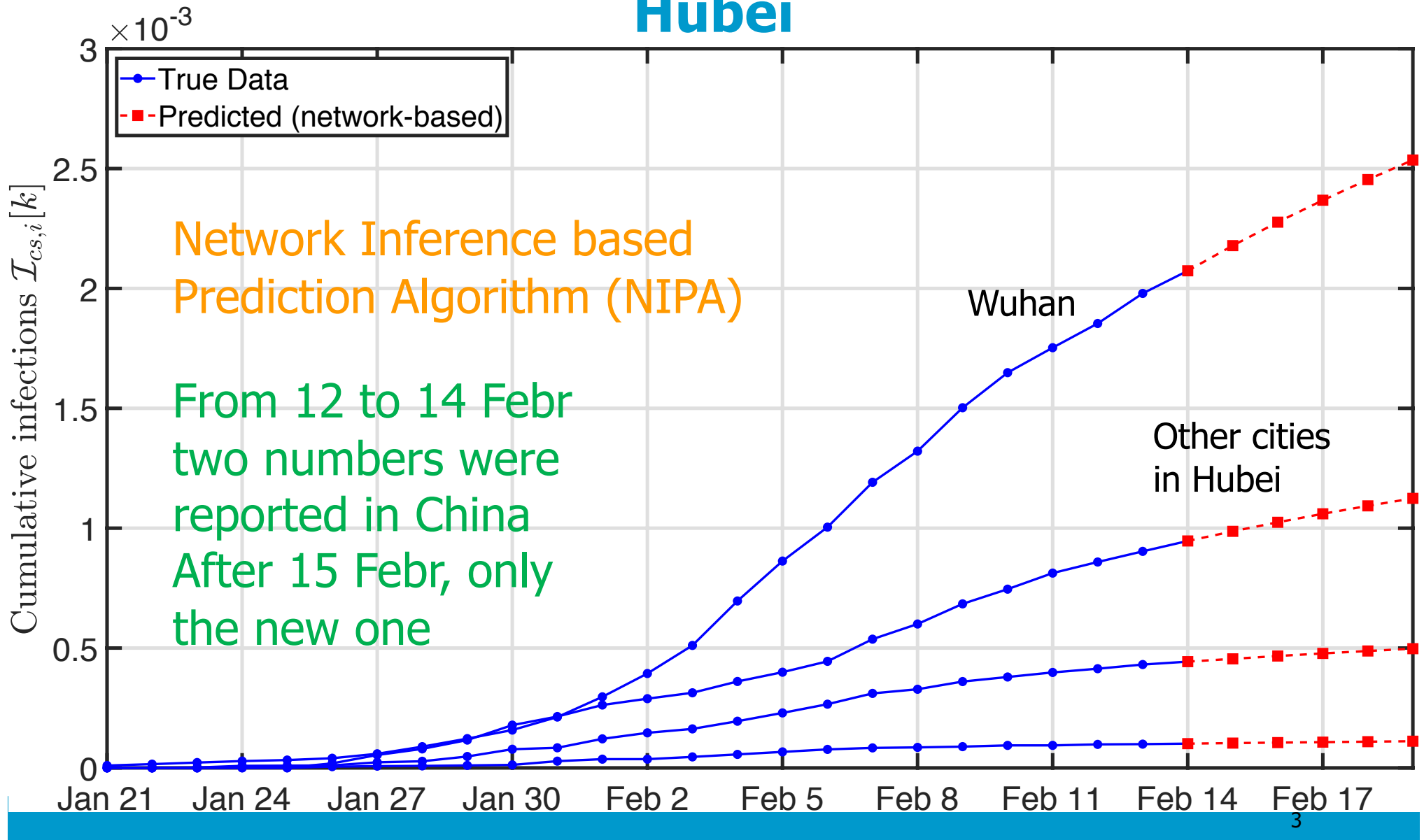
Network Reconstruction from Viral State Observations and Prediction of the Epidemic Nodal State

Bastian Prasse & Piet Van Mieghem

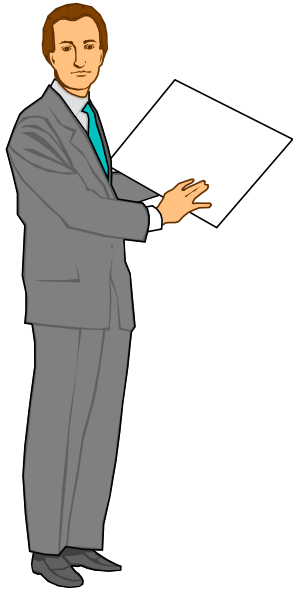
Expected Cum. Fraction of Covid-19 infected in Hubei



Expected Cum. Fraction of Covid-19 infected in Hubei



Outline



Epidemic model

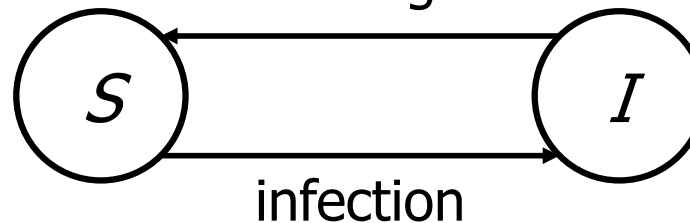
Network reconstruction

Reconstruction and prediction accuracy

Compartmental Epidemic Models

- Simplest epidemic model: every individual is either *Infected (I)* or *Susceptible (S)*

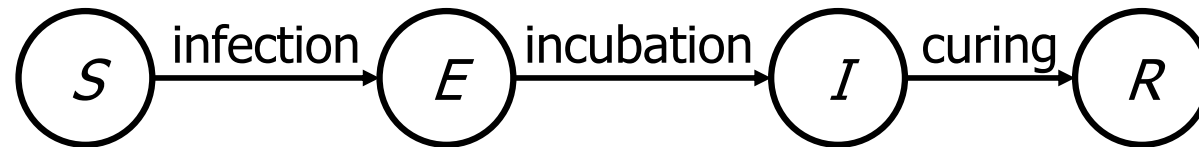
SIS epidemic model:



- More compartments may be more accurate:

E.g., *Exposed (E)* but not contagious and *Recovered (R)*

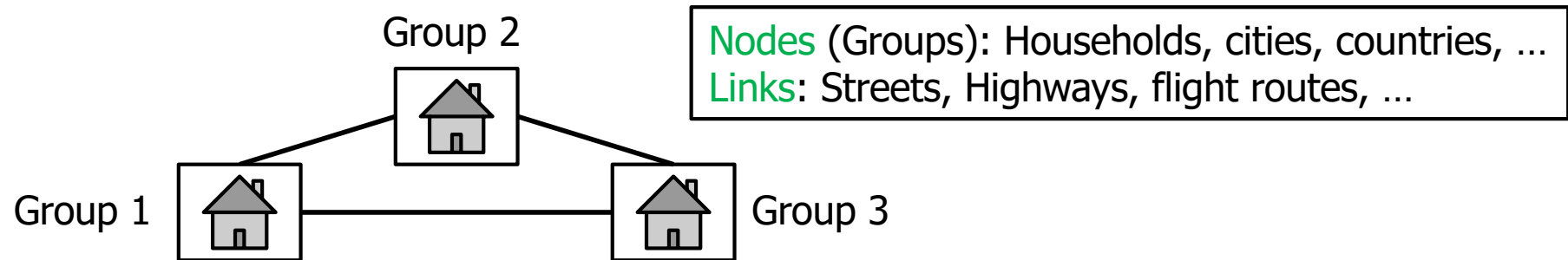
SEIR epidemic model:



- Almost all compartmental epidemic models are special cases of GEMF (Generalized Epidemic Mean-Field)
- Here: focus on SIS model, but results apply to all GEMF models

Group-Based SIS Epidemic Model

Epidemic spread between N groups of individuals



At any *discrete* time $k = 1, 2, \dots$, every group $i = 1, \dots, N$ has a viral state vector $v_i[k] = (S_i[k], I_i[k])^T$

- $S_i[k]$: fraction of **Susceptible** individuals in group i
- $I_i[k]$: fraction of **Infected** individuals in group i
 - $I_i[k] + S_i[k] = 1$ for every group i at every time k

SIS Governing Mean-field Equations

- Nonlinear difference equation of infection probability $I_i[k]$ of node i :

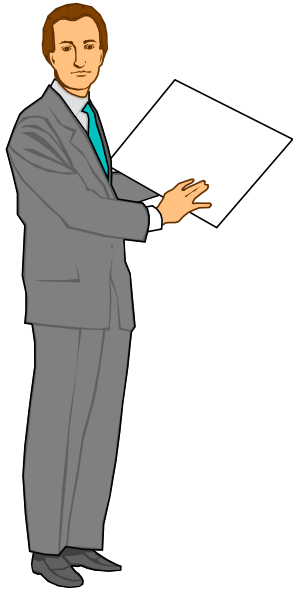
$$I_i[k + 1] = \underbrace{(1 - \delta_i)I_i[k]}_{\text{curing}} + \underbrace{S_i[k] \sum_{j=1}^N \beta_{ij} I_j[k]}_{\text{infections}}$$

- δ_i : Curing probability of group i
- β_{ij} : Infection probability from group j to group i
- Group interactions are specified by the *infection probability matrix*

$$B = \begin{pmatrix} \beta_{11} & \cdots & \beta_{1N} \\ \vdots & \ddots & \vdots \\ \beta_{N1} & \cdots & \beta_{NN} \end{pmatrix}$$

- *Problem* : matrix B unknown for real epidemics
- *Solution attempt* : estimate B from observing the epidemic

Outline

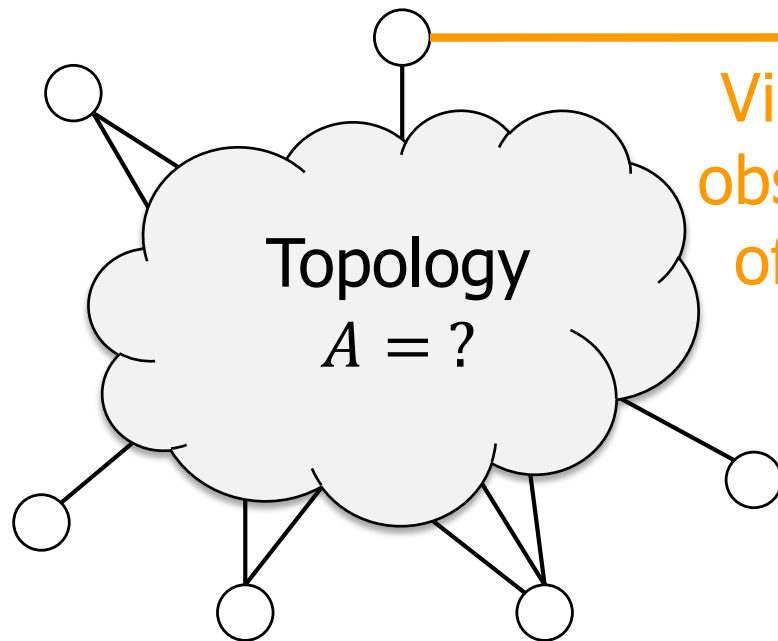


Epidemic model

Network reconstruction

Reconstruction and prediction accuracy

Network Reconstruction from Epidemics (1)



Viral state
observation
of node i

$$\{x_i[1], x_i[2], \dots, x_i[n]\}$$

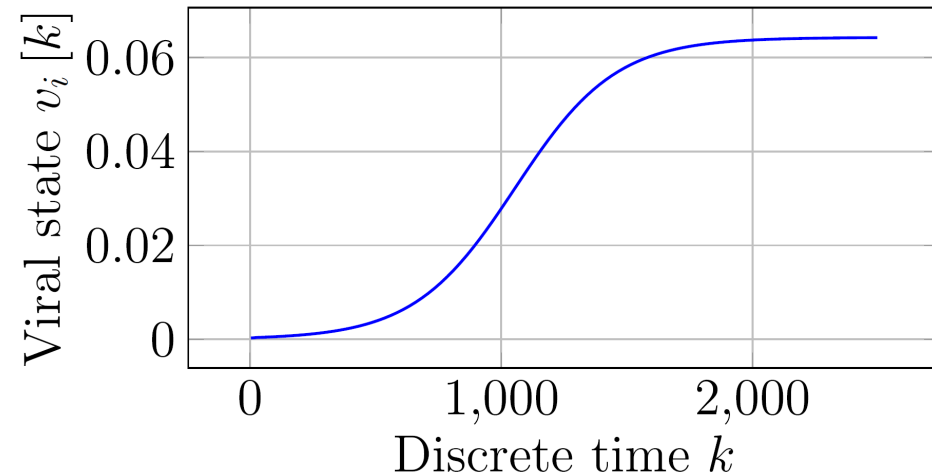
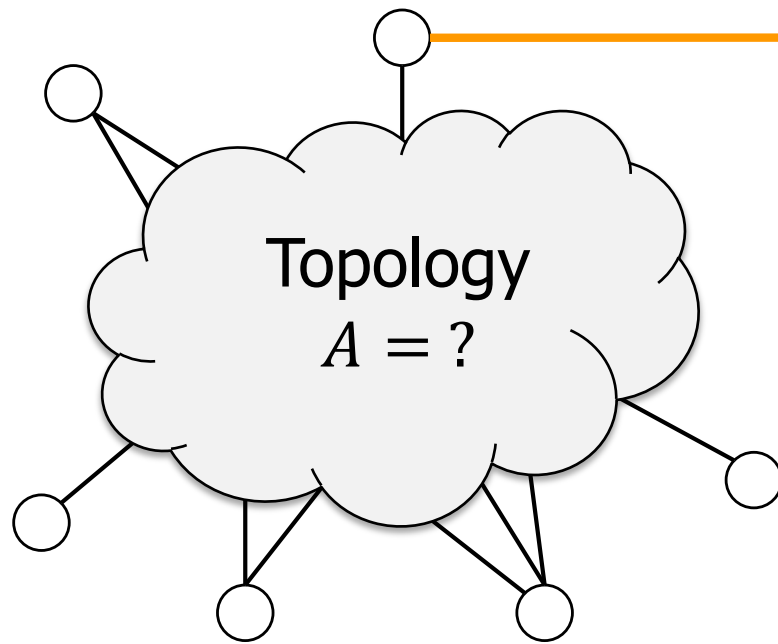
$x_i[k] \in \{0,1\}$ is the viral SIS
state of node i at discrete time k

$x_i[k] = 0$: node i is healthy
 $x_i[k] = 1$: node i is infected

➔ **Objective:** Infer the $N \times N$ adjacency matrix A of the network from viral SIS state observations of all nodes

➔ **Solution:** **Infeasible** (NP-hard)

Network Reconstruction from Epidemics (2)



➔ **Objective:** Infer the $N \times N$ adjacency matrix A of the graph from a set of viral probabilities over time of all nodes

➔ **Solution:** possible when using a mean-field model

Key observation: Governing mean-field equations are linear in the (weighted) adjacency matrix B

Network Inference based Prediction algorithm (NIPA)



n : number of observations

Network reconstruction is *equivalent* to a linear system for every group i :

$$v_i = F_i \begin{pmatrix} \beta_{i1} \\ \vdots \\ \beta_{iN} \end{pmatrix}$$

Vector v_i and matrix F_i follow from the viral state observations $I_i[1], \dots, I_i[n]$

Obtain matrix estimate \hat{B} by solving LASSO:

$$\min_{\beta_{i1}, \dots, \beta_{iN}} \left\| v_i - F_i \begin{pmatrix} \beta_{i1} \\ \vdots \\ \beta_{iN} \end{pmatrix} \right\|^2 + \rho_i \sum_{j=1}^N |\beta_{ij}| \quad \text{for } i = 1, \dots, N$$

Network Reconstruction as Linear System

- Network reconstruction results in set of linear equations:

$$Fb = v$$

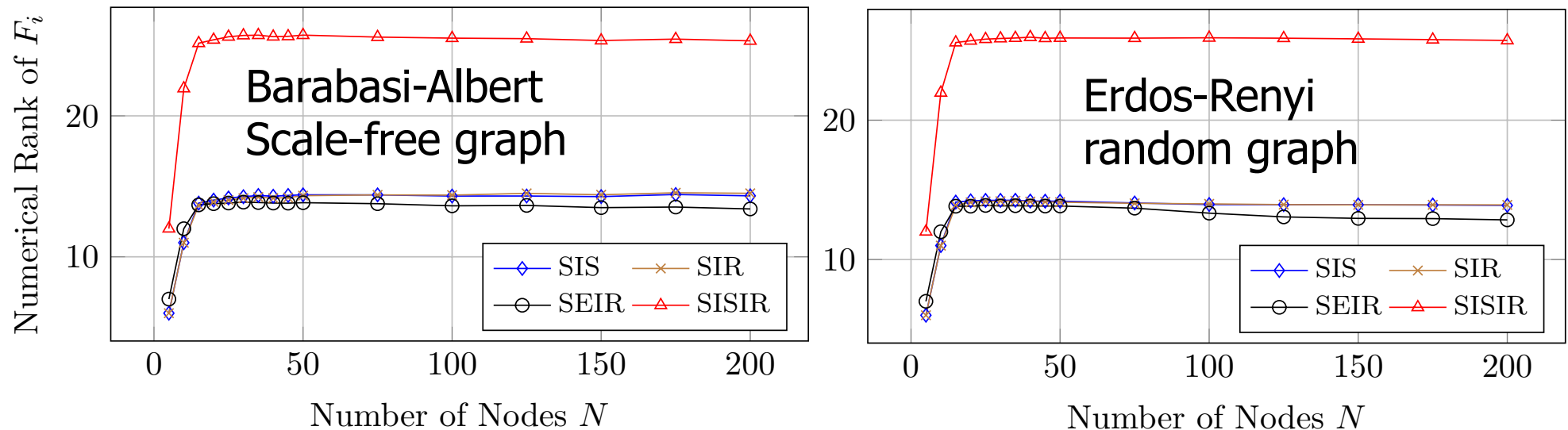
- Matrix F and vector v : transformations of the viral state observations $v_i[0], v_i[1], \dots$ of every node i
- **Problem:** Matrix F is extremely ill-conditioned
 - Lagrangian optimization with constraints (multipliers)
 - A-priori estimates of the matrix B can be taken into account (Bayesian): e.g. mobility pattern between cities
 - convexity (efficient algorithms)
 - *basis pursuit* (L1-norm regularisation)

$$\min_B \|B\|_1$$

$$s. t. \quad Fb = v$$

✧ good without model errors nor a-priori estimate of the graph

Rank problem of Matrix F in $Fb=v$

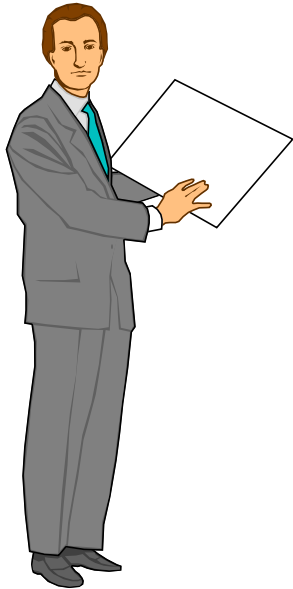


Intuition: only the initial state agitates eigenstates; thereafter the epidemics process autonomously tends to its steady-state mainly steered by the principal eigenvector (corresponding to the largest eigenvalue)

Solution of the rank problem of matrix F :

- Multiple outbreaks/realizations of epidemic on same network
- Assume that the curing rate changes around a mean value (curing rate control)
- Singular values of F decrease exponentially. Hence, model errors have significant effect

Outline

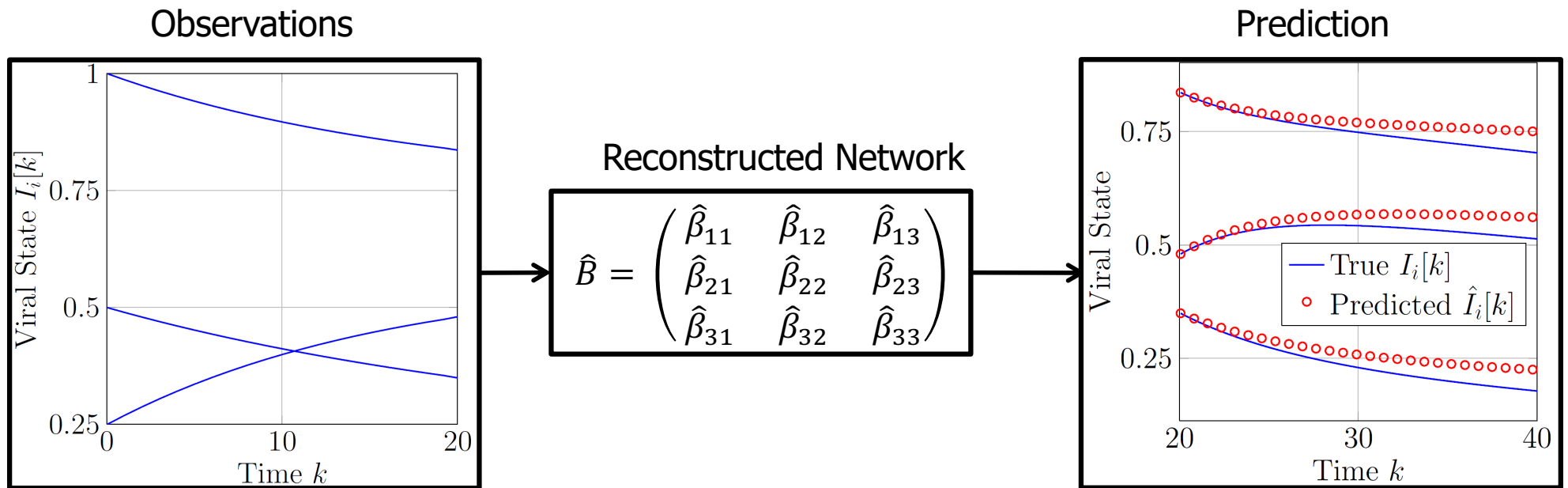


Epidemic model

Network reconstruction

Reconstruction and prediction accuracy

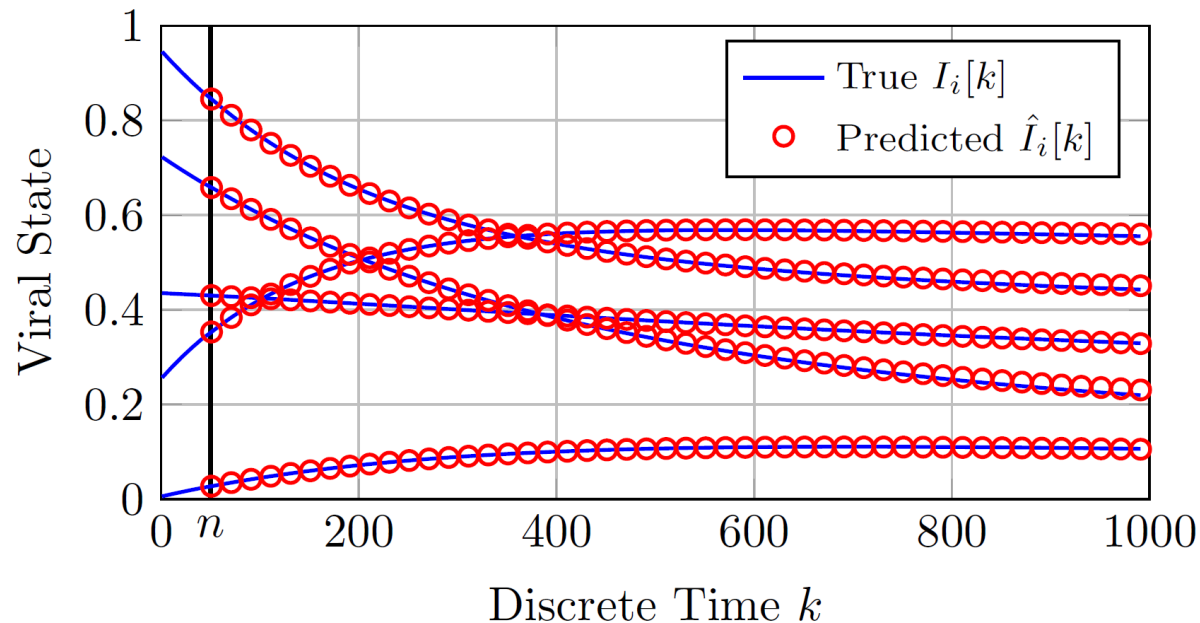
Prediction of Epidemic Outbreaks



- True viral state $I_i[k]$: generated on the true network B
- Predicted viral state $\hat{I}_i[k]$: generated on the reconstructed network \hat{B}

Prediction Accuracy

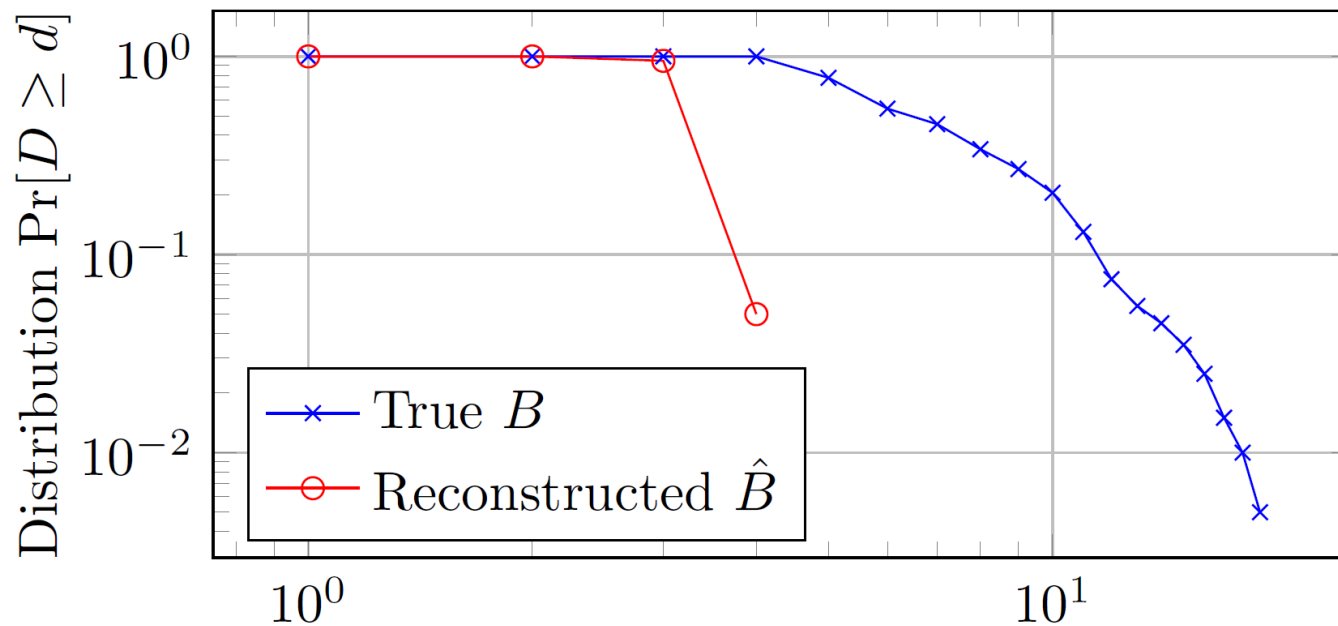
- Barabasi-Albert network with $N = 200$ nodes
- Only few observations: $n = 50$



- Very accurate prediction at every time $k \geq 50$
 - Does the prediction accuracy imply $\hat{B} \approx B$?

Network Reconstruction Accuracy

1. AUC score: **AUC = 0.51**
 - Tossing a coin for reconstructing every link \rightarrow AUC = 0.5
2. In-degree distribution:



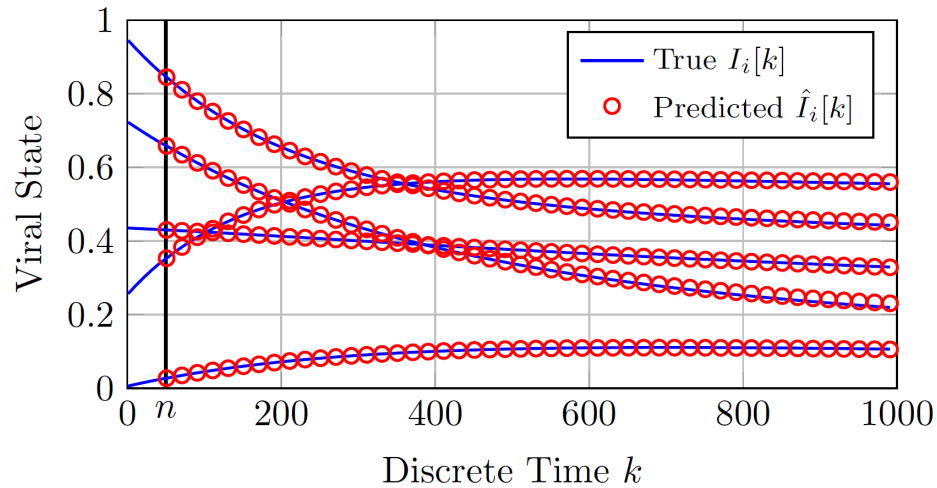
In-Degree d

True network B : Scale-free

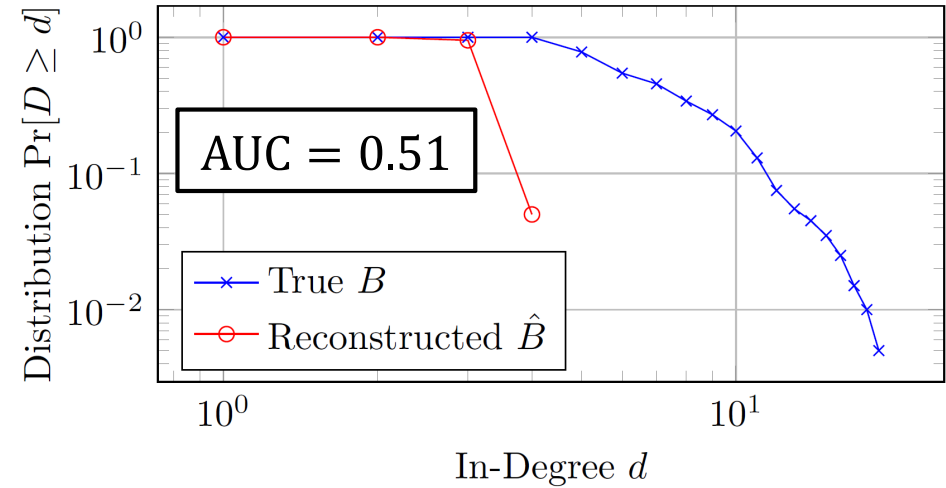
Reconstructed network \hat{B} : Regular

Conclusion

Prediction

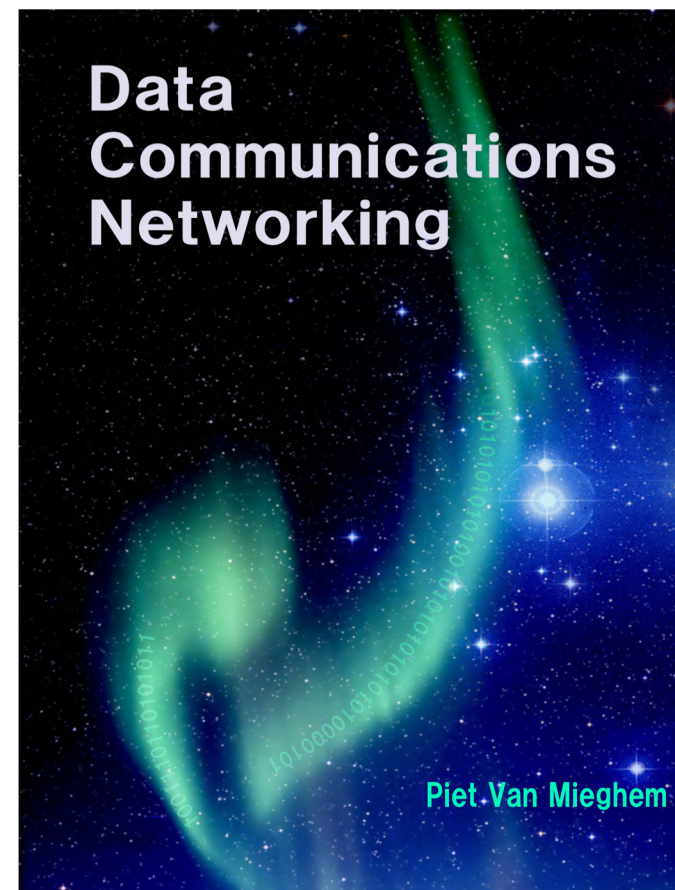
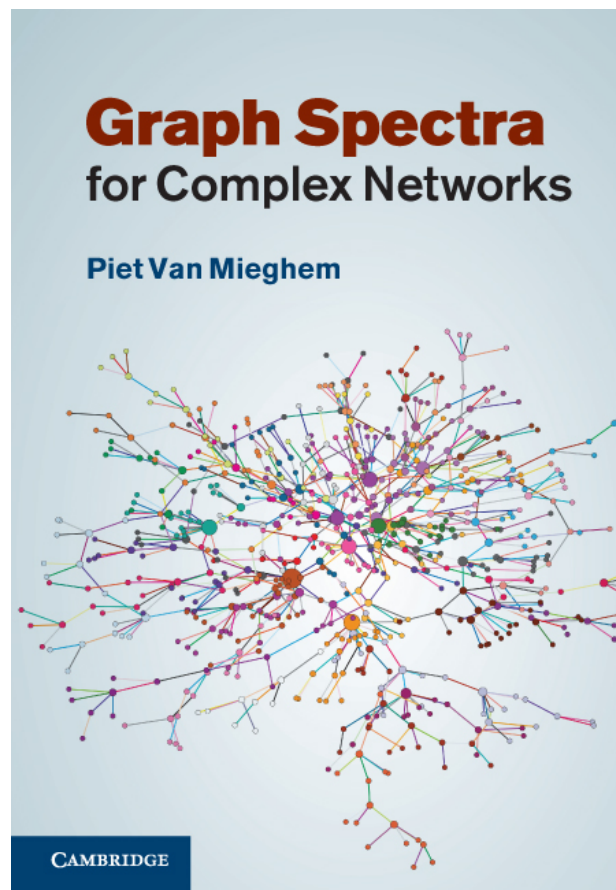
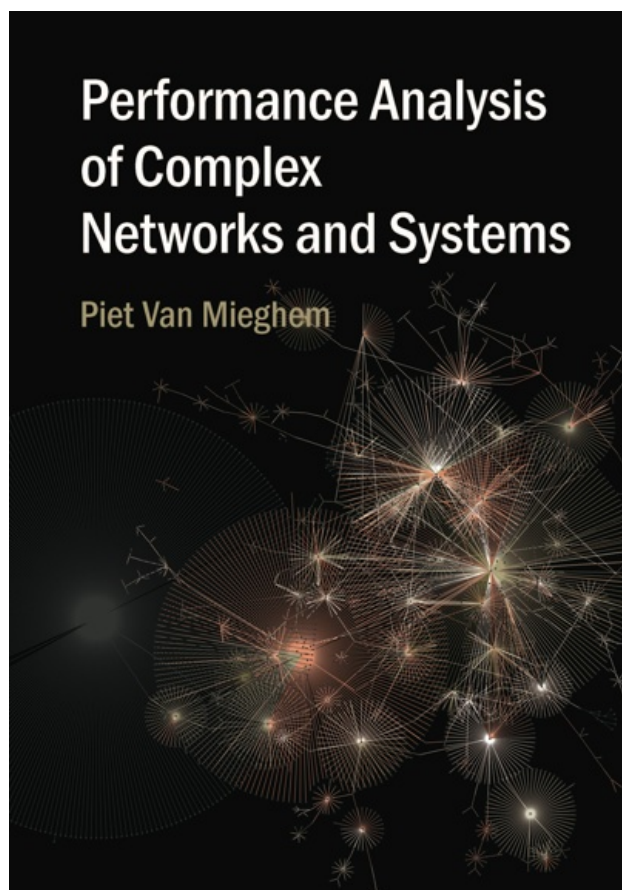


Reconstruction



Accurate prediction of epidemic outbreaks
without accurate network reconstruction

Books



Articles: <http://www.nas.ewi.tudelft.nl>



Thank You

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

20