# Markovian epidemics on a graph: a computational challenge

P. Van Mieghem*

5 February 2024

**Abstract**

After a brief description of the infinitesimal generator matrix of a Markov chain for SIS epidemics on a fixed graph, we motivate its study by pointing to the epidemic phase transition that is exactly embedded in its governing equations. The major computational challenge lies in the size of the Markov state space and its corresponding infinitesimal generator matrix, which grows as $2^N$ for a graph with $N$ nodes.
This note is a quest for *any means,* classical or quantum, to compute the Markovian SIS process for realistic graph sizes $N$.

## 1 Introduction

We consider epidemic spread in a contact graph $G$ that represents a set $\mathcal{N}$ of $N$ individuals as nodes and specifies the $L$ contacts between all pairs of individuals as links [10]. We assume that the graph $G$, characterized [14] by a symmetric adjacency matrix $A$, is fixed and does not change over time. Epidemic spread on a graph is one of the simplest, non-trivial diffusion processes in networks that not only models biological disease spread (e.g. Covid) and digital computer viruses and malware in the Internet, but also social contagion in on-line social platforms (e.g. Twitter and Facebook), rumor spread, cascades of failures in infrastructural networks as the Internet and power grids and brain anomalies such as epileptic seizures [7, 8] and other real-world diffusion applications in graphs.

The class of Susceptible-Infected-Susceptible (SIS) epidemics is the simplest compartmental model of a disease spread with re-infections in a population, in which individuals are either infectious (I) or healthy, but susceptible (S). Other compartmental models can be described [12], both stochastically and deterministically (after a mean-field approximation), similarly as an SIS epidemic.

## 2 Markovian SIS epidemics on a graph

The viral state of a node $k$ at time $t$ is specified by a Bernoulli random variable $X_k(t) \in \{0, 1\}$: $X_k(t) = 0$ for a healthy node and $X_k(t) = 1$ for an infected node. A node $k$ at time $t$ can be in one of the two states: *infected*, with probability $w_k(t) = \Pr[X_k(t) = 1]$ or *healthy*, with probability $1 - w_k(t)$, but susceptible to the virus. We assume that the curing process per node $k$ is a Poisson process

*Delft University of Technology, Faculty of Electrical Engineering, Mathematics and Computer Science, P.O Box 5031, 2600 GA Delft, The Netherlands; *email*: P.F.A.VanMieghem@tudelft.nl

with rate $\delta$ and that the infection process per link is a Poisson process with rate $\beta$. The effective infection rate is $\tau = \frac{\beta}{\delta}$. Obviously, only when a node is infected, it can infect its direct neighbors, that are still healthy. Both the curing and infection Poisson process are independent. This is the general continuous-time description of the simplest type of a Susceptible-Infected-Susceptible (SIS) virus process on a network. Occasionally, a third, independent self-infection process with self-infection rate $\varepsilon$ is considered, which describes background or indirect infections. Infections may happen either through direct contact or indirectly, for example, after touching infected surfaces or inhaling air in a closed room previously contaminated by an infected individual. The Markovian $\varepsilon$-SIS model consists of three, independent Poisson processes: (i) the curing process with rate $\delta$, (ii) infection process with rate $\beta$ and (iii) self-infection process with rate $\varepsilon$.

A description of the $\varepsilon-$SIS epidemic process is as follows [15, Section 17.2-17.3]. Let $I$ denote the set of infected nodes in the graph $G$ and let $a_{ij}$ be the element of the adjacency matrix $A$. Then, the Markov transitions

$$\begin{cases} \text{for } j \notin I\text{:} & I \mapsto I \cup \{j\} & \text{at rate } \beta \sum_{k \in I} a_{kj} + \varepsilon \\ \text{for } i \in I\text{:} & I \mapsto I \backslash \{i\} & \text{at rate } \delta \end{cases} \tag{1}$$

detail the dynamics between the infected subgraph $I$ and its complement $I^c = G \backslash I$, the subgraph of healthy nodes.

The time-dependent $\varepsilon$-SIS process can be described as a continuous-time Markov chain with $2^N$ states [20],[19]. Computationally, enumerating the subgraphs $I$ in $G$ leads to the governing equation (2). Indeed, representing the Markov state $i$ as $i = \sum_{k=1}^{N} x_k(i)\, 2^{k-1}$, where the binary $k$-th digit $x_k(i)$ represents the infectious state of a node $k$ in the network, the time dependence of the probability state vector $s(t)$ in $\varepsilon$-SIS epidemics, with components

$$s_i(t) = \Pr[X_1(t) = x_1(i), X_2(t) = x_2(i), ..., X_N(t) = x_N(i)]$$

and normalization $\sum_{i=0}^{2^N-1} s_i(t) = 1$, obeys the differential equation

$$\frac{ds(t)}{dt} = -Qs(t) \tag{2}$$

where the $2^N \times 2^N$ infinitesimal generator $-Q$ is specified in [20] and [19]. The solution of the matrix differential equation is

$$s(t) = e^{-Qt} s(0) \tag{3}$$

and, for self-infection rate $\varepsilon > 0$, a non-trivial[1] $2^N \times 1$ steady-state vector $s_\infty$ exists, that obeys $Qs_\infty = 0$ and which is the right-eigenvector belonging to zero eigenvalue of $Q$. The left-eigenvector is the all-one vector $u$, resulting in $u^T Q = 0$. Exact analyses for the complete graph are presented in [1] and [16] and for the star in [3]; extensions to fractional calculus in [17].

### 2.1 Epidemic phase transition

Since the start of the research on epidemics in networks [11], the determination of the epidemic phase transition has been key [10]. The epidemic threshold is the value of the effective infection rate $\tau_c$ in

[1] If $\varepsilon = 0$, then the Markov graph possesses an absorbing state (i.e. the overall healhty state in which there is no virus anymore). That absorbing state is also the steady-state vector.

a sufficiently large graph (actually[2] for $N \to \infty$) at which the phase transition occurs. The epidemic threshold separates the epidemic dynamics into two regimes: if the effective infection rate $\tau > \tau_c$, then the epidemic causes that a non-zero fraction of nodes becomes infected, else the epidemic dies out and only a negligible part of the network becomes infected. Already in 1927, Kermack and McKendrick [6] demonstrated the existence of an epidemic phase transition in a homogeneous population, which led to the introduction of the basic reproduction number $R_0$. The precise relation between the Markovian SIR phase transition [2] and the basic reproduction number is difficult, although in a mean-field approximation, it holds that $R_0 = \frac{\tau}{\tau_c}$ so that the phase transition occurs at $R_0 = 1$.

Analogous to crystallization[3] of matter from the liquid to the solid phase, understanding the formation of "epidemic cohesion" in a graph when sweeping the effective infection rate $\tau$ from below to above the epidemic threshold $\tau_c$ (or vice versa) is a major motivation to compute the $2^N \times 1$ probability state vector $s(t)$ for large $N$. In particular, the theory of phase transitions suggests that the joint probability of infection between any subset of nodes in the graph behaves or scales similarly, which implies that the component $i$ in the probability state vector $s(t)$ in an $\varepsilon$-SIS epidemics, which quantifies the probability that a certain subset $i$ of nodes is infected in the graph, cannot be approximated by $s_i(t) = \prod_{k=1;x_k(i)=1}^{N} \Pr[X_k(t) = 1]$. Hence, mean-field theory breaks down around the phase transition. Physically, it is interesting to understand in what topological fashion the epidemic freezes: How do the pairs, triples, quadruples or any particular subset of nodes join the "epidemic ice"? Which subset freezes first? Does the epidemic ice grows similarly as a giant component [5] at the structural phase transition in Erdős-Rényi graphs, etc.?

Although exact for SIS and SIR [2], the solution (3) is hardly computable for a large size $N$ of the contact graph due to the exponential explosion $c^N$ of the state space for an epidemic with $c$ compartments. Nevertheless, all details about the phase transition of the $\varepsilon$-SIS process around the epidemic threshold[4] are embedded in the huge $2^N \times 2^N$ matrix $Q$, still waiting to be unraveled. We believe that a (sufficiently dense) graph of $N = 50$ nodes is already sufficient to exhibit the physics of the epidemic phase transition. As far as I know, only the famous Onsager[5] computation [9] of a phase transition of the Ising model in two dimensions is exactly computable by using elliptic, Jacobian theta-functions: a most remarkable Gaussian [18] tour-de-force!

## 2.2 Extremal probabilities

What is the probability that, during an epidemic, more than $x\%$ of the population is infected at a certain time? Or, what is the probability that a specific group of nodes in a given contact network is infected at the same time? Generally, these questions ask to compute the joint probability $s_i(t)$ of a

[2] The larger the size $N$ of the graph, the faster quantities change around the phase transition. When $N \to \infty$, a zero-one transition occurs at a single point, which defines the phase transition sharply. If $N$ is finite, there is always an interval in which the transition happens.

[3] The most well-known phase transition occurs when water freezes at zero Celsius (or 273 Kelvin): at a temperature $T < 273K$, we can stand on water/ice, while at $T > 273K$, only swimming is possible.

[4] The epidemic threshold lies approximately in a region of the effective infection rate $\tau \in (\frac{1}{\lambda_1}, \frac{1}{\lambda_1} + c)$, where $\lambda_1$ is the largest eigenvalue of the adjacency matrix $A$ of the graph $G$ and $c$ is a yet unknown, positive real number.

[5] The Nobel Prize in Chemistry 1968 was awarded to Lars Onsager "for the discovery of the reciprocal relations bearing his name, which are fundamental for the thermodynamics of irreversible processes"; not for his computation of the exact phase transition!

nodal set $i = \sum_{k=1}^{N} x_k(i)\, 2^{k-1}$ (e.g. set $i = 01010010$ of $N = 8$ labelled nodes) being infected at time $t$. Related questions ask for the time $t$ when the joint probability $s_i(t) > 10^{-a}$ or $s_i(t) < 10^{-a}$, where $a \geq 0$ is an accuracy level. All such problem can be solved from the Markovian SIS model, but are extremely hard to simulate, because Monte Carlo-like simulations will last long to generate sufficient events where a set of three and more nodes are jointly infected.

## 3 Open problem

The $2^N \times 2^N$ infinitesimal generator matrix $Q$ is minus a weighted Laplacian matrix. All Laplacian matrices on graphs are positive semi-definite, with a zero eigenvalue of multiplicity 1 if the graph is connected. Moreover, as shown in [20, Fig. 2 & 3] and [19, Fig. 2], the matrix $Q$ contains structure and is sparse. A matrix recursion formula for the $2^{N+1} \times 2^{N+1}$ infinitesimal generator matrix $Q_{N+1}$ in terms of $Q_N$ exists. Via nodal relabelling, which interchanges rows and columns in $Q$, another structure may be found (see e.g. [4] and references in [2]).

Commercial software such as Matlab and Mathematica is able to provide the solution (3) of the probability state vector roughly up to $N = 12$ (i.e. solving linear equations in matrices up to ca. $4000 \times 4000$). By using Expokit, a software algorithm by Roger Sidje [13] that computes the exponential of a matrix, it seems that the size $N = 12$ can be shifted to $N = 24$.

Notwithstanding the serious increase in graph size $N$ due to Expokit, we would like to explore possibilities to further increase the number $N$ of nodes of the graph for which the $2^N \times 1$ probability state vector $s(t)$ in (3) can be computed:

1. Can we find a numerical computation method of the $2^N \times 1$ probability state vector $s(t)$ as a function of time $t$ for graphs larger than $N = 24$ with about 3 digits accurate (i.e. the numerical computation $s_i^*(t)$ for all $1 \leq i \leq 2^N$ satisfies $|s_i(t) - s_i^*(t)| < 10^{-a}$ with $a = 3$)? What is the maximal size $N$ of the contact graph for which the $2^N \times 1$ probability state vector $s(t)$ can be computed?

2. Anticipating the rise of quantum computers, we expect that quantum algorithms may help? There seems to be a correspondence between the spin-up and spin-down of qubits and the epidemic state $X_k = 1$ or $X_k = 0$.

**Acknowledgement** I am grateful to Scott Dalhgren for informing me about the Expokit software. This research has been funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 101019718).

## References

[1] M. A. Achterberg, B. Prasse, and P. Van Mieghem. An exact analysis of continuous-time Markovian epsilon-SIS epidemics on networks. *Physical Review E*, 105(5):054305, 2022.

[2] M. A. Achterberg and P. Van Mieghem. Analytic solution of Markovian epidemics without re-infections on heterogeneous networks. *ArXiv2311.1672*, 2023.

[3] E. Cator and P. Van Mieghem. Susceptible-Infected-Susceptible epidemics on the complete graph and the star graph: Exact analysis. *Physical Review E*, 87(1):012811, January 2013.

[4] A. Economou, A. Gómez-Corral, and M. López García. A stochastic SIS epidemic model with heterogeneous contacts. *Physica A*, 421:78–97, March 2015.

[5] S. Janson, D. E. Knuth, T. Luczak, and B. Pittel. The birth of the giant component. *Random Structures and Algorithms*, 4(3):233–358, 1993.

[6] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society London, A*, 115:700–721, August 1927.

[7] A. P. Millán, E. C. W. van Straaten, C. J. Stam, I. A. Nissen, S. Idema, J. C. Baayen, P. Van Mieghem, and A. Hillebrand. Epidemic models characterize seizure propagation and the effects of epilepsy surgery in individualized brain networks based on meg and invasive eeg recordings. *Scientific Reports*, 12:art. no. 4086, 2022.

[8] A. P. Millán, E. C. W. van Straaten, C. J. Stam, I. A. Nissen, S. Idema, P. Van Mieghem, and A. Hillebrand. Individualized epidemic spreading models predict epilepsy surgery outcomes: a pseudo-prospective study. *Network Neuroscience*, to appear 2024.

[9] L. Onsager. Crystal statistics: A two-dimensional model with an order-disorder transition. *Physical Review*, 65(3 and 4):117–149, February 1944.

[10] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. Epidemic processes in complex networks. *Review of Modern Physics*, 87(3):925–979, September 2015.

[11] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14):3200–3203, April 2001.

[12] F. D. Sahneh, C. Scoglio, and P. Van Mieghem. Generalized epidemic mean-field model for spreading processes over multi-layer complex networks. *IEEE/ACM Transaction on Networking*, 21(5):1609–1620, October 2013.

[13] R. B. Sidje. Expokit: A software package for computing matrix exponentials. *ACM Transaction on Mathematical Software*, 24(1):130–156, 1998.

[14] P. Van Mieghem. *Graph Spectra for Complex Networks*. Cambridge University Press, Cambridge, U.K., 2011.

[15] P. Van Mieghem. *Performance Analysis of Complex Networks and Systems*. Cambridge University Press, Cambridge, U.K., 2014.

[16] P. Van Mieghem. Explosive phase transition in SIS epidemics with arbitrary small but non-zero self-infection rate. *Physical Review E*, 101(3):032303, March 2020.

[17] P. Van Mieghem. The origin of the fractional derivative and fractional non-Markovian continuous time processes. *Physical Review Research*, 4(2):023242, 2022.

[18] P. Van Mieghem. The Arithmetic-Geometric Mean: A pearl of Gauss. Delft University of Technology, report20230605 (www.nas.ewi.tudelft.nl/people/Piet/TUDelftReports), 2023.

[19] P. Van Mieghem and E. Cator. Epidemics in networks with nodal self-infections and the epidemic threshold. *Physical Review E*, 86(1):016116, July 2012.

[20] P. Van Mieghem, J. Omic, and R. E. Kooij. Virus spread in networks. *IEEE/ACM Transactions on Networking*, 17(1):1–14, February 2009.