Functional Limit Theorems for Non-Markovian Epidemic Models

GUODONG PANG AND ÉTIENNE PARDOUX

Abstract. We study non-Markovian stochastic epidemic models (SIS, SIR, SIRS, and SEIR), in which the infectious (and latent/exposing, immune) periods have a general distribution. We provide a representation of the evolution dynamics using the time epochs of infection (and latency/exposure, immunity). Taking the limit as the size of the population tends to infinity, we prove both a functional law of large number (FLLN) and a functional central limit theorem (FCLT) for the processes of interest in these models. In the FLLN, the limits are a unique solution to a system of deterministic Volterra integral equations, while in the FCLT, the limit processes are multidimensional Gaussian solutions of linear Volterra stochastic integral equations. In the proof of the FCLT, we provide an important Poisson random measures representation of the diffusion-scaled processes converging to Gaussian components driving the limit process.

1. Introduction

There have been extensive studies of Markovian epidemic models, including the SIS, SIR, SIRS and SEIR models, see, e.g., [1, 2, 8] for an overview. Limited work has been done for non-Markovian epidemic models, with general infectious periods, exposing and/or immune periods, etc. Chapter 3 of [8] provides a good review of the existing literature on the non-Markovian closed epidemic models. There is a lack of functional law of large numbers (FLLN) and functional central limit theorems (FCLT) for non-Markovian epidemic models.

In this paper we study some well known non-Markovian epidemic models, including SIR, SIS, SEIR and SIRS models. In all these models, the process counting the cumulative number of individuals becoming infectious is Poisson as usual with a rate depending on the susceptible and infectious populations. In the SIR and SIS models, the infectious periods are assumed to be i.i.d. with a general distribution, including deterministic, exponential, and many non-exponential distributions (see the conditions imposed on the distribution in Assumption 2.1). In the SEIR model, the exposing (latent) and infectious periods are assumed to be i.i.d. random vectors with a general joint distribution (correlation between these two periods for each individual is allowed), see Assumption 3.1. The same conditions are imposed for the infectious and recovered (immune) periods in the SIRS model.

We provide a general representation of the evolution dynamics in these epidemic models, by tracking the time epochs that each individual experiences. In the SIR model, each individual has two time epochs, times of becoming infectious and immune (recovered). In the SEIR model, each individual has three time epochs, times of becoming exposed (latent), infectious and immune (recovered). Then the process counting the number of infectious individuals can be simply represented by using these time epochs. Similarly for the other models of interest.

With these representations, we proceed to prove the FLLN and FCLT for these non-Markovian epidemic models. The results for the SIS model directly follow from those of the SIR model, and similarly, the results for the SIRS model follow from those of the SEIR model, so we focus on the studies of the SIR and SEIR models, and the results of the SIS and SIRS are stated without proofs. The fluid limits for these non-Markovian models are given as the unique deterministic solution to a system of Volterra integral equations. The limits in the FCLT are solutions of multidimensional linear

Date: March 3, 2020.

Key words and phrases. Non-Markovian epidemic models, general infectious periods, functional law of large numbers, functional central limit theorems, Poisson random measure representations.

Volterra stochastic integral equations driven by continuous Gaussian processes. These processes are, of course, non-Markovian, but if the initial quantities converge to Gaussian random variables, then the limit processes are jointly Gaussian. The Gaussian driving force comes from two independent components. One corresponds to the initial quantities: in the SIR model, these are initially infected individuals and in the SEIR model, these are initially exposed and infected individuals. The other corresponds to the newly infected individuals in the SIR model, and the newly exposed individuals in the SEIR model. These are written as functionals of a white noise with two time dimensions (which can be also regarded as space–time white noise). Although the limit processes appear very different in the Markovian case, they are equivalent to the Itô diffusion limit driven by Brownian motions, see, e.g., the proof of Proposition 2.1 for the SIS model.

In the proof of the FCLT for the SIR model, we construct a Poisson random measure (PRM) with mean measure depending on the distribution of the infectious periods, such that the diffusion-scaled processes corresponding to the Gaussian process driving the limit can be represented via integrals of white noises. This helps to establish tightness of these diffusion-scaled processes. For the SEIR model, the PRM has mean measure depending on the joint distribution of the exposing (latent) and infectious periods. It is worth observing the correspondence between the diffusion-scaled processes represented via the PRM and the functionals of the white noise mentioned above. The PRMs are also used to prove tightness in the FLLNs. These PRM representations may turn out to be useful for other studies in future work.

1.1. **Literature review.** One common approach to study non-Markovian epidemic models is by Sellke [21]. He provided a construction to define the epidemic outbreak in continuous time using two sets of i.i.d. random variables, with which one can find the distribution of the number of remaining uninfected individuals in an epidemic affecting a large population. Reinert [20] generalized Sellke's construction, and proved a deterministic limit (LLN) for the empirical measure describing the system dynamics of the SIR model, using Stein's method. From her result, we can derive the fluid model dynamics in Theorem 2.1; however, no FCLTs have been establish using her approach.

Ball [3] provided a unified approach to derive the distribution of the total size and total area under the trajectory of infectives using a Wald's identity for the epidemic process. This was extended to multi-type epidemic models in [4]. See also the LLN and CLT results for the final size of the epidemic in [8]. Barbour [5] proved limit theorems for the distribution of the time between the first infection and the last removal in the closed stochastic epidemic. See also Section 3.4 in [8].

Clancy [9] recently proposed to view the non-Markovian SIR model as a piecewise Markov deterministic process, and derived the joint distribution of the number of survivors of the epidemic and the area under the trajectory of infectives using martingales constructed from the piecewise deterministic Markov process. Gómez-Corral and López-García [13] further study the piecewise deterministic Markov process in [9] and analyze the population transmission number and the infection probability of a given susceptible individual.

For the SIS model with general infectious periods, without proving an FLLN, the Volterra integral equation was developed to describe the proportion of infectious population, see, e.g., [7, 10, 12, 14, 22].

It may be worth mentioning the connection with the infinite-server queueing literature. It may appear that the infectious process in the SIS or SIR model can be regarded as an infinite-server queue with a state-dependent arrival rate, and the infectious process in the SIRS or SEIR model can be regarded as a tandem infinite-server queue with a state-dependent arrival rate; however there are also delicate differences. See detailed discussions in Remark 2.1 and Section 3.1. We refer to the study of $G/GI/\infty$ queues with general i.i.d. service times in [15], [11], [18] and [19]. Although the representation of the epidemic evolution dynamics resembles those in [15], our methods to prove the FLLN and FCLT are new by taking into account the distinct characteristics of the epidemic models.

1.2. **Organization of the paper.** In Section 2, we first describe the SIR model in detail, state the FLLN and the FCLT for the SIR model, and then state the results for the SIS model. This is followed by the studies of the SEIR and SIRS models in Section 3. The proofs of the FLLN and FCLT of the SIR model are given in Sections 4 and 5, respectively. Those for the SEIR model are then given in Sections 6 and 7. In the Appendix, we state the auxiliary result of a system of two linear Volterra equations, and also prove Proposition 2.1.

1.3. **Notation.** Throughout the paper, $\mathbb{N}$ denotes the set of natural numbers, and $\mathbb{R}^k(\mathbb{R}^k_+)$ denotes the space of k-dimensional vectors with real (nonnegative) coordinates, with $\mathbb{R}(\mathbb{R}_+)$ for $k=1$. For $x,y\in\mathbb{R}$, denote $x\wedge y=\min\{x,y\}$ and $x\vee y=\max\{x,y\}$. Let $D=D([0,T],\mathbb{R})$ denote the space of $\mathbb{R}$–valued càdlàg functions defined on $[0,T]$. Throughout the paper, convergence in D means convergence in the Skorohod J_1 topology, see chapter 3 of [6]. Also, D^k stands for the k-fold product equipped with the product topology. Let C be the subset of D consisting of continuous functions. Let C^1 consist of all differentiable functions whose derivative is continuous. For any function $x\in D$, we use $\|x\|_T=\sup_{t\in[0,T]}|x(t)|$. For two functions $x,y\in D$, we use $x\circ y(t)=x(y(t))$ denote their composition. All random variables and processes are defined in a common complete probability space $(\Omega,\mathcal{F},\mathbb{P})$. The notation $\Rightarrow$ means convergence in distribution. We use $\mathbf{1}(\cdot)$ for indicator function. We use small-o notation for real-valued functions f and non-zero g: $f(x)=o(g(x))$ if $\limsup_{x\to\infty}|f(x)/g(x)|=0$.

2. SIR and SIS Models with general infectious period distributions

2.1. **SIR Model with general infectious periods.** In the SIR model, the population consists of susceptible, infectious and recovered (immune) individuals, where susceptible individuals get infected through interaction with infectious ones, and then experience an infectious period until becoming immune (no longer subject to infection). Let n be the population size. Let $S^n(t)$, $I^n(t)$ and $R^n(t)$ represent the susceptible, infectious and recovered individuals, respectively, at time $t\ge 0$. (The processes and random quantities are indexed by n and we let $n\to\infty$ in the asymptotic analysis.) WLOG, assume that $I^n(0)>0$, $S^n(0)=n-I^n(0)$ and $R^n(0)=0$, that is, each individual is either infectious or susceptible at time 0.

An individual i going through the susceptible-infectious-recovered (SIR) process has the following time epochs: τ^n_i and $\tau^n_i+\eta_i$, representing the times of becoming infected and immune, respectively. Here we assume that the infectious period distribution is independent of the population size. For the individuals $I^n(0)$ that are infectious at time 0, let η^0_i be the remaining infectious period. Assume that the η_i's are i.i.d. with c.d.f. F, and η^0_i are also i.i.d. with c.d.f. F_0. Let $F^c=1-F$ and $F^c_0=1-F_0$. Let λ be the rate at which infectious individuals infect susceptible ones.

We make the following assumption on F.

Assumption 2.1. *The c.d.f. F can be written as $F=F_1+F_2$, where $F_1(t)=\sum_i a_i\mathbf{1}(t\ge t_i)$ for a finite or countable number of positive numbers a_i and the corresponding t_i such that $\sum_i a_i\le 1$ and $t_0<t_1<\dots t_k<\dots$, and F_2 is Hölder continuous with exponent $\frac{1}{2}+\theta$ for some $\theta>0$, that is, $F_2(t+\delta)-F_2(t)\le c\delta^{1/2+\theta}$ for some $c>0$.*

Let $A^n(t)$ be the cumulative process of individuals that become infected by time t. Then we can express it as

$$A^n(t)=A_*\left(\lambda n\int_0^t\frac{S^n(s)}{n}\frac{I^n(s)}{n}ds\right) \tag{2.1}$$

where A_* is a unit rate Poisson process. The process $A^n(t)$ has event times τ^n_i, $i\in\mathbb{N}$. Assume that A_*, $I^n(0)$, $\{\eta^0_i\}$ and $\{\eta_i\}$ are mutually independent.

We first observe the following balance equations:

$$n=S^n(t)+I^n(t)+R^n(t),$$

$$S^n(t) = S^n(0) - A^n(t) = n - I^n(0) - A^n(t),$$
$$I^n(t) = I^n(0) + A^n(t) - R^n(t),$$

for each $t \geq 0$. The dynamics of $I^n(t)$ is given by

$$I^n(t) = \sum_{j=1}^{I^n(0)} \mathbf{1}(\eta_j^0 > t) + \sum_{i=1}^{A^n(t)} \mathbf{1}(\tau_i^n + \eta_i > t), \quad t \geq 0. \tag{2.2}$$

Here the first term counts the number of individuals that are initially infected at time 0 and remain infected at time t, and the second term counts the number of individuals that get infected between time 0 and time t, and remain infected at time t. $R^n(t)$ counts the number of recovered individuals, and can be represented as

$$R^n(t) = \sum_{j=1}^{I^n(0)} \mathbf{1}(\eta_j^0 \leq t) + \sum_{i=1}^{A^n(t)} \mathbf{1}(\tau_i^n + \eta_i \leq t), \quad t \geq 0.$$

Remark 2.1. *We remark that the dynamics of $I^n(t)$ resembles that of an $M/GI/\infty$ queue with a "state-dependent" Poisson arrival process $A^n(t)$ and i.i.d. service times $\{\eta_i\}$ under the initial condition $(I^n(0), \{\eta_j^0\})$. However, the "state-dependent" arrival rate $\lambda n \frac{S^n(s)}{n}\frac{I^n(s)}{n}$ not only depends on the infection ("queueing") state $I^n(t)$, but also upon the susceptible state $S^n(t)$. On the other hand, $S^n(t) = n - I^n(0) - A^n(t)$, so the "state-dependent" arrival rate is "self-exciting" in some sense.*

Assumption 2.2. *There exists a deterministic constant $\bar{I}(0) \in (0,1)$ such that $\bar{I}^n(0) \to \bar{I}(0)$ in probability in $\mathbb{R}_+$ as $n \to \infty$.*

Define the fluid-scaled process $\bar{X}^n := n^{-1}X^n$ for any process X^n.

Theorem 2.1. *Under Assumptions 2.1 and 2.2, the processes*

$$(\bar{S}^n, \bar{I}^n, \bar{R}^n) \to (\bar{S}, \bar{I}, \bar{R}) \quad \text{in } D^3$$

in probability as $n \to \infty$, where the limit process $(\bar{S}, \bar{I}, \bar{R})$ is the unique solution to the system of deterministic equations

$$\bar{S}(t) = 1 - \bar{I}(0) - \lambda \int_0^t \bar{S}(s)\bar{I}(s)ds, \tag{2.3}$$

$$\bar{I}(t) = \bar{I}(0)F_0^c(t) + \lambda \int_0^t F^c(t-s)\bar{S}(s)\bar{I}(s)ds, \tag{2.4}$$

$$\bar{R}(t) = \bar{I}(0)F_0(t) + \lambda \int_0^t F(t-s)\bar{S}(s)\bar{I}(s)ds, \tag{2.5}$$

for $t \geq 0$. $\bar{S}$ is in C. If F_0 is continuous, then $\bar{I}$ and $\bar{R}$ are in C; otherwise they are in D.

Remark 2.2. *We remark that the conditions on the c.d.f. F in Assumption 2.1 is required to prove tightness in D, see Lemma 4.2. Without these conditions, we can prove the convergence of $\bar{S}^n$ in D, but of $\bar{I}^n$ and $\bar{R}^n$ only in $L^2(0,T)$, see Remark 4.1 below, and thus in $L^p([0,T])$ for any $p > 0$ since the processes take values in $[0,1]$. Similarly for the FLLN in the SIS, SEIR, SIRS models, which we omit for brevity.*

Remark 2.3. *(the Markovian case) The ODE (2.3) for $\bar{S}(t)$ is the same as in the Markovian case:*

$$\bar{S}'(t) = -\lambda \bar{S}(t)\bar{I}(t), \quad \text{with} \quad \bar{S}(0) = 1 - \bar{I}(0).$$

When $F(t) = F_0(t) = 1 - e^{-\mu t}$, that is, the infectious period is exponentially distributed with parameter $\mu > 0$ (mean μ^{-1}), we have

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

and by taking derivatives,

$$\begin{aligned}\bar{I}'(t) &= -\mu \bar{I}(0)e^{-\mu t} + \lambda \bar{S}(t)\bar{I}(t) - \lambda\mu \int_0^t e^{-\mu(t-s)} \bar{S}(s)\bar{I}(s)ds \\ &= \lambda \bar{S}(t)\bar{I}(t) - \mu \bar{I}(t) = \mu \bar{I}(t)\left(\frac{\lambda}{\mu}\bar{S}(t) - 1\right).\end{aligned}$$

This is the well-known ODE for $\bar{I}(t)$ in the Markovian case. Similarly, we have

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

and its ODE representation reads

$$\bar{R}'(t) = \mu \bar{I}(t), \quad t \geq 0.$$

These ODEs of $(\bar{S}, \bar{I}, \bar{R})$ are referred to as the Kermack-McKendrick equations [1, 8].

Define the diffusion-scaled processes

$$\begin{aligned}\hat{S}^n(t) &:= \sqrt{n}\left(\bar{S}^n(t) - \bar{S}(t)\right) = \sqrt{n}\left(\bar{S}^n(t) - \left(1 - \bar{I}(0) - \lambda \int_0^t \bar{S}(s)\bar{I}(s)ds\right)\right), \\ \hat{I}^n(t) &:= \sqrt{n}\left(\bar{I}^n(t) - \bar{I}(t)\right) = \sqrt{n}\left(\bar{I}^n(t) - \bar{I}(0)F_0^c(t) - \lambda \int_0^t F^c(t-s)\bar{S}(s)\bar{I}(s)ds\right), \\ \hat{R}^n(t) &:= \sqrt{n}\left(\bar{R}^n(t) - \bar{R}(t)\right) = \sqrt{n}\left(\bar{R}^n(t) - \bar{I}(0)F_0(t) - \lambda \int_0^t F(t-s)\bar{S}(s)\bar{I}(s)ds\right). \end{aligned} \tag{2.6}$$

These represent the fluctuations around the fluid dynamics. Observe that

$$\hat{S}^n(t) + \hat{I}^n(t) + \hat{R}^n(t) = 0, \quad t \geq 0.$$

Assumption 2.3. *There exist a deterministic constant $\bar{I}(0) \in (0,1)$ and a random variable $\hat{I}(0)$ such that $\hat{I}^n(0) := \sqrt{n}(\bar{I}^n(0) - \bar{I}(0)) \Rightarrow \hat{I}(0)$ in $\mathbb{R}$ as $n \to \infty$. In addition, $\sup_n \mathbb{E}\big[\hat{I}^n(0)^2\big] < \infty$ and thus by Fatou's lemma, $\mathbb{E}\big[\hat{I}(0)^2\big] < \infty$.*

Theorem 2.2. *Under Assumptions 2.1 and 2.3, we have*

$$(\hat{S}^n, \hat{I}^n, \hat{R}^n) \Rightarrow (\hat{S}, \hat{I}, \hat{R}) \quad \text{in} \quad D^3 \quad \text{as} \quad n \to \infty. \tag{2.7}$$

The limit process $\hat{S}$ is

$$\begin{aligned}\hat{S}(t) &= -\hat{I}(0) - \hat{A}(t) \\ &= -\hat{I}(0) - \lambda \int_0^t \left(\hat{S}(s)\bar{I}(s) + \bar{S}(s)\hat{I}(s)\right) ds - \hat{M}_A(t), \quad t \geq 0,\end{aligned} \tag{2.8}$$

where $\bar{S}(t)$ and $\bar{I}(t)$ are the fluid limits given in Theorem 2.1. The limit process $\hat{I}$ is

$$\hat{I}(t) = -\hat{S}(t) - \hat{R}(t), \quad t \geq 0,$$

which can be represented by

$$\hat{I}(t) = \hat{I}(0)F_0^c(t) + \lambda \int_0^t F^c(t-s)\left(\hat{S}(s)\bar{I}(s) + \bar{S}(s)\hat{I}(s)\right) ds + \hat{I}_0(t) + \hat{I}_1(t), \quad t \geq 0, \tag{2.9}$$

where $\hat{I}_0(t)$ *is a continuous mean-zero Gaussian process with the covariance function*

$$\mathrm{Cov}(\hat{I}_0(t), \hat{I}_0(t')) = \bar{I}(0)(F_0^c(t \vee t') - F_0^c(t)F_0^c(t')), \quad t, t' \geq 0.$$

The limit process $\hat{R}$ *is*

$$\hat{R}(t) = \hat{I}(0)F_0(t) + \lambda \int_0^t F(t-s)\left(\hat{S}(s)\bar{I}(s) + \bar{S}(s)\hat{I}(s)\right) ds + \hat{R}_0(t) + \hat{R}_1(t), \quad t \geq 0, \tag{2.10}$$

where $\hat{R}_0$ *is a continuous mean-zero Gaussian process with covariance function*

$$\mathrm{Cov}(\hat{R}_0(t), \hat{R}_0(t')) = \bar{I}(0)(F_0(t \wedge t') - F_0(t)F_0(t')), \quad t, t' \geq 0,$$

and has the covariance function with $\hat{I}_0$*:*

$$\mathrm{Cov}(\hat{I}_0(t), \hat{R}_0(t')) = (F(t') - F(t))\mathbf{1}(t' \geq t) - F_0^c(t)F_0(t'), \quad t, t' \geq 0.$$

The limit processes $(\hat{M}_A, \hat{I}_1, \hat{R}_1)$ *are a continuous three-dimensional Gaussian process, independent of* $(\hat{I}_0, \hat{R}_0, \hat{I}(0))$*, and have the representation*

$$\begin{aligned}
\hat{M}_A(t) &= W_F([0,t] \times [0,\infty)),\\
\hat{I}_1(t) &= W_F([0,t] \times [t,\infty)),\\
\hat{R}_1(t) &= W_F([0,t] \times [0,t]),
\end{aligned}$$

where W_F *is a Gaussian white noise process on* $\mathbb{R}_+^2$ *with mean zero and*

$$\mathbb{E}\left[W_F((a,b] \times (c,d])^2\right] = \lambda \int_a^b (F(d-s) - F(c-s))\bar{S}(s)\bar{I}(s)ds,$$

for $0 \leq a \leq b$ *and* $0 \leq c \leq d$*.*

The limit process $\hat{S}$ *has continuous sample paths and* $\hat{I}$ *and* $\hat{R}$ *have càdlàg sample paths. If the c.d.f.* F_0 *is continuous, then* $\hat{I}$ *and* $\hat{R}$ *have continuous sample paths. If* $\hat{I}(0)$ *is a Gaussian random variable, then* $(\hat{S}, \hat{I}, \hat{R})$ *is a Gaussian process.*

Remark 2.4. *The processes* $(\hat{S}(t), \hat{I}(t), \hat{R}(t))$ *in* (2.8)*,* (2.9) *and* (2.10) *can be regarded as the solution of a three-dimensional system of Gaussian-driven linear Volterra stochastic integral equations. The existence and uniqueness of a solution is not hard to establish. From the representation of the limit processes* $(\hat{M}_A, \hat{I}_1, \hat{R}_1)$ *using the white noise* W_F*, we easily obtain their covariance functions: for for* $t, t' \geq 0$*,*

$$\begin{aligned}
\mathrm{Cov}(\hat{M}_A(t), \hat{M}_A(t')) &= \lambda \int_0^{t \wedge t'} \bar{S}(s)\bar{I}(s)ds,\\
\mathrm{Cov}(\hat{I}_1(t), \hat{I}_1(t')) &= \lambda \int_0^{t \wedge t'} F^c(t \vee t' - s)\bar{S}(s)\bar{I}(s)ds,\\
\mathrm{Cov}(\hat{R}_1(t), \hat{R}_1(t')) &= \lambda \int_0^{t \wedge t'} F(t \wedge t' - s)\bar{S}(s)\bar{I}(s)ds,\\
\mathrm{Cov}(\hat{M}_A(t), \hat{I}_1(t')) &= \lambda \int_0^{t \wedge t'} F^c(t' - s)\bar{S}(s)\bar{I}(s)ds,\\
\mathrm{Cov}(\hat{M}_A(t), \hat{R}_1(t')) &= \lambda \int_0^{t \wedge t'} F(t' - s)\bar{S}(s)\bar{I}(s)ds,\\
\mathrm{Cov}(\hat{I}_1(t), \hat{R}_1(t')) &= \lambda \int_0^{t} (F(t'-s) - F(t-s))\mathbf{1}(t' > t)\bar{S}(s)\bar{I}(s)ds.
\end{aligned}$$

In addition, if the processes $\hat{M}_A$ and $\hat{R}_1$ are treated individually, we can also use Brownian motions to represent them, that is,

$$\hat{M}_A(t) = B\left(\lambda \int_0^t \bar{S}(s)\bar{I}(s)ds\right),$$
$$\hat{R}_1(t) = B\left(\lambda \int_0^t F(t-s)\bar{S}(s)\bar{I}(s)ds\right),$$

where B is a standard Brownian motion.

Remark 2.5. *For the FCLT, the regularity conditions in Assumption 2.1 are required in our proof of the weak convergence (see Lemmas 5.3 and 5.4). Without that assumption, we would be able to establish weak convergence in $D \times (L^2(0,T))^2$ of the triple $(\hat{S}^n, \hat{I}^n, \hat{R}^n)$, see Remark 5.1 below. The same remark applies to the other models below : SIS, SEIR and SIRS. We will however not repeat this remark, for the sake of brevity.*

Remark 2.6. *When the infectious periods have an exponential distribution of parameter μ, the process $I^n(t)$ can be represented as*

$$I^n(t) = I^n(0) + A^n(t) - L_*\left(\int_0^t \mu I^n(s)ds\right), \quad t \geq 0,$$

and $\bar{I}(t)$ as

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

By martingale convergence theory, we can show that under Assumption 2.3,

$$(\hat{S}^n(t), \hat{I}^n(t)) \Rightarrow (\hat{S}(t), \hat{I}(t)) \quad in \quad (D^2, J_1) \quad as \quad n \to \infty,$$

where

$$\begin{aligned}
\hat{S}(t) &= -\hat{I}(0) - \lambda \int_0^t (\hat{S}(s)\bar{I}(s) + \bar{S}(s)\hat{I}(s))ds - B_A\left(\lambda \int_0^t \bar{S}(s)\bar{I}(s)ds\right) \\
\hat{I}(t) &= \hat{I}(0) + \lambda \int_0^t (\hat{S}(s)\bar{I}(s) + \bar{S}(s)\hat{I}(s))ds - \mu \int_0^t \hat{I}(s)ds \\
&\qquad + B_A\left(\lambda \int_0^t \bar{S}(s)\bar{I}(s)ds\right) - B_I\left(\mu \int_0^t \bar{I}(s)ds\right),
\end{aligned}$$

where B_A and B_I are independent Brownian motions. These diffusion processes are analyzed in [8, Section 2.3]*. In comparison with the limits $(\hat{S}, \hat{I})$ for the SIR model with general infectious periods, the Gaussian-driven Volterra linear stochastic integral equation reduces to the linear SDEs driven by Brownian motion. The proof of their equivalence in distribution follows a similar argument as in the proof of Proposition 2.1.*

Remark 2.7. *The approach in this paper can be slightly modified to allow the rate λ to be non-stationary $\lambda(t)$. In epidemic models, a non-stationary $\lambda(t)$ can represent seasonal effects. The process A^n is written as*

$$A^n(t) = A_*\left(n \int_0^t \lambda(s)\frac{S^n(s)}{n}\frac{I^n(s)}{n}ds\right).$$

For the SIR model, the fluid equation for $\bar{I}$ becomes

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

and the FCLT limit $\hat{I}$ becomes

$$\hat{I}(t) = \hat{I}(0)F_0^c(t) + \int_0^t \lambda(s)F^c(t-s)\left(\hat{S}(s)\bar{I}(s) + \bar{S}(s)\hat{I}(s)\right)ds + \hat{I}_0(t) + \hat{I}_1(t), \quad t \geq 0,$$

where $\hat{I}_0(t)$ is the same as in the stationary case, and $\hat{I}_1(t)$ has covariance function

$$\mathrm{Cov}(\hat{I}_1(t), \hat{I}_1(t')) = \int_0^{t\wedge t'} \lambda(s)F^c(t\vee t'-s)\bar{S}(s)\bar{I}(s)ds, \quad t, t \geq 0.$$

The same applies to the other processes, and the study of other models.

2.2. **SIS Model with general infectious periods.** In the SIS model, individuals become susceptible immediately after they go through the infectious periods. With a population of size n, we have $S^n(t) + I^n(t) = n$ for all $t \geq 0$. The cumulative infectious process A^n has the same expression (2.1) as in the SIR model. Suppose that there are initially $I^n(0)$ infectious individuals whose remaining infectious times are η_j^0, $j = 1, \dots, I^n(0)$, and each individual that become infectious after time 0 has infectious periods η_i, corresponding to the infectious time τ_i^n of A^n. We use F_0 and F for the distributions of η_j^0 and η_i, respectively. Then the dynamics of I^n has the same representation (2.2) as in the SIR model. The only difference is that $S^n(0) = n - I^n(0)$ and $S^n(t) = n - I^n(t)$ so that the dynamics of (S^n, I^n) is determined by the one-dimensional process I^n. Thus we will focus on the process I^n alone. We will impose the same conditions as in Assumptions 2.2–2.1. Define the fluid-scaled process $\bar{I}^n = n^{-1}I^n$.

Theorem 2.3. *Under Assumptions 2.1 and 2.2, the processes*

$$\bar{I}^n \to \bar{I}$$

in probability as $n \to \infty$, where

$$\bar{I}(t) = \bar{I}(0)F_0^c(t) + \lambda\int_0^t F^c(t-s)(1-\bar{I}(s))\bar{I}(s)ds, \quad t \geq 0. \tag{2.11}$$

$\bar{I} \in D$; if F_0 is continuous, then $\bar{I} \in C$.

Remark 2.8. *In the Markovian setting, assuming that $F_0(t) = F(t) = 1 - e^{-\mu t}$, we have the ODE*

$$\bar{I}'(t) = \lambda(1-\bar{I}(t))\bar{I}(t) - \mu\bar{I}(t).$$

Indeed, it is easy to check that by (2.11), we have

$$\begin{aligned}\bar{I}'(t) &= \bar{I}(0)\mu e^{-\mu t} + \lambda(1-\bar{I}(t))\bar{I}(t) - \mu\lambda\int_0^t e^{-\mu(t-s)}(1-\bar{I}(s))\bar{I}(s)ds\\ &= \lambda(1-\bar{I}(t))\bar{I}(t) - \mu\bar{I}(t).\end{aligned}$$

Note that the ODE has two equilibria, $\bar{I}^* = 0$ or $\bar{I}^* = 1 - \mu/\lambda$ if $\mu < \lambda$. For a general distributions F, if F_0 is the equilibrium (stationary excess) distribution of F (F has mean μ^{-1}), that is,

$$F_e(t) := \frac{\int_0^t F^c(s)ds}{\int_0^\infty F^c(s)ds} = \mu\int_0^t F^c(s)ds, \tag{2.12}$$

an equilibrium $\bar{I}^*$ must satisfy

$$\bar{I}^* = \mu\bar{I}^*\int_t^\infty F^c(s)ds + \lambda\bar{I}^*(1-\bar{I}^*)\int_0^t F^c(s)ds,$$

hence either $\bar{I}^* = 0$, or else by differentiating the last expression we find again $\bar{I}^* = 1 - \mu/\lambda$.

Define the diffusion-scaled process $\hat{I}^n = \sqrt{n}(\bar{I}^n - \bar{I})$. Then we have the following FCLT.

Theorem 2.4. *Under Assumptions 2.1 and 2.3, we have*

$$\hat{I}^n \Rightarrow \hat{I} \quad in \quad D \quad as \quad n \to \infty,$$

where

$$\hat{I}(t) = \hat{I}(0)F_0^c(t) + \lambda \int_0^t F^c(t-s)(1-2\bar{I}(s))\hat{I}(s)ds + \hat{I}_0(t) + \hat{I}_1(t), \quad t \geq 0, \tag{2.13}$$

where $\hat{I}_0(t)$ is a continuous mean-zero Gaussian process with the covariance function

$$\text{Cov}(\hat{I}_0(t), \hat{I}_0(t')) = \bar{I}(0)(F_0^c(t \vee t') - F_0^c(t)F_0^c(t')), \quad t, t' \geq 0,$$

and $\hat{I}_1(t)$ is a continuous mean-zero Gaussian process with covariance function

$$\text{Cov}(\hat{I}_1(t), \hat{I}_1(t')) = \lambda \int_0^{t \wedge t'} F^c(t \vee t' - s)(1-\bar{I}(s))\bar{I}(s)ds, \quad t, t' \geq 0.$$

$\hat{I}(0)$, $\hat{I}_0(t)$ and $\hat{I}_1(t)$ are mutually independent. $\hat{I}$ has càdlàg sample paths; if F_0 is continuous, then $\hat{I}$ has continuous sample paths. If $\hat{I}(0)$ is a Gaussian random variable, then $\hat{I}$ is a Gaussian process.

Remark 2.9. *The limit process $\hat{I}(t)$ is the solution of a one-dimensional Gaussian-driven linear Volterra SDE. In the Markovian case with exponential infectious periods of rate μ, we get*

$$\begin{aligned}\hat{I}(t) = \hat{I}(0) &+ \int_0^t \Big(\lambda(1-2\bar{I}(s)) - \mu\Big)\hat{I}(s))ds \\ &+ B_A\left(\lambda \int_0^t (1-\bar{I}(s))\bar{I}(s)ds\right) - B_I\left(\mu \int_0^t \bar{I}(s)ds\right),\end{aligned} \tag{2.14}$$

where B_A and B_I are independent Brownian motions. This is the well known linear SDE driven by Brownian motion for the SIS model. In the next proposition, we show that the two expressions of $\hat{I}(t)$ are equivalent in the Markovian case.

Proposition 2.1. *The expressions of $\hat{I}(t)$ in* (2.13) *and* (2.14) *for the SIS model are equivalent in distribution.*

Remark 2.10. *Suppose the system starts from the equilibrium, that is, $\bar{I}^n(0) = n\bar{I}^* = n(1-\mu/\lambda)$ for $\mu < \lambda$, and $F_0(t) = F_e(t)$ as discussed in Remark 2.8. Then the diffusion-scaled process $\hat{I}^n$ can be defined by $\hat{I}^n = \sqrt{n}(\bar{I}^n - \bar{I}^*)$. The FCLT holds with the limit process*

$$\hat{I}(t) = \lambda(1-2\bar{I}^*) \int_0^t F^c(t-s)\hat{I}(s)ds + \hat{I}_0(t) + \hat{I}_1(t), \quad t \geq 0,$$

where $\hat{I}_0(t)$ and $\hat{I}_1(t)$ have covariance functions

$$\text{Cov}(\hat{I}_0(t), \hat{I}_0(t')) = \bar{I}^*(F_0^c(t \vee t') - F_0^c(t)F_0^c(t')), \quad t, t' \geq 0,$$

and $\hat{I}_1(t)$ is a continuous mean-zero Gaussian process with covariance function

$$\text{Cov}(\hat{I}_1(t), \hat{I}_1(t')) = \mu(1-\mu/\lambda) \int_0^{t \wedge t'} F^c(t \vee t' - s)ds, \quad t, t' \geq 0.$$

In the Markovian case, we have the limiting diffusion

$$\hat{I}(t) = (\mu - \lambda) \int_0^t \hat{I}(s)ds + (2\mu(1-\mu/\lambda))^{1/2}B(t), \quad t \geq 0,$$

where B is a standard Brownian motion.

Remark 2.11. *When the infectious periods are deterministic, that is, η_i is equal to a positive constant η with probability one, the dynamics of $I^n(t)$ can be written as*

$$I^n(t) = \sum_{j=1}^{I^n(0)} \mathbf{1}(\eta_j^0 > t) + A^n(t) - A^n((t-\eta)^+), \quad t \geq 0.$$

We assume that $\eta_j^0 \sim U[0,\eta]$, that is, $F_0(t) = t/\eta$ for $t \in [0,\eta]$, which is the equilibrium (stationary excess) distribution of $F(t) = \mathbf{1}(t \geq \eta)$, $t \geq 0$. The fluid equation $\bar{I}(t)$ becomes

$$\bar{I}(t) = \bar{I}(0)(1-t/\eta)^+ + \lambda \int_{(t-\eta)^+}^t (1-\bar{I}(s))\bar{I}(s)ds$$

which gives

$$\bar{I}'(t) = -\frac{1}{\eta}\bar{I}(0)\mathbf{1}(t<\eta) + \lambda(1-\bar{I}(t))\bar{I}(t) - \lambda\mathbf{1}(t \geq \eta)(1-\bar{I}(t-\eta))\bar{I}(t-\eta).$$

It is clear that the nontrivial equilibrium is $\bar{I}^* = 1 - \frac{1}{\lambda\eta}$.

In the FCLT, we have

$$\hat{I}(t) = \hat{I}(0)(1-t/\eta)^+ + \lambda \int_{(t-\eta)^+}^t (1-2\bar{I}(s))\hat{I}(s)ds + \hat{I}_0(t) + \hat{I}_1(t), \quad t \geq 0, \tag{2.15}$$

where $\hat{I}_0(t)$, $t \in [0,\eta]$, is a continuous mean-zero Gaussian process with the covariance function

$$\mathrm{Cov}(\hat{I}_0(t), \hat{I}_0(t')) = \bar{I}(0)(1-(t\wedge t')/\eta - (1-t/\eta)(1-t'/\eta)), \quad t,t' \in [0,\eta],$$

and $\hat{I}_1(t)$, $t \geq 0$, is a continuous mean-zero Gaussian process with the covariance function

$$\mathrm{Cov}(\hat{I}_1(t), \hat{I}_1(t')) = \lambda \int_0^{t\wedge t'} \mathbf{1}(t \vee t' - s < \eta)(1-\bar{I}(s))\bar{I}(s)ds, \quad t,t' \geq 0.$$

Note that the effect of the initial quantities vanish after time η, that is, in the stochastic integral equation (2.15) of $\hat{I}(t)$, the components $\hat{I}_0(t)$ and $\hat{I}(0)(1-t/\eta)^+$ vanish after η.

If, in addition, the system starts with the equilibrium $I^n(0) = n\bar{I}^*$, then the limit process becomes

$$\hat{I}(t) = \lambda(1-2\bar{I}^*) \int_{(t-\eta)^+}^t \hat{I}(s)ds + \hat{I}_0(t) + \hat{I}_1(t), \quad t \geq 0,$$

where $\hat{I}_0(t)$ has the same covariance function as above with $\bar{I}(0) = \bar{I}^*$, and $\hat{I}_1(t)$ has covariance function

$$\mathrm{Cov}(\hat{I}_1(t), \hat{I}_1(t')) = \frac{1}{\eta}\left(1 - \frac{1}{\lambda\eta}\right)(t \wedge t' - (t \vee t' - \eta)^+)^+, \quad t,t' \geq 0.$$

3. Non-Markovian SEIR and SIRS Models

3.1. **SEIR Model with general exposing and infectious periods.** The SEIR model is described as follows. There are four groups in the population: Susceptible, Exposed, Infectious and Recovered (Immune). Susceptible individuals get infected through interactions with infectious ones. After getting infected, they become exposed and remain so during a latent period of time, and then transit to the infectious period. Afterwards, these individuals become recovered and immune, and will not be susceptible or infected in the future.

Let n be the population size. Let $S^n(t)$, $E^n(t)$, $I^n(t)$ and $R^n(t)$ represent the susceptible, exposed, infectious and recovered individuals, respectively, at time t. Assume that $I^n(0) > 0$, $E^n(0) > 0$, $R^n(0) = 0$, and $S^n(0) = n - I^n(0) - E^n(0)$. An individual i going through the S-E-I-R process has the following time epochs: τ_i^n, $\tau_i^n + \xi_i$, $\tau_i^n + \xi_i + \eta_i$, representing the times of becoming exposed, infectious and recovered (immune), respectively; namely, ξ_i is the exposure period and η_i is the infectious period. (It is reasonable to assume that ξ_i and η_i are independent of the population size n.) For the individuals $I^n(0)$ that are infectious at time 0, let η_j^0 be the remaining infectious period. For the individuals $E^n(0)$ that are exposed at time 0, let ξ_j^0 be the remaining exposure time.

Assume that (ξ_i, η_i)'s are i.i.d. bivariate random vectors with a joint distribution $H(du, dv)$, which has marginal c.d.f.'s G and F for ξ_i and η_i, respectively, and a conditional c.d.f. of η_i, $F(\cdot|u)$ given that $\xi_i = u$. Assume that (ξ_j^0, η_j)'s are i.i.d. bivariate random vectors with a joint distribution $H_0(du, dv)$, which has marginal c.d.f.'s G_0 and F for ξ_i^0 and η_j, respectively, and a conditional c.d.f. of η_j, $F_0(\cdot|u)$ given that $\xi_j^0 = u$. (Note that the pair (ξ_j^0, η_j) is the remaining exposing time and the subsequent infectious period for the i^{th} individual initially being exposed.) In addition, we assume that (ξ_i, η_i) and (ξ_i^0, η_j) are independent for each i, and they are also independent of $\{\eta_j^0\}$ (that is, the remaining infectious times of the initially infected individuals are independent of all the other exposing and infectious times). We use the notation $G^c = 1 - G$, and similarly for G_0^c, F^c and F_0^c. Define

$$\Phi_0(t) := \int_0^t \int_0^{t-u} H_0(du, dv) = \int_0^t \int_0^{t-u} F_0(dv|u) dG_0(u), \tag{3.1}$$

$$\Psi_0(t) := \int_0^t \int_{t-u}^{\infty} H_0(du, dv) = \int_0^t \int_{t-u}^{\infty} F_0(dv|u) dG_0(u) = G_0(t) - \Phi_0(t), \tag{3.2}$$

and

$$\Phi(t) := \int_0^t \int_0^{t-u} H(du, dv) = \int_0^t \int_0^{t-u} F(dv|u) dG(u), \tag{3.3}$$

$$\Psi(t) := \int_0^t \int_{t-u}^{\infty} H(du, dv) = \int_0^t \int_{t-u}^{\infty} F(dv|u) dG(u) = G(t) - \Phi(t). \tag{3.4}$$

Note that in the case of independent ξ_i and η_i, letting $F(dv) = F(dv|u)$, we have

$$\Phi(t) = \int_0^t F(t-u) dG(u), \quad \Psi(t) = \int_0^t F^c(t-u) dG(u) = G(t) - \Phi(t). \tag{3.5}$$

Similarly, with independent ξ_j^0 and η_j, letting $F_0(dv) = F_0(dv|u) = F(dv)$, we have

$$\Phi_0(t) = \int_0^t F(t-u) dG_0(u), \quad \Psi_0(t) = \int_0^t F^c(t-u) dG_0(u) = G_0(t) - \Phi_0(t). \tag{3.6}$$

We make the following assumptions on G and $F(\cdot|u)$.

Assumption 3.1. *The marginal c.d.f. G, and the conditional c.d.f. $F(\cdot|u)$ (uniformly in u) satisfy the conditions in Assumption 2.1, in particular, $F_2(\cdot|u)$ is Hölder continuous with exponent $\frac{1}{2} + \theta$ for some $\theta > 0$ uniformly in $u \ge 0$, that is, $F_2(t+\delta|u) - F_2(t|u) \le c\delta^{1/2+\theta}$ for some $c > 0$ and $\theta > 0$ uniformly for all $u \ge 0$.*

Let $A^n(t)$ be the cumulative process of individuals that become exposed between time 0 and time t. Let λ be the rate of susceptible patients that become exposed. Then we can express it as

$$A^n(t) = A_* \left(\lambda n \int_0^t \frac{S^n(s)}{n} \frac{I^n(s)}{n} ds \right) \tag{3.7}$$

where A_* is a unit rate Poisson process. (This has the same expression as the cumulative process A^n in (2.1) of individuals becoming infectious in the SIR model.) The process $A^n(t)$ has event times τ_i^n, $i \in \mathbb{N}$. Assume that the quantities A_*, $\{(\xi_j^0, \eta_0^j)\}$, $\{(\xi_i, \eta_i)\}$, and the initial quantities $(E^n(0), I^n(0))$ are mutually independent.

We represent the dynamics of (S^n, E^n, I^n, R^n) as follows: for $t \ge 0$,

$$S^n(t) = S^n(0) - A^n(t) = n - I^n(0) - E^n(0) - A^n(t), \tag{3.8}$$

$$E^n(t) = \sum_{j=1}^{E^n(0)} \mathbf{1}(\xi_j^0 > t) + \sum_{i=1}^{A^n(t)} \mathbf{1}(\tau_i^n + \xi_i > t), \tag{3.9}$$

$$\begin{aligned} I^n(t) = &\sum_{j=1}^{I^n(0)} \mathbf{1}(\eta_j^0 > t) + \sum_{j=1}^{E^n(0)} \mathbf{1}(\xi_j^0 \le t)\mathbf{1}(\xi_j^0 + \eta_j > t) \\ &+ \sum_{i=1}^{A^n(t)} \mathbf{1}(\tau_i^n + \xi_i \le t)\mathbf{1}(\tau_i^n + \xi_i + \eta_i > t), \end{aligned} \tag{3.10}$$

$$R^n(t) = \sum_{j=1}^{I^n(0)} \mathbf{1}(\eta_j^0 \le t) + \sum_{j=1}^{E^n(0)} \mathbf{1}(\xi_j^0 + \eta_j \le t) + \sum_{i=1}^{A^n(t)} \mathbf{1}(\tau_i^n + \xi_i + \eta_i \le t). \tag{3.11}$$

Note that we are abusing notation of η_j and η_i in the second and third terms of $I^n(t)$ and $R^n(t)$. The variables η_j (more precisely, η_j^E) in the second term of $I^n(t)$ correspond to the infectious periods of initially exposed individuals that have become infectious by time t, while the variables η_i (more precisely, η_i^A) in the third term correspond to the infectious periods of individuals that has become exposed and infectious after time 0 and before time t. We drop the superscripts E and A, since it should not cause any confusion.

We also let L^n be the cumulative process that counts individuals that have become infectious by time t. Then its dynamics can be represented by

$$L^n(t) = \sum_{j=1}^{E^n(0)} \mathbf{1}(\xi_j^0 \le t) + \sum_{i=1}^{A^n(t)} \mathbf{1}(\tau_i^n + \xi_i \le t), \quad t \ge 0.$$

We have the following balance equations: for each $t \ge 0$,

$$\begin{aligned} n &= S^n(t) + E^n(t) + I^n(t) + R^n(t), \\ E^n(t) &= E^n(0) + A^n(t) - L^n(t), \\ I^n(t) &= I^n(0) + L^n(t) - R^n(t). \end{aligned}$$

Observe that the dynamics of the exposure process $E^n(t)$ is similar to the infectious process $I^n(t)$ in (2.2) in the SIR model. The dynamics of the infectious process $I^n(t)$ resembles the dynamics of the second service station of a tandem infinite-server queue $G/GI/\infty - GI/\infty$, where the arrival process is A^n, and the first station has initial customers $E^n(0)$ with remaining service times $\{\xi_j^0\}$ and the second station has the initial customers $I^n(0)$ with remaining service times $\{\eta_j^0\}$. The processes L^n and R^n correspond to the departure processes from the first and second stations (service completions), respectively. Similar to the SIR model, the arrival process is Poisson with a "state-dependent" arrival rate $\lambda n \frac{S^n(s)}{n}\frac{I^n(s)}{n}$, which depends not only on the state of $I^n(s)$ (state of the second "station" in the tandem queueing model), but also on the state of susceptible individuals, $S^n(s) = n - I^n(0) - E^n(0) - A^n(t)$. However it is independent of the state of the exposure individuals $E^n(t)$.

Assumption 3.2. *There exist deterministic constants $\bar{I}(0) \in (0,1)$ and $\bar{E}(0) \in (0,1)$ such that $\bar{I}(0) + \bar{E}(0) < 1$ and $(\bar{I}^n(0), \bar{E}^n(0)) \to (\bar{I}(0), \bar{E}(0)) \in \mathbb{R}^2$ in probability as $n \to \infty$.*

Define the fluid-scaled processes as in the SIR model. We have the following FLLN for the fluid-scaled processes $(\bar{S}^n, \bar{E}^n, \bar{I}^n, \bar{R}^n)$.

Theorem 3.1. *Under Assumptions 3.1 and 3.2, we have*

$$\big(\bar{S}^n, \bar{E}^n, \bar{I}^n, \bar{R}^n\big) \to \big(\bar{S}, \bar{E}, \bar{I}, \bar{R}\big) \tag{3.12}$$

in probability as $n \to \infty$, where the limit process $(\bar{S}, \bar{E}, \bar{I}, \bar{R})$ is the unique solution to the system of deterministic equations: for each $t \geq 0$,

$$\bar{S}(t) = 1 - \bar{I}(0) - \bar{E}(0) - \bar{A}(t) = 1 - \bar{I}(0) - \bar{E}(0) - \lambda \int_0^t \bar{S}(s)\bar{I}(s)ds, \tag{3.13}$$

$$\bar{E}(t) = \bar{E}(0)G_0^c(t) + \lambda \int_0^t G^c(t-s)\bar{S}(s)\bar{I}(s)ds, \tag{3.14}$$

$$\bar{I}(t) = \bar{I}(0)F_0^c(t) + \bar{E}(0)\Psi_0(t) + \lambda \int_0^t \Psi(t-s)\bar{S}(s)\bar{I}(s)ds, \tag{3.15}$$

$$\bar{R}(t) = \bar{I}(0)F_0(t) + \bar{E}(0)\Phi_0(t) + \lambda \int_0^t \Phi(t-s)\bar{S}(s)\bar{I}(s)ds. \tag{3.16}$$

The limit $\bar{S}$ is in C and $\bar{E}$, $\bar{I}$ and $\bar{R}$ are in D. If G_0 and F_0 are continuous, then they are in C.

We remark that given the input data $\bar{I}(0)$ and $\bar{E}(0)$ and the distribution functions, the solution to the set of equations above can be determined by the two equations (3.13) and (3.15) for $\bar{S}$ and $\bar{I}$, which is a 2–dim system of linear Volterra integral equations. It is easy to check that we have the balance equation for the FLLN limits:

$$1 = \bar{S}(t) + \bar{E}(t) + \bar{I}(t) + \bar{R}(t),$$

As consequence, we have the joint convergence with

$$(\bar{A}^n, \bar{L}^n) \to (\bar{A}, \bar{L})$$

in probability as $n \to \infty$, where

$$\bar{A}(t) = \bar{E}(t) + \bar{L}(t) - \bar{E}(0),$$
$$\bar{L}(t) = \bar{I}(t) + \bar{R}(t) - \bar{I}(0).$$

In particular, we have

$$\bar{A}(t) = \lambda \int_0^t \bar{S}(s)\bar{I}(s)ds,$$
$$\bar{L}(t) = \bar{E}(0)G_0(t) + \lambda \int_0^t G(t-s)\bar{S}(s)\bar{I}(s)ds.$$

Remark 3.1. *In the Markovian case with independent ξ_i and η_i for each i, and independent ξ_j^0 and η_j^0 for each j, assuming that $G_0(t) = G(t) = 1 - e^{-\gamma t}$ and $F_0(t) = F(t) = 1 - e^{-\mu t}$, we obtain*

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

and

$$\begin{aligned}\bar{I}(t) = \bar{I}(0)e^{-\mu t} + \bar{E}(0)\int_0^t e^{-\mu(t-s)}\gamma e^{-\gamma s}ds \\ + \lambda \int_0^t \int_0^{t-s} e^{-\mu(t-s-u)}\gamma e^{-\gamma u}du\bar{S}(s)\bar{I}(s)ds,\end{aligned}$$

which lead to

$$\begin{aligned}\bar{E}'(t) &= -\gamma e^{-\gamma t}\bar{E}(0) + \lambda\bar{S}(t)\bar{I}(t) + \lambda \int_0^t (-\gamma)e^{-\gamma(t-s)}\bar{S}(s)\bar{I}(s)ds \\ &= \lambda\bar{S}(t)\bar{I}(t) - \gamma\bar{E}(t),\end{aligned}$$

and

$$\begin{aligned}\bar{I}'(t) &= -\mu e^{-\mu t}\bar{I}(0) + \bar{E}(0)\gamma e^{-\gamma t} + \bar{E}(0)\int_0^t (-\mu)e^{-\mu(t-s)}\gamma e^{-\gamma s}ds \\ &\quad + \lambda\int_0^t \gamma e^{-\gamma(t-s)}\bar{S}(s)\bar{I}(s)ds + \lambda\int_0^t\int_0^{t-s}(-\mu)e^{-\mu(t-s-u)}\gamma e^{-\gamma u}du\bar{S}(s)\bar{I}(s)ds \\ &= \gamma\bar{E}(t) - \mu\bar{I}(t).\end{aligned}$$

Together with $\bar{S}'(t) = -\lambda\bar{S}(t)\bar{I}(t)$, these ODEs of $(\bar{S},\bar{E},\bar{I},\bar{R})$ are the well-known result for the Markovian SEIR model.

Define the diffusion-scaled processes:

$$\hat{S}^n(t) := \sqrt{n}\left(\bar{S}^n(t) - \bar{S}(t)\right) = \sqrt{n}\left(\bar{S}^n(t) - 1 + \bar{I}(0) + \lambda\int_0^t \bar{S}(s)\bar{I}(s)ds\right), \tag{3.17}$$

$$\hat{E}^n(t) := \sqrt{n}\left(\bar{E}^n(t) - \bar{E}(t)\right) = \sqrt{n}\left(\bar{E}^n(t) - \bar{E}(0)G_0^c(t) - \lambda\int_0^t G^c(t-s)\bar{S}(s)\bar{I}(s)ds\right),$$

$$\hat{I}^n(t) := \sqrt{n}\left(\bar{I}^n(t) - \bar{I}(t)\right) = \sqrt{n}\left(\bar{I}^n(t) - \bar{I}(0)F_0^c(t) - \bar{E}(0)\Psi_0(t) - \lambda\int_0^t \Psi(t-s)\bar{S}(s)\bar{I}(s)ds\right),$$

$$\hat{R}^n(t) := \sqrt{n}\left(\bar{R}^n(t) - \bar{R}(t)\right) = \sqrt{n}\left(\bar{R}^n(t) - \bar{I}(0)F_0(t) - \bar{E}(0)\Phi_0(t) - \lambda\int_0^t \Phi(t-s)\bar{S}(s)\bar{I}(s)ds\right).$$

It is clear that

$$\hat{S}^n(t) + \hat{E}^n(t) + \hat{I}^n(t) + \hat{R}^n(t) = 0, \quad t \geq 0.$$

We will establish a FCLT for the diffusion-scaled processes $(\hat{A}^n, \hat{S}^n, \hat{E}^n, \hat{L}^n, \hat{I}^n, \hat{R}^n)$. For that purpose, we make the following assumption on the initial condition and on the law of the exposure / infectious periods.

Assumption 3.3. *There exist deterministic constants $\bar{I}(0) \in (0,1)$ and $\bar{E}(0) \in (0,1)$ and random variables $\hat{I}(0)$ and $\hat{E}(0)$ such that $\bar{I}(0) + \bar{E}(0) < 1$ and*

$$\left(\sqrt{n}(\bar{I}^n(0) - \bar{I}(0)), \sqrt{n}(\bar{E}^n(0) - \bar{E}(0))\right) \Rightarrow (\hat{I}(0), \hat{E}(0)) \quad \textit{in} \quad \mathbb{R}^2 \quad \textit{as} \quad n \to \infty.$$

In addition, $\sup_n \mathbb{E}\big[\hat{E}^n(0)^2\big] < \infty$ and $\sup_n \mathbb{E}\big[\hat{I}^n(0)^2\big] < \infty$, and thus by Fatou's lemma, $\mathbb{E}\big[\hat{E}(0)^2\big] < \infty$ and $\sup_n \mathbb{E}\big[\hat{I}^n(0)^2\big] < \infty$.

Theorem 3.2. *Under Assumptions 3.1 and 3.3, we have*

$$(\hat{S}^n, \hat{E}^n, \hat{I}^n, \hat{R}^n) \Rightarrow (\hat{S}, \hat{E}, \hat{I}, \hat{R}) \quad \textit{in} \quad D^4 \quad \textit{as} \quad n \to \infty. \tag{3.18}$$

The limit process $\hat{S}$ is

$$\hat{S}(t) = -\hat{I}(0) - \lambda\int_0^t \left(\hat{S}(s)\bar{I}(s) + \bar{S}(s)\hat{I}(s)\right)ds - \hat{M}_A(t), \quad t \geq 0, \tag{3.19}$$

where $\bar{S}(t)$ and $\bar{I}(t)$ are the fluid limit given in Theorem 3.1. The limit process $\hat{E}$ is

$$\hat{E}(t) = \hat{E}(0)G_0^c(t) + \lambda\int_0^t G^c(t-s)\left(\hat{S}(s)\bar{I}(s) + \bar{S}(s)\hat{I}(s)\right)ds + \hat{E}_0(t) + \hat{E}_1(t), \tag{3.20}$$

where $\hat{E}_0(t)$ is a continuous mean-zero Gaussian process with the covariance function

$$\mathrm{Cov}(\hat{E}_0(t), \hat{E}_0(t')) = \bar{E}(0)(G_0^c(t \vee t') - G_0^c(t)G_0^c(t')), \quad t, t' \geq 0.$$

The limit process $\hat{I}$ is given by

$$\hat{I}(t) = \hat{I}(0)F_0^c(t) + \hat{E}(0)\Psi_0(t) + \hat{I}_{0,1}(t) + \hat{I}_{0,2}(t) + \hat{I}_1(t) + \lambda \int_0^t \Psi(t-s)\left(\hat{S}(s)\bar{I}(s) + \bar{S}(s)\hat{I}(s)\right) ds, \tag{3.21}$$

where $\hat{I}_{0,1}(t)$ and $\hat{I}_{0,2}(t)$ are independent continuous mean-zero Gaussian processes: $\hat{I}_{0,1}(t)$ has the covariance function

$$\mathrm{Cov}(\hat{I}_{0,1}(t), \hat{I}_{0,1}(t')) = \bar{I}(0)(F_0^c(t \vee t') - F_0^c(t)F_0^c(t')), \quad t, t' \geq 0,$$

and $\hat{I}_{0,2}(t)$ has the covariance function

$$\mathrm{Cov}(\hat{I}_{0,2}(t), \hat{I}_{0,2}(t')) = \bar{E}(0)\left(\Psi_0(t \wedge t') - \Psi_0(t)\Psi_0(t')\right), \quad t, t' \geq 0.$$

The limit process $\hat{R}$ is given by

$$\hat{R}(t) = \hat{I}(0)F_0(t) + \hat{E}(0)\Phi(t) + \hat{R}_{0,1}(t) + \hat{R}_{0,2}(t) + \hat{R}_1(t) + \lambda \int_0^t \Phi(t-s)\left(\hat{S}(s)\bar{I}(s) + \bar{S}(s)\hat{I}(s)\right) ds, \tag{3.22}$$

where $\hat{R}_{0,1}(t)$ and $\hat{R}_{0,2}(t)$ are independent continuous mean-zero Gaussian processes: $\hat{R}_{0,1}(t)$ has the covariance function

$$\mathrm{Cov}(\hat{R}_{0,1}(t), \hat{R}_{0,1}(t')) = \bar{I}(0)(F_0(t \wedge t') - F_0(t)F_0(t')), \quad t, s \geq 0,$$

and $\hat{R}_{0,2}(t)$ has the covariance function

$$\mathrm{Cov}(\hat{R}_{0,2}(t), \hat{R}_{0,2}(t')) = \bar{E}(0)\left(\Phi_0(t \wedge t') - \Phi_0(t)\Phi_0(t')\right), \quad t, t' \geq 0.$$

The limit processes of the initial quantities have the following covariance functions: for $t, t' \geq 0$,

$$\begin{aligned}
\mathrm{Cov}(\hat{I}_{0,1}(t), \hat{R}_{0,1}(t')) &= \bar{I}(0)\Big((F_0(t') - F_0(t))\mathbf{1}(t' \geq t) - F_0^c(t)F_0(t')\Big), \\
\mathrm{Cov}(\hat{E}_0(t), \hat{I}_{0,2}(t')) &= \bar{E}(0)\left(\int_t^{t'} \mathbf{1}(t' \geq t)F_0^c(t'-s|s)dG_0(s) - G_0^c(t)\Psi_0(t')\right), \\
\mathrm{Cov}(\hat{E}_0(t), \hat{R}_{0,2}(t')) &= \bar{E}(0)\left(\int_t^{t'} F_0(t'-s|s)dG_0(s) - G_0^c(t)\Phi_0(t')\right) \\
\mathrm{Cov}(\hat{I}_{0,2}(t), \hat{R}_{0,2}(t')) &= \bar{E}(0)\left(\int_0^{t\wedge t'} (F_0(t'-s|s) - F_0(t-s|s))dG_0(s) - \Psi_0(t)\Phi_0(t')\right).
\end{aligned} \tag{3.23}$$

The other pairs of limit processes for the initial quantities, $(\hat{E}_0, \hat{I}_{0,1})$, $(\hat{E}_0(t), \hat{R}_{0,1})$, $(\hat{I}_{0,2}(t), \hat{R}_{0,1})$ are independent.

The limit processes $(\hat{M}_A, \hat{E}_1, \hat{I}_1, \hat{R}_1)$ are a four-dimensional continuous Gaussian process, independent of $\hat{E}_0$, $\hat{I}_{0,1}$, $\hat{I}_{0,2}$, $\hat{R}_{0,1}$, $\hat{R}_{0,2}$ and $\hat{I}(0)$, and can be written as

$$\begin{aligned}
\hat{M}_A(t) &= W_H([0,t] \times [0,\infty) \times [0,\infty)), \\
\hat{E}_1(t) &= W_H([0,t] \times [t,\infty) \times [0,\infty)), \\
\hat{I}_1(t) &= W_H([0,t] \times [0,t) \times [t,\infty)), \\
\hat{R}_1(t) &= W_H([0,t] \times [0,t) \times [0,t)).
\end{aligned}$$

where W_H is a continuous Gaussian white noise process on $\mathbb{R}_+^3$ with mean zero and

$$\mathbb{E}\left[W_H([s,t) \times [a,b) \times [c,d))^2\right]$$

$$= \lambda \int_s^t \left(\int_{a-s}^{b-s} (F(d-y-s|y) - F(c-y-s|y)) G(dy) \right) \bar{S}(s) \bar{I}(s) ds, \tag{3.24}$$

for $0 \le s \le t$, $0 \le a \le b$ and $0 \le c \le d$.

The limit process $\hat{S}$ has continuous sample paths and $\hat{E}_1$, $\hat{I}_1$ and $\hat{R}_1$ have càdlàg sample paths. If the c.d.f.'s G_0 and F_0 are continuous, then $\hat{E}_1$, $\hat{I}_1$ and $\hat{R}_1$ have continuous sample paths. If $(\hat{I}(0), \hat{E}(0))$ is a Gaussian random vector, then $(\hat{S}, \hat{E}, \hat{I}, \hat{R})$ is a Gaussian process.

Remark 3.2. *The processes $(\hat{S}(t), \hat{E}(t), \hat{I}(t), \hat{R}(t))$ in (3.19), (3.20), (3.21) and (3.22) can be regarded as the solution of a four-dimensional Gaussian-driven linear Volterra stochastic integral equation. The existence and uniqueness of solution can be easily verified. From the representations of the limit processes $(\hat{M}_A, \hat{E}_1, \hat{I}_1, \hat{R}_1)$ using the white noise W_H, we easily obtain the covariance functions: for $t, t' \ge 0$,*

$$\begin{aligned}
\mathrm{Cov}(\hat{M}_A(t), \hat{M}_A(t')) &= \lambda \int_0^{t\wedge t'} \bar{S}(s)\bar{I}(s) ds, \\
\mathrm{Cov}(\hat{E}_1(t), \hat{E}_1(t')) &= \lambda \int_0^{t\wedge t'} G^c(t \vee t' - s)\bar{S}(s)\bar{I}(s) ds, \\
\mathrm{Cov}(\hat{I}_1(t), \hat{I}_1(t')) &= \lambda \int_0^{t\wedge t'} \Psi(t \vee t' - s)\bar{S}(s)\bar{I}(s) ds, \\
\mathrm{Cov}(\hat{R}_1(t), \hat{R}_1(t')) &= \lambda \int_0^{t\wedge t'} \Phi(t \wedge t' - s)\bar{S}(s)\bar{I}(s) ds, \\
\mathrm{Cov}(\hat{M}_A(t), \hat{E}_1(t')) &= \lambda \int_0^{t\wedge t'} G^c(t' - s)\bar{S}(s)\bar{I}(s) ds, \\
\mathrm{Cov}(\hat{M}_A(t), \hat{I}_1(t')) &= \lambda \int_0^{t\wedge t'} \Psi(t' - s)\bar{S}(s)\bar{I}(s) ds, \\
\mathrm{Cov}(\hat{M}_A(t), \hat{R}_1(t')) &= \lambda \int_0^{t\wedge t'} \Phi(t' - s)\bar{S}(s)\bar{I}(s) ds,
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}(\hat{E}_1(t), \hat{I}_1(t')) &= \lambda \int_0^{t\wedge t'} (G^c(t-s) - \Psi(t'-s))\mathbf{1}(t' \ge t)\bar{S}(s)\bar{I}(s) ds, \\
\mathrm{Cov}(\hat{E}_1(t), \hat{R}_1(t')) &= \lambda \int_0^{t\wedge t'} (G^c(t-s) - \Phi(t'-s))\mathbf{1}(t' \ge t)\bar{S}(s)\bar{I}(s) ds, \\
\mathrm{Cov}(\hat{I}_1(t), \hat{R}_1(t')) &= \lambda \int_0^{t\wedge t'} \int_0^{t'-s} (F(t'-s-y|y) - F(t-s-y|y))\mathbf{1}(t' \ge t) dG(y)\bar{S}(s)\bar{I}(s) ds.
\end{aligned}$$

Remark 3.3. *We remark that the exposing and infectious periods are allowed to be dependent, and the effect of such dependence is exhibited in the covariances of the functions of the limit processes $(\hat{M}_A, \hat{E}_1, \hat{I}_1, \hat{R}_1)$ and in the drift of $\hat{I}$ and $\hat{R}$. Of course, the dependence also affects the fluid equations for $(\bar{S}, \bar{I})$.*

Remark 3.4. *We now recall the Markovian SEIR model, assuming that the exposure and infection periods are exponentially distributed with parameters γ and μ and independent. The processes A^n and S^n remain the same as in (3.7) and (3.8), respectively. The processes E^n and $I^n(t)$ can be described by*

$$E^n(t) = E^n(0) + A^n(t) - K_* \left(\gamma \int_0^t E^n(s) ds \right),$$

$$I^n(t) = I^n(0) + K_*\left(\gamma \int_0^t E^n(s)ds\right) - L_*\left(\mu \int_0^t I^n(s)ds\right),$$

where K_* and L_* are independent rate-one Poisson processes. It is easy to see the fluid limit ODE from these equations:

$$\begin{aligned}\bar{S}'(t) &= -\lambda \bar{S}(t)\bar{I}(t),\\ \bar{E}'(t) &= \lambda \bar{S}(t)\bar{I}(t) - \gamma \bar{E}'(t),\\ \bar{I}'(t) &= \gamma \bar{E}'(t) - \mu \bar{I}'(t).\end{aligned} \tag{3.25}$$

We can show that under Assumption 3.3,

$$(\hat{S}^n, \hat{E}^n, \hat{I}^n) \Rightarrow (\hat{S}, \hat{E}, \hat{I}) \quad in \quad D^3 \quad as \quad n \to \infty,$$

where $\hat{S}$ is as given in Theorem 3.2, in particular,

$$\hat{S}(t) = -\hat{I}(0) - \lambda \int_0^t (\hat{S}(s)\bar{I}(s) + \bar{S}(s)\hat{I}(s))ds - B_A\left(\lambda \int_0^t \bar{S}(s)\bar{I}(s)ds\right), \tag{3.26}$$

and the limit processes $\hat{E}$ and $\hat{I}$ are given by

$$\begin{aligned}\hat{E}(t) = \hat{E}(0) &+ \lambda \int_0^t (\hat{S}(s)\bar{I}(s) + \bar{S}(s)\hat{I}(s))ds - \gamma \int_0^t \hat{E}(s)ds\\ &+ B_A\left(\lambda \int_0^t \bar{S}(s)\bar{I}(s)ds\right) - B_K\left(\gamma \int_0^t \bar{E}(s)ds\right),\end{aligned} \tag{3.27}$$

and

$$\begin{aligned}\hat{I}(t) = \hat{I}(0) &+ \gamma \int_0^t \hat{E}(s)ds - \mu \int_0^t \hat{I}(s)ds\\ &+ B_K\left(\gamma \int_0^t \bar{E}(s)ds\right) - B_L\left(\mu \int_0^t \bar{I}(s)ds\right),\end{aligned} \tag{3.28}$$

where B_A, B_K and B_L are independent Brownian motions. Given the solution to the fluid equation in (3.25)*, the three equations* (3.26)*,* (3.27) *and* (3.28) *form a three-dimensional linear SDE driven by Brownian Motions. This is well known for the Markovian SEIR model, see* [8]*. It can be shown, similarly to the proof of Proposition 2.1, that the Volterra stochastic integral equations are equivalent to these linear SDEs in distribution.*

3.2. **SIRS model with general infectious and immune periods.** In the SIRS model, there are three groups in the population: Susceptible, Infectious, Recovered (Immune). Susceptible individuals get infected through interactions with infectious ones, and they become infectious immediately (no exposure period like in the SEIR model). The infectious individuals become recovered and immune, and after the immune periods, they become susceptible. This has a lot of resemblance with the SEIR model, where the exposure and infectious periods in the SEIR model correspond to the infectious and immune periods in the SIRS model, respectively. We let $S^n(t), I^n(t), R^n(t)$ represent the susceptible, infectious and immune individuals, respectively at each time t in the SIRS model. Note that $I^n(t)$ (resp. $R^n(t)$) in the SIRS model corresponds to $E^n(t)$ (resp. $I^n(t)$) in the SEIR model, and $S^n(t)$ in the SIRS model satisfies the balance equation:

$$n = S^n(t) + I^n(t) + R^n(t), \quad t \geq 0.$$

Since $S^n(t) = n - I^n(t) - R^n(t)$, it suffices to only study the dynamics of the two processes (I^n, R^n). We use the variables ξ_i, η_i represent the infectious and immune periods, respectively, in the SIRS model, and similarly for the initial quantities ξ^0_j, η_j. We also use the same distribution functions associated with these variables as in the SEIR model. We impose the same conditions in Assumptions

3.1–3.2, where the quantities $E^n(0)$ and $I^n(0)$ are replaced by $I^n(0)$ and $R^n(0)$, respectively. To distinguish the differences, we refer to these as Assumptions 3.1'–3.2'.

We first obtain the following FLLN for the fluid-scaled processes $(\bar{I}^n, \bar{R}^n)$.

Theorem 3.3. *Under Assumptions 3.1' and 3.2', we have*

$$\big(\bar{I}^n, \bar{R}^n\big) \to \big(\bar{I}, \bar{R}\big) \tag{3.29}$$

in probability in D as $n \to \infty$, where the limit process $(\bar{I}, \bar{R}) \in D^2$ is the unique solution to the system of deterministic equations: for each $t \geq 0$,

$$\bar{I}(t) = \bar{I}(0)G_0^c(t) + \lambda \int_0^t G^c(t-s)(1-\bar{I}(s)-\bar{R}(s))\bar{I}(s)ds, \tag{3.30}$$

$$\bar{R}(t) = \bar{R}(0)F_0^c(t) + \bar{I}(0)\Psi_0(t) + \lambda \int_0^t \Psi(t-s)(1-\bar{I}(s)-\bar{R}(s))\bar{I}(s)ds. \tag{3.31}$$

If G_0 and F_0 are continuous, then $\bar{I}$ and $\bar{R}$ are in C.

Remark 3.5. *In the case of independent infectious and immune times, assuming $\mathbb{E}[\xi_1] = \gamma^{-1}$ and $\mathbb{E}[\eta_1] = \mu^{-1}$ satisfy $\lambda > \gamma$, if $G_0(t) = G_e(t) := \gamma^{-1}\int_0^t G^c(s)ds$ and $F_0(t) = F_e(t) := \mu^{-1}\int_0^t F^c(s)ds$, the corresponding equilibrium distributions of G and F (see also* (2.12) *in Remark 2.8), there exists a unique nontrivial equilibrium, given by*

$$\bar{S}^* = \frac{\gamma}{\lambda}, \quad \bar{I}^* = \frac{1-\gamma/\lambda}{1+\gamma/\mu}, \quad \text{and} \quad \bar{R}^* = \frac{\gamma}{\mu}\bar{I}^*. \tag{3.32}$$

It is the same as the nontrivial equilibrium in the Markovian setting, using the ODE:

$$\bar{I}'(t) = \lambda(1-\bar{I}(t)-\bar{R}(t))\bar{I}(t) - \gamma\bar{I}(t),$$
$$\bar{R}'(t) = \gamma\bar{I}(t) - \mu\bar{R}(t),$$

and $1 = \bar{S}(t)+\bar{I}(t)+\bar{R}(t)$. From the ODE, it is straightforward to see that the equilibrium $(\bar{S}^*, \bar{I}^*, \bar{R}^*)$ satisfies the two equations

$$\lambda(1-\bar{I}^*-\bar{R}^*) = \gamma, \tag{3.33}$$

$$\mu\bar{R}^* = \gamma\bar{I}^*. \tag{3.34}$$

We now verify that these two identities are also satisfied in the general non-Markovian setting assuming $G_0 = G_e$ and $F_0 = F_e$. This implies that the equilibrium quantities satisfy

$$\bar{I}^* = \gamma\bar{I}^*\int_t^\infty G^c(s)ds + \lambda\bar{I}^*(1-\bar{I}^*-\bar{R}^*)\int_0^t G^c(s)ds,$$
$$\bar{R}^* = \mu\bar{R}^*\int_t^\infty F^c(s)ds + \bar{I}^*\Psi_0(t) + \lambda(1-\bar{I}^*-\bar{R}^*)\bar{I}^*\int_0^t \Psi(s)ds.$$

This system has the trivial solution $\bar{I}^* = \bar{R}^* = 0$. We now look for another solution. Dividing the first identity by $\bar{I}^*$ and differentiating, we recover (3.33)*, and the second identity becomes*

$$\bar{R}^* = \mu\bar{R}^*\int_t^\infty F^c(s)ds + \bar{I}^*\Psi_0(t) + \gamma\bar{I}^*\int_0^t \Psi(s)ds\,.$$

(3.34) *now follows from the identity $\gamma^{-1}\Psi_0(t) + \int_0^t \Psi(s)ds = \int_0^t F^c(s)ds$. To verify this, first note that from the definitions of Ψ_0 in the independent case, and of G_0,*

$$\gamma^{-1}\Psi_0(t) = \int_0^t F^c(t-u)G^c(u)du = \int_0^t F^c(s)ds - \int_0^t F^c(t-u)G(u)du\,.$$

It remains to note that by integration by parts and interchange of orders of integration

$$\begin{aligned}\int_0^t F^c(t-u)G(u)du &= \int_0^t \int_0^{t-u} F^c(v)dvdG(u)\\ &= \int_0^t \int_u^t F^c(v-u)dvdG(u)\\ &= \int_0^t \int_0^v F^c(v-u)dG(u)dv = \int_0^t \Psi(s)ds\end{aligned}$$

We define the diffusion-scaled processes $\hat{I}^n$ and $\hat{R}^n$ as in the SEIR model, but replacing $\bar{S} = 1-\bar{I}-\bar{R}$.

Theorem 3.4. *Under Assumptions 3.1' and 3.3', we have*

$$(\hat{I}^n, \hat{R}^n) \Rightarrow (\hat{I}, \hat{R}) \quad in \quad D^2 \quad as \quad n \to \infty. \tag{3.35}$$

where

$$\hat{I}(t) = \hat{I}(0)G_0^c(t) + \lambda \int_0^t G^c(t-s)\left(-\hat{I}(s)\bar{R}(s) + (1-\bar{I}(s)-2\bar{R}(s))\hat{R}(s)\right)ds + \hat{I}_0(t) + \hat{I}_1(t), \tag{3.36}$$

$$\begin{aligned}\hat{R}(t) = \hat{R}(0)F_0^c(t) + \hat{I}(0)\Psi_0(t) + \lambda \int_0^t \Psi(t-s)\left(-\hat{I}(s)\bar{R}(s) + (1-\bar{I}(s)-2\bar{R}(s))\hat{R}(s)\right)ds\\ + \hat{R}_{0,1}(t) + \hat{R}_{0,2}(t) + \hat{R}_1(t),\end{aligned} \tag{3.37}$$

where $\hat{I}_0(t)$, $\hat{I}_1(t)$, $\hat{R}_{0,1}(t)$ and $\hat{R}_{0,2}(t)$ are as given as $\hat{E}_0(t)$, $\hat{E}_1(t)$, $\hat{I}_{0,1}(t)$ and $\hat{I}_{0,2}(t)$, respectively, in Theorem 3.2. If the c.d.f.'s G_0 and F_0 are continuous, then the limit processes $\hat{I}_1$ and $\hat{R}_1$ have continuous sample paths. If $(\hat{I}(0), \hat{R}(0))$ is a Gaussian random vector, then $(\hat{I}, \hat{R})$ is a Gaussian process.

Remark 3.6. *If the system starts from the equilibrium as discussed in Remark 3.5, then we can define the diffusion-scaled processes $\hat{I}^n = \sqrt{n}(\bar{I}^n - \bar{I}^*)$ and $\hat{R}^n = \sqrt{n}(\bar{R}^n - \bar{R}^*)$ and the FCLT holds with the limit processes $\hat{I}$ and $\hat{R}$ as given in the above theorem where the fluid limits $\bar{I}$ and $\bar{R}$ are replaced by $\bar{I}^*$ and $\bar{R}^*$.*

Remark 3.7. *In the Markovian case, with independent infectious and immune exponential periods with parameters γ and μ, respectively, we obtain diffusion limits as in Remark 3.4 for the SEIR model. We do not repeat them for brevity. If the system starts from the equilibrium, then the FCLT holds as in Remark 3.6 with the limit processes $\hat{I}$ and $\hat{R}$ given by*

$$\hat{I}(t) = \hat{I}(0) - (\lambda\bar{R}^* + \gamma)\int_0^t \hat{I}(s)ds + \lambda(\bar{S}^* - \bar{R}^*)\int_0^t \hat{R}(s)ds + B_A\left(\lambda\bar{S}^*\bar{R}^* t\right) - B_K\left(\gamma\bar{I}^* t\right),$$

and

$$\hat{R}(t) = \hat{R}(0) + \gamma\int_0^t \hat{I}(s)ds - \mu\int_0^t \hat{R}(s)ds + B_K\left(\gamma\bar{I}^* t\right) - B_L\left(\mu\bar{R}^* t\right),$$

where B_A, B_K and B_L are independent Brownian motions, and $\bar{I}^*$ and $\bar{R}^*$ are given in (3.32).

Remark 3.8. *Suppose both the infectious and immune times are deterministic, taking values ξ and η, respectively. The remaining infectious and immune times of the initially infected and immune individuals at time 0, ξ_j^0 and η_j^0, have uniform distributions on the intervals $[0,\xi]$ and $[0,\eta]$, respectively. That is, $G(t) = \mathbf{1}(t \geq \xi)$, $F(t) = \mathbf{1}(t \geq \eta)$, for $t \geq 0$, $G_0(t) = t/\xi$ for $t \in [0,\xi]$ and $F_0(t) = t/\eta$ for $t \in [0,\eta]$. Thus we have $\Psi_0(t) = \xi^{-1}\int_0^t \mathbf{1}(t-u<\eta)du = \xi^{-1}(t-(t-\eta)^+)$, and $\Psi(t) = \mathbf{1}(\xi \leq t < \xi+\eta)$ for $t \geq 0$.*

We can write

$$I^n(t) = \sum_{j=1}^{I^n(0)} \mathbf{1}(\xi_j^0 > t) + A^n(t) - A^n((t-\xi)^+),$$

$$\begin{aligned} R^n(t) = &\sum_{j=1}^{R^n(0)} \mathbf{1}(\eta_j^0 > t) + \sum_{j=1}^{I^n(0)} \mathbf{1}((t-\eta)^+ < \xi_j^0 \le t) \\ &+ A^n((t-\xi)^+) - A^n((t-\xi-\eta)^+). \end{aligned}$$

In the FLLN, we have the fluid equations

$$\bar{I}(t) = \bar{I}(0)(1-t/\xi)^+ + \lambda \int_{((t-\xi)^+,t]} (1-\bar{I}(s)-\bar{R}(s))\bar{I}(s)ds,$$

$$\bar{R}(t) = \bar{R}(0)(1-t/\eta)^+ + \bar{I}(0)\xi^{-1}(t-(t-\eta)^+) + \lambda \int_{((t-\xi-\eta)^+,(t-\xi)^+]} (1-\bar{I}(s)-\bar{R}(s))\bar{I}(s)ds.$$

From the equation of $\bar{R}(t)$, we easily see that their equilibrium satisfies $\bar{R}^* = \bar{I}^*\eta/\xi$, and then from the equation of $\bar{I}(t)$, we obtain $\bar{I}^* = \frac{1-1/(\lambda\xi)}{1+\eta/\xi}$.

In the FCLT, we obtain

$$\hat{I}(t) = \hat{I}(0)(1-t/\xi)^+ + \lambda \int_{((t-\xi)^+,t]} \Big(-\hat{I}(s)\bar{R}(s) + (1-\bar{I}(s)-2\bar{R}(s))\hat{R}(s)\Big) ds + \hat{I}_0(t) + \hat{I}_1(t),$$

$$\begin{aligned} \hat{R}(t) = &\hat{R}(0)(1-t/\eta)^+ + \hat{I}(0)\xi^{-1}(t-(t-\eta)^+) \\ &+ \lambda \int_{((t-\xi-\eta)^+,(t-\xi)^+]} \Big(-\hat{I}(s)\bar{R}(s) + (1-\bar{I}(s)-2\bar{R}(s))\hat{R}(s)\Big) ds \\ &+ \hat{R}_{0,1}(t) + \hat{R}_{0,2}(t) + \hat{R}_1(t), \end{aligned} \tag{3.38}$$

where $\bar{I}$ and $\bar{R}$ are the fluid equations given above, and $\hat{I}_0(t)$, $\hat{I}_1(t)$, $\hat{R}_{0,1}(t)$, $\hat{R}_{0,2}(t)$ and $\hat{R}_1(t)$ have the covariance functions: for $t, t' \ge 0$,

$$\begin{aligned} \mathrm{Cov}(\hat{I}_0(t), \hat{I}_0(t')) &= \bar{I}(0)((1-t\vee t'/\xi)^+ - (1-t/\xi)^+(1-t'/\xi)^+), \\ \mathrm{Cov}(\hat{I}_1(t), \hat{I}_1(t')) &= \lambda \int_0^{t\wedge t'} \mathbf{1}(t\vee t' - s < \xi)(1-\bar{I}(s)-\bar{R}(s))\bar{I}(s)ds, \\ \mathrm{Cov}(\hat{R}_{0,1}(t), \hat{R}_{0,1}(t')) &= \bar{R}(0)((1-t\vee t'/\eta)^+ - (1-t/\eta)^+(1-t'/\eta)^+), \\ \mathrm{Cov}(\hat{R}_{0,2}(t), \hat{R}_{0,2}(t')) &= \bar{I}(0)\xi^{-1}[(t\vee t' - (t\vee t'-\eta)^+) - (t-(t-\eta)^+)(t'-(t'-\eta)^+)], \\ \mathrm{Cov}(\hat{R}_1(t), \hat{R}_1(t')) &= \lambda \int_0^{t\wedge t'} \mathbf{1}(\xi \le t\vee t' - s < \xi+\eta)(1-\bar{I}(s)-\bar{R}(s))\bar{I}(s)ds, \end{aligned}$$

and similarly for the covariances between them. If, in addition, the system starts from the equilibrium, $\bar{I}^n(0) = n\bar{I}^*$ and $\bar{R}^n(0) = n\bar{R}^*$, then the limit processes above will have the fluid quantities $\bar{I}(t)$ and $\bar{R}(t)$ replaced by $\bar{I}^*$ and $\bar{R}^*$, respectively.

4. Proof of the FLLN for the SIR model

In this section we prove Theorem 2.1.

We write the process $\bar{A}^n$ as

$$\bar{A}^n(t) = \frac{1}{\sqrt{n}}\hat{M}_A^n(t) + \bar{\Lambda}^n(t), \tag{4.1}$$

where

$$\bar{\Lambda}^n(t) := \lambda \int_0^t \bar{S}^n(s)\bar{I}^n(s)ds,$$

and

$$\hat{M}_A^n(t) := \frac{1}{\sqrt{n}}\left(A_*\left(n\bar{\Lambda}^n(t)\right) - n\bar{\Lambda}^n(t)\right). \tag{4.2}$$

The process $\{\hat{M}_A^n(t) : t \geq 0\}$ is a square-integrable martingale with respect to the filtration $\{\mathcal{F}_t^n : t \geq 0\}$ defined by

$$\mathcal{F}_t^n := \sigma\left\{I^n(0), A_*\left(n\bar{\Lambda}^n(u)\right) : 0 \leq u \leq t\right\},$$

with the predictable quadratic variation

$$\langle \hat{M}_A^n \rangle(t) = \bar{\Lambda}^n(t), \quad t \geq 0. \tag{4.3}$$

These properties are straightforward to verify; see, e.g. [17] or [8]. Note that by the simple bound

$$\bar{S}^n(t) \leq 1, \quad \bar{I}^n(t) \leq 1, \quad \forall t \geq 0, \tag{4.4}$$

we have, w.p.1., for $0 < s \leq t$,

$$0 \leq \bar{\Lambda}^n(t) - \bar{\Lambda}^n(s) \leq \lambda(t-s). \tag{4.5}$$

Lemma 4.1. *The sequence $\{(\bar{A}^n, \bar{S}^n) : n \geq 1\}$ is tight in D^2.*

Proof. By (4.5), we have $\langle \hat{M}_A^n \rangle(t) \leq \lambda t$, w.p.1. Thus, by [17, Lemma 5.8], the martingale $\{\hat{M}_A^n(t) : t \geq 0\}$ is stochastically bounded in D. Then by [17, Lemma 5.8], we have

$$\frac{1}{\sqrt{n}}\hat{M}_A^n \Rightarrow 0 \quad \text{in} \quad D \quad \text{as} \quad n \to \infty. \tag{4.6}$$

Then, by (4.1), the tightness of the sequence $\{\bar{A}^n : n \geq 1\}$ follows directly by (4.4). Since $\bar{S}^n = 1 - \bar{I}(0) - \bar{A}^n$, we obtain the tightness of $\{\bar{S}^n : n \geq 1\}$ in D immediately. □

We work with a convergent subsequence of $(\bar{A}^n, \bar{S}^n)$. We denote the limit of $\bar{A}^n$ along the subsequence by $\bar{A}$. It is clear from (4.1) that the limit $\bar{A}$ satisfies

$$\bar{A} = \lim_{n\to\infty} \bar{A}^n = \lim_{n\to\infty} \bar{\Lambda}^n = \lim_{n\to\infty} \lambda \int_0^{\cdot} \bar{S}^n(s)\bar{I}^n(s)ds. \tag{4.7}$$

and for $0 < s \leq t$, w.p.1.,

$$0 \leq \bar{A}(t) - \bar{A}(s) \leq \lambda(t-s). \tag{4.8}$$

By definition and Assumption 2.2, we have

$$\bar{S}^n = 1 - \bar{I}^n(0) - \bar{A}^n \Rightarrow \bar{S} = 1 - \bar{I}(0) - \bar{A} \ \text{ in } D, \quad \text{as} \quad n \to \infty. \tag{4.9}$$

We next consider the process $\bar{I}^n$. Recall the expression of I^n in (2.2). Let

$$\bar{I}_0^n(t) := \frac{1}{n}\sum_{j=1}^{n\bar{I}^n(0)} \mathbf{1}(\eta_j^0 > t), \quad t \geq 0.$$

We clearly have

$$\left|\bar{I}_0^n(t) - \frac{1}{n}\sum_{j=1}^{n\bar{I}(0)} \mathbf{1}(\eta_j^0 > t)\right| \leq \frac{1}{n}\sum_{j=n(\bar{I}^n(0)\wedge\bar{I}(0))}^{n(\bar{I}^n(0)\vee\bar{I}(0))} \mathbf{1}(\eta_j^0 > t), \quad t \geq 0. \tag{4.10}$$

Note that by Assumption 2.2, the right–hand side satisfies

$$\mathbb{E}\left[\frac{1}{n}\sum_{j=n(\bar{I}^n(0)\wedge\bar{I}(0))}^{n(\bar{I}^n(0)\vee\bar{I}(0))}\mathbf{1}(\eta_j^0>t)\Big|\mathcal{F}_0^n\right]\le F_0^c(t)|\bar{I}^n(0)-\bar{I}(0)|\to 0$$

in probability as $n\to\infty$. Thus, by the FLLN of empirical processes, we obtain

$$\bar{I}_0^n \Rightarrow \bar{I}_0 = \bar{I}(0)F_0^c(\cdot) \quad \text{in} \quad D \quad \text{as} \quad n\to\infty. \tag{4.11}$$

Let

$$\bar{I}_1^n(t) := \frac{1}{n}\sum_{i=1}^{n\bar{A}^n(t)}\mathbf{1}(\tau_i^n+\eta_i>t), \quad t\ge 0,$$

and its conditional expectation

$$\breve{I}_1^n(t) := \mathbb{E}[\bar{I}_1^n(t)|\mathcal{F}_t^n] = \frac{1}{n}\sum_{i=1}^{n\bar{A}^n(t)}F^c(t-\tau_i^n) = \int_0^t F^c(t-s)d\bar{A}^n(s), \quad t\ge 0.$$

By integration by parts, we have

$$\breve{I}_1^n(t) = \bar{A}^n(t) - \int_0^t \bar{A}^n(s)dF^c(t-s).$$

Here $dF^c(t-s)$ is the differential of the map $s\to F^c(t-s)$. By the continuous mapping theorem,

$$\breve{I}_1^n \Rightarrow \bar{I}_1 \quad \text{in} \quad D \quad \text{as} \quad n\to\infty, \tag{4.12}$$

where

$$\bar{I}_1(t) = \bar{A}(t) - \int_0^t \bar{A}(s)dF^c(t-s) = \int_0^t F^c(t-s)d\bar{A}(s)\,, \quad t\ge 0\,.$$

Let

$$V^n(t) := \bar{I}_1^n(t) - \breve{I}_1^n(t) = \frac{1}{n}\sum_{i=1}^{n\bar{A}^n(t)}\chi_i^n(t), \quad t\ge 0,$$

where

$$\chi_i^n(t) := \mathbf{1}(\tau_i^n+\eta_i>t) - F^c(t-\tau_i^n).$$

We next show the following lemma.

Lemma 4.2. *For any $\epsilon>0$,*

$$P\left(\sup_{t\in[0,T]}|V^n(t)|\ge\epsilon\right)\to 0 \quad \textit{as} \quad n\to\infty. \tag{4.13}$$

Proof. Note that by partitioning $[0,T]$ into intervals of length δ, that is, $[t_i,t_{i+1})$, $i=0,\dots,[T/\delta]$ with $t_0=0$, we have

$$\sup_{t\in[0,T]}|V^n(t)| \le \sup_{i=1,\dots,[T/\delta]}|V^n(t_i)| + \sup_{i=1,\dots,[T/\delta]}\sup_{u\in[0,\delta]}|V^n(t_i+u)-V^n(t_i)|. \tag{4.14}$$

It is easy to check that

$$\mathbb{E}[\chi_i^n(t)|\mathcal{F}_t^n]=0, \quad \forall i; \quad \mathbb{E}[\chi_i^n(t)\chi_j^n(t)|\mathcal{F}_t^n]=0, \quad \forall i\ne j.$$

Thus, we have

$$\mathbb{E}\big[V^n(t)^2\big|\mathcal{F}_t^n\big] = \frac{1}{n^2}\sum_{i=1}^{A^n(t)}E\big[\chi_i^n(t)^2|\mathcal{F}_t^n\big] = \frac{1}{n^2}\sum_{i=1}^{A^n(t)}F(t-\tau_i^n)F^c(t-\tau_i^n)$$

$$
\begin{aligned}
&= \frac{1}{n}\int_0^t F(t-s)F^c(t-s)d\bar{A}^n(s) \\
&= \frac{1}{n^{3/2}}\int_0^t F(t-s)F^c(t-s)d\hat{M}_A^n(s) + \frac{1}{n}\int_0^t F(t-s)F^c(t-s)d\bar{\Lambda}^n(s) \\
&\leq \frac{1}{n^{3/2}}\int_0^t F(t-s)F^c(t-s)d\hat{M}_A^n(s) + \frac{\lambda t}{n},
\end{aligned}
$$

where the inequality follows from (4.4) and (4.5). Thus

$$
\mathbb{E}(|V^n(t)|^2) \leq \frac{\lambda t}{n}, \tag{4.15}
$$

and for any $\epsilon > 0$,

$$
\mathbb{P}(|V^n(t)| > \epsilon) \leq \frac{\lambda t}{n\epsilon^2} \to 0, \quad \text{as } n \to \infty\,.
$$

We now consider $V^n(t+u) - V^n(t)$ for $t, u \geq 0$. By definition, we have

$$
\begin{aligned}
|V^n(t+u) - V^n(t)| &= \left| \frac{1}{n}\sum_{i=1}^{A^n(t+u)} \chi_i^n(t+u) - \frac{1}{n}\sum_{i=1}^{A^n(t)} \chi_i^n(t) \right| \\
&= \left| \frac{1}{n}\sum_{i=1}^{A^n(t)} (\chi_i^n(t+u) - \chi_i^n(t)) + \frac{1}{n}\sum_{i=A^n(t)}^{A^n(t+u)} \chi_i^n(t+u) \right| \\
&\leq \frac{1}{n}\sum_{i=1}^{A^n(t)} \mathbf{1}(t < \tau_i^n + \eta_i \leq t+u) + \int_0^{t+u} (F^c(t-s) - F^c(t+u-s))d\bar{\Lambda}^n(s) \\
&\qquad + \frac{1}{n}\sum_{i=A^n(t)}^{A^n(t+u)} |\chi_i^n(t+u)|.
\end{aligned}
$$

Observing that the first and second terms on the right hand are increasing in u, and that $|\chi_i^n(t)| \leq 1$, we obtain

$$
\begin{aligned}
\sup_{u\in[0,\delta]} |V^n(t+u) - V^n(t)| \leq \frac{1}{n}\sum_{i=1}^{A^n(t)} &\mathbf{1}(t < \tau_i^n + \eta_i \leq t+\delta) \\
&+ \int_0^{t+\delta} (F^c(t-s) - F^c(t+\delta-s))d\bar{A}^n(s) \\
&+ \bar{A}^n(t+\delta) - \bar{A}^n(t).
\end{aligned} \tag{4.16}
$$

Thus, for any $\epsilon > 0$,

$$
\begin{aligned}
&P\left(\sup_{u\in[0,\delta]} |V^n(t+u) - V^n(t)| > \epsilon\right) \\
&\leq \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{A^n(t)} \mathbf{1}(t < \tau_i^n + \eta_i \leq t+\delta) > \epsilon/3\right) \\
&\qquad + \mathbb{P}\left(\int_0^{t+\delta} (F^c(t-s) - F^c(t+\delta-s))d\bar{A}^n(s) > \epsilon/3\right) + \mathbb{P}\left(\bar{A}^n(t+\delta) - \bar{A}^n(t) > \epsilon/3\right) \\
&\leq \frac{9}{\epsilon^2}\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{A^n(t)} \mathbf{1}(t < \tau_i^n + \eta_i \leq t+\delta)\right)^2\right]
\end{aligned}
$$

$$+\frac{9}{\epsilon^2}\mathbb{E}\left[\left(\int_0^{t+\delta}(F^c(t-s)-F^c(t+\delta-s))d\bar{A}^n(s)\right)^2\right]+\frac{9}{\epsilon^2}\mathbb{E}\left[\left(\bar{A}^n(t+\delta)-\bar{A}^n(t)\right)^2\right]. \quad (4.17)$$

We need the following definition to treat the first term on the right hand side of (4.17).

Definition 4.1. *Let $M(ds,dz,du)$ denote a Poisson random measure (PRM) on $[0,T]\times\mathbb{R}_+\times\mathbb{R}_+$ with mean measure $\nu(ds,dz,du)=dsF(dz)du$, and $\overline{M}(ds,dz,du)$ denote the associated compensated PRM.*

We have

$$\begin{aligned}
&\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{A^n(t)}\mathbf{1}(t<\tau_i^n+\eta_i\le t+\delta)\right)^2\right]\\
&=\mathbb{E}\left[\left(\frac{1}{n}\int_0^t\int_{t-s}^{t+\delta-s}\int_0^\infty\mathbf{1}(u\le\lambda n\bar{S}^n(s)\bar{I}^n(s))M(ds,dz,du)\right)^2\right]\\
&\le 2\mathbb{E}\left[\left(\frac{1}{n}\int_0^t\int_{t-s}^{t+\delta-s}\int_0^\infty\mathbf{1}(u\le\lambda n\bar{S}^n(s)\bar{I}^n(s))\overline{M}(ds,dz,du)\right)^2\right]\\
&\qquad+2\mathbb{E}\left[\left(\int_0^t(F^c(t-s)-F^c(t+\delta-s))d\bar{\Lambda}^n(s)\right)^2\right]\\
&=\frac{2}{n}\mathbb{E}\left[\int_0^t(F^c(t-s)-F^c(t+\delta-s))d\bar{\Lambda}^n(s)\right]\\
&\qquad+2\mathbb{E}\left[\left(\int_0^t(F^c(t-s)-F^c(t+\delta-s))d\bar{\Lambda}^n(s)\right)^2\right]\\
&\le\frac{2}{n}\lambda\int_0^{t+\delta}(F^c(t-s)-F^c(t+\delta-s))ds\\
&\qquad+2\left(\lambda\int_0^t(F^c(t-s)-F^c(t+\delta-s))ds\right)^2 \qquad (4.18)
\end{aligned}$$

The last inequality follows from (4.5). The first term on the right hand converges to zero as $n\to\infty$, and for the second term, as will be shown below in Lemma 4.3, by Assumption 2.1, we have

$$\frac{1}{\delta}\left(\int_0^t(F^c(t-s)-F^c(t+\delta-s))ds\right)^2\to 0\quad\text{as}\quad\delta\to 0. \qquad (4.19)$$

For the second term on the right hand side of (4.17), by (4.1), we have

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^{t+\delta}(F^c(t-s)-F^c(t+\delta-s))d\bar{A}^n(s)\right)^2\right]\\
&\le 2\mathbb{E}\left[\left(\frac{1}{\sqrt{n}}\int_0^{t+\delta}(F^c(t-s)-F^c(t+\delta-s))d\hat{M}_A^n(s)\right)^2\right]\\
&\qquad+2\mathbb{E}\left[\left(\int_0^{t+\delta}(F^c(t-s)-F^c(t+\delta-s))d\bar{\Lambda}^n(s)\right)^2\right].
\end{aligned}$$

By (4.6), the first term converges to zero as $n\to\infty$. By (4.5), the second term is bounded by

$$2\left(\lambda\int_0^{t+\delta}(F^c(t-s)-F^c(t+\delta-s))ds\right)^2.$$

to which (4.19) again applies.

By (4.1) and (4.5), we have

$$\bar{A}^n(t+\delta) - \bar{A}^n(t) \le \frac{1}{\sqrt{n}}(\hat{M}^n(t+\delta) - \hat{M}^n(t)) + \lambda\delta.$$

Thus, for the third term on the right hand side of (4.17), we have

$$\mathbb{E}\left[\left(\bar{A}^n(t+\delta) - \bar{A}^n(t)\right)^2\right] \le 2\mathbb{E}\left[\left(\frac{1}{\sqrt{n}}(\hat{M}^n(t+\delta) - \hat{M}^n(t))\right)^2\right] + 2\lambda^2\delta^2. \tag{4.20}$$

Again, by (4.6), the first term converges to zero as $n \to \infty$.

By (4.14), we have for $\delta > 0$,

$$\mathbb{P}\left(\sup_{t\in[0,T]} |V^n(t)| \ge \epsilon\right) \le \left[\frac{T}{\delta}\right] \sup_{t\in[0,T]} \mathbb{P}\left(|V^n(t)| \ge \epsilon\right) + \left[\frac{T}{\delta}\right] \sup_{t\in[0,T]} \mathbb{P}\left(\sup_{u\in[0,\delta]} |V^n(t+u) - V^n(t)| > \epsilon\right). \tag{4.21}$$

The first term converges to zero as $n \to \infty$ by (4.15). By (4.17)–(4.20) and the above arguments, we obtain

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

Therefore, we have shown that (4.13) holds. □

It remains to establish the following technical Lemma.

Lemma 4.3. *The convergence in* (4.19) *follows from Assumption 2.1.*

Proof. Recalling the notations in Assumption 2.1, it suffices to prove (4.19) in both cases $F = F_1$ and $F = F_2$. Consider first the case $F = F_1$. We have

$$\begin{aligned}\left(\int_0^t (F^c(t-s) - F^c(t+\delta-s))ds\right)^2 &= \left(\int_0^t (F(r+\delta) - F(r))dr\right)^2 \\ &= \left(\sum_i a_i \int_0^t \left(\mathbf{1}(r+\delta \ge t_i) - \mathbf{1}(r \ge t_i)\right) dr\right)^2 \\ &\le \delta^2,\end{aligned}$$

since $\sum_i a_i \le 1$, from which (4.19) follows. In the case $F = F_2$, we have

$$\left(\int_0^t (F(r+\delta) - F(r))dr\right)^2 \le c^2 t^2 \delta^{1+2\theta},$$

from which again (4.19) follows. □

By the convergence of $\breve{I}_1^n$ in (4.12) and Lemma 4.2, we obtain

$$\bar{I}_1^n(t) \Rightarrow \bar{I}_1(t) = \int_0^t F^c(t-s)d\bar{A}(s) \quad \text{in} \quad D \quad \text{as} \quad n \to \infty.$$

Combining this with (4.11), we have

$$\bar{I}^n = \bar{I}_0^n + \bar{I}_1^n \Rightarrow \bar{I} := \bar{I}_0 + \bar{I}_1 = \bar{I}(0)F_0^c(\cdot) + \int_0^{\cdot} F^c(\cdot - s)d\bar{A}(s)$$

in D as $n \to \infty$.

We now show the joint convergence

$$(\bar{S}^n, \bar{I}^n) \Rightarrow (\bar{S}, \bar{I}) \quad \text{in} \quad D^2 \quad \text{as} \quad n \to \infty. \tag{4.22}$$

We first have the joint convergence of $\bar{I}^n(0)$, and the two terms in the expression of $\bar{I}^n_0(t)$ in (4.10) since $\bar{I}^n(0)$ and the first term are independent and the second term converges to zero. We then have the joint convergence of $\bar{A}^n$ and $\breve{I}^n_1$, since we can apply the continuous mapping theorem applied to the map $x \in D \to (x, x - \int_0^\cdot x(s)dF^c(\cdot - s)) \in D^2$. Then the claim in (4.22) holds by independence of these two groups of processes, combined with (4.13).

Thus we obtain

$$\int_0^\cdot \bar{S}^n(s)\bar{I}^n(s)ds \Rightarrow \int_0^\cdot \bar{S}(s)\bar{I}(s)ds \quad \text{in} \quad D \quad \text{as} \quad n \to \infty. \tag{4.23}$$

By (4.1) and (4.6), this implies that

$$\bar{A}^n \Rightarrow \bar{A} = \lambda \int_0^\cdot \bar{S}(s)\bar{I}(s)ds \quad \text{in} \quad D \quad \text{as} \quad n \to \infty.$$

Therefore, the limits $\bar{S}$ and $\bar{I}$ satisfy the integral equations given in (2.3) and (2.4).

We next prove uniqueness of the solution to the system of equations (2.3) and (2.4). The two equations (2.3) and (2.4) can be regarded as Volterra integral equations of the second kind for two functions. For uniqueness, suppose there are two solutions $(\bar{S}_1, \bar{I}_1)$ and $(\bar{S}_2, \bar{I}_2)$. Then we have

$$\begin{aligned}
\bar{S}_1(t) - \bar{S}_2(t) &= -\lambda \int_0^t \Big((\bar{S}_1(s) - \bar{S}_2(s))\bar{I}_1(s) + \bar{S}_2(s)(\bar{I}_1(s) - \bar{I}_2(s)) \Big) ds, \\
\bar{I}_1(t) - \bar{I}_2(t) &= \lambda \int_0^t F^c(t-s)\Big((\bar{S}_1(s) - \bar{S}_2(s))\bar{I}_1(s) + \bar{S}_2(s)(\bar{I}_1(s) - \bar{I}_2(s)) \Big) ds.
\end{aligned}$$

Hence,

$$|\bar{S}_1(t) - \bar{S}_2(t)| + |\bar{I}_1(t) - \bar{I}_2(t)| \le 2\lambda \int_0^t \Big(|\bar{S}_1(s) - \bar{S}_2(s)| + |\bar{I}_1(s) - \bar{I}_2(s)| \Big) ds,$$

where we use the simple bounds $\bar{S}_i(s) \le 1$ and $\bar{I}_i(s) \le 1$. The uniqueness follows from applying Gronwall's inequality.

Since the system of integral equations (2.3) and (2.4) has a unique deterministic solution (existence is easily established by a standard Picard iteration argument, identical to the classical one for Lipschitz ODEs), the whole sequence converges, and we have convergence in probability.

Remark 4.1. Convergence in $L^2(0,T)$ *Here we reconsider the above proof, without requiring Assumption 2.1, which has been used in the proof of Lemma 4.3, i.e., it was necessary only for the proof of Lemma 4.2. We first note that Lemma 4.1 still holds true, hence $\bar{S}^n \Rightarrow \bar{S}$ in D, at least along a subsequence. We now consider $\bar{I}^n$.*

It clearly follows from (4.15) that

$$V^n \to 0 \quad \textit{in} \quad L^2([0,T] \times \Omega) \quad \textit{as} \quad n \to \infty.$$

This, together with (4.11) and (4.12), implies that

$$\bar{I}^n \Rightarrow \bar{I} \quad \textit{in} \quad L^2(0,T) \quad \textit{as} \quad n \to \infty,$$

since convergence in law in D implies convergence in law in $L^2(0,T)$. The joint convergence of $(\bar{S}^n, \bar{I}^n)$, and that of the whole sequence of the triplets $(\bar{S}^n, \bar{I}^n, \bar{R}^n)$ as stated in Remark 2.2 is then easily established, following some of the arguments from the above proof.

5. Proof of the FCLT for the SIR model

In this section we prove Theorem 2.2. Recall the definitions of the diffusion-scaled processes $(\hat{S}^n, \hat{I}^n, \hat{R}^n)$ in (2.6), and $\hat{M}_A^n$ defined in (4.2). We also define

$$\hat{A}^n(t) := \sqrt{n}\left(\bar{A}^n(t) - \bar{A}(t)\right) = \sqrt{n}\left(\bar{A}^n(t) - \lambda\int_0^t \bar{S}(s)\bar{I}(s)ds\right).$$

Note that under Assumption 2.3, we have $\bar{I}^n(0) \Rightarrow \bar{I}(0)$ in $\mathbb{R}$ as $n \to \infty$, and thus the convergence of the fluid-scaled processes holds in Theorem 2.1. This is taken as given in the proceeding proof of the FCLT.

By the definitions of the diffusion-scaled processes in (2.6), we have

$$\hat{A}^n(t) = \hat{M}_A^n(t) + \lambda\int_0^t (\hat{S}^n(s)\bar{I}^n(s) + \bar{S}(s)\hat{I}^n(s))ds, \tag{5.1}$$

$$\begin{aligned}\hat{S}^n(t) &= -\hat{I}^n(0) - \hat{A}^n(t)\\ &= -\hat{I}^n(0) - \hat{M}_A^n(t) - \lambda\int_0^t (\hat{S}^n(s)\bar{I}^n(s) + \bar{S}(s)\hat{I}^n(s))ds,\end{aligned} \tag{5.2}$$

$$\begin{aligned}\hat{I}^n(t) &= \hat{I}^n(0)F_0^c(t) + \hat{I}_0^n(t) + \hat{I}_1^n(t)\\ &\qquad + \lambda\int_0^t F^c(t-s)\left(\hat{S}^n(s)\bar{I}^n(s) + \bar{S}(s)\hat{I}^n(s)\right)ds,\end{aligned} \tag{5.3}$$

and

$$\begin{aligned}\hat{R}^n(t) &= \hat{I}^n(0)F_0(t) + \hat{R}_0^n(t) + \hat{R}_1^n(t)\\ &\qquad + \lambda\int_0^t F(t-s)(\hat{S}^n(s)\bar{I}^n(s) + \bar{S}(s)\hat{I}^n(s))ds,\end{aligned} \tag{5.4}$$

where

$$\hat{I}_0^n(t) := \frac{1}{\sqrt{n}}\sum_{j=1}^{n\bar{I}^n(0)}\left(\mathbf{1}(\eta_j^0 > t) - F_0^c(t)\right), \tag{5.5}$$

$$\hat{I}_1^n(t) := \frac{1}{\sqrt{n}}\sum_{i=1}^{n\bar{A}^n(t)}\mathbf{1}(\tau_i^n + \eta_i > t) - \lambda\sqrt{n}\int_0^t F^c(t-s)\bar{S}^n(s)\bar{I}^n(s)ds, \tag{5.6}$$

and

$$\hat{R}_0^n(t) := \frac{1}{\sqrt{n}}\sum_{j=1}^{n\bar{I}^n(0)}\left(\mathbf{1}(\eta_j^0 \le t) - F_0(t)\right), \tag{5.7}$$

$$\hat{R}_1^n(t) := \frac{1}{\sqrt{n}}\sum_{i=1}^{n\bar{A}^n(t)}\mathbf{1}(\tau_i^n + \eta_i \le t) - \lambda\sqrt{n}\int_0^t F(t-s)\bar{S}^n(s)\bar{I}^n(s)ds. \tag{5.8}$$

We first establish the following joint convergence of the initial quantities.

Lemma 5.1. *Under Assumption 2.3, we have*

$$(\hat{I}^n(0)F_0^c(\cdot), \hat{I}^n(0)F_0(\cdot), \hat{I}_0^n, \hat{R}_0^n) \Rightarrow \left(\hat{I}(0)F_0^c(\cdot), \hat{I}(0)F_0(\cdot), \hat{I}_0, \hat{R}_0\right) \tag{5.9}$$

in D^4 as $n \to \infty$, where the limit processes $\hat{I}_0$ and $\hat{R}_0$ are as defined in Theorem 2.2.

Proof. We define

$$\widetilde{I}_0^n(t) := \frac{1}{\sqrt{n}} \sum_{j=1}^{n\bar{I}(0)} \big(\mathbf{1}(\eta_j^0 > t) - F_0^c(t)\big),$$

$$\widetilde{R}_0^n(t) := \frac{1}{\sqrt{n}} \sum_{j=1}^{n\bar{I}(0)} \big(\mathbf{1}(\eta_j^0 \le t) - F_0(t)\big).$$

By the FCLT for empirical processes, see, e.g., [6, Theorem 14.3], we have the joint convergence

$$(\hat{I}^n(0)F_0^c(\cdot), \hat{I}^n(0)F_0(\cdot), \widetilde{I}_0^n, \widetilde{R}_0^n) \Rightarrow \Big(\hat{I}(0)F_0^c(\cdot), \hat{I}(0)F_0(\cdot), \hat{I}_0, \hat{R}_0\Big)$$

in D^4 as $n \to \infty$. The claim then follows by showing that $\widetilde{I}_0^n - \hat{I}_0^n \Rightarrow 0$ in D as $n \to \infty$, and $\widetilde{R}_0^n - \hat{R}_0^n \Rightarrow 0$ in D as $n \to \infty$. We focus on $\widetilde{I}_0^n - \hat{I}_0^n \Rightarrow 0$. We have for each $t \ge 0$, $\mathbb{E}[\widetilde{I}_0^n(t) - \hat{I}_0^n(t)] = 0$ and

$$\mathbb{E}[|\widetilde{I}_0^n(t) - \hat{I}_0^n(t)|^2] = F_0^c(t)F_0(t)E[|\bar{I}^n(0) - \bar{I}(0)|] \to 0 \quad \text{as} \quad n \to \infty,$$

where the convergence follows from Assumption 2.3. It then suffices to show that $\{\widetilde{I}_0^n - \hat{I}_0^n : n \ge 1\}$ is tight. We have

$$\begin{aligned} \operatorname{sign}(\bar{I}(0) - \bar{I}^n(0)) \Big(\widetilde{I}_0^n(t) - \hat{I}_0^n(t)\Big) &= \frac{1}{\sqrt{n}} \sum_{j=n(\bar{I}^n(0)\wedge\bar{I}(0))}^{n(\bar{I}^n(0)\vee\bar{I}(0))} \big(\mathbf{1}(\eta_j^0 > t) - F_0^c(t)\big) \\ &= |\hat{I}^n(0)|F_0(t) - \frac{1}{\sqrt{n}} \sum_{j=n(\bar{I}^n(0)\wedge\bar{I}(0))}^{n(\bar{I}^n(0)\vee\bar{I}(0))} \mathbf{1}(\eta_j^0 \le t). \end{aligned}$$

By Assumption 2.3, the first term on the right hand side is tight. Denoting the second term by $\Theta_0^n(t)$, since it is increasing in t, by the Corollary on page 83 in [6], see also the use of (4.14) in the proof of Lemma 4.2 above, its tightness will follow from the fact that for any $\epsilon > 0$,

$$\limsup_{n\to\infty} \frac{1}{\delta}\mathbb{P}\big(\big|\Theta_0^n(t+\delta) - \Theta_0^n(t)\big| \ge \epsilon\big) \to 0 \quad \text{as} \quad \delta \to 0.$$

This is immediate since by Assumption 2.3,

$$\mathbb{E}\left[\big|\Theta_0^n(t+\delta) - \Theta_0^n(t)\big|^2\right] = E[|\bar{I}^n(0) - \bar{I}(0)|]|F_0(t+\delta) - F_0(t)| \to 0 \quad \text{as} \quad n \to \infty.$$

This completes the proof. $\square$

Recall the PRM $M(ds, dz, du)$ and the compensated PRM $\overline{M}(ds, dz, du)$ in Definition 4.1.

Definition 5.1. *Let $M_1(ds, dz, du)$ be the PRM on $[0,T]\times\mathbb{R}_+\times\mathbb{R}_+$ with mean measure $\tilde{\nu}(ds, dz, du) = dsF_s(dz)du$, where $F_s((a,b]) = F((a+s, b+s])$. Denote the associated compensated PRM by $\widetilde{M}(ds, dz, du)$.*

We can rewrite the processes $\hat{I}_1^n$ and $\hat{R}_1^n$ as

$$\begin{aligned} \hat{I}_1^n(t) &= \frac{1}{\sqrt{n}} \int_0^t \int_{t-s}^\infty \int_0^\infty \varphi_n(s,u)\overline{M}(ds,dz,du) = \frac{1}{\sqrt{n}} \int_0^t \int_t^\infty \int_0^\infty \varphi_n(s,u)\widetilde{M}(ds,dz,du), \\ \hat{R}_1^n(t) &= \frac{1}{\sqrt{n}} \int_0^t \int_0^{t-s} \int_0^\infty \varphi_n(s,u)\overline{M}(ds,dz,du) = \frac{1}{\sqrt{n}} \int_0^t \int_0^t \int_0^\infty \varphi_n(s,u)\widetilde{M}(ds,dz,du), \end{aligned}$$

where

$$\varphi_n(s,u) = \mathbf{1}\left(u \le n\lambda\bar{S}^n(s)\bar{I}^n(s)\right).$$

We also observe that the process $\hat{M}_A^n$ can also be represented by the same PRMs:

$$\hat{M}_A^n(t) = \frac{1}{\sqrt{n}}\int_0^t\int_0^\infty\int_0^\infty \varphi_n(s,u)\overline{M}(ds,dz,du) = \frac{1}{\sqrt{n}}\int_0^t\int_0^\infty\int_0^\infty \varphi_n(s,u)\widetilde{M}(ds,dz,du),$$

and that

$$\hat{M}_A^n(t) = \hat{I}_1^n(t) + \hat{R}_1^n(t), \quad t \ge 0.$$

We define the auxiliary processes $\widetilde{I}_1^n$ and $\widetilde{R}_1^n$ by

$$\begin{aligned}
\widetilde{I}_1^n(t) &= \frac{1}{\sqrt{n}}\int_0^t\int_t^\infty\int_0^\infty \widetilde{\varphi}_n(s,u)\widetilde{M}(ds,dz,du),\\
\widetilde{R}_1^n(t) &= \frac{1}{\sqrt{n}}\int_0^t\int_0^t\int_0^\infty \widetilde{\varphi}_n(s,u)\widetilde{M}(ds,dz,du),\\
\widetilde{M}_A^n(t) &= \frac{1}{\sqrt{n}}\int_0^t\int_0^\infty\int_0^\infty \widetilde{\varphi}_n(s,u)\widetilde{M}(ds,dz,du),
\end{aligned}$$

where

$$\widetilde{\varphi}_n(s,u) = \mathbf{1}\left(u \le n\lambda \bar{S}(s)\bar{I}(s)\right).$$

Note that in the definitions of $\widetilde{I}_1^n(t)$ and $\widetilde{R}_1^n(t)$, we have replaced $\bar{S}^n(s)$ and $\bar{I}^n(s)$ in the integrands $\varphi_n(s,u)$ by the deterministic fluid functions $\bar{S}(s)$ and $\bar{I}(s)$. Also, it is clear that

$$\widetilde{M}_A^n(t) = \widetilde{I}_1^n(t) + \widetilde{R}_1^n(t), \quad t \ge 0.$$

We first prove the following result.

Lemma 5.2.

$$\sup_n \sup_{t\in[0,T]} \mathbb{E}[|\hat{S}^n(t)|^2] < \infty, \quad \sup_n \sup_{t\in[0,T]} \mathbb{E}[|\hat{I}^n(t)|^2] < \infty, \quad \sup_n \sup_{t\in[0,T]} \mathbb{E}[|\hat{R}^n(t)|^2] < \infty.$$

Proof. We have

$$\sup_{t\in[0,T]} \mathbb{E}[\hat{M}_A^n(t)^2] \le \lambda T.$$

It is clear that there exists a constant C such that for all $n \ge 1$,

$$\sup_{t\in[0,T]} \mathbb{E}[(\hat{I}^n(0)F_0^c(t))^2] \le \mathbb{E}[\hat{I}^n(0)^2] \le C,$$

$$\sup_{t\in[0,T]} \mathbb{E}[(\hat{I}_0^n(t))^2] = \sup_{t\in[0,T]} \mathbb{E}[\bar{I}^n(0)]F_0(t)F_0^c(t) \le \mathbb{E}[\bar{I}^n(0)] \le C,$$

and

$$\sup_{t\in[0,T]} \mathbb{E}[(\hat{I}_1^n(t))^2] = \sup_{t\in[0,T]} \lambda\int_0^t F^c(t-s)\bar{S}^n(s)\bar{I}^n(s)ds \le \lambda T.$$

Then by taking the square of the representations of $\hat{S}^n(t)$ in (5.2) and $\hat{I}^n(t)$ in (5.4), then using Cauchy-Schwartz inequality and the simple bounds $\bar{I}^n(t) \le 1$ and $\bar{S}(t) \le 1$, we can apply Gronwall's inequality to conclude the claim. □

We next show that the differences of the processes $\hat{M}_A^n, \hat{R}_1^n, \hat{I}_1^n$ with their corresponding $\widetilde{M}_A^n, \widetilde{R}_1^n, \widetilde{I}_1^n$ are asymptotically negligible, stated in the next Lemma.

Lemma 5.3. *Under Assumptions 2.3 and 2.1,*

$$(\hat{M}_A^n - \widetilde{M}_A^n, \hat{R}_1^n - \widetilde{R}_1^n, \hat{I}_1^n - \widetilde{I}_1^n) \Rightarrow 0 \quad in \quad D^3 \quad as \quad n \to \infty.$$

Proof. It suffices to prove the convergence of each coordinate separately. We focus on the convergence $\hat{R}_1^n - \widetilde{R}_1^n \Rightarrow 0$, since the convergence $\hat{M}_A^n - \widetilde{M}_A^n$ follows similarly, and then the convergence $\hat{I}_1^n - \widetilde{I}_1^n \Rightarrow 0$ follows by the facts that $\hat{M}_A^n(t) = \hat{I}_1^n(t) + \hat{R}_1^n(t)$ and $\widetilde{M}_A^n(t) = \widetilde{I}_1^n(t) + \widetilde{R}_1^n(t)$, for each $t \geq 0$.

Let $\widetilde{\Xi}^n := \hat{R}_1^n - \widetilde{R}_1^n$. It is easy to see that for each $t \geq 0$, $\mathbb{E}[\widetilde{\Xi}_1^n(t)] = 0$, and

$$\mathbb{E}\left[\widetilde{\Xi}^n(t)^2\right] = \int_0^t F(t-s)\mathbb{E}\left[|\bar{S}^n(s)\bar{I}^n(s) - \bar{S}(s)\bar{I}(s)|\right] ds \to 0 \quad \text{as} \quad n \to \infty,$$

where the convergence holds by Theorem 2.1 and the dominated convergence theorem. Then it suffices to show that the sequence $\{\widetilde{\Xi}^n : n \geq 1\}$ is tight. Note that $\widetilde{\Xi}^n$ can be written as $\widetilde{\Xi}^n(t) = \widetilde{\Xi}_1^n(t) - \widetilde{\Xi}_2^n(t)$, where

$$\widetilde{\Xi}_1^n(t) := \frac{1}{\sqrt{n}} \int_0^t \int_0^t \int_{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\wedge\bar{S}(s)\bar{I}(s))}^{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\vee\bar{S}(s)\bar{I}(s))} \operatorname{sign}(\bar{S}^n(s)\bar{I}^n(s) - \bar{S}(s)\bar{I}(s))M_1(ds,dz,du),$$

$$\widetilde{\Xi}_2^n(t) := \lambda\sqrt{n}\int_0^t F(t-s)\big(\bar{S}^n(s)\bar{I}^n(s) - \bar{S}(s)\bar{I}(s)\big)ds.$$

Both processes $\widetilde{\Xi}_1^n(t)$ and $\widetilde{\Xi}_2^n(t)$ are differences of two processes, each increasing in t, that is,

$$\begin{aligned}\widetilde{\Xi}_1^n(t) = {} & \frac{1}{\sqrt{n}} \int_0^t \int_0^t \int_{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\wedge\bar{S}(s)\bar{I}(s))}^{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\vee\bar{S}(s)\bar{I}(s))} \mathbf{1}(\bar{S}^n(s)\bar{I}^n(s) - \bar{S}(s)\bar{I}(s) > 0)M_1(ds,dz,du) \\ & - \frac{1}{\sqrt{n}} \int_0^t \int_0^t \int_{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\wedge\bar{S}(s)\bar{I}(s))}^{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\vee\bar{S}(s)\bar{I}(s))} \mathbf{1}(\bar{S}^n(s)\bar{I}^n(s) - \bar{S}(s)\bar{I}(s) < 0)M_1(ds,dz,du),\end{aligned}$$

and

$$\widetilde{\Xi}_2^n(t) = \lambda\sqrt{n}\int_0^t F(t-s)\big(\bar{S}^n(s)\bar{I}^n(s) - \bar{S}(s)\bar{I}(s)\big)^+ ds - \lambda\sqrt{n}\int_0^t F(t-s)\big(\bar{S}^n(s)\bar{I}^n(s) - \bar{S}(s)\bar{I}(s)\big)^+ ds.$$

Define Ξ_1^n and Ξ_2^n by

$$\Xi_1^n(t) := \frac{1}{\sqrt{n}} \int_0^t \int_0^t \int_{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\wedge\bar{S}(s)\bar{I}(s))}^{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\vee\bar{S}(s)\bar{I}(s))} M_1(ds,dz,du),$$

and

$$\Xi_2^n(t) := \lambda\sqrt{n}\int_0^t F(t-s)\big|\bar{S}^n(s)\bar{I}^n(s) - \bar{S}(s)\bar{I}(s)\big|ds.$$

Tightness of $\Xi_1^n(t)$ and $\Xi_2^n(t)$ implies tightness of the four components in the above expressions of $\widetilde{\Xi}_1^n(t)$ and $\widetilde{\Xi}_2^n(t)$. By the increasing property of $\Xi_1^n(t)$ and $\Xi_2^n(t)$, we only need to verify the following (see the Corollary on page 83 in [6] or the use of (4.14) in the proof of Lemma 4.2): for any $\epsilon > 0$, and $i = 1,2$,

$$\limsup_{n\to\infty} \frac{1}{\delta}\mathbb{P}\big(\big|\Xi_i^n(t+\delta) - \Xi_i^n(t)\big| \geq \epsilon\big) \to 0 \quad \text{as} \quad \delta \to 0. \tag{5.10}$$

For the process $\Xi_2^n(t)$, we have

$$\begin{aligned}&\mathbb{E}\big[|\Xi_2^n(t+\delta) - \Xi_2^n(t)|^2\big] \\ &= \mathbb{E}\left[\lambda^2\left(\int_t^{t+\delta} F(t+\delta-s)\Delta^n(s)ds + \int_0^t (F(t+\delta-s) - F(t-s))\Delta^n(s)ds\right)^2\right] \\ &\leq 2\lambda^2\mathbb{E}\left[\left(\int_t^{t+\delta} F(t+\delta-s)\Delta^n(s)ds\right)^2\right] \\ &\quad + 2\lambda^2\mathbb{E}\left[\left(\int_0^t (F(t+\delta-s) - F(t-s))\Delta^n(s)ds\right)^2\right],\end{aligned} \tag{5.11}$$

where

$$\Delta^n(s) := \sqrt{n}\left|\bar{S}^n(s)\bar{I}^n(s) - \bar{S}(s)\bar{I}(s)\right| = |\hat{S}^n(s)\bar{I}^n(s) + \bar{S}(s)\hat{I}^n(s)| \le |\hat{S}^n(s)| + |\hat{I}^n(s)|. \tag{5.12}$$

For the first term on the right hand side of (5.11), by Cauchy-Schwartz inequality, we have

$$\mathbb{E}\left[\left(\int_t^{t+\delta} F(t+\delta-s)\Delta^n(s)ds\right)^2\right] \le \delta\int_t^{t+\delta}\mathbb{E}[\Delta^n(s)^2]ds \le \delta^2\sup_{s\in[0,T]}\mathbb{E}[\Delta^n(s)^2].$$

For the second term, we need

$$\limsup_{n\to\infty}\frac{1}{\delta}\mathbb{E}\left[\left(\int_0^t (F(t+\delta-s)-F(t-s))\Delta^n(s)ds\right)^2\right] \to 0 \quad \text{as} \quad \delta\to 0. \tag{5.13}$$

This is implied by Assumption 2.1, Lemma 5.2 and (5.12), as we now show, by an argument which slightly extends that in Lemma 4.3. If $F = F_1$, we have

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^t (F(t+\delta-s)-F(t-s))\Delta^n(s)ds\right)^2\right]\\
&\le \sum_i a_i\mathbb{E}\left[\left(\int_{t_i-\delta}^{t_i}\Delta_n(t-r)dr\right)^2\right]\\
&\le \delta\sum_i a_i\mathbb{E}\left[\int_{t_i-\delta}^{t_i}\Delta_n(t-r)^2dr\right] \le \delta^2\sup_{s\in[0,T]}\mathbb{E}[\Delta^n(s)^2].
\end{aligned}$$

If $F = F_2$, we have

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^t (F(t+\delta-s)-F(t-s))\Delta^n(s)ds\right)^2\right]\\
&\le t\mathbb{E}\left[\int_0^t (F(t+\delta-s)-F(t-s))^2\Delta^n(s)^2ds\right] \le T^2c\delta^{1+2\theta}\sup_{s\in[0,T]}\mathbb{E}[\Delta^n(s)^2].
\end{aligned}$$

Thus, (5.10) holds for $\Xi_2^n(t)$.

For the process $\Xi_1^n(t)$, we have

$$\begin{aligned}
&\mathbb{E}\left[|\Xi_1^n(t+\delta)-\Xi_1^n(t)|^2\right]\\
&= \mathbb{E}\Bigg[\Bigg(\frac{1}{\sqrt{n}}\int_t^{t+\delta}\int_0^{t+\delta}\int_{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\wedge\bar{S}(s)\bar{I}(s))}^{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\vee\bar{S}(s)\bar{I}(s))} M_1(ds,dz,du)\\
&\quad + \frac{1}{\sqrt{n}}\int_0^t\int_t^{t+\delta}\int_{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\wedge\bar{S}(s)\bar{I}(s))}^{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\vee\bar{S}(s)\bar{I}(s))} M_1(ds,dz,du)\Bigg)^2\Bigg]\\
&\le 2\mathbb{E}\left[\left(\frac{1}{\sqrt{n}}\int_t^{t+\delta}\int_0^{t+\delta}\int_{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\wedge\bar{S}(s)\bar{I}(s))}^{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\vee\bar{S}(s)\bar{I}(s))} M_1(ds,dz,du)\right)^2\right]\\
&\quad + 2\mathbb{E}\left[\left(\frac{1}{\sqrt{n}}\int_0^t\int_t^{t+\delta}\int_{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\wedge\bar{S}(s)\bar{I}(s))}^{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\vee\bar{S}(s)\bar{I}(s))} M_1(ds,dz,du)\right)^2\right]\\
&=: B_1^n + B_2^n.
\end{aligned}$$

Note that we can write

$$\frac{1}{\sqrt{n}}\int_t^{t+\delta}\int_0^{t+\delta}\int_{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\wedge\bar{S}(s)\bar{I}(s))}^{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\vee\bar{S}(s)\bar{I}(s))} M_1(ds,dz,du)$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{n}} \int_t^{t+\delta} \int_0^{t+\delta} \int_{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\wedge \bar{S}(s)\bar{I}(s))}^{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\vee \bar{S}(s)\bar{I}(s))} \widetilde{M}(ds,dz,du) \\
&\quad + \lambda \int_t^{t+\delta} F(t+\delta-s)\Delta^n(s)ds.
\end{aligned}
$$

Thus, we have the following bound

$$
\begin{aligned}
B_1^n &\le 2\mathbb{E}\left[\left(\frac{1}{\sqrt{n}} \int_t^{t+\delta} \int_0^{t+\delta} \int_{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\wedge \bar{S}(s)\bar{I}(s))}^{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\vee \bar{S}(s)\bar{I}(s))} \widetilde{M}(ds,dz,du)\right)^2\right] \\
&\quad + 2\mathbb{E}\left[\left(\lambda \int_t^{t+\delta} F(t+\delta-s)\Delta^n(s)ds\right)^2\right] \\
&\le 2\lambda \int_t^{t+\delta} F(t+\delta-s)\mathbb{E}\left[|\bar{S}^n(s)\bar{I}^n(s)-\bar{S}(s)\bar{I}(s)|\right]ds \\
&\quad + 2\lambda^2\delta^2 \sup_{s\in[0,T]} \mathbb{E}[|\Delta^n(s)|^2]. \qquad (5.14)
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
B_2^n &\le 2\mathbb{E}\left[\left(\frac{1}{\sqrt{n}} \int_0^{t} \int_t^{t+\delta} \int_{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\wedge \bar{S}(s)\bar{I}(s))}^{n\lambda(\bar{S}^n(s)\bar{I}^n(s)\vee \bar{S}(s)\bar{I}(s))} d\widetilde{M}(ds,dz,du)\right)^2\right] \\
&\quad + 2\mathbb{E}\left[\left(\lambda \int_0^{t} (F(t+\delta-s)-F(t-s))\Delta^n(s)ds\right)^2\right] \\
&\le 2\lambda \int_0^{t} (F(t+\delta-s)-F(t-s))\mathbb{E}\left[|\bar{S}^n(s)\bar{I}^n(s)-\bar{S}(s)\bar{I}(s)|\right]ds \\
&\quad + 2\lambda^2 \mathbb{E}\left[\left(\int_0^{t} (F(t+\delta-s)-F(t-s))\Delta^n(s)ds\right)^2\right]. \qquad (5.15)
\end{aligned}
$$

It is straightforward that the first terms on the right hand sides of (5.14) and (5.15) converges to zero as $n\to\infty$ since $\mathbb{E}\left[|\bar{S}^n(s)\bar{I}^n(s)-\bar{S}(s)\bar{I}(s)|\right]\to 0$ as $n\to\infty$ by Theorem 2.1, and by the dominated convergence theorem. Thus, by (5.13), we have shown (5.10) for $\Xi_1^n(t)$. This completes the proof. ☐

Let

$$\mathcal{G}_t^A := \sigma\left\{\widetilde{M}([0,u]\times\mathbb{R}_+^2) : 0\le u\le t\right\}, \quad t\ge 0,$$

and

$$\mathcal{G}_t^R := \sigma\left\{\widetilde{M}([0,u]\times[0,u]\times\mathbb{R}_+) : 0\le u\le t\right\}, \quad t\ge 0.$$

Then $\widetilde{M}_A^n$ is a $\{\mathcal{G}_t^A : t\ge 0\}$-martingale with quadratic variation

$$\langle \widetilde{M}_A^n\rangle(t) = \lambda\int_0^t \bar{S}(s)\bar{I}(s)ds, \quad t\ge 0,$$

and $\widetilde{R}_1^n$ is a $\{\mathcal{G}_t^R : t\ge 0\}$-martingale, with quadratic variation

$$\langle \widetilde{R}_1^n\rangle(t) = \lambda\int_0^t F(t-s)\bar{S}(s)\bar{I}(s)ds, \quad t\ge 0.$$

Note that we do not have a martingale property for $\widetilde{I}^n$. It is important to observe that the joint process $(\widetilde{M}_A^n, \widetilde{R}_1^n)$ is not a martingale with respect to a common filtration, and therefore we cannot prove the joint convergence of them using FCLT of martingales. However, they play the role of establishing tightness of the processes $\{\hat{M}_A^n\}$, $\{\hat{I}_1^n\}$, and $\{\hat{R}_1^n\}$. Moreover, while $\{\hat{M}_A^n\}$ is a $\mathcal{F}_t^n$ martingale, $\{\hat{R}_1^n\}$ is not a martingale, the point being that the intensity $\lambda n\bar{S}^n(t)\bar{I}^n(t)$ is not $\mathcal{G}^R$–adapted. In fact, for the sake of establishing tightness, one can exploit the martingale property of $\{\hat{M}_A^n\}$, so that the introduction of $\widetilde{M}_A^n$ is not necessary. And since the tightness of $\hat{I}_A^n$ follows from those of both $\hat{M}_A^n$ and $\hat{R}_1^n$, only $\widetilde{R}_1^n$ really needs to be introduced for proving tightness. However, in the proof of Lemma 5.4, we shall now need the full strength of Lemma 5.3.

Lemma 5.4. *Under Assumptions 2.3 and 2.1,*

$$(\hat{M}_A^n, \hat{I}_1^n, \hat{R}_1^n) \Rightarrow (\hat{M}_1, \hat{I}^n, \hat{R}_1) \quad in \quad D^3 \quad as \quad n\to\infty,$$

where $(\hat{M}_A, \hat{I}_1, \hat{R}_1)$ are given in Theorem 2.2.

Proof. In view of Lemma 5.3, all we need to show is that

$$(\widetilde{M}_A^n, \widetilde{I}_1^n, \widetilde{R}^n) \Rightarrow (\hat{M}_A, \hat{I}_1, \hat{R}_1) \quad \text{in} \quad D^3 \quad \text{as} \quad n\to\infty\,.$$

Exploiting the martingale property of both $\widetilde{M}_A^n$ and $\widetilde{R}_1^n$, we can show that each of these two processes is tight in D. Since moreover any limit of a converging subsequence of either of these processes is continuous, the difference $\widetilde{I}_1^n(t) = \widetilde{M}_A^n(t) - \widetilde{R}_1^n(t)$ is also tight. Thus, by Lemma 5.3, $\{\hat{M}_A^n\}$, $\{\hat{I}_1^n\}$, and $\{\hat{R}_1^n\}$ are tight. It remains to show (i) convergence of finite dimensional distributions and (ii) the limits are continuous.

To prove the convergence of finite dimensional distributions, by the independence of the restrictions of a PRM to disjoint subsets, it suffices to show that for $0\le t'\le t$ and $0\le a\le b<\infty$,

$$\begin{aligned}
&\lim_{n\to\infty}\mathbb{E}\left[\exp\left(i\frac{\vartheta}{\sqrt{n}}\int_{t'}^t\int_a^b\int_0^\infty \widetilde{\varphi}_n(s,u)d\widetilde{M}(ds,dz,du)\right)\right]\\
&= \exp\left(-\frac{\vartheta^2}{2}\lambda\int_{t'}^t (F(b-s)-F(a-s))\bar{S}(s)\bar{I}(s)ds\right). \qquad (5.16)
\end{aligned}$$

Recall that for a compensated PRM $\bar{N}$ with mean measure ν and a deterministic function ϕ, we have

$$\mathbb{E}\left[\exp(i\vartheta\bar{N}(\phi))\right] = e^{-i\vartheta\nu(\phi)}\exp\left(\nu(e^{i\vartheta\phi}-1)\right), \qquad (5.17)$$

where $\nu(\phi) := \int\phi d\nu$. As a consequence, the left hand side of (5.16) is equal to

$$\begin{aligned}
&\exp\left(-i\frac{\vartheta}{\sqrt{n}}\int_{t'}^t (F(b-s)-F(a-s))\lambda n\bar{S}(s)\bar{I}(s)ds\right)\\
&\times\exp\left((e^{i\vartheta/\sqrt{n}}-1)\int_{t'}^t (F(b-s)-F(a-s))\lambda n\bar{S}(s)\bar{I}(s)ds\right).
\end{aligned}$$

Then the claim (5.16) is immediate by applying Taylor expansion.

Given the consistent finite dimensional distributions of $\hat{R}_1$, to show that the limit process $\hat{R}_1$ have a continuous version in C, it suffices to show that

$$\mathbb{E}\left[(\hat{R}_1(t+\delta)-\hat{R}_1(t)))^4\right] \le c\delta^{1+\theta}. \qquad (5.18)$$

This is immediate since as a consequence of (5.16),

$$\begin{aligned}
&\mathbb{E}\left[(\hat{R}_1(t+\delta)-\hat{R}_1(t)))^4\right] = 3\left(\mathbb{E}\left[(\hat{R}_1(t+\delta)-\hat{R}_1(t)))^2\right]\right)^2\\
&= 3\left(\lambda\int_t^{t+\delta}F(t+\delta-s)\bar{S}(s)\bar{I}(s)ds + \lambda\int_0^t (F(t+\delta-s)-F(t-s))\bar{S}(s)\bar{I}(s)ds\right)^2
\end{aligned}$$

$$\leq 6\lambda\delta^2 + 6\lambda \left(\int_0^t (F(t+\delta-s) - F(t-s))\bar{S}(s)\bar{I}(s)ds \right)^2 .$$

Then the claim follows from Assumption 2.1. This property holds analogously for the processes $\hat{M}_A$ and $\hat{I}_1$. This completes the proof. $\square$

Completing the proof of Theorem 2.2. By Lemmas 5.1 and 5.4, we first obtain the joint convergence

$$(\hat{I}^n(0)F_0^c(\cdot), \hat{I}^n(0)F_0(\cdot), \hat{I}_0^n, \hat{R}_0^n, \hat{M}_A^n, \hat{I}_1^n, \hat{R}_1^n) \Rightarrow \left(\hat{I}(0)F_0^c(\cdot), \hat{I}(0)F_0(\cdot), \hat{I}_0, \hat{R}_0, \hat{M}_A, \hat{I}_1, \hat{R}_1\right)$$

in D^7 as $n \to \infty$. Since the limit processes $\hat{I}_0, \hat{R}_0, \hat{M}_A, \hat{I}_1, \hat{R}_1$ are continuous, we have the convergence:

$$(-\hat{M}_A^n, \hat{I}^n(0)F_0^c(\cdot) + \hat{I}_0^n + \hat{I}_1^n, \hat{I}^n(0)F_0(\cdot) + \hat{R}_0^n + \hat{R}_1^n) \Rightarrow (-\hat{M}_A, \hat{I}(0)F_0^c(\cdot) + \hat{I}_0 + \hat{I}_1, \hat{I}(0)F_0^c(\cdot) + \hat{R}_0 + \hat{R}_1),$$

in D^3 as $n \to \infty$. It follows from (5.3), (5.4), Theorem 2.1, Lemma 5.1, 5.2 and 5.4 that $(\hat{I}^n, \hat{R}^n)$ is tight in D^2, and any limit of a converging subsequence satisfies (2.9) and (2.10), where we may replace $\hat{S}$ by $-\hat{I} - \hat{R}$, since $\hat{S}^n = -\hat{I}^n - \hat{R}^n$ for all n. From Lemma 8.1, this characterizes uniquely the limit, hence the whole sequence converges, and finally (2.7), (2.8) follow readily from the above, and again the fact that $\hat{S}^n = -\hat{I}^n - \hat{R}^n$ for all n. $\square$

Remark 5.1. *We now justify the claim stated in Remark 2.5. We reconsider the above proof, without requiring Assumption 2.1. Note that this Assumption has been used only in the proof of Lemma 5.3. An inspection of* (5.2) *reveals that the tightness of $\hat{S}^n$ in D follows readily from the martingale property of $\hat{M}_A^n$ together with* (4.3), (4.5)*, and Lemma 5.2. From the identity $\hat{S}^n + \hat{I}^n + \hat{R}^n = 0$, we note that it suffices to consider the tightness of $\hat{R}^n$. $\hat{R}_0^n$ is not a problem. We need to consider $\hat{R}_1^n$. As argued at the beginning of the proof of Lemma 5.4, tightness of $\widetilde{R}_1^n$ in D is not hard to establish. The argument where we made use of Assumption 2.1 is the convergence of $\hat{R}_1^n - \widetilde{R}_1^n$ to* 0 *in D. The fifth line of the proof of Lemma 5.4 establishes that $\mathbb{E}[(\hat{R}_1^n(t) - \widetilde{R}_1^n(t))^2] \to 0$. The same argument shows that the integral from $t=0$ to $t=T$ of this last quantity goes to zero as $n \to \infty$, hence in particular $\hat{R}_1^n - \widetilde{R}_1^n \to 0$ in $L^2(0,T)$ in probability, and $\hat{R}^n$ is tight in $L^2(0,T)$. It is now easy to conclude the proof of the claim stated in Remark 2.5.*

6. Proof of the FLLN for the SEIR model

In this section we prove Theorem 3.1. The expressions and claims in (4.1)–(4.9) hold by the same arguments, which we assume from now on. By slightly modifying the argument as for the process $\bar{I}^n$ in the SIR model, we obtain that

$$\bar{E}^n(\cdot) \Rightarrow \bar{E}(0)G_0^c(\cdot) + \int_0^{\cdot} G^c(\cdot - s)d\bar{A}(s)$$

in D as $n \to \infty$.

Recall $I^n(t)$ in (3.10). Define

$$\bar{I}_{0,1}^n(t) := \frac{1}{n}\sum_{j=1}^{I^n(0)} \mathbf{1}(\eta_j^0 > t),$$

$$\bar{I}_{0,2}^n(t) := \frac{1}{n}\sum_{j=1}^{E^n(0)} \mathbf{1}(\xi_j^0 \leq t)\mathbf{1}(\xi_j^0 + \eta_j > t),$$

$$\bar{I}_1^n(t) := \frac{1}{n}\sum_{i=1}^{A^n(t)} \mathbf{1}(\tau_i^n + \xi_i \leq t)\mathbf{1}(\tau_i^n + \xi_i + \eta_i > t).$$

By the FLLN of empirical processes, and by Assumption 3.2, we have

$$(\bar{I}^n_{0,1}, \bar{I}^n_{0,2}) \Rightarrow (\bar{I}_{0,1}, \bar{I}_{0,2}) \quad \text{in} \quad D^2 \quad \text{as} \quad n \to \infty, \tag{6.1}$$

where

$$\bar{I}_{0,1} := \bar{I}(0) G_0^c(\cdot), \quad \bar{I}_{0,2} := \bar{E}(0)\Psi_0(\cdot).$$

For the study of the process $\bar{I}^n_1$, we first consider

$$\begin{aligned}\breve{I}^n_1(t) := \mathbb{E}[\bar{I}^n_1(t)|\mathcal{F}^n_t] = \frac{1}{n}\sum_{i=1}^{A^n(t)} \Psi(t-\tau^n_i) = \int_0^t \Psi(t-s) d\bar{A}^n(s) \\ = \bar{A}^n(t) - \int_0^t \bar{A}^n(s) d\Psi(t-s).\end{aligned}$$

Applying the continuous mapping theorem, we obtain

$$\breve{I}^n_1 \Rightarrow \bar{I}_1 \quad \text{in} \quad D \quad \text{as} \quad n \to \infty. \tag{6.2}$$

where

$$\bar{I}_1(t) := \bar{A}(t) - \int_0^t \bar{A}(s) d\Psi(t-s) = \int_0^t \Psi(t-s) d\bar{A}(s), \quad t \ge 0. \tag{6.3}$$

We now consider the difference

$$V^n(t) := \bar{I}^n_1(t) - \breve{I}^n_1(t) = \frac{1}{n}\sum_{i=1}^{A^n(t)} \kappa^n_i(t),$$

where

$$\kappa^n_i(t) = \mathbf{1}(\tau^n_i + \xi_i \le t)\mathbf{1}(\tau^n_i + \xi_i + \eta_i > t) - \Psi(t-\tau^n_i).$$

We next show the following lemma.

Lemma 6.1. *For any $\epsilon > 0$,*

$$\mathbb{P}\left(\sup_{t\in[0,T]} |V^n(t)| > \epsilon\right) \to 0 \quad \text{as} \quad n \to \infty. \tag{6.4}$$

Proof. We partition $[0,T]$ into intervals of length $\delta > 0$, and have the bound for $\sup_{t\in[0,T]} |V^n(t)|$ as in (4.14).

First, we have

$$\mathbb{E}[\kappa^n_i(t)|\mathcal{F}^n_t] = 0, \quad \forall\, i; \quad \mathbb{E}[\kappa^n_i(t)\kappa^n_j(t)|\mathcal{F}^n_t] = 0, \quad \forall\, i \ne j.$$

Thus

$$\begin{aligned}&\mathbb{E}[V^n(t)^2|\mathcal{F}^n_t] = \frac{1}{n^2}\sum_{i=1}^{A^n(t)} \mathbb{E}[\kappa^n_i(t)^2|\mathcal{F}^n_t] \\ &= \frac{1}{n^2}\sum_{i=1}^{A^n(t)} \Psi(t-\tau^n_i)(1-\Psi(t-\tau^n_i)) = \frac{1}{n}\int_0^t \Psi(t-s)\left(1-\Psi(t-s)\right) d\bar{A}^n(s) \\ &= \frac{1}{n^{3/2}}\int_0^t \Psi(t-s)\left(1-\Psi(t-s)\right) d\hat{M}^n_A(s) + \frac{1}{n}\int_0^t \Psi(t-s)\left(1-\Psi(t-s)\right) d\bar{\Lambda}^n(s) \\ &\le \frac{1}{n^{3/2}}\int_0^t \Psi(t-s)\left(1-\Psi(t-s)\right) d\hat{M}^n_A(s) + \frac{\lambda t}{n},\end{aligned}$$

where the inequality follows from (4.4) and (4.5). Thus

$$\mathbb{E}[V^n(t)^2] \leq \frac{\lambda t}{n}, \quad \mathbb{P}(|V^n(t)| > \epsilon) \leq \frac{\lambda t}{\epsilon^2 n}. \tag{6.5}$$

Next we have

$$\begin{aligned}
&|V^n(t+u) - V^n(u)| \\
&= \left| \frac{1}{n} \sum_{i=1}^{A^n(t+u)} \kappa_i^n(t+u) - \frac{1}{n} \sum_{i=1}^{A^n(t)} \kappa_i^n(t) \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^{A^n(t)} (\kappa_i^n(t+u) - \kappa_i^n(t)) + \frac{1}{n} \sum_{i=A^n(t)}^{A^n(t+u)} \kappa_i^n(t+u) \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^{A^n(t)} (\mathbf{1}(\tau_i^n + \xi_i \leq t+u)\mathbf{1}(\tau_i^n + \xi_i + \eta_i > t+u) - \mathbf{1}(\tau_i^n + \xi_i \leq t)\mathbf{1}(\tau_i^n + \xi_i + \eta_i > t)) \right| \\
&\quad + \left| \int_0^t (\Psi(t+u-s) - \Psi(t-s))\, d\bar{A}^n(s) \right| + \frac{1}{n} \sum_{i=A^n(t)}^{A^n(t+u)} |\kappa_i^n(t+u)| \\
&\leq \frac{1}{n} \sum_{i=1}^{A^n(t)} \mathbf{1}(\tau_i^n + \xi_i \leq t+u)(\mathbf{1}(\tau_i^n + \xi_i + \eta_i > t) - \mathbf{1}(\tau_i^n + \xi_i + \eta_i > t+u)) \\
&\quad + \frac{1}{n} \sum_{i=1}^{A^n(t)} (\mathbf{1}(\tau_i^n + \xi_i \leq t+u) - \mathbf{1}(\tau_i^n + \xi_i \leq t))\mathbf{1}(\tau_i^n + \xi_i + \eta_i > t) \\
&\quad + \int_0^t \left(\int_0^{t-s+u} (F^c(t-s-v|v) - F^c(t+u-s-v|v))dG(v) \right) d\bar{A}^n(s) \\
&\quad + \int_0^t \left(\int_{t-s}^{t-s+u} F^c(t-s-v|v)dG(v) \right) d\bar{A}^n(s) + \frac{1}{n} \sum_{i=A^n(t)}^{A^n(t+u)} |\kappa_i^n(t+u)|.
\end{aligned}$$

Observing that the first four terms on the right hand side are all increasing in u, and that $|\kappa_i^n(t)| \leq 1$ for all t, i, n, we obtain that

$$\begin{aligned}
&\sup_{u\in[0,\delta]} |V^n(t+u) - V^n(u)| \\
&\leq \frac{1}{n} \sum_{i=1}^{A^n(t)} \mathbf{1}(\tau_i^n + \xi_i \leq t+\delta)(\mathbf{1}(\tau_i^n + \xi_i + \eta_i > t) - \mathbf{1}(\tau_i^n + \xi_i + \eta_i > t+\delta)) \\
&\quad + \frac{1}{n} \sum_{i=1}^{A^n(t)} (\mathbf{1}(\tau_i^n + \xi_i \leq t+\delta) - \mathbf{1}(\tau_i^n + \xi_i \leq t))\mathbf{1}(\tau_i^n + \xi_i + \eta_i > t) \\
&\quad + \int_0^t \left(\int_0^{t-s+\delta} (F^c(t-s-v|v) - F^c(t+\delta-s-v|v))dG(v) \right) d\bar{A}^n(s) \\
&\quad + \int_0^t \left(\int_{t-s}^{t-s+\delta} F^c(t-s-v|v)dG(v) \right) d\bar{A}^n(s) + (\bar{A}^n(t+\delta) - \bar{A}^n(t)).
\end{aligned} \tag{6.6}$$

Thus, for any $\epsilon > 0$,

$$\begin{aligned}
&\mathbb{P}\left(\sup_{u\in[0,\delta]} |V^n(t+u) - V^n(u)| > \epsilon\right) \qquad (6.7)\\
&\le \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{A^n(t)} \mathbf{1}(\tau_i^n + \xi_i \le t+\delta)\mathbf{1}(t < \tau_i^n + \xi_i + \eta_i \le t+\delta) > \epsilon/5\right)\\
&\quad + \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{A^n(t)} \mathbf{1}(t < \tau_i^n + \xi_i \le t+\delta)\mathbf{1}(\tau_i^n + \xi_i + \eta_i > t) > \epsilon/5\right)\\
&\quad + \mathbb{P}\left(\int_0^t \left(\int_0^{t-s+\delta} (F^c(t-s-v|v) - F^c(t+\delta-s-v|v))dG(v)\right) d\bar{A}^n(s) > \epsilon/5\right)\\
&\quad + \mathbb{P}\left(\int_0^t \left(\int_{t-s}^{t-s+\delta} F^c(t-s-v|v)dG(v)\right) d\bar{A}^n(s) > \epsilon/5\right) + \mathbb{P}\left((\bar{A}^n(t+\delta) - \bar{A}^n(t)) > \epsilon/5\right).
\end{aligned}$$

We need the following definition to treat the first two terms on the right hand side of (6.7)

Definition 6.1. *Define a PRM $M(ds, dy, dz, du)$ on $[0,T] \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+$ with mean measure $\nu(ds, dy, dz, du) = dsH(dy, dz)du$. Denote the compensated PRM by $\overline{M}(ds, dy, dz, du)$.*

For the first term on the right hand side of (6.7), we have

$$\begin{aligned}
&\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{A^n(t)} \mathbf{1}(\tau_i^n + \xi_i \le t+\delta)\mathbf{1}(t < \tau_i^n + \xi_i + \eta_i \le t+\delta)\right)^2\right]\\
&= \mathbb{E}\left[\left(\frac{1}{n}\int_0^t \int_0^{t+\delta-s}\int_{t-s-y}^{t+\delta-s-y}\int_0^{n\lambda\bar{S}^n(s)\bar{I}^n(s)} M(ds, dy, dz, du)\right)^2\right]\\
&\le 2\mathbb{E}\left[\left(\frac{1}{n}\int_0^t \int_0^{t+\delta-s}\int_{t-s-y}^{t+\delta-s-y}\int_0^{n\lambda\bar{S}^n(s)\bar{I}^n(s)} \overline{M}(ds, dy, dz, du)\right)^2\right]\\
&\quad + 2\mathbb{E}\left[\left(\int_0^t \left(\int_0^{t-s+\delta} (F^c(t-s-v|v) - F^c(t+\delta-s-v|v))dG(v)\right) d\bar{\Lambda}^n(s)\right)^2\right]\\
&= \frac{2}{n}\mathbb{E}\left[\int_0^t \left(\int_0^{t-s+\delta} (F^c(t-s-v|v) - F^c(t+\delta-s-v|v))dG(v)\right) d\bar{\Lambda}^n(s)\right]\\
&\quad + 2\mathbb{E}\left[\left(\int_0^t \left(\int_0^{t-s+\delta} (F^c(t-s-v|v) - F^c(t+\delta-s-v|v))dG(v)\right) d\bar{\Lambda}^n(s)\right)^2\right]\\
&\le \frac{2}{n}\lambda\int_0^t \left(\int_0^{t-s+\delta} (F^c(t-s-v|v) - F^c(t+\delta-s-v|v))dG(v)\right) ds\\
&\quad + 2\left(\lambda\int_0^t \left(\int_0^{t-s+\delta} (F^c(t-s-v|v) - F^c(t+\delta-s-v|v))dG(v)\right) ds\right)^2\\
&= \frac{2}{n}\lambda\int_0^{t+\delta} \left(\int_0^{t-v+\delta} (F^c(t-s-v|v) - F^c(t+\delta-s-v|v))ds\right) dG(v)\\
&\quad + 2\left(\lambda\int_0^{t+\delta} \left(\int_0^{t-v+\delta} (F^c(t-s-v|v) - F^c(t+\delta-s-v|v))ds\right) dG(v)\right)^2. \qquad (6.8)
\end{aligned}$$

Here the second inequality uses (4.5). The first term on the right hand side of (6.8) converges to zero as $n \to \infty$. By Assumption 3.1, we have

$$\frac{1}{\delta}\left(\int_0^{t+\delta}\left(\int_0^{t-v+\delta}(F^c(t-s-v|v)-F^c(t+\delta-s-v|v))ds\right)dG(v)\right)^2 \to 0 \quad \text{as} \quad \delta \to 0. \tag{6.9}$$

Indeed, in the case $F = F_1$, the left hand side is equal to

$$\frac{1}{\delta}\left(\int_0^{t+\delta}\left(\int_0^{t-v+\delta}\sum_i a_i(\mathbf{1}(t-s-v<t_i)-\mathbf{1}(t+\delta-s-v<t_i))ds\right)dG(v)\right)^2 \le \delta G(t+\delta)^2,$$

since $\sum_i a_i \le 1$. In the case $F = F_2$, the left hand side is bounded by

$$\frac{1}{\delta}\left(\int_0^{t+\delta}\left(c\delta^{1/2+\theta}(t-v+\delta)\right)dG(v)\right)^2 \le c^2\delta^{2\theta}(t+\delta)^2 G(t+\delta)^2.$$

It is then clear that in both cases (6.9) holds.

Similarly, for the second term on the right hand side of (6.7), we have

$$\begin{aligned}
&\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{A^n(t)}\mathbf{1}(t<\tau_i^n+\xi_i\le t+\delta)\mathbf{1}(\tau_i^n+\xi_i+\eta_i>t)\right)^2\right]\\
&=\mathbb{E}\left[\left(\frac{1}{n}\int_0^t\int_{t-s}^{t+\delta-s}\int_{t-s-y}^{\infty}\int_0^{n\lambda\bar{S}^n(s)\bar{I}^n(s)}M(ds,dy,dz,du)\right)^2\right]\\
&\le 2\mathbb{E}\left[\left(\frac{1}{n}\int_0^t\int_{t-s}^{t+\delta-s}\int_{t-s-y}^{\infty}\int_0^{n\lambda\bar{S}^n(s)\bar{I}^n(s)}\overline{M}(ds,dy,dz,du)\right)^2\right]\\
&\quad+2\mathbb{E}\left[\left(\int_0^t\left(\int_{t-s}^{t-s+\delta}F^c(t-s-v|v)dG(v)\right)d\bar{\Lambda}^n(s)\right)^2\right]\\
&=\frac{2}{n}\mathbb{E}\left[\int_0^t\left(\int_{t-s}^{t-s+\delta}F^c(t-s-v|v)dG(v)\right)d\bar{\Lambda}^n(s)\right]\\
&\quad+2\mathbb{E}\left[\left(\int_0^t\left(\int_{t-s}^{t-s+\delta}F^c(t-s-v|v)dG(v)\right)d\bar{\Lambda}^n(s)\right)^2\right]\\
&\le\frac{2}{n}\lambda\int_0^t\left(\int_{t-s}^{t-s+\delta}F^c(t-s-v|v)dG(v)\right)ds\\
&\quad+2\left(\lambda\int_0^t\left(\int_{t-s}^{t-s+\delta}F^c(t-s-v|v)dG(v)\right)ds\right)^2.
\end{aligned}\tag{6.10}$$

Again, here the second inequality uses (4.5). The first term on the right hand side of (6.10) converges to zero as $n \to \infty$. By Assumption 3.1, we have

$$\frac{1}{\delta}\left(\lambda\int_0^t\left(\int_{t-s}^{t-s+\delta}F^c(t-s-v|v)dG(v)\right)ds\right)^2 \to 0 \quad \text{as} \quad \delta \to 0. \tag{6.11}$$

For the third term on the right hand side of (6.7), by (4.1), we have

$$\mathbb{E}\left[\left(\int_0^t\left(\int_0^{t-s+\delta}(F^c(t-s-v|v)-F^c(t+\delta-s-v|v))dG(v)\right)d\bar{A}^n(s)\right)^2\right]$$

$$\leq 2\mathbb{E}\left[\left(\frac{1}{\sqrt{n}}\int_0^t\left(\int_0^{t-s+\delta}(F^c(t-s-v|v)-F^c(t+\delta-s-v|v))dG(v)\right)d\hat{M}_A^n(s)\right)^2\right]$$
$$+2\mathbb{E}\left[\left(\int_0^t\left(\int_0^{t-s+\delta}(F^c(t-s-v|v)-F^c(t+\delta-s-v|v))dG(v)\right)d\bar{\Lambda}^n(s)\right)^2\right]. \qquad (6.12)$$

Then by (4.6) the first term converges to zero as $n\to\infty$, and the second term can be treated similarly as the second term in (6.8). The fourth term in (6.7) can be treated similarly. The last term in (6.7) is the same as in (4.20). Therefore, by combining the above arguments and (6.7)–(6.12), we obtain

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\left[\frac{T}{\delta}\right]\sup_{0\leq t\leq T}\mathbb{P}\left(\sup_{u\in[0,\delta]}|V^n(t+u)-V^n(t)|\geq\epsilon\right)=0.$$

Then by (4.21) and (6.5), we conclude that (6.4) holds. □

By (6.2) and (6.4), we have

$$\bar{I}_1^n\Rightarrow\bar{I}_1 \quad\text{in}\quad D \quad\text{as}\quad n\to\infty.$$

Combining this with the convergences of $(\bar{I}_{0,1}^n,\bar{I}_{0,2}^n)$ in (6.1), by independence of $(\bar{I}_{0,1}^n,\bar{I}_{0,2}^n)$ and $\bar{I}_1^n$, we have

$$\bar{I}^n=\bar{I}_{0,1}^n+\bar{I}_{0,2}^n+\bar{I}_1^n\Rightarrow\bar{I}=\bar{I}_{0,1}+\bar{I}_{0,2}+\bar{I}_1 \quad\text{in}\quad D\quad\text{as}\quad n\to\infty.$$

Similar to the SIR model, we can show the joint convergence

$$(\bar{S}^n,\breve{I}^n)\Rightarrow(\bar{S},\bar{I}) \quad\text{in}\quad D^2\quad\text{as}\quad n\to\infty.$$

Thus, using a similar argument as in the SIR model, we have shown that the limits $(\bar{S},\bar{I})$ of $(\bar{S}^n,\bar{I}^n)$ satisfy the integral equations (3.13) and (3.15). Similarly to the SIR model, these two equations have a unique solution. Once the solutions of $(\bar{S},\bar{I})$ are uniquely determined, the other limits $\bar{A},\bar{E},\bar{L},\bar{R}$ are also uniquely determined by the corresponding integral equations. This proves the convergence in probability. Therefore the proof of Theorem 3.1 is complete.

7. Proof of the FCLT for the SEIR model

In this section we prove Theorem 3.2, for the diffusion-scaled processes $(\hat{S}^n,\hat{E}^n,\hat{I}^n,\hat{R}^n)$ defined in (3.17). Similarly to the SIR model, under Assumption 3.3, we have $(\bar{I}^n(0),\bar{E}^n(0))\Rightarrow(\bar{I}(0),\bar{E}(0))\in\mathbb{R}_+^2$ as $n\to\infty$, and thus the FLLN Theorem 3.1 holds, which will be taken as given in the proof below. Recall the martingale $\hat{M}_A^n$ defined in (4.2).

We have the following representation of the diffusion-scaled processes. We have the same representation of $\hat{S}^n$ in (5.2) for the SIR model. For the ease of exposition, we repeat the following expression for the process $\hat{S}^n$:

$$\hat{S}^n(t)=-\hat{I}^n(0)-\hat{M}_A^n(t)-\lambda\int_0^t\left(\hat{S}^n(s)\bar{I}^n(s)+\bar{S}(s)\hat{I}^n(s)\right)ds.$$

For the process $\hat{E}^n$,

$$\begin{aligned}\hat{E}^n(t)=\hat{E}^n(0)G_0^c(t)+\hat{E}_0^n(t)+\hat{E}_1^n(t)&\\ +\lambda\int_0^t G^c(t-s)\left(\hat{S}^n(s)\bar{I}^n(s)+\bar{S}(s)\hat{I}^n(s)\right)ds,&\end{aligned}$$

where

$$\hat{E}_0^n(t):=\frac{1}{\sqrt{n}}\sum_{j=1}^{n\bar{E}^n(0)}\left(\mathbf{1}(\xi_j^0>t)-G_0^c(t)\right),$$

$$\hat{E}_1^n(t) := \frac{1}{\sqrt{n}} \sum_{i=1}^{n\bar{A}^n(t)} \mathbf{1}(\tau_i^n + \xi_i > t) - \sqrt{n}\lambda \int_0^t G^c(t-s)\bar{S}^n(s)\bar{I}^n(s)ds.$$

For the process $\hat{I}^n$,

$$\begin{aligned}\hat{I}^n(t) = \hat{I}^n(0)F_0^c(t) + \hat{E}^n(0)\Psi_0(t) + \hat{I}_{0,1}^n(t) + \hat{I}_{0,2}^n(t) + \hat{I}_1^n(t) \\ + \lambda \int_0^t \Psi(t-s)\left(\hat{S}^n(s)\bar{I}^n(s) + \bar{S}(s)\hat{I}^n(s)\right) ds,\end{aligned}$$

where

$$\hat{I}_{0,1}^n(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{I^n(0)} \big(\mathbf{1}(\eta_j^0 > t) - F_0^c(t)\big),$$

$$\hat{I}_{0,2}^n(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{E^n(0)} \left(\mathbf{1}(\xi_j^0 \le t)\mathbf{1}(\xi_j^0 + \eta_j > t) - \Psi_0(t)\right),$$

and

$$\hat{I}_1^n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{A^n(t)} \mathbf{1}(\tau_i^n + \xi_i \le t)\mathbf{1}(\tau_i^n + \xi_i + \eta_i > t) - \lambda\sqrt{n}\int_0^t \Psi(t-s)\bar{S}^n(s)\bar{I}^n(s)ds.$$

For the process $\hat{R}^n$,

$$\begin{aligned}\hat{R}^n(t) = \hat{I}^n(0)F_0(t) + \hat{E}^n(0)\Phi_0(t) + \hat{R}_{0,1}^n(t) + \hat{R}_{0,2}^n(t) + \hat{R}_1^n(t) \\ + \lambda \int_0^t \Phi(t-s)\left(\hat{S}^n(s)\bar{I}^n(s) + \bar{S}(s)\hat{I}^n(s)\right) ds,\end{aligned}$$

where

$$\hat{R}_{0,1}^n(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{I^n(0)} \big(\mathbf{1}(\eta_j^0 \le t) - F_0(t)\big),$$

$$\hat{R}_{0,2}^n(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{E^n(0)} \left(\mathbf{1}(\xi_j^0 + \eta_j \le t) - \Phi_0(t)\right),$$

and

$$\hat{R}_1^n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{A^n(t)} \mathbf{1}(\tau_i^n + \xi_i + \eta_i \le t) - \lambda\sqrt{n}\int_0^t \Phi(t-s)\bar{S}^n(s)\bar{I}^n(s)ds.$$

To facilitate the proof, we also define the process $\hat{L}^n$ (recall that $L^n(t) = I^n(t) + R^n(t) - I^n(0)$) :

$$\hat{L}^n(t) := \sqrt{n}\left(\bar{L}^n(t) - \bar{L}(t)\right) = \sqrt{n}\left(\bar{L}^n(t) - \left(\bar{E}(0)G_0(t) + \lambda\int_0^t G(t-s)\bar{S}(s)\bar{I}(s)ds\right)\right).$$

It has the following representation:

$$\begin{aligned}\hat{L}^n(t) = \hat{E}^n(0)G_0(t) + \hat{L}_0^n(t) + \hat{L}_1^n(t) \\ + \lambda \int_0^t G(t-s)\left(\hat{S}^n(s)\bar{I}^n(s) + \bar{S}(s)\hat{I}^n(s)\right) ds,\end{aligned}$$

where

$$\hat{L}_0^n(t) := \frac{1}{\sqrt{n}} \sum_{j=1}^{n\bar{E}^n(0)} \big(\mathbf{1}(\xi_j^0 \le t) - G_0(t)\big),$$

$$\hat{L}_1^n(t) := \frac{1}{\sqrt{n}} \sum_{i=1}^{n\bar{A}^n(t)} \mathbf{1}(\tau_i^n + \xi_i \le t) - \sqrt{n}\lambda \int_0^t G(t-s)\bar{S}^n(s)\bar{I}^n(s)ds.$$

We have the following joint convergence for the initial quantities similar to Lemma 5.1 for the SIR model. Its proof is omitted for brevity.

Lemma 7.1. *Under Assumption 3.3, we have*

$$\Big(\hat{E}^n(0)G_0^c(\cdot), \hat{E}_0^n, \hat{E}^n(0)G_0(\cdot), \hat{L}_0^n, \hat{I}^n(0)F_0^c(\cdot), \hat{E}^n(0)\Psi_0(\cdot), \hat{I}_{0,1}^n, \hat{I}_{0,2}^n, \hat{I}^n(0)F_0(\cdot), \hat{E}^n(0)\Phi_0(\cdot), \hat{R}_{0,1}^n, \hat{R}_{0,2}^n\Big)$$

$$\Rightarrow \Big(\hat{E}(0)G_0^c(\cdot), \hat{E}_0, \hat{E}(0)G_0(\cdot), \hat{L}_0, \hat{I}(0)F_0^c(\cdot), \hat{E}(0)\Psi_0(\cdot), \hat{I}_{0,1}, \hat{I}_{0,2}, \hat{I}(0)F_0(\cdot), \hat{E}(0)\Phi_0(\cdot), \hat{R}_{0,1}, \hat{R}_{0,2}\Big)$$

in D^{12} as $n \to \infty$, where the limit processes $\hat{E}_0$, $\hat{I}_{0,1}$, $\hat{I}_{0,2}$, $\hat{R}_{0,1}$ and $\hat{R}_{0,2}$ are given in Theorem 2.2, and $\hat{L}_0$ is a continuous mean-zero Gaussian process with the covariance function

$$\mathrm{Cov}(\hat{L}_0(t), \hat{L}_0(s)) = \bar{E}(0)(G_0(t \wedge s) - G_0(t)G_0(s)), \quad t, s \ge 0.$$

In addition,

$$\mathrm{Cov}(\hat{E}_0(t), \hat{L}_0(t')) = \bar{I}(0)\Big((G_0(t') - F_0(t))\mathbf{1}(t' \ge t) - G_0^c(t)G_0(t')\Big),$$

$$\mathrm{Cov}(\hat{L}_0(t), \hat{I}_{0,2}(t')) = \bar{E}(0)\left(\int_t^{t'} \mathbf{1}(t' \ge t)F_0(t'-s|s)dG_0(s) - G_0(t)\Psi_0(t')\right),$$

$$\mathrm{Cov}(\hat{L}_0(t), \hat{R}_{0,2}(t')) = \bar{E}(0)\left(\int_t^{t'} F_0(t'-s|s)dG_0(s) - G_0(t)\Phi_0(t')\right),$$

and $\hat{L}_0$ is independent with the other limit processes of the initial quantities.

Recall the definition of PRM $M(ds, dy, dz, du)$ and its compensated PRM in Definition 6.1.

Definition 7.1. *Let $M_1(ds, dy, dz, du)$ be a PRM on $[0,T] \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+$ with mean measure $\tilde{\nu}(ds, dy, dz, du) = ds\tilde{H}_s(dy, dz)du$ such that the first marginal of $\tilde{H}_s$ is $\tilde{G}_s((a,b]) = G((a+s, b+s])$ and the conditional distribution $\tilde{F}_s((a,b]|y) = F((a+s+y, b+s+y]|y)$. Denote the compensated PRM by $\widetilde{M}(ds, dy, dz, du)$.*

We use again the notation $\varphi_n(s,u) = \mathbf{1}\left(u \le n\lambda\bar{S}^n(s)\bar{I}^n(s)\right)$. We can rewrite the processes $\hat{M}_A^n$, $\hat{E}_1^n$, $\hat{L}_1^n$, $\hat{I}_1^n$ and $\hat{R}_1^n$ as

$$\begin{aligned}
\hat{M}_A^n(t) &= \frac{1}{\sqrt{n}} \int_0^t \int_0^\infty \int_0^\infty \int_0^\infty \varphi_n(s,u)\overline{M}(ds, dy, dz, du) \\
&= \frac{1}{\sqrt{n}} \int_0^t \int_0^\infty \int_0^\infty \int_0^\infty \varphi_n(s,u)\widetilde{M}(ds, dy, dz, du), \\
\hat{E}_1^n(t) &= \frac{1}{\sqrt{n}} \int_0^t \int_{t-s}^\infty \int_0^\infty \int_0^\infty \varphi_n(s,u)\overline{M}(ds, dy, dz, du) \\
&= \frac{1}{\sqrt{n}} \int_0^t \int_t^\infty \int_0^\infty \int_0^\infty \varphi_n(s,u)\widetilde{M}(ds, dy, dz, du), \\
\hat{L}_1^n(t) &= \frac{1}{\sqrt{n}} \int_0^t \int_0^{t-s} \int_0^\infty \int_0^\infty \varphi_n(s,u)\overline{M}(ds, dy, dz, du) \\
&= \frac{1}{\sqrt{n}} \int_0^t \int_0^t \int_0^\infty \int_0^\infty \varphi_n(s,u)\widetilde{M}(ds, dy, dz, du), \\
\hat{I}_1^n(t) &= \frac{1}{\sqrt{n}} \int_0^t \int_0^{t-s} \int_{t-s-y}^\infty \int_0^\infty \varphi_n(s,u)\overline{M}(ds, dy, dz, du)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{n}} \int_0^t \int_0^t \int_t^\infty \int_0^\infty \varphi_n(s,u) \widetilde{M}(ds,dy,dz,du), \\
\hat{R}_1^n(t) &= \frac{1}{\sqrt{n}} \int_0^t \int_0^{t-s} \int_0^{t-s-y} \int_0^\infty \varphi_n(s,u) \overline{M}(ds,dy,dz,du) \\
&= \frac{1}{\sqrt{n}} \int_0^t \int_0^t \int_0^t \int_0^\infty \varphi_n(s,u) \widetilde{M}(ds,dy,dz,du).
\end{aligned}
$$

Observe that

$$
\hat{M}_A^n(t) = \hat{E}_1^n(t) + \hat{L}_1^n(t), \quad t \geq 0, \tag{7.1}
$$

and

$$
\hat{L}_1^n(t) = \hat{I}_1^n(t) + \hat{R}_1^n(t), \quad t \geq 0. \tag{7.2}
$$

We define the auxiliary processes $\widetilde{M}_A^n$, $\widetilde{E}_1^n$, $\widetilde{L}_1^n$, $\widetilde{I}_1^n$ and $\widetilde{R}_1^n$ by replacing $\varphi_n(s,u)$ by

$$
\widetilde{\varphi}_n(s,u) = \mathbf{1}\left(u \leq n\lambda \bar{S}(s)\bar{I}(s)\right),
$$

in the corresponding processes using the compensated PRM $\widetilde{M}(ds,dy,dz,du)$. Then we have

$$
\widetilde{M}_A^n(t) = \widetilde{E}_1^n(t) + \widetilde{L}_1^n(t), \quad t \geq 0, \tag{7.3}
$$

and

$$
\widetilde{L}_1^n(t) = \widetilde{I}_1^n(t) + \widetilde{R}_1^n(t), \quad t \geq 0. \tag{7.4}
$$

Similar to Lemma 5.2 for the SIR model, we have the following result. We omit its proof for brevity.

Lemma 7.2.

$$
\sup_{t\in[0,T]} \mathbb{E}[|\hat{S}^n(t)|^2] < \infty, \quad \sup_{t\in[0,T]} \mathbb{E}[|\hat{E}^n(t)|^2] < \infty, \quad \sup_{t\in[0,T]} \mathbb{E}[|\hat{I}^n(t)|^2] < \infty, \quad \sup_{t\in[0,T]} \mathbb{E}[|\hat{R}^n(t)|^2] < \infty.
$$

Then, following an analogous argument as in the proof of Lemma 5.3, we obtain the following.

Lemma 7.3. *Under Assumption 2.3,*

$$
(\hat{M}_A^n - \widetilde{M}_A^n, \hat{E}_1^n - \widetilde{E}_1^n, \hat{L}_1^n - \widetilde{L}_1^n, \hat{I}_1^n - \widetilde{I}_1^n, \hat{R}_1^n - \widetilde{R}_1^n) \Rightarrow 0 \quad in \quad D^5 \quad as \quad n \to \infty.
$$

Proof. By the same argument as in the proof for the SIR model, we obtain the convergence $\hat{M}_A^n - \widetilde{M}_A^n \Rightarrow 0$, and $\hat{L}_1^n - \widetilde{L}_1^n \Rightarrow 0$, and thus, by (7.1) and (7.3), we have $\hat{E}_1^n - \widetilde{E}_1^n \Rightarrow 0$. We then show that $\hat{R}_1^n - \widetilde{R}_1^n \Rightarrow 0$, which will imply $\hat{I}_1^n - \widetilde{I}_1^n \Rightarrow 0$ by (7.2) and (7.4). On the other hand, the proof of $\hat{R}_1^n - \widetilde{R}_1^n \Rightarrow 0$ follows essentially the same argument as that in the SIR model, if we replace the infectious periods by the sum of the exposing and infectious periods. In the analysis we simply replace the distribution function F by the convolution of F and G. In particular, the difference process $\Xi^n = \hat{R}_1^n - \widetilde{R}_1^n$, has $\mathbb{E}[\Xi_1^n(t)] = 0$, and

$$
\mathbb{E}\left[\Xi^n(t)^2\right] = \frac{1}{n} \int_0^t \Phi(t-s) \mathbb{E}\left[|\bar{S}^n(s)\bar{I}^n(s) - \bar{S}(s)\bar{I}(s)|\right] ds,
$$

for each $t \geq 0$. To show that the sequence $\{\Xi^n : n \geq 1\}$ is tight, as in the proof of the SIR model, it suffices to show the tightness of the processes $\Xi_1^n(t)$ and $\Xi_2^n(t)$:

$$
\begin{aligned}
\Xi_1^n(t) &= \frac{1}{\sqrt{n}} \int_0^t \int_0^t \int_0^t \int_{n\lambda(\bar{S}^n(s)\bar{I}^n(s) \wedge \bar{S}(s)\bar{I}(s))}^{n\lambda(\bar{S}^n(s)\bar{I}^n(s) \vee \bar{S}(s)\bar{I}(s))} M_1(ds,dy,dz,du), \\
\Xi_2^n(t) &= \lambda\sqrt{n} \int_0^t \Phi(t-s) \big|\bar{S}^n(s)\bar{I}^n(s) - \bar{S}(s)\bar{I}(s)\big| ds.
\end{aligned}
$$

It suffices to show that (5.10) holds for each process. Both processes $\Xi_1^n(t)$ and $\Xi_2^n(t)$ are increasing in t. The proof then follows step by step and it requires the condition:

$$\limsup_{n\to\infty}\frac{1}{\delta}\mathbb{E}\left[\left(\int_0^t(\Phi(t+\delta-s)-\Phi(t-s))\Delta^n(s)ds\right)^2\right]\to 0 \tag{7.5}$$

as $\delta\to 0$. We observe that

$$\begin{aligned}
&\Phi(t+\delta-s)-\Phi(t-s)\\
&=\int_0^{t+\delta-s}F(t+\delta-s-u|u)dG(u)-\int_0^{t-s}F(t-s-u|u)dG(u)\\
&=\int_{t-s}^{t+\delta-s}F(t+\delta-s-u|u)dG(u)+\int_0^{t-s}(F(t+\delta-s-u|u)-F(t-s-u|u))dG(u).
\end{aligned}$$

Thus, we have

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^t(\Phi(t+\delta-s)-\Phi(t-s))\Delta^n(s)ds\right)^2\right]\\
&\le 2\mathbb{E}\left[\left(\int_0^t\int_{t-s}^{t+\delta-s}F(t+\delta-s-u|u)dG(u)\Delta^n(s)ds\right)^2\right]\\
&\quad+2\mathbb{E}\left[\left(\int_0^t\int_0^{t-s}(F(t+\delta-s-u|u)-F(t-s-u|u))dG(u)\Delta^n(s)ds\right)^2\right].
\end{aligned}$$

The first term can be bounded by

$$2\mathbb{E}\left[\left(\int_0^t(G(t+\delta-s)-G(t-s))\Delta^n(s)ds\right)^2\right]$$

which can be dealt with in the same way as the SIR model. For the second term, by interchanging the order of integration and using Jensen's inequality, we have

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^t\int_0^{t-s}(F(t+\delta-s-u|u)-F(t-s-u|u))\Delta^n(s)dsdG(u)\right)^2\right]\\
&\le \mathbb{E}\left[\int_0^t\left(\int_0^{t-u}(F(t+\delta-s-u|u)-F(t-s-u|u))\Delta^n(s)ds\right)^2dG(u)\right].
\end{aligned}$$

Then given the assumptions satisfied by the conditional distribution function $F(\cdot|u)$, we can show that this term is at most of the order of $o(\delta)$ as in the SIR model. This completes the proof. □

Let

$$\mathcal{G}_t^A:=\sigma\left\{\widetilde{M}([0,u]\times\mathbb{R}_+^3):0\le u\le t\right\},\quad t\ge 0,$$

$$\mathcal{G}_t^L:=\sigma\left\{\widetilde{M}([0,u]\times[0,u]\times\mathbb{R}_+^2):0\le u\le t\right\},\quad t\ge 0,$$

and

$$\mathcal{G}_t^R:=\sigma\left\{\widetilde{M}([0,u]\times[0,u]\times[0,u]\times\mathbb{R}_+):0\le u\le t\right\},\quad t\ge 0.$$

It is clear that $\widetilde{M}_A^n$ is a $\{\mathcal{G}_t^{A,n}:t\ge 0\}$-martingale with quadratic variation

$$\langle\widetilde{M}_A^n\rangle(t)=\lambda\int_0^t\bar{S}(s)\bar{I}(s)ds,\quad t\ge 0,$$

$\widetilde{L}_1^n$ is a $\{\mathcal{G}_t^{L,n}: t \geq 0\}$-martingale with quadratic variation

$$\langle \widetilde{L}_1^n \rangle(t) = \lambda \int_0^t G(t-s)\bar{S}(s)\bar{I}(s)ds, \quad t \geq 0,$$

and $\widetilde{R}_1^n$ is a $\{\mathcal{G}_t^{R,n}: t \geq 0\}$-martingale with quadratic variation

$$\langle \widetilde{R}_1^n \rangle(t) = \lambda \int_0^t \Phi(t-s)\bar{S}(s)\bar{I}(s)ds, \quad t \geq 0,$$

Note that we do not have a martingale property for $\widetilde{E}^n$ nor $\widetilde{I}^n$, and like in the SIR model, it is important to observe that the joint process $(\widetilde{M}_A^n, \widetilde{L}_A^n, \widetilde{R}_1^n)$ is not a martingale with respect to a common filtration, and we only use their individual martingale property to conclude their tightness.

Lemma 7.4. *Under Assumptions 3.3 and 3.1,*

$$(\hat{M}_A^n, \hat{E}_1^n, \hat{L}_1^n, \hat{I}_1^n, \hat{R}_1^n) \Rightarrow (\hat{M}_1, \hat{E}_1, \hat{L}_1, \hat{I}^n, \hat{R}_1) \quad in \quad D^5 \quad as \quad n \to \infty,$$

where $(\hat{M}_A, \hat{E}_1, \hat{I}_1, \hat{R}_1)$ are given in Theorem 2.2, and $\hat{L}_1$ is a continuous Gaussian process with covariance function: for $t, t' \geq 0$,

$$\mathrm{Cov}(\hat{L}_1(t), \hat{L}_1(t')) = \lambda \int_0^{t\wedge t'} G(t \vee t' - s)\bar{S}(s)\bar{I}(s)ds,$$

and it has covariance functions with the other processes: for $t, t' \geq 0$,

$$\mathrm{Cov}(\hat{M}_A(t), \hat{L}_1(t')) = \lambda \int_0^{t\wedge t'} G(t' - s)\bar{S}(s)\bar{I}(s)ds,$$

$$\mathrm{Cov}(\hat{E}_1(t), \hat{L}_1(t')) = \lambda \int_0^{t\wedge t'} (G(t'-s) - G(t-s))\mathbf{1}(t' \geq t)\bar{S}(s)\bar{I}(s)ds,$$

$$\mathrm{Cov}(\hat{L}_1(t), \hat{I}_1(t')) = \lambda \int_0^{t\wedge t'} (G(t-s) - \Psi(t'-s))\mathbf{1}(t' \geq t)\bar{S}(s)\bar{I}(s)ds,$$

$$\mathrm{Cov}(\hat{L}_1(t), \hat{I}_R(t')) = \lambda \int_0^{t\wedge t'} (G(t-s) - \Phi(t'-s))\mathbf{1}(t' \geq t)\bar{S}(s)\bar{I}(s)ds.$$

Proof. In view of Lemma 7.3, it suffices to prove that

$$(\widetilde{M}_A^n, \widetilde{E}_1^n, \widetilde{L}_1^n, \widetilde{I}_1^n, \widetilde{R}_1^n) \Rightarrow (\hat{M}_A, \hat{E}_1, \hat{L}_1, \hat{I}_1, \hat{R}_1) \quad \text{in} \quad D^5 \quad \text{as} \quad n \to \infty.$$

Using the martingale property of $\widetilde{M}_A^n$, $\widetilde{L}_1^n$ and $\widetilde{R}_1^n$, we establish tightness of each of these processes in D. Moreover each of the possible limit being continuous, the differences $\widetilde{I}_1^n(t) = \widetilde{L}_1^n(t) - \widetilde{R}_1^n(t)$, and $\widetilde{E}_1^n(t) = \widetilde{M}_A^n(t) - \widetilde{L}_1^n(t)$ are tight. Lemma 7.3 now implies that $\{\hat{M}_A^n\}$, $\{\hat{E}_A^n\}$, , $\{\hat{I}_1^n\}$, and $\{\hat{R}_1^n\}$ are tight. We next show (i) convergence of finite dimensional distributions and (ii) the limits are continuous.

To prove the convergence of finite dimensional distributions, by the independence of the restrictions of a PRM to disjoint subsets, it suffices to show that for $0 \leq t' \leq t$, $0 \leq a \leq b < \infty$ and $0 \leq c \leq d < \infty$,

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}\left[\exp\left(i\frac{\vartheta}{\sqrt{n}}\int_{t'}^t\int_a^b\int_c^d\int_0^\infty \widetilde{\varphi}^n(s)\widetilde{M}(ds,dy,dz,du)\right)\right] \\
&= \exp\left(-\frac{\vartheta^2}{2}\lambda\int_{t'}^t\left(\int_a^b\int_c^d \tilde{H}_s(dy,dz)\right)\bar{S}(s)\bar{I}(s)ds\right). \qquad (7.6)
\end{aligned}$$

where

$$\int_a^b\int_c^d \tilde{H}_s(dy,dz) = \int_{a-s}^{b-s}(F(d-y-s|y) - F(c-y-s|y))G(dy).$$

By (5.17), the left hand side of (7.6) is equal to

$$\exp\left(-i\frac{\vartheta}{\sqrt{n}}\int_{t'}^{t}\left(\int_a^b\int_c^d \tilde{H}_s(dy,dz)\right)\lambda n\bar{S}(s)\bar{I}(s)ds\right)$$
$$\times\exp\left((e^{i\vartheta/\sqrt{n}}-1)\int_{t'}^{t}\left(\int_a^b\int_c^d \tilde{H}_s(dy,dz)\right)\lambda n\bar{S}(s)\bar{I}(s)ds\right).$$

Then the claim in (5.16) is immediate by applying Taylor expansion.

We next show that there exists a continuous version of the limit processes $\hat{M}_A$, $\hat{E}_1$, $\hat{I}_1$ and $\hat{R}_1$ in C. Taking $\hat{R}_1$ as an example, we need to show (5.18) holds. By (7.6), we have

$$\begin{aligned}
&\mathbb{E}\left[(\hat{R}_1(t+\delta)-\hat{R}_1(t)))^4\right] = 3\left(E\left[(\hat{R}_1(t+\delta)-\hat{R}_1(t)))^2\right]\right)^2\\
&= 3\left(\lambda\int_t^{t+\delta}\Phi(t+\delta-s)\bar{S}(s)\bar{I}(s)ds+\lambda\int_0^t(\Phi(t+\delta-s)-\Phi(t-s))\bar{S}(s)\bar{I}(s)ds\right)^2\\
&\leq 6\lambda\delta^2+6\lambda\left(\int_0^t(\Phi(t+\delta-s)-\Phi(t-s))\bar{S}(s)\bar{I}(s)ds\right)^2.
\end{aligned}$$

This implies that (5.18) holds thanks to Assumption 3.1, see the computations for the proof of (7.5) above. This completes the proof. □

Completing the proof of Theorem 2.2. By Lemmas 7.1 and 7.4, we first obtain the joint convergence

$$\begin{aligned}
&\Big(-\hat{I}^n(0)-\hat{M}_A^n,\hat{E}^n(0)G_0^c(\cdot)+\hat{E}_0^n+\hat{E}_1^n,\hat{I}^n(0)F_0^c(\cdot)+\hat{E}^n(0)\Psi_0(\cdot)+\hat{I}_{0,1}^n+\hat{I}_{0,2}^n+\hat{I}_1^n,\\
&\qquad\hat{I}^n(0)F_0(\cdot)+\hat{E}^n(0)\Phi_0(\cdot)+\hat{R}_{0,1}^n+\hat{R}_{0,2}^n+\hat{R}_1^n\Big)\\
&\Rightarrow\Big(-\hat{I}(0)-\hat{M}_A,\hat{E}(0)G_0^c(\cdot)+\hat{E}_0+\hat{E}_1,\hat{I}(0)F_0^c(\cdot)+\hat{E}(0)\Psi_0(\cdot)+\hat{I}_{0,1}+\hat{I}_{0,2}+\hat{I}_1,\\
&\qquad\hat{I}(0)F_0(\cdot)+\hat{E}(0)\Phi_0(\cdot)+\hat{R}_{0,1}+\hat{R}_{0,2}+\hat{R}_1\Big)
\end{aligned}$$

in D^4 as $n\to\infty$. Then by Lemma 8.1 and the continuous mapping theorem, we obtain (3.18). □

As a consequence of the above proof, we also obtain the convergence $\hat{L}^n\Rightarrow\hat{L}$ in D as $n\to\infty$, jointly with the processes in (3.18), where

$$\begin{aligned}
\hat{L}(t) &= \hat{E}(0)G_0(t)+\hat{L}_0(t)+\hat{L}_1(t)\\
&\qquad+\lambda\int_0^t G(t-s)\left(\hat{S}(s)\bar{I}(s)+\bar{S}(s)\hat{I}(s)\right)ds,\quad t\geq 0.
\end{aligned}$$

8. Appendix

8.1. **A system of two linear Volterra integral equations.** Define the mapping $\Gamma:(a,x,y,z)\to(\phi,\psi)$ define by the integral equations:

$$\begin{aligned}
\phi(t) &= a+x(t)+c\int_0^t(\phi(s)z(s)+w(s)\psi(s))ds\\
\psi(t) &= y(t)+c\int_0^t K(t-s)(\phi(s)z(s)+w(s)\psi(s))ds,
\end{aligned}\tag{8.1}$$

where $(a,x,y,z)\in\mathbb{R}\times D^3$, and $c>0$ and $w\in C$. (Here c and w are given and fixed.) We study the existence and uniqueness of its solution and the continuity property in the Skorohod J_1 topology.

Lemma 8.1. *Assume that $K(0)=0$ and $K(\cdot)$ is measurable, bounded and continuous, and let $c>0$ and $w \in C$ be given. There exist a unique solution $(\phi,\psi)\in D^2$ to the integral equations (8.1). The mapping Γ is continuous in the Skorohod topology, that is, if $a^n \to a$ in $\mathbb{R}$ and $(x^n,y^n,z^n)\to(x,y,z)$ in D^3 as $n\to\infty$ with $(x,z)\in C^2$ and $y\in D$, then $(\phi^n,\psi^n)\to(\phi,\psi)$ in D^2 as $n\to\infty$. In addition, if $y\in C$, then $(\phi,\psi)\in C^2$, and the mapping is continuous uniformly on compact sets in $[0,T]$.*

Proof. By Theorems 1.2 and 2.3 in Chapter II of [16], if $x,y\in C$, we have existence and uniqueness of a solution $(\phi,\psi)\in C^2$ to the integral equations (8.1). The proof can be easily extended to the case where $x,y\in D$ by applying the Schauder-Tychonoff fixed point theorem.

We next show the continuity in the Skorohod J_1 topology. Note that the functions in D are necessarily bounded. For the given $(x,z)\in C^2$ and $y\in D$, let the interval right end point T be a continuity point of y. Since $(x,z)\in C^2$, the convergence $(x^n,y^n,z^n)\to(x,y,z)$ in D^3 in the product J_1 topology is equivalent to convergence $(x^n,y^n,z^n)\to(x,y,z)$ in $D([0,T],\mathbb{R}^3)$ in the strong J_1 topology. Then there exist increasing homeomorphisms λ^n on $[0,T]$ such that $\|\lambda^n-e\|_T\to0$, $\|x^n-x\circ\lambda^n\|_T\to0$, $\|y^n-y\circ\lambda^n\|_T\to0$, and $\|z^n-z\circ\lambda^n\|_T\to0$, as $n\to\infty$. Here $e(t):=t$ for all $t\ge0$. Moreover, it suffices to consider homeomorphisms λ^n that are absolutely continuous with resect to the Lebesgue measure on $[0,T]$ having derivatives $\dot\lambda^n$ satisfying $\|\dot\lambda^n-1\|_T\to0$ as $n\to\infty$. Let $\sup_{t\in[0,T]}|K(t)|\le c_K$.

We have

$$\begin{aligned}
|\phi^n(t)-\phi(\lambda^n(t))| &\le |a^n-a|+\|x^n-x\circ\lambda^n\|_T\\
&\qquad + c\left|\int_0^t(\phi^n(s)z^n(s)+w(s)\psi^n(s))ds-\int_0^{\lambda^n(t)}(\phi(s)z(s)+w(s)\psi(s))ds\right|\\
&\le |a^n-a|+\|x^n-x\circ\lambda^n\|_T\\
&\qquad + c\bigg|\int_0^t(\phi^n(s)z^n(s)+w(s)\psi^n(s))ds\\
&\qquad\quad -\int_0^t(\phi(\lambda^n(s))z(\lambda^n(s))+w(\lambda^n(s))\psi(\lambda^n(s)))\dot\lambda^n(s)ds\bigg|\\
&\le |a^n-a|+\|x^n-x\circ\lambda^n\|_T\\
&\qquad + c\|\dot\lambda^n-1\|_T\int_0^T|\phi(s)z(s)+w(s)\psi(s)|ds\\
&\qquad + c\int_0^t\Big(|\phi^n(s)-\phi(\lambda^n(s))||z^n(s)|+|\phi(\lambda^n(s))||z^n(s)-z(\lambda^n(s))|\\
&\qquad\quad +|w(s)-w(\lambda^n(s))||\psi^n(s)|+|w(\lambda^n(s))||\psi^n(s)-\psi(\lambda^n(s))|\Big)ds
\end{aligned}$$

and similarly,

$$\begin{aligned}
|\psi^n(t)-\psi(\lambda^n(t))| &\le \|y^n-y\circ\lambda^n\|_T\\
&\qquad + c\bigg|\int_0^t K(t-s)(\phi^n(s)z^n(s)+w(s)\psi^n(s))ds\\
&\qquad\quad -\int_0^{\lambda^n(t)}K(t-s)(\phi(s)z(s)+w(s)\psi(s))ds\bigg|\\
&\le \|y^n-y\circ\lambda^n\|_T+c\times c_K\|\dot\lambda^n-1\|_T\int_0^T|\phi(s)z(s)+w(s)\psi(s)|ds
\end{aligned}$$

$$+ c \times c_K \int_0^t \Big(|\phi^n(s) - \phi(\lambda^n(s))||z^n(s)| + |\phi(\lambda^n(s))||z^n(s) - z(\lambda^n(s))|$$
$$+ |w(s) - w(\lambda^n(s))||\psi^n(s)| + |w(\lambda^n(s))||\psi^n(s) - \psi(\lambda^n(s))| \Big) ds$$

By first applying Gronwall's inequality and then using the convergence of $a^n \to a$ in $\mathbb{R}$ and $(x^n, y^n, z^n) \to (x, y, z)$ in D^3, and $w \in C$, we obtain

$$\|\phi^n - \phi \circ \lambda^n\|_T + \|\psi^n - \psi \circ \lambda^n\|_T \to 0 \quad \text{as} \quad n \to \infty.$$

This completes the proof of the continuity property in the Skorohod J_1 topology. If $y \in C$, the continuity property is straightforward. □

8.2. **Proof of Proposition 2.1.**

Proof of Proposition 2.1. Recall that the unique solution of the linear differential equation: $x(t) = x(0) + a\int_0^t x(s)ds + y(t)$ with $y(0) = 0$, is given by the formula $x(t) = e^{at}x(0) + \int_0^t ae^{a(t-s)}y(s)ds + y(t)$, for $t \geq 0$, and if $y \in C^1$, we have $x(t) = e^{at}x(0) + \int_0^t e^{a(t-s)}\dot{y}(s)ds$.

Let $X_1(t) = \hat{I}(0)e^{-\mu t}$. We have

$$X_1(t) = -\mu \int_0^t X_1(s)ds + \hat{I}(0). \tag{8.2}$$

Let

$$X_2(t) = \lambda \int_0^t e^{-\mu(t-s)}(1 - 2\bar{I}(s))\hat{I}(s)ds.$$

We have

$$X_2(t) = -\mu \int_0^t X_2(s)ds + \lambda \int_0^t (1 - 2\bar{I}(s))\hat{I}(s)ds. \tag{8.3}$$

For $\hat{I}_0(t)$, its covariance is

$$\text{Cov}(\hat{I}_0(t), \hat{I}_0(t')) = \bar{I}(0)(e^{-\mu(t \vee t')} - e^{-\mu t}e^{-\mu t'}), \quad t, t' \geq 0.$$

It is easy to verify that

$$\hat{I}_0(t) = -\mu \int_0^t \hat{I}_0(s)ds + W_0(t) \tag{8.4}$$

where $W_0(t) = \bar{I}(0)^{1/2}B_0(1 - e^{-\mu t})$ for a standard Brownian motion B_0. We can represent $W_0(t) = \bar{I}(0)^{1/2}\int_0^t \sqrt{\mu e^{-\mu s}}d\tilde{B}_0(s)$ for another Brownian motion B_0, and thus write

$$\hat{I}_0(t) = \bar{I}(0)^{1/2} \int_0^t e^{-\mu(t-s)}\sqrt{\mu e^{-\mu s}}d\tilde{B}_0(s), \quad t \geq 0,$$

which gives the same covariance as above by Itô's isometry property.

For $\hat{I}_1$, its covariance is

$$\text{Cov}(\hat{I}_1(t), \hat{I}_1(t')) = \lambda \int_0^{t \wedge t'} e^{-\mu(t \vee t' - s)}(1 - \bar{I}(s))\bar{I}(s)ds, \quad t, t' \geq 0. \tag{8.5}$$

We next show that

$$\hat{I}_1(t) = -\mu \int_0^t \hat{I}_1(s)ds + W_1(t) \tag{8.6}$$

where $W_1(t)$ is a continuous Gaussian process, independent of $W_0(t)$, with the covariance function

$$\text{Cov}(W_1(t), W_1(t')) = \int_0^{t \wedge t'} \theta(r)dr$$

where

$$\theta(r) := \lambda(1 - \bar{I}(r))\bar{I}(r) + \mu\bar{I}(r) - \bar{I}(0)\mu e^{-\mu r}.$$

We have

$$\hat{I}_1(t) = -\mu \int_0^t e^{-\mu(t-s)} W_1(s)ds + W_1(t), \quad t \geq 0.$$

We compute the covariance $\mathrm{Cov}(\hat{I}_1(t), \hat{I}_1(s))$ using this expression: for $t > s$,

$$\begin{aligned}
\mathrm{Cov}(\hat{I}_1(t), \hat{I}_1(s)) = \mathbb{E}\left[W_1(t)W_1(s)\right] &- \mu \mathbb{E}\left[W_1(t)\int_0^s e^{-\mu(s-r)} W_1(r)dr\right] \\
&- \mu \mathbb{E}\left[W_1(s)\int_0^t e^{-\mu(t-r)} W_1(r)dr\right] \\
&+ \mu^2 \mathbb{E}\left[\int_0^t \left(\int_0^s e^{-\mu(t-r)} e^{-\mu(s-r')} W_1(r)W_1(r')dr'\right) dr)\right].
\end{aligned}$$

The first term is

$$\mathbb{E}\left[W_1(t)W_1(s)\right] = \int_0^s \theta(u)du.$$

The second term is

$$\begin{aligned}
-\mu \int_0^s e^{-\mu(s-r)} \mathbb{E}\left[W_1(t)W_1(r)\right] dr &= -\mu \int_0^s e^{-\mu(s-r)} \left(\int_0^r \theta(u)du\right) dr \\
&= -\int_0^s (1 - e^{-\mu(s-r)})\theta(r)dr.
\end{aligned}$$

The third term is

$$\begin{aligned}
&- \mu \int_0^t e^{-\mu(t-r)} \mathbb{E}\left[W_1(s)W_1(r)\right] dr \\
&= -\mu \int_0^s e^{-\mu(t-r)} \left(\int_0^r \theta(u)du\right) dr - \mu \int_s^t e^{-\mu(t-r)} \left(\int_0^s \theta(u)du\right) dr \\
&= -e^{-\mu(t-s)} \int_0^s (1 - e^{-\mu(s-r)})\theta(r)dr - (1 - e^{-\mu(t-s)}) \int_0^s \theta(u)du \\
&= -\int_0^s (1 - e^{-\mu(t-r)})\theta(r)dr.
\end{aligned}$$

The fourth term is

$$\begin{aligned}
&\mu^2 \int_0^t \left(\int_0^s e^{-\mu(t-r)} e^{-\mu(s-r')} \mathbb{E}[W_1(r)W_1(r')]dr'\right) dr \\
&= \mu^2 \int_s^t \left(\int_0^s e^{-\mu(t-r)} e^{-\mu(s-r')} \mathbb{E}[W_1(r)W_1(r')]dr'\right) dr \\
&\quad + \mu^2 \int_0^s \left(\int_0^s e^{-\mu(t-r)} e^{-\mu(s-r')} \mathbb{E}[W_1(r)W_1(r')]dr'\right) dr \\
&= \mu^2 \int_s^t \left(\int_0^s e^{-\mu(t-r)} e^{-\mu(s-r')} \left(\int_0^{r'} \theta(u)du\right) dr'\right) dr \\
&\quad + 2\mu^2 \int_0^s \left(\int_0^r e^{-\mu(t-r)} e^{-\mu(s-r')} \left(\int_0^{r'} \theta(u)du\right) dr'\right) dr \\
&= (1 - e^{-\mu(t-s)}) \int_0^s (1 - e^{-\mu(s-r)})\theta(r)dr \\
&\quad + e^{-\mu(t-s)} \int_0^s (1 - 2e^{-\mu(s-r)} + e^{-2\mu(s-r)})\theta(r)dr
\end{aligned}$$

$$= \int_0^s (1 - e^{-\mu(s-r)} - e^{-\mu(t-r)} + e^{-\mu(t-r)-\mu(s-r)})\theta(r)dr.$$

Combining the four terms, we obtain

$$e^{-\mu(t-s)} \int_0^s e^{-2\mu(s-r)}\theta(r)dr.$$

Now we check that this is equal to the covariance of $\hat{I}_1(t)$ in (8.5). Taking the difference between the last expression and the right–hand side of (8.5) with $t' = s < t$, we obtain

$$\begin{aligned} & e^{-\mu(t-s)} \int_0^s e^{-2\mu(s-r)}\theta(r)dr - \lambda \int_0^s e^{-\mu(t-r)}(1-\bar{I}(r))\bar{I}(r)dr \\ &= e^{-\mu(t-s)} \int_0^s e^{-2\mu(s-r)}(\lambda(1-\bar{I}(r))\bar{I}(r) + \mu\bar{I}(r))dr - \lambda e^{-\mu(t-s)} \int_0^s e^{-\mu(s-r)}(1-\bar{I}(r))\bar{I}(r)dr \\ &\quad - e^{-\mu(t-s)} \int_0^s e^{-2\mu(s-r)}\bar{I}(0)\mu e^{-\mu r}dr. \end{aligned} \tag{8.7}$$

Observe that the fluid equation for $I(t)$ can be written as

$$\bar{I}'(t) = -\mu\bar{I}(t) + \lambda\bar{I}(t)(1-\bar{I}(t))$$

and

$$\bar{I}'(t) = -2\mu\bar{I}(t) + \lambda\bar{I}(t)(1-\bar{I}(t)) + \mu\bar{I}(t).$$

These two equations give the following representations of $\bar{I}(t)$:

$$\bar{I}(t) = \bar{I}(0)e^{-\mu s} + \lambda \int_0^s e^{-\mu(s-r)}\bar{I}(r)(1-\bar{I}(r))dr,$$

and

$$\bar{I}(t) = \bar{I}(0)e^{-2\mu s} + \int_0^s e^{-2\mu(s-r)}\left(\lambda\bar{I}(r)(1-\bar{I}(r)) + \mu\bar{I}(r)\right)dr.$$

Also notice that $\int_0^s e^{-2\mu(s-r)}\mu e^{-\mu r}dr = e^{-\mu s} - e^{-2\mu s}$. Using these equations, we verify that (8.7) is equal to zero, and thus the equation for $\hat{I}_1$ in (8.6) is established. Therefore, by combining (8.2), (8.3) (8.4) and (8.6), we obtain the equivalence of the non-Markovian and Markovian representations of $\hat{I}$ for the SIS model. □

References

[1] Roy M. Anderson and Robert M. May. *Infectious diseases of humans: dynamics and control*. Oxford university press, 1992.

[2] Hakan Andersson and Tom Britton. *Stochastic epidemic models and their statistical analysis*, volume 151. Springer Science & Business Media, 2012.

[3] Frank Ball. A unified approach to the distribution of total size and total area under the trajectory of infectives in epidemic models. *Advances in Applied Probability*, 18(2):289–310, 1986.

[4] Frank Ball and Damian Clancy. The final size and severity of a generalised stochastic multitype epidemic model. *Advances in applied probability*, 25(4):721–736, 1993.

[5] Andrew D Barbour. The duration of the closed stochastic epidemic. *Biometrika*, 62(2):477–482, 1975.

[6] Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons, 1999.

[7] Fred Brauer. On a nonlinear integral equation for population growth problems. *SIAM Journal on Mathematical Analysis*, 6(2):312–317, 1975.

[8] Tom Britton and Etienne Pardoux. Stochastic epidemics in a homogeneous community. *arXiv preprint arXiv:1808.05350*, 2018.

[9] Damian Clancy. Sir epidemic models with general infectious period distribution. *Statistics & Probability Letters*, 85:1–5, 2014.

[10] Kenneth L Cooke. An epidemic equation with immigration. *Mathematical Biosciences*, 29(1-2):135–158, 1976.

[11] Laurent Decreusefond, Pascal Moyal, et al. A functional central limit theorem for the $M/GI/\infty$ queue. *The Annals of Applied Probability*, 18(6):2156–2178, 2008.

[12] Odo Diekmann. Limiting behaviour in an epidemic model. *Nonlinear Analysis: Theory, Methods & Applications*, 1(5):459–470, 1977.

[13] A Gómez-Corral and M López-García. On SIR epidemic models with generally distributed infectious periods: Number of secondary cases and probability of infection. *International Journal of Biomathematics*, 10(02):1750024, 2017.

[14] Herbert W Hethcote and Pauline van den Driessche. An SIS epidemic model with variable population size and a delay. *Journal of mathematical biology*, 34(2):177–194, 1995.

[15] Elena V Krichagina and Anatolii A Puhalskii. A heavy-traffic analysis of a closed queueing system with a GI/∞ service center. *Queueing Systems*, 25(1-4):235–280, 1997.

[16] R. K. Miller. *Nolinear Volterra Integral Equations.* Benjiamin Press, Menlo Park, Cal., 1971.

[17] Guodong Pang, Rishi Talreja, and Ward Whitt. Martingale proofs of many-server heavy-traffic limits for markovian queues. *Probability Surveys*, 4:193–267, 2007.

[18] Guodong Pang and Ward Whitt. Two-parameter heavy-traffic limits for infinite-server queues. *Queueing Systems*, 65(4):325–364, 2010.

[19] Josh Reed, Rishi Talreja, et al. Distribution-valued heavy-traffic limits for the $G/G/\infty$ queue. *The Annals of Applied Probability*, 25(3):1420–1474, 2015.

[20] Gesine Reinert. The asymptotic evolution of the general stochastic epidemic. *The Annals of Applied Probability*, 5(4):1061–1086, 1995.

[21] Thomas Sellke. On the asymptotic distribution of the size of a stochastic epidemic. *Journal of Applied Probability*, 20(2):390–394, 1983.

[22] Pauline van den Driessche and James Watmough. A simple sis epidemic model with a backward bifurcation. *Journal of Mathematical Biology*, 40(6):525–540, 2000.

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

Aix–Marseille Université, CNRS, Centrale Marseille, I2M, UMR 7373 13453 Marseille, France
Email address: etienne.pardoux@univ.amu.fr