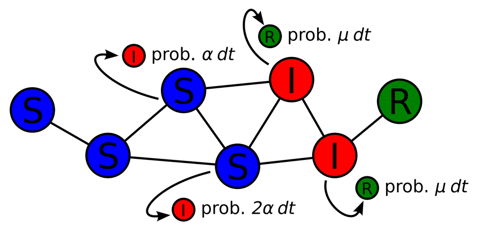
Exact(?) SIR dynamics on networks of finite size

Pierre-André Noël, Antoine Allard and Louis J. Dubé

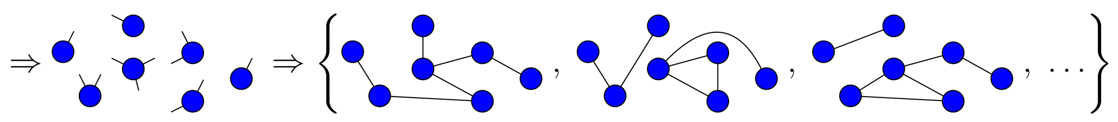
Département de Physique, de Génie Physique, et d'Optique, Université Laval, Québec City, Canada

Goal



- We are interested in the **time evolution** of SIR dynamics on networks. During the time period $[t, t+dt)$, a **susceptible** (S) node has probability $\alpha\,dt$ for each of its **infectious** (I) neighbour to itself become infectious. Moreover, an infectious node has probability $\mu\,dt$ to become **removed** (R). In the appropriate time unit, we may choose $\alpha = 1$.
- A **network ensemble** is specified by its **degree sequence**; the stubs corresponding to these degrees are attached at random. The probability for each outcome on a given structure is weighted by the probability for that structure in the network ensemble.

Degree	0	1	2	3
Number	0	3	3	1



- We want to take into account the **finite size** of the system since some questions cannot be answered by a formalism assuming infinite size.
  - How long does an epidemic take to invade the system?
  - If an epidemic would affect 10% of an infinite population, how many would be infected in a population of 100 nodes?
- We require results to be under the form of **probability distributions**. The principal alternative, *i.e.* a treatment of mean values, does not convey sufficient information about the wide variety of possible outcomes (ranging from outbreaks to epidemics).

The approach...

We note S_k and I_k the number of susceptible and infectious nodes with k **unassigned stubs**, *i.e.* stubs for which we do not yet know which node is at the other end of the link. We define the **state** of a SI system through the numbers S_0, S_1, S_2, ... and I_0, I_1, I_2, ...

At time t, each of these states has a probability to be the one that will be achieved by the stochastic SI dynamics over one of the possible structure of the network ensemble. The complete **distribution** of all those probabilities is what we are looking for.

The time evolution of this distribution is provided by a high-dimensional **differential equation** system. The transition rate matrix (or operator) L completely specify the probability flows between states. Provided an initial condition, the solution is readily obtained.

The "diagonal part" of L takes into account the probabilities leaving each state. This is proportional to the total number of stubs belonging to infectious nodes.

When we decide what is at the other end of a stub belonging to an infectious node, this node has now one less unassigned stub.

Similarly, a susceptible node acquiring the infection becomes an infectious node with one less unassigned stub (the one from which it acquired the infection).

Unassigned stubs from infectious nodes form links with other stubs at random. In order to preserve **normalization**, the rate at which this is done must be divided by the total number of unassigned stubs.

The contributions to L for a **SI system** are thus the following.

- One of the k unassigned stubs of an infectious node may get assigned ($-k\,b_k^\dagger b_k$).
- This stub may target a susceptible node ($k'\,b_{k'-1}^\dagger a_{k'} \Omega\, k\, b_{k-1}^\dagger b_k$)
- or an infectious node with k' unassigned stubs ($k'\,b_{k'-1}^\dagger b_{k'} \Omega\, k\, b_{k-1}^\dagger b_k$).

In a **SIR system**, additional contributions must be taken into account.

- An infectious may target a removed ($k'\,c_{k'-1}^\dagger c_{k'} \Omega\, k\, b_{k-1}^\dagger b_k$).
- An infectious may become removed ($\mu c_k^\dagger b_k - \mu b_k^\dagger b_k$).

...and the maths

$$\left|\begin{matrix} S_0 & S_1 & S_2 & S_3 & \cdots \\ I_0 & I_1 & I_2 & I_3 & \cdots \end{matrix}\right\rangle$$

$$|\psi(t)\rangle = \sum P\left(\begin{matrix} S_0 & S_1 & \cdots \\ I_0 & I_1 & \cdots \end{matrix}\,\middle|\, t\right) \left|\begin{matrix} S_0 & S_1 & \cdots \\ I_0 & I_1 & \cdots \end{matrix}\right\rangle$$

$$\frac{d}{dt}|\psi(t)\rangle = L\,|\psi(t)\rangle\,, \quad |\psi(t)\rangle = \mathrm{e}^{Lt}\,|\psi(0)\rangle$$

$$\sum_k k I_k \left|\begin{matrix} S_0 & S_1 & \cdots \\ I_0 & I_1 & \cdots \end{matrix}\right\rangle = \sum_k k b_k^\dagger b_k \left|\begin{matrix} S_0 & S_1 & \cdots \\ I_0 & I_1 & \cdots \end{matrix}\right\rangle$$

$$\sum_k k I_k \left|\begin{matrix} \cdots & S_{k-1} & S_k & \cdots \\ \cdots & I_{k-1}+1 & I_k - 1 & \cdots \end{matrix}\right\rangle = \sum_k k b_{k-1}^\dagger b_k \left|\begin{matrix} S_0 & S_1 & \cdots \\ I_0 & I_1 & \cdots \end{matrix}\right\rangle$$

$$\sum_k k S_k \left|\begin{matrix} \cdots & S_{k-1} & S_k - 1 & \cdots \\ \cdots & I_{k-1}+1 & I_k & \cdots \end{matrix}\right\rangle = \sum_k k b_{k-1}^\dagger a_k \left|\begin{matrix} S_0 & S_1 & \cdots \\ I_0 & I_1 & \cdots \end{matrix}\right\rangle$$

$$\frac{1}{\sum_{k'} k' \left(S_{k'} + I_{k'}\right)} \left|\begin{matrix} S_0 & S_1 & \cdots \\ I_0 & I_1 & \cdots \end{matrix}\right\rangle = \Omega \left|\begin{matrix} S_0 & S_1 & \cdots \\ I_0 & I_1 & \cdots \end{matrix}\right\rangle$$

$$\text{with } \Omega^{-1} = \sum_k k\left(a_k^\dagger a_k + b_k^\dagger b_k + \underbrace{c_k^\dagger c_k}_{SIR\text{ only}}\right)$$

$$L = \sum_{k,k'} k k' \left(b_{k'-1}^\dagger a_{k'} + b_{k'-1}^\dagger b_{k'}\right) \Omega\; b_{k-1}^\dagger b_k - \sum_k k b_k^\dagger b_k$$

$$L = \sum_{k,k'} k k' \left(b_{k'-1}^\dagger a_{k'} + b_{k'-1}^\dagger b_{k'} + c_{k'-1}^\dagger c_{k'}\right) \Omega\; b_{k-1}^\dagger b_k + \sum_k \left(\mu c_k^\dagger b_k - (k+\mu)\, b_k^\dagger b_k\right)$$

Annihilation and creation operators

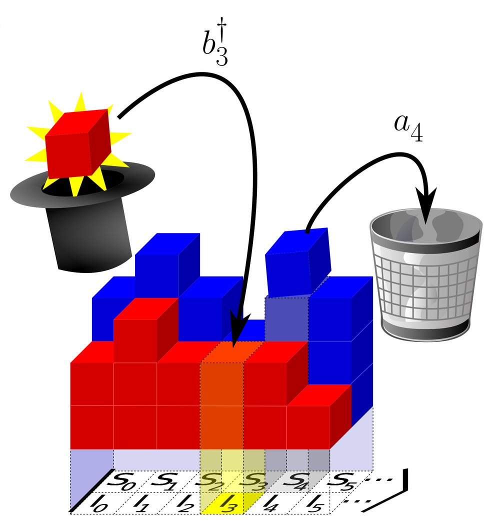
We define the **annihilation** (a_k and b_k $\forall\, k$) and **creation** ($a_k^\dagger$ and $b_k^\dagger$ $\forall\, k$) operators through their effects on a state

$$a_k \left|\begin{matrix} S_0 & S_1 & \cdots & S_k & \cdots \\ I_0 & I_1 & \cdots & I_k & \cdots \end{matrix}\right\rangle = S_k \left|\begin{matrix} S_0 & S_1 & \cdots & S_k - 1 & \cdots \\ I_0 & I_1 & \cdots & I_k & \cdots \end{matrix}\right\rangle$$

$$a_k^\dagger \left|\begin{matrix} S_0 & S_1 & \cdots & S_k & \cdots \\ I_0 & I_1 & \cdots & I_k & \cdots \end{matrix}\right\rangle = \left|\begin{matrix} S_0 & S_1 & \cdots & S_k + 1 & \cdots \\ I_0 & I_1 & \cdots & I_k & \cdots \end{matrix}\right\rangle$$

$$b_k \left|\begin{matrix} S_0 & S_1 & \cdots & S_k & \cdots \\ I_0 & I_1 & \cdots & I_k & \cdots \end{matrix}\right\rangle = I_k \left|\begin{matrix} S_0 & S_1 & \cdots & S_k & \cdots \\ I_0 & I_1 & \cdots & I_k - 1 & \cdots \end{matrix}\right\rangle$$

$$b_k^\dagger \left|\begin{matrix} S_0 & S_1 & \cdots & S_k & \cdots \\ I_0 & I_1 & \cdots & I_k & \cdots \end{matrix}\right\rangle = \left|\begin{matrix} S_0 & S_1 & \cdots & S_k & \cdots \\ I_0 & I_1 & \cdots & I_k + 1 & \cdots \end{matrix}\right\rangle \quad .$$



Their **commutators** (defined as $[X,Y] = XY - YX$) are

$$\left[a_k, a_{k'}^\dagger\right] = \delta_{kk'} \qquad \left[b_k, b_{k'}^\dagger\right] = \delta_{kk'}$$

$$\left[a_k, a_{k'}\right] = \left[a_k, b_{k'}\right] = \left[a_k, b_{k'}^\dagger\right] = \left[b_k, b_{k'}\right] = \left[a_k^\dagger, a_{k'}^\dagger\right] = \left[a_k^\dagger, b_{k'}^\dagger\right] = \left[a_k^\dagger, b_{k'}\right] = \left[b_k^\dagger, b_{k'}^\dagger\right] = 0 \quad .$$

For SIR systems, c_k and $c_k^\dagger$ are defined similarly except that they act on removed nodes.

Familiar with PGFs?

If you are more familiar with Probability Generating Functions (PGFs) than with creation and annihilation operators, you may find the following equivalences useful.

$$a_k^\dagger \leftrightarrow x_k \qquad b_k^\dagger \leftrightarrow y_k$$

$$a_k \leftrightarrow \frac{\partial}{\partial x_k} \qquad b_k \leftrightarrow \frac{\partial}{\partial y_k}$$

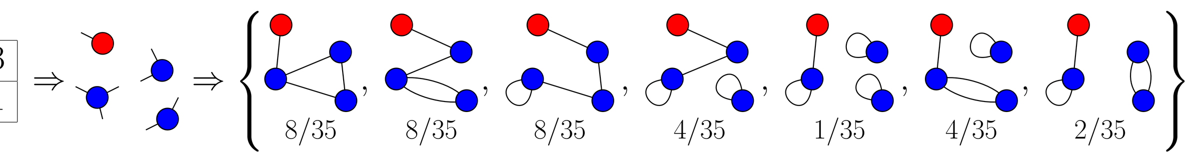
$$\left|\begin{matrix} S_0 & S_1 & S_2 & \ldots \\ I_0 & I_1 & I_2 & \ldots \end{matrix}\right\rangle \leftrightarrow \prod_k x_k^{S_k} y_k^{I_k}$$

$$|\psi(t)\rangle \leftrightarrow \psi(\mathbf{x}, \mathbf{y}; t)$$

A simple example

We are interested in an SI dynamics on a network ensemble specified by a degree sequence

Degree	0	1	2	3
Number	0	1	2	1



If patient zero (the first infectious) is of degree 1, the system is initially in the state $|\psi(0)\rangle = \left|\begin{smallmatrix} 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 0 \end{smallmatrix}\right\rangle$.

The differential equation provided by L gives the flow of probabilities between states. A convenient analogy may be done with the flow of **water** (representing probabilities) between **buckets** (representing states). Each bucket may have some **holes** of varying size (representing the rates specified by L) from which water may leak to a bucket placed under it.

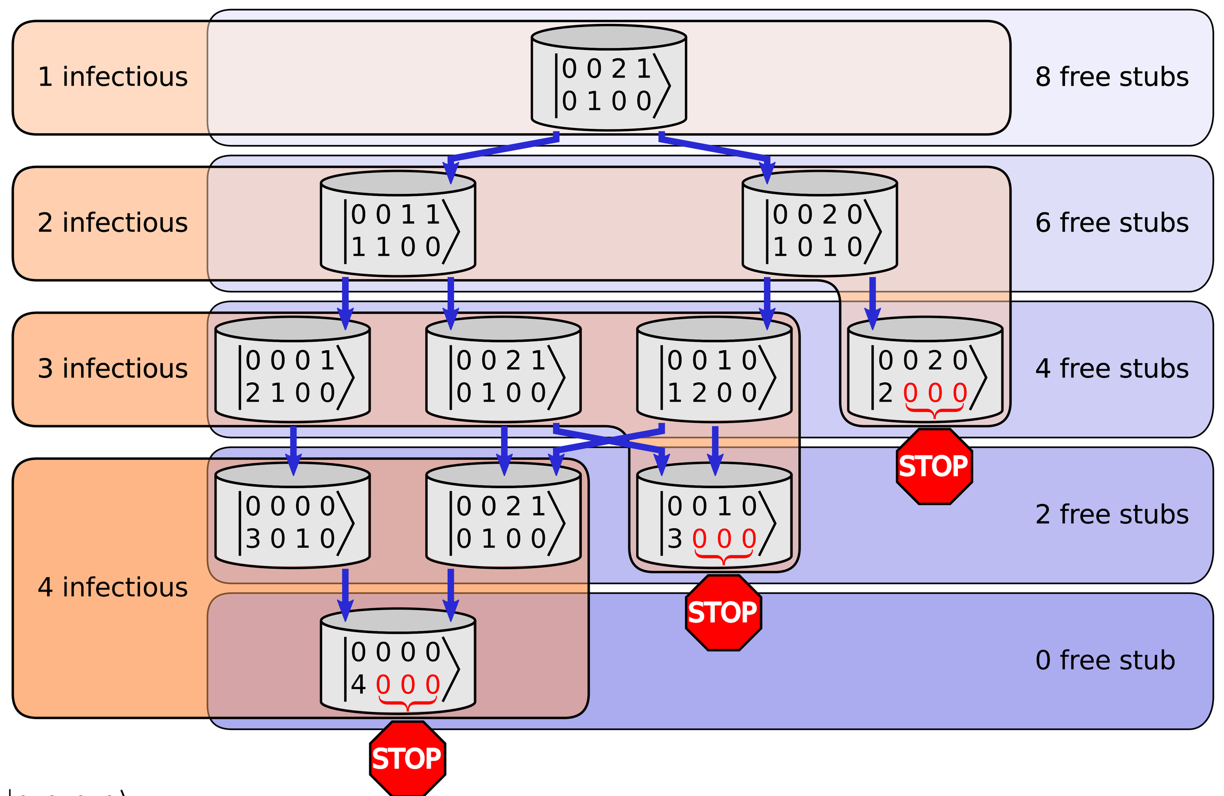
The first bucket has two holes leading to the $\left|\begin{smallmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{smallmatrix}\right\rangle$ and $\left|\begin{smallmatrix} 0 & 0 & 2 & 0 \\ 1 & 0 & 1 & 0 \end{smallmatrix}\right\rangle$ buckets at rates 4/7 and 3/7 respectively

$$L\left|\begin{matrix} 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 0 \end{matrix}\right\rangle = \frac{4}{7}\left|\begin{matrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{matrix}\right\rangle + \frac{3}{7}\left|\begin{matrix} 0 & 0 & 2 & 0 \\ 1 & 0 & 1 & 0 \end{matrix}\right\rangle \,.$$

Since these are constants, we know that 4/7 of the water will eventually pass through the first bucket and 3/7 through the second.

Some buckets have no holes. This occurs for states with no infectious nodes bearing free stubs since applying L onto them gives 0. The long term distribution $|\psi(\infty)\rangle$ may only be composed of such states

$$|\psi(\infty)\rangle = \frac{3}{35}\left|\begin{matrix} 0 & 0 & 2 & 0 \\ 2 & 0 & 0 & 0 \end{matrix}\right\rangle + \frac{8}{35}\left|\begin{matrix} 0 & 0 & 1 & 0 \\ 3 & 0 & 0 & 0 \end{matrix}\right\rangle + \frac{24}{35}\left|\begin{matrix} 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 \end{matrix}\right\rangle \,.$$



Obtaining this long term distribution does not require to solve the differential equation system. Indeed, one may instead determine for each bucket the fraction of water that will eventually pass through each of its holes. The process is repeated until only buckets without holes contain water.

When the time evolution is required, the distribution of probability for each state $|\psi(t)\rangle = \mathrm{e}^{Lt}|\psi(0)\rangle$ can be computed. If the quantity of interest is the number of infectious nodes, the distribution of probability for states can be converted to a distribution for the number of infectious nodes. This is simply done by summing the probability for each states bearing the same number of infectious.

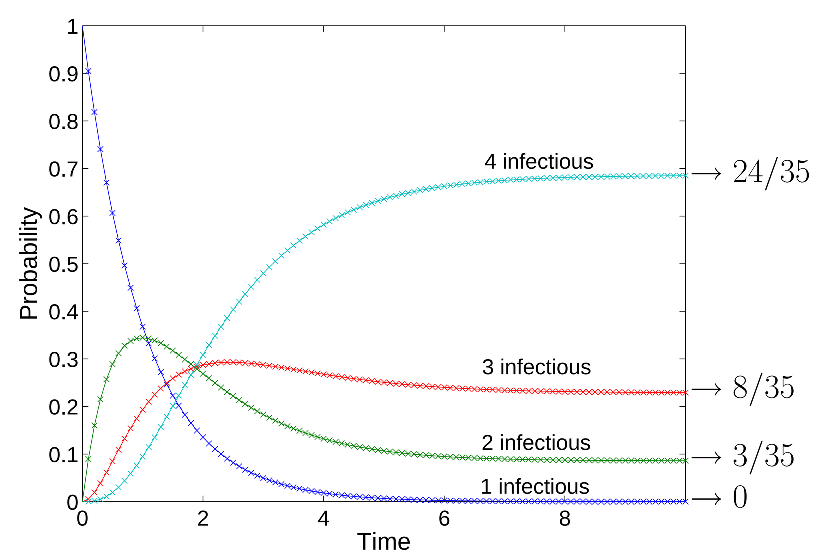


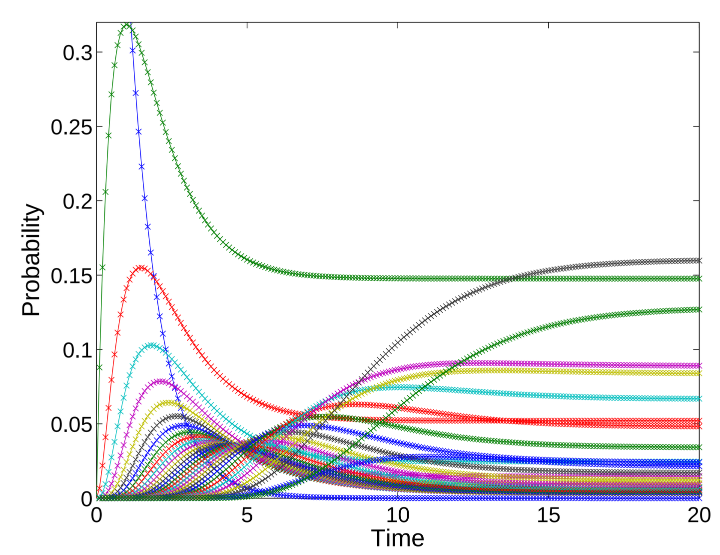
Figure: *Time evolution of the probabilities for the number of infectious nodes. The lines show the solution of the differential equations while the symbols are obtained through Monte-Carlo numerical simulations.*

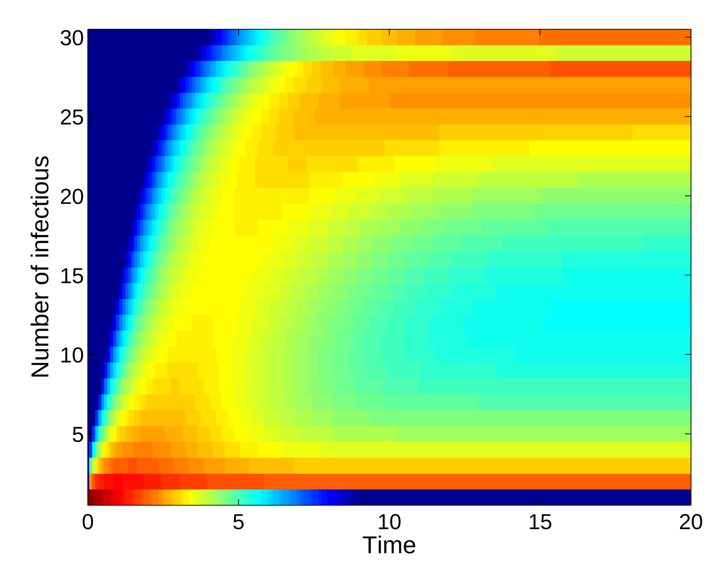
A more complex case

Degree	0	1	2	3	4
Number	0	10	10	6	4

Patient zero has degree 1.

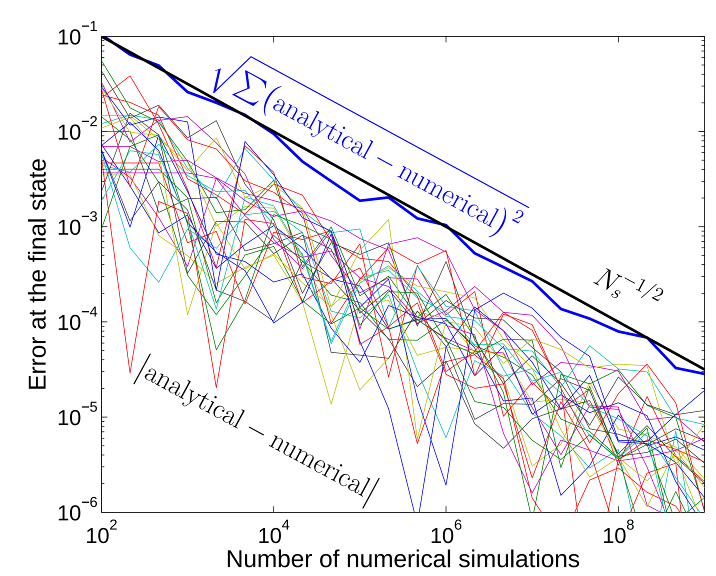
$\Rightarrow$ Initial state: $\left|\begin{matrix} 0 & 9 & 10 & 6 & 4 \\ 0 & 1 & 0 & 0 & 0 \end{matrix}\right\rangle$





Exact?

The difference between the analytical results and the Monte-Carlo numerical simulations decreases with the number N_s of simulations. Moreover, the rate of this decrease ($N_s^{-1/2}$) is the one that would be expected if the analytical results were exact.



The network ensembles used for both the numerical simulations and the analytical approach are such that **self loops** and **repeated links** are **allowed**.

When such structures are forbidden, the formalism presented here is **not exact**. However, these structures become less likely with an increase in the network size. Hence, the formalism will produces very good results for networks that are not too small.

NSERC CRSNG
CIHR IRSC
Fonds de recherche sur la nature et les technologies Québec