Household epidemic models and McKean-Vlasov Poisson driven SDEs

R. Forien and É. Pardoux

Abstract

This paper presents a new view of household epidemic models, where we exploit the fact that the interaction between the households is of mean field type. We thus obtain in the limit of infinitely many households a nonlinear Markov process solution of a McKean–Vlasov type Poisson driven SDE, and a propagation of chaos result. We also define a basic reproduction number R_0, and show that if $R_0 > 1$, then the nonlinear Markov process has a unique non trivial ergodic invariant probability measure, whereas if $R_0 \leq 1$, it converges to 0 as t tends to infinity.

1 Introduction

In this paper, we present a new view of household epidemic models. Motivated by its simplicity, we present it in the particular case of the SIS model, but the same approach can be developed for other types of epidemic models, like the SIR, SIRS, SIR model with demography, and others. We recall that S stands for susceptible, I for infected and R for removed.

Household models, which have been mainly presented in the framework of the SIR model so far, is a key example of two–level mixing models. A very natural step in changing homogeneous epidemics models into more realistic models is to include households, which are small groups of individuals who interact more frequently within their group than with other individuals in the population. This describes both the situation of human populations, but also of many domestic animal populations, where cages/sheds in poultry farms of pens/fields in sheep/cattle farms play the role of households.

Household models can be roughly described as follows. The total population is the union of households of relatively small (and varying) size. Each infectious individual infects any other individuals in the same household at a "local rate" λ_L, and any other individual in the total population at a "global rate" λ_G divided by the total population size. In the last sentence, "any other" means "chosen uniformly at random". The infectious periods are i.i.d., in our case exponential with a given parameter γ (since we want to have a Markov model).

The first papers on epidemics models with two level of mixing go back to the 1950's, with Rushton and Mautner [9] who study deterministic models, Bartlet [3] and Daley [4] who study stochastic models. We refer to Ball, Mollison and Scalia–Tomba [1] who give a deep study of stochastic SIR epidemic models with two levels of mixing, as well as to Ball and Sirl [2] for an up–to–date presentation of stochastic SIR epidemics in structured populations, and for more references.

Our viewpoint in this paper is to study asymptotic results as the number of households (and hence also the total population size) tends to infinity, while the household sizes remain unchanged. It is easy to see that the interaction between the various households in of *mean field* type. This is reminiscent of the situation of particle systems which was studied by Sznitman [10]. We establish a result of *propagation of chaos*, and prove that in the limit of an infinite number of households, the typical epidemic in a household is a so–called nonlinear Markov process, whose transition depends not only upon the situation of the epidemic in the household, but also upon its probability law (through its mean, which is the limiting effect of the infections coming from the other households). Similar non–linear Markov processes have a long history, with in particular the work of McKean [7]. The SDE of those nonlinear Markov processes are called McKean–Vlasov SDEs. Most of the literature on that topic treats Brownian driven SDEs. However, Léonard [6] considers an epidemic model where the infection is the effect of a mean field interaction, and he obtains a McKean–Vlasov type SDE of Poissonian type as a law of large numbers limit.

Should we assume that the household sizes were bounded, then the existence and uniqueness of the nonlinear Markov process would be very elementary. Indeed, the Fokker–Planck equation for the evolution of its law would be a finite dimensional system of ODEs with locally Lipschitz coefficients, whose solution cannot explode since it is a probability distribution. Once all the marginal laws of the process are specified, then the SDE for the nonlinear Markov process becomes a classical easy to solve Poisson driven SDE. However, we only assume that the household size is a square integrable random variable. We prove existence and uniqueness of a solution of the nonlinear Fokker–Planck equation by a fixed point argument and approximation by both a decreasing sequence of supersolutions, and an increasing sequence of subsolutions.

We next study the large time behavior of our limiting SDE. We define the basic reproduction number R_0, which is the mean number of households infected as a result of a local infection in a typical infected household, started with one infectious. We give an explicit formula for R_0. If $R_0 \leq 1$, then the number of infectious individuals in a typical household tends to 0 as $t \to \infty$, whereas if $R_0 > 1$, the law of that number converges to an invariant measure which is not the Dirac measure at zero. Our results extend well–known classical results concerning the case where all households have size 1 (the homogeneous model). Note that we shall study the fluctuations around the law of large numbers obtained in the present paper in another publication.

The paper is organised as follows. The model is defined precisely in section 2. Section 3 states the three main results of the paper, namely Theorem 3.1 which gives the existence and uniqueness of the nonlinear Markov process, Theorem 3.2 which states the propagation of chaos result (which might be considered as a law of large numbers), and finally Theorem 3.3 which gives the large time behavior of the nonlinear Markov process. Section 4 studies what we call the "forced process", which is our nonlinear Markov process, where we replace the unknown quantity $\mathbb{E}[X(t)]$ by a given function $m(t)$. In particular, we establish the monotonicity property of the forced process as a function of m. That property is exploited in an essential way in section 5 for the proof of Theorem 3.1. Section 6 is devoted to the proof of Theorem 3.2 and finally section 7 to the definition and computation of R_0, and the proof of Theorem 3.3. In this last section, we use in particular a comparison with a non–Markov continuous time branching process.

2 Definition of the model

We consider an SIS household epidemic model. In our model, the population consists of N households, with sizes $\nu_1, \nu_2, \dots, \nu_N$, where the ν_i's are i.i.d. $\mathbb{N}$–valued random variables. Let $X_i^N(t)$ denote the number of infectious individuals in the i–th household at time t.

We suppose that each infected individual can infect another individual within the same household at rate λ_L, for some $\lambda_L > 0$ (the infected individual is chosen uniformly from those in the household, and if it is already infected, nothing happens). Moreover, each infected individual can infect another individual chosen *uniformly from the whole population* at rate λ_G, for some λ_G (again, if it is already infected nothing happens). Finally, each infected individual becomes suceptible at rate γ, for $\gamma > 0$. The parameters λ_L and λ_G are the rates of local (respectively global) infections. We note that, for each global infection, choosing an individual uniformly from the population is equivalent to first choosing a household from the size-biased distribution and then choosing an individual uniformly in this household.

Below is a more formal definition of this process. Let

$$\mathcal{X} = \{(n,k) \in \mathbb{N} \times \mathbb{Z}_+ : n \geq 1, 0 \leq k \leq n\}.$$

Definition 2.1 (SIS household epidemic model)**.** *Fix $\lambda_L > 0$, $\lambda_G > 0$ and $\gamma > 0$. Let $\{(\nu_i, X_i(0)), i \geq 1\}$ be i.i.d. $\mathcal{X}$-valued random variables such that $\mathbb{E}[\nu_1^2] < +\infty$ and let $(P_{inf,i}(t), t \geq 0, i \geq 1)$ and $(P_{rec,i}(t), t \geq 0, i \geq 0)$ be mutually independent standard Poisson processes, which are also independent of $\{(\nu_i, X_i(0)), i \geq 1\}$. We define $\overline{\nu}^N = \frac{1}{N}\sum_{i=1}^N \nu_i$. For $N \geq 1$, let $(X_1^N(t), \ldots, X_N^N(t), t \geq 0)$ be the solution of the following SDE:*

(1)

$$X_i^N(t) = X_i(0) + P_{inf,i}\left(\int_0^t \left(1 - \frac{X_i^N(s)}{\nu_i}\right)\left[\lambda_L X_i^N(s) + \lambda_G \frac{\nu_i}{\overline{\nu}^N}\frac{1}{N}\sum_{j=1}^N X_j^N(s)\right] ds\right) - P_{rec,i}\left(\gamma \int_0^t X_i^N(s) ds\right).$$

We call this process the SIS household model with N households.

The fact that there exists a unique solution to (1) follows from a standard arguement which exploits the fact that the jumps are isolated, and the process remains constant between its jumps. The distribution of the ν_i's will be fixed throughout the paper, and we set

$$\pi(n) = \mathbb{P}(\nu_1 = n), \qquad \overline{\pi} = \mathbb{E}[\nu_1].$$

We shall also use the size-biased distribution of the ν_i's and its first moment, which we define as

$$\pi^+(n) = \frac{n\pi(n)}{\overline{\pi}}, \qquad \overline{\pi}^+ = \sum_{n\geq 1} n\pi^+(n) = \frac{\mathbb{E}[\nu^2]}{\mathbb{E}[\nu]}.$$

We note that the different households only interact through the mean number of infected individuals in the N households, *i.e.* it is a *mean-field* interaction. We thus expect that, as the number of households N becomes very large, any finite subset of households are asymptotically mutually independent and each one evolves according to the following SDE:

(2)

$$X(t) = X(0) + P_{inf}\left(\int_0^t \left(1 - \frac{X(s)}{\nu}\right)\left[\lambda_L X(s) + \lambda_G \frac{\nu}{\overline{\pi}}\mathbb{E}[X(s)]\right] ds\right) - P_{rec}\left(\gamma \int_0^t X(s)ds\right),$$

where $(\nu, X(0))$ has the same law as $(\nu_1, X_1^N(0))$ and P_{inf} and P_{rec} are two independent standard Poisson processes which are also independent of $(\nu, X(0))$. This is what is called *propagation of chaos* [10], and will be made more precise in Theorem 3.2 below.

This equation is a McKean-Vlasov Poisson driven SDE, because the transition rates of $(X(t), t \geq 0)$ depend on the law of $X(t)$ (specifically on its expectation). We refer to McKean [7] for the study of similar Brownian driven SDEs. As we will see later, this equation defines a semigroup acting on probability distributions on $\mathbb{N} \times \mathbb{Z}_+$ but, contrary to ordinary Markov processes, this semigroup is non-linear (because of the term $\mathbb{E}[X(s)]$ appearing on the right hand side of (2)). For this reason we will call $(X(t), t \geq 0)$ the *non-linear* Markov process.

3 Main results

Existence and uniqueness of the non-linear process. It is not clear *a priori* that there exists a process solving (2), much less that it is unique.

Suppose for a moment that it exists and set

$$\mu_{n,k}(t) = \mathbb{P}(X(t) = k, \nu = n).$$

Then, $\mu(t) = \{\mu_{n,k}(t), (n,k) \in \mathcal{X}\}$ is the law of $(\nu, X(t))$ and

$$\forall n \geq 1, \quad \sum_{k=0}^{n} \mu_{n,k}(t) = \pi(n). \tag{3}$$

Equation (2) then implies that $\{\mu_{n,k}(t), t \geq 0, (n,k) \in \mathcal{X}\}$ solves the following non-linear Fokker-Planck equation:

(4)

$$\frac{d\mu_{n,k}(t)}{dt} = \mu_{n,k-1}(t)\left(1 - \frac{k-1}{n}\right)\left[\lambda_L(k-1) + \lambda_G \frac{n}{\overline{\pi}} \sum_{i=1}^{\infty}\sum_{j=1}^{i} j\mu_{i,j}(t)\right]$$
$$-\mu_{n,k}(t)\left\{\left(1-\frac{k}{n}\right)\left[\lambda_L k + \lambda_G \frac{n}{\overline{\pi}} \sum_{i=1}^{\infty}\sum_{j=1}^{i} j\mu_{i,j}(t)\right] + \gamma k\right\} + \mu_{n,k+1}(t)\gamma(k+1),$$

with the convention that $\mu_{n,-1}(t) = \mu_{n,n+1}(t) = 0$. Note that (4) defines an infinite system of coupled ordinary differential equations. We then have the following theorem.

Theorem 3.1. *Assume that the second moment of the probability distribution π is finite. Then, given a probability measure $\mu(0) = \{\mu_{n,k}(0), (n,k) \in \mathcal{X}\}$ satisfying* (3)*, there exists a unique time dependent probability measure $(\mu(t),\ t \geq 0)$ on $\mathcal{X}$ which solves the system of ODEs* (4)*. Moreover, given a random variable (ν, X_0) which is such that $\mathbb{P}(X_0 = k, \nu = n) = \mu_{n,k}(0)$ for $(n,k) \in \mathcal{X}$, the SDE* (2) *has a unique solution $(X(t),\ t \geq 0)$ which is such that for each $t \geq 0$, $\mathbb{P}(X(t) = k, \nu = n) = \mu_{n,k}(t)$ for each $(n,k) \in \mathcal{X}$.*

We prove this theorem in Section 5.

Propagation of chaos. We now deal with the limiting behaviour of the household model of Defintion 2.1 as the number of households N tends to infinity. For $T > 0$, let $\mathcal{P}(D([0,T], \mathcal{X}))$ denote the space of probability measures on the sample paths space $D([0,T], \mathcal{X})$. Also let $\mu \in \mathcal{P}(D([0,T], \mathcal{X}))$ denote the law of the non-linear Markov process $((\nu, X(t)), t \in [0,T])$, given by Theorem 3.1.

Theorem 3.2 (Propagation of chaos in the SIS household model)**.** *Assume that $\{(\nu_i, X_i(0)), i \geq 1\}$ are independent and identically distributed $\mathcal{X}$-valued random variables such that $\mathbb{E}[\nu_1^2] < +\infty$. For all*

$N \geq 1$, let $(X_i^N(t), t \geq 0, 1 \leq i \leq N)$ be the solution of equation (1). Define $\mu_N \in \mathcal{P}(D([0,T], \mathcal{X}))$ by

$$\mu_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{(\nu_i, X_i^N(\cdot))}.$$

Then the random measure μ_N converges weakly to μ as $N \to \infty$ in probability. Moreover, for any $k \geq 1$,

$$\mathit{Law}\left((\nu_1, X_1^N(\cdot)), \ldots, (\nu_k, X_k^N(\cdot)\right) \Rightarrow \mu^{\otimes k} \quad \text{as } N \to \infty$$

in $\mathcal{P}\left(D([0,T], \mathcal{X}^k)\right)$.

We prove Theorem 3.2 in Section 6. Note that by Proposition 2.2 in [10], the second part of the theorem follows from the convergence of the empirical measures μ_N. Theorem 3.2 says two things: as N becomes large, any finite subset of households behaves asymptotically as independent copies of the non-linear Markov process (2), and the global epidemic, as measured through the empirical measure μ_N, becomes asymptotically deterministic and equal to the law of the non-linear Markov process. It is then natural to ask whether the epidemic has an endemic equilibrium and if it is stable in the non-linear Markov process.

Large time behaviour of the non-linear Markov process. As is usual in SIS epidemic models, there is in our model a basic reproduction number R_0 such that if $R_0 > 1$, there exists a unique stable endemic equilibrium (*i.e.* the epidemic survives forever) and if $R_0 \leq 1$, the disease free equilibrium is globally asymptotically stable (the epidemic eventually dies out). This number is usually defined as the number of secondary infections produced by a single infected individual. Here, however, this number will be defined as the mean number of households which are infected by a single household, in which there is initially one infected individual and whose size is chosen according to the size-biased distribution π^+.

To do this, let $(I(t), t \geq 0)$ be the solution to the following SDE:

(5)

$$I(t) = I(0) + P_{inf}\left(\int_0^t \lambda_L \left(1 - \frac{I(s)}{\nu}\right) I(s) ds\right) - P_{rec}\left(\int_0^t \gamma I(s) ds\right),$$

where ν is distributed according to the probability distribution π and P_{inf} and P_{rec} are two independent standard Poisson processes, which are independent of $(\nu, I(0))$. Then $(I(t), t \geq 0)$ is the number of infected individuals in an isolated household.

We then define

$$R_0 = \lambda_G \sum_{n=1}^{\infty} \pi^+(n) \mathbb{E}\left[\int_0^{\infty} I(t)dt \middle| I(0) = 1, \nu = n \right]. \tag{6}$$

The large time behaviour of the non-linear process $(X(t), t \geq 0)$ of Theorem 3.1 is then given by the following result.

Theorem 3.3 (Large time behaviour of the non-linear Markov process)**.** *Let $(X(t), t \geq 0)$ be the unique solution to equation* (2)*, and assume that $\mathbb{E}[\nu^2] < +\infty$.*

i) If $R_0 > 1$, then there exists a unique probability distribution μ_∞ on $\mathcal{X}$ such that, if $\mathbb{P}(X(0) \geq 1) > 0$, $(\nu, X(t))$ converges in distribution to μ_∞ as $t \to \infty$. Moreover μ_∞ is non-trivial in the sense that $\mu_\infty \neq \pi \otimes \delta_0$.

ii) If $R_0 \leq 1$, then $X(t) \to 0$ in probability as $t \to \infty$.

We prove Theorem 3.3 in Section 7. This result should be seen as an analogue of the fact that the solution of the ODE

$$\frac{di(t)}{dt} = \lambda i(t) \left(1 - \frac{i(t)}{n}\right) - \gamma i(t) \tag{7}$$

converges as $t \to \infty$ to $n\left(1 - \frac{\gamma}{\lambda}\right)$ if $\lambda > \gamma$ and to 0 otherwise.

In the proof of Theorem 3.3, we shall also prove the following formula for R_0, which is of independent interest:

$$R_0 = \frac{\lambda_G}{\gamma} \sum_{n=1}^{\infty} \pi^+(n) \left(1 + \sum_{\ell=1}^{n-1} \left(\frac{\lambda_L}{\gamma}\right)^{\ell} \prod_{j=1}^{\ell} \left(1 - \frac{j}{n}\right)\right), \tag{8}$$

(see in particular the proof of Lemma 7.4 in Subsection 7.3).

Remark 3.4. *a) If $\pi = \delta_1$, every household is of size 1, and* (1) *reduces to the homogeneous SIS epidemic model, with parameters λ_G and γ (see [2]). We can then check that $\mathbb{E}[X(t)]$ solves the ODE* (7) *with $\lambda = \lambda_G$, and that* (6) *reduces to $R_0 = \lambda_G/\gamma$, as expected.*

b) The same is true if we take $\lambda_L = 0$ and keep π very general, the only infections in the system are global infections and the model reduces to the standard SIS epidemic model.

c) Another interesting case is when the size of all the households is very large. In that case, if we approximate $(I(t), t \geq 0)$ by a branching process, we see that R_0 should be approximated by $+\infty$ if $\lambda_L \geq \gamma$ and by $\lambda_G/(\gamma - \lambda_L)$ if $\lambda_L < \gamma$, and we see that $R_0 > 1$ is equivalent to $\lambda_G + \lambda_L > \gamma$.

4 The forced process

It is worth noting that if we replace $\mathbb{E}[X(s)]$ in (2) by any deterministic measurable function $s \mapsto m(s)$, then $(X(t), t \geq 0)$ becomes a time-inhomogeneous Markov process.

Definition 4.1 (The forced process)**.** *Let $m : \mathbb{R}_+ \to [0, \bar{\pi}]$ be a measurable function, (ν, X_0) an $\mathcal{X}$-valued random variable and P_{inf} and P_{rec} two independent standard Poisson processes which are also independent of (ν, X_0). Then the forced process $(X_t(m), t \geq 0)$ is defined as the solution to*

(9)

$$X_t(m) = X_0 + P_{inf}\left(\int_0^t \left[\lambda_L X_s(m) + \lambda_G \frac{\nu}{\bar{\pi}} m(s)\right]\left(1 - \frac{X_s(m)}{\nu}\right) ds\right) - P_{rec}\left(\int_0^t \gamma X_s(m) ds\right).$$

We call this process the *forced* process because we fix the intensity of global infections to be $\lambda_G \nu m(t)/\bar{\pi}$. The fact that there exists a unique strong solution to (9) follows from standard arguments similar to that used in (1).

Comparing (2) and (9), we see that solving (2) is equivalent to finding a measurable function m such that $m(t) = \mathbb{E}[X_t(m)]$ for all $t \geq 0$.

4.1 Graphical construction of the forced process

We are going to show that we can construct this process with the following procedure. Let $c(dk)$ denote the counting measure on $\mathbb{N}$. Conditionally on (ν, X_0), let Π_{rec}, Π_L and Π_G be three mutually independent Poisson point processes such that

- Π_{rec} is a Poisson point process on $\mathbb{R}_+ \times [\![1,\nu]\!]$ with intensity $\gamma dt \otimes c(dk)$,
- Π_L is a Poisson point process on $\mathbb{R}_+ \times [\![1,\nu]\!] \times [\![1,\nu]\!]$ with intensity $\frac{\lambda_L}{\nu} dt \otimes c(dk) \otimes c(dk)$,
- Π_G is a Poisson point process on $\mathbb{R}_+ \times [\![1,\nu]\!] \times [0,\overline{\pi}]$ with intensity $\frac{\lambda_G}{\overline{\pi}} dt \otimes c(dk) \otimes du$.

Let us describe the effect of these different processes before formally constructing the forced process. A point (t,i) in Π_{rec} means that if the individual i was infected at time t^-, it becomes suceptible at time t (it undergoes a remission). A point (t,i,j) in Π_L means that individual i can infect individual j at time t. This occurs if i is infected while j is suceptible at time t^-. Finally a point (t,i,u) in Π_G means that individual i can be infected from a global infection. We allow this infection to take place only if i is suceptible at time t^- and if $u \leq m(t)$.

In fact, we can view the total set of infected individuals at any time as the union of several local infections, each resulting from a previous global infection or from the individuals infected at time 0. To do this, note that Π_G is almost surely locally finite, so we can order its points according to their time coordinate. Thus let

$$\Pi_G = \{(t_k, i_k, u_k), k \geq 1, 0 < t_1 < t_2 < \ldots\}.$$

Let us then define a random set $I^k(t) \subset [\![1,\nu]\!]$ for all $t \geq 0$ as follows.

- For $t < t_k$, $I^k(t) = \emptyset$.
- At $t = t_k$, we set $I^k(t_k) = \{i_k\}$.
- For each $(t,i,j) \in \Pi_L$, if $i \in I^k(t^-)$, then $I^k(t) = I^k(t^-) \cup \{j\}$.
- For each $(t,i) \in \Pi_{rec}$, $I^k(t) = I^k(t^-) \cap \{i\}^c$.

We define in the same way the local infection resulting from the initially infected individuals $(I^0(t), t \geq 0)$, *i.e.* $I^0(0) = \{i : 1 \leq i \leq X_0\}$ and I^0 evolves according to the same rules as I^k for $k \geq 1$. We note that, for all $k \geq 0$, $(I^k(t), t \geq 0)$ is right-continuous with left limits.

Proposition 4.2. *For all $t \geq 0$, let*

$$X_t(m) = \left| I^0(t) \cup \bigcup_{k \geq 1} \{I^k(t) : u_k \leq m(t_k)\} \right|, \tag{10}$$

where $|\cdot|$ denotes the cardinal of a set. Then the process $(X_t(m), t \geq 0)$ is a solution to the SDE (9).

Proof. Clearly $X_0(m) = |I^0(t)| = X_0$. It remains to check that the waiting times between upward and downward jumps of $X_t(m)$ are distributed as exponential variables with the correct rates.

If the current value of $X_t(m)$ is x, then the next remission takes place at the next point $(t, i) \in \Pi_{rec}$ with $i \in I^k(t^-)$ for some $k \geq 0$ with $u_k \leq m(t_k)$ (we can set $u_0 = \overline{\pi}$ and $t_0 = 0$). This happens at instantaneous rate γx, as in (9).

Likewise, the next time an individual currently infected infects a suceptible individual is given by the next point $(t, i, j) \in \Pi_L$ such that $i \in I^k(t^-)$ for some $k \geq 0$ with $u_k \leq m(t_k)$ and $j \notin I^k(t^-)$ for all such k. This happens at rate $\frac{\lambda_L}{\nu} x(\nu - x)$, as in (9).

Finally, the next time a suceptible individual becomes infected due to a global infection is the next $(t, i, u) \in \Pi_G$ such that $i \notin I^k(t^-)$ for all $k \geq 0$ such that $u_k \leq m(t_k)$ and $u \leq m(t)$. This happens at instantaneous rate $\frac{\lambda_G}{\overline{\pi}}(\nu - x)m(t)$, as in (9). □

4.2 Monotonicity of the forced process

With this construction, the next lemma is straightforward.

Lemma 4.3 (Monotonicity of the forced process)**.** *Suppose that $X_0^{(1)}$ and $X_0^{(2)}$ are two random variables such that $X_0^{(1)} \leq X_0^{(2)}$ almost surely. Also let m_1 and m_2 be two measurable functions from $\mathbb{R}_+$ to $[0, \overline{\pi}]$ such that $m_1(t) \leq m_2(t)$ for almost every $t \geq 0$. Then there exists a process $(X_t(m_1), t \geq 0)$ solving (9) with $m = m_1$ and $X_0 = X_0^{(1)}$, and a process $(X_t(m_2), t \geq 0)$ solving (9) with $m = m_2$ and $X_0 = X_0^{(2)}$, defined on the same probability space, such that, almost surely,*

$$X_t(m_1) \leq X_t(m_2), \quad \forall t \geq 0.$$

Proof. We use Proposition 4.2 to construct both processes with the same Poisson point processes Π_{rec}, Π_L and Π_G. We define $(I^{0,i}(t), t \geq 0)$ for $i \in \{1, 2\}$ as above with

$$I^{0,i}(0) = \{k : 1 \leq k \leq X_0^{(i)}\},$$

so that, almost surely, $I^{0,1}(0) \subset I^{0,2}(0)$. Then, from the evolution of $(I^{0,i}(t), t \geq 0)$, we deduce that $I^{0,1}(t) \subset I^{0,2}(t)$ for all $t \geq 0$. Furthermore, since $m_1 \leq m_2$,

$$\{k : u_k \leq m_1(t_k)\} \subset \{k : u_k \leq m_2(t_k)\}.$$

It then follows from equation (10) that $X_t(m_1) \le X_t(m_2)$. □

The following lemma will also be useful in the proof of existence and uniqueness of the non-linear process. For $t \ge 0$, $m : \mathbb{R}_+ \to [0, \overline{\pi}]$ measurable and μ_0 a probability measure on $\mathcal{X}$ whose first marginal is π, let

$$\overline{\mu}_t(m, \mu_0) = \mathbb{E}[X_t(m)], \tag{11}$$

where (ν, X_0) is distributed according to μ_0.

Lemma 4.4. *Suppose that μ_0 is as above. If m_1 and m_2 are two measurable functions from $\mathbb{R}_+$ to $[0, \overline{\pi}]$ satisfying $m_1(t) \le m_2(t)$ for almost every $t \ge 0$, then*

$$0 \le \overline{\mu}_t(m_2, \mu_0) - \overline{\mu}_t(m_1, \mu_0) \le \overline{\pi}^+ \lambda_G \int_0^t (m_2(s) - m_1(s)) ds.$$

Proof. The fact that $\overline{\mu}_t(m_2) - \overline{\mu}_t(m_1) \ge 0$ follows from Lemma 4.3. To prove the second inequality, we construct $(X_t(m_1), t \ge 0)$ and $(X_t(m_2), t \ge 0)$ as in Proposition 4.2. Then

$$0 \le X_t(m_2) - X_t(m_1) \le \left| \cup_{k \ge 1} \{ I^k(t) : m_1(t_k) \le u_k \le m_2(t_k) \} \right|.$$

Moreover, we can restrict the union to the values of k for which $t_k \le t$. Since $|I^k(t)| \le \nu$ for all $t \ge 0$, we can write

$$0 \le X_t(m_2) - X_t(m_1) \le \nu \left| \{ k \ge 1 : m_1(t_k) < u_k \le m_2(t_k), t_k \le t \} \right|. \tag{12}$$

Now, by the definition of Π_G, the right hand side is, conditionally on ν, ν times a Poisson random variable with parameter

$$\lambda_G \frac{\nu}{\overline{\pi}} \int_0^t (m_2(s) - m_1(s)) ds.$$

As a result, taking expectations in (12) (first conditionally on ν and then over the law of ν), we obtain

$$0 \le \overline{\mu}_t(m_2, \mu_0) - \overline{\mu}_t(m_1, \mu_0) \le \lambda_G \, \overline{\pi}^+ \int_0^t (m_2(s) - m_1(s)) ds,$$

and the lemma is proved. □

We shall come back to the forced process in the proof of Theorem 3.3, as it will be used to characterize the possible stationary distributions of the non-linear process.

5 Existence and uniqueness of the non-linear Markov process

We now set out to prove Theorem 3.1. We note that finding a solution to (4) is equivalent to finding a fixed point of

$$m(\cdot) \mapsto \overline{\mu}_{\cdot}(m, \mu_0). \tag{13}$$

Indeed, if m_* is a fixed point of this function, then $(X_t(m_*), t \geq 0)$ is a solution to (2). We thus need to prove that, given μ_0, there exists a unique fixed point of (13).

Proof of Theorem 3.1. Fix μ_0 and assume that (ν, X_0) is distributed according to μ_0. Let $(m^{+,k}, k \geq 0)$ and $(m^{-,k}, k \geq 0)$ be two sequences of functions defined by

$$m^{+,0}(t) = \overline{\pi}, \qquad m^{+,k+1}(t) = \overline{\mu}_t(m^{+,k}, \mu_0),$$
$$m^{-,0}(t) = 0, \qquad m^{-,k+1}(t) = \overline{\mu}_t(m^{-,k}, \mu_0),$$

where $\overline{\mu}_t(m, \mu_0)$ was defined in (11). Clearly, since $0 \leq \mathbb{E}[X_t(m)] \leq \overline{\pi}$,

$$m^{+,1}(t) \leq m^{+,0}(t), \qquad m^{-,1}(t) \geq m^{-,0}(t).$$

Then by induction, using Lemma 4.3, we obtain that, for all $k \geq 0$,

$$m^{-,k}(t) \leq m^{-,k+1}(t) \leq m^{+,k+1}(t) \leq m^{+,k}(t).$$

Hence $m^{+,k}$ and $m^{-,k}$ both converge pointwise. Let $m^{+,\infty}$ and $m^{-,\infty}$ be their respective limits. Then, using Lemma 4.4 with $m_1 = m^{+,\infty}$ and $m_2 = m^{+,k}$,

$$\left|\overline{\mu}_t(m^{+,\infty}) - m^{+,\infty}(t)\right| \leq \left|m^{+,k+1}(t) - m^{+,\infty}(t)\right| + \lambda_G \, \overline{\pi}^+ \int_0^t (m^{+,k}(s) - m^{+,\infty}(s)) ds.$$

The integral on the right hand side vanishes as $k \to \infty$ by dominated convergence and the first term vanishes because $m^{+,k}$ converges pointwise to $m^{+,\infty}$. As a result, $m^{+,\infty}$ (and also $m^{-,\infty}$ by the same argument) is a fixed point of (13). This shows existence of solutions to (4) (and thus to (2)).

To prove uniqueness, first note that, by induction and using Lemma 4.3, any fixed point m_* satisfies

$$m^{-,k}(t) \leq m_*(t) \leq m^{+,k}(t),$$

for all $k \geq 0$ and $t \geq 0$. Hence we also have

$$m^{-,\infty}(t) \leq m_*(t) \leq m^{+,\infty}(t).$$

To prove uniqueness, it is thus enough to prove that $m^{+,\infty}(t) = m^{-,\infty}(t)$ for all $t \geq 0$. Using Lemma 4.4 with $m_1 = m^{-,0}$ and $m_2 = m^{+,0}$, we obtain

$$0 \leq m^{+,1}(t) - m^{-,1}(t) \leq \overline{\pi}\,\overline{\pi}^+\lambda_G t,$$

and by induction, we deduce that, for $k \geq 1$,

$$0 \leq m^{+,k}(t) - m^{-,k}(t) \leq \overline{\pi}\frac{(\overline{\pi}^+\lambda_G t)^k}{k!}.$$

Leting $k \to \infty$, it follows that $m^{+,\infty}(t) = m^{-,\infty}(t)$ for all $t \geq 0$ and the theorem is proved. □

6 Propogation of chaos for the SIS household model

The aim of this section is to prove Theorem 3.2. As we have said before, using Proposition 2.2 in [10], the second part of the statement follows from the convergence of the empirical measures μ_N to the law of the non-linear process μ. We establish this convergence by showing that the sequence $\{\mu_N, N \geq 1\}$ is tight in $\mathcal{P}(D([0,T],\mathcal{X}))$, and identifying its possible limit points.

Lemma 6.1. *The sequence $\{\mu_N, N \geq 1\}$ is tight in $\mathcal{P}(D([0,T],\mathcal{X}))$.*

Proof. By Proposition 2.2(ii) in [10], the sequence $\{\mu_N, N \geq 1\}$ is tight if and only if the laws of $(\nu_1, X_1^N(\cdot))$ are tight, but this is straightforward from (1) where we see that the rate of increase is bounded by $(\lambda_L + \lambda_G)\nu_i$, while the rate of decrease is bounded by $\gamma\nu_i$. □

Next we note that equation (1) can be reformulated as follows. Let $\{\mathcal{M}_{inf,i}, i \geq 1\}$ and $\{\mathcal{M}_{rec,i}, i \geq 1\}$ be mutually independent random Poisson measures on $\mathbb{R}_+^2$ with intensity measure the Lebesgue measure, which are also independent of $\{(\nu_i, X_i(0)), i \geq 1\}$. Then, with the notation

$$\overline{\mu}_t^N = \frac{1}{N}\sum_{i=1}^N X_i^N(t)\,,$$

$$X_i^N(t) = X_i(0) + \int_0^t \int_0^\infty \mathbf{1}_{u \le \left(1 - \frac{X_i^N(s^-)}{\nu_i}\right)\left[\lambda_L X_i^N(s^-) + \lambda_G \frac{\nu_i}{\overline{\nu}^N} \overline{\mu}^N_{s-}\right]} \mathcal{M}_{inf,i}(ds\, du)$$
$$- \int_0^t \int_0^\infty \mathbf{1}_{u \le \gamma X_i^N(s^-)} \mathcal{M}_{rec,i}(ds\, du).$$

Clearly, for any $\phi : \mathcal{X} \mapsto \mathbb{R}$,

$$\phi(\nu_i, X_i^N(t)) = \phi(\nu_i, X_i^N(0))$$
$$+ \int_0^t [\phi(\nu_i, X_i^N(s^-)+1) - \phi(\nu_i, X_i^N(s^-))] \int_0^\infty \mathbf{1}_{u \le \left(1 - \frac{X_i^N(s^-)}{\nu_i}\right)\left[\lambda_L X_i^N(s^-) + \lambda_G \frac{\nu_i}{\overline{\nu}^N} \overline{\mu}^N_{s-}\right]} \mathcal{M}_{inf,i}(ds, du)$$
$$+ \int_0^t [\phi(\nu_i, X_i^N(s^-)-1) - \phi(\nu_i, X_i^N(s^-))] \int_0^\infty \mathbf{1}_{u \le \gamma X_i^N(s^-)} \mathcal{M}_{rec,i}(ds, du).$$

Let $\overline{\mathcal{M}}_{inf,i}$ and $\overline{\mathcal{M}}_{rec,i}$ denote the compensated measures

$$\overline{\mathcal{M}}_{inf,i}(ds, du) = \mathcal{M}_{inf,i}(ds, du) - dsdu,$$
$$\overline{\mathcal{M}}_{rec,i}(ds, du) = \mathcal{M}_{rec,i}(ds, du) - dsdu.$$

Then setting

$$M_i^\phi(t) = \int_0^t [\phi(\nu_i, X_i^N(s^-)-1) - \phi(\nu_i, X_i^N(s^-))] \int_0^\infty \mathbf{1}_{u \le \gamma X_i^N(s^-)} \overline{\mathcal{M}}_{rec,i}(ds, du)$$
$$+ \int_0^t [\phi(\nu_i, X_i^N(s^-)+1) - \phi(\nu_i, X_i^N(s^-))] \int_0^\infty \mathbf{1}_{u \le \left(1 - \frac{X_i^N(s^-)}{\nu_i}\right)\left[\lambda_L X_i^N(s^-) + \lambda_G \frac{\nu_i}{\overline{\nu}^N} \overline{\mu}^N_{s-}\right]} \overline{\mathcal{M}}_{inf,i}(ds, du),$$

we have

$$\phi(\nu_i, X_i^N(t)) = \phi(\nu_i, X_i^N(0))$$
$$+ \int_0^t [\phi(\nu_i, X_i^N(s)+1) - \phi(\nu_i, X_i^N(s))] \left(1 - \frac{X_i^N(s)}{\nu_i}\right) \left[\lambda_L X_i^N(s) + \lambda_G \frac{\nu_i}{\overline{\nu}^N} \overline{\mu}^N_s\right] ds$$
$$+ \gamma \int_0^t [\phi(\nu_i, X_i^N(s) - 1) - \phi(\nu_i, X_i^N(s))] X_i^N(s) ds + M_i^\phi(t).$$

We rewrite this identity in the form
(14)

$$\phi(\nu_i, X_i^N(t)) = \phi(\nu_i, X_i^N(0)) + \int_0^t [\mathcal{L}\phi](\nu_i, X_i^N(s), \overline{\nu}^N, \overline{\mu}^N_s) ds + M_i^\phi(t)$$

where for $n \ge 1$, $x \in \{0, 1, \dots, n\}$, $y \ge 0$ and $0 \le m \le y$,

$$\mathcal{L}\phi(n, x, y, m) = [\phi(n, x+1) - \phi(n, x)] \left(1 - \frac{x}{n}\right) \left[\lambda_L x + \lambda_G \frac{n}{y} m\right] + [\phi(n, x-1) - \phi(n, x)] \gamma x.$$

Proof of Theorem 3.2. Let μ_∞ be a limit point of the sequence μ^N. First note that, by the classical law of large numbers, for any bounded and measurable $\phi : \mathcal{X} \to \mathbb{R}$,

$$\mathbb{E}_{\mu_\infty}[\phi(\nu, X(0))] = \mathbb{E}[\phi(\nu_1, X_1(0))].$$

In order to identify the possible limit points of μ^N, we define, for $\mu \in \mathcal{P}(D([0,T], \mathcal{X}))$ and $0 \le s \le t \le T$,

$$\Phi_{s,t}(\mu) = \mathbb{E}_\mu \left[\left(\phi(\nu, X(t)) - \phi(\nu, X(s)) - \int_s^t \mathcal{L}\phi(\nu, X(r), \overline{\nu}, \overline{\mu}_r) dr \right) \psi_s(\nu, X(\cdot)) \right],$$

where

$$\overline{\mu}_t = \mathbb{E}_\mu \left[X(t) \right], \qquad \overline{\nu} = \mathbb{E}_\mu[\nu],$$

and ϕ is any bounded function from $\mathcal{X}$ to $\mathbb{R}$ and ψ_s is of the form

$$\psi_s(\nu, X(\cdot)) = \phi_1(\nu, X(s_1)) \ldots \phi_k(\nu, X(s_k))$$

with $0 \le s_1 \le \ldots \le s_k \le s$ and $\phi_1, \ldots, \phi_k$ are bounded functions from $\mathcal{X}$ to $\mathbb{R}$. By Theorem 3.1, the result will be proved if we show that

$$\Phi_{s,t}(\mu_\infty) = 0,$$

almost surely for any such function $\Phi_{s,t}$.

Using (14),

$$\Phi_{s,t}(\mu^N) = \frac{1}{N} \sum_{i=1}^N (M_i^\phi(t) - M_i^\phi(s)) \psi_s(\nu_i, X_i^N(\cdot)).$$

From the definition of M_i^ϕ,

$$\langle M_i^\phi, M_j^\phi \rangle_t = 0, \quad \forall i \neq j,$$

and

$$\langle M_i^\phi \rangle_t = \int_0^t \mathcal{G}\phi(\nu_i, X_i^N(s), \overline{\nu}^N, \overline{\mu}_s^N) ds,$$

where

$$\mathcal{G}\phi(n,x,y,m) = [\phi(n, x+1) - \phi(n,x)]^2 \left(1 - \frac{x}{n}\right) \left[\lambda_L x + \lambda_G \frac{n}{y} m\right] + [\phi(n, x-1) - \phi(n,x)]^2 \gamma x.$$

Note that, for $m \leq y$

$$\mathcal{G}\phi(n,x,y,m) \leq 4 \sup_{\mathcal{X}} |\phi|^2 \left(\lambda_L + \lambda_G + \gamma\right) n.$$

As a result,

$$\begin{aligned}\mathbb{E}\left[\Phi_{s,t}(\mu^N)^2\right] &= \frac{1}{N^2}\sum_{i=1}^N \mathbb{E}\left[(\langle M_i^\phi\rangle_t - \langle M_i^\phi\rangle_s)\psi_s(\nu_i, X_i^N(\cdot))^2\right] \\ &\leq \frac{C}{N}\mathbb{E}[\nu_1],\end{aligned}$$

for some $C > 0$. It follows that

$$\Phi_{s,t}(\mu^N) \to 0,$$

in L^2 as $N \to \infty$, hence μ_∞ is equal to μ, the distribution of the non-linear process of (2). This proves Theorem 3.2. □

7 Large time behaviour of the non-linear Markov process

Let us start this section by noting that if the non-linear process of (2) with initial distribution μ_0 is stationary, then the forced process with initial distribution μ_0 and with $m(t) = \overline{\mu}_0 = \mathbb{E}_{\mu_0}[X(0)]$ is also stationary. Thus to study the possible stationary distributions of the non-linear process, we first study the large-time behaviour of the forced process.

7.1 The large-time behaviour of the forced process

Suppose that we take $m(t) = m$ for all $t \geq 0$ for some $m \in [0, \overline{\pi}]$. Then $(X_t(m), t \geq 0)$ becomes a homogeneous continuous-time Markov process. On the event $\{\nu = n\}$, it takes values in $[\![1, n]\!]$. If $m > 0$, then it is positive recurrent on this set, while if $m = 0$, 0 is the only absorbing state for $X_t(m)$. As a result, conditionally on $\nu = n$, $(X_t(m), t \geq 0)$ admits a unique stationary distribution. It follows that $((\nu, X_t(m)), t \geq 0)$ admits a unique stationary distribution $\mu_\infty(m)$.

This distribution can be obtained as in Proposition 4.2 in the following way. Let $\overleftarrow{\Pi}_{rec}$, $\overleftarrow{\Pi}_L$ and $\overleftarrow{\Pi}_G$ be independent Poisson point processes as above, but on $\mathbb{R}_-$ instead of $\mathbb{R}_+$ for the first coordinate. We can then order the points in $\overleftarrow{\Pi}_G$ in decreasing order:

$$\overleftarrow{\Pi}_G = \{(t_k, i_k, u_k), k \geq 1, 0 > t_1 > t_2 > \ldots\}.$$

The points in $\overleftarrow{\Pi}_G$ represent global infection which took place in the past. We then perform the same construction of $I^k(t)$, this time for $t \leq 0$, and we set

$$X_\infty(m) = \left| \cup_{k\geq 1}\{I^k(0) : u_k \leq m\} \right|.$$

Proposition 7.1. *For each $m \in [0, \overline{\pi}]$, $X_\infty(m)$ is distributed according to $\mu_\infty(m)$.*

Proof. For $t \geq 0$, let

$$\tilde{X}_t(m) = \left| \cup_{k\geq 1}\{I^k(0) : u_k \leq m, t_k \geq -t\} \right|.$$

In other words, we only consider the local epidemics which started after time $-t$. Then from Proposition 4.2, we see that for each $t \geq 0$, $\tilde{X}_t(m)$ is distributed as $X_t(m)$, where $X_t(m)$ is the solution of (9) with $m(t) = m$ and $X_0 = 0$. By the ergodic theorem for homogeneous Markov processes, $X_t(m)$, and hence $\tilde{X}_t(m)$, converge in distribution as $t \to \infty$ to $\mu_\infty(m)$. At the same time, we see from the definition of $\tilde{X}_t(m)$ and $X_\infty(m)$ that

$$\tilde{X}_t(m) = \sum_{i=1}^{\nu} 1_{\left\{\exists k\geq 1: i \in I^k(0), u_k \leq m, t_k \geq -t\right\}}, \quad X_\infty(m) = \sum_{i=1}^{\nu} 1_{\left\{\exists k\geq 1: i \in I^k(0), u_k \leq m\right\}}.$$

Hence by monotone convergence,

$$\tilde{X}_t(m) \to X_\infty(m) \text{ as } t \to \infty,$$

almost surely, and the lemma is proved. □

The next lemma says that $m(t) \to m_\infty$ as $t \to \infty$ is sufficient for $X_t(m)$ to converge in distribution to $\mu_\infty(m_\infty)$.

Lemma 7.2 (Large time behaviour of the forced process)**.** *Suppose that $m : \mathbb{R}_+ \to [0, \overline{\pi}]$ is measurable and that*

$$m_\infty = \lim_{t\to\infty} m(t)$$

exists. Then $X_t(m)$ converges in distribution to $\mu_\infty(m_\infty)$ as t tends to infinity.

Proof. Suppose for now that $0 < m_\infty < \overline{\pi}$. Then for all $\varepsilon > 0$, there exists t_ε such that, for all $t \geq t_\varepsilon$,

$$m_\infty - \varepsilon \leq m(t) \leq m_\infty + \varepsilon.$$

We choose ε small enough that $0 \leq m_\infty - \varepsilon$ and $m_\infty + \varepsilon \leq \overline{\pi}$. We then define two functions m^+ and m^- by

$$m^+(t) = \overline{\pi} 1_{\{t<t_\varepsilon\}} + (m_\infty + \varepsilon) 1_{\{t \geq t_\varepsilon\}}, \quad m^-(t) = (m_\infty - \varepsilon) 1_{\{t \geq t\varepsilon\}}.$$

Then $m^- \leq m \leq m^+$, so by Lemma 4.3, we can construct jointly the three processes $(X_t(m^-), t \geq 0)$, $(X_t(m), t \geq 0)$ and $(X_t(m^+), t \geq 0)$ such that, almost surely,

$$X_t(m^-) \leq X_t(m) \leq X_t(m^+), \quad \forall t \geq 0.$$

It follows that, for each $t \geq 0$ and each $k \in \mathbb{N}$,

$$\mathbb{P}\left(X_t(m^+) \leq k\right) \leq \mathbb{P}\left(X_t(m) \leq k\right) \leq \mathbb{P}\left(X_t(m^-) \leq k\right).$$

Since m^+ and m^- are both constant after time t_ε (which is deterministic), as $t \to \infty$, $X_t(m^+)$ and $X_t(m^-)$ respectively converge in distribution to $\mu_\infty(m_\infty + \varepsilon)$ and $\mu_\infty(m_\infty - \varepsilon)$. Thus, letting $t \to \infty$ above,

$$\begin{aligned} \mu_\infty(m_\infty + \varepsilon)\left(\{0, \ldots, k\}\right) &\leq \liminf_{t \to \infty} \mathbb{P}\left(X_t(m) \leq k\right) \\ &\leq \limsup_{t \to \infty} \mathbb{P}\left(X_t(m) \leq k\right) \leq \mu_\infty(m_\infty - \varepsilon)\left(\{0, \ldots, k\}\right). \end{aligned}$$

But, as $\varepsilon \downarrow 0$, the measure $\mu_\infty(m_\infty \pm \varepsilon)$ converges weakly to $\mu_\infty(m_\infty)$ (in fact the construction in Proposition 7.1 gives a construction of $X_\infty(m \pm \varepsilon)$ and $X_\infty(m)$ such that $X_\infty(m \pm \varepsilon) \to X_\infty(m)$ almost surely as $\varepsilon \downarrow 0$, using monotone convergence as in the proof of Proposition 7.1). Hence letting $\varepsilon \downarrow 0$ above, we obtain, for any $k \geq 0$,

$$\mathbb{P}\left(X_t(m) \leq k\right) \to \mu_\infty(m_\infty)\left(\{0, \ldots, k\}\right) \text{ as } t \to \infty,$$

and the lemma is proved. If $m_\infty = 0$, then we can take instead $m^-(t) = 0$, and if $m_\infty = \overline{\pi}$, then we take $m^+(t) = \overline{\pi}$, and the rest of the proof is essentially identical. □

7.2 The stationary distribution of the forced process

We now study in more detail the family of distributions $\mu_\infty(\cdot)$. For $m \in [0, \overline{\pi}]$, we set

$$\overline{\mu}_\infty(m) = \mathbb{E}\left[X_\infty(m)\right].$$

Lemma 7.3. *The function $m \mapsto \overline{\mu}_\infty(m)$ is continuous, non-decreasing and strictly concave on $[0, \overline{\pi}]$.*

Proof. Fix $m_1 \leq m_2$. Then, using the construction in Proposition 7.1, we have, almost surely,

$$X_\infty(m_1) \leq X_\infty(m_2).$$

Taking expectations, we obtain

$$\overline{\mu}_\infty(m_1) \leq \overline{\mu}_\infty(m_2).$$

Hence $m \mapsto \overline{\mu}_\infty(m)$ is non-decreasing.

The continuity follows from Proposition 7.1 and the monotone convergence theorem.

To show that it is concave, we will construct two random variables $X^\delta_\infty(m_1)$ and $X^\delta_\infty(m_2)$ distributed according to $\mu_\infty(m_1+\delta)$ and $\mu_\infty(m_2+\delta)$ such that

$$X^\delta_\infty(m_1) - X_\infty(m_1) \geq X^\delta_\infty(m_2) - X_\infty(m_2),$$

almost surely. To do this, we will add the same set of global infections (with rate $\lambda_G \delta\nu/\overline{\pi}$) to both processes.

Fix $\delta > 0$ and let Π^δ_G be an independent Poisson point process on $\mathbb{R}_- \times [\![1, \nu]\!]$ with intensity $\delta \frac{\lambda_G}{\overline{\pi}} dt \otimes c(dk)$. We then order the points in Π^δ_G as above,

$$\Pi^\delta_G = \{(t^\delta_k, i^\delta_k) : k \geq 1, 0 > t^\delta_1 > t^\delta_2 > \ldots\},$$

and we define $I^{k,\delta}(t)$ for $t \leq 0$ as above, using the same Poisson point processes of local infections and remission as before, *i.e.* $\overleftarrow{\Pi}_{rec}$ and $\overleftarrow{\Pi}_L$. We then define

$$X^\delta_\infty(m) = \left| \cup_{k\geq 1}\{I^k(0) : u_k \leq m\} \bigcup \cup_{k\geq 1}\{I^{k,\delta}(0) : k \geq 1\} \right|.$$

From Proposition 7.1, $X^\delta_\infty(m)$ is distributed according to $\mu_\infty(m+\delta)$. Furthermore,

$$X^\delta_\infty(m) - X_\infty(m) = \left| \cup_{k\geq 1}\{I^{k,\delta}(0), k\geq 1\} \bigcap \Big(\cup_{k\geq 1}\{I^k(0) : u_k \leq m\}\Big)^c \right|.$$

Then, since $m_1 \leq m_2$, we have

$$\cup_{k\geq 1}\{I^k(0) : u_k \leq m_1\} \subset \cup_{k\geq 1}\{I^k(0) : u_k \leq m_2\},$$

and we deduce that, almost surely,

$$X^\delta_\infty(m_1) - X_\infty(m_1) \geq X^\delta_\infty(m_2) - X_\infty(m_2).$$

Taking expectations, we obtain, for $m_1 \leq m_2$,

$$\overline{\mu}_\infty(m_1+\delta) - \overline{\mu}_\infty(m_1) \geq \overline{\mu}_\infty(m_2+\delta) - \overline{\mu}_\infty(m_2).$$

This shows that $m \mapsto \overline{\mu}_\infty(m)$ is concave. To show that it is strictly concave, it is sufficient to show that the above inequality is strict with positive probability for any $\delta > 0$, which is obvious from our construction. This concludes the proof of the lemma. □

7.3 The basic reproduction number R_0

Since the non-linear process solves (9) with $m(t) = \mathbb{E}[X_t(m)]$, if it admits a stationary distribution, we expect that it should be of the form $\mu_\infty(m)$ with m satisfying

$$\overline{\mu}_\infty(m) = m. \tag{15}$$

We note that $m = 0$ is always a solution to (15), but, given Lemma 7.3, another solution may exist if

$$\frac{d\overline{\mu}_\infty}{dm}(0) > 1.$$

Lemma 7.4. *Recall the definition of R_0 in* (6)*, then*

$$\frac{d\overline{\mu}_\infty}{dm}(0) = R_0.$$

Corollary 7.5. *If $R_0 \leq 1$, then $m = 0$ is the unique solution to* (15)*. If $R_0 > 1$, then there exists a unique $m_\star \in (0, \overline{\pi}]$ satisfying* (15)*.*

Proof. This is straightforward from Lemma 7.4 and Lemma 7.3 and the inequality $X_\infty(m) \le \nu$. $\square$

Let us now prove Lemma 7.4.

Proof of Lemma 7.4. We prove this result by showing that both terms are equal to the expression given in (8). If we set

$$\mu_\infty^{n,k}(m) = \mathbb{P}(\nu = n, X_\infty(m) = k),$$

then the measure $\mu_\infty(m)$ is characterized by

$$\sum_{k=0}^{n} \mathcal{L}\phi(n,k,\overline{\pi},m)\mu_\infty^{n,k}(m) = 0.$$

Choosing $\phi(n,k) = 1_{\{k\le \ell\}}$ for $0 \le \ell \le n-1$ yields

$$\left(1-\frac{\ell}{n}\right)\left[\lambda_L \ell + \lambda_G \frac{n}{\overline{\pi}} m\right]\mu_\infty^{n,\ell}(m) = \gamma(\ell+1)\mu_\infty^{n,\ell+1}(m).$$

This, together with the obvious condition $\sum_{k=0}^{n}\mu_\infty^{n,k}(m) = \pi(n)$ (see (3)) leads to the following expression

$$\mu_\infty^{n,\ell}(m) = \mu_\infty^{n,0}(m)\frac{1}{\gamma^\ell}\prod_{k=0}^{\ell-1}\left\{\frac{1-\frac{k}{n}}{k+1}\left(\lambda_L k + \lambda_G \frac{n}{\overline{\pi}} m\right)\right\}, \quad 1 \le \ell \le n,$$

$$\mu_\infty^{n,0}(m) = \pi(n)\left(1+\sum_{\ell=1}^{n}\frac{1}{\gamma^\ell}\prod_{k=0}^{\ell-1}\left\{\frac{1-\frac{k}{n}}{k+1}\left(\lambda_L k + \lambda_G \frac{n}{\overline{\pi}} m\right)\right\}\right)^{-1}.$$

From this we deduce easily that

$$\frac{d\overline{\mu}_\infty}{dm}(0) = \frac{\lambda_G}{\gamma}\sum_{n=1}^{\infty}\pi^+(n)\left(1+\sum_{\ell=1}^{n-1}\left(\frac{\lambda_L}{\gamma}\right)^\ell\prod_{j=1}^{\ell}\left(1-\frac{j}{n}\right)\right).$$

We now turn to the quantity R_0 defined in (6). Note that, by the definition of the process $(I(t), t \ge 0)$ in (5),

$$\phi(\nu, I(t)) - \phi(\nu, I(0)) - \int_0^t \mathcal{L}\phi(\nu, I(s), \overline{\pi}, 0)ds \tag{16}$$

is a martingale with respect to the natural filtration of $\{(\nu, I(t)), t \geq 0\}$. Thus if we find a function ϕ such that $\mathcal{L}\phi(n, x, \overline{\pi}, 0) = x$, we will have

$$\mathbb{E}\left[\int_0^T I(s)ds \middle| I(0) = 1, \nu = n\right] = n\{\phi(n, 0) - \phi(n, 1)\}, \tag{17}$$

where $T = \inf\{t \geq 0 : I(t) = 0\}$ (to obtain this, take the expectation of (16) at time $t \wedge T$ and let $t \to \infty$, using monotone convergence in the integral and dominated convergence in the other term). Setting $\psi(n, x) = \gamma(\phi(n, x-1) - \phi(n, x))$, $\mathcal{L}\phi(n, x, \overline{\pi}, 0) = x$ translates into

$$\begin{cases} \psi(n, x) = 1 + \frac{\lambda_L}{\gamma}\left(1 - \frac{x}{n}\right)\psi(n, x+1), \ 1 \leq x \leq n-1, \\ \psi(n, n) = 1. \end{cases}$$

We deduce from this that

$$\psi(n, 1) = \gamma(\phi(n, 0) - \phi(n, 1)) = 1 + \sum_{\ell=1}^{n-1}\left(\frac{\lambda_L}{\gamma}\right)^\ell \prod_{j=1}^{\ell}\left(1 - \frac{j}{n}\right).$$

Together with (17), this proves the lemma. □

Let us quickly mention another avenue for proving Lemma 7.4, which makes use of Proposition 7.1. For $\varepsilon > 0$, let us write

$$\{k \geq 1 : u_k \leq \varepsilon\} = \{1 \leq k_1(\varepsilon) < k_2(\varepsilon) < \ldots\}.$$

Then we write

$$X_\infty(\varepsilon) = \left|I^{k_1(\varepsilon)}(0)\right| + \left|\cup_{j \geq 2} I^{k_j(\varepsilon)}(0) \cap I^{k_1(\varepsilon)}(0)^c\right|.$$

Then, noting that $-t_{k_1(\varepsilon)}$ is distributed as an exponential variable with parameter $\lambda_G \nu \varepsilon / \overline{\pi}$, it is possible to see that

$$\mathbb{E}\left[\left|I^{k_1(\varepsilon)}(0)\right| \middle| \nu = n\right] = \varepsilon \lambda_G \frac{n}{\overline{\pi}} \int_0^\infty \mathbb{E}\left[\left|I^1(t_1 + t)\right| \middle| \nu = n\right] dt + o(\varepsilon),$$

and that

$$\mathbb{E}\left[\left|\cup_{j \geq 2} I^{k_j(\varepsilon)}(0) \cap I^{k_1(\varepsilon)}(0)^c\right| \middle| \nu = n\right] = o(\varepsilon).$$

We then finish by noting that $I^1(t_1 + t)$ is distributed as $I(t)$ conditionally on $I(0) = 1$ and that

$$\frac{d\overline{\mu}_\infty}{dm}(0) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \mathbb{E}[X_\infty(\varepsilon)].$$

7.4 Large-time behaviour of the non-linear Markov process

We now prove Theorem 3.3. We split the proof in two parts, first dealing with the case $R_0 \leq 1$ and then with $R_0 > 1$.

Proof of Theorem 3.3, $R_0 \leq 1$. Let $m_0^+(t) = \overline{\pi}$ and set, for $k \geq 0$,

$$m_{k+1}^+(t) = \overline{\mu}_t(m_k^+, \mu_0).$$

Clearly $\mathbb{E}[X(t)] \leq m_0^+(t)$ for all $t \geq 0$. Since $(\mathbb{E}[X(t)], t \geq 0)$ is a fixed point of $m(\cdot) \mapsto \overline{\mu}_\cdot(m, \mu_0)$ and using Lemma 4.3, for every $k \geq 0$,

$$0 \leq \mathbb{E}[X(t)] \leq m_k^+(t). \tag{18}$$

Furthermore, by Lemma 7.2, for all $k \geq 0$,

$$\lim_{t\to\infty} m_k^+(t) = \overline{\mu}_\infty^{\circ k}(\overline{\pi}),$$

where $\overline{\mu}_\infty^{\circ k}(\cdot) = \overline{\mu}_\infty(\overline{\mu}_\infty(\ldots))$ is the k-th iterate of $m \mapsto \overline{\mu}_\infty(m)$. Letting $t \to \infty$ in (18),

$$0 \leq \liminf_{t\to\infty} \mathbb{E}[X(t)] \leq \limsup_{t\to\infty} \mathbb{E}[X(t)] \leq \overline{\mu}_\infty^{\circ k}(\overline{\pi}).$$

But, by Lemma 7.3 and Lemma 7.4, since $R_0 \leq 1$,

$$\overline{\mu}_\infty^{\circ k}(\overline{\pi}) \to 0 \text{ as } k \to \infty.$$

As a result,

$$\lim_{t\to\infty} \mathbb{E}[X(t)] = 0,$$

and the result follows. □

Before proving the result when $R_0 > 1$, we state the following lemma, whose proof we delay until Subsection 7.5.

Lemma 7.6. *Suppose that $R_0 > 1$ and that $\mathbb{E}[X(0)] > 0$, then*

$$\liminf_{t\to\infty} \mathbb{E}[X(t)] > 0.$$

Let us now finish the proof of Theorem 3.3.

Proof of Theorem 3.3, $R_0 > 1$. The strategy of the proof is similar to the case $R_0 \leq 1$, but we now define two functions

$$m_0^+(t) = \overline{\pi}, \qquad\qquad m_0^-(t) = \inf_{s \geq 0} \mathbb{E}[X(s)].$$

Note that by Lemma 7.6, $\lim_{t\to\infty} m_0^-(t) > 0$. As before, we set, for $k \geq 0$,

$$m_{k+1}^+(t) = \overline{\mu}_t(m_k^+, \mu_0), \qquad\qquad m_{k+1}^-(t) = \overline{\mu}_t(m_k^-, \mu_0).$$

By the same argument as before, since $m_0^-(t) \leq \mathbb{E}[X(t)] \leq m_0^+(t)$, we have, for every $k \geq 0$,

$$m_k^-(t) \leq \mathbb{E}[X(t)] \leq m_k^+(t).$$

Using Lemma 7.2 and letting $t \to \infty$, we obtain

$$\overline{\mu}_\infty^{\circ k}(\inf_{t\geq 0} \mathbb{E}[X(t)]) \leq \liminf_{t\to\infty} \mathbb{E}[X(t)] \leq \limsup_{t\to\infty} \mathbb{E}[X(t)] \leq \overline{\mu}_\infty^{\circ k}(\overline{\pi}). \tag{19}$$

But, by Lemma 7.3 and the fact that $R_0 > 1$, we have

$$\lim_{k\to\infty} \overline{\mu}_\infty^{\circ k}(\inf_{t\geq 0} \mathbb{E}[X(t)]) = \lim_{k\to\infty} \overline{\mu}_\infty^{\circ k}(\overline{\pi}) = m_\star,$$

where $m_\star \in (0, \overline{\pi}]$ is defined by Corollary 7.5 (also using Lemma 7.6 and the fact that $\mathbb{E}[X(0)] > 0$). Hence, letting $k \to \infty$ in (19),

$$\mathbb{E}[X(t)] \to m_\star,$$

as $t \to \infty$. Finally by Lemma 7.2, since the non-linear process is the forced process with $m(t) = \mathbb{E}[X(t)]$,

$$(\nu, X(t)) \to \mu_\infty(m_\star),$$

in distribution as $t \to \infty$, and the theorem is proved. □

Note that, without Lemma 7.6, we would not have been able to bound $\mathbb{E}[X(t)]$ from below by anything useful, since $\overline{\mu}_\infty(0) = 0$.

7.5 Branching process minoration

Proof of Lemma 7.6. Since $R_0 > 1$, we can choose $p, q \in \mathbb{Q}$ such that $0 < q < p < 1$ and

$$(1-p)R_0 > 1.$$

Without loss of generality, we can assume that there exist N_0, N_1 in $\mathbb{N}$ such that $p = 1/N_0$ and $q = 1/N_1$. For the rest of this proof, we restrict N to multiples of both N_0 and N_1.

Let Z_t^N denote the process of infections between households:

$$Z_t^N = \sum_{i=1}^{N} \mathbf{1}_{X_i^N \geq 1},$$

where $\{X_i^N(t),\ t \geq 0;\ 1 \leq i \leq N\}$ is the solution of the model (1).

We now define a continuous-time non-Markovian branching process of infections as follows. Start with $Y_0^N = Nq$ infected households, each with a single infected individual, and whose sizes are chosen according to the size-biased distribution π^+. If there are currently k infected households with $x_1, \dots, x_k$ infected individuals, at rate $(1-p)\lambda_G \sum_{i=1}^{k} x_i$, a new household, whose size is chosen according to the size-biased distribution π^+, is added to the process with a single infected individual. Apart from this, each household undergoes a local epidemic with rates λ_L and γ, independently from the others. Then Y_t^N denotes the number of infected households at time $t \geq 0$.

The corresponding discrete time branching process is supercritical, since the expected number of "offspring" of each household is $(1-p)R_0 > 1$. Then from Lemma 2.1 in Doney [5], if $r > 0$ denotes the real number such that

$$\lambda_G(1-p)\sum_{n=1}^{+\infty} \pi^+(n) \int_0^{\infty} e^{-rt}\mathbb{E}_1[I(t)|\nu = n]dt = 1,$$

where $(I(t), t \geq 0)$ is the process defined in (5) and $\mathbb{E}_1$ means that we take the expectation under the initial condition $I(0) = 1$, then

$$\mathbb{E}[Y_t^N] \sim Nae^{rt} \text{ as } t \to \infty, \tag{20}$$

where a is given by the formula

$$a = q\frac{\int_0^{\infty} e^{-rt}L(t)dt}{\frac{\lambda_G}{\overline{\pi}}(1-p)\int_0^{\infty} te^{-rt}\mathbb{E}_1[\nu I(t)]dt},$$

with $L(t) = \mathbb{P}(\mathcal{I} > t)$ and $\mathcal{I}$ denotes the duration of the infection of a local household epidemic starting with one infectious, where the size of the household is chosen according to the size-biased distribution π^+.

Suppose that $Nq \le Z_0^N$ and define

$$T_{N,p} = \inf\left\{ t \ge 0 : \frac{\sum_{i=1}^N \mathbf{1}_{X_i^N(t)\ge 1}}{\sum_{i=1}^N \nu_i} > p \right\}.$$

Then we claim that, on the interval $[0, T_{N,p})$, Z_t^N stochastically dominates Y_t^N (*i.e.* we can defined $(Y_t^N, t \ge 0)$ such that $Y_t^N \le Z_t^N$ for $t \in [0, T_{N,p}))$. To see this, note that $Y_0^N \le Z_0^N$ and that, since each household in Y_t^N starts with a single infected individual, the number of infected individuals in each household is larger in Z_0^N than in Y_0^N. This stays true until the first time at which a *new* household is infected in either process, since the local infection parameters are the same in both processes, and in Z^N, there are additional infections due to global infections between already infected households. Furthermore, in the process $(Z_t^N, t \ge 0)$, a *new* household is infected at rate

$$\lambda_G \frac{1}{N} \sum_{j=1}^N X_j^N(t) \sum_{i=1}^N \frac{\nu_i}{\overline{\nu}^N} \mathbf{1}_{X_i^N(t)=0} = \lambda_G \left(1 - \frac{\sum_{i=1}^N \nu_i \mathbf{1}_{X_i^N(t)\ge 1}}{\sum_{i=1}^N \nu_i} \right) \sum_{j=1}^N X_j^N(t).$$

and for $t \in [0, T_{N,p})$, this rate is larger than the rate at which a new household is infected in the process $(Y_t^N, t \ge 0)$. We can thus couple the two processes in such a way that

$$Y_t^N \le Z_t^N, \quad \forall t \in [0, T_{N,p}),$$

almost surely for all $N \ge 1$.

Now, by Theorem 3.2, as $N \to \infty$, for any $T > 0$,

$$\frac{Z_t^N}{N} \to p(t) := \mathbb{P}(X_t \ge 1), \tag{21}$$

and

$$\frac{\sum_{i=1}^N \nu_i \mathbf{1}_{X_i^N(t)\ge 1}}{\sum_{i=1}^N \nu_i} \to \frac{1}{\pi} \mathbb{E}[\nu \mathbf{1}_{X(t)\ge 1}], \tag{22}$$

uniformly on $[0,T]$, in probability. Furthermore, there exists a deterministic function $f : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\frac{Y_t^N}{N} \to f(t),$$

uniformly on $[0,T]$ as $N \to \infty$, in probability. Furthermore, by (20),

$$f(t) \sim ae^{rt} \text{ as } t \to \infty. \tag{23}$$

For any $p' < p$, let

$$T_{p'} = \inf\left\{t \geq 0 : \frac{1}{\overline{\pi}}\mathbb{E}[\nu \mathbf{1}_{X(t)\geq 1}] > p'\right\}.$$

By (22), choosing $T > T_{p'}$, for any $t \leq T_{p'}$, $t < \liminf_{N\to\infty} T_{N,p}$, and consequently for N large enough,

$$\frac{Z_t^N}{N} \geq \frac{Y_t^N}{N}, \ \forall t \leq T_{p'} .$$

Letting $N \to \infty$, we obtain

$$p(t) \geq f(t), \ \forall t \leq T_{p'}.$$

Now define

$$T_b^f = \inf\{t \geq 0 : f(t) > b\}.$$

Then, if $T_b^f < T_{p'}$, $p(T_b^f) \geq b$. If however $T_{p'} \leq T_b^f$, then, by the Cauchy-Schwarz inequality,

$$\mathbb{E}[\nu \mathbf{1}_{X(t)\geq 1}] \leq \sqrt{\mathbb{E}[\nu^2]}\sqrt{p(t)},$$

and thus,

$$p(T_{p'}) \geq (p')^2 \frac{\overline{\pi}}{\overline{\pi}^+}.$$

As a consequence, if for some $t \geq 0$, $p(t) = q$, then $p(t+s)$ reaches $b \wedge (p')^2\overline{\pi}/\overline{\pi}^+$ for some $s \leq T_b^f$. Moreover, by (23), f is uniformly bounded away from 0. This proves the Lemma. □

Acknowledgement

The authors wish to thank Frank Ball for pointing out to them that their first version of this work was not consistent with the classical household models studied in the literature.

References

[1] F. Ball, D. Mollison, G. Scalia–Tomba, Epidemics with two level of mixing, *The Annals of Applied Probability* **7**, 46–89, 1997.

[2] F. Ball, D. Sirl, Stochastic SIR epidemics in structured populations, in *Stochastic Epidemic Models with Inference*, T. Britton and E. Pardoux eds., Springer, to appear.

[3] M. S. Bartlett, Measles periodicity and community size. *J. Roy. Statist. Soc. Ser. A* **120** 48–70, 1957.

[4] D. J. Daley, Some aspects of Markov chains in queueing theory and epidemiology, Ph.D. thesis, Cambridge Univ. 1967.

[5] R.A. Doney, A limit theorem for a class of supercritical branching processes *J. Appl. Prob.* **9**, 707–724, 1972.

[6] C. Léonard, Some epidemic systems are long range interacting particle systems, in *Stochastic Processes in Epidemic Theory*, J.P. Gabriel et al. eds., Springer Verlag, 1990.

[7] H. P. McKean, A class of Markov processes associated with nonlinear parabolic equations, *Proc. Nat. Acad. Sci.* **56**, 1907–1911, 1966.

[8] E. Pardoux, *Probabilistic models of population evolution. Scaling limits and interactions*, Springer, 2016

[9] S. Rushton and A.J. Mautner, The deterministic model of a simple epidemic for more than one community. *Biometrika* **42**, 126–132, 1955.

[10] A.S. Sznitman, Topics in propagation of chaos, in *Ecole d'été de Probabilités de Saint-Flour XIX* 1989, 165–261, Lecture Notes in Math. **1464**, Springer, Berlin, 1991.