EPIDEMIC MODELS WITH VARYING INFECTIVITY

RAPHAËL FORIEN*, GUODONG PANG†, AND ÉTIENNE PARDOUX ‡

Abstract. We introduce an epidemic model with varying infectivity and general exposed and infectious periods, where the infectivity of each individual is a random function of the elapsed time since infection, those function being i.i.d. for the various individuals in the population. This approach models infection-age dependent infectivity, and extends the classical SIR and SEIR models. We focus on the infectivity process (total force of infection at each time), and prove a functional law of large number (FLLN). In the deterministic limit of this FLLN, the evolution of the mean infectivity and of the proportion of susceptible individuals are determined by a two-dimensional deterministic integral equation. From its solutions, we then obtain expressions for the evolution of the proportions of exposed, infectious and recovered individuals. For the early phase, we study the stochastic model directly by using an approximate (non–Markovian) branching process, and show that the epidemic grows at an exponential rate on the event of non-extinction, which matches the rate of growth derived from the deterministic linearized equations. We also use these equations to derive the expression for the basic reproduction number R_0 during the early stage of an epidemic, in terms of the average individual infectivity function and the exponential rate of growth of the epidemic, and apply our results to the Covid–19 epidemic.

Key words. epidemic model, varying infectivity, infection-age dependent infectivity, deterministic integral equations, early phase of an epidemic, basic reproduction number R_0, Poisson random measure

AMS subject classifications. 92D30, 60F17, 60J80

1. Introduction.

Most of the literature on epidemic models is based upon ODE models which assume that the length of time during which a given individual is infectious follows an exponential distribution. More precisely, those deterministic models are law of large numbers limits, as the size of the population tends to infinity, of stochastic models where all transitions from one compartment to the next have exponential distributions, see [6] for a recent account. However, it is largely recognized that for most diseases, the durations of the exposed and infectious periods are far from following an exponential distribution. In the case of influenza, a deterministic duration would probably be a better approximation. Recently in [21], the last two authors of the present paper have obtained the functional law of large numbers (FLLN) limits for SIS, SIR, SEIR and SIRS models where in the stochastic model the duration of the stay in the I compartment (resp. both in the E and the I, resp. both in the I and the R compartments) follow a very arbitrary distribution. Of course, in this case the stochastic model is not a Markov model, which makes some of the proofs more delicate. Indeed, the fluctuating part of a Markov process is a martingale, and many tools exist to study tightness and limits of martingales, which are missing in the non–Markovian setting. Nevertheless, we were able in [21] to use *ad hoc* techniques in order to circumvent that difficulty, and we proved not only FLLNs, but also functional central limit theorems (FCLTs). While the classical "Markovian" deterministic models are ODEs, our more general and more realistic "non–Markovian" deterministic models are Volterra type integral equations of the same dimension as the classical ODE models, i.e., equations with memory. Recently in [11], the authors used the approach in [21] to

*INRAE, Centre INRAE PACA, Domaine St-Paul - Site Agroparc 84914 Avignon Cedex FRANCE (raphael.forien@inrae.fr).

†The Harold and Inge Marcus Department of Industrial and Manufacturing Engineering, College of Engineering, Pennsylvania State University, University Park, PA 16802 USA (gup3@psu.edu).

‡Aix–Marseille Université, CNRS, Centrale Marseille, I2M, UMR 7373 13453 Marseille, France (etienne.pardoux@univ.amu.fr)

This manuscript is for review purposes only.

describe the Covid-19 epidemic in France. The flexibility of the choice for the law of the infectious period was very helpful in order to write a realistic model with very few compartments, and our model follows better the data than Markov models.

The aim of the present paper is to go a step further in the direction of realistic models of epidemics, and to consider the case where the infectivity of infectious individuals depends upon their time since infection. It has been established in [14] that in the case of the Covid-19 disease, the infectivity of infectious individuals decreases after symptom onset. In fact it is believed that in most infectious diseases, the infectivity of infectious individuals depends upon the time since infection. This was already argued almost a century ago by Kermack and McKendrick, two of the founders of epidemic modeling in [17]. In that paper, the authors assume both an infection age infectivity, and an infection age recovery rate. The latter can be thought of as the hazard function of the duration of the infectious period, which then has a general absolutely continuous distribution. Like in the present paper, their model is a Volterra integral equation. The same deterministic model has also been described as an "age of infection epidemic model" in [4] and in the recent book [5, Chapter 4.5]. See also two recent papers in the study of Covid-19 pandemic [13, 12], which use a transport PDE model (it is worth noting that PDEs have been commonly used to capture the effect of age of infection in the epidemic literature, see, e.g., [15, 25, 16, 20]). The novelty of the present paper is that we prove that our integral equation deterministic model is the law of large numbers limit of a well specified individual based stochastic model.

The most realistic assumption is probably that this infectivity first increases continuously from 0, and then decreases back to 0. We shall however allow jumps in the random infectivity function, in order in particular to include the classical case of a constant infectivity during the infectious period. We also want to allow a very arbitrary law for the infectious (or exposed/infectious) period(s), as was done in [21]. In this work again, the FLLN limiting deterministic model is a Volterra type integral equation, which is of the same dimension as the corresponding classical ODE model, see Theorem 2.7. We treat only the case of SIR and SEIR models (see also Remark 2.10 on the SIS and SIRS models), but we intend to extend in later publications our approach to other types of models, including models with age classes and spatial distribution, see already [22] for multi–patch models with general exposed and infectious durations. We have also established in a separate publication the FCLT associated to the FLLN established in the present paper, see [23].

Our approach in this paper is to assume that in the original stochastic finite population model, the infectivity of each individual is a random function of the time elapsed since his/her infection, those functions associated to various individuals being independent and identically distributed (i.i.d.). The total force of infection at each time is the aggregate infectivity of all the individuals that are currently infectious. We assume that the infectivity random functions are piecewise continuous with a finite number of discontinuities, which includes all the commonly seen examples, in particular, constant infectivity over a given time interval as a special case. They are also allowed to start with a value zero for a period of time to generalize the SEIR model. These random functions then determine the durations of the exposed and infectious periods, and therefore, their corresponding probability distributions, which can be very general.

Under the i.i.d. assumptions of these infectivity random functions of the various individuals, we prove a FLLN for the infectivity process, together with the counting processes for the susceptible, exposed, infectious and recovered individuals. The mean infectivity and the proportion of susceptible individuals in the limit are uniquely

determined by a two-dimensional Volterra integral equation. Given these two functions, the proportions of exposed, infectious and recovered individuals in the limit are expressed in terms of the two above quantities. They generalize the integral equations in the standard SIR/SEIR models with general exposed and infectious periods in [21]. Our proofs are based upon Poisson random measures associated with the infectivity process, which help us to establish tightness and convergence. This paper further develops the techniques in [21], since for establishing the mean infectivity equation, we cannot integrate by parts as was done in [21]. See below Lemmas 4.4 and 4.5, which give a key argument for the proof of Lemma 4.6.

Our limiting integral equations can be easily solved numerically. For the standard SIR/SEIR model with general exposed and infectious periods, the integral equations are implemented to estimate the state of the Covid-19 pandemic in France in [11]. In another recent work, Fodor et al. [10] argue that integral equations (in the case of deterministic infectious periods) should be used instead of ODEs since the latter may significantly underestimate the initial basic reproduction number R_0. We claim that our model may be used to better predict the trajectory of the epidemic, especially at the beginning of the epidemic and when certain control measures like lockdown and reopening are implemented.

We also study the early phase of the epidemic, during which the proportion of susceptible individuals remains close to 1, which allows to linearize the system of equations. However, typically the epidemic starts with a very small number of infected individuals, so that we need to go back to the stochastic model if we want to describe that early phase. Thanks to a comparison with (non–Markov) branching processes, we are able to show that, conditioned upon non-extinction, the epidemic grows at an exponential rate ρ, reaching a given proportion of infected individuals in the population after a length of time of the order of $\rho^{-1}\log(N)$, if N is the total population size. After that time, we can follow the linearized deterministic model, whose rate of growth is the same ρ.

The rate ρ is easily estimated from the data (if d denote the "doubling time", i.e., the number of days necessary for the number of cases to double, $\rho = d^{-1}\log(2)$). It is then interesting to express the basic reproduction number R_0 in terms of ρ and of the average infectivity function, a formula which we deduce from the linearized Volterra equation, as was already done by [26], see their formula (2.7). We compute explicitly the value of R_0 for different values of two unknown parameters for the case of the early phase of the Covid–19 epidemic in France, assuming a decrease of the infectivity compatible with the results in [14]. We see that the decrease of the infectivity with infection–age induces a decrease of R_0.

The paper is organized as follows. In Section 2.1, we formulate our stochastic model, and make precise all the assumptions. In Section 2.2, we state the FLLN, Theorem 2.7. Section 2.3 is devoted to the early phase of the epidemic: we state Theorem 2.11 which describes the behavior of the stochastic model, and Theorem 2.13, which describes the behavior of the deterministic linearized model. In Section 2.4, we express R_0 in terms of the exponential growth rate and the mean infectivity function, and in Section 2.5 we apply our techniques to the French Covid–19 epidemic during 2020. Section 3 is devoted to the proofs of Theorem 2.11 and Theorem 2.13, and Section 4 to the proof of Theorem 2.7.

2. Model and Results.

2.1. Model description. All random variables and processes are defined in a common complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We consider a generalized SEIR

epidemic model where each infectious individual has an infectivity that is randomly varying with the time elapsed since infection. As usual, the population consists of four groups of individuals, susceptible, exposed, infectious and recovered. Let N be the population size, and $S^N(t), E^N(t), I^N(t), R^N(t)$ denote the sizes of the four groups, respectively. We have the balance equation $N = S^N(t) + E^N(t) + I^N(t) + R^N(t)$ for $t \geq 0$. Assume that $R^N(0) = 0$, $S^N(0) > 0$ and $E^N(0) + I^N(0) > 0$ such that $S^N(0) + E^N(0) + I^N(0) = N$. Let $A^N(t)$ be the cumulative number of individuals that become infected in $(0, t]$ for $t \geq 0$ and denote the associated event times by τ_i^N, $i = 1, \ldots, A^N(t)$.

Note that an infected individual is either exposed or infectious. More precisely, he/she is first exposed, then infectious. Let us first consider those individuals who are infected after time 0 (i.e. they are in the S compartment at time 0). The i–th infected individual is infected at time τ_i^N. He/she is first exposed during the time interval $[\tau_i^N, \tau_i^N + \zeta_i)$. Then he/she is infectious during the time interval $(\tau_i^N + \zeta_i, \tau_i^N + \zeta_i + \eta_i)$, and finally removed on the time interval $[\tau_i^N + \zeta_i + \eta_i, +\infty)$. To this individual is attached an infectivity process $\{\lambda_i(t) : t \geq 0\}$, which is a random right–continuous function such that

$$\lambda_i(t) \begin{cases} = 0, & \text{if } 0 \leq t < \zeta_i, \\ > 0, & \text{if } \zeta_i < t < \zeta_i + \eta_i, \\ = 0, & \text{if } t \geq \zeta_i + \eta_i. \end{cases} \tag{2.1}$$

We shall formulate some assumptions on the functions λ_i below. Let us just say for now that the collection of the functions $\{\lambda_i(\cdot)\}_{i \geq 1}$ are i.i.d. Since

$$\zeta_i = \inf\{t > 0,\ \lambda_i(t) > 0\}, \quad \text{and } \zeta_i + \eta_i = \inf\{t > 0,\ \lambda_i(r) = 0,\ \forall r \geq t\}, \tag{2.2}$$

the collection of random vectors $(\zeta_i, \eta_i)_{i \geq 1}$ is also i.i.d.

Each initially exposed individual is associated with an infectivity process $\lambda_j^0(t)$, $j = 1, \ldots, E^N(0)$, with a càdlàg path; the λ_j^0's are assumed to be i.i.d. and such that

$$\zeta_j^0 = \inf\{t > 0,\ \lambda_j^0(t) > 0\} > 0 \text{ a.s.} \quad \text{and } \zeta_j^0 + \eta_j^0 = \inf\{t > 0,\ \lambda_j^0(r) = 0,\ \forall r \geq t\}. \tag{2.3}$$

Each initially infectious individual is associated with an infectivity process $\lambda_k^{0,I}(t)$, $k = 1, \ldots, I^N(0)$, with a càdlàg path; the $\lambda_k^{0,I}$'s are also assumed to be i.i.d. and such that

$$\inf\{t > 0,\ \lambda_k^{0,I}(t) > 0\} = 0 \text{ a.s.} \quad \text{and } \eta_k^{0,I} = \inf\{t > 0,\ \lambda_k^{0,I}(r) = 0,\ \forall r \geq t\}. \tag{2.4}$$

We will write (ζ, η) (resp. (ζ^0, η^0), resp. $\eta^{0,I}$) for a vector which has the same law as (ζ_i, η_i) (resp. (ζ_j^0, η_j^0), resp. $\eta_k^{0,I}$). Let $H(du, dv)$ denote the law of (ζ, η), $H_0(du, dv)$ that of (ζ^0, η^0) and $F_{0,I}$ the c.d.f. of $\eta^{0,I}$. Moreover, we define

$$\Phi(t) := \int_0^t \int_0^{t-u} H(du, dv) = \mathbb{P}(\zeta + \eta \leq t),$$

$$\Psi(t) := \int_0^t \int_{t-u}^{\infty} H(du, dv) = \mathbb{P}(\zeta \leq t < \zeta + \eta),$$

$$\Phi_0(t) := \int_0^t \int_0^{t-u} H_0(du, dv) = \mathbb{P}(\zeta^0 + \eta^0 \leq t),$$

$$\Psi_0(t) := \int_0^t \int_{t-u}^{\infty} H_0(du, dv) = \mathbb{P}(\zeta^0 \leq t < \zeta^0 + \eta^0),$$

and $F_{0,I}(t) := \mathbb{P}(\eta^{0,I} \le t)$. We shall also write

$$H(du, dv) = G(du)F(dv|u), \quad H_0(du, dv) = G_0(du)F_0(dv|u),$$

i.e., G is the c.d.f. of ζ and $F(\cdot|u)$ is the conditional law of η, given that $\zeta = u$, G_0 is the c.d.f. of ζ^0 and $F_0(\cdot|u)$ is the conditional law of η^0, given that $\zeta^0 = u$. In the case of independent exposed and infectious periods, it is reasonable that the infectious periods of the initially exposed individuals have the same distribution as the newly exposed ones, that is, $F_0 = F$. Note that $\Psi(t) = G(t) - \Phi(t)$ and $\Psi_0(t) = G_0(t) - \Phi_0(t)$. Also, let $G_0^c = 1 - G_0$, $G^c = 1 - G$, $F_{0,I}^c = 1 - F_{0,I}$, and $F^c = 1 - F$.

We remark that our framework allows very general random infectivity functions $\lambda(t)$, which can be piecewise continuous (see Assumption 2.1) and can also generate dependent and independent ζ and η variables for each individual. We give an example of independent ζ and η variables. Let ζ, η and h be random objects so that ζ is independent of the pair (η, h), where ζ and η are $\mathbb{R}_+$ valued and h is a random element of $C([0,1];\mathbb{R}_+)$ satisfying $h(0) = h(1) = 0$ and $h(t) > 0$ for $0 < t < 1$, a.s. (η and h can be dependent). We extend h as an element of $C(\mathbb{R};\mathbb{R}_+)$ by specifying that $h(t) = 0$ if $t \notin [0,1]$. Define $\lambda(t) = h(\zeta\eta^{-1}(\zeta^{-1}t - 1))$ for any $t \ge 0$. Then $\lambda(t) = 0$ on $[0, \zeta]$, and again on $[\zeta + \eta, +\infty)$, where $\lambda(t) > 0$ if $\zeta < t < \zeta + \eta$. By construction, ζ and η are independent.

The total force of infection which is exerted on the susceptibles at time t can be written as

$$\mathfrak{I}^N(t) = \sum_{j=1}^{E^N(0)} \lambda_j^0(t) + \sum_{k=1}^{I^N(0)} \lambda_k^{0,I}(t) + \sum_{i=1}^{A^N(t)} \lambda_i(t - \tau_i^N), \quad t \ge 0. \tag{2.5}$$

Thus, the instantaneous infectivity rate function at time t is

$$\Upsilon^N(t) = \frac{S^N(t)}{N}\mathfrak{I}^N(t), \quad t \ge 0. \tag{2.6}$$

The infection process $A^N(t)$ can be expressed by

$$A^N(t) = \int_0^t \int_0^\infty \mathbf{1}_{u \le \Upsilon^N(s^-)} Q(ds, du), \quad t \ge 0, \tag{2.7}$$

where Q is a standard Poisson random measure (PRM) on $\mathbb{R}_+^2$, and we use $\mathbf{1}_{\{\cdot\}}$ for the indicator function. One may observe that besides the PRM Q, the randomness in the epidemic dynamics comes only from the infectivity processes $\{\lambda_j^0(t)\}$, $\{\lambda_k^{0,I}(t)\}$ and $\{\lambda_i(t)\}$ (the infectious periods $\{\eta_j^0\}$, $(\eta_k^{0,I})$ and $\{\eta_i\}$ are induced from them).

The epidemic dynamics of the model can be described by

$$S^N(t) = S^N(0) - A^N(t), \tag{2.8}$$

$$E^N(t) = \sum_{j=1}^{E^N(0)} \mathbf{1}_{\zeta_j^0 > t} + \sum_{i=1}^{A^N(t)} \mathbf{1}_{\tau_i^N + \zeta_i > t}, \tag{2.9}$$

$$I^N(t) = \sum_{j=1}^{E^N(0)} \mathbf{1}_{\zeta_j^0 \le t < \zeta_j^0 + \eta_j^0} + \sum_{k=1}^{I^N(0)} \mathbf{1}_{\eta_k^{0,I} > t} + \sum_{i=1}^{A^N(t)} \mathbf{1}_{\tau_i^N + \zeta_i \le t < \tau_i^N + \zeta_i + \eta_i}, \tag{2.10}$$

$$R^N(t) = \sum_{j=1}^{E^N(0)} \mathbf{1}_{\zeta_j^0 + \eta_j^0 \le t} + \sum_{k=1}^{I^N(0)} \mathbf{1}_{\eta_k^{0,I} \le t} + \sum_{i=1}^{A^N(t)} \mathbf{1}_{\tau_i^N + \zeta_i + \eta_i \le t}. \tag{2.11}$$

In the case where $\zeta_j^0 = 0$ and $\zeta_i = 0$, the model is a generalized SIR model, and $E^N(t) \equiv 0$.

We now make the following assumptions on the infectivity functions and the initial quantities. We first state our assumptions on λ^0, $\lambda^{0,I}$ and λ.

Assumption 2.1. The random functions $\lambda(t)$ (resp. $\lambda^0(t)$ and resp. $\lambda^{0,I}(t)$), of which $\lambda_1(t), \lambda_2(t), \dots$ (resp. $\lambda_1^0(t), \lambda_2^0(t), \dots$ and resp. $\lambda_1^{0,I}(t), \lambda_2^{0,I}(t), \dots$) are i.i.d. copies, satisfy the following assumptions. There exists a constant $\lambda^* < \infty$ such that $\sup_{t\in[0,T]} \max\{\lambda^0(t), \lambda^{0,I}(t), \lambda(t)\} \le \lambda^*$ almost surely, and in addition there exist a given number $k \ge 1$, a random sequence $0 = \xi^0 \le \xi^1 \le \cdots \le \xi^k = \eta$ and random functions $\lambda^j \in C(\mathbb{R}_+; \mathbb{R}_+)$, $1 \le j \le k$ such that

$$\lambda(t) = \sum_{j=1}^{k} \lambda^j(t) \mathbf{1}_{[\xi^{j-1}, \xi^j)}(t)\,. \tag{2.12}$$

We define

$$\varphi_T(r) := \sup_{1\le j\le k} \sup_{0\le s,t\le T, |t-s|\le r} |\lambda^j(t) - \lambda^j(s)|\,.$$

It is clear that for each $T > 0$, φ_T is continuous and $\varphi_T(0) = 0$.

Let $\bar\lambda^0(t) = \mathbb{E}[\lambda^0(t)]$, $\bar\lambda^{0,I}(t) = \mathbb{E}[\lambda^{0,I}(t)]$ and $\bar\lambda(t) = \mathbb{E}[\lambda(t)]$ for $t \ge 0$.

It is clear that $\bar\lambda^0(t), \bar\lambda^{0,I}(t)$ and $\bar\lambda(t)$ are all càdàg, and they are also uniformly bounded by λ^*.

Remark 2.2. We think that $\lambda(t)$ being continuous is a good model of reality. However, the early phase of the function $\lambda(t)$ is not well known, since patients are tested only after symptom onset, and usually (this is the case in particular for the Covid–19) they may have been infectious (i.e., with $\lambda(t) > 0$) prior to that. Consequently we should not exclude the possibility that $\lambda(t)$ jumps to its maximum at time ζ, and the decreases continuously to 0.

Moreover, in order to include the "classical" models where $\lambda(t)$ is first 0 during the exposed period, and then equal to a positive constant during the infectious period, as well as possible models of infectivity that would be piecewise constant, we allow $\lambda(t)$ to have a given number of jumps.

For one of our results, we shall need the following assumption.

Assumption 2.3. Assume that

$$\mathbb{E}\left[\left(\int_0^\infty \lambda(t)dt\right)^2\right] < \infty, \qquad \mathbb{E}\left[\left(\int_0^\infty \lambda^0(t)dt\right)^2\right] < \infty.$$

Remark 2.4. The assumption on the second moment of $\int_0^\infty \lambda(t)dt$ will be necessary in order to apply Theorem 3.2 from [9] to the branching process approximation of the stochastic model for the early phase of the epidemic. Since we assume that $\lambda(t) \le \lambda^*$, for this second moment condition to be satisfied, it is sufficient that the duration of the infectious period η satisfies $\mathbb{E}[\eta^2] < \infty$, which certainly is not a serious restriction in practice. In our application to the Covid–19 in Section 2.5, we choose a law with compact support for η.

Let $\bar X^N := N^{-1} X^N$ for any process X^N. Let $D = D(\mathbb{R}_+; \mathbb{R})$ denote the space of $\mathbb{R}$–valued càdlàg functions defined on $\mathbb{R}_+$. Throughout the paper, convergence in D means convergence in the Skorohod J_1 topology, see Chapter 3 of [3]. Also, D^k stands for the k-fold product equipped with the product topology.

Assumption 2.5. Assume that there exist deterministic constants $\bar{E}(0), \bar{I}(0) \in [0,1]$ such that $0 < \bar{E}(0) + \bar{I}(0) < 1$, and $(\bar{E}^N(0), \bar{I}^N(0)) \to (\bar{E}(0), \bar{I}(0)) \in \mathbb{R}^2_+$ in probability as $N \to \infty$.

Finally we make the following independence assumption.

Assumption 2.6. Assume that the triple $(\lambda_i(\cdot), i \geq 1;\ \lambda^0_j(\cdot), j \geq 1;\ \lambda^{0,I}_k(\cdot), k \geq 1)$, $(E^N(0), I^N(0))$ and Q (the PRM upon which the construction of the process $A^N(\cdot)$ is based) are independent.

2.2. FLLN. We now state the main result of this paper.

THEOREM 2.7. *Under Assumptions 2.1, 2.5 and 2.6,*

$$\left(\bar{S}^N, \bar{\mathfrak{I}}^N, \bar{E}^N, \bar{I}^N, \bar{R}^N\right) \to \left(\bar{S}, \bar{\mathfrak{I}}, \bar{E}, \bar{I}, \bar{R}\right) \quad in \quad D^5 \quad as \quad N \to \infty, \tag{2.13}$$

in probability, locally uniformly in t. The limits $\bar{S}$ and $\bar{\mathfrak{I}}(t)$ are the unique solution of the following system of Volterra integral equations

$$\bar{S}(t) = 1 - \bar{E}(0) - \bar{I}(0) - \int_0^t \bar{S}(s)\bar{\mathfrak{I}}(s)ds\,, \tag{2.14}$$

$$\bar{\mathfrak{I}}(t) = \bar{E}(0)\bar{\lambda}^0(t) + \bar{I}(0)\bar{\lambda}^{0,I}(t) + \int_0^t \bar{\lambda}(t-s)\bar{S}(s)\bar{\mathfrak{I}}(s)ds\,, \tag{2.15}$$

and the limit $(\bar{E}, \bar{I}, \bar{R})$ is given by the following integral equations:

$$\bar{E}(t) = \bar{E}(0)G^c_0(t) + \int_0^t G^c(t-s)\bar{S}(s)\bar{\mathfrak{I}}(s)ds\,, \tag{2.16}$$

$$\bar{I}(t) = \bar{I}(0)F^c_{0,I}(t) + \bar{E}(0)\Psi_0(t) + \int_0^t \Psi(t-s)\bar{S}(s)\bar{\mathfrak{I}}(s)ds\,, \tag{2.17}$$

$$\bar{R}(t) = \bar{I}(0)F_{0,I}(t) + \bar{E}(0)\Phi_0(t) + \int_0^t \Phi(t-s)\bar{S}(s)\bar{\mathfrak{I}}(s)ds\,. \tag{2.18}$$

The limit $\bar{S}$ is in C, and the limits $\bar{\mathfrak{I}}, \bar{E}, \bar{I}, \bar{R}$ are in D. If $\bar{\lambda}^0$ and $\bar{\lambda}^{0,I}$ are continuous, then $\bar{\mathfrak{I}}$ is in C, and if G_0 and $F_{0,I}$ are continuous, then $\bar{E}, \bar{I}, \bar{R}$ are in C.

Remark 2.8. If we suppose only that Assumptions 2.5 and 2.6 are valid, and $\sup_{t\in[0,T]} \max\{\lambda^0(t),$
$\lambda^{0,I}(t), \lambda(t)\} \leq \lambda^*$ almost surely, then Theorem 2.7 remains valid, but with the convergence in probability in D^5 being replaced by the convergence in probability in $L^p_{loc}(\mathbb{R}_+; \mathbb{R}^5)$, for any $p \geq 1$.

The SEIR/SIR model. Suppose now we do not want to follow the disease progression in the detail adopted so far. Rather, we merge the compartments E (exposed) and I (infectious) into a single compartment I, where now I stands for infected, whether exposed or infectious. Doing this, we do not modify at all our model. Each newly infected individual belongs to the I compartment from the time of infection τ^N_i until the end of the infectious period $\tau^N_i + \zeta_i + \eta_i$, where again $\zeta_i + \eta_i = \inf\{t > 0,\ \lambda_i(r) = 0,\ \forall r \geq t\}$. Of course, between time τ^N_i and time $\tau^N_i + \zeta_i$, $\lambda_i(t) = 0$ (recall that $\zeta_i = \inf\{t,\ \lambda_i(t) > 0\}$), so that he/she is not infectious, but exposed. Likewise, each initially infected individual belongs to the I compartment from time 0 up to time $\zeta^0_j + \eta^0_j$, where $\zeta^0_j + \eta^0_j = \inf\{t \geq 0 : \lambda^0_j(r) = 0, \forall r \geq t\}$. Note

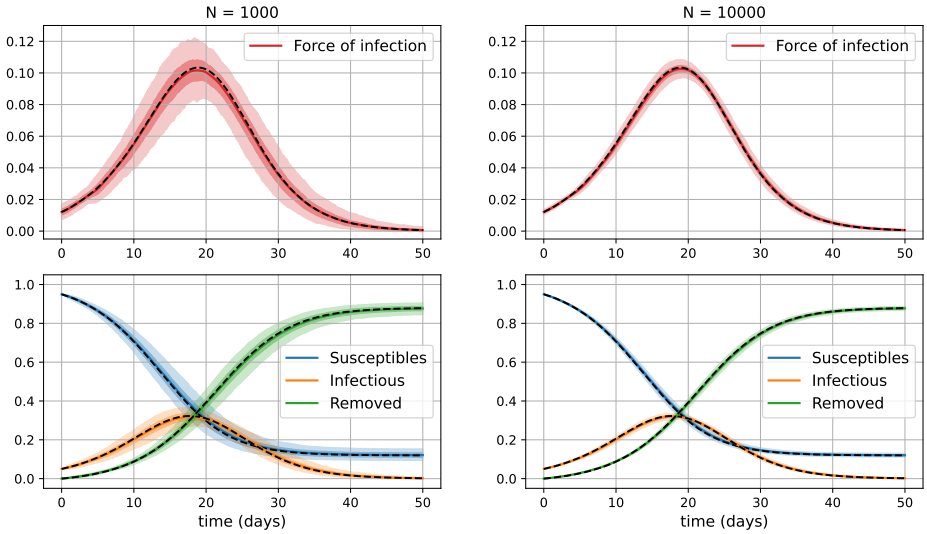

FIG. 1. *Numerical illustration of the FLLN obtained in Theorem 2.7 for the SEIR/SIR model (see below). Each graphic shows the mean of 1,000 independent simulations of the stochastic SEIR/SIR model (continuous lines) and the corresponding deterministic solution to (2.14)-(2.18) (black dashed lines), each started with* $\overline{I}^N(0) = \overline{I}(0) = 0.05$. *For each curve, the dark (resp. light) shaded areas around the curves represent the intervals containing 50% (resp. 95%) of the simulations. The two compartments E and I have been merged so as not to burden the graphic with another pair of curves (see below). The population size* $N = 10^3$ *on the left,* $N = 10^4$ *on the right. The model and the distribution of* (ζ, η, λ) *are as described in Subsection 2.5 below, with* $p_R = 0.8$, $\alpha = 0.7$.

that $\zeta_j^0 = 0$ if $\lambda_j^0(0) > 0$ (if the individual is already infectious at time 0). As a result, (2.9) and (2.10) are replaced by

$$I^N(t) = \sum_{k=1}^{I^N(0)} \mathbf{1}_{t<\zeta_k^0+\eta_k^0} + \sum_{i=1}^{A^N(t)} \mathbf{1}_{t<\tau_i^N+\zeta_i+\eta_i}, \tag{2.19}$$

and $E^N(t) = 0$ in all the other equations. The force of infection is then

$$\mathfrak{I}^N(t) = \sum_{k=1}^{I^N(0)} \lambda_k^0(t) + \sum_{i=1}^{A^N(t)} \lambda_i(t - \tau_i^N). \tag{2.20}$$

We call this model the SEIR/SIR model, since it is an SIR model, but with I meaning "infected", and the state E is implicit, i.e. we do not exclude that individuals, when they become infected, are first exposed, then later infectious. Define

$$F(t) = \mathbb{P}(\zeta + \eta \le t), \quad \text{where } \zeta + \eta = \inf\{t > 0,\ \lambda(r) = 0,\ \forall r \ge t\},$$
$$F_0(t) = \mathbb{P}(\zeta^0 + \eta^0 \le t), \quad \text{where } \zeta^0 + \eta^0 = \inf\{t > 0,\ \lambda^0(r) = 0,\ \forall r \ge t\}.$$

With those notations, the deterministic LLN SEIR/SIR model reads as follows.

$$\bar{S}(t) = 1 - \bar{I}(0) - \int_0^t \bar{S}(s)\bar{\mathfrak{I}}(s)ds\,, \tag{2.21}$$

$$\bar{\mathfrak{I}}(t) = \bar{I}(0)\bar{\lambda}^{0,I}(t) + \int_0^t \bar{\lambda}(t-s)\bar{S}(s)\bar{\mathfrak{I}}(s)ds\,, \tag{2.22}$$

$$\bar{I}(t) = \bar{I}(0)F_0^c(t) + \int_0^t F^c(t-s)\bar{S}(s)\bar{\mathfrak{I}}(s)ds\,, \tag{2.23}$$

$$\bar{R}(t) = \bar{I}(0)F_0(t) + \int_0^t F(t-s)\bar{S}(s)\bar{\mathfrak{I}}(s)ds\,. \tag{2.24}$$

Now in the particular case where $\lambda^0(\cdot)$ and $\lambda(\cdot)$ are such that $\zeta = \zeta^0 = 0$ a.s. (i.e., an infected individual is immediately infectious), there is no exposed period, then the above model is the generalized SIR model with varying infectivity.

Figure 1 illustrates the FLLN of Theorem 2.7 for the SEIR/SIR model, for two values of the population size (10^3 and 10^4). Each figure displays the mean of 1,000 independent simulations, the trajectory of the deterministic equations (2.14)-(2.18), and the intervals containing 50% and 95% of the trajectories. The details of the model and the distribution of (ζ, η, λ) used in the simulations are described in Subsection 2.5 below. In each case, the mean of the simulations is almost superposed with the solution to the deterministic equations, and for $N = 10^4$, the envelopes are very concentrated around the means. This is not surprising in view of the FCLT proved in [23]. Indeed, this theorem implies that the trajectory of the (renormalised) stochastic process $(\overline{S}^N(t), \overline{\mathfrak{I}}^N(t), \overline{I}^N(t), \overline{R}^N(t), t \geq 0)$ is (with high probability) at a distance of the order of $N^{-1/2}$ from that of the deterministic limit. The simulations obtained in Figure 1 confirm this, and the width of the 50% and 95% intervals are exactly proportional to $N^{-1/2}$.

Remark 2.9. The above result generalizes both our SIR and our SEIR FLLN results in [21].

The SIR model in [21] is the particular case of the present result, where $\lambda(t) = \lambda \mathbf{1}_{t<\eta}$, η being the random duration of the infectious period. In this case, $\bar{\lambda}(t) = \lambda F^c(t)$, if F is the c.d.f. of η, and $F^c = 1 - F$. Note that in this case $\bar{\mathfrak{I}}(t) = \lambda \bar{I}(t)$. Therefore, if we divide the $\bar{\mathfrak{I}}$ equation by λ, we find equation (2.17), which is also equation (2.4) in [21]. If we assume that the law of η is exponential, then we are in the case of the classical SIR model.

The SEIR model in [21] corresponds to the situation where $\lambda(t) = \lambda \mathbf{1}_{\zeta \leq t < \zeta + \eta}$, where ζ is the duration of the exposed period (the time when the individual is infected, but not yet infectious), and η is as above, while $\lambda^0(t) = \lambda \mathbf{1}_{\zeta^0 \leq t < \zeta^0 + \eta^0}$. Then $\bar{\lambda}(t) = \lambda[\mathbb{P}(\zeta \leq t) - \mathbb{P}(\zeta + \eta \leq t)] = \lambda \Psi(t)$. If we divide the $\bar{\mathfrak{I}}$ equation by λ, we find equation (2.17), which is also (3.15) in [21]. If moreover ζ and η are independent exponential random variables, then we are reduced to the classical SEIR model.

Remark 2.10. For the generalized SIS model, since $\bar{S}(t) = 1 - \bar{I}(t)$, it is clear that the epidemic dynamics in the FLLN is determined by the two–dimensional functions

$\big(\bar{\mathfrak{I}}, \bar{I}\big)$ via the following integral equations:

$$\bar{\mathfrak{I}}(t) = \bar{I}(0)\bar{\lambda}^{0,I}(t) + \int_0^t \bar{\lambda}(t-s)(1-\bar{I}(s))\bar{\mathfrak{I}}(s)ds\,,$$

$$\bar{I}(t) = \bar{I}(0)F^c_{0,I}(t) + \int_0^t F^c(t-s)(1-\bar{I}(s))\bar{\mathfrak{I}}(s)ds\,.$$

Recall that as shown in Theorem 2.3 of [21], in the SIS with general infectious periods, $\bar{\mathfrak{I}}(s) = \lambda\bar{I}(s)$, and the epidemic dynamics is determined by the one–dimensional integral equation for $\bar{I}$.

For the generalized SIRS model, the variables (ζ_i, η_i) in our setup represent the infectious and recovered/immune periods of newly infected individuals, and similarly the variables (ζ^0_j, η^0_j) represent the infectious and immune periods of initially infectious individuals. We assume that there is no initially immune individuals. Let I^N, R^N be the processes counting infectious and recovered/immune individuals (corresponding to the notation E^N and I^N in the SEIR model). Of course, instead of (2.1), the infectivity function $\lambda(t)$ should be positive only in the infectious periods $[0, \zeta_i)$. Similarly, $\lambda^0_j(t)$ should be positive only over $[0, \zeta^0_j)$. The definitions of the variables (ζ_i, η_i), (ζ^0_j, η^0_j) in (2.2) and (2.3) also need to be modified accordingly in the natural way. The distribution functions $G_0, F_{0,R}$ are for initially infectious and immune periods, and G, F for newly infectious and immune periods, similarly for the notation $\Psi, \Psi_0, \Phi, \Phi_0$. Then the epidemic dynamics of the generalized SIRS model in the FLLN is determined by the three–dimensional functions $\big(\bar{\mathfrak{I}}, \bar{I}, \bar{R}\big)$ via the following integral equations:

$$\bar{\mathfrak{I}}(t) = \bar{I}(0)\bar{\lambda}^0(t) + \int_0^t \bar{\lambda}(t-s)\big(1-\bar{I}(s)-\bar{R}(s)\big)\bar{\mathfrak{I}}(s)ds\,,$$

$$\bar{I}(t) = \bar{I}(0)G^c_0(t) + \int_0^t G^c(t-s)\big(1-\bar{I}(s)-\bar{R}(s)\big)\bar{\mathfrak{I}}(s)ds\,,$$

$$\bar{R}(t) = \bar{I}(0)\Psi_0(t) + \int_0^t \Psi(t-s)\big(1-\bar{I}(s)-\bar{R}(s)\big)\bar{\mathfrak{I}}(s)ds\,.$$

Also recall that as shown in Theorem 3.3 of [21], in the SIRS model with general infectious and recovered periods, $\bar{\mathfrak{I}}(s) = \lambda\bar{I}(s)$, and the epidemic dynamics is determined by the two–dimensional integral equation for $\big(\bar{I}, \bar{R}\big)$.

2.3. The early phase of the epidemic. Theorem 2.7 shows that the deterministic system of equations (2.14)-(2.15) accurately describes the evolution of the stochastic process defined in Subsection 2.1 when the initial number of infectious individuals is of the order of N. But epidemics typically start with only a handful of infectious individuals, and it takes some time before the epidemic enters the regime of Theorem 2.7. Exactly how long this takes depends on the population size N and on the growth rate of the epidemic. To determine this growth rate, we study the behavior of the stochastic process when the initial number of infectious individuals is kept fixed as $N \to \infty$.

In order to simplify the notations, we shall use the reduced model introduced in (2.19) and (2.20), where exposed and infectious individuals are merged in a single infected compartment I. We now suppose that $I^N(0) = I(0)$ is a fixed random variable taking values in $\{1, \dots, N_0\}$ for some $N_0 \geq 1$, and we take $N \geq N_0$ throughout this section.

Let

$$R_0 = \int_0^\infty \overline{\lambda}(t)dt, \tag{2.25}$$

and let $\rho \in \mathbb{R}$ be the unique solution of

$$\int_0^\infty \overline{\lambda}(t)e^{-\rho t}dt = 1. \tag{2.26}$$

The quantity R_0 is the well–known basic reproduction number, *i.e.*, the average number of individuals infected by a typical infected individual in a large, fully susceptible population. It is also well known that, if $R_0 \leq 1$, the total number of infections remains small as $N \to \infty$, *i.e.*, $\limsup_{t\to\infty} A^N(t)$ converges in probability as $N \to \infty$ to a random variable Z taking values in $\mathbb{N}$, almost surely, see Corollary 1.2.6 in [6]. If $R_0 > 1$, however, with positive probability, a major outbreak takes place, *i.e.*, a positive fraction of the N individuals is infected at some point during the course of the epidemic. The time needed in order to observe this major outbreak has been studied for Markovian epidemic models in [1]. More precisely, it has been shown that, starting from a fixed number of individuals, on the event that there is a major outbreak, the first time at which the proportion of infected individuals is at least $\varepsilon > 0$ is

$$\frac{1}{\rho}\log(N) + \mathcal{O}(1),$$

as $N \to \infty$, for any $\varepsilon > 0$ small enough, where $\rho > 0$ is given by (2.26) (it can easily be seen that $\rho > 0$ if and only if $R_0 > 1$). The aim of this section is to extend this result to our non–Markovian setting.

We thus let, for $\varepsilon \in (0, 1)$,

$$T_\varepsilon^N := \inf\{t \geq 0 : A^N(t) \geq \varepsilon N\}$$

and, for any $\alpha \in (0, 1)$,

$$\mathcal{T}_\alpha^N := \inf\{t \geq 0 : A^N(t) \geq N^\alpha\}.$$

Here and in what follows, we shall use $X^N \Rightarrow X$ to denote the convergence in distribution of a sequence of random variables $(X^N, N \geq 1)$ to a random variable X as $N \to \infty$, *i.e.*, $X^N \Rightarrow X$ if and only if, for any continuous and bounded real-valued function Φ, $\mathbb{E}\left[\Phi(X^N)\right] \to \mathbb{E}\left[\Phi(X)\right]$ as $N \to \infty$. We then have the following result, which we prove in Section 3.

THEOREM 2.11. *Under Assumptions 2.1 and 2.3, for any $\varepsilon > 0$ such that $\varepsilon < 1 - \frac{1}{R_0}$, as $N \to \infty$,*

$$\frac{T_\varepsilon^N}{\log(N)} \Rightarrow \frac{1}{\rho}X,$$

where $X = +\infty$ with probability q and $X = 1$ otherwise, for some $q \in (0, 1)$. Moreover, for any $\alpha \in (0, 1)$,

$$\frac{\mathcal{T}_\alpha^N}{\log(N)} \Rightarrow \frac{\alpha}{\rho}X.$$

Theorem 2.11 essentially says that, on an event of probability close to $1-q$, $t \mapsto A^N(t)$ grows approximately like (a constant times) $t \mapsto e^{\rho t}$ until it becomes of the order of N. This exponential growth comes from the fact that, as long as $\overline{S}^N(t) \approx 1$, the infected individuals behave almost like a branching process (which in our case is non–Markovian, and is of the type studied in [8, 9]). Since $A^N(t) \approx e^{\rho t}$, this approximation is good as long as $t \ll \frac{1}{\rho}\log(N)$, at which time the proportion of susceptible individuals is no longer close to one, and the branching process approximation breaks down. We shall also see in the proof of Theorem 2.11 that q is equal to the extinction probability of this approximating branching process.

Remark 2.12. The condition $\varepsilon < 1 - \frac{1}{R_0}$ comes from the fact that, as long as $\overline{S}(t) < \frac{1}{R_0}$, each infected individual infects on average more than one susceptible individual. Hence the proportion of susceptible individuals needs to become lower than this threshold for the epidemic to die out (on the event that there is a major outbreak). As a result, $A^N(t)$ has to exceed εN for some time $t < \infty$ for any $\varepsilon < 1 - \frac{1}{R_0}$.

The fact that the number of infected individuals grows exponentially at rate ρ as long as the proportion of susceptible individuals stays close to one can also be seen from the deterministic equations by taking $\overline{S}(t) = 1$ in (2.22) (as well as (2.23) and (2.24)). This substitution leads to the following (linear) system (recall that in this section F is the distribution function of the r.v. $\zeta + \eta$):

$$
\begin{aligned}
\mathfrak{I}(t) &= I(0)\bar{\lambda}^0(t) + \int_0^t \bar{\lambda}(t-s)\mathfrak{I}(s)ds\,,\\
I(t) &= I(0)F_0^c(t) + \int_0^t F^c(t-s)\mathfrak{I}(s)ds\,,\\
R(t) &= R(0) + I(0)F_0(t) + \int_0^t F(t-s)\mathfrak{I}(s)ds\,.
\end{aligned}
\tag{2.27}
$$

We prove the following in Section 3.

THEOREM 2.13. *Assume that Assumption 2.1 holds true. For $\rho \in \mathbb{R}$, suppose that $\mathbb{E}\big[e^{-\rho(\zeta+\eta)}\big] < \infty$ and define*

$$
\boldsymbol{i} := \int_0^\infty F^c(s)\rho e^{-\rho s}ds, \qquad\qquad \boldsymbol{r} := 1 - \boldsymbol{i}, \tag{2.28}
$$

and

$$
\overline{\lambda}_\rho(t) := \frac{\int_0^\infty \overline{\lambda}(t+s)e^{-\rho s}ds}{\int_0^\infty F^c(s)e^{-\rho s}ds}, \qquad\qquad F^c_\rho(t) := \frac{\int_0^\infty F^c(t+s)e^{-\rho s}ds}{\int_0^\infty F^c(s)e^{-\rho s}ds}.
$$

Suppose first that $R_0 > 1$ and that $\rho > 0$ is the solution to (2.26). Then, if $\overline{\lambda}^0 = \overline{\lambda}_\rho$ and $F_0 = F_\rho$, the linear system (2.27) admits the following solution

$$
\mathfrak{I}(t) = \rho\, e^{\rho t}, \quad I(t) = \boldsymbol{i}\, e^{\rho t}, \quad R(t) = \boldsymbol{r}\, e^{\rho t} \quad t \geq 0. \tag{2.29}
$$

If, however, $R_0 < 1$ and $\rho < 0$ (still satisfying (2.26)), then the linear system (2.27) (with $\overline{\lambda}^0 = \overline{\lambda}_\rho$ and $F_0 = F_\rho$) admits the following solution

$$
\mathfrak{I}(t) = -\rho e^{\rho t}, \quad I(t) = -\boldsymbol{i} e^{\rho t}, \quad R(t) = R(0) + \boldsymbol{r}(1 - e^{\rho t}), \quad t \geq 0.
$$

The deterministic system (2.27) can be thought of as an approximation of the expectation of the stochastic process $(\mathfrak{I}^N(t), I^N(t), R^N(t))$ when $\overline{S}^N(t) \approx 1$. Note that if we take the exponentially growing solution (2.29) and if we set

$$A(t) := I(t) + R(t) - (I(0) + R(0))$$

(which corresponds to the number of newly infected individuals up to time t), then, since $\boldsymbol{i} + \boldsymbol{r} = 1$, $A(t) = e^{\rho t} - 1$ and

$$A\left(\frac{\alpha}{\rho}\log(N)\right) = N^\alpha - 1 \sim N^\alpha. \tag{2.30}$$

Hence Theorems 2.11 and 2.13 show that the stochastic model and the linear deterministic system (2.27) have the same asymptotical behavior, on the event that there is a major outbreak, for times of the form $\frac{\alpha}{\rho}\log(N)$, $\alpha \in (0,1)$. This is further illustrated in Figure 2, which displays the mean of a subset 1,000 independent copies of $t \mapsto I(0) + A^N(t)$ for which the epidemic didn't go extinct at the beginning. We see on the figure that, after an initial stochastic phase, whose duration may vary between different realizations, the cumulative number of infected individuals indeed grows at the expected rate ρ. We also see that the slope of $t \mapsto I(0) + A^N(t)$ starts to decline when $A^N(t)$ exceeds $N/10$ (hence when $\bar{S}^N(t)$ becomes less than 0.9), which is to be expected from the deterministic model.

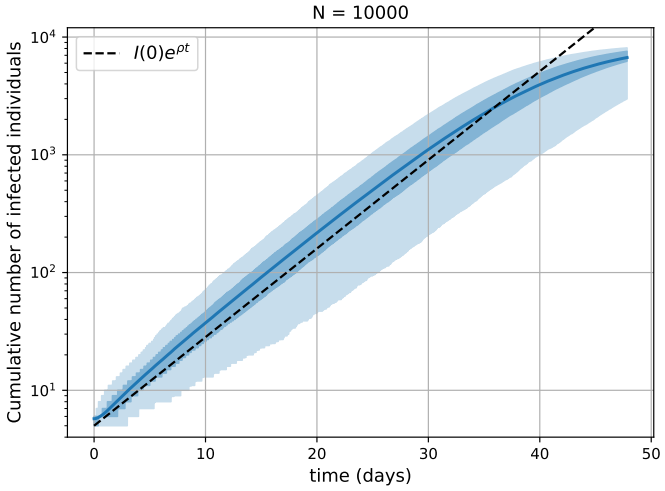

FIG. 2. *Exponential growth of the cumulative number of infected individuals $t \mapsto I(0) + A^N(t)$ in the stochastic model. The figure shows the mean (blue line), 50% envelope (dark blue region) and 95% envelope (light blue region) of the subset of 1,000 independent simulations for which the epidemic did not go extinct at the beginning. Each simulation was started with $I(0) = 5$ infectious individuals and a population size of $N = 10^4$. The dashed black line shows the expected exponential growth during this early phase $t \mapsto I(0)e^{\rho t}$ (the factor $I(0)$ arises from the branching property). The mean of the sample is slightly above the dashed line, owing to the bias resulting from the fact that only trajectories leading to a major outbreak were kept.*

In the case of Markovian (SIR) epidemic models, Theorem 2 of [1] states that the full duration of the epidemic (*i.e.*, the time to extinction of the I population) T_N, when starting from a single infected individual, satisfies

$$\mathbb{P}\left(T_N - a\log(N) - c \geq x\right) \to (1-q)\,\mathbb{P}\left(W \geq x\right), \quad N \to \infty,$$

for some constants $a > 0$ and $c \in \mathbb{R}$, where W is a linear combination of two independent Gumbel random variables. Moreover, $a = \frac{1}{\rho} + \frac{1}{\rho'}$, where ρ is the same as in Theorem 2.11 and ρ' is the rate of decay of the number of infected individuals during the final stage of the epidemic. In addition, Theorem 1.1 in [2] shows that the stochastic process can be coupled with a branching process so that the two follow the same trajectory up to the time $\min(T_0^N, \mathcal{T}_\alpha^N)$, for $\alpha = 7/12$, except on an event of asymptotical negligible probability. Moreover, Theorem 1.1 in [2] also says that, for times of the form $\mathcal{T}_\alpha^N + t$, for $0 \le t \le \frac{1-\alpha}{\rho}\log(N) + T$, the trajectory of the stochastic process is, with high probability, at most at distance $kN^{-\gamma}$ of the trajectory of a solution of the deterministic (non-linear) equations (2.21)–(2.24), whose initial condition is of the form

$$\overline{S}(0) = 1 - \frac{I(0)}{N}, \qquad \overline{I}(0) = \frac{I(0)}{N},$$

up to a time shift which stays of the order of 1 as $N \to \infty$, and which accounts for the stochastic fluctuations when the number of infected individuals is small. We expect that a similar result holds in our non-Markovian setting, but proving this would require a careful comparison of the stochastic model with the deterministic model started from an $\mathcal{O}(1/N)$ initial proportion of infected individuals over timescales of the order of $\log(N)$, and this would go beyond the scope of this paper.

The second part of the statement (when $R_0 < 1$) describes what takes place when the daily number of new infections is decreasing, either because a large fraction of the population has been infected (or vaccinated) or because effective containment measures have been put into place (*e.g.*, a strict lockdown). In the former case, $\overline{S}(t)$ is not close to one, and $\overline{\lambda}$ should be replaced by $\overline{S}(t)\overline{\lambda}$ in order to determine ρ and $\overline{\lambda}_\rho$ (assuming that $\overline{S}(t)$ varies slowly at this point).

Note that if we replace $I(0)$, $R(0)$, $\overline{\lambda}^0$ and F_0 by their values in Theorem 2.13, and if we set, for $t < 0$,

$$\mathfrak{I}(t) = \rho e^{\rho t}, \qquad I(t) = \boldsymbol{i} e^{\rho t}, \qquad R(t) = \boldsymbol{r} e^{\rho t},$$

then we have

$$\mathfrak{I}(t) = \int_{-\infty}^t \overline{\lambda}(t-s)\mathfrak{I}(s)ds, \quad I(t) = \int_{-\infty}^t F^c(t-s)\mathfrak{I}(s)ds,$$
$$R(t) = \int_{-\infty}^t F(t-s)\mathfrak{I}(s)ds.$$

Hence (2.27) can also be interpreted as the (expected) behavior of an epidemic which has started from an infinitesimal number of infected individuals very far back in the past. Incidentally, substituting $\mathfrak{I}(t) = \rho e^{\rho t}$ in the first equation yields exactly (2.26).

2.4. Estimating the basic reproduction number for an ongoing an epidemic. The function $\overline{\lambda}$ (as well as F) depends on many factors. Some of these factors are related to the evolution of the pathogen inside an infected individual's organism, and how easily it can be transmitted to neighboring individuals, and some of these factors depend on the intensity of social contacts in the population, in particular on physical contacts between individuals when they meet (hand shaking, kiss, hug, or none of those). This function is affected by changes in social contacts and collective behaviors, including public policies aimed at mitigating the effects of the epidemic,

and the use of face masks. For example, during the Covid-19 pandemic, many countries implemented strict lockdowns in order to curb the spread of the disease, which drastically reduced the rate of infectious contacts and significantly affected the growth rate of the number of newly infected individuals. In order to estimate the impact of such policies in terms of the dynamics of the epidemic, we thus need to be able to gather some information on the contact rate $\overline{\lambda}$ from the available data at some given time.

Let us suppose that $\overline{\lambda}$ is only known up to a constant factor $\mu > 0$, *i.e.*,

$$\overline{\lambda}(t) = \mu\, \overline{g}(t), \quad t \geq 0,$$

where μ is unknown but $\overline{g}$ is known (for example from medical data on viral shedding). We can then estimate μ (and R_0) from the growth rate ρ, which can be measured easily at the beginning of the epidemic ($\rho = \log(2)/d$, where d is the doubling time of the daily number of newly infected individuals), using the relation (2.26). The following is thus a corollary of Theorem 2.11.

Corollary 2.14. Let ρ be the growth rate of the number of infected individuals. Then

$$\mu = \left(\int_0^\infty \overline{g}(t) e^{-\rho t} ds \right)^{-1},$$

and the basic reproduction number R_0 is given by

$$R_0 = \frac{\int_0^\infty \overline{g}(t) dt}{\int_0^\infty \overline{g}(t) e^{-\rho t} dt}. \tag{2.31}$$

In the literature, $(\int_0^\infty \overline{g}(t)dt)^{-1}\overline{g}(t)$ is called the generation interval distribution (it is the distribution of the interval between the time at which an individual is infected and the time at which its "children" are infected). The relation (2.31) is thus (2.7) in [26]. Note that R_0 is the mean multiplicative factor of the epidemic from one generation to the next, while ρ is a growth factor in continuous time.

Note that, by the second part of Theorem 2.13, (2.31) remains valid on any interval during which $\overline{S}(t) \approx \overline{S}(t_0)$ remains approximately constant (but not necessarily close to 1), even when $\rho \leq 0$. In that case, one should add a factor $\overline{S}(t_0)$ in front of $\mathfrak{I}(s)$ on the right hand sides of (2.27), and we obtain

$$\mu\, \overline{S}(t_0) \int_0^\infty \bar{g}(s) e^{-\rho_e s} ds = 1.$$

Hence if we define the *effective* reproduction number R_e by $R_e := \overline{S}(t_0) \int_0^\infty \overline{\lambda}(t) dt$ (*i.e.*, the average number of secondary infections when $\overline{S}(t) = \overline{S}(t_0)$), we have

$$R_e = \overline{S}(t_0) R_0 = \frac{\int_0^\infty \overline{g}(s) ds}{\int_0^\infty \overline{g}(s) e^{-\rho_e s} ds}.$$

Remark 2.15. Note that the exponent ρ is a quantity which is deduced from the observation of the epidemic (it is closely related to the "doubling time" of the number of cases). The above results give us μ and R_0 in terms of ρ and the function $\bar{g}(t)$. If $\lambda(t)$ is deterministic, so are $g(t)$ and η and thus

$$R_0 = \frac{\int_\zeta^{\zeta+\eta} g(s) ds}{\int_\zeta^{\zeta+\eta} g(s) e^{-\rho s} ds}.$$

If, in addition, $\bar{g}(t) \equiv g > 0$ for $\zeta \leq t < \zeta + \eta$, then this simplifies to the well–known result

$$R_0 = \frac{\rho\eta}{e^{-\rho\zeta}(1 - e^{-\rho\eta})}.$$

Remark 2.16. Theorem 2.13 and its Corollary generalize Proposition 2 and Corollary 3 in [11], in the case $\lambda(t) = \lambda \mathbf{1}_{\zeta \leq t < \zeta+\eta}$ for some constant $\lambda > 0$, and the pair (ζ, η) is an arbitrary $\mathbb{R}^2_+$–valued random vector. In that case, our formula for R_0 reduces to

$$R_0 = \frac{\rho\, \mathbb{E}[\eta]}{\mathbb{E}[e^{-\rho\zeta}(1 - e^{-\rho\eta})]}\,.$$

In the particular case where ζ and η are independent exponential random variables, with parameters ν and γ, the above formula becomes

$$R_0 = \left(1 + \frac{\rho}{\nu}\right)\left(1 + \frac{\rho}{\gamma}\right)\,.$$

From this we deduce the formula in the classical SIR case by choosing $\nu = +\infty$, i.e.,

$$R_0 = 1 + \frac{\rho}{\gamma}.$$

2.5. Application to the Covid–19 epidemic. We now want to explain how the type of model described in this paper can be used to model the Covid–19 epidemic. As we have seen, the increase in realism with respect to the classical "Markovian" models (where the infectivity is constant and fixed across the population, and the Exposed and Infectious periods follow an exponential distribution) is paid by replacing a system of ODEs by a system of Volterra integral equations. However, we have a small benefit in that the flexibility induced by the fact that the law of λ is arbitrary allows us to reduce the number of compartments in the model, so that we can replace a system of ODEs by a system of Volterra type equations of smaller dimension.

To be more specific, let us describe the SEIRU model of [19]. An individual who is infected is first "Exposed" E, then "Infectious" I. Soon after, the infectious individual either develops significant symptoms, and then will be soon "Reported" R, and isolated so that he/she does not infect any more; while the alternative is that this infectious individual is asymptomatic: he/she develops no or very mild symptoms, so remains "Unreported" U, and continues to infect susceptible individuals for a longer period. Both unreported and reported cases eventually enter the "Removed" (Rem.) compartment. In this model, there are 6 compartments: S like susceptible, E like exposed, I like infectious, R like reported, U like unreported, and Rem like removed.

Our approach allows us to have a more realistic version of this model with only 3 compartments (see Figure 3): S like susceptible, I like infected (first exposed, then infectious), R like removed (which includes the Reported individuals, since they do not infect any more, and will recover soon or later). As already explained, we do not need to distinguish between the exposed and infectious, since the function λ is allowed to remain equal to zero during a certain time interval starting from the time of infection. More importantly, since the law of λ is allowed to be bimodal, we can accommodate in the same compartment I individuals who remain infectious for a short duration of time, and others who will remain infectious much longer (but probably with a lower infectivity). Moreover, since we know, see [14], that the infectivity decreases after a maximum which in the case of symptomatic individuals, seems to take place shortly before symptom onset, our varying infectivity model allows us to use a model

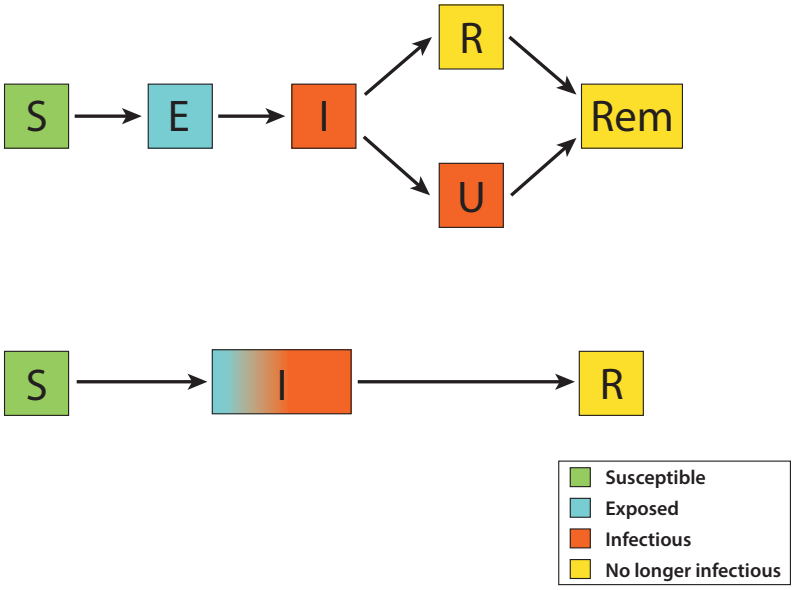

FIG. 3. *Flow chart of the SEIRU model of [19] and of our SIR model. We are able to replace the six compartments of the SEIRU model with only three compartments by using the equations described in Theorem 2.7.*

corresponding to what the medical science tells us about this illness. Note that our version of the SEIRU model from [19] is the same as the one which we have already used in [11] (except that there we had to distinguish the E and the I compartments). However, the main novelty here is that the infectivity decreases after a maximum near the beginning of the infectious period.

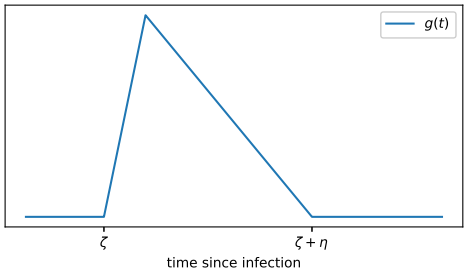

FIG. 4. *Profile of the function $g(t)$ used in our computation of R_0 as a function of ζ and η. The function increases linearly (up to a value 1 or α depending on whether the individual is reported or unreported) on the interval $[\zeta, \zeta + \eta/5]$ and then decreases linearly on $[\zeta + \eta/5, \zeta + \eta]$.*

More precisely, we consider that $t \mapsto g(t)$ increases linearly on the interval $[\zeta, \zeta + \eta/5]$, from 0 to 1 for reported individuals, and from 0 to α for unreported individuals, and that it then decreases linearly to 0 on the interval $[\zeta + \eta/5, \zeta + \eta]$, as shown on Figure 4. We then take (X_1, X_2) a pair of independent Beta random

variables with parameters (2, 2) and we assume that

$$\zeta = 2 + 2X_1, \qquad \eta = \begin{cases} 3 + X_2 & \text{for reported individuals,} \\ 8 + 4X_2 & \text{for unreported individuals.} \end{cases}$$

This joint law of (ζ, η) is the one that was used in [11] to study the Covid–19 epidemic in France (where the infectivity was assumed to be constant and uniform among individuals in this work), and these values are compatible with the results described in [14].

Numerical results are presented in Figure 5 for three growth rates (0.277, -0.06, 0.032) which are derived from the doubling/halving times of the number of hospital deaths during the first wave (doubling time of 2.5 days), the first lockdown (halving time of 11.6 days) and the second wave (doubling time of 21.4 days) of the Covid–19 epidemic in France [11]. We note that, when $\rho > 0$ (resp. when $\rho < 0$), R_0 is increasing (resp. decreasing) with the proportion of unreported individuals and with α. We also note that with the same durations of the exposed and infectious periods, but with $\lambda(t)$ constant, R_0 would be larger, which is not surprising, since in the present model the decrease of $\bar{\lambda}(t)$ reduces the effect of the factor $e^{-\rho t}$ in the integrals in the denominator, which makes $R_0 > 1$ for $\rho > 0$.

3. The early phase of the epidemic. The aim of this section is to prove Theorem 2.11 and Theorem 2.13. In particular, we assume in this section that $\mathbb{E}\left[\left(\int_0^\infty \lambda(t)dt\right)^2\right] < \infty$ and that Assumption 2.1 is satisfied. The first step is to couple the stochastic process $(A^N(t), \mathfrak{I}^N(t), t \geq 0)$ with two branching processes such that, at least up to some stopping time, the stochastic process A^N stays between the two branching processes. To do this, we redefine the model of Subsection 2.1 in the following way. Let $(\lambda_k^0(\cdot), k \geq 1)$ be as before and let Q be a PRM on $\mathbb{R}_+^2 \times D$ with intensity $ds \otimes du \otimes P(d\lambda)$, where P is the probability distribution of $\lambda(\cdot)$. We then set

$$\mathfrak{I}^N(t) := \sum_{k=1}^{I(0)} \lambda_k^0(t) + \int_0^t \int_0^\infty \int_D \lambda(t-s) \mathbf{1}_{u \leq \Upsilon^N(s^-)} Q(ds, du, d\lambda),$$

$$A^N(t) := \int_0^t \int_0^\infty \int_D \mathbf{1}_{u \leq \Upsilon^N(s^-)} Q(ds, du, d\lambda),$$

with $\Upsilon^N(t) = \frac{S^N(t)}{N} \mathfrak{I}^N(t)$ and $S^N(t) = N - I(0) - A^N(t)$ as before. Then, for $\varepsilon \in [0, 1)$, we define

$$\mathfrak{I}_\varepsilon(t) := \sum_{k=1}^{I(0)} \lambda_k^0(t) + \int_0^t \int_0^\infty \int_D \lambda(t-s) \mathbf{1}_{u \leq (1-\varepsilon)\mathfrak{I}_\varepsilon(s^-)} Q(ds, du, d\lambda),$$

$$A_\varepsilon(t) := \int_0^t \int_0^\infty \int_D \mathbf{1}_{u \leq (1-\varepsilon)\mathfrak{I}_\varepsilon(s^-)} Q(ds, du, d\lambda).$$

Recall that, for any $\varepsilon \in [0, 1)$,

$$T_\varepsilon^N = \inf\{t \geq 0 : A^N(t) \geq \varepsilon N\}.$$

LEMMA 3.1. *For each $N \geq N_0$, the process $(\mathfrak{I}^N(t), S^N(t), A^N(t), t \geq 0)$ has the same distribution as the one defined in Subsection 2.1. Moreover,*

$$\forall t \geq 0, \quad \mathfrak{I}^N(t) \leq \mathfrak{I}_0(t), \quad A^N(t) \leq A_0(t), \tag{3.1}$$

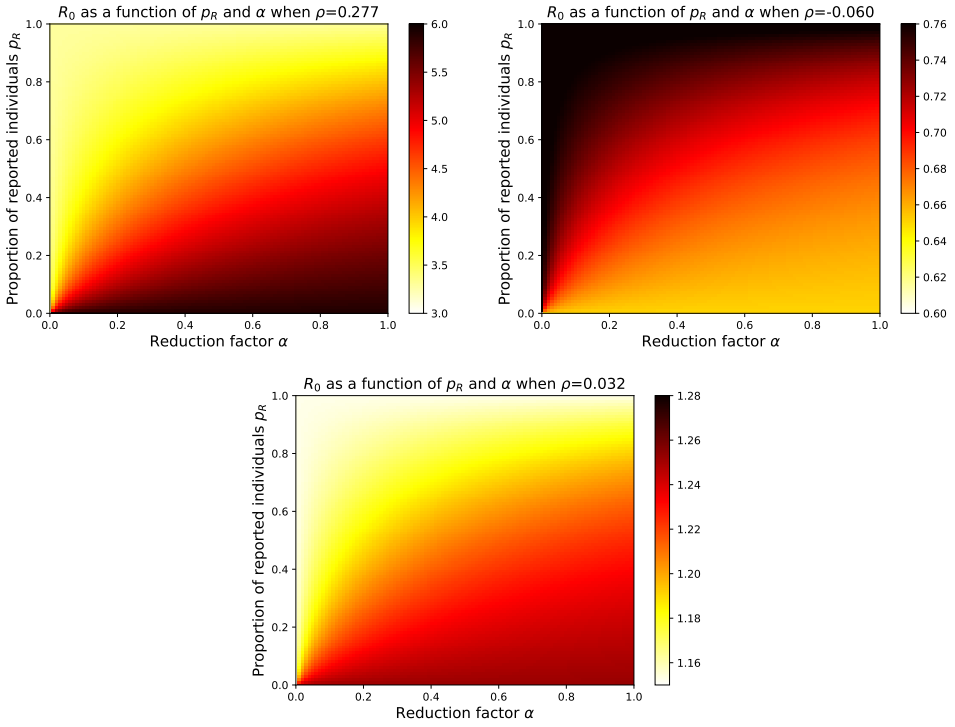

FIG. 5. *Heatmap of the value of R_0 for three growth rates: 0.277 (doubling time of 2.5 days), -0.06 (halving time of 11.6 days) and 0.032 (doubling time of 21.4 days), corresponding to three phases of the Covid–19 epidemic in France. In each graphic, the horizontal coordinate is the factor α (which is the relative infectivity of unreported individuals compared to reported individuals), and the vertical coordinate is the proportion of reported individuals p_R. Note that the range of values varies significantly with the growth rate ρ (from 3 up to 6 in the leftmost graphic, from 0.6 to 0.76 in the middle one and from 1.15 up to 1.28 in the rightmost graphic).*

and, for all $0 < \varepsilon < \varepsilon'$, for $N \geq \frac{N_0+1}{\varepsilon'-\varepsilon}$, almost surely,

$$\forall t \leq T_\varepsilon^N, \quad \mathfrak{I}^N(t) \geq \mathfrak{I}_{\varepsilon'}(t), \quad A^N(t) \geq A_{\varepsilon'}(t). \tag{3.2}$$

We note that, even though the distribution of $(\mathfrak{I}^N, A^N, S^N)$ is the same as in Subsection 2.1, this construction yields a different coupling between $(\mathfrak{I}^{N_1}, A^{N_1}, S^{N_1})$ and $(\mathfrak{I}^{N_2}, A^{N_2}, S^{N_2})$ for $N_1 \neq N_2$.

Proof. The fact that this new construction does not change the law of the process $(\mathfrak{I}^N, S^N, A^N)$ is straightforward. For the second part of the statement, let

$$\tau_0 := \inf\{t \geq 0 : \mathfrak{I}^N(t) > \mathfrak{I}_0(t)\}.$$

By construction, if $\tau_0 < \infty$, there exist $s \leq \tau_0$ and $u > 0$ such that

$$Q\left(\{s\} \times \{u\} \times D\right) = 1$$

and

$$\mathfrak{I}_0(s^-) < u \leq \Upsilon^N(s^-).$$

Since $\Upsilon^N(t) \le \mathfrak{I}^N(t)$, this implies $\mathfrak{I}_0(s^-) < \mathfrak{I}^N(s^-)$ for some $s \le \tau_0$. This contradicts the definition of τ_0, hence $\tau_0 = +\infty$ and $\mathfrak{I}^N(t) \le \mathfrak{I}_0(t)$ for all $t \ge 0$. By the definition of A^N and A_0, this also implies $A^N(t) \le A_0(t)$ for all $t \ge 0$.

For the lower bound (3.2), we note that, for $t \le T_\varepsilon^N$,

$$\begin{aligned}\Upsilon^N(t) &= \left(1 - \frac{I(0) + A^N(t)}{N}\right)\mathfrak{I}^N(t)\\ &\ge \left(1 - \frac{N_0 + 1}{N} - \varepsilon\right)\mathfrak{I}^N(t)\\ &\ge (1-\varepsilon')\mathfrak{I}^N(t),\end{aligned}$$

for $N \ge (N_0+1)/(\varepsilon' - \varepsilon)$. The lower bound then follows by a similar argument as above. ∎

We note that the process $A_\varepsilon(\cdot)$ does not depend on N, and that it is a branching process which belongs to the class of processes studied in [8, 9]. The following result is then Theorem 3.2 in [9].

LEMMA 3.2. *Under Assumptions 2.1 and 2.3, for each $\varepsilon \in [0,1)$, there exists a random variable $W_\varepsilon \ge 0$ such that*

$$A_\varepsilon(t)e^{-\rho_\varepsilon t} \to W_\varepsilon, \qquad \text{almost surely as } t \to \infty,$$

where $\rho_\varepsilon \in \mathbb{R}$ is the (unique) solution to

$$(1-\varepsilon)\int_0^\infty \overline{\lambda}(t)e^{-\rho_\varepsilon t}dt = 1. \tag{3.3}$$

Proof. We need to check the conditions of Theorem 3.2 in [9]. First, since $\lambda(t) \le \lambda^*$, for any $p > 1$,

$$\int_0^\infty (\overline{\lambda}(t))^p dt \le (\lambda^*)^{p-1}\int_0^\infty \overline{\lambda}(t)dt = (\lambda^*)^{p-1}R_0,$$

which we have assumed to be finite. On the other hand, if N is the number of offsprings of a given individual, then, using the properties of the Poisson distribution,

$$\mathbb{E}\left[N^2\right] = \mathbb{E}\left[\int_0^\infty \lambda(t)dt\right] + \mathbb{E}\left[\left(\int_0^\infty \lambda(t)dt\right)^2\right] < \infty,$$

by assumption (this is also true if the individual was initially infected, replacing λ by λ^0 above). This concludes the proof. ∎

Remark 3.3. The condition $\varepsilon \le 1 - \frac{1}{R_0}$ in Theorem 2.11 ensures that there exists a positive solution $\rho_\varepsilon > 0$ to the equation (3.3), i.e., that the branching process $A_\varepsilon(\cdot)$ is supercritical. This will be used in the proof of Theorem 2.11. See also Remark 2.12.

LEMMA 3.4. *If ρ satisfies (2.26) and ρ_ε is given by (3.3), then, for all $\varepsilon \in (0,1)$,*

$$0 \le \rho - \rho_\varepsilon \le \frac{\varepsilon}{1-\varepsilon}\left(\int_0^\infty \overline{\lambda}(t)te^{-\rho t}dt\right)^{-1}.$$

Proof. From the definitions of ρ and ρ_ε,

$$\int_0^\infty \overline{\lambda}(t)\left(e^{-\rho_\varepsilon t} - e^{-\rho t}\right)dt = \frac{\varepsilon}{1-\varepsilon}.$$

Hence it is clear that $\rho \geq \rho_\varepsilon$. In addition, $e^{-\rho_\varepsilon t} - e^{-\rho t} \geq te^{-\rho t}(\rho - \rho_\varepsilon)$, from which the stated inequality follows. ∎

LEMMA 3.5. *Let* $(W_\varepsilon, \varepsilon \in [0,1))$ *be the family of random variables defined in Lemma 3.2. Then*

$$\lim_{\varepsilon \downarrow 0} \mathbb{P}(W_\varepsilon = 0) = \mathbb{P}(W_0 = 0).$$

Proof. In [9], it is shown that $\mathbb{P}(W_\varepsilon = 0)$ is the probability of extinction of a branching process in which each individual born after time 0 leaves a conditionally Poisson number of offsprings with parameter $(1-\varepsilon)\int_0^\infty \lambda(t)dt$. Thus if X_0 denote the random variable corresponding to the number of offsprings of the $I(0)$ individuals alive at time 0, then

$$\mathbb{P}(W_\varepsilon = 0) = \mathbb{E}\left[q_\varepsilon^{X_0}\right], \tag{3.4}$$

where q_ε is the unique fixed point in $(0,1)$ of the function $s \mapsto h_\varepsilon(s)$ defined by

$$h_\varepsilon(s) := \mathbb{E}\left[s^{X_\varepsilon}\right],$$

where X_ε is conditionally Poisson with parameter $(1-\varepsilon)\int_0^\infty \lambda(t)dt$. It is then straightforward to see that h_ε converges to h_0 locally uniformly when $\varepsilon \downarrow 0$, and, as a result, $q_\varepsilon \to q_0$. We then conclude from (3.4) and the dominated convergence theorem. ∎

We can now prove Theorem 2.11.

Proof of Theorem 2.11. We begin by a lower bound on T_ε^N. By (3.1), for any $\delta \in (0,1)$,

$$A^N\left(\frac{1-\delta}{\rho}\log(N)\right) \leq A_0\left(\frac{1-\delta}{\rho}\log(N)\right).$$

Noting that $\rho_0 = \rho$, by Lemma 3.2, almost surely, for all N large enough,

$$A_0\left(\frac{1-\delta}{\rho}\log(N)\right) \leq N^{1-\delta}(W_0 + \delta).$$

But $N^{1-\delta}(W_0 + \delta) < \varepsilon N$ for N large enough. It follows that, for any $\delta \in (0,1)$,

$$\liminf_{N\to\infty} \frac{T_\varepsilon^N}{\log(N)} \geq \frac{1-\delta}{\rho}, \quad \text{almost surely.} \tag{3.5}$$

By the same argument, for any $\delta \in (0,\alpha)$ and $\alpha \in (0,1)$,

$$\liminf_{N\to\infty} \frac{\mathcal{T}_\alpha^N}{\log(N)} \geq \frac{\alpha-\delta}{\rho}. \tag{3.6}$$

On the event $\{W_0 = 0\}$, the branching process $(A_0, \mathfrak{I}_0)$ goes extinct (*i.e.*, $\mathfrak{I}_0(t) = 0$ for all t large enough), and

$$\lim_{t\to\infty} A_0(t) < +\infty.$$

As a result, for any $t > 0$,

$$\begin{aligned} A^N\left(t\log(N)\right) &\leq A_0(t\log(N)) \\ &\leq \lim_{s\to\infty} A_0(s). \end{aligned}$$

Hence $\mathcal{T}_\alpha^N > t\log(N)$ for all $t > 0$ for all N such that $N^\alpha > \lim_{t\to\infty} A_0(t)$. Hence

$$\liminf_{N\to\infty} \frac{\mathcal{T}_\alpha^N}{\log(N)} = +\infty, \tag{3.7}$$

almost surely on the event $\{W_0 = 0\}$ for any $\alpha \in (0,1)$. Since $T_\varepsilon^N \geq \mathcal{T}_\alpha^N$ for $\alpha \in (0,1)$ and N large enough, we also obtain

$$\liminf_{N\to\infty} \frac{T_\varepsilon^N}{\log(N)} = +\infty, \tag{3.8}$$

almost surely on the same event.

We now prove the upper bound on $\mathcal{T}_\alpha^N$ on the event $\{W_0 > 0\}$. By Lemma 3.1, for any $\delta \in (0, 1-\alpha)$ and $\varepsilon \in (0, 1/2)$, for N large enough,

$$A^N\left(\frac{\alpha+\delta}{\rho}\log(N) \wedge T_\varepsilon^N\right) \geq A_{2\varepsilon}\left(\frac{\alpha+\delta}{\rho}\log(N) \wedge T_\varepsilon^N\right).$$

By (3.5), $T_\varepsilon^N \geq \frac{\alpha+\delta}{\rho}\log(N)$ for all N large enough (choosing a different δ in (3.5) if needed) and, by Lemma 3.2,

$$A_{2\varepsilon}\left(\frac{\alpha+\delta}{\rho}\log(N)\right) \geq \frac{W_{2\varepsilon}}{2} N^{\frac{\rho_{2\varepsilon}}{\rho}(\alpha+\delta)},$$

almost surely for N large enough. By Lemma 3.4, we can choose ε small enough that

$$\frac{\rho_{2\varepsilon}}{\rho}(\alpha+\delta) > \alpha.$$

As a result,

$$\mathbb{P}\left(\left\{\limsup_{N\to\infty} \frac{\mathcal{T}_\alpha^N}{\log(N)} > \frac{\alpha+\delta}{\rho}\right\} \cap \{W_0 > 0\}\right) \leq \mathbb{P}\left(\{W_{2\varepsilon} = 0\} \cap \{W_0 > 0\}\right). \tag{3.9}$$

Since, by construction, $A_{2\varepsilon}(t) \leq A_0(t)$,

$$\mathbb{P}\left(\{W_{2\varepsilon} = 0\} \cap \{W_0 > 0\}\right) = \mathbb{P}(W_0 > 0) - \mathbb{P}(W_{2\varepsilon} > 0).$$

The right hand side can then be made arbitrarily small by choosing ε small enough by Lemma 3.5. Since the left hand side in (3.9) does not depend on ε, we conclude that

$$\limsup_{N\to\infty} \frac{\mathcal{T}_\alpha^N}{\log(N)} \leq \frac{\alpha+\delta}{\rho}, \tag{3.10}$$

almost surely on $\{W_0 > 0\}$. Combining (3.6), (3.7) and (3.10), we obtain that, for any $\alpha \in (0,1)$, almost surely,

$$\frac{\mathcal{T}_\alpha^N}{\log(N)} \to \begin{cases} \frac{\alpha}{\rho} & \text{if } W_0 > 0 \\ +\infty & \text{otherwise.} \end{cases}$$

This convergence thus holds in distribution for the original model defined in Subsection 2.1.

We now prove the upper bound on T_ε^N on the event $\{W_0 > 0\}$ for $\varepsilon < 1 - \frac{1}{R_0}$. To do this, we define, for $\delta \in (0,1)$, $\varepsilon' \in (\varepsilon, 1 - \frac{1}{R_0})$ and $\eta \in (0,1)$,

$$\mathfrak{I}_-^N(t) := \sum_{k=1}^{I(0)} \lambda_k^0(t) + \int_0^t \int_0^\infty \int_D \lambda(t-s) \mathbf{1}_{u \le q^N(s)\mathfrak{I}_-^N(s^-)} Q(ds, du, d\lambda),$$

$$A_-^N(t) := \int_0^t \int_0^\infty \int_D \mathbf{1}_{u \le q^N(s)\mathfrak{I}_-^N(s^-)} Q(ds, du, d\lambda),$$

where

$$q^N(t) = \begin{cases} 1-\eta & \text{if } 0 \le t \le \frac{1-\delta}{\rho}\log(N) \\ 1-\varepsilon' & \text{otherwise.} \end{cases}$$

We note that, for $t \le \frac{1-\delta}{\rho}\log(N)$, $(\mathfrak{I}_-^N(t), A_-^N(t)) = (\mathfrak{I}_\eta(t), A_\eta(t))$ and, by a similar argument as in Lemma 3.1, for all N large enough, using (3.5),

$$\forall t \le T_\varepsilon^N, \quad \mathfrak{I}^N(t) \ge \mathfrak{I}_-^N(t), \quad A^N(t) \ge A_-^N(t). \tag{3.11}$$

In addition, for any $\delta' > 0$,

$$A_-^N\left(\frac{1+\delta'}{\rho}\log(N)\right) = A_\eta\left(\frac{1-\delta}{\rho}\log(N)\right) \frac{A_-^N\left(\frac{1-\delta}{\rho}\log(N) + \frac{\delta+\delta'}{\rho}\log(N)\right)}{A_\eta\left(\frac{1-\delta}{\rho}\log(N)\right)}.$$

By Lemma 3.2, for all N large enough

$$A_\eta\left(\frac{1-\delta}{\rho}\log(N)\right) \ge \frac{W_\eta}{2} N^{\frac{\rho\eta}{\rho}(1-\delta)}. \tag{3.12}$$

Next we note that we can write, for $t \ge 0$,

$$A_-^N\left(\frac{1-\delta}{\rho}\log(N) + t\right) = \sum_{i=1}^{A_\eta\left(\frac{1-\delta}{\rho}\log(N)\right)} \tilde{A}_i(t),$$

where $(\tilde{A}_i(t), t \ge 0)_{i \ge 1}$ is a family of i.i.d. branching processes of the form

$$\tilde{A}_i(t) = \int_0^t \int_0^\infty \int_D \mathbf{1}_{u \le (1-\varepsilon')\tilde{I}_i(s^-)} \tilde{Q}_i(ds, du, d\lambda),$$

$$\tilde{\mathfrak{I}}_i(t) = \tilde{\lambda}_i^0(t) + \int_0^t \int_0^\infty \int_D \lambda(t-s) \mathbf{1}_{u \le (1-\varepsilon')\tilde{\mathfrak{I}}(s^-)} \tilde{Q}_i(ds, du, d\lambda),$$

where $\{Q, \tilde{Q}_1, \tilde{Q}_2, \ldots\}$ are i.i.d., and Q is the PRM which was used in the definition of the branching process A_η up to time $\frac{1-\delta}{\rho}\log(N)$. Since $\varepsilon' < 1 - \frac{1}{R_0}$, $\tilde{A}_i$ is supercritical and has growth rate $\rho_{\varepsilon'} > 0$. Moreover, by Lemma 3.2, $e^{-\rho_{\varepsilon'} t}\tilde{A}_i(t) \to \tilde{W}_i$ as $t \to \infty$, where the $\tilde{W}_i$ are i.i.d. and such that $\mathbb{P}(\tilde{W}_i > 0) > 0$. As a result, on $\{W_\eta > 0\}$, from (3.12),

$$A_\eta\left(\frac{1-\delta}{\rho}\log(N)\right) \to \infty$$

and, by the law of large numbers, as $N \to \infty$,

$$\frac{A_-^N\left(\frac{1-\delta}{\rho}\log(N)+\frac{\delta+\delta'}{\rho}\log(N)\right)}{A_\eta\left(\frac{1-\delta}{\rho}\log(N)\right)}N^{-\frac{\rho_{\varepsilon'}}{\rho}(\delta+\delta')} \to \mathbb{E}[\tilde{W}_1] > 0.$$

Hence on the event $\{W_\eta > 0\}$, for some constant $C > 0$ and for N large enough,

$$A_-^N\left(\frac{1+\delta'}{\rho}\log(N)\right) \geq \frac{C\,W_\eta}{4}N^{\frac{\rho_\eta}{\rho}(1-\delta)+\frac{\rho_{\varepsilon'}}{\rho}(\delta+\delta')}.$$

But by Lemma 3.4, for any $\delta' > 0$ and $\varepsilon' < 1 - \frac{1}{R_0}$ (which ensures that $\rho_{\varepsilon'} > 0$), we can choose η and δ small enough that

$$\frac{\rho_\eta}{\rho}(1-\delta)+\frac{\rho_{\varepsilon'}}{\rho}(\delta+\delta') > 1.$$

For such a choice of η and δ,

$$A_-^N\left(\frac{1+\delta'}{\rho}\log(N)\right) > N$$

for all N large enough, almost surely on the event $\{W_\eta > 0\}$. By (3.11), this implies

$$\mathbb{P}\left(\left\{\limsup_{N\to\infty}\frac{T_\varepsilon^N}{\log(N)} > \frac{1+\delta'}{\rho}\right\}\cap\{W_0>0\}\right) \leq \mathbb{P}(W_0>0)-\mathbb{P}(W_\eta>0),$$

for all $\eta > 0$ small enough. Letting $\eta \to 0$ and using Lemma 3.5, we thus obtain

$$\limsup_{N\to\infty}\frac{T_\varepsilon^N}{\log(N)} \leq \frac{1+\delta'}{\rho},$$

almost surely on $\{W_0 > 0\}$, for any $\delta' > 0$. Combining this with (3.5) and (3.8) yields the result. ∎

Let us now prove Theorem 2.13.

Proof of Theorem 2.13. Plugging (2.29) into (2.27), and replacing $\overline{\lambda}^0$ and F_0 by $\overline{\lambda}_\rho$ and F_ρ, we obtain

$$I(0)\overline{\lambda}^0(t)+\int_0^t\overline{\lambda}(t-s)\mathfrak{I}(s)ds = \int_0^\infty\overline{\lambda}(t+s)\rho e^{-\rho s}ds+\int_0^t\overline{\lambda}(t-s)\rho e^{\rho s}ds\,,$$
$$I(0)F_0^c(t)+\int_0^t F^c(t-s)\mathfrak{I}(s)ds = \int_0^\infty F^c(t+s)\rho e^{-\rho s}ds+\int_0^t F^c(t-s)\rho e^{\rho s}ds\,.$$

Changing variables in each integral and then summing them together, we obtain

$$\begin{aligned}\int_0^\infty\overline{\lambda}(t+s)\rho e^{-\rho s}ds+\int_0^t\overline{\lambda}(t-s)\rho e^{\rho s}ds &= \int_t^\infty\overline{\lambda}(s)\rho e^{\rho(t-s)}ds+\int_0^t\overline{\lambda}(s)\rho e^{\rho(t-s)}ds\\ &= \rho e^{\rho t},\end{aligned}$$

where we have used (2.26) in the last line. The same calculation with F^c instead of $\overline{\lambda}$ yields

$$\int_0^\infty F^c(t+s)\rho e^{-\rho s}ds+\int_0^t F^c(t-s)\rho e^{\rho s}ds = \int_0^\infty F^c(s)\rho e^{\rho(t-s)}ds = \boldsymbol{i}e^{\rho t},$$

using the definition of $\boldsymbol{i}$ in (2.28). In the case $\rho < 0$, these calculations are unchanged, and we simply multiply each line by -1. Finally, the equation for $R(t)$ follows from the fact that

$$\begin{aligned} I(t) + R(t) &= I(0) + R(0) + \int_0^t \mathfrak{I}(s)ds \\ &= R(0) + I(0) + \int_0^t |\rho| e^{\rho s} ds. \end{aligned}$$

Subtracting $I(t) = |\boldsymbol{i}|e^{\rho t}$, we obtain

$$R(t) = R(0) + \operatorname{sign}(\rho)(1 - \boldsymbol{i})(e^{\rho t} - 1).$$

Since $\boldsymbol{r} = 1 - \boldsymbol{i}$, this concludes the proof (we choose $R(0) = \boldsymbol{r}$ in the case $\rho > 0$). ∎

4. Proof of the FLLN. In this section, for a sequence $\{X^N, N \geq 1\}$ of random elements of D, and X a random element of D, $X^N \Rightarrow X$ in D means that X^N converges weakly (i.e., in law) towards X in D, that is, for any $\Phi \in C_b(D; \mathbb{R})$, $\mathbb{E}[\Phi(X^N)] \to \mathbb{E}[\Phi(X)]$ as $N \to \infty$.

4.1. Convergence of $(\bar{S}^N, \bar{\mathfrak{I}}^N)$.. For the process $A^N(t)$, we have the decomposition

$$A^N(t) = M_A^N(t) + \int_0^t \Upsilon^N(s)ds, \tag{4.1}$$

where

$$M_A^N(t) = \int_0^t \int_0^\infty \mathbf{1}_{u \leq \Upsilon^N(s^-)} \overline{Q}(ds, du),$$

with $\overline{Q}(ds, du) = Q(ds, du) - dsdu$ being the compensated PRM. It is clear that the process $\{M_A^N(t) : t \geq 0\}$ is a square-integrable martingale (see, e.g., [7, Chapter VI]) with respect to the filtration $\{\mathcal{F}_t^N : t \geq 0\}$ defined by

$$\begin{aligned} \mathcal{F}_t^N := \sigma\Big\{ & E^N(0), I^N(0), \{\lambda_j^0(\cdot)\}_{j\geq 1}, \{\lambda_k^{0,I}(\cdot)\}_{k\geq 1}, \{\lambda_i(\cdot)\}_{i\geq 1}, \\ & \int_0^{t'} \int_0^\infty \mathbf{1}_{u \leq \Upsilon^N(s^-)} Q(ds, du) : t' \leq t \Big\}. \end{aligned}$$

It has a finite quadratic variation

$$\langle M_A^N \rangle(t) = \int_0^t \Upsilon^N(s)ds, \quad t \geq 0.$$

Under Assumption 2.1, we have

$$0 \leq N^{-1} \int_s^t \Upsilon^N(u)du \leq \lambda^*(t - s), \quad \text{w.p. } 1 \quad \text{for} \quad 0 \leq s \leq t. \tag{4.2}$$

Thus, this implies that, in probability as $N \to \infty$,

$$\langle \overline{M}_A^N \rangle(t) = N^{-2} \int_0^t \Upsilon^N(s)ds \to 0 \quad \text{in} \quad D,$$

and by Doob's inequality,

$$\overline{M}_A^N(t) \to 0 \tag{4.3}$$

in mean square, locally uniformly in t, hence in probability in D. As a consequence, we obtain the following lemma.

LEMMA 4.1. *Under Assumptions 2.1, 2.5 and 2.6, the sequence $\{(\bar{A}^N, \bar{S}^N)\}_{N\geq 1}$ is tight in D^2. The limit of any converging subsequence of $\{\bar{A}^N\}$, denoted by $\bar{A}$, satisfies*

$$\bar{A} = \lim_{N\to\infty} \bar{A}^N = \lim_{N\to\infty} \int_0^{\cdot} \bar{\Upsilon}^N(u)du, \tag{4.4}$$

and

$$0 \leq \bar{A}(t) - \bar{A}(s) \leq \lambda^*(t-s), \quad \text{w.p. 1} \quad \text{for} \quad 0 \leq s \leq t. \tag{4.5}$$

Given the limit $\bar{A}$ of a converging subsequence of $\{\bar{A}^N\}$, along the same subsequence, $\bar{S}^N \Rightarrow \bar{S} := \bar{S}(0) - \bar{A} = 1 - \bar{I}(0) - \bar{A}$ in D as $N \to \infty$.

Let

$$\bar{\mathfrak{I}}_{0,1}^N(t) := N^{-1} \sum_{k=1}^{I^N(0)} \lambda_k^{0,I}(t), \quad \bar{\mathfrak{I}}_{0,2}^N(t) := N^{-1} \sum_{j=1}^{E^N(0)} \lambda_j^0(t), \quad t \geq 0.$$

LEMMA 4.2. *Under Assumptions 2.1 and 2.5, as $N \to \infty$,*

$$\left(\bar{\mathfrak{I}}_{0,1}^N, \bar{\mathfrak{I}}_{0,2}^N\right) \to \left(\bar{\mathfrak{I}}_{0,1}, \bar{\mathfrak{I}}_{0,2}\right) \quad \text{in} \quad D^2 \text{ in probability,} \tag{4.6}$$

where

$$\bar{\mathfrak{I}}_{0,1}(t) := \bar{I}(0)\bar{\lambda}^{0,I}(t), \quad \bar{\mathfrak{I}}_{0,2}(t) := \bar{E}(0)\bar{\lambda}^0(t), \quad t \geq 0.$$

Proof. Define the processes

$$\widetilde{\mathfrak{I}}_{0,1}^N(t) := N^{-1} \sum_{k=1}^{N\bar{I}(0)} \lambda_k^{0,I}(t), \quad \widetilde{\mathfrak{I}}_{0,2}^N(t) := N^{-1} \sum_{j=1}^{N\bar{E}(0)} \lambda_j^0(t), \quad t \geq 0. \tag{4.7}$$

By the i.i.d. assumptions for the sequences $\{\lambda_j^0(t)\}$ and $\{\lambda_k^{0,I}(t)\}$, and their independence, and by the LLN for random elements in D (see Theorem 1 in [24] or Corollary 7.10 in [18]), we directly obtain that, as $N \to \infty$,

$$\left(\widetilde{\mathfrak{I}}_{0,1}^N, \widetilde{\mathfrak{I}}_{0,2}^N\right) \to \left(\bar{\mathfrak{I}}_{0,1}, \bar{\mathfrak{I}}_{0,2}\right) \quad \text{in} \quad D^2 \text{ in probability.}$$

It then suffices to show that, as $N \to \infty$,

$$\left(\widetilde{\mathfrak{I}}_{0,1}^N - \bar{\mathfrak{I}}_{0,1}^N, \widetilde{\mathfrak{I}}_{0,2}^N - \bar{\mathfrak{I}}_{0,2}^N\right) \to 0 \quad \text{in} \quad D^2 \text{ in probability.} \tag{4.8}$$

We have

$$\widetilde{\mathfrak{I}}_{0,1}^N(t) - \bar{\mathfrak{I}}_{0,1}^N(t) = \text{sign}(\bar{I}(0) - \bar{I}^N(0)) N^{-1} \sum_{k=N(\bar{I}^N(0)\wedge\bar{I}(0))}^{N(\bar{I}^N(0)\vee\bar{I}(0))} \lambda_k^{0,I}(t), \tag{4.9}$$

and thus

$$\sup_{0\le t\le T}\left|\widetilde{\mathfrak{I}}^N_{0,1}(t)-\bar{\mathfrak{I}}^N_{0,1}(t)\right|\le\lambda^*\left|\bar{I}^N(0)-\bar{I}(0)\right|.$$

By the convergence $\bar{I}^N(0)-\bar{I}(0)\to 0$ in probability under Assumption 2.5, we obtain that $\widetilde{\mathfrak{I}}^N_{0,1}-\bar{\mathfrak{I}}^N_{0,1}\to 0$ in D in probability. A similar argument yields the convergence $\widetilde{\mathfrak{I}}^N_{0,2}-\bar{\mathfrak{I}}^N_{0,2}\to 0$ in D in probability. This completes the proof. ∎

Let

$$\bar{\mathfrak{I}}^N_1(t):=N^{-1}\sum_{i=1}^{A^N(t)}\lambda_i(t-\tau^N_i),\quad t\ge 0.$$

Before we prove the convergence of $\bar{\mathfrak{I}}^N_1$ in D, let us first establish three technical results which will be useful in the next proof. The first of those results was implicitly used in [21].

LEMMA 4.3. *Let $\{X^N\}_{N\ge1}$ be a sequence of random elements in D. If the two conditions*

(i) for all $\epsilon>0$, $0\le t\le T$, $\mathbb{P}\big(|X^N(t)|>\epsilon\big)\to 0$, as $N\to\infty$, and

(ii) for all $\epsilon>0$, $\limsup_N\sup_{0\le t\le T}\frac{1}{\delta}\mathbb{P}\big(\sup_{0\le u\le\delta}|X^N(t+u)-X^N(t)|>\epsilon\big)\to 0$, as $\delta\to 0$

are satisfied for all $T>0$, then $X^N(t)\to 0$ in probability locally uniformly in t.

Proof. We partition the interval $[0,T]$ into subintervals of length δ, that is, we define $t_i=i\delta\wedge T$, $i=0,1,\dots,\lfloor T/\delta\rfloor$, and obtain

$$\sup_{t\in[0,T]}|X^N(t)|\le\sup_{i=0,\dots,\lfloor T/\delta\rfloor}|X^N(t_i)|+\sup_{i=0,\dots,\lfloor T/\delta\rfloor}\sup_{u\in[0,\delta]}|X^N(t_i+u)-X^N(t_i)|\,.$$

We immediately obtain the following inequality

$$\mathbb{P}\left(\sup_{0\le t\le T}|X^N(t)|>\varepsilon\right)\le\sum_{i=0}^{\lfloor T/\Delta\rfloor}\mathbb{P}(|X^N(t_i)|>\varepsilon/2)+\left(\frac{T}{\delta}+1\right)\sup_{0\le t\le T}\mathbb{P}\left(\sup_{0\le u\le\delta}|X^N(t+u)-X^N(t)|>\epsilon/2\right).$$

From condition (i), $\limsup_N$ of the first term on the right is zero for any $\delta>0$, while by condition (ii), $\limsup_N$ of the second term tends to zero as $\delta\to 0$. The result follows. ∎

In the next statement, $D_\uparrow(\mathbb{R}_+)$ (resp. $C_\uparrow(\mathbb{R}_+)$) denotes the set of real-valued nondecreasing function on $\mathbb{R}_+$, which belong to $D(\mathbb{R}_+)$ (resp. $C(\mathbb{R}_+)$).

LEMMA 4.4. *Let $f\in D(\mathbb{R}_+)$ and $\{g_N\}_{N\ge1}$ be a sequence of elements of $D_\uparrow(\mathbb{R}_+)$ which is such that $g_N\to g$ locally uniformly as $N\to\infty$, where $g\in C_\uparrow(\mathbb{R}_+)$. Then, for any $t>0$, as $N\to\infty$,*

$$\int_{[0,t]}f(s)g_N(ds)\to\int_{[0,t]}f(s)g(ds)\,.$$

Proof. The assumption implies that the sequence of measures $g_N(ds)$ converges weakly, as $N\to\infty$, towards the measure $g(ds)$. Since, moreover, f is bounded and the set of discontinuities of f is of $g(ds)$ measure 0, the convergence is essentially a minor improvement of the Portmanteau theorem, see Theorem 2.1 in [3]. ∎

LEMMA 4.5. *Let $\{X^N,\ N \geq 1\}$ be a sequence of random elements in D, which is such that for all $k \geq 1$, $0 \leq t_1 < t_2 < \cdots < t_k$, as $N \to \infty$, $(X^N(t_1), \ldots, X^N(t_k)) \Rightarrow (X(t_1), \ldots, X(t_k))$, and moreover the sequence $\{X^N\}$ satisfies condition (ii) of Lemma 4.3. Then $X^N \Rightarrow X$ in D, and moreover $X \in C$ a.s. If, in addition, for all $t \geq 0$, $X^N(t) \to X(t)$ in probability, then $X^N(t) \to X(t)$ in probability locally uniformly in t.*

Proof. Define the modulus of continuity on $[0, T]$ of a function x as

$$\omega_x(T, \delta) = \sup_{0 \leq s < t \leq T,\ t-s \leq \delta} |x(t) - x(s)| .$$

It is clear (see the proof of Theorem 7.4 in [3]) that

$$\mathbb{P}(\omega_{X^N}(T, \delta) > 3\epsilon) \leq \sup_{0 \leq t \leq T} \left(\frac{T}{\delta} + 1 \right) \mathbb{P}\left(\sup_{0 \leq u \leq \delta} |X^N(t+u) - X^N(t)| > \epsilon \right)$$

Since the "D–modulus of continuity" $\omega'_x(T, \delta)$ satisfies $\omega'_x(T, \delta) \leq \omega_x(T, 2\delta)$ (see (12.7) in [3]), we conclude from Theorem 13.2 and its Corollary in [3] that $\{X^N\}$ is tight in D. Since all finite dimensional distributions of X^N converge to those of X, all converging subsequences of the sequence $\{X^N\}$ converge to X, and the whole sequence converges to X. Moreover, it follows from our assumptions that for any $T > 0$, $\omega_X(T, \delta) \to 0$, as $\delta \to 0$, hence $X \in C$ a.s. Concerning the convergence in probability, we note that under the additional assumption, $Y^N(t) := X^N(t) - X(t)$ satisfies the conditions of Lemma 4.3, hence the result. ∎

LEMMA 4.6. *Under Assumptions 2.1 and 2.6, if $\bar{A}$ is the limit of a converging subsequence of $\{\bar{A}^N\}$, then along the same subsequence,*

$$\bar{\mathfrak{I}}_1^N \Rightarrow \bar{\mathfrak{I}}_1 \quad \text{in} \quad D \quad \text{as} \quad N \to \infty, \tag{4.10}$$

where

$$\bar{\mathfrak{I}}_1(t) := \int_0^t \bar{\lambda}(t-s) d\bar{A}(s), \quad t \geq 0.$$

Proof. Let

$$\breve{\mathfrak{I}}_1^N(t) := N^{-1} \sum_{i=1}^{A^N(t)} \bar{\lambda}(t - \tau_i^N) = \int_0^t \bar{\lambda}(t-s) d\bar{A}^N(s), \quad t \geq 0. \tag{4.11}$$

The proof will be split in two steps.

STEP 1. CONVERGENCE OF $\breve{\mathfrak{I}}_1^N$

Under Assumption 2.1, applying Lemmas 4.1 and 4.4 and the continuous mapping theorem, we obtain that, as $N \to \infty$, all finite dimensional distributions of $\breve{\mathfrak{I}}_1^N$ converge to those of $\bar{\mathfrak{I}}_1$. It remains to establish condition (ii) from Lemma 4.3 in order to deduce from Lemma 4.5 that

$$\breve{\mathfrak{I}}_1^N \Rightarrow \bar{\mathfrak{I}}_1 \quad \text{in} \quad D \quad \text{as} \quad N \to \infty. \tag{4.12}$$

That is, we need to show that

$$\lim_{\delta \to 0} \limsup_{N \to \infty} \frac{1}{\delta} \sup_{t \in [0,T]} \mathbb{P}\left(\sup_{u \in [0,\delta]} \left| \breve{\mathfrak{I}}_1^N(t+u) - \breve{\mathfrak{I}}_1^N(t) \right| > \epsilon \right) = 0. \tag{4.13}$$

We have for $t, u \geq 0$,

$$\begin{aligned}\left|\check{\mathfrak{I}}_1^N(t+u) - \check{\mathfrak{I}}_1^N(t)\right| &\leq \left| N^{-1} \sum_{i=1}^{A^N(t)} \left(\bar{\lambda}(t+u-\tau_i^N) - \bar{\lambda}(t-\tau_i^N)\right)\right| \\ &\quad + N^{-1} \sum_{i=A^N(t)+1}^{A^N(t+u)} \bar{\lambda}(t+u-\tau_i^N) \\ &=: \Delta_{t,u}^{N,1} + \Delta_{t,u}^{N,2}.\end{aligned}$$

We first note that by (4.2),

$$\begin{aligned}\sup_{0\leq u\leq \delta} \Delta_{t,u}^{N,2} &\leq \lambda^*\left(\bar{A}^N(t+\delta) - \bar{A}^N(t)\right) \\ &\leq (\lambda^*)^2\delta + \lambda^*\left(\bar{M}_A^N(t+\delta) - \bar{M}_A^N(t)\right),\end{aligned}$$

so that by (4.3), for any $T > 0$, $\epsilon > 0$, provided $\delta < \varepsilon/(4(\lambda^*)^2)$,

$$\begin{aligned}\mathbb{P}\left(\sup_{0\leq u\leq\delta} \Delta_{t,u}^{N,2} > \epsilon/2\right) &\leq \mathbb{P}\left(\left|\bar{M}_A^N(t+\delta) - \bar{M}_A^N(t)\right| > \varepsilon/4\lambda^*\right) \\ &\to 0, \quad \text{as } N \to \infty,\end{aligned}$$

and consequently,

$$\limsup_{N\to\infty} \frac{1}{\delta} \sup_{t\in[0,T]} \mathbb{P}\left(\sup_{u\in[0,\delta]} \left|\Delta_{t,u}^{N,2}\right| > \epsilon/2\right) = 0. \tag{4.14}$$

We now consider the first term $\Delta_{t,u}^{N,1}$. Let

$$\Lambda_\delta(t) := \sup_{u\leq\delta} \left|\bar{\lambda}(t+u) - \bar{\lambda}(t)\right|.$$

We have

$$\sup_{u\leq\delta} \Delta_{t,u}^{N,1} \leq \int_0^t \Lambda_\delta(t-s) d\bar{A}^N(s)\,,$$

and

$$\begin{aligned}\mathbb{P}\left(\sup_{u\leq\delta} |\Delta_{t,u}^{N,1}| > \frac{\epsilon}{2}\right) &\leq \mathbb{P}\left(\int_0^t \Lambda_\delta(t-s) d\bar{A}^N(s) > \frac{\epsilon}{2}\right) \\ &\leq \mathbb{P}\left(\left|\int_0^t \Lambda_\delta(t-s) d\bar{M}_A^N(s)\right| > \frac{\epsilon}{4}\right) \\ &\quad + \mathbb{P}\left(\int_0^t \Lambda_\delta(t-s)\bar{\Upsilon}^N(s) ds > \frac{\epsilon}{4}\right).\end{aligned}$$

It is not hard to show that for any $\delta > 0$,

$$\limsup_{N\to+\infty} \frac{1}{\delta} \sup_{t\in[0,T]} \mathbb{P}\left(\left|\int_0^t \Lambda_\delta(t-s) d\bar{M}_A^N(s)\right| > \frac{\epsilon}{4}\right) = 0\,.$$

Next we note that for any $t \in [0,T]$,

$$\int_0^t \Lambda_\delta(t-s)\bar{\Upsilon}^N(s)ds \le \lambda^* \int_0^t \Lambda_\delta(t-s)ds \le \lambda^* \int_0^T \Lambda_\delta(s)ds\,.$$

Since $\bar{\lambda}$ is right continuous and bounded by λ^*, this last expression tends to 0 as $\delta \to 0$. Consequently, for $\delta > 0$ small enough,

$$\sup_N \sup_{t\in[0,T]} \mathbb{P}\left(\int_0^t \Lambda_\delta(t-s)\bar{\Upsilon}^N(s)ds > \frac{\epsilon}{4}\right) = 0\,.$$

It follows that (4.14) holds true with $\Delta^{N,2}_{t,u}$ replaced by $\Delta^{N,1}_{t,u}$. We have completed the proof of (4.13), hence of (4.12).

Step 2. $\mathfrak{I}_1^N - \check{\mathfrak{I}}_1^N \to 0$

Now it remains to show that, as $N \to \infty$,

$$V^N := \bar{\mathfrak{I}}_1^N - \check{\mathfrak{I}}_1^N \to 0 \quad \text{in} \quad D \ \text{ in probability.} \tag{4.15}$$

We have

$$V^N(t) = N^{-1}\sum_{i=1}^{A^N(t)} \chi_i^N(t), \quad \chi_i^N(t) := \lambda_i(t-\tau_i^N) - \bar{\lambda}(t-\tau_i^N).$$

$\chi_i^N(t)$ clearly satisfies $\mathbb{E}\big[\chi_i^N(t)\big] = 0$ and $\mathbb{E}\big[\chi_i^N(t)\chi_j^N(t)|\tau_i^N, \tau_j^N\big] = 0$. Thus,

$$\mathbb{E}\big[V^N(t)^2\big] = N^{-2}\mathbb{E}\left[\sum_{i=1}^{A^N(t)} \nu(t-\tau_i^N)\right] = N^{-1}\mathbb{E}\left[\int_0^t \nu(t-s)d\bar{A}^N(s)\right],$$

where $\nu(t) := E[(\lambda_i(t) - \bar{\lambda}(t))^2]$ and $\nu(t) < (\lambda^*)^2$ under Assumption 2.1. We easily obtain that for each $t \ge 0$,

$$V^N(t) \to 0 \ \text{ in probability,} \quad \text{as} \quad N \to \infty\,.$$

It remains to establish condition (ii) of Lemma 4.3, i.e., that for any $T > 0$, $\epsilon > 0$,

$$\lim_{\delta\to 0} \limsup_{N\to\infty} \frac{1}{\delta} \sup_{t\in[0,T]} \mathbb{P}\left(\sup_{u\in[0,\delta]} \big|V^N(t+u) - V^N(t)\big| > \epsilon\right) = 0. \tag{4.16}$$

We have for $t, u \ge 0$,

$$\begin{aligned}\big|V^N(t+u) - V^N(t)\big| \le{}& \left|N^{-1}\sum_{i=1}^{A^N(t)} \big(\lambda_i(t+u-\tau_i^N) - \lambda_i(t-\tau_i^N)\big)\right| \\ &+ \left|N^{-1}\sum_{i=1}^{A^N(t)} \big(\bar{\lambda}(t+u-\tau_i^N) - \bar{\lambda}(t-\tau_i^N)\big)\right| \\ &+ \left|N^{-1}\sum_{i=A^N(t)+1}^{A^N(t+u)} \big(\lambda_i(t+u-\tau_i^N) - \bar{\lambda}(t+u-\tau_i^N)\big)\right|.\end{aligned}$$

The second term has already been treated in STEP 1, and the treatment of the third term is the same as that of the second term in the analogous inequality in STEP 1 in (4.14). It remains to treat the first term, which we denote by $\Phi^{N,1}_{t,u}$. By Assumption 2.1,

$$\begin{aligned}\Phi^{N,1}_{t,u} &\le N^{-1}\sum_{i=1}^{A^N(t)}\sum_{j=1}^{k}|\lambda_i^j(t+u-\tau_i^N)-\lambda_i^j(t-\tau_i^N)|\mathbf{1}_{\xi_i^{j-1}\le t-\tau_i^N<t+u-\tau_i^N<\xi_i^j}\\ &\quad+\lambda^* N^{-1}\sum_{i=1}^{A^N(t)}\sum_{j=1}^{k}\mathbf{1}_{t-\tau_i^N\le\xi_i^j<t+u-\tau_i^N}\\ &\le \varphi_{T+\delta}(u)\bar{A}^N(t)+\lambda^*\sum_{j=1}^{k}N^{-1}\sum_{i=1}^{A^N(t)}\mathbf{1}_{t-\tau_i^N\le\xi_i^j<t+u-\tau_i^N}\,.\end{aligned}$$

The right hand side being nondecreasing in u, we deduce that

$$\sup_{0\le u\le\delta}\Phi^{N,1}_{t,u}\le\varphi_{T+\delta}(\delta)\bar{A}^N(t)+\lambda^*\sum_{j=1}^{k}N^{-1}\sum_{i=1}^{A^N(t)}\mathbf{1}_{t-\tau_i^N\le\xi_i^j<t+\delta-\tau_i^N}\,.$$

The first term on the right is the same as the one which appeared in the upper bound of $\Delta^{N,1}_{t,u}$ in STEP 1. We need only consider the second term. We have

$$\begin{aligned}&\mathbb{P}\left(\lambda^*\sum_{j=1}^{k}N^{-1}\sum_{i=1}^{A^N(t)}\mathbf{1}_{t-\tau_i^N\le\xi_i^j<t+\delta-\tau_i^N}>\epsilon\right)\\ &\le\frac{1}{\epsilon^2}\mathbb{E}\left[\left(\lambda^*\sum_{j=1}^{k}N^{-1}\sum_{i=1}^{A^N(t)}\mathbf{1}_{t-\tau_i^N\le\xi_i^j<t+\delta-\tau_i^N}\right)^2\right]\\ &\le\frac{2}{\epsilon^2}\mathbb{E}\left[\left(\lambda^*\sum_{j=1}^{k}N^{-1}\int_0^t\int_0^\infty\int_{t-s}^{t+\delta-s}\mathbf{1}_{u\le\Upsilon^N(s^-)}\overline{Q}_j(ds,du,d\xi)\right)^2\right]\\ &\quad+\frac{2}{\epsilon^2}\mathbb{E}\left[\left(\lambda^*\sum_{j=1}^{k}N^{-1}\int_0^t\big(F_j(t+\delta-s)-F_j(t-s)\big)\Upsilon^N(s)ds\right)^2\right],\end{aligned}\tag{4.17}$$

where $Q_j(ds,du,d\xi)$ is a PRM on $\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{R}_+$ with mean measure $dsduF_j(d\xi)$, and $\overline{Q}_j(ds,du,d\xi)$ is the corresponding compensated PRM. Observe that

$$\begin{aligned}&\mathbb{E}\left[\left(N^{-1}\int_0^t\int_0^\infty\int_{t-s}^{t+\delta-s}\mathbf{1}_{u\le\Upsilon^N(s^-)}\overline{Q}_j(ds,du,d\xi)\right)^2\right]\\ &=N^{-2}\mathbb{E}\left[\int_0^t\big(F_j(t+\delta-s)-F_j(t-s)\big)\Upsilon^N(s)ds\right]\\ &\le N^{-1}\lambda^*\int_0^t\big(F_j(t+\delta-s)-F_j(t-s)\big)ds,\end{aligned}$$

which tends to 0 as $N \to \infty$, for any $\delta > 0$. Moreover,

$$\begin{aligned}
&\mathbb{E}\left[\left(N^{-1}\int_0^t \big(F_j(t+\delta-s)-F_j(t-s)\big)\Upsilon^N(s)ds\right)^2\right]\\
&\le \left(\lambda^*\int_0^t \big(F_j(t+\delta-s)-F_j(t-s)\big)ds\right)^2\\
&\le \left(\lambda^*\left(\int_t^{t+\delta} F_j(u)du - \int_0^{\delta} F_j(u)du\right)\right)^2\\
&\le (\lambda^*\delta)^2\,.
\end{aligned}$$

We deduce that for any $\epsilon > 0$,

$$\limsup_{N\to\infty}\frac{1}{\delta}\sup_{t\in[0,T]}\mathbb{P}\Big(\sup_{u\in[0,\delta]}\big|\Phi^{N,1}_{t,u}\big| > \epsilon\Big) \to 0, \quad \text{as} \quad \delta \to 0. \tag{4.18}$$

We have proved (4.16). This completes the proof of the lemma. ∎

From the proof of Lemma 4.6, clearly $(\bar{A}^N, \breve{\mathfrak{I}}^N_1) \Rightarrow (\bar{A}, \bar{\mathfrak{I}}_1)$ along a subsequence. It also follows from Lemma 4.2 and the proof of Lemma 4.6 that $\bar{\mathfrak{I}}^N - \breve{\mathfrak{I}}^N_1 \to \bar{\mathfrak{I}}_{0,1} + \bar{\mathfrak{I}}_{0,2}$ in probability in D, as $N \to \infty$. Hence $(\bar{A}^N, \mathfrak{I}^N) \Rightarrow (\bar{A}, \bar{\mathfrak{I}})$ along the same subsequence as above, where $\bar{\mathfrak{I}} = \bar{\mathfrak{I}}_{0,1} + \bar{\mathfrak{I}}_{0,2} + \bar{\mathfrak{I}}_1$. It follows that, along that subsequence,

$$\int_0^{\cdot} \bar{\Upsilon}^N(s)ds = \int_0^{\cdot} \bar{S}^N(s)\bar{\mathfrak{I}}^N(s)ds \Rightarrow \int_0^{\cdot} \bar{S}(s)\bar{\mathfrak{I}}(s)ds \quad \text{in} \quad D, \tag{4.19}$$

and also

$$\bar{A}^N \Rightarrow \bar{A} = \int_0^{\cdot} \bar{S}(s)\bar{\mathfrak{I}}(s)ds \quad \text{in} \quad D. \tag{4.20}$$

Therefore, the limits $\big(\bar{S}, \bar{\mathfrak{I}}\big)$ satisfy the integral equations (2.14) and (2.15) in Theorem 2.7. Finally, the existence and uniqueness of a deterministic solution to the integral equations follows from applying Gronwall's inequality in a straightforward way, and the whole sequence converges in probability. This completes the proof of the convergence of $\big(\bar{S}^N, \bar{\mathfrak{I}}^N\big) \to \big(\bar{S}, \bar{\mathfrak{I}}\big)$ in D^2 in probability.

4.2. Convergence of $(\bar{E}^N, \bar{I}^N, \bar{R}^N)$. The proof for the convergence of the processes $(\bar{E}^N, \bar{I}^N, \bar{R}^N)$ in D^3 will be similar to the previous step.

For the initially exposed and infectious individuals, let

$$\begin{aligned}
&\bar{E}^N_0(t) := N^{-1}\sum_{j=1}^{E^N(0)} \mathbf{1}_{\zeta^0_j > t}\,,\\
&\bar{I}^N_{0,1}(t) := N^{-1}\sum_{k=1}^{I^N(0)} \mathbf{1}_{\eta^{0,I}_k > t}\,, \quad \bar{I}^N_{0,2}(t) := N^{-1}\sum_{j=1}^{E^N(0)} \mathbf{1}_{\zeta^0_j+\eta^0_j > t}\,,\\
&\bar{R}^N_{0,1}(t) := N^{-1}\sum_{k=1}^{I^N(0)} \mathbf{1}_{\eta^{0,I}_k \le t}\,, \quad \bar{R}^N_{0,2}(t) := N^{-1}\sum_{j=1}^{E^N(0)} \mathbf{1}_{\zeta^0_j+\eta^0_j \le t}\,.
\end{aligned}$$

By the FLLN for empirical processes, we obtain the following lemma.

LEMMA 4.7. *Under Assumption 2.5, as* $N \to \infty$,

$$\left(\bar{E}_0^N, \bar{I}_{0,1}^N, \bar{I}_{0,2}^N, \bar{R}_{0,1}^N, \bar{R}_{0,2}^N\right) \to \left(\bar{E}_0, \bar{I}_{0,1}, \bar{I}_{0,2}, \bar{R}_{0,1}, \bar{R}_{0,2}\right) \quad \text{in} \quad D^5 \text{ in probability,} \tag{4.21}$$

where

$$\bar{E}_0(t) = \bar{E}(0)G_0^c(t), \quad \bar{I}_{0,1}(t) = \bar{I}(0)F_{0,I}^c(t), \quad \bar{I}_{0,2}(t) = \bar{E}(0)\Psi_0(t),$$
$$\bar{R}_{0,1}(t) = I(0)F_{0,I}(t), \quad \bar{R}_{0,2}(t) = \bar{E}(0)\Phi_0(t).$$

Proof. Recall the definition of $\left(\widetilde{\mathfrak{I}}_{0,1}^N, \widetilde{\mathfrak{I}}_{0,2}^N\right)$ in (4.7). Similarly, define $\left(\widetilde{E}_0^N, \widetilde{I}_{0,1}^N, \widetilde{I}_{0,2}^N, \widetilde{R}_{0,1}^N, \widetilde{R}_{0,2}^N\right)$ by replacing $E^N(0)$ and $I^N(0)$ with $N\bar{E}(0)$ and $N\bar{I}(0)$, respectively, in the definitions of $\left(\bar{E}_0^N, \bar{I}_{0,1}^N, \bar{I}_{0,2}^N, \bar{R}_{0,1}^N, \bar{R}_{0,2}^N\right)$. By the i.i.d. assumption of $\{\lambda_k^{0,I}\}_{k\geq 1}$ and the definition of $\eta_k^{0,I}$ from $\lambda_k^{0,I}$ in (2.4), we obtain that, as $N \to \infty$,

$$\left(\widetilde{\mathfrak{I}}_{0,1}^N, \widetilde{I}_{0,1}^N, \widetilde{R}_{0,1}^N\right) \to \left(\bar{\mathfrak{I}}_{0,1}, \bar{I}_{0,1}, \bar{R}_{0,1}\right) \quad \text{in} \quad D^3 \text{ in probability.}$$

Similarly, by the i.i.d. assumption of $\{\lambda_j^0\}_{j\geq 1}$ and the definition of (ζ_j^0, η_j^0) from λ_j^0 in (2.3), we obtain that, as $N \to \infty$,

$$\left(\widetilde{E}_0^N, \widetilde{I}_{0,2}^N, \widetilde{R}_{0,2}^N\right) \to \left(\bar{E}_0, \bar{I}_{0,2}, \bar{R}_{0,2}\right) \quad \text{in} \quad D^3 \text{ in probability.}$$

Then it remains to show that, as $N \to \infty$,

$$\left(\widetilde{E}_0^N - \bar{E}_0^N, \widetilde{I}_{0,1}^N - \bar{I}_{0,1}^N, \widetilde{I}_{0,2}^N - \bar{I}_{0,2}^N, \widetilde{R}_{0,1}^N - \bar{R}_{0,1}^N, \widetilde{R}_{0,2}^N - \bar{R}_{0,2}^N\right) \to 0 \quad \text{in} \quad D^5 \text{ in probability.}$$

Similarly as in the proof of Lemma 4.2, we have

$$\widetilde{I}_{0,2}^N(t) - \bar{I}_{0,2}^N(t) = \text{sign}(\bar{E}(0) - \bar{E}^N(0))N^{-1} \sum_{j=N(\bar{E}^N(0)\wedge\bar{E}(0))}^{N(\bar{E}^N(0)\vee\bar{E}(0))} \mathbf{1}_{\zeta_j^0+\eta_j^0>t},$$

and

$$\mathbb{E}\left[N^{-1} \sum_{j=N(\bar{E}^N(0)\wedge\bar{E}(0))}^{N(\bar{E}^N(0)\vee\bar{E}(0))} \mathbf{1}_{\zeta_j^0+\eta_j^0>t} \,\middle|\, \mathcal{F}_0^N\right] \leq \Psi_0(t)|\bar{E}(0) - \bar{E}^N(0)| \to 0 \quad \text{as} \quad N \to \infty.$$

The other convergences follow by a similar argument. This completes the proof. ∎

For the newly infected individuals, let

$$\bar{E}_1^N(t) := N^{-1} \sum_{i=1}^{A^N(t)} \mathbf{1}_{\tau_i^N+\zeta_i>t}, \quad \bar{I}_1^N(t) := N^{-1} \sum_{i=1}^{A^N(t)} \mathbf{1}_{\tau_i^N+\zeta_i\leq t<\tau_i^N+\zeta_i+\eta_i},$$
$$\bar{R}_1^N(t) := N^{-1} \sum_{i=1}^{A^N(t)} \mathbf{1}_{\tau_i^N+\zeta_i+\eta_i\leq t}.$$

LEMMA 4.8. *Under Assumptions 2.1 , 2.5 and 2.6, as* $N \to \infty$,

$$\left(\bar{E}_1^N, \bar{I}_1^N, \bar{R}_1^N\right) \to \left(\bar{E}_1, \bar{I}_1, \bar{R}_1\right) \quad \text{in} \quad D^3 \text{ in probability,} \tag{4.22}$$

where

$$\bar{E}_1(t) := \int_0^t G^c(t-s)\bar{S}(s)\bar{\mathfrak{I}}(s)ds\,, \quad \bar{I}_1(t) := \int_0^t \Psi(t-s)\bar{S}(s)\bar{\mathfrak{I}}(s)ds\,,$$
$$\bar{R}_1(t) := \int_0^t \Phi(t-s)\bar{S}(s)\bar{\mathfrak{I}}(s)ds\,.$$

Proof. We first note that we have the two identities $\bar{A}^N(t) = \bar{E}_1^N(t)+\bar{I}_1^N(t)+\bar{R}_1^N(t)$ and $\bar{A}(t) = \bar{E}_1(t) + \bar{I}_1(t) + \bar{R}_1(t)$, which reflects the two facts:

$$1 = \mathbf{1}_{\zeta_i \le t-\tau_i^N < \zeta_i+\eta_i} + \mathbf{1}_{\zeta_i > t-\tau_i^N} + \mathbf{1}_{\zeta_i+\eta_i \le t-\tau_i^N},$$
$$1 = \Psi(t-s) + G^c(t-s) + \Phi(t-s)\,.$$

Consequently, since we already know that $\bar{A}^N(t) \to \bar{A}(t)$ in probability locally uniformly in t, we only need to establish the two convergences $\bar{E}_1^N \to \bar{E}_1$ and $\bar{R}_1^N \to \bar{R}_1$, from which the convergence $\bar{I}_1^N \to \bar{I}_1$ will follow as a corollary.

We shall apply the same argument as in Lemma 4.6, but now we know that $\bar{A}^N \to \bar{A}$ in probability. Define

$$\breve{E}_1^N(t) := N^{-1}\sum_{i=1}^{A^N(t)} G^c(t-\tau_i^N) = \int_0^t G^c(t-s)d\bar{A}^N(s)\,,$$
$$\breve{R}_1^N(t) := N^{-1}\sum_{i=1}^{A^N(t)} \Phi(t-\tau_i^N) = \int_0^t \Phi(t-s)d\bar{A}^N(s)\,.$$

Let us establish that $\bar{E}_1^N \to \bar{E}_1$. We shall then discuss why the same arguments work in the case of $\bar{R}_1^N$.

Step 1 It follows from Lemma 4.4 that for all $t > 0$, $\breve{E}_1^N(t) \to \bar{E}_1(t)$ in probability. In order to establish that the convergence is in fact locally uniform in t, according to Lemma 4.5, it remains to prove that condition (ii) in Lemma 4.3 is satisfied, namely that

$$\lim_{\delta\to 0}\limsup_{N\to\infty}\frac{1}{\delta}\sup_{t\in[0,T]}\mathbb{P}\left(\sup_{u\in[0,\delta]}\left|\breve{E}_1^N(t+u)-\breve{E}_1^N(t)\right| > \epsilon\right) = 0. \tag{4.23}$$

We have

$$\breve{E}_1^N(t+u)-\breve{E}_1^N(t) = \int_0^t [G^c(t+u-s)-G^c(t-s)]d\bar{A}^N(s) + \int_t^{t+u} G^c(t+u-s)d\bar{A}^N(s)\,,$$
$$\sup_{0<u\le\delta}|\breve{E}_1^N(t+u)-\breve{E}_1^N(t)| \le \int_0^t [G^c(t-s)-G^c(t+\delta-s)]d\bar{A}^N(s) + \bar{A}^N(t+\delta)-\bar{A}^N(t)\,.$$

The second term in the right hand side satisfies

$$\bar{A}^N(t+\delta)-\bar{A}^N(t) \le \lambda^*\delta + \bar{M}_A^N(t+\delta)-\bar{M}_A^N(t),$$

and since $\bar{M}_A^N$ tends to 0 locally uniformly in t,

$$\limsup_N \sup_{t\in[0,T]} \frac{1}{\delta}\mathbb{P}\big(\bar{A}^N(t+\delta) - \bar{A}^N(t) > \epsilon\big) = 0,$$

as soon as $\delta < \epsilon/\lambda^*$. Moreover

$$\begin{aligned}
&\mathbb{P}\left(\int_0^t [G^c(t-s) - G^c(t+\delta-s)]d\bar{A}^N(s) > \epsilon\right)\\
&\le \mathbb{P}\left(\left|\int_0^t [G^c(t-s) - G^c(t+\delta-s)]d\bar{M}_A^N(s)\right| > \epsilon/2\right)\\
&\quad + \mathbb{P}\left(\int_0^t [G^c(t-s) - G^c(t+\delta-s)]\bar{\Upsilon}^N(s)ds > \epsilon/2\right).
\end{aligned}$$

It is not hard to show that for any $\delta > 0$,

$$\limsup_N \frac{1}{\delta}\sup_{t\in[0,T]} \mathbb{P}\left(\left|\int_0^t [G^c(t-s) - G^c(t+\delta-s)]d\bar{M}_A^N(s)\right| > \epsilon/2\right) = 0\,.$$

Next we note that for any $t \in [0,T]$,

$$\begin{aligned}
\int_0^t [G^c(t-s) - G^c(t+\delta-s)]\bar{\Upsilon}^N(s)ds &\le \lambda^* \int_0^t [G^c(s) - G^c(s+\delta)]ds\\
&\le \lambda^* \int_0^T [G^c(s) - G^c(s+\delta)]ds\,.
\end{aligned}$$

Since G^c is right continuous and bounded by 1, this last expression tends to 0 as $\delta \to 0$. Consequently, for $\delta > 0$ small enough,

$$\sup_N \sup_{t\in[0,T]} \mathbb{P}\left(\int_0^t [G^c(t-s) - G^c(t+\delta-s)]\bar{\Upsilon}^N(s)ds > \epsilon/2\right) = 0\,.$$

Thus, (4.23) has been established, hence $\breve{E}_1^N(t) \to \bar{E}_1(t)$ in probability locally uniformly in t. It remains to consider $\bar{E}_1^N - \breve{E}_1^N$, which we do in the next step.

Step 2 Consider

$$W^N(t) := \bar{E}_1^N(t) - \breve{E}_1^N(t) = \frac{1}{N}\sum_{i=1}^{A^N(t)} \left(\mathbf{1}_{\zeta_i > t-\tau_i^N} - G^c(t-\tau_i^N)\right).$$

It is not hard to see that if $i \neq j$,

$$\mathbb{E}\Big[\left(\mathbf{1}_{\zeta_i > t-\tau_i^N} - G^c(t-\tau_i^N)\right)\left(\mathbf{1}_{\zeta_j > t-\tau_j^N} - G^c(t-\tau_j^N)\right)\Big|\tau_i^N, \tau_j^N\Big] = 0\,.$$

Consequently,

$$\begin{aligned}
\mathbb{E}\big[\,\big(W^N(t)\big)^2\,\big] &= \frac{1}{N^2}\mathbb{E}\left[\sum_{i=1}^{A^N(t)} G^c(t-\tau_i^N)(1-G^c(t-\tau_i^N))\right]\\
&= \frac{1}{N}\mathbb{E}\left[\int_0^t G^c(t-s)(1-G^c(t-s))d\bar{A}^N(s)\right]\\
&\to 0, \quad \text{as } N\to\infty\,.
\end{aligned}$$

It remains to show that condition (ii) of Lemma 4.3 holds, namely that

$$\lim_{\delta\to 0}\limsup_{N\to\infty}\frac{1}{\delta}\sup_{t\in[0,T]}\mathbb{P}\left(\sup_{u\in[0,\delta]}\left|W^N(t+u)-W^N(t)\right|>\epsilon\right)=0. \tag{4.24}$$

We have

$$\begin{aligned}|W^N(t+u)-W^N(t)| \le{}& \frac{1}{N}\sum_{i=1}^{A^N(t)}\left(\mathbf{1}_{\zeta_i>t-\tau_i^N}-\mathbf{1}_{\zeta_i>t+u-\tau_i^N}\right)\\ &+\frac{1}{N}\sum_{i=1}^{A^N(t)}\left(G^c(t-\tau_i^N)-G^c(t+u-\tau_i^N)\right)\\ &+\left|\frac{1}{N}\sum_{i=A^N(t)+1}^{A^N(t+u)}\left(\mathbf{1}_{\zeta_i>t+u-\tau_i^N}-G^c(t+u-\tau_i^N)\right)\right|.\end{aligned}$$

The second term has already been treated in Step 1, as well as $\bar{A}^N(t+\delta)-\bar{A}^N(t)$, which bounds the third term. It remains to treat the first term. Let

$$\Delta_1^N(t,u):=\frac{1}{N}\sum_{i=1}^{A^N(t)}\mathbf{1}_{t-\tau_i^N<\zeta_i\le t+u-\tau_i^N},$$

$$\sup_{u\le\delta}\Delta_1^N(t,u)=\frac{1}{N}\sum_{i=1}^{A^N(t)}\mathbf{1}_{t-\tau_i^N<\zeta_i\le t+\delta-\tau_i^N},$$

$$\mathbb{P}\left(\sup_{u\le\delta}\Delta_1^N(t,u)>\epsilon\right)\le\frac{1}{\epsilon^2}\mathbb{E}\left[\left(\frac{1}{N}\sum_{i=1}^{A^N(t)}\mathbf{1}_{t-\tau_i^N<\zeta_i\le t+\delta-\tau_i^N}\right)^2\right].$$

Let $P(ds,du,d\zeta)$ be a PRM on $\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{R}_+$ with mean measure $dsduG(d\zeta)$, and $\bar{P}$ the associated compensated measure. We have

$$\begin{aligned}&\mathbb{E}\left[\left(\frac{1}{N}\sum_{i=1}^{A^N(t)}\mathbf{1}_{t-\tau_i^N<\zeta_i\le t+\delta-\tau_i^N}\right)^2\right]\\ &=\mathbb{E}\left[\left(\frac{1}{N}\int_0^t\int_0^\infty\int_{t-s}^{t+\delta-s}\mathbf{1}_{u\le\Upsilon^N(s^-)}P(ds,du,d\zeta)\right)^2\right]\\ &\le 2\mathbb{E}\left[\left(\frac{1}{N}\int_0^t\int_0^\infty\int_{t-s}^{t+\delta-s}\mathbf{1}_{u\le\Upsilon^N(s^-)}\bar{P}(ds,du,d\zeta)\right)^2\right]\\ &\quad+2\mathbb{E}\left[\left(\frac{1}{N}\int_0^t(G^c(t-s)-G^c(t+\delta-s))\Upsilon^N(s)ds\right)^2\right].\end{aligned}$$

The first term is of ordre N^{-1}, and tends to 0 as $N\to\infty$. The second term is bounded by $2(\lambda^*)^2$ times

$$\begin{aligned}\left(\int_0^t(G(t+\delta-s)-G(t-s))ds\right)^2&\le\left(\int_t^{t+\delta}G(u)du-\int_0^\delta G(u)du\right)^2\\ &\le\delta^2.\end{aligned}$$

Consequently

$$\limsup_N \frac{1}{\delta} \sup_{t\le T} \mathbb{P}\left(\sup_{u\le\delta} \Delta_1^N(t,u) > \epsilon\right) \to 0, \quad \text{as } \delta \to 0\,.$$

STEP 3. THE CASE OF $\bar{R}_1^N$. Essentially the same argument will work in the case of $\bar{R}_1^N$ (G^c was decreasing, Φ is increasing). The details are left to the reader. ∎

Remark 4.9. A proof of Lemma 4.8 can be found in [21]. There the authors use the fact that the integral of $G^c(t-s)$ (resp. $\Phi(t-s)$) can be integrated by parts, since G^c (resp. Φ) is decreasing (resp. increasing), thus simplifying step 1 of the proof. However, the present version of step 1, which follows the same argument as Lemma 4.6, allows to shorten step 2.

Acknowledgments. The authors want to thank two anonymous Referees, whose criticisms and suggestions on a first version of this work have led to significant improvements, in particular to the addition of the analysis of the stochastic model during the early phase, namely Theorem 2.11. G. Pang was supported in part by the US National Science Foundation grant DMS-1715875 and Army Research Office grant W911NF-17-1-0019.

REFERENCES

[1] A. D. BARBOUR, *The duration of the closed stochastic epidemic*, Biometrika, 62 (1975), pp. 477–482.
[2] A. D. BARBOUR, K. HAMZA, H. KASPI, AND F. C. KLEBANER, *Escape from the boundary in Markov population processes*, Advances in Applied Probability, 47 (2015), pp. 1190–1211, https://doi.org/10.1239/aap/1449859806.
[3] P. BILLINGSLEY, *Convergence of probability measures*, John Wiley & Sons, 1999.
[4] F. BRAUER, *Age-of-infection and the final size relation*, Mathematical Biosciences & Engineering, 5 (2008), p. 681.
[5] F. BRAUER, C. CASTILLO-CHAVEZ, AND Z. FENG, *Mathematical Models in Epidemiology*, Springer, 2019.
[6] T. BRITTON AND E. PARDOUX, *Stochastic epidemics in a homogeneous community*, Stochastic Epidemic Models with Inference (T. Britton and E. Pardoux eds). Part I. Lecture Notes in Math. 2255, (2019), pp. 1–120.
[7] E. ÇINLAR, *Probability and Stochastics*, vol. 261, Springer Science & Business Media, 2011.
[8] K. S. CRUMP AND C. J. MODE, *A general age-dependent branching process. I*, Journal of mathematical analysis and applications, 24 (1968), pp. 494–508.
[9] K. S. CRUMP AND C. J. MODE, *A general age-dependent branching process. II*, Journal of mathematical analysis and applications, 25 (1969), pp. 8–17.
[10] S. D. K. FODOR, ZOLTAN AND T. G. KOVACS, *Why integral equations should be used instead of differential equations to describe the dynamics of epidemics*, arXiv:2004.07208, (2020).
[11] R. FORIEN, G. PANG, AND É. PARDOUX, *Estimating the state of the covid–19 epidemic in France using a model with memory*, Royal Society Open Science, 8:202327 (2021), https://doi.org/10.1098/rsos.202327.
[12] F. FOUTEL-RODIER, F. BLANQUART, P. COURAU, P. CZUPPON, J.-J. DUCHAMPS, J. GAMBLIN, É. KERDONCUFF, R. KULATHINAL, L. RÉGNIER, L. VUDUC, A. LAMBERT, AND E. SCHERTZER, *From individual-based epidemic models to McKendrick-von Foerster PDEs: A guide to modeling and inferring COVID-19 dynamics*, arXiv:2007.09622 [q-bio.PE], (2020), https://arxiv.org/abs/2007.09622.
[13] S. GAUBERT, M. AKIAN, X. ALLAMIGEON, M. BOYET, B. COLIN, ET AL., *Understanding and monitoring the evolution of the Covid-19 epidemic from medical emergency calls: the example of the Paris area*, medRxiv preprint, (2020), https://doi.org/10.1101/2020.03.27.20045575.
[14] X. HE, E. H. LAU, P. WU, X. DENG, J. WANG, X. HAO, Y. C. LAU, J. Y. WONG, Y. GUAN, X. TAN, ET AL., *Temporal dynamics in viral shedding and transmissibility of COVID-19*, Nature medicine, 26 (2020), pp. 672–675.

[15] F. Hoppensteadt, *An age dependent epidemic model*, Journal of the Franklin Institute, 297 (1974), pp. 325–333.
[16] H. Inaba and H. Sekine, *A mathematical model for Chagas disease with infection-age-dependent infectivity*, Mathematical Biosciences, 190 (2004), pp. 39–69.
[17] W. O. Kermack and A. G. McKendrick, *A contribution to the mathematical theory of epidemics*, Proceedings of the Royal Society of London. Series A, 115 (1927), pp. 700–721.
[18] M. Ledoux and M. Talagrand, *Probability in Banach Spaces: Isoperimetry and Processes*, Springer Science & Business Media, 2013.
[19] Z. Liu, P. Magal, O. Seydi, and G. Webb, *A COVID-19 epidemic model with latency period*, Infectious Disease Modelling, 5 (2020), pp. 323–337.
[20] P. Magal and C. McCluskey, *Two-group infection age model including an application to nosocomial infection*, SIAM Journal on Applied Mathematics, 73 (2013), pp. 1058–1095.
[21] G. Pang and É. Pardoux, *Functional limit theorems for non-Markovian epidemic models*, arXiv:2003.03249, (2020).
[22] G. Pang and É. Pardoux, *Multi–patch epidemic models with general infectious periods*, arXiv:2006.14412, (2020).
[23] G. Pang and É. Pardoux, *Functional central limit theorems for epidemic models with varying infectivity*, arXiv:2009.11861, (2021).
[24] R. R. Rao, *The law of large numbers for $D[0,1]$-valued random variables*, Theory of Probability & Its Applications, 8 (1963), pp. 70–74.
[25] H. R. Thieme and C. Castillo-Chavez, *How may infection-age-dependent infectivity affect the dynamics of HIV/AIDS?*, SIAM Journal on Applied Mathematics, 53 (1993), pp. 1447–1479.
[26] J. Wallinga and M. Lipsitch, *How generation intervals shape the relationship between growth rates and reproductive numbers*, Proceedings of the Royal Society B: Biological Sciences, 274 (2007), pp. 599–604, https://doi.org/10.1098/rspb.2006.3754.