Stochastic Modelling and Simulation
Mean Field and Fluid Approximation

Luca Bortolussi[1]

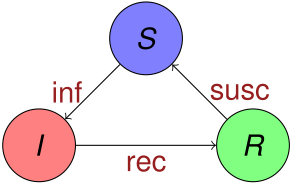
[1]Dipartimento di Matematica e Geoscienze
Università degli studi di Trieste

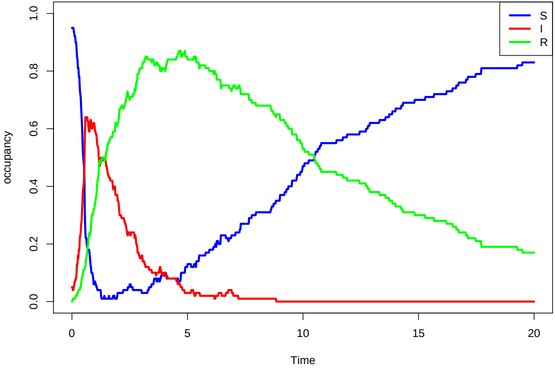
Office 328, third floor, H2bis
lbortolussi@units.it

Trieste, Data Science and Scientific Computing

EXAMPLE: SIR EPIDEMICS

Population model
$N = 100$ agents

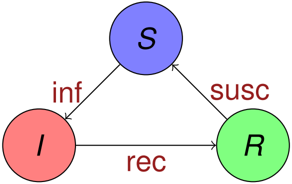


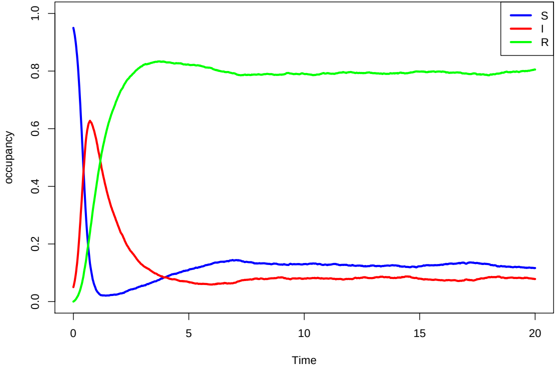


(1 run)

EXAMPLE: SIR EPIDEMICS

Population model
$N = 10000$ agents





(1 run)

OVERVIEW

We will consider Markov models of population processes: systems composed of populations of interacting agents, whose behaviour is a collective emergent property.

MEAN FIELD/ FLUID APPROXIMATION

Approximation by a deterministic system (differential/ difference equations).

MEAN FIELD (ORIGINALLY)/ FAST SIMULATION

Approximation by another, simpler, stochastic model.

OVERVIEW: FLUID APPROXIMATION

LIMIT THEOREM POINT OF VIEW

Considers the deterministic model as the limit of the stochastic process for large populations/ system size:

- CTMC to ODE
- DTMC to Difference Equations
- DTMC to ODE
- CTMC to Gaussian processes (central limit)
- CTMC to hybrid system
- CTMC to SDE (diffusion limit)

MOMENT CLOSURE POINT OF VIEW

Considers the deterministic model as an approximation of the mean of the stochastic process.
Equations for higher order moments can be given as well.

OVERVIEW: MEAN FIELD

Approximation by another, simpler, stochastic model.

FAST SIMULATION

Approximate the behaviour of one or few agents by another stochastic process depending on the mean of the rest of the system.

HARTREE APPROXIMATION (MEAN FIELD)

Approximates the process (at transient/ steady state) by assuming a product form (w.r.t. variables). The decoupling is obtained by averaging the rates of transitions acting on a variable X with respect to the other variables.

MENU À LA CARTE

- Fluid approximation (CTMC + ODE)
- Steady state limits
- Fluid equation and moments
- Central Limit and linear noise approximation

Outline

Outline

POPULATION CTMC

If we want to describe population processes, with many agents, representing the CTMC by its Q-matrix is unfeasible, as the state space blows up.

A population CTMC model is a tuple $\mathcal{X} = (\mathbf{X}, \mathcal{D}, \mathcal{T}, \mathbf{x_0})$, where:

1. $\mathbf{X}$ — vector of *variables* counting how many individuals in each state.
2. $\mathcal{D} = \prod_i \mathcal{D}_i$ — (countable) state space.
3. $\mathbf{x_0} \in \mathcal{D}$ —*initial state*.
4. $\eta_i \in \mathcal{T}$ — *global transitions*, $\eta_i = (\mathbf{v}, r(\mathbf{X}))$
    1. $\mathbf{v} \in \mathbb{R}^n$ — *update vector* (from $\mathbf{X}$ to $\mathbf{X} + \mathbf{v}$)
    2. $r : \mathcal{D} \to \mathbb{R}_{\geq 0}$ — rate function.

MASTER EQUATION

The Kolmogorov equation in the context of Population Processes is often know as master equation.

There is one equation per state $\mathbf{x} \in \mathcal{D}$, for the probability mass $P(\mathbf{x}, t)$, which considers the inflow and outflow of probability at time t.

$$\frac{dP(\mathbf{x}, t)}{dt} = \sum_{\eta \in \mathcal{T}} r_\eta(\mathbf{x} - \mathbf{v}_\eta) P(\mathbf{x} - \mathbf{v}_\eta, t) - \sum_{\eta \in \mathcal{T}} r_\eta(\mathbf{x}) P(\mathbf{x}, t)$$

POISSON REPRESENTATION

Population CTMC admit a simple description in terms of Poisson processes (random time change).
Essentially, we introduce variables $R_\eta(t)$ counting how many times each transition η has fired up to time t. Hence we can write:

$$X(t) = X(0) + \sum_{\eta \in \mathcal{T}} \mathbf{v}_\eta R_\eta(t).$$

It turns out that $R_\eta(t)$ is a time-inhomogeneous Poisson process with cumulative rate $\int_0^t r_\eta(X(s))ds$, independent from the other $R_{\eta'}$.
Hence, let $\mathcal{N}_\eta$ be independent Poisson processes. For each $t \geq 0$:

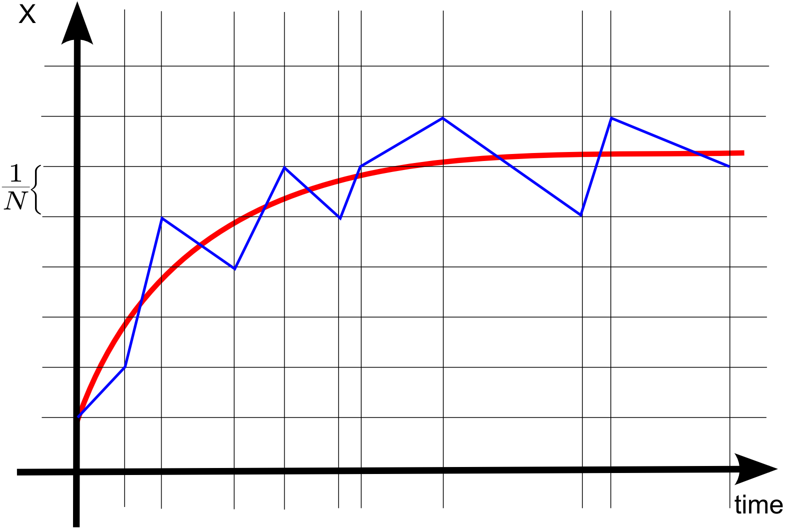
$$X(t) = X(0) + \sum_{\eta \in \mathcal{T}} \mathbf{v}_\eta \mathcal{N}_\eta \left(\int_0^t r_\eta(X(s))ds \right).$$

Fluid Approximation

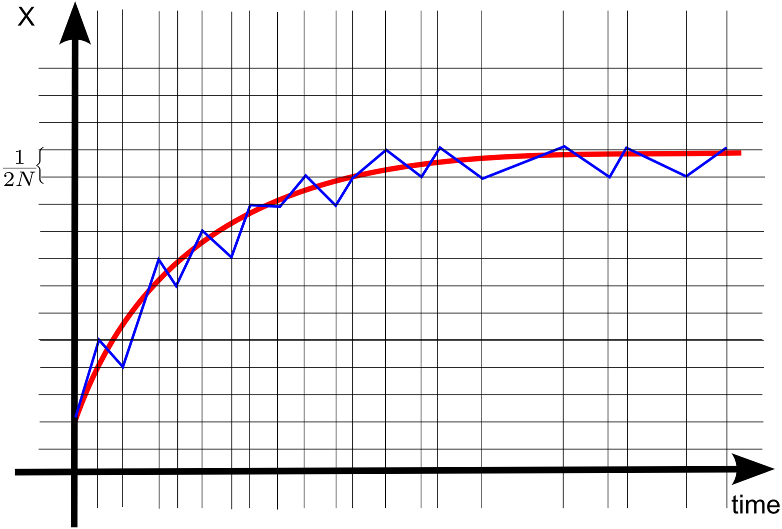
Basics

- It applies to CTMC models of population dynamics with large population size N (studies the limit as $N \to \infty$)
- It works on scaled variables, to treat uniformly different population levels.
- Requires proper scaling and regularity assumptions on rates.
- The method works by constructing an ODE from the sequence of population dependent CTMC.
- It can be proved that, in any finite time horizon, the trajectories of the CTMC become indistinguishable from the solution of the ODE.

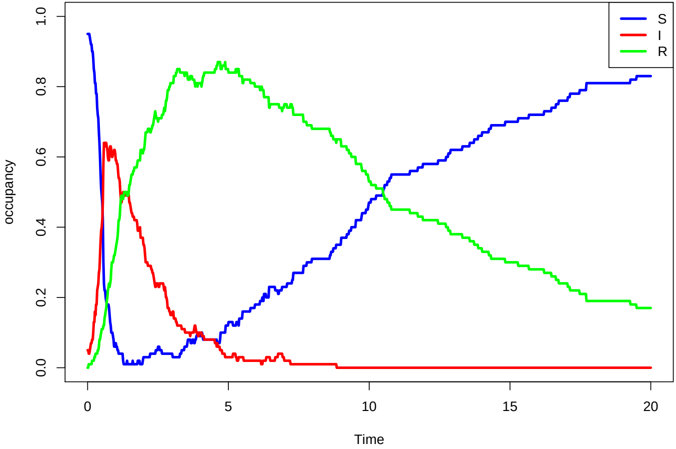
AN INTUITION



AN INTUITION

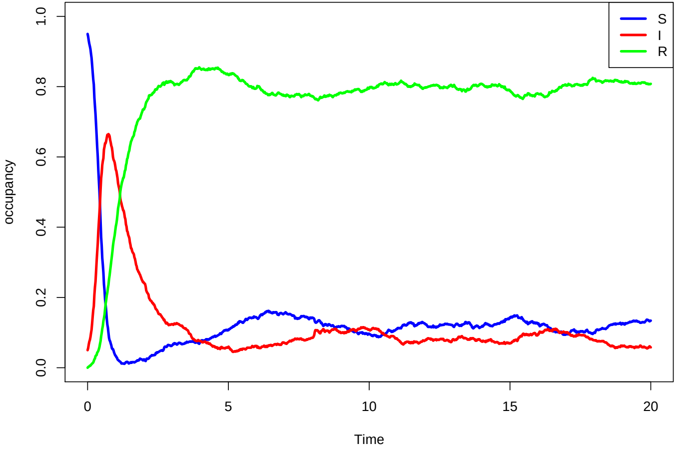


EXAMPLE CONTINUED



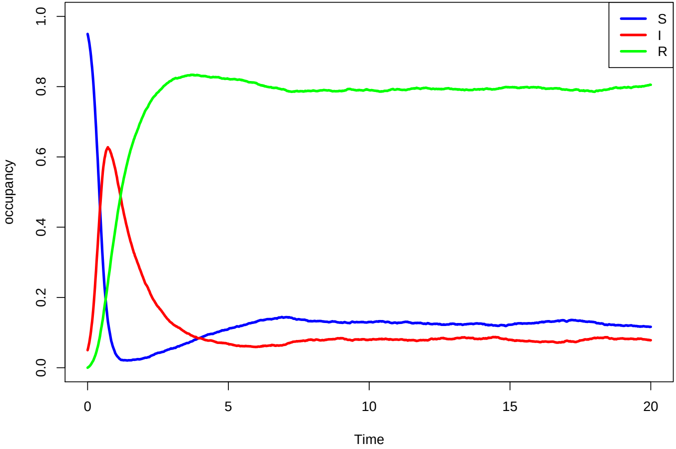
CTMC $N = 100$

EXAMPLE CONTINUED



CTMC $N = 1000$

EXAMPLE CONTINUED



CTMC $N = 10000$

SCALING CONDITIONS

BASICS

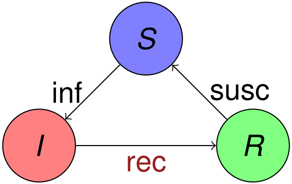
- We have a sequence $\mathcal{X}^{(N)}$ of models, for increasing system size (e.g. total population N).
- We normalize such models in order to bring them to the same scale (divide variables by size N).
- $\mathbf{X}^{(N)}(t)$ is the Markov process (in continuous time) defined by $\mathcal{X}^{(N)}$.

NORMALIZATION

The normalized model $\hat{\mathcal{X}}^{(N)} = (\hat{\mathbf{X}}, \hat{\mathcal{D}}^{(N)}, \hat{\mathcal{T}}^{(N)}, \hat{\mathbf{X}}_0^{(N)})$ associated with $\mathcal{X}^{(N)} = (\mathbf{X}, \mathcal{D}^{(N)}, \mathcal{T}^{(N)}, \mathbf{X}_0^{(N)})$ is defined by:

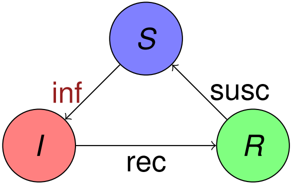
- Variables: $\hat{\mathbf{X}} = \frac{\mathbf{X}}{N}$
- Domain: $\hat{\mathcal{D}}^{(N)} = \{N^{-1}\mathbf{x} \mid \mathbf{x} \in \mathcal{D}\}$.
- Initial conditions: $\hat{\mathbf{X}}_n^{(N)} = \frac{\mathbf{X}_0^{(N)}}{N}$
- Normalized transition $\hat{\tau} = (\frac{\mathbf{v}_\tau}{N}, \hat{r}_\tau^{(N)}(\hat{\mathbf{X}}))$ associated with $\tau \in \mathcal{T}^{(N)}$:
  - Update: $\frac{\mathbf{v}_\tau}{N}$;
  - Rates: $r_\tau^{(N)}(\mathbf{X}) = N f_\tau^{(N)}\left(\frac{\mathbf{X}}{N}\right) = \hat{r}_\tau^{(N)}\left(\frac{X}{N}\right)$

EXAMPLE: SIR EPIDEMICS



- $r_{rec}^{(N)}(\mathbf{X}) = k_R X_I = N k_R \frac{X_I}{N} = N k_R \hat{X}_I$
  $\hat{r}_{rec}^{(N)}(\hat{\mathbf{X}}) = N k_R \hat{X}_I$, $f_{rec}(\hat{\mathbf{X}}) = k_R \hat{X}_I$
- $r_{inf}^{(N)}(\mathbf{X}) = \frac{k_I}{N} X_S X_I = N k_I \frac{X_S}{N} \frac{X_I}{N} = N k_I \hat{X}_S \hat{X}_I$
  $\hat{r}_{inf}^{(N)}(\hat{\mathbf{X}}) = N k_I \hat{X}_S \hat{X}_I$, $f_{inf}(\hat{\mathbf{X}}) = k_I \hat{X}_S \hat{X}_I$

EXAMPLE: SIR EPIDEMICS



- $r_{rec}^{(N)}(\mathbf{X}) = k_R X_I = N k_R \frac{X_I}{N} = N k_R \hat{X}_I$
  $\hat{r}_{rec}^{(N)}(\hat{\mathbf{X}}) = N k_R \hat{X}_I$, $f_{rec}(\hat{\mathbf{X}}) = k_R \hat{X}_I$

- $r_{inf}^{(N)}(\mathbf{X}) = \frac{k_I}{N} X_S X_I = N k_I \frac{X_S}{N} \frac{X_I}{N} = N k_I \hat{X}_S \hat{X}_I$
  $\hat{r}_{inf}^{(N)}(\hat{\mathbf{X}}) = N k_I \hat{X}_S \hat{X}_I$, $f_{inf}(\hat{\mathbf{X}}) = k_I \hat{X}_S \hat{X}_I$

SCALING ASSUMPTIONS: STATE SPACE

- Consider the normalised state space $\hat{\mathcal{D}}^{(N)}$ of $\hat{\mathbf{X}}^{(N)}(t)$.
- We need to find a set $E \subset \mathbb{R}^n$ (open or compact) which contains $\hat{\mathcal{D}}^{(N)}$ for each N. This will be the set in which the fluid limit will live.

EXAMPLE: SIR EPIDEMICS

In this case, the normalised variables take values in a discrete grid between 0 and 1:
$\hat{\mathcal{D}}_i^{(N)} = \{\frac{j}{N} \mid j = 1, \ldots, N\}$.

Hence, we can take E to be the unit cube $[0,1]^3$.

However, the total population is conserved, so we can restrict to the unit simplex $E = \{\mathbf{x} \in [0,1]^3 \mid \sum_i x_i = 1\}$.

SCALING ASSUMPTIONS

$f_\tau^{(N)}$ is required to converge uniformly to a locally Lipschitz continuous and locally bounded function f_τ:

$$\sup_{\mathbf{x}\in E} \|f_\tau^{(N)}(\mathbf{x}) - f_\tau(\mathbf{x})\| \to 0.$$

If $f_\tau^{(N)} = f_\tau$ does not depend on N, the rate satisfies the density dependence condition.

f locally Lipschitz iff $\forall \mathbf{x}, \exists B(\mathbf{x}), L > 0, \forall \mathbf{y} \in B(\mathbf{x})\ \|f(\mathbf{x}) - f(\mathbf{y})\| \leq L\|\mathbf{x} - \mathbf{y}\|$
f locally bounded iff $\forall \mathbf{x}, \exists B(\mathbf{x}), M > 0,\ \|f(\mathbf{x})\| \leq M\|\mathbf{x} - \mathbf{y}\|$

The following theorem works also under less restrictive assumptions (e.g. random increments with bounded variance and average).

DRIFT AND LIMIT VECTOR FIELD

DRIFT

The drift or mean increment at level N is

$$F^{(N)}(\mathbf{x}) = \sum_{\tau \in \mathcal{T}} \mathbf{v}_\tau f_\tau^{(N)}(\mathbf{x})$$

By the scaling assumptions, $F^{(N)}$ converges uniformly to F, the limit vector field:

$$F(\mathbf{x}) = \sum_{\tau \in \mathcal{T}} \mathbf{v}_\tau f_\tau(\mathbf{x}).$$

FLUID ODE

The fluid ODE is

$$\frac{d\mathbf{x}(t)}{dt} = F(\mathbf{x}(t))$$

DETERMINISTIC APPROXIMATION THEOREM

HYPOTHESIS

- $\hat{\mathbf{X}}^{(N)}(t)$: sequence of Markov processes that satisfy the conditions above.
- F Lipschitz continuous in E.
- $\exists \mathbf{x_0} \in S$ such that $\hat{\mathbf{X}}^{(N)}(0) \to \mathbf{x_0}$ in probability (or almost surely)

$\mathbf{x}(t)$: solution of $\dot{\mathbf{x}} = F(\mathbf{x})$, $\mathbf{x}(0) = \mathbf{x_0}$, living in E for all $t \geq 0$.

DETERMINISTIC APPROXIMATION THEOREM

THEOREM (KURTZ)

For any finite time horizon $T < \infty$, it holds that:

$$\sup_{0 \leq t \leq T} \|\hat{\mathbf{X}}^{(N)}(t) - \mathbf{x}(t)\| \to 0 \text{ in probability,}$$

meaning, for each $\delta > 0$, that

$$\lim_{N\to\infty} \mathbb{P}\left\{ \sup_{0 \leq t \leq T} \|\hat{\mathbf{X}}^{(N)}(t) - \mathbf{x}(t)\| > \delta \right\} = 0$$

REMARK

Convergence holds also almost surely:

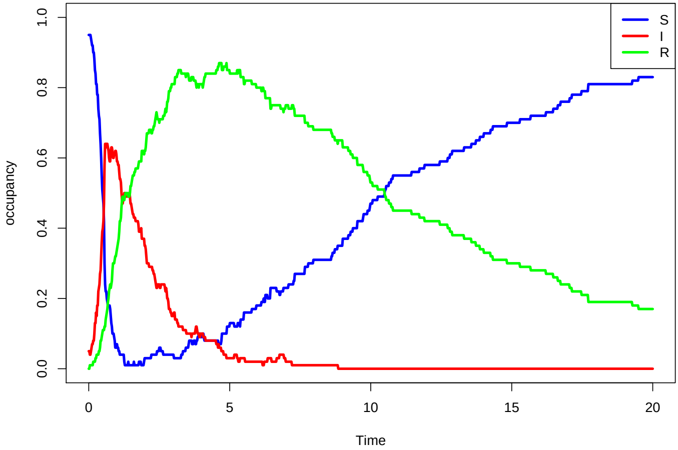
$$\mathbb{P}\left\{ \lim_{N\to\infty} \sup_{0 \leq t \leq T} \|\hat{\mathbf{X}}^{(N)}(t) - \mathbf{x}(t)\| = 0 \right\} = 1$$

EPIDEMICS EXAMPLE CONTINUED

The CTMC $\mathbf{X}^{(N)}(t)$ of the epidemics model satisfies all the hypothesis of fluid limit theorem, so it converges in probability to the solution of the following set of ODEs:

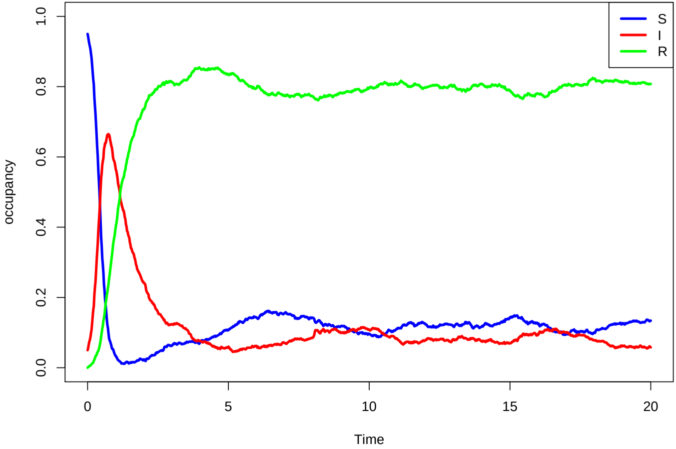
$$\begin{cases} \frac{dx_S}{dt} = k_S x_R - k_I x_I x_S \\ \frac{dx_I}{dt} = k_I x_I x_S - k_R x_I \\ \frac{dx_R}{dt} = k_R x_I - k_S x_R \end{cases}$$

EPIDEMICS EXAMPLE CONTINUED



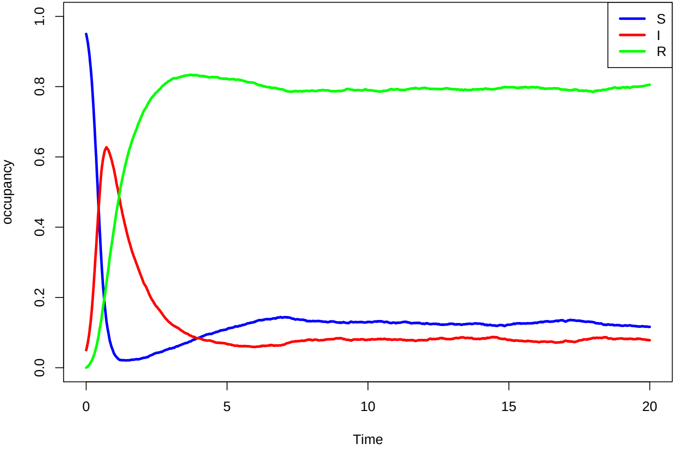
CTMC $N = 100$

EPIDEMICS EXAMPLE CONTINUED



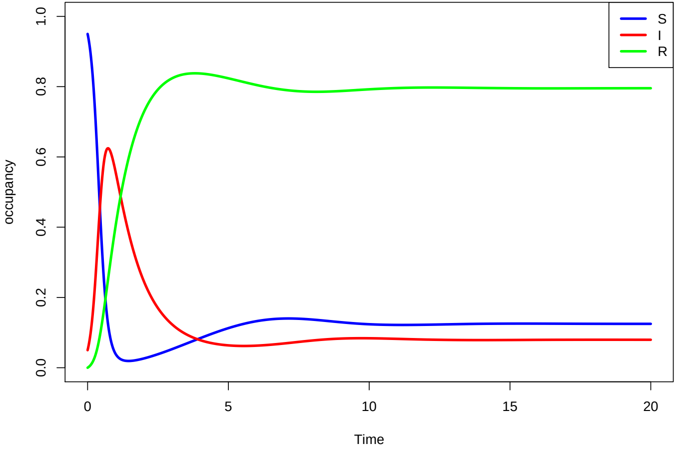
CTMC $N = 1000$

EPIDEMICS EXAMPLE CONTINUED



CTMC $N = 10000$

EPIDEMICS EXAMPLE CONTINUED



Limit ODE

Outline

STATIONARY REGIME

- The fluid approximation and mean field theorems provide conditions for the convergence up to any finite time horizon.
- They do not predict convergence of the stationary regime.
- This is because they hold for any possible trajectory of the ODE, including unstable ones.
- In order to provide some result for the stationary behaviour, one has to look at the Phase Space Properties of the system of ODEs.

SOME DEFINITIONS

DEFINITIONS

- Flow of the ODE: $\xi(t, x)$
- Orbit of the flow, starting from x: $\gamma(x)$
- Forward orbit of the flow, starting from x: $\gamma^+(x)$
- Invariant set A iff $\gamma(x) \subset A$, for $x \in A$
- Attractor: invariant set A such that there is a neighborhood U of A with $\lim_{t\to\infty} d_H(\xi(t, x), A) = 0$ uniformly for $x \in U$
- Basin of attraction of A:
  $B(A) = \{x \in E \mid \lim_{t\to\infty} d_H(\xi(t, x), A) = 0\}$

BIRKHOFF CENTRE AND INVARIANT MEASURES

BIRKHOFF CENTRE OF A FLOW

The Birkhoff centre $B(\xi)$ of a flow ξ is, informally, the set of limit points of the flow (steady states, limit circles, etc.).

INVARIANT MEASURE OF A FLOW

A probability measure μ on $(E\,\mathcal{B})$ is invariant for the flow ξ iff for each $A \in \mathcal{B}$ and $t \geq 0$

$$\mu(\xi^{-1}(t, A)) = \mu(A).$$

INVARIANT MEASURES AND BIRKHOFF CENTRE

Any invariant probability measure μ for the flow ξ has support contained in $B(\xi)$.

Convergence of Invariant Measures

Theorem

Let $\mu^{(N)}$ be an invariant measure for $\mathbf{X}^{(N)}(t)$. Any limit point μ (w.r.t. the weak topology) of the sequence $\mu^{(N)}$ is an invariant measure of the flow ξ.

In other words: $\mathbf{X}^{(N)}(t)$ spends most of its time close to the Birkhoff centre $B(\xi)$ of the flow.

Corollary

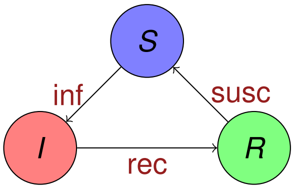
If $X^{(N)}(t)$ are irreducible and the ODE have a unique globally attracting stable fixed point $\bar{x}$, then $\mu^{(N)} \to \mu$, where μ concentrates the mass on $\bar{x}$.

Fixed Point Method

The fixed point method for mean field analysis approximates the stationary distribution with the value of the occupancy measure of the ODE fixed, if it is unique.
However, global attractiveness has to be proved.

EXAMPLE: SIR EPIDEMICS

Global attractiveness is a crucial property. Consider again the SIR model and the set of fluid equations.



$$\begin{cases} \frac{dx_S}{dt} = k_S x_R - k_I x_I x_S \\ \frac{dx_I}{dt} = k_I x_I x_S - k_R x_I \\ \frac{dx_R}{dt} = k_R x_I - k_S x_R \end{cases}$$

ODEs have two fixed points: $\left(\frac{k_R}{k_I}, \frac{k_S(k_I - k_R)}{k_I(k_S + k_R)}, \frac{k_R(k_I - k_R)}{k_I(k_S + k_R)}\right)$, if $\frac{k_R}{k_I} < 1$, and $(1, 0, 0)$

No matter how large is N, all trajectories of the CTMC will eventually reach the state in which the epidemics is extinct: the steady state measure of $\hat{\mathbf{X}}^{(N)}$ is the Dirac delta on $(1, 0, 0)$.

Outline

1. Fluid Approximation

2. Steady State approximation

3. Rewards

REWARDS

Reward measures are a very useful companion of CTMC population models. They allow us to capture useful measures, like the throughput of a system, or the energy consumption.

We consider here two classes of reward measures, all state-based.

REWARD FUNCTION

$\rho : E \to \mathbb{R}_{\geq 0}$ is the reward associated to a state $\mathbf{x} \in E$.
We assume ρ is continuous in E.

We assume rewards depend on the normalised state.

INSTANTANEOUS AND CUMULATIVE REWARDS

INSTANTANEOUS REWARD

The expected value of ρ at time t

$$R_I^{(N)}(t) = \mathbb{E}\left[\rho(\hat{\mathbf{X}}^{(N)}(t))\right]$$

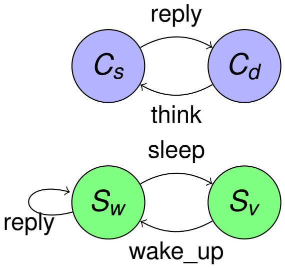
CUMULATIVE REWARDS

The expected reward accumulated up to time t

$$R_C^{(N)}(t) = \mathbb{E}\left[\int_0^t \rho(\hat{\mathbf{X}}^{(N)}(s))ds\right]$$

EXAMPLE: QUEUE MODEL WITH SERVER VACATION

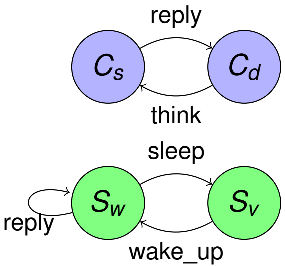
Consider a model of a closed queue network, with a (M/M/m) service station and a delay station, and assume servers can take a vacation, to save energy.



Four variables: C_s, C_d, S_w, S_v.

- $(\mathit{reply}, \top, (-1, +1, 0, 0), k_r \min\{C_s, S_w\})$
- $(\mathit{think}, \top, (+1, -1, 0, 0), k_t C_d)$
- $(\mathit{sleep}, \top, (0, 0, -1, +1), k_v S_w)$
- $(\mathit{wake\_up}, \top, (0, 0, +1, -1), k_w S_v)$

EXAMPLE: QUEUE MODEL WITH SERVER VACATION



Four variables: C_s, C_d, S_w, S_v.
αN clients, βN servers

- $(\mathit{reply}, \top, (-1, +1, 0, 0), k_r \min\{C_s, S_w\})$
- $(\mathit{think}, \top, (+1, -1, 0, 0), k_t C_d)$
- $(\mathit{sleep}, \top, (0, 0, -1, +1), k_v S_w)$
- $(\mathit{wake\_up}, \top, (0, 0, +1, -1), k_w S_v)$

$$(c_s, c_d, c_w, c_v) = (C_s, C_d, S_w, S_v)/N$$

REWARDS

THROUGHPUT: $\rho_t(c_s, c_d, c_w, c_v) = k_r \min\{c_s, s_w\})$

ENERGY CONSUMPTION: $\rho_u(c_s, c_d, c_w, c_v) = u_s \cdot s_w$

Convergence of Instantaneous Rewards

$\rho : E \to \mathbb{R}$ is a (bounded) continuous function

Continuous Mapping Theorem

If $\mathbf{X}^{(N)} \to \mathbf{X}$ (a.s./ in prob.) and f is $\mathbf{X}$-a.s. continuous (i.e. $f(\mathbf{X})$ is continuous with probability one), then $f(\mathbf{X}^{(N)}) \to f(\mathbf{X})$.

Bounded convergence

If $\mathbf{X}^{(N)} \to \mathbf{X}$ (a.s./ in prob.) and $\mathbb{E}[\mathbf{X}] < \infty$ and $\|\mathbf{X}^{(N)}\| \leq M$ for each N, then $\mathbb{E}[\|\mathbf{X}^{(N)} - \mathbf{X}\|] \to 0$ (convergence in mean).

Corollary (of Kurtz theorem)

$$\mathbb{E}[\rho(\hat{\mathbf{X}}^{(N)}(t))] \to \mathbb{E}[\rho(\mathbf{x}(t))] = \rho(\mathbf{x}(t))$$

CONVERGENCE OF CUMULATIVE REWARDS

$f_C(\mathbf{x}) = \int_0^t \rho(\mathbf{x}(s))ds$, for fixed t, can be seen as a functionals of a trajectory $\mathbf{x}$ of the CTMC, which is a cadlag function with values in E. Call $\mathcal{E}$ this set, then $f_C : \mathcal{E} \to \mathbb{R}$

WEAK CONVERGENCE

Let $\mathbf{X}^{(N)}, \mathbf{X}$ have values in $\mathcal{E}$. $\mathbf{X}^{(N)} \Rightarrow \mathbf{X}$ (weakly) if and only if, for each continuous and bounded functional $f : \mathcal{E} \to \mathbb{R}$, it holds that

$$\mathbb{E}[f(\mathbf{X}^{(N)})] \to \mathbb{E}[f(\mathbf{X})]$$

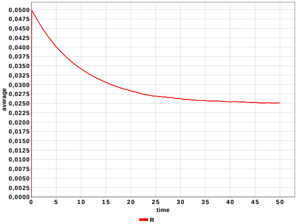
COROLLARY (OF KURTZ THEOREM)

By Kurtz theorem $\hat{\mathbf{X}}^{(N)} \Rightarrow \mathbf{x}$ (weakly), and (if E is compact) f_C is a continuous and bounded functional, so that:

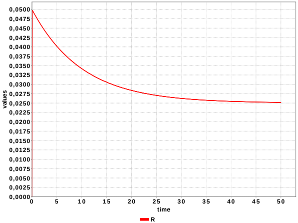
$$\mathbb{E}\left[\int_0^t \rho(\hat{\mathbf{X}}^{(N)}(s))ds\right] \to \mathbb{E}\left[\int_0^t \rho(\mathbf{x}(s))ds\right] = \int_0^t \rho(\mathbf{x}(s))ds$$

EXAMPLE: QUEUE MODEL WITH SERVER VACATION

Throughput

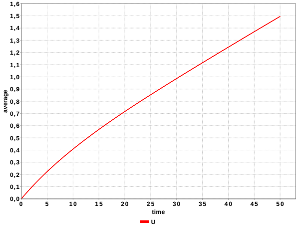


$N = 1000$

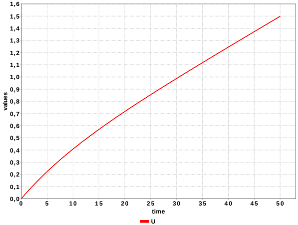


fluid

Energy consumption



$N = 1000$



fluid

REFERENCES

- T. Kurtz, Solutions of ordinary differential equations as limits of pure jump Markov processes, Journal of Applied Probability 7 (1970) 49-58.
- T. Kurtz, Approximation of population processes, SIAM, 1981.
- T. Kurtz, S. Ethier, Markov Processes - Characterisation and Convergence, Wiley, 1986.
- R. Darling, Fluid limits of pure jump Markov processes: A practical guide, ArXiv Mathematics e-printsarXiv:arXiv:math/0210109
- R. Darling, J. Norris, Differential equation approximations for markov chains, Probability Surveys 5 (2008) 37-79.
- Le Boudec et al. A generic mean field convergence result for systems of interacting objects. *QEST*, 2007.
- Benaïm and Le Boudec. A class of mean field interaction models for computer and communication systems. *Performance Evaluation*, n. 65, 2008.
- Benaïm and Weibull. Deterministic approximation of stochastic evolution in games. *Econometrica*, n. 71, 2003.

REFERENCES

- Mirco Tribastone, Jie Ding, Stephen Gilmore, and Jane Hillston. Fluid Rewards for a Stochastic Process Algebra, TSE, 2012.
- Anton Stefanek, Richard A. Hayden, Mark Mac Gonagle, Jeremy T. Bradley: Mean-Field Analysis of Markov Models with Reward Feedback. ASMTA 2012: 193-211.
- Anton Stefanek, Richard A. Hayden, Jeremy T. Bradley: Fluid analysis of energy consumption using rewards in massively parallel markov models. ICPE 2011: 121-132