Stochastic SIR model with individual heterogeneity and infection-age dependent infectivity on large non-homogeneous random graphs

GUODONG PANG, ÉTIENNE PARDOUX, AND AURÉLIEN VELLERET

Abstract. We study an individual-based stochastic SIR epidemic model with infection-age dependent infectivity on a large random graph, capturing individual heterogeneity and non-homogeneous connectivity. Each individual is associated with particular characteristics (for example, spatial location and age structure), which may not be i.i.d., and represented by a particular node. The connectivities among the individuals are given by a non-homogeneous random graph, whose connecting probabilities may depend on the individual characteristics of the edge. Each individual is associated with a random varying infectivity function, which is also associated with the individual characteristics. We use measure-valued processes to describe the epidemic evolution dynamics, tracking the infection age of all individuals, and their associated characteristics. We consider the epidemic dynamics as the population size grows to infinity under a specific scaling of the connectivity graph related to the convergence to a graphon. In the limit, we obtain a system of measure-valued equations, which can be also represented as a PDE model on graphon, which reflects the heterogeneities in individual characteristics and social connectivity.

1. Introduction

In the current paper, we are interested in the large population limit of a stochastic model of an epidemic that spreads among an heterogeneous population with random connectivities. Such models have been studied to an extent in the literature. For example, in [11], an epidemic model with an age structure and various social activity levels is studied to understand the effect of population heterogeneity on herd immunity. In addition to 'age' as an obvious heterogeneous factor, spatial locations and other characteristics/features may also indicate individual heterogeneity. Another source of heterogeneity arises from the individual connectivities (such as households, communities and social activities), which is often modeled as a non-homogeneous random graph. For example, epidemic models on random graphs with given degrees, typically under the configuration model are studied in [1, 52, 49, 9, 5, 32, 24, 39, 40, 15], on weighted (configuration) graphs [12, 18, 13, 50], on dynamic (evolving) graphs [3, 2, 33, 22, 14, 23, 28]. (See also the relevant models on random networks with household structures in [7, 4, 6, 38], and on multilayer networks in [29, 38].)

A few papers have established large population limits for epidemic models on random graphs. In particular, [17, 31] proved a functional law of large numbers (FLLN) for a Markovian SIR model on a configuration model graph with specified degree distributions and edges being randomly matched, and established the measure-valued limit as a systems of nonlinear differential equations, which verifies the conjecture in [52]. In [37], a functional

Date: May 22, 2025.

Key words and phrases. Individual-based stochastic epidemic models, SIR, infection-age dependent infectivity, individual heterogeneity, large non-homogeneous random graph, graphon, measure-valued processes, PDE model on graphon.

central limit theorem (FCLT) is established for a similar Markovian SI model for the total count processes. In [10], an FLLN is established for a Markovian SIR model on a stochastic block model. In [17, 31, 37, 10], the degree distribution converges to a finite limit as the number of nodes tends to infinity and the number of edges is thus of the same order as the number n of nodes. Such a description with very sparse graphs leads to limits that are of distinct nature from the one of dynamics on a graphon. On the other hand, for dense connectivities among individuals, one can derive large population approximations of the epidemic dynamics as PDEs on graphons. In [36], an FLLN is established for density-dependent Markov processes with finite state space on large random graphs, which includes the Markovian SIS model with individual heterogeneity on graphs as a special case, and a PDE on graphon is derived as the limit. In [20], an FLLN is established for a Markovian SIS model with a general form of individual heterogeneity on random graphs, where a PDE on graphon for the measure-valued state descriptors is also derived. In the two later references ([36, 20]) the scalings of the parameters allow for the random generation of graphs whose number of edges is of order n^{1+a}, where n is the number of nodes and a can be freely taken in $(0, 1]$. This extends the classical assumption of a dense graph for the convergence to a graphon, where the number of edges is of order n^2. However, all these FLLN results are about epidemic models on random graphs that are Markovian, where the infectious durations or recovery times are exponentially distributed. In the present paper, we establish an FLLN for a non-Markovian SIR model on large random graphs that results in a measure-valued limit on a graphon.

Non-Markovian stochastic epidemic models with a homogeneous population (no age structure, homogeneous social connectivity, etc.) have been recently studied, see the recent survey [27]. Although these models offer much more possibilities to fit observational data on infection profiles, many of the classical tools in probability cannot be directly exploited. In particular, the standard epidemic models with a constant infection rate and a general infection duration distribution are recently studied in [42] (see also [54, 53] for Gaussian approximations and [48] for measure-valued state descriptors and PDE limit for the FLLN), and epidemic models with varying infectivity are studied in [25, 43, 26], where deterministic or stochastic integral equations are derived for the FLLNs and FCLTs, respectively. By tracking the elapsed infectious times, measure-valued processes are used to describe the epidemic evolution dynamics in models with such infection-age dependent infectivity, and the corresponding PDE and SPDE limits are established for the FLLNs and FCLTs in [44] and [45], respectively. An epidemic model with contact tracing and general infection duration distributions is studied in [21], where measure-valued processes are used to describe the dynamics and the FLLN is established with a PDE limit. In [38], an individual-based multi-layer SIR model with households and workplaces to take into account social connectivity heterogeneity is studied, where measure-valued processes tracking remaining infectious times are used to describe the epidemic dynamics and a PDE limit is established for the FLLN.

In this paper we consider an SIR model with infection-age dependent infectivity on a large random graph that captures both individual and connectivity heterogeneities. In particular, connectivities among individuals are given by a non-homogeneous random graph. Each node on this graph represents a single individual, each associated with an individual characteristic/feature. These may account for spatial locations, age and/or social belongings. So they may be distributed on a compact subset of $\mathbb{R}^d$, as in the spatial SIR model considered in [35], which we generalize in the simplest setting of pairwise interactions. In the case of age structure or social activity levels, our model generalizes [11] in that they only consider a finite number of compartments while our model includes a possibly continuous age model. The random connectivity of each edge is then assumed to depend on the individuals' characteristics in a general manner. The connectivities indicate various levels of social activities, which may for instance depend on the spatial locations or age (again more generally than the model with age and social activity heterogeneities in [11]). The random graphs that we consider are of the same form as the ones in [20]. The main difference is that the connectivity of each edge was assumed in [20] to be a deterministic function of the pair of individuals' characteristics, while we allow for an additional degree of randomness. In [36], there was no heterogeneity in the contact rate and the dependency of the graph on individual types is more restrictive. Our model extends the framework of stochastic block models considered in [10] except that the scalings of the contact probability are compatible only in the degenerate cases.

Each individual is moreover associated with a random varying infectivity function, which reflects the propensity for/ force of infection of each individual at any elapsed time of infection, and depends as well on the individual characteristics. As a result, the infectious duration is also individualized (we express the corresponding distribution as a function $F_x(\cdot)$ of the individual characteristic x).

Individuals are grouped into three compartments: Susceptible, Infected and Recovered. Infections are generated by the interactions between susceptible and infected individuals. Each individual is thus exposed to a specific force of infection which changes over time and which is given by aggregating the weighted infectivity levels of the individuals that are connected to him/her in the graph. The whole dynamic of the epidemic depends on the non-homogeneous random graph and on the random infectivity functions through the expression of this total force of infection for each individual.

To describe the evolution dynamics, we use three measure-valued processes for the three compartments, where the measure for the infected process is over both the individual characteristics and the infection ages, while the measure for the susceptible and recovered processes is only over the individual characteristics. We prove a FLLN for these measure-valued processes when the population size goes to infinity. As the population size increases, so does the connectivity graph. We impose conditions on the connectivity probabilities to keep the graph consistency as it grows, which notably covers the case of a graphon in the limit. The limits for the FLLN are given by a set of measure-valued equations, from which we further derive a PDE model for the measure-valued infection process. The PDE model is linear, but with a boundary condition which is given by the product of the measure-valued susceptible process and the aggregate force of infection. To note, the dynamics described from the set of measure-valued equations and the PDE may be seen as evolving on a graphon.

The proof for the FLLN is challenging with this level of realism because of the complicated dependence among the variables and processes. First of all, the individual characteristics are not assumed to be i.i.d., and they affect the random infectivity functions and the edge connectivities. Second, the measure-valued processes all depend on the aggregate forces of infection that act specifically on each individual. Notably, these forces of infection are not functions of the three measure-valued state descriptors, but more generally expressed in terms of the detailed graph structure. Thus it is difficult to prove the convergence of the processes directly.

Our approach consists in constructing an intermediate model for which the infection rates are derived from the limiting model of the FLLN rather than the interactions through the connectivity graph. This approach facilitates the proof of convergence by enabling specific decoupling of the dependence. It extends the one that was first developed for the homogeneous model in [26], which was motivated by ideas from the theory of propagation of chaos [51]. In this way, the regularity conditions imposed on the random varying infectivity functions in [25] could be relaxed to only require their boundedness. The current extension is highly non-trivial however because of the heterogeneities among both the individuals and the connectivities. Moreover, we consider measure-valued processes in this paper, while in [26], only total count processes are considered. Different weak convergence criteria are therefore required for this setting.

In addition, we have made the effort to allow for the individual characteristics to take values in a state space as general as possible instead of the classical choice of $[0,1]$ for graphon kernels. This is particularly of importance regarding the regularity assumptions of the interaction kernel in the limit, which we only assume to be almost everywhere continuous. One of our motivations was to allow for kernels on multidimensional state spaces that possibly display discontinuity thresholds in terms of the distance between the two parameters, as discussed in [20]. We prove in Appendix A a generic convergence result of independent interest, to deal with these weak regularity properties in conjunction with weak convergence of random measures in probability. The approach demonstrates its potential by extending very efficiently to this level of heterogeneity, even with the randomness additionally considered for the varying infectivity functions. The conditions that we assume on the average infection rate per edge and on the average variance in this infection rate seem very close to optimal. We refer the reader to Section 5 for the detailed construction of the intermediate processes and for the proofs of weak convergence.

1.1. **Organization of the paper.** The rest of the paper is organized as follows. We give a summary of some common notations used throughout the paper in the next subsection. We provide a detailed model description and assumptions in Section 2. In Section 3, we present the main result of the paper. We establish the uniqueness of the solution to the limiting deterministic PDE model and establish some properties of that solution in Section 4. The proofs for the convergence of the FLLN are given in Section 5, with additional technical supporting results proved in Appendix A.

1.2. **Notation.** We denote by $\mathbb{N}^*$ the set of positive integers and $\mathbb{N} = \mathbb{N}^* \cup \{0\}$. We also set $[\![1, n]\!] = \{k \in \mathbb{N} :\ 1 \le k \le n\}$ for $n \in \mathbb{N}^*$. For $a, b \in \mathbb{R}$, we write $a \wedge b$ for the minimum between a and b and $a \vee b$ for the maximum between a and b. For a real-valued function defined on a set $\mathcal{X}$, we denote by $\|f\|_\infty$ its supremum norm with

$$\|f\|_\infty = \sup_{\mathcal{X}} |f|.$$

We denote by $C_b(\mathcal{X})$ the set of continuous and bounded functions on a metric space $\mathcal{X}$.

For a measurable set $(\mathcal{X}, \mathcal{F})$, we denote by $\mathcal{M}_1(\mathcal{X})$ and $\mathcal{M}(\mathcal{X})$ the sets of respectively probability and non-negative finite measures on $\mathcal{X}$. For $\mu \in \mathcal{M}(\mathcal{X})$ and a real-valued measurable function f defined on $\mathcal{X}$, we will sometimes denote the integral of f with respect to the measure μ, if well-defined, by $\langle \mu,\, f\rangle = \int_{\mathcal{X}} f(x)\, \mu(\mathrm{d}x) = \int f\, \mathrm{d}\mu$. For a metric space $\mathcal{X}$ endowed with its Borel σ-field, we endow $\mathcal{M}_1(\mathcal{X})$ with the topology of weak convergence.

We also denote by $\mathcal{D}(\mathbb{R}_+, \mathcal{X})$ the space of right-continuous left-limited (càd-làg) paths from $\mathbb{R}_+$ to $\mathcal{X}$. This space is endowed with the Skorokhod topology (see, *e.g.*, [8, Chapter 3]).

In what follows $(\mathbb{X}, \mathrm{d})$ will denote a Polish metric space (complete and separable).

2. Model description

We consider a population consisting of N individuals, who are in either of the three states – susceptible, infected or recovered at each time, and may interact with each other according to a random graph described below. Susceptible individuals can be infected randomly through interactions with the infected ones. Once infected, an individual will become infectious for a random period of time until recovery, and once recovered, he or she will no longer infect any susceptibles, or become infected again. Let $\mathcal{S}^N(t), \mathcal{I}^N(t), \mathcal{R}^N(t)$ be the subsets of $[\![1, N]\!]$ that denote the sets of susceptible, infected or recovered individuals at each time t, respectively. The corresponding processes $S^N(t) = |\mathcal{S}^N(t)|$, $I^N(t) = |\mathcal{I}^N(t)|$ and $R^N(t) = |\mathcal{R}^N(t)|$ denote the numbers of susceptible, infected or recovered individuals at each time t.

In the limit of a large population, our goal is to relate the spread of the disease in the population of N individuals to a dynamics acting on a Polish space $\mathbb{X}$ that accounts for the heterogeneity of individuals. Any individual $i \in [\![1, N]\!]$ is actually characterized by some random variable $X_i^N \in \mathbb{X}$ (which is fixed over time). The characteristic X_i^N of individual i is allowed to affect the infectious contact rates with the other individuals as well as the evolution of its own infectivity level.

2.1. Infection-age dependent infectivity.

For an individual $i \in \mathcal{S}^N(0)$, let $\tau_i^N > 0$ be the time at which he/she starts being infected, and let $A_i^N(t) = t - \tau_i^N$ be the infection age of the individual i at time t (by default it is equal to zero for $t < \tau_i^N$). For an initially infected individual $j \in \mathcal{I}^N(0)$, let $\tau_j^N = -A_j^N(0)$ denote the infection time before time 0, so that $A_j^N(t) = t + A_j^N(0)$ is the infection age at time t.

Let $\mathcal{F}_0^N$ be the σ-field generated by the $(X_i^N, i \le N)$, $\mathcal{S}^N(0)$, $\mathcal{I}^N(0)$ and $(A_j^N(0), j \in \mathcal{I}^N(0))$. We shall generally consider the dynamics conditional on this sigma-field $\mathcal{F}_0^N$. Thus, we shall use the notations $\mathbb{P}_0^N$ and $\mathbb{E}_0^N$ to express probabilities and expectations conditional on $\mathcal{F}_0^N$, respectively.

At time t, each individual $i \in \mathcal{I}^N(t)$ who is infectious has an infectivity function depending randomly on the infection age:

$$\lambda_i^N(A_i^N(t)), \quad t \ge 0. \tag{2.1}$$

For the initially infected individuals $j \in \mathcal{I}^N(0)$, we also define their infectivity level at a given time $t > 0$ as $\lambda_j^{N,0}(t)$:

$$\lambda_j^{N,0}(t) := \lambda_j^N(t + A_j^N(0))\,.$$

Remark 2.1. *A typical way to define these random functions $\lambda_i^N(\cdot)$ is specified as follows, although we do not rely in our proofs on this representation. Let $\mathbb{Y}$ be a Polish space. Let $\mathcal{Y}^N = \{Y_i^N\}_{i \le N}$ be an i.i.d. sequence of random variables taking values in $\mathbb{Y}$, with common distribution μ_Y. We assume that the sequence $\mathcal{Y}^N = \{Y_i^N\}_{i \le N}$ is independent from $\mathcal{F}_0^N$. Let $\widehat{\lambda}$ be a deterministic measurable function from $\mathbb{X} \times \mathbb{Y} \times \mathbb{R}_+$ to $\mathbb{R}_+$. Then we can define the random functions $\lambda_i^N(\cdot)$ for each i as*

$$\lambda_i^N(a) = \widehat{\lambda}(X_i^N, Y_i^N, a)\,. \tag{2.2}$$

The variables $\{Y_i^N\}$ capture the random factors associated with the infectivity of the individuals apart from their characteristics and their infection age.

Assumption 2.2. *There exists a global upper-bound $\lambda^* < \infty$ on the random functions λ_i^N taking values in $\mathcal{D}(\mathbb{R}_+, \mathbb{R}_+)$, in the sense that the following holds almost surely:*

$$\lambda_i^N(a) \leq \lambda^*, \quad \forall N,\ \forall i \in [\![1, N]\!],\ \forall a \in \mathbb{R}_+ .$$

Remark 2.3. *For Assumption 2.2 to hold under the formalism presented in Remark 2.1, it is sufficient to require the function $\widehat{\lambda}$ to be uniformly upper-bounded by this value λ^*.*

The infection duration η_i^N of individual $i \in \mathcal{S}^N(0)$ is defined as

$$\eta_i^N := \sup\{a > 0 : \lambda_i^N(a) > 0\} .$$

Likewise, we consider the infection duration $\eta_j^{N,0}$ of an initially infected individual $j \in \mathcal{I}^N(0)$ with initial infection age $A^N(0)$:

$$\eta_j^{N,0} := \sup\{a > 0 : \lambda_j^N(A^N(0) + a) > 0\} .$$

We specify the individual distribution of infection duration through its cumulative distribution. Under the formalism of Remark 2.1 it is to be thought as depending only on the variables (X_i^N, Y_i^N).

Assumption 2.4. *There exists a measurable function $(F_x(a))_{x\in\mathbb{X}, a\in\mathbb{R}_+}$ with values in $[0, 1]$ which is right-continuous, non-decreasing in the variable 'a' such that $\lim_{a\to\infty} F_x(a) = 1$ for all $x \in \mathbb{X}$ and that the two following identities hold for any $N \geq 1$, $t > 0$, $i \in \mathcal{S}^N(0)$ and $j \in \mathcal{I}^N(0)$ with the notation $F_x^c(a) = 1 - F_x(a)$:*

$$\mathbb{P}_0^N(\eta_i^N > t) = F^c_{X_i^N}(t),$$
$$\mathbb{P}_0^N(\eta_j^{N,0} > t) = \frac{F^c_{X_j^N}(A_j^N(0) + t)}{F^c_{X_j^N}(A_j^N(0))} .$$

Recall that $\mathbb{P}_0^N(\eta_i^N > t)$ equals by definition $\mathbb{P}\big(\eta_i^N > t \,\big|\, \mathcal{F}_0^N\big)$ and similarly that $\mathbb{P}_0^N(\eta_j^{N,0} > t) = \mathbb{P}\big(\eta_j^{N,0} > t \,\big|\, \mathcal{F}_0^N\big)$.

Remark 2.5. *This assumption appear more explicitly stated under the formalism presented in Remark 2.1. By definition, there exists a deterministic measurable function $g : \mathbb{X}\times\mathbb{Y} \to \mathbb{R}_+$ such that $\eta_i^N = g(X_i^N, Y_i^N)$. For each $x \in \mathbb{X}$, we define the cumulative distribution:*

$$F_x(t) = \mathbb{P}(g(x, Y_1^N) \leq t) = \int_{\mathbb{Y}} \mathbb{1}_{\{g(x,y)\leq t\}} \mu_Y(dy)$$

and set $F_x^c(\cdot) = 1 - F_x(\cdot)$. F_x^c is clearly non-increasing for any $x \in \mathbb{X}$, with values in $[0, 1]$. As a consequence, we check the first identity in Assumption 2.4, for any $N \geq 1$, any $i \in \mathcal{S}^N(0)$, and any $t > 0$:

$$\mathbb{P}_0^N(\eta_i^N > t) = \mathbb{P}\big(g(X_i^N, Y_i^N) > t \,\big|\, X_i^N\big) = F^c_{X_i^N}(t).$$

For any N and any initially infected individual $j \in \mathcal{I}^N(0)$, the constraint that $\eta_j^{N,0} > 0$, which translates into $g(X_j^N, Y_j^N) > A_j^N(0)$, explains the form of the second identity in

Assumption 2.4, where for any $t > 0$,

$$\begin{aligned}\mathbb{P}_0^N(\eta_j^{N,0} > t) &= \mathbb{P}\big(g(X_j^N, Y_j^N) > t + A_j^N(0) \,\big|\, X_j^N, A_j^N(0), \mathbb{1}_{\{g(X_j^N,Y_j^N)>A_j^N(0)\}}\big)\\ &= \frac{F^c_{X_j^N}(A_j^N(0)+t)}{F^c_{X_j^N}(A_j^N(0))}.\end{aligned}$$

Assumption 2.6. *There exists a function $\bar\lambda$ in $\mathcal{D}(\mathbb{X}\times\mathbb{R}_+,\mathbb{R}_+)$ such that the two following identities hold for any $N \ge 1$, any $a \ge 0$, any $t \ge 0$, any $i \in \mathcal{S}^N(0)$ and any $j \in \mathcal{I}^N(0)$:*

$$\begin{aligned}\mathbb{E}_0^N\big[\lambda_i^N(a)\big] &= \bar\lambda(X_i^N, a),\\ \mathbb{E}_0^N\big[\lambda_j^{N,0}(t)\big] &= \frac{\bar\lambda(X_j^N, A_j^N(0)+t)}{F^c_{X_j^N}(A_j^N(0))},\end{aligned}$$

where $F^c_{X_j^N}(A_j^N(0)) > 0$ holds a.s. for any $j \in \mathcal{I}^N(0)$. Moreover, the function $\chi : \mathbb{X}\times\mathbb{R}_+ \mapsto \mathbb{R}_+$ defined as $\chi(x,a) = \bar\lambda(x,a)/F^c_x(a)$ if $F^c_x(a) > 0$ and $\chi(x,a) = 0$ otherwise, is upper-bounded by λ_*.

Recall that $\mathbb{E}_0^N\big[\lambda_i^N(a)\big]$ equals by definition $\mathbb{E}\big[\lambda_i^N(a) \,\big|\, \mathcal{F}_0^N\big]$ and similarly for $\mathbb{E}_0^N\big[\lambda_j^{N,0}(t)\big]$.

Remark 2.7. *Under the formalism presented in Remark 2.1, $\bar\lambda$ is directly expressed as follows:*

$$\bar\lambda(x,a) = \int_{\mathbb{Y}} \widehat\lambda(x,y,a)\, \mu_Y(\mathrm{d}y)\,,$$

and the first identity in Assumption 2.6 is automatic. We can also check the second identity:

$$\begin{aligned}\mathbb{E}_0^N\big[\lambda_j^{N,0}(t)\big] &= \mathbb{E}\big[\widehat\lambda(X_j^N, Y_j^N, t + A_j^N(0)) \,\big|\, X_j^N, A_j^N(0), \mathbb{1}_{\{g(X_j^N,Y_j^N)>A_j^N(0)\}}\big]\\ &= \frac{\int_{\mathbb{Y}} \widehat\lambda(X_j^N, y, t + A_j^N(0))\,\mu_Y(\mathrm{d}y)}{F^c_{X_j^N}(A_j^N(0))} = \frac{\bar\lambda(X_j^N, t + A_j^N(0))}{F^c_{X_j^N}(A_j^N(0))},\end{aligned}$$

where we have exploited in the second line the fact that $\widehat\lambda(X_j^N, y, t + A_j^N(0)) > 0$ entails $g(X_j^N, Y_j^N) > t + A_j^N(0)$ and a fortiori $g(X_j^N, Y_j^N) > A_j^N(0)$. In addition, for any $x \in \mathbb{X}$ and any $a \ge 0$ such that $F^c_x(a) > 0$, since $\widehat\lambda(x,y,a) = 0$ on the set $\{g(x,y) \le a\}$,

$$\frac{\bar\lambda(x,a)}{F^c_x(a)} = \frac{\int_{\mathbb{Y}} \widehat\lambda(x,y,a)\, \mathbb{1}_{\{g(x,y)>a\}}\, \mu_Y(\mathrm{d}y)}{\int_{\mathbb{Y}} \mathbb{1}_{\{g(x,y)>a\}}\, \mu_Y(\mathrm{d}y)}\,,$$

and is actually upper-bounded by λ_* as soon as $\widehat\lambda$ is itself upper-bounded by λ_*.

Remark 2.8. *In this work we generally allow for intricate dependencies between the possible value of $\lambda_i^N(a)$ and the infection duration of individual i. Though, a classical choice for the function $\widehat\lambda$ is to be defined in terms of two given deterministic functions $\widetilde\lambda(x,a)$ and $g(x,y)$ as follows: $\widehat\lambda(x,y,a) = \widetilde\lambda(x,a)\mathbb{1}_{\{a<g(x,y)\}}$. In this case, the expression for $\bar\lambda$ simplifies as follows:*

$$\bar\lambda(X_i^N, a) = \widetilde\lambda(X_i^N, a)\cdot F^c_{X_i^N}(a).$$

Remark 2.9. *As an instance of the case presented in the above Remark 2.8, the most standard Markovian description of a constant infectivity $\check{\lambda} > 0$ with an exponential random duration with mean $\check{g}(x)$ is obtained as follows. We can choose μ_Y to be the law of a standard exponential random variable with mean 1, on $\mathbb{Y} = \mathbb{R}_+$, and take $\widetilde{\lambda}(a, x) \equiv \check{\lambda}$ and $g(x, y) = \check{g}(x) \cdot y$. The expression for $\bar{\lambda}$ further simplifies:*

$$\bar{\lambda}(X_i^N, a) = \check{\lambda} \cdot \exp\Big(- a/\check{g}(X_i^N)\Big).$$

2.2. **Construction of the connectivity graph.** We consider the underlying connectivity graph among the individuals, which is denoted as a graph $(\mathcal{G}^N, \mathcal{E}^N)$ with $\mathcal{G}^N$ being the set of N nodes and $\mathcal{E}^N$ being the set of edges (which are undirected). We write $i \overset{N}{\sim} j$ to indicate that nodes i and j are connected, whose connectivity depends on N. We assume that the connectivity probability of nodes $i, j \in \mathcal{G}^N$ is given by the deterministic symmetric measurable function $\kappa^N : \mathbb{X} \times \mathbb{X} \to [0, 1]$:

$$\mathbb{P}_0^N\big(i \overset{N}{\sim} j\big) = \mathbb{P}\big(i \overset{N}{\sim} j \,\big|\, \mathcal{F}_0^N\big) = \kappa^N(X_i^N, X_j^N), \tag{2.3}$$

in terms of the characteristic variables X_i and X_j for individuals i and j. We assume in addition that the events $\{i \overset{N}{\sim} j\}_{i \neq j}$ are mutually independent and independent of the sequence $(Y_i^N)_{i \leq N}$ conditionally on $\mathcal{F}_0^N$.

The typical construction of such a random graph starts from a graphon kernel. For example, assuming here that $\mathbb{X} = [0, 1]$, let $\bar{\kappa}(x, x')$ be a deterministic function on $[0, 1]^2$, such as $\bar{\kappa}(x, x') = x \cdot x'$. One can sample a N-node random graph with $\kappa^N = \bar{\kappa}$ so that $\mathbb{P}_0^N\big(i \overset{N}{\sim} j\big) = \bar{\kappa}(X_i^N, X_j^N)$ for nodes $i, j = 1, \ldots, N$, where the X_i^N are uniformly and independently sampled in $[0, 1]$. Another very natural choice of X_i^N is also $X_i^N = i/N$ for $i = 1, \ldots, N$, which also leads to $\bar{\mu}_X$ being the Lebesgue measure on $[0, 1]$. In the case $\bar{\kappa}(x, x') = x \cdot x'$, the degree of individual i in the graph then scales linearly with N and with the individual type X_i^N. The fact that the degree scales linearly in N is typically how a "dense" graph is defined. Yet, we allow for more generality in the graph density with a sequence (κ^N) of functions that may scale with N. The first way to do so is to introduce a scaling factor $\varepsilon^N > 0$, typically $\varepsilon^N = N^{-\alpha}$ with $\alpha \in (0, 1)$, so that $\kappa^N = \varepsilon^N \cdot \bar{\kappa}$ as in [36]. Following [20], this form of scaling is however not assumed to allow for even more general dependencies between the pair of individual characteristics (X_i^N, X_j^N) and N. The role of the denseness of the graph will be discussed in Subsection 3.1, after the statement of our main theorem.

2.3. **Epidemic dynamics.** To describe the force of infection, we introduce a random weight function $(\omega^N(i, j))_{i,j \in [\![1,N]\!]}$ that is non-negative and equal to zero except for the edges (i, j) of $\mathcal{E}^N$. $\omega^N(i, j)$ is meant to represent the scaling factor that translates the infectivity value of individual j, typically $\lambda_j^N(A_j^N(t))$ at time t, into the rate at which individual i is subject to an infectious contact with individual j, which becomes $\omega^N(i, j) \cdot \lambda_j^N(A_j^N(t))$ at time t.

Note that heterogeneity in the susceptibility to infection can be included in the model through this values $\omega^N(i, j)$, such heterogeneity between individuals being possibly partly explained by the characteristic X_i^N. This is a major reason for us not to assume that ω^N is symmetric. Such a framework for $\omega^N(i, j)$ allows in addition more realistic and intricate relations between the contact rates of individuals i and j depending on their respective characteristics X_i^N and X_j^N. Even more generally than contact matrices typically exploited for epidemiological predictions, see, e.g., [41], we allow for the interplay between characteristics that are continuous and for an additional degree of independent randomness.

The values $(\omega^N(i,j))_{i,j\in[\![1,N]\!]}$ are assumed in addition to be mutually independent between different edges and independent of the $(\lambda_i^N)_{i\le N}$ conditionally on $\mathcal{F}_0^N$, that is, of the (Y_i^N) if we consider the formalism in Remark 2.1. The deterministic measurable function γ^N : $\mathbb{X}\times\mathbb{X}\to\mathbb{R}_+$ captures the expected infectious rate of interaction of active contacts, in that the following identity is assumed:

$$\mathbb{E}\big[\omega^N(i,j)\,\big|\,\mathcal{F}_0^N,\mathcal{E}^N\big] = \begin{cases}\gamma^N(X_i^N,X_j^N) & \text{if } (i,j)\in\mathcal{E}^N,\\ 0 & \text{otherwise.}\end{cases} \tag{2.4}$$

We may also allow for a degree of variability in the infectious rate of interaction, that we synthetize with the following variance term υ^N, for any $(i,j)\in\mathcal{E}^N$:

$$\upsilon^N(i,j) = \mathrm{Var}\big[\omega^N(i,j)\,\big|\,\mathcal{F}_0^N,\mathcal{E}^N\big]\,. \tag{2.5}$$

Given the above definition of γ^N, this definition of υ^N shall be interpreted as follows:

$$\mathbb{E}\big[\omega^N(i,j)^2\,\big|\,\mathcal{F}_0^N,\mathcal{E}^N\big] = \begin{cases}\upsilon^N(i,j)+\gamma^N(X_i^N,X_j^N)^2 & \text{if } (i,j)\in\mathcal{E}^N,\\ 0 & \text{otherwise.}\end{cases} \tag{2.6}$$

Aa a consequence of these definitions, we have

$$N\cdot\mathbb{E}_0^N\left[\omega^N(i,j)\right] = \bar{\omega}^N(X_i^N,X_j^N), \quad \text{where } \ \bar{\omega}^N := N\cdot\kappa^N\cdot\gamma^N\,. \tag{2.7}$$

Remark that γ^N, thus $\bar{\omega}^N$, are not required to be symmetric, given that the allowed asymmetry of $\omega^N(i,j)$ has no good reason to vanish after taking conditional expectation. However, recall that the construction imposes κ^N to be symmetric.

For each individual i, the aggregated force of infection at time t acting upon i is given by

$$\overline{\mathfrak{F}}_i^N(t) = \sum_{j\in\mathcal{I}^N(t)} \omega^N(i,j)\cdot\lambda_j^N(A_j^N(t))\,. \tag{2.8}$$

For each $i\in\mathcal{S}^N(0)$, we describe the progression of the disease through the following process:

$$D_i^N(t) = \int_0^t\int_0^\infty \mathbb{1}_{\left\{D_i^N(s^-)=0\right\}}\mathbb{1}_{\left\{u\le\overline{\mathfrak{F}}_i^N(s^-)\right\}}Q_i(\mathrm{d}s,\mathrm{d}u)\,, \tag{2.9}$$

where $Q_i(\mathrm{d}s,\mathrm{d}u)$ is a standard Poisson random measure on $\mathbb{R}_+^2$ with mean measure $\mathrm{d}s\,\mathrm{d}u$, the $(Q_j)_{j\in\mathbb{N}}$ being globally independent and independent of $\mathcal{F}_0^N$, the random graph generation $(\omega^N(j,k))_{j,k\le N}$, and the random infectivity functions $(\lambda_j^N)_{j\le N}$ (in the sense that they are independent of the $\mathcal{Y}^N$ in the formulation (2.2)). Let

$$\tau_i^N = \inf\{t\ge 0;\, D_i^N(t)=1\}; \quad \mathcal{D}^N(t) := \{i\in\mathcal{S}^N(0);\, \tau_i^N\le t\}, \tag{2.10}$$

so that the infection time τ_i^N is the unique jump time of D_i^N (which takes the value ∞ in the absence of such a jump), while $\mathcal{D}^N(t)$ is the subset of initially susceptible individuals infected by the disease by time t after time 0, or equivalently those $i\in\mathcal{S}^N(0)$ such that $D_i^N(t)=1$. Note that $A_i^N(t)>0$ means $D_i^N(t^-)=1$, which is not exactly $D_i^N(t)=1$. In other words, $\mathcal{D}^N(t) = (\mathcal{I}^N(t)\cup\mathcal{R}^N(t))\setminus(\mathcal{I}^N(0)\cup\mathcal{R}^N(0))$.

Using $D_i^N(t)$ in (2.8), we also obtain the following alternative expression for $\mathfrak{F}_i^N(t)$ for each $i = 1, \dots, N$:

$$\overline{\mathfrak{F}}_i^N(t) = \sum_{k \in \mathcal{I}^N(0)} \omega^N(i,k) \lambda_k^N(A_k^N(0)+t) + \sum_{j \in \mathcal{S}^N(0)} \omega^N(i,j) \int_0^t \int_0^\infty \lambda_j^N(t-s) \mathbb{1}_{\{D_j^N(s^-)=0\}} \mathbb{1}_{\{u \le \overline{\mathfrak{F}}_j^N(s^-)\}} Q_j(\mathrm{d}s, \mathrm{d}u) \,. \tag{2.11}$$

Define the following measure-valued processes associated with the susceptible, infectious and recovered individuals:

$$\mu_t^{S,N}(\mathrm{d}x) = \sum_{i \in \mathcal{S}^N(t)} \delta_{X_i^N}(\mathrm{d}x) \,, \tag{2.12}$$

$$\mu_t^{I,N}(\mathrm{d}x, \mathrm{d}a) = \sum_{i \in \mathcal{I}^N(t)} \delta_{X_i^N}(\mathrm{d}x) \delta_{A_i^N(t)}(\mathrm{d}a) \,, \tag{2.13}$$

$$\mu_t^{R,N}(\mathrm{d}x) = \sum_{i \in \mathcal{R}^N(t)} \delta_{X_i^N}(\mathrm{d}x) \,. \tag{2.14}$$

For any $t > 0$, $\mu_t^{S,N}$ and $\mu_t^{R,N}$ are regarded as elements in the set $\mathcal{M}(\mathbb{X})$ of non-negative finite measure on $\mathbb{X}$, which is equipped with the topology of weak convergence, and similarly for $\mu_t^{I,N}$ belonging to the set $\mathcal{M}(\mathbb{X} \times \mathbb{R}_+)$.

The dynamics of these measure-valued processes can be represented using $D_i^N(t)$ and $\overline{\mathfrak{F}}_i^N(t)$ as follows, in terms of test functions $\varphi \in C_b(\mathbb{X})$ and $\psi \in C_b(\mathbb{X} \times \mathbb{R}_+)$:

- for the susceptible process:

$$\langle \mu_t^{S,N}, \varphi \rangle = \langle \mu_0^{S,N}, \varphi \rangle - \sum_{i \in \mathcal{D}^N(t)} \varphi(X_i^N) \,, \tag{2.15}$$

  with

$$\sum_{i \in \mathcal{D}^N(t)} \varphi(X_i^N) = \sum_{i \in \mathcal{S}^N(0)} \int_0^t \int_0^\infty \mathbb{1}_{\{D_i^N(s^-)=0\}} \mathbb{1}_{\{u \le \overline{\mathfrak{F}}_i^N(s^-)\}} \varphi(X_i^N) Q_i(\mathrm{d}s, \mathrm{d}u),$$

- for the infected process:

$$\langle \mu_t^{I,N}, \psi \rangle = \sum_{j \in \mathcal{I}^N(0)} \mathbb{1}_{\{\eta_j^{N,0} > t\}} \psi(X_j^N, A_j^N(0)+t) + \sum_{i \in \mathcal{D}^N(t)} \mathbb{1}_{\{\tau_i^N + \eta_i^N > t\}} \psi(X_i^N, t - \tau_i^N) \,, \tag{2.16}$$

  with

$$\sum_{i \in \mathcal{D}^N(t)} \mathbb{1}_{\{\tau_i^N + \eta_i^N > t\}} \psi(X_i^N, t-\tau_i^N) = \sum_{i \in \mathcal{S}^N(0)} \int_0^t \int_0^\infty \mathbb{1}_{\{D_i^N(s^-)=0\}} \mathbb{1}_{\{u \le \overline{\mathfrak{F}}_i^N(s^-)\}} \mathbb{1}_{\{\eta_i^N > t-s\}} \psi(X_i^N, t-s) Q_i(\mathrm{d}s, \mathrm{d}u) \,,$$

- for the recovered process:

$$\langle \mu_t^{R,N}, \varphi \rangle = \langle \mu_0^{R,N}, \varphi \rangle + \sum_{j \in \mathcal{I}^N(0)} \mathbb{1}_{\{\eta_j^{N,0} \le t\}} \varphi(X_j^N) + \sum_{i \in \mathcal{D}^N(t)} \mathbb{1}_{\{\tau_i^N + \eta_i^N \le t\}} \varphi(X_i^N) \,. \tag{2.17}$$

Note that the processes $S^N(t)$, $I^N(t)$ and $R^N(t)$ corresponding to respectively the numbers of susceptible, infected and recovered individuals at time t can be obtained from these measure-valued processes: $S^N(t) = \langle \mu_t^{S,N}, \mathbb{1}\rangle$, $I^N(t) = \langle \mu_t^{I,N}, \mathbb{1}\rangle$ and $R^N(t) = \langle \mu_t^{R,N}, \mathbb{1}\rangle$ with $\mathbb{1}(x) \equiv 1$ and $\mathbb{1}(a,x) \equiv 1$.

3. Functional law of large numbers

We consider the LLN-scaled measure-valued processes derived from respectively (2.12), (2.13) and (2.14):

$$\bar{\mu}^{S,N} = N^{-1} \cdot \mu^{S,N}, \quad \bar{\mu}^{I,N} = N^{-1} \cdot \mu^{I,N}, \quad \bar{\mu}^{R,N} = N^{-1} \cdot \mu^{R,N}\,. \tag{3.1}$$

We make the following assumptions on the initial quantities.

Assumption 3.1. *There exist finite measures $\bar{\mu}_0^S(\mathrm{d}x)$ on $\mathbb{X}$, $\bar{\mu}_0^I(\mathrm{d}x, \mathrm{d}a)$ on $\mathbb{X} \times \mathbb{R}_+$ and $\bar{\mu}_0^R(\mathrm{d}x)$ on $\mathbb{X}$ that are the weak limits in probability as $N \to \infty$ of respectively $\bar{\mu}_0^{S,N}$, $\bar{\mu}_0^{I,N}$, and $\bar{\mu}_0^{R,N}$. In addition, the following inequality holds between the measures $\bar{\mu}_0^R$ and $\bar{\mu}_0^I$:*

$$\bar{\mu}_0^R(\mathrm{d}x) \geq \int_0^\infty \frac{F_x(a)}{F_x^c(a)} \bar{\mu}_0^I(\mathrm{d}x, da). \tag{3.2}$$

We consider also the two complete distributions μ_X^N and μ_X of characteristics on $\mathbb{X}$:

$$\bar{\mu}_X^N(\mathrm{d}x) = \bar{\mu}_0^{S,N}(\mathrm{d}x) + \langle \bar{\mu}_0^{I,N}(\mathrm{d}x, .), \mathbf{1}\rangle + \bar{\mu}_0^{R,N}(\mathrm{d}x) = \frac{1}{N} \sum_{i \leq N} \delta_{X_i^N}(\mathrm{d}x), \tag{3.3}$$

$$\bar{\mu}_X(\mathrm{d}x) = \bar{\mu}_0^S(\mathrm{d}x) + \langle \bar{\mu}_0^I(\mathrm{d}x, .), \mathbf{1}\rangle + \bar{\mu}_0^R(\mathrm{d}x). \tag{3.4}$$

Here $\mathbf{1}(a) \equiv 1$. A direct consequence of Assumption 3.1 is that $\bar{\mu}_X$ is a probability distribution, and $\bar{\mu}_X^N$ converges weakly to $\bar{\mu}_X$.

Remark 3.2. *The property that relates $\bar{\mu}_0^R$ to $\bar{\mu}_0^I$ in Assumption 3.1 is justified by the following guiding idea: individuals infected with age a at time 0 should be interpreted as the outcomes of infection events at time $-a$ that produced a larger proportion of infected (with a scaling factor $[F_x^c(a)]^{-1}$), from which a proportion $F_x(a)$ got to recover.*

In addition, we make the following two assumptions to capture the behavior of the graph structure $(\omega^N(i,j))_{i,j \in [\![1,N]\!]}$ as N tends to infinity, firstly in terms of conditional expectations.

Assumption 3.3. *Recalling the relation $\bar{\omega}^N = N \cdot \kappa^N \cdot \gamma^N$, the following convergence as $N \to \infty$ holds in probability uniformly over $\mathbb{X} \times \mathbb{X}$:*

$$\bar{\omega}^N \to \bar{\omega},$$

where $\bar{\omega} : \mathbb{X} \times \mathbb{X} \to \mathbb{R}_+$ is some deterministic and bounded measurable function that is $\bar{\mu}_X^{\otimes 2}$-almost everywhere (a.e.) continuous.

Complementary to Assumptions 2.4 and 2.6, the function : $(x, a) \mapsto F_x(a)$ is $\bar{\mu}_0^I$-a.e. continuous while the function : $x \mapsto \bar{\lambda}(x, .)$ taking values in $\mathcal{D}(\mathbb{R}_+, \mathbb{R}_+)$ equipped with the uniform norm topology is $\bar{\mu}_X$-a.e. continuous.

The sequence $\bar{\omega}^N$ is uniformly upper-bounded by virtue of Assumption 3.3:

Lemma 3.4. *Under Assumption 3.3, there exists a constant $\omega^* > 0$ that is an upper-bound on $\bar{\omega}^N$ uniformly on $\mathbb{X}^2$ and for N sufficiently large.*

Without loss of generality regarding our convergence results in $N \to \infty$, we assume in the following that $\omega^* > 0$ is actually an upper-bound uniformly on $\mathbb{X}^2$ and on any N.

The next assumption provides the crucial estimate to deal with the variability of the random graph generation. Its relation to the denseness of the graph and to possibly high levels of infection rates is discussed after the statement of the main result.

Assumption 3.5. *The two following quantities $\bar{\gamma}^N$ and Υ^N converge in probability to zero as $N \to \infty$:*

$$\begin{aligned} \bar{\gamma}^N &:= \frac{1}{N^2} \sum_{i,j} \gamma^N(X_i^N, X_j^N), \\ \Upsilon^N &:= \frac{1}{N} \sum_{i,j} \mathbb{E}_0^N \left[\upsilon^N(i,j); (i,j) \in \mathcal{E}^N \right] \\ &= \frac{1}{N} \sum_{i,j} \kappa^N(X_i^N, X_j^N) \cdot \mathbb{E}_0^N \left[\upsilon^N(i,j) \,\middle|\, (i,j) \in \mathcal{E}^N \right]. \end{aligned} \tag{3.5}$$

To be clear with our notations of conditional expectations, $\mathbb{E}_0^N\big[\upsilon^N(i,j) \,\big|\, (i,j) \in \mathcal{E}^N\big]$ describes the ratio of $\mathbb{E}\big[\upsilon^N(i,j); (i,j) \in \mathcal{E}^N \,\big|\, \mathcal{F}_0^N\big]$ by $\mathbb{P}\big[(i,j) \in \mathcal{E}^N \,\big|\, \mathcal{F}_0^N\big] = \kappa^N(X_i^N, X_j^N)$.

For brevity, we denote $\mathcal{D}_1 := \mathcal{D}(\mathbb{R}_+, \mathcal{M}(\mathbb{X}))$ and $\mathcal{D}_2 := \mathcal{D}(\mathbb{R}_+, \mathcal{M}(\mathbb{X} \times \mathbb{R}_+))$.

Theorem 3.6. *Under Assumptions 2.2, 2.4, 2.6, 3.1, 3.3 and 3.5,*

$$(\bar{\mu}_\cdot^{S,N}, \bar{\mu}_\cdot^{I,N}, \bar{\mu}_\cdot^{R,N}) \to (\bar{\mu}_\cdot^S, \bar{\mu}_\cdot^I, \bar{\mu}_\cdot^R)$$

in $\mathcal{D}_1 \times \mathcal{D}_2 \times \mathcal{D}_1$ as $N \to \infty$. In the above limit, $\bar{\mu}_\cdot^S$ is the first component of the unique solution $(\bar{\mu}_\cdot^S, \overline{\mathfrak{F}}(\cdot))$ to the following set of equations,

$$\bar{\mu}_t^S(\mathrm{d}x) = \bar{\mu}_0^S(\mathrm{d}x) - \int_0^t \overline{\mathfrak{F}}(s,x) \bar{\mu}_s^S(\mathrm{d}x)\mathrm{d}s\,, \tag{3.6}$$

and

$$\begin{aligned} \overline{\mathfrak{F}}(t,x) = \int_0^\infty \!\!\int_{\mathbb{X}} \bar{\omega}(x,x') \frac{\bar{\lambda}(x', a'+t)}{F_{x'}^c(a')} \bar{\mu}_0^I(\mathrm{d}x', da') \\ + \int_0^t \!\!\int_{\mathbb{X}} \bar{\omega}(x,x') \bar{\lambda}(x', t-s) \overline{\mathfrak{F}}(s,x') \bar{\mu}_s^S(\mathrm{d}x')\mathrm{d}s\,, \end{aligned} \tag{3.7}$$

with $\bar{\lambda}$, $\bar{\omega}$ being given respectively in Assumptions 2.6 and 3.3. The uniqueness of the above system is stated among the potential candidates $(\bar{\mu}_t^S, \overline{\mathfrak{F}}(t))_{t \ge 0}$ in which $\bar{\mu}^S$ is a $\mathcal{M}(\mathbb{X})$-valued càd-làg process such that $\mu_t(\mathbb{X}) \le 1$ for any $t \ge 0$, while $\overline{\mathfrak{F}}(\cdot)$ is a càd-làg process whose values are bounded measurable functions from $\mathbb{X}$ to $\mathbb{R}_+$, with local in time upper-bounds.

Given the pair $(\bar{\mu}_\cdot^S, \overline{\mathfrak{F}}(\cdot))$, $\langle \bar{\mu}_t^I, \psi \rangle$ and $\langle \bar{\mu}_t^R, \varphi \rangle$ are given as follows, with the test functions $\psi \in C_b(\mathbb{X} \times \mathbb{R}_+)$ and $\varphi \in C_b(\mathbb{R}_+)$,

$$\begin{aligned} \langle \bar{\mu}_t^I, \psi \rangle = \int_0^\infty \!\!\int_{\mathbb{X}} \psi(x, a+t) \frac{F_x^c(a+t)}{F_x^c(a)} \bar{\mu}_0^I(\mathrm{d}x, \mathrm{d}a) \\ + \int_0^t \!\!\int_{\mathbb{X}} \psi(x, t-s) F_x^c(t-s) \overline{\mathfrak{F}}(s,x) \bar{\mu}_s^S(\mathrm{d}x)\mathrm{d}s\,, \end{aligned} \tag{3.8}$$

and

$$\begin{aligned}\langle \bar{\mu}^R_t, \varphi\rangle = \langle \bar{\mu}^R_0, \varphi\rangle + \int_0^\infty\!\!\int_{\mathbb{X}} \varphi(x)\Big(1 - \frac{F^c_x(a+t)}{F^c_x(a)}\Big)\bar{\mu}^I_0(\mathrm{d}x, \mathrm{d}a)\\ + \int_0^t\!\!\int_{\mathbb{X}} \varphi(x) F_x(t-s)\overline{\mathfrak{F}}(s,x)\bar{\mu}^S_s(\mathrm{d}x)\mathrm{d}s\,.\end{aligned} \tag{3.9}$$

Note that the unique solution to (3.6) is given by

$$\bar{\mu}^S_t(\mathrm{d}x) = \exp\Big(-\int_0^t \overline{\mathfrak{F}}(s,x)\mathrm{d}s\Big)\bar{\mu}^S_0(\mathrm{d}x)\,. \tag{3.10}$$

3.1. **Discussions of the main result.**

About the limiting description of heterogeneity.
Our conditions on the limiting kernel of interaction $\bar{\omega}$ encompasses most, if not all, of the pairwise kernel interactions. Contact matrices over a discrete space are allowed, as, e.g., in [11], as well as spatial interactions as considered in [35] with characteristics that are distributed on a compact subset of $\mathbb{R}^d$, or even combination of both settings. A possible extension of our result would be to consider individual types that evolves in time, typically when the individual locations move in space. It is another natural objective to establish the FLLN in instances where the limiting kernel $\bar{\omega}$ incorporates an additional dependency on the overall distribution $\bar{\mu}_X$ as in the general setting of [35].

About the boundedness of the kernel.
For the existence and uniqueness of the limiting system, the boundedness conditions on $\bar{\omega}$ could certainly be relaxed, with conditions that involve the integral over $\bar{\mu}_X$ (see [19] for similar results in this direction). The uniform boundedness of $\big(\bar{\omega}(x,x')\big)$ is however really helpful for the convergence of the random N-model. Further insights are required to extend the result in the context of superspreading, defined as the occurrence of highly heterogeneous transmission patterns across the individual types (X^N_i).

About the relation to dynamics on a graphon.
It is through the expression of the limiting force of infection $\overline{\mathfrak{F}}$ in terms of the limiting kernel $\bar{\omega}$ that we can relate our result to previous works about dynamics on a graphon.

Let us recall the previously introduced example where $\bar{\kappa}^N = \bar{\kappa}$ for any N and $\bar{\kappa}(x,x') = x{\cdot}x'$ while the (X^N_i) are uniformly sampled in $[0,1]$. Let us assume in addition that the contact rate function γ^N is fixed and does not depend on the types of the individuals i and j in contact while $\bar{\omega}$ is non-zero. In this case, $\Upsilon^N \equiv 0$.

If we denote by $\bar{\gamma}^N > 0$ the value taken by the constant contact rate, then Assumption 3.3 holds if and only if $\bar{\gamma}^N$ is equivalent to $\bar{\gamma}/N$ for some value $\bar{\gamma} > 0$, in which case $\bar{\omega}(x,x') = \bar{\gamma}\cdot\bar{\kappa}(x,x') = \bar{\gamma}\cdot x\cdot x'$. Assumption 3.5 is then automatically satisfied. The proposed kernel structure displays a convenient property of separation between three contributions to the value: the role (i) of the susceptible type, (ii) of the infector type and (iii) of the contact rate. This property is then inherited by the force of infection in that $\overline{\mathfrak{F}}(t,x) = \bar{\gamma}\cdot x\cdot\check{\mathfrak{F}}(t)$ where $\check{\mathfrak{F}}(t)$ describes the aggregated contribution of the potential infectors at time t:

$$\check{\mathfrak{F}}(t) = \int_0^\infty\!\!\int_{\mathbb{X}} x'\cdot\frac{\bar{\lambda}(x', a'+t)}{F^c_{x'}(a')}\bar{\mu}^I_0(\mathrm{d}x', da') + \int_0^t\!\!\int_{\mathbb{X}} x'\cdot\bar{\lambda}(x', t-s)\overline{\mathfrak{F}}(s,x')\bar{\mu}^S_s(\mathrm{d}x')\mathrm{d}s\,.$$

As considered in [36], such a setting for the dynamics on a graphon can be generalized by assuming a scaling factor $\varepsilon^N > 0$ depending on N, for instance $\varepsilon^N = N^{-\alpha}$ with $\alpha \in (0,1)$, in that $\bar{\kappa}^N = \varepsilon^N \cdot \bar{\kappa}$. Provided we again assume the contact rate to be fixed constant at a value $\bar{\gamma}^N > 0$, we can follow the same reasoning except that $\bar{\gamma}^N$ is then equivalent to $\bar{\gamma}/(N\varepsilon^N)$ instead of $\bar{\gamma}/N$ and that $N\varepsilon^N \to \infty$ is required (and sufficient) for Assumption 3.5 to hold.

Note also that $N\varepsilon^N \to \infty$ typically means that the node degrees are asked to go to infinity, without condition on the relation to the number N of nodes. We refer to [20] for more details on the relation to the denseness level of the graph (in terms notably of the number of edges or the node degrees). The additional condition proposed in [36] corresponding to the convergence to infinity of $N\varepsilon^N / \log(N)$ appears related to their technique of proof rather than to the convergence of the epidemic process itself. The case where $\varepsilon^N = N^{-1}$ leads to limiting equations of a different nature, as hinted with exploratory simulations in [20] and explicited in the result of [10] for instances of stochastic block models.

In practice, different contact patterns could induce different scalings of $\gamma^N(X_i^N, X_j^N)$ in relation to the two types X_i^N and X_j^N of individuals in contact. The interest in (3.5) is then to quantify the convergence in terms of an aggregate quantity that is more directly accessible than a global scaling factor to be inferred.

In the setting of [20] still, the variability in the contact rate is purely determined by the individual types. We see with Theorem 3.6, which could be adapted to the simpler epidemiological behavior considered in [20], that the criteria can be efficiently extended to random contact rates. The same average over pairs should then be evaluated in terms of conditional expectations together with a similarly averaged criterion on the variance, as stated in Assumption 3.5.

About the role of a negligible subset of edges.
The two statistics introduced in Assumption 3.5 are averages over the total set of edges. Therefore, the introduction of a high rate of contagion on a negligible fraction of edges would be insufficient to compromise the convergence result deduced from the other edges. In particular, the introduction of a high rate of contagion from or to a negligible fraction of individuals would lead to the same conclusion.

About heterogeneous denseness of the graph.
In the setting of Assumption 3.3, we ask for the convergence of the product $N\gamma^N\kappa^N$ without asking for separate convergence properties of γ^N on the one hand and of κ^N on the other hand. This choice was made to emphasize that the scaling in N between the contact probability and the contact rate could itself be mediated by various kinds of interactions reflected through the individual types X_i^N and X_j^N.

For example, we may consider the (X_i^N) as spatial coordinates, still uniformly sampled in $[0,1]$ yet with the norm-distance $d(x,y) = |x-y|$ on the circle so that all the locations are equivalent (i.e., on the one-dimensional torus with 0 and 1 identified). Let us then distinguish in the expressions of κ^N and γ^N a local interaction pattern (specified by the index $\mathcal{L}$) with radius $\delta \in (0,1/2)$ and a global interaction pattern (with index $\mathcal{G}$) over the whole domain:

$$\kappa^N(x,y) = \begin{cases} \kappa^{N,\mathcal{L}} \\ \kappa^{N,\mathcal{G}} \end{cases}, \quad \gamma^N(x,y) = \begin{cases} \gamma^{N,\mathcal{L}} & \text{if } |x-y| \le \delta, \\ \gamma^{N,\mathcal{G}} & \text{otherwise.} \end{cases}$$

Typically, we could expect $\kappa^{N,\mathcal{L}} \gg \kappa^{N,\mathcal{G}}$ while $\gamma^{N,\mathcal{G}} \gg \gamma^{N,\mathcal{L}}$, that is many small local contacts as compared to rare yet strongly connected global contacts. Such a framework on a network is captured by our model.

Both types of contacts remain in the limit provided both $\kappa^{N,\mathcal{L}} \cdot \gamma^{N,\mathcal{L}}$ and $\kappa^{N,\mathcal{G}} \cdot \gamma^{N,\mathcal{G}}$ scale as N^{-1}. Actually we see that the contribution of the local contacts then outcompetes the one of global contacts under the following condition: $(\kappa^{N,\mathcal{L}} \cdot \gamma^{N,\mathcal{L}})/(\kappa^{N,\mathcal{G}} \cdot \gamma^{N,\mathcal{G}}) \geq (2\delta)/(1-2\delta)$.

Such a variability could be introduced more generally to capture different kinds of interactions, for instance with more denseness but less frequent contacts for the infections during travels than at the workplaces.

About the randomness in the contact rate.
To fix ideas regarding the variance term in Assumption 3.5, let us consider the case where the heterogeneity in individual contacts is purely neutral, so independent of the (X_i^N). Let us assume the existence of three positive parameters $\bar{\kappa}^N$, $\bar{\gamma}^N$ and $\bar{\sigma}^N \in (0, \bar{\gamma}^N)$ such that the two following conditions hold in addition to the above-described construction:

(i) the graph of active contacts is an Erdös-Rényi graph with parameter $\bar{\kappa}^N$, that is $\mathbb{P}_0^N\big(i \overset{N}{\sim} j\big) = \bar{\kappa}^N$ for any i, j.
(ii) on the event $\{(i,j) \in \mathcal{E}^N\}$ and conditionally on $\mathcal{F}_0^N$, $\omega^N(i,j)$ is distributed as a uniform random variable between $\bar{\gamma}^N - \bar{\sigma}^N$ and $\bar{\gamma}^N + \bar{\sigma}^N$.

Then (2.3) and (2.4) are satisfied with

$$\kappa^N(x,y) \equiv \bar{\kappa}^N, \quad \gamma^N(x,y) \equiv \bar{\gamma}^N, \quad \upsilon^N(i,j) \equiv \frac{(\bar{\sigma}^N)^2}{3}.$$

Assumption 3.3 then translates into the convergence of the product $N\,\bar{\kappa}^N\,\bar{\gamma}^N$ to some limiting value $\bar{\omega}$. The notation $\bar{\gamma}^N$ is coherent with the formula given in Assumption 3.5 while $\Upsilon^N = N\,\bar{\upsilon}^N\,\bar{\kappa}^N = N \cdot (\bar{\sigma}^N)^2 \cdot \bar{\kappa}^N/3$. Let us assume the system to be non-degenerate in that $\bar{\omega} > 0$ and $\bar{\gamma}^N > 0$ for any N.

Then, $\Upsilon^N \sim \bar{\omega}\ \cdot (\bar{\sigma}^N)^2/(3\gamma^N)$ converges to zero as required in Assumption 3.5 if and only if $(\bar{\sigma}^N)^2 \ll \gamma_N$. Since $\bar{\sigma}^N < \gamma_N$ in this case (to keep ω^N non-negative), both properties hold true if and only if γ_N tends to zero, which is exactly the first condition in Assumption 3.5. So we do not have any additional restriction on $\bar{\sigma}_N$ in this model.

For another example with possibly high levels of heterogeneity, let us replace the uniform distribution in (ii) by a gamma distribution, whose two parameters we fix by taking $\bar{\gamma}_N$ as the mean and $\bar{\sigma}_N \cdot \bar{\gamma}_N$ as the variance. Then, the convergence of Υ^N to zero corresponds exactly to the convergence of $\bar{\sigma}_N$ to zero.

We remark that $\bar{\sigma}_N$ is classically described as the scale parameter. If it would not converge to zero, then the convergence to 0 of the mean $\bar{\gamma}^N$ would entail the convergence to 0 of the corresponding shape parameter $\bar{\alpha}^N = \bar{\gamma}^N/\bar{\sigma}_N$.

Generally with $\bar{\upsilon}^N$ the variance of the corresponding distribution in (ii), Υ^N converges to zero if and only if the standard deviation $\sqrt{\bar{\upsilon}^N}$ is negligible against the root of the average rate $\sqrt{\bar{\gamma}^N}$. Given that $\bar{\gamma}^N$ itself is expected to tend to 0, we stress that our result covers fluctuation levels in the rate of transmission that are quite large in comparison to the expected value.

4. Properties of the limiting system of equations

4.1. Existence and uniqueness of solution to the limiting equations.

Proposition 4.1. *Under Assumptions 2.2 and 3.3, the set of equations (3.6)–(3.7) has a unique solution* $(\bar{\mu}^S_\cdot, \overline{\mathfrak{F}}_\cdot)$.

Proof. We first prove the uniqueness. Before proceeding, we establish some useful bounds. Suppose that $(\bar{\mu}^S_\cdot, \overline{\mathfrak{F}}_\cdot)$ is a solution. By Assumptions 2.2 and 3.3, we have $\bar{\lambda}(x,t) \le \lambda^*$ and $\bar{\omega}(x,x') \le \omega^*$ for some $\omega^* > 0$. Recalling (3.7) and Assumption 3.1, we derive

$$\begin{aligned}\overline{\mathfrak{F}}(t,x) &\le \lambda^* \int_{\mathbb{X}} \bar{\omega}(x,x') \int_0^\infty \left[1 + \frac{F_{x'}(a')}{F^c_{x'}(a')}\right] \bar{\mu}^I_0(\mathrm{d}x', \mathrm{d}a') \\ &\quad + \lambda^* \int_{\mathbb{X}} \bar{\omega}(x,x') \int_0^t \overline{\mathfrak{F}}(s,x') \exp\left(- \int_0^s \overline{\mathfrak{F}}(r,x')\mathrm{d}r\right) \mathrm{d}s \; \bar{\mu}^S_0(\mathrm{d}x') \\ &\le \lambda^* \int_{\mathbb{X}} \bar{\omega}(x,x') \cdot \left[\langle \bar{\mu}^I_0(\mathrm{d}x', .), \mathbb{1}\rangle + \bar{\mu}^R_0(\mathrm{d}x') + \bar{\mu}^S_0(\mathrm{d}x')\right] \\ &\le \lambda^* \omega^* < \infty,\end{aligned} \tag{4.1}$$

where we exploited that 1 is a natural upper-bound of the integral over s in the second line and that $\bar{\mu}_X$ in (3.4) is a probability measure.

On the other hand by (3.6),

$$\bar{\mu}^S_s(\mathrm{d}x') \le \bar{\mu}^S_0(\mathrm{d}x') \le \bar{\mu}_X(\mathrm{d}x). \tag{4.2}$$

Suppose now that there are two solutions $(\bar{\mu}^{S,\ell}_\cdot, \overline{\mathfrak{F}}^\ell_\cdot)$, $\ell = 1, 2$. By (3.6) and (3.7), for $\varphi \in C_b(\mathbb{R}_+)$, we obtain

$$\langle \bar{\mu}^{S,1}_t - \bar{\mu}^{S,2}_t, \varphi\rangle = - \int_0^t \langle \bar{\mu}^{S,1}_s - \bar{\mu}^{S,2}_s, \overline{\mathfrak{F}}^1(s)\varphi\rangle \mathrm{d}s - \int_0^t \langle \bar{\mu}^{S,2}_s, (\overline{\mathfrak{F}}^1(s) - \overline{\mathfrak{F}}^2(s))\varphi\rangle \mathrm{d}s\,, \tag{4.3}$$

and

$$\begin{aligned}\overline{\mathfrak{F}}^1(t,x) - \overline{\mathfrak{F}}^2(t,x) &= \int_0^t\!\!\int_{\mathbb{X}} \bar{\omega}(x,x')\bar{\lambda}(x', t-s)\overline{\mathfrak{F}}^1(s,x')(\bar{\mu}^{S,1}_s(\mathrm{d}x') - \bar{\mu}^{S,2}_s(\mathrm{d}x'))\mathrm{d}s \\ &\quad + \int_0^t\!\!\int_{\mathbb{X}} \bar{\omega}(x,x')\bar{\lambda}(x', t-s)(\overline{\mathfrak{F}}^1(s,x') - \overline{\mathfrak{F}}^2(s,x'))\bar{\mu}^{S,2}_s(\mathrm{d}x')\mathrm{d}s\,.\end{aligned} \tag{4.4}$$

Let

$$D_t = (\lambda^*\omega^*) \cdot \left\|\bar{\mu}^{S,1}_t - \bar{\mu}^{S,2}_t\right\|_{TV} + \|\overline{\mathfrak{F}}^1(t,.) - \overline{\mathfrak{F}}^2(t,.)\|_\infty.$$

By (4.3), (4.4), (4.1) and (4.2):

$$D_t \le C \int_0^t D_s \mathrm{d}s, \tag{4.5}$$

where $C = 2\lambda^*\omega^*$. Thanks to Gronwall's inequality, the proof of the uniqueness is concluded.

Finally, the existence can be proved by Picard iteration (which will invoke similar estimates as above). □

4.2. **Regularity of $\overline{\mathfrak{F}}(\cdot)$.**

Lemma 4.2. *The function : $x \mapsto \overline{\mathfrak{F}}(.,x)$ with values in the set of bounded measurable functions from $\mathbb{R}_+$ to itself equipped with the uniform norm topology is $\bar{\mu}_X$-almost everywhere continuous.*

Proof. Recalling (3.7), with the uniform upper-bound of $\bar{\lambda}$ by λ^*, we first deduce the following inequality for any $t \geq 0$ and any $x_1, x_2 \in \mathbb{X}$:

$$\Big|\overline{\mathfrak{F}}(t,x_1) - \overline{\mathfrak{F}}(t,x_2)\Big| \leq \lambda^* \int_{\mathbb{X}} |\bar{\omega}(x_1,x') - \bar{\omega}(x_2,x')| \Big\{ \Big(1 + \frac{F_{x'}(a')}{F^c_{x'}(a')}\Big) \bar{\mu}^I_0(\mathrm{d}x', \mathrm{d}a') + \int_0^t \overline{\mathfrak{F}}(s,x') \exp\Big(-\int_0^s \overline{\mathfrak{F}}(r,x')\mathrm{d}r\Big) \mathrm{d}s\, \bar{\mu}^S_0(\mathrm{d}x')\Big\}.$$

With similar arguments as in the proof of Proposition 4.1 to deduce (4.1), notably with Assumption 3.1, we obtain

$$\|\overline{\mathfrak{F}}(.,x_1) - \overline{\mathfrak{F}}(.,x_2)\|_\infty \leq \lambda^* \int_{\mathbb{X}} |\bar{\omega}(x_1,x') - \bar{\omega}(x_2,x')|\, \bar{\mu}_X(\mathrm{d}x'). \tag{4.6}$$

As stated in Appendix A in Proposition A.2 and proved just afterwards, the function $x \mapsto \bar{\omega}(x,.)$ is $\bar{\mu}_X$ a.e. continuous with the $L^1(\bar{\mu}_X)$ distance, provided that $\bar{\omega}$ is $\bar{\mu}_X^{\otimes 2}$-a.e. continuous as in Assumption 3.3. This concludes the proof of Lemma 4.2. □

4.3. **Alternative representation of the limit as PDEs.**

Provided that the hazard rate function corresponding to the durations before remission is regular enough, one can describe the solution $\bar{\mu}^I$ of the limiting system in Theorem 3.6 in terms of a PDE as stated in the next proposition.

Proposition 4.3. *Assume that F_x is absolutely continuous with density f_x for each $x \in \mathbb{X}$ and that $F^c_x(a) > 0$ for any $a \in \mathbb{R}_+$. Assume that $h_x(a) = f_x(a)/F^c_x(a)$, the hazard rate function, is continuous and bounded uniformly in both $x \in \mathbb{X}$ and $a \in \mathbb{R}_+$. Assume moreover that $\bar{\mu}^I_0(\mathbb{X} \times \{0\}) = 0$. Then, $\bar{\mu}^I_t(\mathrm{d}x, \mathrm{d}a)$ in (3.8) is the unique solution to the following equation, defined for any $\psi \in C^{0,1}(\mathbb{X} \times \mathbb{R}_+)$ and $t > 0$,*

$$\frac{d}{dt}\langle \bar{\mu}^I_t, \psi\rangle = \langle \bar{\mu}^I_t, \partial_a \psi - h\psi\rangle + \int_{\mathbb{X}} \psi(x,0)\overline{\mathfrak{F}}(t,x)\bar{\mu}^S_t(\mathrm{d}x)\,. \tag{4.7}$$

Hence, $\bar{\mu}^I_t(\mathrm{d}x, \mathrm{d}a)$ is the unique solution to the following PDE:

$$\langle \partial_t \bar{\mu}^I_t, \psi\rangle + \langle \partial_a \bar{\mu}^I_t, \psi\rangle = -\langle \bar{\mu}^I_t, h\psi\rangle \tag{4.8}$$

with the initial condition $\bar{\mu}^I_0$ given in Assumption 3.1 and the boundary condition at $a = 0$: $\tilde{\mu}^I_t(\mathrm{d}x, 0) = \overline{\mathfrak{F}}(t,x)\bar{\mu}^S_t(\mathrm{d}x)$, where $\tilde{\mu}^I_t(\mathrm{d}x, a)$ is the "density" of $\bar{\mu}^I_t(\mathrm{d}x, \mathrm{d}a)$, that is, $\bar{\mu}^I_t(\mathrm{d}x, \mathrm{d}a) = \tilde{\mu}^I_t(\mathrm{d}x, a)\mathrm{d}a$ and $a \mapsto \tilde{\mu}^I_t(\mathrm{d}x, a)$ is continuous at 0^+.

Recalling (3.10), remark that we can also write the last term in (4.7) as

$$\int_{\mathbb{X}} \psi(x,0)\overline{\mathfrak{F}}(t,x) \exp\Big(-\int_0^t \overline{\mathfrak{F}}(s,x)\mathrm{d}s\Big)\bar{\mu}^S_0(\mathrm{d}x)\,.$$

which only involves the initial $\bar{\mu}^S_0(\mathrm{d}x)$ and $\overline{\mathfrak{F}}(t,x)$.

Proof. By taking derivative with respect to t in the expression of $\langle \bar{\mu}^I_t, \psi\rangle$ in (3.8), we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle \bar{\mu}^I_t, \psi\rangle &= \int_0^\infty\!\!\int_{\mathbb{X}} \partial_a\psi(x, a+t)\frac{F^c_x(a+t)}{F^c_x(a)}\bar{\mu}^I_0(\mathrm{d}x, \mathrm{d}a)\\ &\quad - \int_0^\infty\!\!\int_{\mathbb{X}} \psi(x, a+t)\frac{f_x(a+t)}{F^c_x(a)}\bar{\mu}^I_0(\mathrm{d}x, \mathrm{d}a)\\ &\quad + \int_{\mathbb{X}} \psi(x,0)\overline{\mathfrak{F}}(t,x)\bar{\mu}^S_t(\mathrm{d}x)\\ &\quad - \int_0^t\!\!\int_{\mathbb{X}} \psi(x, t-s) f_x(t-s)\overline{\mathfrak{F}}(s,x)\bar{\mu}^S_s(\mathrm{d}x)\mathrm{d}s\\ &\quad + \int_0^t\!\!\int_{\mathbb{X}} \partial_a\psi(x, t-s) F^c_x(t-s)\overline{\mathfrak{F}}(s,x)\bar{\mu}^S_s(\mathrm{d}x)\mathrm{d}s\,.\end{aligned}$$

Next, from (3.8), we observe that

$$\begin{aligned}\langle \bar{\mu}^I_t, \partial_a\psi\rangle &= \int_0^\infty\!\!\int_{\mathbb{X}} \partial_a\psi(x, a+t)\frac{F^c_x(a+t)}{F^c_x(a)}\bar{\mu}^I_0(\mathrm{d}x, \mathrm{d}a)\\ &\quad + \int_0^t\!\!\int_{\mathbb{X}} \partial_a\psi(x, t-s) F^c_x(t-s)\overline{\mathfrak{F}}(s,x)\bar{\mu}^S_s(\mathrm{d}x)\mathrm{d}s\,,\end{aligned}$$

and

$$\begin{aligned}\langle \bar{\mu}^I_t, h\psi\rangle &= \int_0^\infty\!\!\int_{\mathbb{X}} \psi(x, a+t)\frac{f_x(a+t)}{F^c_x(a)}\bar{\mu}^I_0(\mathrm{d}x, \mathrm{d}a)\\ &\quad + \int_0^t\!\!\int_{\mathbb{X}} \psi(x, t-s) f_x(t-s)\overline{\mathfrak{F}}(s,x)\bar{\mu}^S_s(\mathrm{d}x)\mathrm{d}s\,.\end{aligned}$$

Hence, the last three identities lead to the expression in (4.7).

Next, (3.8) entails the following formula

$$\begin{aligned}\langle \bar{\mu}^I_t, \psi\rangle &= \int_t^\infty \int_{\mathbb{X}} \psi(x,a)\frac{F^c_x(a)}{F^c_x(a-t)}\bar{\mu}^I_0(\mathrm{d}x, \mathrm{d}a-t)\\ &\quad + \int_0^t \int_{\mathbb{X}} \psi(x,a)F^c_x(a)\overline{\mathfrak{F}}(t-a,x)\bar{\mu}^S_{t-a}(\mathrm{d}x)\mathrm{d}a\,.\end{aligned} \tag{4.9}$$

We see from this formula that the restriction of the measure $\bar{\mu}^I_t$ to the set $\mathbb{X}\times[0,t)$ is absolutely continuous w.r.t. the measure $\bar{\mu}^S_{t-a}(\mathrm{d}x)\mathrm{d}a$, hence the existence of the "density" $\tilde{\mu}^I_s(\mathrm{d}x, a)$, whose value at $a = 0$ is specified above by the boundary condition.

Now integrating (4.7) over the interval $[0,t]$, we obtain

$$\langle \bar{\mu}^I_t, \psi\rangle = \langle \bar{\mu}^I_0, \psi\rangle + \int_0^t \langle \bar{\mu}^I_s, \partial_a\psi - h\psi\rangle ds + \int_0^t \int_{\mathbb{X}} \psi(x,0)\overline{\mathfrak{F}}(s,x)\bar{\mu}^S_s(\mathrm{d}x)\mathrm{d}s\,. \tag{4.10}$$

We now choose in (4.10) $\psi_n(x,a) = \varsigma(x)(1-na)^+$ for some $\varsigma \in C^1(\mathbb{R}_+)$. Since $\bar{\mu}^I_0(\mathbb{X}\times\{0\}) = 0$, we deduce from (4.9) that $\bar{\mu}^I_t(\mathbb{X}\times\{0\}) = 0$ holds for any $t \geq 0$. We have

$$\langle \bar{\mu}^I_t, \psi_n\rangle \to 0 \quad \text{and} \quad \int_0^t \langle \bar{\mu}^I_s, h\psi_n\rangle \mathrm{d}s \to 0, \quad \text{as } n \to \infty.$$

Also $\psi_n(x,0)=\varsigma(x)$, while

$$\int_0^t \langle \bar{\mu}^I_s, \partial_a \psi_n\rangle \mathrm{d}s = -n\int_0^{1/n}\int_{\mathbb{X}} \varsigma(x)\bar{\mu}^I_s(\mathrm{d}x,\mathrm{d}a)\,.$$

We deduce from the above computations that, as $n\to\infty$,

$$n\int_0^t \mathrm{d}s \int_0^{1/n} \bar{\mu}^I_s(\mathrm{d}x,\mathrm{d}a) \Rightarrow \int_0^t \overline{\mathfrak{F}}(s,x)\bar{\mu}^S_s(\mathrm{d}x)\mathrm{d}s,$$

in the sense of weak convergence of measures on $\mathbb{X}$. If we denote by $\tilde{\mu}^I_t(\mathrm{d}x,0)$ the limit as $n\to\infty$ of $n\int_0^{1/n}\bar{\mu}^I_s(\mathrm{d}x,\mathrm{d}a)$, we have

$$\int_0^t \tilde{\mu}^I_s(\mathrm{d}x,0)\mathrm{d}s = \int_0^t \overline{\mathfrak{F}}(s,x)\bar{\mu}^S_s(\mathrm{d}x)\mathrm{d}s, \quad \text{hence also}$$
$$\tilde{\mu}^I_s(\mathrm{d}x,0) = \overline{\mathfrak{F}}(s,x)\bar{\mu}^S_s(\mathrm{d}x)\,.$$

From this, by the integration by parts formula, we obtain

$$\int_0^t \langle \bar{\mu}^I_s, \partial_a\psi\rangle\mathrm{d}s + \int_0^t\int_{\mathbb{X}} \psi(x,0)\overline{\mathfrak{F}}(s,x)\bar{\mu}^S_s(\mathrm{d}x)\mathrm{d}s = -\int_0^t \langle \partial_a\bar{\mu}^I_s, \psi\rangle \mathrm{d}s. \tag{4.11}$$

Hence, from (4.7) and (4.11), we obtain the PDE model in (4.8) with the boundary condition at $a=0$: $\tilde{\mu}^I_s(\mathrm{d}x,0)=\overline{\mathfrak{F}}(s,x)\bar{\mu}^S_s(\mathrm{d}x)$.

Reciprocally, let us consider in addition to $\bar{\mu}^I$ any arbitrary solution $(\widetilde{\mu}^I_t)$ to the PDE in (4.7) that satisfies the initial condition $\widetilde{\mu}^I_0=\bar{\mu}^I_0$, and define $\Delta\mu^I(\mathrm{d}x)=\widetilde{\mu}^I(\mathrm{d}x)-\bar{\mu}^I(\mathrm{d}x)$. For any $t>0$, we define as follows the function $\Psi_t\in C^{0,1}(\mathbb{X}\times\mathbb{R}_+)$ in terms of $\psi_0\in C^{0,1}(\mathbb{X}\times\mathbb{R}_+)$ and $\varphi_0\in C^{0,1}(\mathbb{X}\times[0,t])$:

$$\Psi_t(x,a) = \begin{cases} \psi_0(x,a-t)\cdot\exp\left[\int_{a-t}^a h(x,a')\mathrm{d}a'\right] & \text{for any } x\in\mathbb{X}, a\in(t,\infty),\\ \varphi_0(x,t-a)\cdot\exp\left[\int_0^a h(x,a')\mathrm{d}a'\right] & \text{for any } x\in\mathbb{X}, a\in[0,t]. \end{cases} \tag{4.12}$$

The interest of this definition lies in the relation between the time-derivatives in t and in a, that makes the process $\langle\Delta\mu^I_t,\Psi_t\rangle$ stay constant as stated next in (4.15).

Given that h is a bounded continuous function, $(\Psi_t(x,a))_{t,x,a}\in\mathcal{C}^{1,0,1}(\mathbb{R}_+\times\mathbb{X}\times\mathbb{R}_+)$. Let us compute the relevant partial derivatives for our concern, first for any $x\in\mathbb{X}$ and any $a\in(t,\infty)$,

$$\partial_t\Psi_t(x,a) = -\partial_a\psi_0(x,a-t)\exp\Big[\int_{a-t}^a h(x,a')\mathrm{d}a'\Big] + \Psi_t(x,a)\cdot h(x,a-t), \tag{4.13}$$
$$\partial_a\Psi_t(x,a) = \partial_a\psi_0(x,a-t)\;\exp\Big[\int_{a-t}^a h(x,a')\mathrm{d}a'\Big] + \Psi_t(x,a)\cdot\big[h(x,a)-h(x,a-t)\big].$$

On the other hand, for any $x\in\mathbb{X}$ and any $a\in[0,t]$,

$$\begin{aligned}\partial_t\Psi_t(x,a) &= \partial_a\varphi_0(x,t-a)\exp\Big[\int_0^a h(x,a')\mathrm{d}a'\Big],\\ \partial_a\Psi_t(x,a) &= -\partial_a\psi_0(x,t-a)\exp\Big[\int_0^a h(x,a')\mathrm{d}a'\Big] + \Psi_t(x,a)\cdot h(x,a).\end{aligned} \tag{4.14}$$

Thanks to (4.13) and (4.14), we obtain

$$\partial_t\Psi_t + \partial_a\Psi_t - h\cdot\Psi_t \equiv 0.$$

Since $\widetilde{\mu}^I, \bar{\mu}^I$ are solutions to the PDE in (4.8) and satisfy the same initial condition $\widetilde{\mu}^I_0 = \bar{\mu}^I_0$, the above identity implies the following one:

$$\langle \Delta\mu^I_t, \Psi_t \rangle = \langle \Delta\mu^I_0, \psi_0 \rangle = 0, \tag{4.15}$$

which, for any $t > 0$, holds for any $\psi_0, \varphi_0 \in \mathcal{C}^{0,1}(\mathbb{X} \times \mathbb{R}_+)$.

We next verify that for any $\psi \in \mathcal{C}^{0,1}(\mathbb{X} \times \mathbb{R}_+)$, we can choose ψ_0 and φ_0 such that Ψ given by (4.12) satisfies $\Psi_T(x,a) = \psi(x,a)$. Since $h(x,a) = -\partial_a \log F^c_x(a)$, we know that $\exp\left[-\int_0^a h(x,a')\mathrm{d}a'\right] = F^c_x(a) > 0$. Let $T > 0$ and $\psi \in \mathcal{C}^{0,1}(\mathbb{X} \times \mathbb{R}_+)$. For ψ to agree with Ψ_T, we define for any $x \in \mathbb{X}$, any $a \in \mathbb{R}_+$ and any $r \in [0,T]$:

$$\psi_0(x,a) = \psi(x,a+T) \cdot \frac{F^c_x(a+T)}{F^c_x(a)}, \qquad \varphi_0(x,r) = \frac{\psi(x,T-r)}{F^c_x(T-r)}.$$

We then check, for any $a > T$, that $\psi_0(x,a-T) \cdot \exp\left[\int_{a-T}^a h(x,a')\mathrm{d}a'\right]$ agrees with $\psi(x,a)$ and similarly for any $a \in [0,T]$ with $\varphi_0(x,T-a) \cdot \exp\left[\int_0^a h(x,a')\mathrm{d}a'\right]$ instead. Thanks to (4.15), with the specific choice of $t = T$, we deduce that $\langle \widetilde{\mu}^I_T, \psi \rangle = \langle \bar{\mu}^I_T, \psi \rangle$ holds for any $\psi \in \mathcal{C}^{0,1}(\mathbb{X} \times \mathbb{R}_+)$. Since this identity is valid for any T provided h is bounded continuous, the uniqueness of the solution to the PDE in (4.8) is deduced.

Finally, if $(\widetilde{\mu}^I_t)$ is instead assumed to be any solution to the PDE in (4.8), with the boundary condition at $a = 0$ as specified in Proposition 4.3. Then, the integration by parts formula in (4.11) holds with $\widetilde{\mu}^I$ instead of $\bar{\mu}^I$, and so (4.10) similarly for any $t > 0$. Therefore, $\widetilde{\mu}^I$ is actually solution to the PDE in (4.8). So it coincides with $\bar{\mu}^I$ by the preceding uniqueness result. □

5. Proof of Theorem 3.6, our main result

5.1. Convergence of $(\bar{\mu}^{S,N}_{\cdot}, \overline{\mathfrak{F}}^N(\cdot))$.

We start by constructing an auxiliary model using the limit $\overline{\mathfrak{F}}$. Recall $D^N_i(t)$ in (2.9). Define

$$\widetilde{D}^N_i(t) = \int_0^t \int_0^\infty \mathbb{1}_{\left\{\widetilde{D}^N_i(s^-)=0\right\}} \mathbb{1}_{\left\{u \le \overline{\mathfrak{F}}(s,X^N_i)\right\}} Q_i(\mathrm{d}s,\mathrm{d}u)\,, \tag{5.1}$$

and

$$\langle \widetilde{\mu}^{S,N}_t, \varphi \rangle = \frac{1}{N} \sum_{i \in \mathcal{S}^N(0)} \mathbb{1}_{\left\{\widetilde{D}^N_i(t^-)=0\right\}} \varphi(X^N_i)\,. \tag{5.2}$$

Considering this approximation considerably helps to justify the proximity with the limiting measure $\bar{\mu}^S$, as we can see from the next Lemma 5.1.

Lemma 5.1. *As $N \to \infty$, the following convergence holds in probability for any t and any bounded continuous function φ from $\mathbb{X}$ to $\mathbb{R}$:*

$$\langle \widetilde{\mu}^{S,N}_t, \varphi \rangle \to \langle \bar{\mu}^S_t, \varphi \rangle,$$

where the limit $\bar{\mu}^S$ is the first term of the unique solution of (3.6)-(3.7).

Proof. We start with the formula (5.2). Noting that

$$\mathbb{P}^N_0(\widetilde{D}^N_i(t) = 0) = \exp\left(-\int_0^t \overline{\mathfrak{F}}(s,X^N_i)\mathrm{d}s\right),$$

we deduce

$$\mathbb{E}_0^N\Big[\langle\widetilde{\mu}_t^{S,N},\varphi\rangle\Big]=\langle\bar{\mu}_0^{S,N},\varphi\cdot\exp\big(-\int_0^t\overline{\mathfrak{F}}(s,.)\mathrm{d}s\big)\rangle. \tag{5.3}$$

Recall from Assumption 3.1 that $\bar{\mu}_0^{S,N}$ converges in probability to $\bar{\mu}_0^S\le\bar{\mu}_X$, and from Lemma 4.2 that the (deterministic) function $\varphi\cdot\exp\big(-\int_0^t\overline{\mathfrak{F}}(s,.)\mathrm{d}s\big)$ is $\bar{\mu}_X$-a.e. continuous and bounded. Thanks to the Portmanteau theorem, we thus deduce the convergence in probability of the expectation in (5.3) to

$$\langle\bar{\mu}_0^S,\varphi\cdot\exp\big(-\int_0^t\overline{\mathfrak{F}}(s,.)\mathrm{d}s\big)\rangle=\langle\bar{\mu}_t^S,\varphi\rangle. \tag{5.4}$$

On the other hand, we control the fluctuations through the variance, by exploiting the independence property of $\widetilde{D}_i^N$ between individuals i. Similarly as $\mathbb{E}_0^N$ denotes the expectation conditional on $\mathcal{F}_0^N$, Var_0^N denotes the variance conditional on $\mathcal{F}_0^N$, in the sense that $\mathrm{Var}_0^N(Z)=\mathbb{E}_0^N[Z^2]-\mathbb{E}_0^N[Z]^2$ for any random variable Z. We have

$$\mathrm{Var}_0^N(\langle\widetilde{\mu}_t^{S,N},\varphi\rangle)=\frac{1}{N^2}\sum_{i\in\mathcal{S}^N(0)}\varphi(X_i^N)^2\cdot\mathrm{Var}_0^N\Big(\mathbb{1}_{\{\widetilde{D}_i^N(t)=0\}}\Big)\le\frac{\|\varphi\|_\infty^2}{N}. \tag{5.5}$$

Let $\varepsilon>0$. By choosing N larger than $\varepsilon^{-3}\cdot\|\varphi\|_\infty^2$, we deduce as a consequence of Chebyshev's inequality:

$$\mathbb{P}\Big(\big|\langle\widetilde{\mu}_t^{S,N},\varphi\rangle-\mathbb{E}_0^N\big[\langle\widetilde{\mu}_t^{S,N},\varphi\rangle\big]\big|\ge\varepsilon\Big)\le\varepsilon.$$

So this sequence of centered random variables converges in probability to 0. With the convergence in probability of the conditional expectation stated in (5.4), this concludes the proof of Lemma 5.1. □

In order to relate $\widetilde{\mu}^{S,N}$ to our original process $\bar{\mu}^{S,N}$, we will study the convergence of the following quantity, notably with the forthcoming Proposition 5.11:

$$\overline{\mathfrak{D}}^N(t):=\mathbb{E}_0^N\left[\frac{1}{N}\sum_{i\in\mathcal{S}^N(0)}\sup_{r\in[0,t]}|D_i^N(r)-\widetilde{D}_i^N(r)|\right], \tag{5.6}$$

We relate this convergence to the differences $\overline{\mathfrak{F}}_i^N(s)-\overline{\mathfrak{F}}(s,X_i^N)$ for $s\ge0$ and $i\in\mathcal{S}^N(0)$. We deduce from (2.8) and (3.7) that those differences can then be decomposed into seven terms, by exploiting the following definitions.

We first define

$$\overline{\mathfrak{A}}_i^N(t)=\sum_{j\in\mathcal{S}^N(0)}\omega^N(i,j)\cdot\big[\lambda_j^N(t-\tau_j^N)-\lambda_j^N(t-\widetilde{\tau}_j^N)\big], \tag{5.7}$$

where $\widetilde{\tau}_j^N$ is the first jump time corresponding to the process $(\widetilde{D}_j^N(s))_{s>0}$ in (5.1), so that $\overline{\mathfrak{A}}_i^N(t)$ captures the discrepancies between the jumps of (D_j^N) and those of $(\widetilde{D}_j^N)$.

Remark 5.2. *Similarly to $\mathcal{D}^N(t)$ in (2.10), we can define $\widetilde{\mathcal{D}}^N(t)$ accordingly to $(\widetilde{\tau}_j^N)$:*

$$\widetilde{\mathcal{D}}^N(t)=\big\{j\in\mathcal{S}_0^N;\,\widetilde{\tau}_j^N\le t\big\}, \tag{5.8}$$

that is the subset of individuals infected by time t, while being affected by the mean-field infection rate. For any $j\in\widetilde{\mathcal{D}}^N(t)$, $\widetilde{A}_j^N(t)=t-\widetilde{\tau}_j^N$ can be interpreted as the corresponding infection age, of individual j at time t, like $A_j^N(t)=t-\tau_j^N$ for any $j\in\mathcal{D}^N(t)$. Possibly $\widetilde{A}_j^N(t)>\eta_j^N$, thus the individual j has recovered by time t and $\lambda_j^N(t-\widetilde{\tau}_j^N)=0$. For any

$j \in \mathcal{S}_0^N \setminus \widetilde{\mathcal{D}}^N(t)$ on the other hand, $t - \widetilde{\tau}_j^N < 0$, which entails also $\lambda_j^N(t - \widetilde{\tau}_j^N) = 0$. Similar observations hold for the value of $\lambda_j^N(t - \tau_j^N)$ depending on whether $j \in \mathcal{D}^N(t)$ or not, and if yes whether $A_j^N(t) > \eta_j^N$ holds or not. It justifies the statement that $\overline{\mathfrak{A}}_i^N(t)$ corresponds to the component due to the discrepancies between infection ages.

We next define

$$\overline{\mathfrak{V}}_i^{N,1}(t) = \frac{1}{N} \sum_{j \in \mathcal{S}^N(0)} \Big[N\omega^N(i,j) \cdot \lambda_j^N(t - \widetilde{\tau}_j^N) - \bar{\omega}^N(X_i^N, X_j^N) \cdot \mathbb{E}_0^N[\bar{\lambda}(X_j^N, t - \widetilde{\tau}_j^N)] \Big], \quad (5.9)$$

so that $\overline{\mathfrak{V}}_i^{N,1}$ concerns the approximation of the transmission rate by its average, where the expectation in the last term averages the randomness of the infection time $\widetilde{\tau}_j^N$ with the exterior field $\overline{\mathfrak{F}}(., X_j^N)$ acting on individual j.

For any $x \in \mathbb{X}$ and $t \geq 0$, we define

$$\overline{\mathfrak{L}}^{N,1}(t,x) = \int_{\mathbb{X}} \Big[\bar{\omega}^N(x,x') - \bar{\omega}(x,x') \Big] \cdot \mathbb{E}[\bar{\lambda}(x', t - \widetilde{\tau}_{x'})] \bar{\mu}_0^{S,N}(\mathrm{d}x') \,, \quad (5.10)$$

where we define the random time $\widetilde{\tau}_{x'}$ whatever $x' \in \mathbb{X}$ as follows in term of some Poisson random measure Q on $\mathbb{R}_+^2$ with intensity dsdu:

$$\widetilde{\tau}_{x'} = \inf \Big\{ t \geq 0; \int_0^t \int_0^\infty \mathbb{1}_{\{u \leq \overline{\mathfrak{F}}(s,x')\}} Q(\mathrm{d}s, \mathrm{d}u) \geq 1 \Big\} . \quad (5.11)$$

Thus, $\overline{\mathfrak{L}}^{N,1}$ concerns the approximation of $\bar{\omega}^N$ by the kernel $\bar{\omega}$.

Remark 5.3. *The time $\widetilde{\tau}_{x'}$ will only be considered through expectations taken at fixed x' value, so that we are not concerned about letting the random measure Q depend on x'.*

The approximation of the initial condition $\bar{\mu}_0^{S,N}$ by $\bar{\mu}_0^S$ is treated separately with the next term, defined also for any $x \in \mathbb{X}$ and $t \geq 0$,

$$\overline{\mathfrak{E}}^{N,1}(t,x) = \int_{\mathbb{X}} \bar{\omega}(x,x') \cdot \mathbb{E}[\bar{\lambda}(x', t - \widetilde{\tau}_{x'})] \Big[\bar{\mu}_0^{S,N} - \bar{\mu}_0^S \Big](\mathrm{d}x'). \quad (5.12)$$

We will see in Remark 5.5 that $\overline{\mathfrak{E}}^N(t, X_i^N)$ can also be related to the approximation of the empirical measure process $(\widetilde{\mu}_s^{S,N})$ by the limiting $(\bar{\mu}_s^S)$.

We then define

$$\overline{\mathfrak{V}}_i^{N,0}(t) = \frac{1}{N} \sum_{j \in \mathcal{I}^N(0)} \Big[N\omega^N(i,j)\lambda_j^N(A_j^N(0) + t) - \bar{\omega}^N(X_i^N, X_j^N) \frac{\bar{\lambda}(X_j^N, A_j^N(0) + t)}{F^c_{X_j^N}(A_j^N(0))} \Big], \quad (5.13)$$

in a similar way as $\overline{\mathfrak{V}}_i^N(t)$ for the initially infected individuals. We recall from Assumption 2.6 that for any $j \in \mathcal{I}^N(0)$ the value of the conditional expectation $\bar{\lambda}(X_j^N, A_j^N(0)+t)$ is amplified by the denominator $F^c_{X_j^N}(A_j^N(0))$ to account for the bias in $\lambda_j^N(A_j^N(0)+t)$ due to the fact that individual j has not recovered by time 0. The next terms $\overline{\mathfrak{L}}^{N,0}$ and $\overline{\mathfrak{E}}^{N,0}$ are similarly the analogs of respectively $\overline{\mathfrak{L}}^{N,1}$ and $\overline{\mathfrak{E}}^{N,1}$ for the initially infected individuals. For any $x \in \mathbb{X}$

and $t \geq 0$,

$$\overline{\mathfrak{L}}^{N,0}(t,x) = \int_{\mathbb{X}}\int_0^\infty \left[\bar\omega^N(x,x') - \bar\omega(x,x')\right] \cdot \frac{\bar\lambda(x',a'+t)}{F^c_{x'}(a')}\, \bar\mu_0^{I,N}(\mathrm{d}x',\mathrm{d}a')\,, \tag{5.14}$$

and finally,

$$\overline{\mathfrak{E}}^{N,0}(t,x) = \int_{\mathbb{X}}\int_0^\infty \bar\omega(x,x') \cdot \frac{\bar\lambda(x',a'+t)}{F^c_{x'}(a')} \left[\bar\mu_0^{I,N}(\mathrm{d}x',\mathrm{d}a') - \bar\mu_0^{I}(\mathrm{d}x',\mathrm{d}a')\right], \tag{5.15}$$

so that $\overline{\mathfrak{E}}^{N,0}(t,X_i^N)$ accounts for the approximation of the initial condition $\bar\mu_0^{I,N}$ by $\bar\mu_0^I$.

Lemma 5.4. *With the above definitions, for any $N \geq 1$, $i \in \mathcal{S}^N(0)$, $t \geq 0$, we have*

$$\begin{aligned}&\overline{\mathfrak{F}}_i^N(t) - \overline{\mathfrak{F}}(t,X_i^N)\\ &= \overline{\mathfrak{V}}_i^{N,0}(t) + \overline{\mathfrak{L}}^{N,0}(t,X_i^N) + \overline{\mathfrak{E}}^{N,0}(t,X_i^N) + \overline{\mathfrak{A}}_i^N(t) + \overline{\mathfrak{V}}_i^{N,1}(t) + \overline{\mathfrak{L}}^{N,1}(t,X_i^N) + \overline{\mathfrak{E}}^{N,1}(t,X_i^N).\end{aligned}$$

Proof. From (5.11) it follows that $\mathbb{P}(\widetilde\tau_{x'} > s) = \exp\big(\int_0^s \overline{\mathfrak{F}}(r,x')\mathrm{d}r\big)$, hence the law of $\widetilde\tau_{x'}$ has the density $\overline{\mathfrak{F}}(s,x')\exp[\int_0^s \overline{\mathfrak{F}}(r,x')\mathrm{d}r]$. This fact combined with (3.10) leads to

$$\int_0^t\int_{\mathbb{X}} \bar\omega(x,x')\bar\lambda(x',t-s)\overline{\mathfrak{F}}(s,x')\bar\mu_s^S(\mathrm{d}x')\mathrm{d}s = \int_{\mathbb{X}} \bar\omega(x,x')\mathbb{E}[\bar\lambda(x',t-\widetilde\tau_{x'})]\bar\mu_0^S(\mathrm{d}x')\,. \tag{5.16}$$

Plugging this identity in (3.7), we deduce the decomposition $\overline{\mathfrak{F}}(t,x) = \overline{\mathfrak{F}}^0(t,x) + \overline{\mathfrak{F}}^1(t,x)$, where

$$\overline{\mathfrak{F}}^0(t,x) = \int_{\mathbb{X}}\int_0^\infty \bar\omega(x,x') \cdot \frac{\bar\lambda(x',a'+t)}{F^c_{x'}(a')}\, \bar\mu_0^I(\mathrm{d}x',\mathrm{d}a')\,, \tag{5.17}$$

$$\overline{\mathfrak{F}}^1(t,x) = \int_{\mathbb{X}} \bar\omega(x,x') \cdot \mathbb{E}[\bar\lambda(x',t-\widetilde\tau_{x'})]\, \bar\mu_0^S(\mathrm{d}x')\,. \tag{5.18}$$

We aim at a similar decomposition for $\overline{\mathfrak{F}}_i^N(t)$. Recall (2.10) where $\mathcal{D}^N(t)$ has been defined as the subset in $\mathcal{S}_0^N$ of individuals that have been infected by the disease by time t. If $j \in \mathcal{I}^N(t) \cap \mathcal{S}_0^N$, then $j \in \mathcal{D}^N(t)$ and thus $A_j^N(t) = t - \tau_j^N$. If, on the other hand, $j \in \mathcal{S}_0^N \setminus \mathcal{D}^N(t)$, then $A_j^N(t) = 0$, and thus $\lambda_j^N(A_j^N(t)) = 0$.

Recalling (2.11), we thus get the decomposition $\overline{\mathfrak{F}}_i^N(t) = \overline{\mathfrak{F}}_i^{N,0}(t) + \overline{\mathfrak{F}}_i^{N,1}(t)$, where

$$\overline{\mathfrak{F}}_i^{N,0}(t) = \sum_{j\in\mathcal{I}^N(0)} \omega^N(i,j)\cdot\lambda_j^N(A_j^N(0)+t) \tag{5.19}$$

$$\overline{\mathfrak{F}}_i^{N,1}(t) = \sum_{j\in\mathcal{S}^N(0)} \omega^N(i,j)\cdot\lambda_j^N(t-\tau_j^N)\,. \tag{5.20}$$

We first show that $\overline{\mathfrak{F}}_i^{N,0}(t) - \overline{\mathfrak{F}}^0(t,X_i^N) = \overline{\mathfrak{V}}_i^{N,0}(t) + \overline{\mathfrak{L}}^{N,0}(t,X_i^N) + \overline{\mathfrak{E}}^{N,0}(t,X_i^N)$, and next that $\overline{\mathfrak{F}}_i^{N,1}(t) - \overline{\mathfrak{F}}^1(t,X_i^N) = \overline{\mathfrak{A}}_i^N(t) + \overline{\mathfrak{V}}_i^{N,1}(t) + \overline{\mathfrak{L}}^{N,1}(t,X_i^N) + \overline{\mathfrak{E}}^{N,1}(t,X_i^N)$.

By combining (5.13) with (5.19), then with the definition of $\bar\mu_0^{I,N}$ in (3.1), we have

$$\begin{aligned}\overline{\mathfrak{F}}_i^{N,0}(t) - \overline{\mathfrak{V}}_i^{N,0}(t) &= \frac{1}{N}\sum_{j\in\mathcal{I}^N(0)} \bar\omega^N(X_i^N,X_j^N)\frac{\bar\lambda(X_j^N,A_j^N(0)+t)}{F^c_{X_j^N}(A_j^N(0))}\\ &= \int_{\mathbb{X}}\int_0^\infty \bar\omega^N(X_i^N,x') \cdot \frac{\bar\lambda(x',a'+t)}{F^c_{x'}(a')}\, \bar\mu_0^{I,N}(\mathrm{d}x',\mathrm{d}a')\,.\end{aligned}$$

Plugging (5.14) into this expression, then exploiting (5.15) and (5.17), we obtain

$$\begin{aligned} \overline{\mathfrak{F}}_i^{N,0}(t) - \overline{\mathfrak{V}}_i^{N,0}(t) - \overline{\mathfrak{L}}^{N,0}(t, X_i^N) &= \int_{\mathbb{X}} \int_0^\infty \bar{\omega}(X_i^N, x') \cdot \frac{\bar{\lambda}(x', a'+t)}{F_{x'}^c(a')} \bar{\mu}_0^{I,N}(\mathrm{d}x', \mathrm{d}a') \\ &= \overline{\mathfrak{E}}^{N,0}(t, X_i^N) + \overline{\mathfrak{F}}^0(t, X_i^N)\,, \end{aligned} \tag{5.21}$$

which concludes our first claim. For the second claim, there is a first additional step where $A_i^N(t)$ is related to $\widetilde{A}_i^N(t) = t - \widetilde{\tau}_i^N$, through $\overline{\mathfrak{A}}_i^N(t)$, before we can exploit the same arguments. We first combine (5.7) with (5.20), and hence,

$$\overline{\mathfrak{F}}_i^{N,1}(t) - \overline{\mathfrak{A}}_i^N(t) = \sum_{j \in \mathcal{S}^N(0)} \omega^N(i,j) \lambda_j^N(t - \widetilde{\tau}_j^N)\,.$$

Plugging (5.9) into this expression and exploiting the definitions of $\bar{\mu}^{S,N}$ in (3.1) and of $\widetilde{\tau}_{x'}$ in (5.11), we deduce that

$$\begin{aligned} \overline{\mathfrak{F}}_i^{N,1}(t) - \overline{\mathfrak{A}}_i^N(t) - \overline{\mathfrak{V}}_i^{N,1}(t) &= \frac{1}{N} \sum_{j \in \mathcal{S}^N(0)} \bar{\omega}^N(X_i^N, X_j^N) \cdot \mathbb{E}_0^N[\bar{\lambda}(X_j^N, t - \widetilde{\tau}_j^N)] \\ &= \int_{\mathbb{X}} \bar{\omega}^N(X_i^N, x') \mathbb{E}[\bar{\lambda}(x', t - \widetilde{\tau}_{x'})] \bar{\mu}_0^{S,N}(\mathrm{d}x')\,. \end{aligned} \tag{5.22}$$

Next plugging (5.10) into this expression, then exploiting (5.12) and (5.18), we get

$$\begin{aligned} \overline{\mathfrak{F}}_i^{N,1}(t) - \overline{\mathfrak{A}}_i^N(t) - \overline{\mathfrak{V}}_i^{N,1}(t) - \overline{\mathfrak{L}}^{N,1}(t, X_i^N) &= \int_{\mathbb{X}} \bar{\omega}(X_i^N, x') \mathbb{E}[\bar{\lambda}(x', t - \widetilde{\tau}_{x'})] \bar{\mu}_0^{S,N}(\mathrm{d}x') \\ &= \overline{\mathfrak{E}}^{N,1}(t, X_i^N) + \overline{\mathfrak{F}}^1(t, X_i^N). \end{aligned} \tag{5.23}$$

Since $\overline{\mathfrak{F}}(t,x) = \overline{\mathfrak{F}}^0(t,x) + \overline{\mathfrak{F}}^1(t,x)$ and $\overline{\mathfrak{F}}_i^N(t) = \overline{\mathfrak{F}}_i^{N,0}(t) + \overline{\mathfrak{F}}_i^{N,1}(t)$, combining (5.21) and (5.23) concludes the proof of Lemma 5.4. □

Remark 5.5. *In the expression of $\overline{\mathfrak{E}}^{N,1}(t,x)$ in (5.12), we decided to relate as directly as possible to the difference between $\bar{\mu}_0^{S,N}$ and $\bar{\mu}_0^S$. That being said, this term has the following alternative interpretation in terms of the difference between the processes $\widetilde{\mu}_s^{S,N}$ and $\bar{\mu}_s^S$:*

$$\overline{\mathfrak{E}}^{N,1}(t,x) = \mathbb{E}_0^N\left[\int_0^t \int_{\mathbb{X}} \bar{\omega}(x,x') \bar{\lambda}(x', t-s) \cdot \overline{\mathfrak{F}}(s,x') \left[\widetilde{\mu}_s^{S,N} - \bar{\mu}_s^S \right](\mathrm{d}x')\, \mathrm{d}s \right]. \tag{5.24}$$

Proof of (5.24). Recall the identity (5.16). Similarly, we express $\widetilde{\tau}_j^N$ through the Poisson random measure Q_j:

$$\begin{aligned} &\frac{1}{N} \sum_{j \in \mathcal{S}^N(0)} \bar{\omega}(x, X_j^N) \mathbb{E}_0^N[\bar{\lambda}(X_j^N, t - \widetilde{\tau}_j^N)] \qquad (5.25) \\ &= \mathbb{E}_0^N\left[\frac{1}{N} \sum_{j \in \mathcal{S}^N(0)} \int_0^t \int_0^\infty \bar{\omega}(x, X_j^N) \bar{\lambda}(X_j^N, t-s) \mathbb{1}_{\left\{\widetilde{D}_j^N(s^-) = 0\right\}} \mathbb{1}_{\left\{u \le \overline{\mathfrak{F}}(s^-, X_j^N)\right\}} Q_j(\mathrm{d}s, \mathrm{d}u) \right]. \end{aligned}$$

In the expression in the second line, we can replace Q_j by its intensity since the integrant is predictable with respect to its filtration $(\widetilde{\mathcal{F}}_t^N)$, then express the sum over $j \in \mathcal{S}^N(0)$ in terms of the empirical measure $\widetilde{\mu}_s^{S,N}$, recall (5.2), which leads to

$$\mathbb{E}_0^N\Big[\int_0^t \int_{\mathbb{X}} \bar{\omega}(x,x')\bar{\lambda}(x',t-s)\overline{\mathfrak{F}}(s,x')\widetilde{\mu}_s^{S,N}(\mathrm{d}x')\mathrm{d}s\Big]. \tag{5.26}$$

Recalling (5.22) in addition to (5.16), (5.25) and (5.26), we deduce identity (5.24). □

In the following, we treat separately the various terms distinguished in Lemma 5.4, first $\overline{\mathfrak{A}}_i^N$, see Lemma 5.6, second $\overline{\mathfrak{V}}^{N,1}$ and $\overline{\mathfrak{V}}^{N,0}$, see Lemma 5.7, third $\overline{\mathfrak{L}}^{N,1}$ and $\overline{\mathfrak{L}}^{N,0}$, see Lemma 5.8, and finally $\overline{\mathfrak{E}}^{N,1}$ and $\overline{\mathfrak{E}}^{N,0}$, see Lemma 5.10. We start with the processes $(\overline{\mathfrak{A}}_i^N)$ defined in (5.7).

Lemma 5.6. *Under Assumptions 2.2, 3.1, 3.3 and 3.5, there exist a constant $C > 0$ such that*

$$\mathbb{E}_0^N\left[\frac{1}{N}\sum_{i\in\mathcal{S}^N(0)}\left|\overline{\mathfrak{A}}_i^N(t)\right|\right] \le C\cdot\overline{\mathfrak{D}}^N(t)+\overline{\mathfrak{V}}^N$$

holds a.s. for any $t \ge 0$, where $\overline{\mathfrak{D}}^N(t)$ is defined in (5.6), and $\overline{\mathfrak{V}}^N$ is given by

$$\overline{\mathfrak{V}}^N := \lambda^*\cdot\sqrt{\Upsilon^N+\omega^*\cdot\bar{\gamma}^N}\,. \tag{5.27}$$

We recall the defining property of ω^* given just after Lemma 3.4.

Proof. Recalling (5.7) and exploiting the upper-bound λ^* of the functions (λ_j^N), we get

$$\sum_{i\in\mathcal{S}^N(0)}\left|\overline{\mathfrak{A}}_i^N(t)\right| \le \lambda^*\cdot\sum_{j\in\mathcal{S}^N(0)}\sup_{r\le t}|D_j^N(r)-\widetilde{D}_j^N(r)|\cdot\sum_{i\in\mathcal{S}^N(0)}\omega^N(i,j). \tag{5.28}$$

We observe for any $j\in\mathcal{S}^N(0)$, the following identity by virtue of (2.7):

$$\mathbb{E}_0^N\Big[\sum_{i\in\mathcal{S}^N(0)}\omega^N(i,j)\Big] = \int_{\mathbb{X}}\bar{\omega}^N(x,X_j^N)\bar{\mu}_0^{S,N}(\mathrm{d}x).$$

Since $\bar{\omega}^N$ is upper-bounded by ω^* and $\bar{\mu}_0^{S,N}(\mathbb{X})\le 1$, we obtain the upper bound

$$\frac{1}{N}\sum_{i\in\mathcal{S}^N(0)}\left|\overline{\mathfrak{A}}_i^N(t)\right| \le \frac{\lambda^*}{N}\sum_{j\in\mathcal{S}^N(0)}\sup_{r\le t}|D_j^N(r)-\widetilde{D}_j^N(r)|\cdot(\widetilde{S}_j^N+\omega^*)\,, \tag{5.29}$$

where

$$\widetilde{S}_j^N := \sum_{i\in\mathcal{S}^N(0)}\omega^N(i,j)-\mathbb{E}_0^N\Big[\sum_{i\in\mathcal{S}^N(0)}\omega^N(i,j)\Big]\,. \tag{5.30}$$

Since $\sup_{r\le t}|D_j^N(r)-\widetilde{D}_j^N(r)|$ is upper-bounded by 1 for any $j\in\mathcal{S}^N(0)$, (5.29) entails

$$\mathbb{E}_0^N\left[\frac{1}{N}\sum_{i\in\mathcal{S}^N(0)}\left|\overline{\mathfrak{A}}_i^N(t)\right|\right] \le C\cdot\overline{\mathfrak{D}}^N(t)+\overline{\mathfrak{U}}^N\,,$$

where $C=\lambda^*\cdot\omega^*$ and

$$\overline{\mathfrak{U}}^N := \frac{\lambda^*}{N}\mathbb{E}_0^N\left[\sum_{j\in\mathcal{S}^N(0)}\left|\widetilde{S}_j^N\right|\right]. \tag{5.31}$$

So Lemma 5.6 will be concluded by showing that $\overline{\mathfrak{U}}^N \leq \overline{\mathfrak{V}}^N$. Thanks twice to the Cauchy-Schwartz inequality, with the fact that the cardinality of $\mathcal{S}^N(0)$ is less than N:

$$\begin{aligned}\mathbb{E}_0^N\left[\frac{1}{N}\sum_{j\in\mathcal{S}^N(0)}\left|\widetilde{S}_j^N\right|\right] &\leq \frac{1}{N}\sum_{j\in\mathcal{S}^N(0)}\sqrt{\mathbb{E}_0^N\left[\left(\widetilde{S}_j^N\right)^2\right]} \\ &\leq \sqrt{\frac{1}{N}\sum_{j\in\mathcal{S}^N(0)}\mathbb{E}_0^N\left[\left(\widetilde{S}_j^N\right)^2\right]}\,.\end{aligned} \tag{5.32}$$

For any $j \in \mathcal{S}^N(0)$ and conditionally on $\mathcal{F}_0^N$, $\widetilde{S}_j^N$ is the sum of independent centered variables, which leads to the following identity:

$$\mathbb{E}_0^N\left[\left(\widetilde{S}_j^N\right)^2\right] = \sum_{i\in\mathcal{S}^N(0)} \mathrm{Var}_0^N\left[\omega^N(i,j)\right].$$

We recall that $\mathrm{Var}_0^N(Z) = \mathbb{E}_0^N(Z^2) - \mathbb{E}_0^N(Z)^2$ by definition of this variance conditional on $\mathcal{F}_0^N$ for any random variable Z. As we will need later the conditional second moment of $\omega^N(i,j)$ instead of its variance, we rather consider it as the upper-bound, then exploit (2.6) and (2.3) to deduce the following inequality:

$$\begin{aligned}\mathbb{E}_0^N\left[\left(\widetilde{S}_j^N\right)^2\right] &\leq \sum_{i\in\mathcal{S}^N(0)} \mathbb{E}_0^N\left[\omega^N(i,j)^2\right] \\ &= \sum_{i\in\mathcal{S}^N(0)} \mathbb{E}_0^N\left[\upsilon^N(i,j);(i,j)\in\mathcal{E}^N\right] + \sum_{i\in\mathcal{S}^N(0)} \kappa^N(X_i^N,X_j^N)\cdot\gamma^N(X_i^N,X_j^N)^2\,.\end{aligned} \tag{5.33}$$

Remark that for any $i,j \in \mathcal{S}^N(0)$, $\kappa^N(X_i^N,X_j^N)\cdot\gamma^N(X_i^N,X_j^N) = \bar{\omega}^N(X_i^N,X_j^N)/N \leq \omega^*/N$. Recalling the definitions given in Assumption 3.5, we thus deduce

$$\frac{1}{N}\sum_{j\in\mathcal{S}^N(0)}\mathbb{E}_0^N\left[\left(\widetilde{S}_j^N\right)^2\right] \leq \Upsilon^N + \omega^*\cdot\bar{\gamma}^N\,. \tag{5.34}$$

Recalling (5.31) and (5.32), this entails $\overline{\mathfrak{U}}^N \leq \overline{\mathfrak{V}}^N$ a.s. and concludes the proof of Lemma 5.6. □

For the upper-bound of $\overline{\mathfrak{V}}_i^{N,1}$ and $\overline{\mathfrak{V}}_i^{N,0}$ as defined in respectively (5.9) and (5.13), the proof of the next lemma follows similar principles as in the previous one.

Lemma 5.7. *The following upper-bound holds for any $t > 0$ with the sequence $(\bar{\mathfrak{V}}^N)_{N\geq1}$ defined in (5.27):*

$$\mathbb{E}_0^N\Big[\frac{1}{N}\sum_{i\in\mathcal{S}^N(0)}\left|\overline{\mathfrak{V}}_i^{N,1}(t)\right|\Big] \vee \mathbb{E}_0^N\Big[\frac{1}{N}\sum_{i\in\mathcal{S}^N(0)}\left|\overline{\mathfrak{V}}_i^{N,0}(t)\right|\Big] \leq \overline{\mathfrak{V}}^N\,.$$

Proof. We treat this component similarly as $\mathfrak{U}^N$, first thanks to the Cauchy-Schwartz inequality:

$$\mathbb{E}_0^N\Big[\frac{1}{N}\sum_{i\in\mathcal{S}^N(0)}\big|\overline{\mathfrak{V}}_i^{N,1}(t)\big|\Big] \leq \frac{1}{N}\sum_{i\in\mathcal{S}^N(0)}\sqrt{\mathbb{E}_0^N\Big[\big(\overline{\mathfrak{V}}_i^{N,1}(t)\big)^2\Big]}$$
$$\leq \sqrt{\frac{1}{N}\sum_{i\in\mathcal{S}^N(0)}\mathbb{E}_0^N\Big[\big(\overline{\mathfrak{V}}_i^{N,1}(t)\big)^2\Big]},$$

Recalling (5.9), we note that the random variables $\overline{\mathfrak{V}}_i^{N,1}(t)$ are centered conditionally on $\mathcal{F}_0^N$, so that the term under the square root is actually a variance. $\overline{\mathfrak{V}}_i^{N,1}(t)$ is a sum of r.v.'s which are orthogonal in $L^2(\Omega,\mathbb{P}_0^N)$. Thus, for any $i\in\mathcal{S}^N(0)$,

$$\mathbb{E}_0^N\Big[\big(\overline{\mathfrak{V}}_i^{N,1}(t)\big)^2\Big] = \sum_{j\in\mathcal{S}^N(0)}\mathrm{Var}_0^N\big[\omega^N(i,j)\cdot\lambda_j^N(t-\widetilde{\tau}_j^N)\big].$$

For any $i,j\in\mathcal{S}^N(0)$, since $\omega^N(i,j)$ and $\lambda_j^N(t-\widetilde{\tau}_j^N)$ are independent conditionally on $\mathcal{F}_0^N$ and since the functions (λ_j^N) are uniformly upper-bounded by λ^*:

$$\mathrm{Var}_0^N\big[\omega^N(i,j)\cdot\lambda_j^N(t-\widetilde{\tau}_j^N)\big] \leq (\lambda^*)^2\cdot\mathbb{E}_0^N\big[\omega^N(i,j)^2\big]. \tag{5.35}$$

The argument for the following inequality is then the same as for (5.34):

$$\frac{1}{N}\sum_{i\in\mathcal{S}^N(0)}\mathbb{E}_0^N\Big[\big(\overline{\mathfrak{V}}_i^{N,1}(t)\big)^2\Big] \leq (\lambda^*)^2\cdot\Big[\Upsilon^N+\omega^*\cdot\bar{\gamma}^N\Big]. \tag{5.36}$$

Concerning the sequence $(\overline{\mathfrak{V}}_i^{N,0}(t))$, we first recall the following identity for any $j\in\mathcal{I}^N(0)$ as part of Assumption 2.6:

$$\mathbb{E}_0^N\Big[\lambda_j^N(A_j^N(0)+t)\Big] = \frac{\bar{\lambda}(X_j^N,A_j^N(0)+t)}{F^c_{X_j^N}(A_j^N(0))}. \tag{5.37}$$

Since $\omega^N(i,j)$ and $\lambda_j^N(A_j^N(0)+t)$ are independent conditionally on $\mathcal{F}_0^N$, we deduce that $\overline{\mathfrak{V}}_i^{N,0}(t)$ is also conditionally centered, whatever $i\in\mathcal{S}^N(0)$ and $t\geq 0$. We can then exploit the same argument for $\overline{\mathfrak{V}}_i^{N,0}(t)$ as for $\overline{\mathfrak{V}}_i^{N,1}(t)$, thanks to the Cauchy-Schwartz inequality and replacing (5.35) by

$$\mathrm{Var}_0^N\big[\omega^N(i,j)\cdot\lambda_j^N(A_j^N(0)+t)\big] \leq (\lambda^*)^2\cdot\mathbb{E}_0^N\big[\omega^N(i,j)^2\big].$$

Lemma 5.7 is therefore concluded with $\overline{\mathfrak{V}}^N$ in (5.27). □

$\overline{\mathfrak{L}}^{N,1}$ and $\overline{\mathfrak{L}}^{N,0}$ as defined respectively in (5.10) and (5.14) are quite directly upper-bounded.

Lemma 5.8. *By virtue of Assumptions 2.6 and 3.1:*

$$\Big|\overline{\mathfrak{L}}^{N,1}(t,x)\Big| + \Big|\overline{\mathfrak{L}}^{N,0}(t,x)\Big| \leq \lambda^*\cdot\|\bar{\omega}^N-\bar{\omega}\|_\infty$$

holds for any $t>0$ and $x\in\mathbb{X}$.

Proof. We first upper-bound $|\bar{\omega}^N(x,x') - \bar{\omega}(x,x')|$ by $\|\bar{\omega}^N - \bar{\omega}\|_\infty$. By virtue of Assumption 2.6, we exploit λ^* as the upper-bound for any $x' \in \mathbb{X}$ and any $a', t \in \mathbb{R}_+$ of $\mathbb{E}[\bar{\lambda}(x', t - \widetilde{\tau}_{x'})]$ in (5.10) and of $\bar{\lambda}(x', a'+t)/F^c_{x'}(a'+t)$. Since $F^c_{x'}$ is non-increasing for any such x' by virtue of Assumption 2.4, the latter upper-bound entails the one of $\bar{\lambda}(x', a'+t)/F^c_{x'}(a')$ in (5.14). Finally, by virtue of the considered scaling of $\bar{\mu}^{S,N}(\mathrm{d}x')$ and of $\bar{\mu}^{I,N}(\mathrm{d}x', \mathrm{d}a')$, recall (3.1), their added masses is upper-bounded by 1, which concludes the proof of Lemma 5.8. ☐

In order to finally deal with the upper-bound of both $\overline{\mathfrak{E}}^{N,1}$ and $\overline{\mathfrak{E}}^{N,0}$, defined respectively in (5.12) and (5.15), we exploit the following proposition which will be proved in Appendix A.

Proposition 5.9. *Let $\mathcal{X}$ and $\mathcal{Y}$ be Polish spaces. Let $\mu \in \mathcal{M}(\mathcal{X})$, $\nu \in \mathcal{M}(\mathcal{Y})$ and $k : \mathcal{X} \times \mathcal{Y} \mapsto \mathbb{R}$ be a bounded measurable function that is continuous $\mu \otimes \nu$ almost everywhere. Let in addition $(\mu^N)_{N\geq 1}$ and $(\nu^N)_{N\geq 1}$ be two sequences of possibly random measures in $\mathcal{M}(\mathcal{X})$ and $\mathcal{M}(\mathcal{Y})$, respectively, that converge in probability respectively to μ and ν, for the topology of weak convergence. Then, the following quantity converges to zero in probability as N tends to infinity:*

$$\int_{\mathcal{X}} \Big| \int_{\mathcal{Y}} k(x,y)[\nu^N - \nu](dy) \Big| \mu^N(\mathrm{d}x).$$

We are now ready for the estimate of $\overline{\mathfrak{E}}^{N,1}$ and $\overline{\mathfrak{E}}^{N,0}$ provided in the next lemma.

Lemma 5.10. *For any $t > 0$, the following random variable converges to 0 in probability as N tends to infinity:*

$$\overline{\mathfrak{E}}^N(t) := \langle \bar{\mu}_0^{S,N}, |\overline{\mathfrak{E}}^{N,1}(t,.)| + |\overline{\mathfrak{E}}^{N,0}(t,.)| \rangle \,.$$

Proof. Recalling (5.15), the fact that $\langle \bar{\mu}_0^{S,N}, |\overline{\mathfrak{E}}^{N,0}(t,.)| \rangle$ converges to zero is a direct consequence of Proposition 5.9 with $\mathcal{X} = \mathbb{X}$, $\mathcal{Y} = \mathbb{X} \times \mathbb{R}_+$ and

$$\mu^N := \bar{\mu}_0^{S,N}, \quad \mu := \bar{\mu}_0^S, \qquad k(x,x',a') = \begin{cases} \bar{\omega}(x,x') \cdot \frac{\bar{\lambda}(x',a'+t)}{F^c_{x'}(a')} & \text{if } F^c_{x'}(a'+t) > 0 \\ 0 & \text{otherwise,} \end{cases}$$

$$\nu^N(\mathrm{d}x', \mathrm{d}a') := \bar{\mu}_0^{I,N}(\mathrm{d}x', \mathrm{d}a'), \quad \nu(\mathrm{d}x', \mathrm{d}a') := \bar{\mu}_0^I(\mathrm{d}x', \mathrm{d}a')\,.$$

Assumption 3.1 states that μ^N and ν^N converge weakly to respectively μ and ν. Exploiting first the fact that $a \mapsto F^c_{x'}(a)$ is non increasing, and next Assumption 2.6, we have that for all $x' \in \mathbb{X}, a', t \geq 0$,

$$\frac{\bar{\lambda}(x', a'+t)}{F^c_{x'}(a')} \leq \frac{\bar{\lambda}(x', a'+t)}{F^c_{x'}(a'+t)} \leq \lambda^*\,.$$

We recall also the uniform upper-bound ω^* of $\bar{\omega}^N$, as defined just after Lemma 3.4. $\omega^* \cdot \lambda^*$ thus defines a global upper-bound of k.

Note that $\bar{\lambda}(x', a') < \lambda^* \cdot F^c_{x'}(a')$ as long as $F^c_{x'}(a') > 0$, by virtue of Assumption 2.6. By virtue of Assumptions 2.4, 2.6 and 3.3, k is $(\mu \otimes \nu)$ a.e. continuous, those points (x, x', a') that satisfy $F^c_{x'}(a') = 0$ being handled specifically thanks to the above property. Indeed, thanks to the second part of Assumption 3.3 for $(\mu \otimes \nu)$ almost every such realization, since $t > 0$, we know that for any $(\check{x}, \check{x}', \check{a}')$ in a small enough neighborhood of (x, x', a') that $F^c_{\check{x}'}(\check{a}' + t) = 0$, thus $\bar{\lambda}(\check{x}', \check{a}' + t) = 0$ and $k(\check{x}, \check{x}', \check{a}') = 0$.

Recalling (5.12), the fact that $\langle \bar{\mu}_0^{S,N}, |\overline{\mathfrak{E}}^{N,1}(t,.)| \rangle$ converges to zero is a consequence as well of Proposition 5.9 with this time $\mathcal{Y} = \mathbb{X}$, still $\mathcal{X} = \mathbb{X}$ and

$$\mu^N = \nu^N := \bar{\mu}_0^{S,N}, \quad \mu = \nu = \bar{\mu}_0^S, \quad k(x,x') = \bar{\omega}(x,x') \cdot \mathbb{E}[\bar{\lambda}(x', t - \widetilde{\tau}_{x'})]\,.$$

The fact that μ^N and ν^N converge weakly in probability to $\mu = \nu$ follows from Assumption 3.1. k is bounded under Assumptions 2.2 and 3.3. To check that k is $(\mu \otimes \nu)$ a.e. continuous, we exploit the following alternative expression:

$$k(x, x') = \bar{\omega}(x, x') \int_0^t \bar{\lambda}(x', t-s) \cdot \overline{\mathfrak{F}}(s, x') \cdot \exp\Big[- \int_0^s \overline{\mathfrak{F}}(r, x') \mathrm{d}r \Big],$$

derived with the same argument as for (5.24). Thanks to Assumptions 2.2 and 3.3 and Lemma 4.2, we then conclude that k is $(\mu \otimes \nu)$ a.e. continuous. □

With Lemmas 5.6, 5.7, 5.8 and 5.10, we are ready to prove the following comparison result with the original model.

Proposition 5.11. *As $N \to \infty$, $\overline{\mathfrak{D}}^N(t)$ defined in (5.6) converges in probability to 0 locally uniformly in t.*

Proof. First note that it suffices to prove the convergence for any fixed t. The locally uniform convergence then follows from Lemma A.1 in Appendix A since : $t \mapsto \overline{\mathfrak{D}}^N(t)$ is a.s. non-decreasing for any N. For any i and t,

$$\sup_{r \le t} |D_i^N(r) - \widetilde{D}_i^N(r)| \le \int_0^t \int_0^\infty \mathbb{1}_{\{u \in \mathcal{B}_i^N(s)\}} Q_i(\mathrm{d}s, \mathrm{d}u),$$

where the interval $\mathcal{B}_i^N(s)$ is defined as follows:

$$\mathcal{B}_i^N(s) = \big[\overline{\mathfrak{F}}_i^N(s) \wedge \overline{\mathfrak{F}}(s, X_i^N),\, \overline{\mathfrak{F}}_i^N(s) \vee \overline{\mathfrak{F}}(s, X_i^N)\big],$$

with a length equal to $\big|\overline{\mathfrak{F}}_i^N(s) - \overline{\mathfrak{F}}(s, X_i^N)\big|$. Summing over i and taking expectation on the $(Q_i)_{i \le N}$, we obtain

$$\overline{\mathfrak{D}}^N(t) \le \mathbb{E}_0^N \Big[\frac{1}{N} \sum_{i \in \mathcal{S}^N(0)} \int_0^t |\overline{\mathfrak{F}}_i^N(s) - \overline{\mathfrak{F}}(s, X_i^N)| \mathrm{d}s \Big]. \tag{5.38}$$

Starting from (5.38), for any i and t, we decompose the integrand into seven terms according to Lemma 5.4. Five among these terms are first treated thanks to Lemmas 5.6, 5.7 and 5.8, so that there exists $C > 0$, $\overline{\mathfrak{V}}^N = \lambda^* \cdot \sqrt{\Upsilon^N + \omega^* \cdot \bar{\gamma}^N}$ and $\overline{\mathfrak{L}}^N = \lambda^* \cdot \|\bar{\omega}^N - \bar{\omega}\|_\infty$, the later two converging in probability to zero by virtue respectively of Assumptions 3.5 and 3.3, such that

$$\overline{\mathfrak{D}}^N(t) \le C \cdot \int_0^t \overline{\mathfrak{D}}^N(s) + t \cdot (2\overline{\mathfrak{V}}^N + \overline{\mathfrak{L}}^N) + \int_0^t \langle \bar{\mu}_0^{S,N}, |\overline{\mathfrak{E}}^{N,1}(s,.)| + |\overline{\mathfrak{E}}^{N,0}(s,.)| \rangle \mathrm{d}s. \tag{5.39}$$

Thanks to Lemma 5.10, the integrand $(\langle \bar{\mu}_0^{S,N}, |\overline{\mathfrak{E}}^{N,1}(s,.)| + |\overline{\mathfrak{E}}^{N,0}(s,.)| \rangle)_{s \in [0,t]}$ converges to 0 pointwise in probability for any $s > 0$. By virtue of Assumption 2.6 and Lemma 3.4, for the defining properties of λ^* and ω^*, we obtain that

$$|\overline{\mathfrak{E}}^{N,1}(s, x)| \vee |\overline{\mathfrak{E}}^{N,0}(s,.)| \le \omega^* \lambda^*,$$

holds for any $s \ge 0$ and any $x \in \mathbb{X}$. Since $\bar{\mu}_0^{S,N}(\mathbb{X}) \le 1$, the integrand is itself upper-bounded by $\omega^* \lambda^*$. By Lebesgue's dominated convergence theorem, we deduce the convergence to 0 in probability as N tends to infinity of the following random variable for any t, the r.v. being non-decreasing with t:

$$\int_0^t \langle \bar{\mu}_0^{S,N}, |\overline{\mathfrak{E}}^{N,1}(s,.)| + |\overline{\mathfrak{E}}^{N,0}(s,.)|\rangle \mathrm{d}s\,.$$

With (5.39), we are therefore in situation to apply Gronwall's inequality and conclude Proposition 5.11 in that $\overline{\mathfrak{D}}^N(t)$ converges to 0 in probability for any t as $N \to \infty$. □

Before proceeding, let us state a result in Proposition 5.12 which is exactly Theorem II.4.1 in [46] and will be needed in the next proof. Let us generally consider a Polish space $\mathcal{X}$ and the set $\mathcal{C}_b(\mathcal{X})$ of bounded continuous functions on $\mathcal{X}$. We say that a subset $\mathfrak{M}$ of $\mathcal{C}_b(\mathcal{X})$ is separating if (i) it includes the constant function : $x \mapsto 1$ and (ii) it discriminates elements of $\mathcal{M}_1(\mathcal{X})$, in that for any $(\nu,\nu') \in \mathcal{M}_1(\mathcal{X})^2$, $\nu = \nu'$ is equivalent to the property that $\langle \nu, \phi\rangle = \langle \nu, \phi\rangle$ for any $\phi \in \mathfrak{M}$. It is classical that $\mathcal{C}_b(\mathcal{X})$ itself is separating.

Proposition 5.12. *A sequence of processes $(\nu^N)_{n\in\mathbb{N}^*}$ is C-tight in $\mathcal{D}$ if and only if:*

(a) ***Compact Containment Condition (CCC).*** *For all $\varepsilon > 0$ and T, there exists a compact set K_ε in $\mathcal{X}$ such that:*

$$\sup_{N\in\mathbb{N}^*} \mathbb{P}\left(\sup_{t\le T} \nu_t^N(K_\varepsilon^c) > \varepsilon\right) < \varepsilon.$$

(b) ***Tightness of the projections.*** *The sequence $(\langle \nu_.^N, \varphi\rangle)_{N\in\mathbb{N}^*}$ is C-tight in $\mathcal{D}(\mathbb{R}_+,\mathbb{R})$ for any function φ in a separating class $\mathfrak{M}$.*

We can then conclude to the convergence of the measure $\bar{\mu}^{S,N}$:

Proposition 5.13. *As $N \to \infty$, the following convergence holds in probability*

$$\bar{\mu}^{S,N} \to \bar{\mu}^S \quad \text{in} \quad \mathcal{D}(\mathbb{R}_+, \mathcal{M}(\mathbb{X}))\,.$$

Proof. We apply Proposition 5.12 with $\mathcal{X} = \mathbb{X}$, $\nu^N = \bar{\mu}^{(S,N)}$ and $\mathfrak{M} = \mathcal{C}_{b+}(\mathbb{X})$ the set of bounded continuous functions from $\mathbb{X}$ to $\mathbb{R}_+$. $\mathcal{C}_{b+}(\mathbb{X})$ is separating since $\mathcal{C}_b(\mathbb{X})$ is itself separating. Point (a) follows readily from the two facts: $\bar{\mu}_t^{(S,N)}(K_\varepsilon^c) \le \bar{\mu}_0^{(S,N)}(K_\varepsilon^c)$, and $\mu_0^{(S,N)} \Rightarrow \mu_0^S$ in probability (exploiting for instance the Lévy-Prokhorov metric).

Concerning Point (b), let us consider any $\varphi \in C_b(\mathbb{X})$, $\varphi \ge 0$. Recall that from Lemma 5.1 $\langle \widetilde{\mu}_t^{(S,N)}, \varphi\rangle$ converges in probability to $\langle \bar{\mu}_t^S, \varphi\rangle$ for any $t > 0$. Moreover, by definition of $\widetilde{\mu}^{(S,N)}$ and $\overline{\mathfrak{D}}^N$ in respectively (5.2) and (5.6):

$$\big|\langle \bar{\mu}_t^{(S,N)}, \varphi\rangle - \langle \widetilde{\mu}_t^{(S,N)}, \varphi\rangle\big| \le \|\varphi\|_\infty \cdot \overline{\mathfrak{D}}_t^N\,, \tag{5.40}$$

with an upper-bound that converges to 0 in probability locally uniformly in t thanks to Proposition 5.11. Therefore, $\langle \bar{\mu}_t^{(S,N)}, \varphi\rangle$ converges in probability to $\langle \bar{\mu}_t^S, \varphi\rangle$ pointwise for any $t > 0$. In addition, since φ is non-negative, $t \to \langle \bar{\mu}_t^{(S,N)}, \varphi\rangle$ is non-increasing for each $N \ge 1$. Also, $t \to \langle \bar{\mu}_t^S, \varphi\rangle$ is continuous for any N. Thanks to Lemma A.1 in Appendix A, the convergence in probability of $\langle \bar{\mu}_t^{(S,N)}, \varphi\rangle$ to $\langle \bar{\mu}_t^S, \varphi\rangle$ therefore holds locally uniform in t.

Thanks to Proposition 5.12, the sequence $(\bar{\mu}^{(S,N)})_{N\in\mathbb{N}^*}$ is thus C-tight. By the convergence of the projection, any limit point is necessarily $\bar{\mu}_.^{(S)}$, which concludes that $\bar{\mu}^{(S,N)}$ converges to $\bar{\mu}^{(S)}$ in $\mathcal{D}(\mathbb{R}_+, \mathcal{M}(\mathbb{X}))$. □

The above arguments directly entail the following pointwise in time convergence result on the potential force of infection.

Proposition 5.14. *The following convergence to 0 in conditional expectation holds for any $t > 0$:*

$$\lim_{N\to\infty} \mathbb{E}_0^N\left[\frac{1}{N}\sum_{i\in\mathcal{S}^N(0)} \left|\overline{\mathfrak{F}}_i^N(t) - \overline{\mathfrak{F}}(t, X_i^N)\right|\right] = 0. \tag{5.41}$$

Proof. As a consequence of Lemmas 5.4 , 5.6, 5.7, 5.8 and 5.10, we obtain

$$\mathbb{E}_0^N\left[\frac{1}{N}\sum_{i\in\mathcal{S}^N(0)} \left|\overline{\mathfrak{F}}_i^N(t) - \overline{\mathfrak{F}}(t, X_i^N)\right|\right] \leq C\cdot \overline{\mathfrak{D}}^N(t) + 3\overline{\mathfrak{V}}^N + \lambda^*\cdot\|\bar{\omega}^N - \bar{\omega}\|_\infty + \overline{\mathfrak{E}}^N(t)$$

where $C, \lambda^* < \infty$ while $\overline{\mathfrak{D}}^N(t)$, $\overline{\mathfrak{V}}^N$, $\|\bar{\omega}^N - \bar{\omega}\|_\infty$ and $\overline{\mathfrak{E}}^N(t)$ all tend to 0 thanks respectively to Proposition 5.11, Assumption 3.5, Assumption 3.3, and Lemma 5.10. This concludes the proof of Proposition 5.14. □

As a consequence of Propositions 5.11 and 5.14, by exploiting a similar approach as for the convergence to $\bar{\mu}_t^S$ in Lemma 5.1, we could typically prove the following pointwise in time convergence in probability in terms of any test function $\varphi \in C_b(\mathbb{X})$:

$$\sum_{i\leq N} \overline{\mathfrak{F}}_i^N(t)\, \mathbb{1}_{\{D_i^N(t)=0\}}\, \varphi(X_i^N) \to \langle \overline{\mathfrak{F}}(t)\bar{\mu}_t^S, \varphi\rangle\,.$$

The test function φ evaluates here the convergence of a distribution on $\mathbb{X}$ that we call the activated force of infection. The measure on the left-hand side can be interpreted as minus the derivative of the process $\bar{\mu}^S$ at time t.

5.2. **Convergence of** $(\bar{\mu}_\cdot^{I,N}, \bar{\mu}_\cdot^{R,N})$**.** Because the proof is simpler and more related to the one of Proposition 5.13 we first justify in the next proposition the convergence of the LLN-scaled recovered process $\bar{\mu}_t^{R,N}$, before we treat similarly in Proposition 5.16 the process $\bar{\mu}_t^{I,N}$.

Proposition 5.15. *As $N \to \infty$, the following convergence holds in probability*

$$\bar{\mu}^{R,N} \to \bar{\mu}^R \quad \text{in} \quad \mathcal{D}(\mathbb{R}_+, \mathcal{M}(\mathbb{X}))\,.$$

Proof. We distinguish three components depending on the initial condition of the individuals:

$$\bar{\mu}_t^{R,N} = \bar{\mu}_0^{R,N} + \bar{\mu}_t^{R,N,0} + \bar{\mu}_t^{R,N,1}\,, \tag{5.42}$$

where the measures $\bar{\mu}^{R,N,0}$ and $\bar{\mu}^{R,N,1}$ act as follows on test functions $\varphi \in C_b(\mathbb{R}_+)$ and time $t \geq 0$:

$$\langle \bar{\mu}_t^{R,N,0}, \varphi\rangle = \frac{1}{N}\sum_{j\in\mathcal{I}^N(0)} \mathbb{1}_{\{\eta_j^{N,0}\leq t\}}\varphi(X_j^N)\,, \tag{5.43}$$

and

$$\langle \bar{\mu}_t^{R,N,1}, \varphi\rangle = \frac{1}{N}\sum_{i\in\mathcal{D}^N(t)} \mathbb{1}_{\{\tau_i^N+\eta_i^N\leq t\}}\varphi(X_i^N)\,. \tag{5.44}$$

We recall that $\mathcal{D}^N(t)$ is defined in (2.10) as the subset in $\mathcal{S}_0^N$ of individuals infected by the disease by time t. The convergence for the first term $\langle \bar{\mu}_0^{R,N}, \varphi\rangle$ to $\langle \bar{\mu}_0^R, \varphi\rangle$ follows directly from Assumption 3.1.

Concerning $\bar{\mu}^{R,N,0}$, we will justify the convergence in probability through the computation of the expectations and variances, conditional on $\mathcal{F}_0^N$.

$$\mathbb{E}_0^N\left[\langle\bar{\mu}_t^{R,N,0},\varphi\rangle\right] = \int_0^\infty\int_{\mathbb{X}} \varphi(x)\left(1-\frac{F_x^c(a+t)}{F_x^c(a)}\right)\bar{\mu}_0^{I,N}(\mathrm{d}x,\mathrm{d}a) \tag{5.45}$$

Thanks to Assumption 3.1, the above conditional expectation converges in probability to

$$\langle\bar{\mu}_t^{R,0},\varphi\rangle = \int_0^\infty\int_{\mathbb{X}} \varphi(x)\left(1-\frac{F_x^c(a+t)}{F_x^c(a)}\right)\bar{\mu}_0^{I}(\mathrm{d}x,\mathrm{d}a)\,. \tag{5.46}$$

By the independence of the $\eta_j^{N,0}$ for $j\in\mathcal{I}^N(0)$ conditionally on $\mathcal{F}_0^N$, we obtain

$$\begin{aligned}\mathrm{Var}_0^N\left[\langle\bar{\mu}_t^{R,N,0},\varphi\rangle\right] &= \frac{1}{N}\int_0^\infty\int_{\mathbb{X}} \varphi(x)^2\cdot\frac{F_x^c(a+t)}{F_x^c(a)}\cdot\left(1-\frac{F_x^c(a+t)}{F_x^c(a)}\right)\bar{\mu}_0^{I,N}(\mathrm{d}x,\mathrm{d}a)\\ &\le \frac{\|\varphi\|_\infty^2}{N}\,.\end{aligned} \tag{5.47}$$

Finally we conclude the convergence in probability of $\bar{\mu}_t^{R,N,0}$ to $\bar{\mu}_t^{R,0}$ as defined in (5.46).

We then look at the mean-field approximation of $\bar{\mu}^{R,N,1}$ as defined in (5.44):

$$\langle\widetilde{\mu}_t^{R,N,1},\varphi\rangle = \frac{1}{N}\sum_{i\in\widetilde{\mathcal{D}}^N(t)} \mathbb{1}_{\{\widetilde{\tau}_i^N+\eta_i^N\le t\}}\varphi(X_i^N)\,. \tag{5.48}$$

We recall that $\widetilde{\mathcal{D}}^N(t)$ is defined in (5.8) as the subset of individuals infected according to $(\widetilde{\tau}_i^N)$ during the time-interval $(0,t]$.

$$\mathbb{E}_0^N\left[\langle\widetilde{\mu}_t^{R,N,1},\varphi\rangle\right] = \int_{\mathbb{X}}\varphi(x)\int_0^t \overline{\mathfrak{F}}(s,x)\cdot\exp\left[-\int_0^s\overline{\mathfrak{F}}(r,x)\mathrm{d}r\right]\cdot F_x(t-s)\mathrm{d}s\,\bar{\mu}_0^{S,N}(\mathrm{d}x)\,. \tag{5.49}$$

Thanks to Assumption 3.1 and (3.10), the above conditional expectation converges in probability to:

$$\begin{aligned}\langle\bar{\mu}_t^{R,1},\varphi\rangle &= \int_0^t\int_{\mathbb{X}}\varphi(x)\cdot F_x(t-s)\cdot\overline{\mathfrak{F}}(s,x)\cdot\exp\left[-\int_0^s\overline{\mathfrak{F}}(r,x)\mathrm{d}r\right]\bar{\mu}_0^S(\mathrm{d}x)\,\mathrm{d}s\\ &= \int_0^t\int_{\mathbb{X}}\varphi(x)\cdot F_x(t-s)\cdot\overline{\mathfrak{F}}(s,x)\bar{\mu}_s^S(\mathrm{d}x)\,\mathrm{d}s\,.\end{aligned} \tag{5.50}$$

By the independence of the η_i^N and of the $\widetilde{\tau}_i^N$ for $i\in\mathcal{S}^N(0)$ conditionally on $\mathcal{F}_0^N$, we obtain

$$\mathrm{Var}_0^N\left[\langle\widetilde{\mu}_t^{R,N,1},\varphi\rangle\right] \le \frac{\|\varphi\|_\infty^2}{N}\,. \tag{5.51}$$

Recalling (5.6), since the two events $\{\widetilde{\tau}_i^N+\eta_i^N\le t\}$ and $\{\tau_i^N+\eta_i^N\le t\}$ agree on the event $\{D_i^N(r)=\widetilde{D}_i^N(r),\forall r\in[0,t]\}$:

$$|\langle\widetilde{\mu}_t^{R,N,1}-\bar{\mu}_t^{R,N,1},\varphi\rangle| \le \|\varphi\|_\infty\cdot\overline{\mathfrak{D}}_t^N\,. \tag{5.52}$$

The right-hand side converges in probability to 0 as $N\to\infty$ thanks to Proposition 5.11. With (5.49), (5.50), (5.51) and (5.52) we deduce the convergence in probability of $\langle\widetilde{\mu}_t^{R,N,1},\varphi\rangle$ to $\langle\widetilde{\mu}_t^{R,1},\varphi\rangle$ as defined in (5.48). By recalling (5.42), the convergence of $\bar{\mu}_t^{R,N,0}$ and (3.9), we conclude the convergence in probability of $\langle\bar{\mu}_t^{R,N},\varphi\rangle$ to $\langle\bar{\mu}_t^{R},\varphi\rangle$.

For non-negative φ, the function : $t \mapsto \langle \bar{\mu}_t^{R,N}, \varphi \rangle$ is non-decreasing. We can thus easily adapt the argument given in Proposition 5.13, which notably involves the tightness criteria given in Proposition 5.12 with $\nu^N = \bar{\mu}^{R,N}$, still $\mathcal{X} = \mathbb{X}$ and $\mathfrak{M} = \mathcal{C}_{b+}(\mathbb{X})$ so as to apply the Second Dini theorem. For any t and any compact set K, $\bar{\mu}_t^{R,N}(K^c) \le \bar{\mu}_X^N(K^c)$ as defined in (3.3). Since the latter converges in probability to $\bar{\mu}_X$, recalling (3.3), we can find for any $\varepsilon > 0$ some compact set K_ε such that

$$\sup_{N \ge 1} \mathbb{P}\Big(\bar{\mu}_X^N(K_\varepsilon^c) > \varepsilon \Big) < \varepsilon.$$

Point (a) in Proposition 5.12 is therefore verified as well. So we deduce that the convergence in probability extends to the function $\bar{\mu}^{R,N}$ in $\mathcal{D}_1$, which concludes the proof of Proposition 5.15. □

Proposition 5.16. *As $N \to \infty$, the following convergence holds in probability*

$$\bar{\mu}^{I,N} \to \bar{\mu}^I \quad in \quad \mathcal{D}(\mathbb{R}_+, \mathcal{M}(\mathbb{X} \times \mathbb{R}_+)) .$$

Proof. Although the proof is more technical than the one of Proposition 5.15, we exploit similar arguments. In order to exploit the Second Dini theorem, we wish to consider non-increasing projections. This is why we will study the following extended measure $\bar{\mu}_t^{SI,N}$ for test functions ψ in the set $\mathfrak{M}(\mathbb{X} \times [-1, \infty))$ of functions from $\mathbb{X} \times [-1, \infty)$ to $\mathbb{R}_+$ that are continuous, bounded, non-negative and non-increasing in the second variable:

$$\begin{aligned}\langle \bar{\mu}_t^{SI,N}, \psi \rangle &:= \langle \bar{\mu}_t^{S,N}, \psi(\,.\,, -1) \rangle + \langle \bar{\mu}_t^{I,N}, \psi_{|\mathbb{X} \times [0,\infty)} \rangle \\ &= \frac{1}{N} \sum_{i \in \mathcal{S}^N(0)} \Big(\mathbb{1}_{\{D_i^N(t) = 0\}} \psi(X_i^N, -1) + \mathbb{1}_{\{D_i^N(t)=1\}} \, \mathbb{1}_{\{\eta_i^N > t - \tau_i^N\}} \psi(X_i^N, t - \tau_i^N) \Big) \\ &\quad + \frac{1}{N} \sum_{j \in \mathcal{I}^N(0)} \mathbb{1}_{\{\eta_i^{N,0} > t\}} \psi(X_j^N, A_j^N(0) + t) .\end{aligned}$$

In words, $\bar{\mu}^{SI,N}$ is derived from the addition of both $\bar{\mu}^{S,N}$ and $\bar{\mu}^{I,N}$ where the first measure on $\mathbb{X}$ is projected with a fixed component -1 according to the age variable. Intuitively, what we are doing is considering susceptible individuals as infected with infection age -1. The fact that ψ is non-negative implies that any recovery event leads to a reduction of $\langle \bar{\mu}^{SI,N}, \psi \rangle$ at this particular time. The fact that ψ is non-increasing in the second variable implies that the aging of the actively infected leads as well to a reduction of $\langle \bar{\mu}^{SI,N}, \psi \rangle$ over time. With this trick of combining $\bar{\mu}^{S,N}$ to $\bar{\mu}^{I,N}$ into $\bar{\mu}^{SI,N}$, any infection event leads as well to a reduction of $\langle \bar{\mu}^{SI,N}, \psi \rangle$ at the infection time.

We will make use of Proposition 5.12 in combination with the following lemma by considering for $\nu^N = \bar{\mu}^{SI,N}$ the set $\mathcal{X} = \mathbb{X} \times [-1, \infty)$ and the proposed set $\mathfrak{M}(\mathbb{X} \times [-1, \infty))$ as the separating class.

Lemma 5.17. *The set $\mathfrak{M}(\mathbb{X} \times [-1, \infty))$ of functions that are continuous, bounded, non-negative and non-increasing in the second variable is a separating class.*

Proof. The fact that the constant function equal to 1 is part of $\mathfrak{M}(\mathbb{X} \times [-1, \infty))$ comes readily from the definition. If ν, ν' are such that $\langle \nu, \phi \rangle = \langle \nu', \phi \rangle$ for any $\phi \in \mathfrak{M}$, then the classical approximation scheme of indicator functions by bounded continuous functions leads to the identity $\nu(A \times [-1, a]) = \nu'(A \times [-1, a])$, for any $a \in [-1, \infty)$ and measurable subset A of $\mathbb{X}$. The sets of this form $A \times [-1, a]$ form a π-system of subsets of the product space

$\mathbb{X}\times[-1,\infty)$ that contains $\mathbb{X}\times[-1,\infty)$ itself and generates the Borel σ-field of $\mathbb{X}\times[-1,\infty)$. The identity $\nu=\nu'$ is thus deduced thanks e.g. to [34, Lemma 1.17]. This concludes the proof of Lemma 5.17. □

Then, it mainly remains to adapt the computations of conditional expectations and variances from the proof of Proposition 5.15. $\bar{\mu}^{I,N}$ is similarly decomposed into the sum of $\bar{\mu}^{I,N,0}$ and $\bar{\mu}^{I,N,1}$, that are represented as follows for any $\psi\in C_b(\mathbb{X}\times\mathbb{R}_+)$ and $t\geq 0$:

$$\langle\bar{\mu}^{I,N,0}_t,\psi\rangle=\frac{1}{N}\sum_{j\in\mathcal{I}^N(0)}\mathbb{1}_{\{\eta_j^{N,0}>t\}}\psi(X_j^N,A_j^N(0)+t)\,,\tag{5.53}$$

and

$$\langle\bar{\mu}^{I,N,1}_t,\psi\rangle=\frac{1}{N}\sum_{i\in\mathcal{D}^N(t)}\mathbb{1}_{\{\tau_i^N+\eta_i^N>t\}}\psi(X_i^N,t-\tau_i^N)\,.\tag{5.54}$$

Concerning $\bar{\mu}^{I,N,0}$, we have

$$\mathbb{E}_0^N\Big[\langle\bar{\mu}^{I,N,0}_t,\psi\rangle\Big]=\int_0^\infty\!\!\int_{\mathbb{X}}\psi(x,a+t)\cdot\frac{F_x^c(a+t)}{F_x^c(a)}\bar{\mu}_0^{I,N}(\mathrm{d}x,\mathrm{d}a)$$

Thanks to Assumption 3.1, the above conditional expectation converges in probability to:

$$\langle\bar{\mu}^{I,0}_t,\psi\rangle=\int_0^\infty\!\!\int_{\mathbb{X}}\psi(x,a+t)\cdot\frac{F_x^c(a+t)}{F_x^c(a)}\,\bar{\mu}_0^{I}(\mathrm{d}x,\mathrm{d}a)\,.\tag{5.55}$$

By the independence of the $\eta_j^{N,0}$ for $j\in\mathcal{I}^N(0)$ conditionally on $\mathcal{F}_0^N$, we obtain

$$\begin{aligned}\mathrm{Var}_0^N\Big[\langle\bar{\mu}^{I,N,0}_t,\psi\rangle\Big]&=\frac{1}{N}\int_0^\infty\!\!\int_{\mathbb{X}}\psi(x,a+t)^2\cdot\frac{F_x^c(a+t)}{F_x^c(a)}\cdot\left(1-\frac{F_x^c(a+t)}{F_x^c(a)}\right)\bar{\mu}_0^{I,N}(\mathrm{d}x,\mathrm{d}a)\\&\leq\frac{\|\psi\|_\infty^2}{N}\,.\end{aligned}$$

So we conclude to the convergence in probability of $\langle\bar{\mu}^{I,N,0}_t,\psi\rangle$ to $\langle\bar{\mu}^{I,N,0}_t,\psi\rangle$ as defined in (5.55), valid for any $t>0$ and any $\psi\in\mathcal{C}_b(\mathbb{X}\times\mathbb{R}_+)$.

We then consider the mean-field approximation of $\bar{\mu}^{I,N,1}$ as defined in (5.54), exploiting the notations $\widetilde{\tau}_i^N$ and $\widetilde{\mathcal{D}}^N(t)$ from (5.8):

$$\langle\widetilde{\mu}^{I,N,1}_t,\psi\rangle=\frac{1}{N}\sum_{i\in\widetilde{\mathcal{D}}^N(t)}\mathbb{1}_{\{\widetilde{\tau}_i^N+\eta_i^N>t\}}\psi(X_i^N,t-\widetilde{\tau}_i^N)\,.\tag{5.56}$$

$$\mathbb{E}_0^N\Big[\langle\widetilde{\mu}^{I,N,1}_t,\psi\rangle\Big]=\int_{\mathbb{X}}\int_0^t\overline{\mathfrak{F}}(s,x)\cdot\exp\Big[-\int_0^s\overline{\mathfrak{F}}(r,x)\mathrm{d}r\Big]\cdot F_x(t-s)\cdot\psi(x,t-s)\mathrm{d}s\,\bar{\mu}_0^{S,N}(\mathrm{d}x)\,.\tag{5.57}$$

Thanks to Assumption 3.1 and (3.10), the above conditional expectation converges in probability to

$$\begin{aligned}\langle\bar{\mu}^{I,1}_t,\psi\rangle&=\int_{\mathbb{X}}\!\!\int_0^t\psi(x,t-s)\cdot F_x(t-s)\cdot\overline{\mathfrak{F}}(s,x)\cdot\exp\Big[-\int_0^s\overline{\mathfrak{F}}(r,x)\mathrm{d}r\Big]\mathrm{d}s\,\bar{\mu}_0^S(\mathrm{d}x)\\&=\int_0^t\!\!\int_{\mathbb{X}}\psi(x,t-s)\cdot F_x(t-s)\cdot\overline{\mathfrak{F}}(s,x)\bar{\mu}_s^S(\mathrm{d}x)\,\mathrm{d}s\,.\end{aligned}\tag{5.58}$$

By the independence of the η_i^N and of the $\widetilde{\tau}_i^N$ for $i \in \mathcal{S}^N(0)$ conditionally on $\mathcal{F}_0^N$, we obtain

$$\mathrm{Var}_0^N\Big[\langle \widetilde{\mu}_t^{I,N,1}, \psi\rangle\Big] \leq \frac{\|\psi\|_\infty^2}{N}. \tag{5.59}$$

Recalling (5.6), since the two events $\{\widetilde{\tau}_i^N + \eta_i^N > t\}$ and $\{\tau_i^N + \eta_i^N > t\}$ agree on the event $\{D_i^N(r) = \widetilde{D}_i^N(r), \forall t \in [0,t]\}$, we get

$$|\langle \widetilde{\mu}_t^{I,N,1} - \bar{\mu}_t^{I,N,1}, \psi\rangle| \leq \|\psi\|_\infty \cdot \overline{\mathfrak{D}}_t^N. \tag{5.60}$$

The right-hand side converges in probability to 0 as N tends to infinity thanks to Proposition 4.1. With (5.57), (5.58), (5.59) and (5.60) we deduce the convergence in probability of $\langle \widetilde{\mu}_t^{I,N,1}, \psi\rangle$ to $\langle \widetilde{\mu}_t^{I,1}, \psi\rangle$ as defined in (5.58).

This concludes the proof that $\langle \bar{\mu}_t^{I,N}, \psi\rangle$ converges in probability to $\langle \bar{\mu}_t^{I}, \psi\rangle$, for any $t > 0$ and any $\psi \in \mathcal{C}_b(\mathbb{X} \times \mathbb{R}_+)$. Recalling Proposition 5.13, we deduce specifically that $\langle \bar{\mu}_t^{SI,N}, \psi\rangle$ converges in probability to $\langle \bar{\mu}_t^{SI}, \psi\rangle$, for any $t > 0$ and any $\psi \in \mathfrak{M}$, where

$$\langle \bar{\mu}_t^{SI}, \psi\rangle := \langle \bar{\mu}_t^{S}, \psi(\,.\,, -1)\rangle + \langle \bar{\mu}_t^{I}, \psi_{|\mathbb{X}\times[0,\infty)}\rangle. \tag{5.61}$$

Note about the above definition of $\bar{\mu}_t^{SI}$ through test functions $\psi \in \mathfrak{M}$ that it already uniquely specifies $\widetilde{\mu}_t^{SI}$ due to $\mathfrak{M}$ being a separating class. The extension of this definition to any $\psi \in \mathcal{C}_b(\mathbb{X} \times [-1, \infty))$ is yet very natural.

With the crucial arguments given at the beginning of this proof of Proposition 5.16, recall that $\langle \bar{\mu}_t^{SI,N}, \psi\rangle$ is for any $\psi \in \mathfrak{M}$ non-increasing as a function of t. On the other hand, $\langle \bar{\mu}_t^{SI}, \psi\rangle$ is deterministic, continuous and non-increasing as a function of t. We can thus adapt the argument given in Proposition 5.13 to show that the convergence in probability of $\langle \bar{\mu}_t^{SI,N}, \psi\rangle$ to $\langle \bar{\mu}_t^{SI}, \psi\rangle$ is locally uniform in t. This concludes Point (b) in Proposition 5.12. Concerning Point (a), we remark for any compact set K in $\mathbb{X}$ and any $A > 0$ that

$$\bar{\mu}_t^{SI,N}\big[(K \times [-1, A+t])^c\big] \leq \bar{\mu}_0^{S,N}\big[K^c\big] + \bar{\mu}_0^{I,N}\big[(K \times [0,A])^c\big]. \tag{5.62}$$

Since $\bar{\mu}_0^{S,N}$ and $\bar{\mu}_0^{I,N}$ converge in probability to respectively $\bar{\mu}_0^{S}$ and $\bar{\mu}_0^{I}$, there exists for any $\varepsilon > 0$ such a compact set K in $\mathbb{X}$ and $A > 0$ that satisfy

$$\sup_N \mathbb{P}(\bar{\mu}_0^{S,N}\big[K^c\big] + \bar{\mu}_0^{I,N}\big[(K \times [0,A])^c\big] > \varepsilon) < \varepsilon.$$

Recalling (5.62), this entails Point (a), for any $T > 0$ with $K^{(\mathcal{X})} = K \times [-1, A+T]$. We then exploit Proposition 5.12 and conclude the proof of Proposition 5.16 that $\bar{\mu}^{(I,N)}$, as the restriction of $\bar{\mu}^{(SI,N)}$ to $\mathbb{X}\times\mathbb{R}_+$, converges in probability to $\bar{\mu}^{(I,N)}$ in $\mathcal{D}(\mathbb{R}_+, \mathcal{M}(\mathbb{X}\times\mathbb{R}_+))$. □

Appendix A. Technical supporting results

In Appendix A, we prove three technical results that are exploited in the current paper. Lemma A.1 is used to deduce local uniform convergence in probability from pointwise estimates. Proposition A.2 is used to deduce the a.e. continuity of sections of a.e. continuous kernels. The more technical proof of proposition 5.9 is given afterwards.

A.1. **From pointwise to locally uniform convergence in probability.**
In Lemma A.1, we state that the second Dini theorem extends to convergences in probability of random functions.

Lemma A.1. *Let ψ be a possibly random non-decreasing and continuous function from $\mathbb{R}_+$ to $\mathbb{R}$. Let also ψ^n be a possibly random sequence of non-decreasing functions from $\mathbb{R}_+$ to $\mathbb{R}$ that converges pointwise in probability to ψ. Then, ψ^n converges in probability to ψ locally uniformly.*

Proof. We exploit the relation between the convergence in probability and the a.s. convergence along sequence extractions as stated in [34, Lemma 4.2].

Let $(N[k])_{k\geq 1} \in \mathbb{N}^{\mathbb{N}}$ be an increasing sequence and for any T, let $(t_i)_{i\geq 1}$ be a countable dense subset of $[0,T]$. By a triangular argument, we can then define an extraction $(\widetilde{N}_n)_{n\geq 1} = (N[k_n])_{n\geq 1} \in \mathbb{N}^{\mathbb{N}}$, with the sequence $(k_n)_{n\geq 1}$ being increasing, such that for any $i \geq 1$, $\psi^{\widetilde{N}_n}(t_i)$ converges a.s. to $\psi(t_i)$ as n tends to infinity. Thanks to the second Dini theorem (see Exercise 127 on page 81, and its solution on page 270 in Polya and Szegö [47]), this entails the a.s. convergence of $\psi^{\widetilde{N}_n}(t)$ to $\psi(t)$ uniformly in $t \in [0,T]$. Since the sequence $(\widetilde{N}_n)$ is an extraction of any initial subsequence and T can be freely chosen, this concludes thanks to [34, Lemma 4.2] that $\psi^{\widetilde{N}_n}(t)$ converges in probability to $\psi(t)$ locally uniformly in t. □

A.2. **Almost everywhere continuity.**

Proposition A.2. *Assume that $\bar{\omega}$ is $\bar{\mu}_X^{\otimes 2}$-a.e. continuous. Then the function $: x \mapsto \bar{\omega}(x,.)$ is $\bar{\mu}_X$ a.e. continuous from $\mathbb{X}$ with values in $L^1(\bar{\mu}_X)$.*

Proof. Let us define the following function of $x \in \mathbb{X}$ and $\eta > 0$:

$$\Psi(x;\eta) = \sup\Big\{ \int_{\mathbb{X}} |\bar{\omega}(x_1,x') - \bar{\omega}(x_2,x')|\, \bar{\mu}_X(\mathrm{d}x');\ x_1,x_2 \in B(x;\eta)\Big\},$$

where $B[x;\eta]$ denotes the open ball centered in x of radius η. It follows from (4.6) that the irregularities of $\overline{\mathfrak{F}}$ in x will be located through the following subsets $\Psi[\delta,\eta]$ of $\mathbb{X}$, defined for any $\delta,\eta > 0$:

$$\Psi[\delta,\eta] := \{x \in \mathbb{X};\ \Psi(x;\eta) > \delta\}. \tag{A.1}$$

On the other hand, the $\bar{\mu}_X$-a.e. continuity of the kernel $\bar{\omega}$ is related to the two following similar definitions. First the function Φ is defined for $(x,x') \in \mathbb{X}^2$ and $\eta > 0$:

$$\Phi((x,x');\eta) = \sup\Big\{|\bar{\omega}(x_1,x_1') - \bar{\omega}(x_2,x_2')|;\ (x_1,x_1'),(x_2,x_2') \in B((x,x');\eta)\Big\},$$

with $B[(x,x');\eta]$ the open ball centered in (x,x') of radius η (with the supremum norm between the two components). Then, the sets $\Phi[\delta,\eta]$ are defined for $\delta,\eta > 0$ as:

$$\Phi[\delta,\eta] := \{(x,x') \in \mathbb{X}^2;\ \Phi((x,x');\eta) > \delta\}. \tag{A.2}$$

The measurability of the sets $\Phi[\delta,\eta]$ and $\Psi[\delta,\eta]$, for any $\delta,\eta > 0$, is proven for completeness in Lemma A.3 just after we conclude the proof of Proposition A.2.

Note that $\cup_{\delta>0} \cap_{\eta>0} \Phi[\delta,\eta]$ is exactly the set of discontinuity points in $\mathbb{X}^2$ of the kernel $\bar{\omega}$. $\cup_{\delta>0} \cap_{\eta>0} \Psi[\delta,\eta]$ similarly covers the set of discontinuity points in $\mathbb{X}$ of the function $: x \mapsto \int \bar{\omega}(x,x')\, \bar{\mu}_X(\mathrm{d}x')$. Since $\Phi[\delta,\eta]$ is non-increasing in η and increasing in δ, the fact that $\bar{\omega}$ is $\bar{\mu}_X^{\otimes 2}$ a.e. continuous translates into the following convergence property to 0 for any $\delta > 0$:

$$\lim_{\eta \to 0} \bar{\mu}_X^{\otimes 2}(\Phi[\delta, \eta]) = 0. \tag{A.3}$$

We will also consider the sections $\Phi_x[\delta, \eta]$ of $\Phi[\delta, \eta]$ with the fixed first component x, i.e.,

$$\Phi_x[\delta, \eta] := \{x' \in \mathbb{X};\ (x, x') \in \Phi[\delta, \eta]\}. \tag{A.4}$$

Then, by exploiting ω^* as the uniform upper-bound of the non-negative kernel $\bar{\omega}$ and recalling (A.2), we deduce the following inequality for any $x \in \mathbb{X}$, $\delta, \eta > 0$,

$$\varPsi(x; \eta) \leq \omega^* \cdot \bar{\mu}_X(\Phi_x[\delta, \eta]) + \delta.$$

As a consequence, for any $x \in \Psi[2\delta, \eta]$, it holds that $\bar{\mu}_X(\Phi_x[\delta, \eta]) \geq \delta/\omega^*$. Since $\bar{\mu}_X$ is a probability measure, it entails the following for any $\delta, \eta > 0$,

$$\bar{\mu}_X^{\otimes 2}(\Phi[\delta, \eta]) \geq \frac{\delta}{\omega^*} \bar{\mu}_X(\Psi[2\delta, \eta]).$$

Recalling from (A.3) that the l.h.s. tends to zero as η tends to 0, so does the sequence $\bar{\mu}_X(\Psi[\delta, \eta])$ for any δ. Up to the next Lemma A.3, this concludes the proof of Proposition A.2. □

Lemma A.3. *For any $\delta, \eta > 0$, the sets $\Phi[\delta, \eta]$ and $\Psi[\delta, \eta]$ as defined respectively in (A.2) and (A.1) are measurable.*

Proof. The measurability of the set $\Phi[\delta, \eta]$ can be verified through the following definition, which happens to be equivalent to the one given in (A.2):

$$\Phi[\delta, \eta] = \cup_{\{q \in \mathbf{Q}_+, m \geq 1\}} \bar{\omega}^{-1}([0, q])^\eta \cap \bar{\omega}^{-1}([q + \delta + 2^{-m}, \omega^*])^\eta,$$

where A^η is defined as follows for any Borel subset A of $\mathbb{X}^2$ and any $\eta > 0$:

$$A^\eta := \{(x, x') \in \mathbb{X}^2;\ B[(x, x'); \eta] \cap A \neq \emptyset\}, \tag{A.5}$$

so that A^η denotes the η-vicinity of A.

The measurability of the sets $\Psi[\delta, \eta]$ is a bit more technical than the one of $\Phi[\delta, \eta]$, yet follows similar ideas. We exploit here the density in $L^1(\bar{\mu}_X)$ of a countable subset $\mathfrak{Q}$, according to [16, Proposition 3.4.5]. Let $\varepsilon_m = 2^{-m}$ and $\mathcal{W} : x \in \mathbb{X} \mapsto \omega(x, .)$. Then, we claim that the following definition is equivalent to the one of $\Psi[\delta, \eta]$ in (A.1):

$$\overline{\Psi}[\delta, \eta] = \cup_{\{m \geq 1\}} \cup_{\{(\omega_1, \omega_2) \in \mathfrak{Q}^2 \cap \mathfrak{D}_{\delta, m}\}} \mathcal{W}^{-1}\Big(B_{L^1(\bar{\mu}_X)}(\omega_1, \varepsilon_m)\Big)^\eta \cap \mathcal{W}^{-1}\Big(B_{L^1(\bar{\mu}_X)}(\omega_2, \varepsilon_m)\Big)^\eta,$$

where the couple $(\omega_1, \omega_2) \in \mathfrak{Q}^2$ belong to $\mathfrak{D}_{\delta, m}$ provided that

$$\int_{\mathbb{X}} |\omega_1 - \omega_2|\, \bar{\mu}_X \geq \delta + 2\varepsilon_m. \tag{A.6}$$

It is clear from this definition that $\overline{\Psi}[\delta, \eta]$ is measurable. We next prove the equality with $\Psi[\delta, \eta]$. If $x \in \overline{\Psi}[\delta, \eta]$, then it means that there exists $m \geq 1$, $(\omega_1, \omega_2) \in \mathfrak{Q}^2 \cap \mathfrak{D}_{\delta, m}$ and $(x_1, x_2) \in B(x, \eta)^2$ such that

$$\int_{\mathbb{X}} |\omega(x_1, .) - \omega_1|\, \bar{\mu}_X < \varepsilon_m, \quad \int_{\mathbb{X}} |\omega(x_2, .) - \omega_2|\, \bar{\mu}_X < \varepsilon_m.$$

This entails the following property from which $x \in \Psi[\delta, \eta]$ is a direct consequence:

$$\int_{\mathbb{X}} |\omega(x_1, .) - \omega(x_2, .)|\, \bar{\mu}_X > \delta. \tag{A.7}$$

Reciprocally, if we assume that there exist a couple $(x_1, x_2) \in B(x,\eta)^2$ which satisfies (A.7), then we can first choose $m \geq 1$ such that

$$\int_{\mathbb{X}} |\omega(x_1,.) - \omega(x_2,.)|\, \bar{\mu}_X \geq \delta + 4\varepsilon_m.$$

We can next identify two elements $(\omega_1, \omega_2) \in \mathfrak{Q}^2$ in respectively $B_{L^1(\bar{\mu}_X)}(\omega(x_1,.), \varepsilon_m)$ and $B_{L^1(\bar{\mu}_X)}(\omega(x_2,.), \varepsilon_m)$. By construction, $(\omega_1,\omega_2) \in \mathfrak{D}_{\delta,m}$, $x_1 \in \mathcal{W}^{-1}\Big(B_{L^1(\bar{\mu}_X)}(\omega_1, \varepsilon_m)\Big)$ and similarly for x_2 (with ω_2 instead of ω_1). This entails that $x \in \Psi[\delta,\eta]$. From this we have concluded that $\Psi[\delta,\eta] = \overline{\Psi}[\delta,\eta]$, so that the former is indeed measurable. □

A.3. **Proof of Proposition 5.9.** We first consider (μ^N) and (ν^N) as deterministic sequences, and then extend the result to random sequences in the last fifth step of the proof.

Proof. By linearity of the above quantity in the function k and in the pair (ν^N, ν), we may assume without loss of generality that k is non-negative and bounded by 1, while $\nu^N(\mathcal{Y}) \leq 1$ and $\nu(\mathcal{Y}) \leq 1$. Let us define the integrand in x as $\varepsilon^N(x)$:

$$\varepsilon^N(x) = \Big| \int_{\mathcal{Y}} k(x,y)[\nu^N - \nu](dy) \Big| . \tag{A.8}$$

In the degenerate case where $\nu \equiv 0$, the convergence of $\langle \mu^N, \varepsilon^N \rangle$ to 0 can be directly deduced with the upper-bound of ε^N by $\|k\|_\infty \cdot \nu^N(\mathcal{Y})$ which converges to 0. In the degenerate case where $\mu \equiv 0$, it suffices to take $3\|k\|_\infty \cdot \nu(\mathcal{Y})$ as the uniform upper-bound of ε^N for any N sufficiently large, as $\mu^N(\mathcal{X})$ then tends to 0. In the following, we can thus assume that both $\nu(\mathcal{Y}) > 0$ and $\mu(\mathcal{X}) > 0$.

The irregularities of k will be located through the following subset $\mathcal{G}[\delta,\eta]$ of $\mathcal{X} \times \mathcal{Y}$, defined for any $\delta, \eta > 0$ similarly as in (A.2):

$$\mathcal{G}[\delta,\eta] := \{(x,y) \in \mathcal{X} \times \mathcal{Y};\ Diam_k(B[(x,y);\eta]) > \delta\}, \tag{A.9}$$

where the diameter function $Diam_k$ corresponding to the kernel k is defined as follows for any measurable subset A of $\mathcal{X} \times \mathcal{Y}$:

$$Diam_k(A) := \sup\{|k(z) - k(z')|; z, z' \in A\}, \tag{A.10}$$

while $B[(x,y);\eta]$ denotes the open ball centered in (x,y) of radius η. The size η of the vicinities in (A.9) shall be considered sufficiently small to ensure that discrepancies of order δ are exceptional. The measurability of $\mathcal{G}[\delta,\eta]$, for any $\delta,\eta > 0$, is proved as in Lemma A.3.

Step 1: Convergence of $\langle \mu, \varepsilon^N \rangle$ to zero.
Since k is continuous $\mu \otimes \nu$ almost everywhere, in particular, $k(x,.)$ is continuous ν almost everywhere for x on a measurable set $\mathcal{A} \subset \mathcal{X}$ such that $\mu(\mathcal{A}) = 1$. For any $x \in \mathcal{A}$, thanks to the Portmanteau theorem, see e.g. [30], Subsection IV.3a on the "Weak Convergence of Probability Measures", $\varepsilon^N(x)$ converges to 0. Remark as compared to the classical version of Portmanteau theorem that we allow ν^N and ν to be general non-negative finite measure rather than probability measures, given that the proof is not difficult to adapt for this setting. As a consequence of Lebesgue's dominated convergence theorem, recalling that ε^N is bounded (by 1 under our assumption), we deduce

$$\lim_{N\to\infty} \langle \mu, \varepsilon^N \rangle = 0. \tag{A.11}$$

Step 2: Convergence of $[\mu\otimes\nu](\mathcal{G}[\delta,\eta])$ and $[\mu\otimes\nu^N](\mathcal{G}[\delta,\eta])$ to zero.
Remark that the points of discontinuity of the kernel k are identified as follows in terms of the sets $\mathcal{G}[\delta,\eta]$:

$$\cup_{n\geq 1}\cap_{m\geq 1}\mathcal{G}[2^{-n},2^{-m}].$$

Note also that the sets $\mathcal{G}[\delta,\eta]$ are increasing as δ decreases and non-increasing as η decreases. Therefore, due to the fact that k is $\mu\otimes\nu$ almost everywhere continuous, the following convergence to zero holds for any δ:

$$\lim_{\eta\to 0}[\mu\otimes\nu](\mathcal{G}[\delta,\eta])=0. \tag{A.12}$$

Secondly, we remark that the weak convergence of ν^N to ν implies the weak convergence of $\mu\otimes\nu^N$ to $\mu\otimes\nu$. For any N sufficiently large, thanks to the Portmanteau theorem:

$$[\mu\otimes\nu^N](\mathcal{G}[\delta,\eta])\leq[\mu\otimes\nu](\mathcal{G}[\delta,\eta]^\eta)+\eta, \tag{A.13}$$

where we recall the notation A^η from (A.5). Since $B[(x,y);\eta]\subset B[(x',y');2\eta]$ holds true for any $(x',y')\in B[(x,y);\eta]$, it is a straightforward consequence of definition (A.9) that $\mathcal{G}[\delta,\eta]^\eta\subset\mathcal{G}[\delta,2\eta]$. Recalling (A.12) and coming back to (A.13), we have proved the following convergence to zero for any δ:

$$\lim_{\eta\to 0}[\mu\otimes\nu^N](\mathcal{G}[\delta,\eta])=0. \tag{A.14}$$

Step 3: Relation between the level sets of ε^N to $\mathcal{G}[\delta,\eta]$.
We consider for any value $\delta>0$ the corresponding level-set of ε^N:

$$\mathcal{H}^N[\delta]:=\left\{x\in\mathcal{X};\,\varepsilon^N(x)\geq\delta\right\}. \tag{A.15}$$

Let $\theta:=2+2\nu(\mathcal{Y})>0$. For any η sufficiently small, we will relate in the following lemma the intersection $\mathcal{H}^N[\delta]^c\cap\mathcal{H}^N[\theta\,\delta]^\eta$ to conditions on the following subsets of $\mathcal{Y}$:

$$\mathcal{G}_x[\delta,\eta]:=\{y\in\mathcal{Y};\,(x,y)\in\mathcal{G}[\delta,\eta]\}, \tag{A.16}$$

namely the restriction of $\mathcal{G}[\delta,\eta]$, recall (A.9), with x as the first coordinate.

Lemma A.4. *The following inclusion holds for any $\delta,\eta>0$ and $N\geq 1$:*

$$\mathcal{H}^N[\delta]^c\cap\mathcal{H}^N[\theta\,\delta]^\eta\subset\Big\{x\in\mathcal{X};\,\nu^N(\mathcal{G}_x[\delta,\eta])\geq\frac{\delta}{2}\Big\}\cup\Big\{x\in\mathcal{X};\,\nu(\mathcal{G}_x[\delta,\eta])\geq\frac{\delta}{2}\Big\}.$$

Proof. Let us consider any $x\in\mathcal{H}^N[\delta]^c\cap\mathcal{H}^N[\theta\,\delta]^\eta$. We can thus choose some $x'\in\mathcal{H}^N[\theta\,\delta]\cap B(x,\eta)$. By virtue of (A.16), for N sufficiently large, we obtain

$$\begin{aligned}\Big|\int_{\mathcal{G}_x[\delta,\eta]^c}|k(x,y)-k(x',y)|\,[\nu^N-\nu](dy)\Big|&\leq[\nu^N\vee\nu](\mathcal{G}_x[\delta,\eta]^c)\cdot\delta\\&\leq 2\delta\cdot\nu(\mathcal{Y})\,,\end{aligned} \tag{A.17}$$

where we have exploited that k, ν^N and ν are non-negative and that $\nu^N(\mathcal{Y})$ converges to $\nu(\mathcal{Y})>0$. On the other hand,

$$\Big|\int_{\mathcal{G}_x[\delta,\eta]}|k(x,y)-k(x',y)|\,[\nu^N-\nu](dy)\Big|\leq[\nu^N\vee\nu](\mathcal{G}_x[\delta,\eta]), \tag{A.18}$$

where we recall the assumption that k is non-negative and bounded by 1. Since $x\in\mathcal{H}^N[\delta]^c$ while $x'\in\mathcal{H}^N[\theta\,\delta]$:

$$|\varepsilon^N(x)-\varepsilon^N(x')|\geq\varepsilon^N(x')-\varepsilon^N(x)\geq\delta\cdot(1+2\nu(\mathcal{Y})).$$

Recalling (A.8) to combine this result with (A.17) and (A.18), we deduce:

$$[\nu \vee \nu^N](\mathcal{G}_x[\delta, \eta]) \geq \delta.$$

This inequality implies either $\nu^N(\mathcal{G}_x[\delta, \eta]) \geq \delta/2$ or $\nu(\mathcal{G}_x[\delta, \eta]) \geq \delta/2$. This concludes the proof of Lemma A.4. □

Step 4: Proof of Proposition 5.9 in the particular case where (μ^N) and (ν^N) are deterministic.
For any $\delta > 0$ and $N \geq 1$ sufficiently large, since the mass of μ^N converges to the one of μ, we obtain

$$\langle \mu^N, \varepsilon^N \rangle \leq 2\theta\, \delta \cdot \mu(\mathcal{X}) + \mu^N(\mathcal{H}^N[\theta\, \delta]), \tag{A.19}$$

where we recall (A.15).

We choose $\eta \in (0, \delta)$ sufficiently small thanks to Step 2, to ensure both that $[\mu \otimes \nu](\mathcal{G}[\delta, \eta])$ is smaller than $\delta^2/2$ and similarly for $[\mu \otimes \nu^N](\mathcal{G}[\delta, \eta])$ for any N sufficiently large. Thanks to the Markov inequality, we have

$$\mu\Big(\Big\{x \in \mathcal{X};\, \nu(\mathcal{G}_x[\delta, \eta]) \geq \frac{\delta}{2}\Big\}\Big) \leq \frac{2[\mu \otimes \nu](\mathcal{G}[\delta, \eta])}{\delta} \leq \delta. \tag{A.20}$$

Similarly,

$$\mu\Big(\Big\{x \in \mathcal{X};\, \nu^N(\mathcal{G}_x[\delta, \eta]) \geq \frac{\delta}{2}\Big\}\Big) \leq \frac{2[\mu \otimes \nu^N](\mathcal{G}[\delta, \eta])}{\delta} \leq \delta. \tag{A.21}$$

Since μ^N converges weakly to μ, for any N sufficiently large we have

$$\mu^N(\mathcal{H}^N[\theta\, \delta]) \leq \mu(\mathcal{H}^N[\theta\, \delta]^\eta) + \eta, \tag{A.22}$$

where we adapt the definition of η vicinity given in (A.5) to subsets of $\mathcal{X}$. As we expect $\mathcal{H}^N[\theta\, \delta]^\eta$ to be mostly comprised into $\mathcal{H}^N[\delta]$, we make the following distinction

$$\mu(\mathcal{H}^N[\theta\, \delta]^\eta) \leq \mu(\mathcal{H}^N[\delta]) + \mu(\mathcal{H}^N[\delta]^c \cap \mathcal{H}^N[\theta\, \delta]^\eta). \tag{A.23}$$

Thanks to the Markov inequality, we obtain

$$\mu(\mathcal{H}^N[\delta]) \leq \delta^{-1} \cdot \langle \mu, \varepsilon^N \rangle,$$

which converges to 0 as N tends to infinity as stated in (A.11). We thus restrict to N sufficiently large in order to ensure that

$$\mu(\mathcal{H}^N[\delta]) \leq \delta. \tag{A.24}$$

On the other hand, as a consequence of Step 3, see Lemma A.4, we obtain

$$\begin{aligned} &\mu(\mathcal{H}^N[\delta]^c \cap \mathcal{H}^N[\theta\, \delta]^\eta) \\ &\qquad \leq \mu\Big(\Big\{x \in \mathcal{X};\, \nu^N(\mathcal{G}_x[\delta, \eta]) \geq \frac{\delta}{2}\Big\}\Big) + \mu\Big(\Big\{x \in \mathcal{X};\, \nu(\mathcal{G}_x[\delta, \eta]) \geq \frac{\delta}{2}\Big\}\Big). \end{aligned} \tag{A.25}$$

For the next upper-bound, valid for η sufficiently small then N sufficiently large, we recall (A.20), (A.21), (A.23),(A.24), (A.25) and get

$$\mu(\mathcal{H}^N[\theta\, \delta]^\eta) \leq 3\delta.$$

It remains to combine this result with (A.19) and (A.22) to conclude the proof that $(\langle \mu^N, \varepsilon^N \rangle)$ tends to 0 as N tends to infinity, since δ can be taken arbitrarily small. This concludes the proof of Proposition 5.9 in the particular case where the sequences (μ^N) and (ν^N) are deterministic.

Step 5: Proof of Proposition 5.9 in the general case where (μ^N) *and* (ν^N) *are random.* For this final step, we no longer require the sequences (μ^N) and (ν^N) to be a priori deterministic, though our approach consists in referring to this convenient situation. We exploit [34, Lemma 4.2] to relate the convergence of probability to a.s. convergence of sequence extractions. Let $(N[k])_{k\geq 1} \in \mathbb{N}^{\mathbb{N}}$ be an increasing sequence. Since μ^N and ν^N converge in probability, we can extract a subsequence $(\check{N}[\ell])_{\ell\geq 1} = (N[K[\ell]])_{\ell\geq 1}$ from this sequence such that $(\check{\mu}^\ell) = (\mu^{\check{N}[\ell]})$ and $(\check{\nu}^\ell) = (\nu^{\check{N}[\ell]})$ converge a.s. respectively to μ and ν. On this event of probability 1, we deduce from Step 4 that $\langle \check{\mu}^\ell, \check{\varepsilon}^\ell \rangle$ converges to 0 as ℓ tends to infinity. Since the convergence of such an extraction of the sequence $(\langle \mu^N, \varepsilon^N \rangle)$ holds whatever the initial extraction, this concludes the proof of Proposition 5.9 in that $(\langle \mu^N, \varepsilon^N \rangle)$ converges in probability to 0. □

References

[1] H. Andersson. Limit theorems for a random graph epidemic model. *Annals of Applied Probability*, 8(4):1331–1349, 1998.

[2] F. Ball and T. Britton. Epidemics on networks with preventive rewiring. *Random Structures & Algorithms*, 61(2):250–297, 2022.

[3] F. Ball, T. Britton, K. Y. Leung, and D. Sirl. A stochastic SIR network epidemic model with preventive dropping of edges. *Journal of Mathematical Biology*, 78(6):1875–1951, 2019.

[4] F. Ball and D. Sirl. An SIR epidemic model on a population with random network and household structure, and several types of individuals. *Advances in Applied Probability*, 44(1):63–86, 2012.

[5] F. Ball and D. Sirl. Acquaintance vaccination in an epidemic on a random graph with specified degree distribution. *Journal of Applied Probability*, 50(4):1147–1168, 2013.

[6] F. Ball and D. Sirl. Evaluation of vaccination strategies for SIR epidemics on random networks incorporating household structure. *Journal of Mathematical Biology*, 76(1):483–530, 2018.

[7] F. Ball, D. Sirl, and P. Trapman. Threshold behaviour and final outcome of an epidemic on a random network with household structure. *Advances in Applied Probability*, 41(3):765–796, 2009.

[8] P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, Inc., New York, second edition, 1999.

[9] T. Bohman and M. Picollelli. SIR epidemics on random graphs with a fixed degree sequence. *Random Structures & Algorithms*, 41(2):179–214, 2012.

[10] C. Borgs, K. Huang, and C. Ikeokwu. A law of large numbers for SIR on the stochastic block model: A proof via herd immunity. *arXiv preprint arXiv:2410.07097*, 2024.

[11] T. Britton, F. Ball, and P. Trapman. A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2. *Science*, 369(6505):846–849, 2020.

[12] T. Britton, M. Deijfen, and F. Liljeros. A weighted configuration model and inhomogeneous epidemics. *Journal of Statistical Physics*, 145:1368–1384, 2011.

[13] T. Britton and D. Lindenstrand. Inhomogeneous epidemics on weighted networks. *Mathematical Biosciences*, 240(2):124–131, 2012.

[14] W. Chen, Y. Hou, and D. Yao. SIR epidemics on evolving Erdös Rényi graphs. *arXiv preprint arXiv:2208.11923*, 2022.

[15] D. Clancy Jr. Epidemics on critical random graphs with heavy-tailed degree distribution. *Stochastic Processes and their Applications*, 179:104510, 2025.

[16] D. L. Cohn. *Measure Theory: Second Edition*. Birkhäuser Advanced Texts Basler Lehrbücher. Springer, New York, NY, 2013.

[17] L. Decreusefond, J.-S. Dhersin, P. Moyal, and V. C. Tran. Large graph limit for an SIR process in random network with heterogeneous connectivity. *Ann. Appl. Probab.*, 22(2):541–575, 2012.

[18] M. Deijfen. Epidemics and vaccination on weighted graphs. *Mathematical Biosciences*, 232(1):57–65, 2011.

[19] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. An infinite-dimensional metapopulation SIS model. *J. Differ. Equ.*, 313:1–53, 2022.

[20] J.-F. Delmas, P. Frasca, F. Garin, C. Tran, A. Velleret, and P.-A. Zitt. Individual based SIS models on (not so) dense large random networks. *ALEA: Probab. Math. Stat.*, 21:1375–1405, 2024.
[21] J.-J. Duchamps, F. Foutel-Rodier, and E. Schertzer. General epidemiological models: Law of large numbers and contact tracing. *Electronic Journal of Probability*, 28:1–37, 2023.
[22] R. Durrett and D. Yao. Susceptible–infected epidemics on evolving graphs. *Electronic Journal of Probability*, 27:1–66, 2022.
[23] J. Fernley, P. Mörters, and M. Ortgiese. The contact process on a graph adapting to the infection. *arXiv preprint arXiv:2312.06251*, 2023.
[24] E. J. Ferreyra, M. Jonckheere, and J. P. Pinasco. SIR dynamics with vaccination in a large configuration model. *Applied Mathematics & Optimization*, 84:1769–1818, 2021.
[25] R. Forien, G. Pang, and É. Pardoux. Epidemic models with varying infectivity. *SIAM Journal on Applied Mathematics*, 81(5):1893–1930, 2021.
[26] R. Forien, G. Pang, and É. Pardoux. Multi-patch multi-group epidemic model with varying infectivity. *Probability, Uncertainty and Quantitative Risk*, 7(4):333–364, 2022.
[27] R. Forien, G. Pang, and É. Pardoux. Recent advances in epidemic modeling: Non-Markov stochastic models and their scaling limits. *The Graduate Journal of Mathematics*, 7(2):19–75, 2022.
[28] Y. Huang and A. Röllin. The SIR epidemic on a dynamic Erdös Rényi random graph. *arXiv preprint arXiv:2404.12566*, 2024.
[29] K. A. Jacobsen, M. G. Burch, J. H. Tien, and G. A. Rempała. The large graph limit of a stochastic epidemic model on a dynamic multilayer network. *Journal of Biological Dynamics*, 12(1):746–788, 2018.
[30] J. Jacod and A. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer-Verlag, Berlin, 2 edition, 2003.
[31] S. Janson, M. Luczak, and P. Windridge. Law of large numbers for the SIR epidemic on a random graph with given degrees. *Random Structures & Algorithms*, 45(4):726–763, 2014.
[32] S. Janson, M. Luczak, P. Windridge, and T. House. Near-critical SIR epidemic on a random graph with given degrees. *Journal of Mathematical Biology*, 74:843–886, 2017.
[33] Y. Jiang, R. Kassem, G. York, M. Junge, and R. Durrett. SIR epidemics on evolving graphs. *arXiv preprint arXiv:1901.06568*, 2019.
[34] O. Kallenberg. *Foundations of Modern Probability*. Probability and Its Applications. Springer New York, New York, NY, 2002.
[35] A. Kanga and É. Pardoux. Spatial SIR epidemic model with varying infectivity without movement of individuals: Law of Large Numbers. *arXiv preprint arXiv:2412.01673*, 2024.
[36] D. Keliger, I. Horváth, and B. Takács. Local-density dependent Markov processes on graphons with epidemiological applications. *Stoch. Proc. Appl.*, 148:324–352, 2022.
[37] W. R. Khudabukhsh, C. Woroszylo, G. A. Rempała, and H. Koeppl. A functional central limit theorem for SI processes on configuration model graphs. *Advances in Applied Probability*, 54(3):880–912, 2022.
[38] M. Kubasch. Large population limit for a multilayer SIR model including households and workplaces. *arXiv preprint arXiv:2305.17064*, 2023.
[39] A. A. Lashari, A. Serafimović, and P. Trapman. The duration of a supercritical SIR epidemic on a configuration model. *Electronic Journal of Probability*, 26:1–49, 2021.
[40] M. Luczak. A note on the Markovian SIR epidemic on a random graph with given degrees. *Journal of Theoretical Probability*, 37(2):1039–1051, 2024.
[41] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinska, and W. J. Edmunds. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLOS Medicine*, 5(3):1–1, 03 2008.
[42] G. Pang and É. Pardoux. Functional limit theorems for non-Markovian epidemic models. *The Annals of Applied Probability*, 32(3):1615–1665, 2022.
[43] G. Pang and E. Pardoux. Functional central limit theorems for epidemic models with varying infectivity. *Stochastics*, 95(5):819–866, 2023.
[44] G. Pang and É. Pardoux. Functional law of large numbers and PDEs for epidemic models with infection-age dependent infectivity. *Applied Mathematics and Optimization*, 87:Article 50, 2023.
[45] G. Pang and É. Pardoux. SPDE for stochastic SIR epidemic models with infection-age dependent infectivity. *arXiv:2410.17200*, 2024.

[46] E. A. Perkins. Dawson-Watanabe superprocesses and mesure-valued diffusions. In Springer, editor, *Ecole d'Eté de Probabilités de Saint-Flour*, volume 1781 of *Lectures Notes in Math.*, pages 125–329, New York, 1999.
[47] G. Pólya and G. Szegö. *Problems and Theorems in Analysis I: Series, integral calculus, theory of functions*. Springer Verlag, Berlin, Heidelberg, 1978.
[48] G. Reinert. The asymptotic evolution of the general stochastic epidemic. *The Annals of Applied Probability*, 5(4):1061–1086, 1995.
[49] Y. Shang. Mixed SI (R) epidemic dynamics in random graphs with general degree distributions. *Applied Mathematics and Computation*, 219(10):5042–5048, 2013.
[50] K. Spricer and T. Britton. An SIR epidemic on a weighted network. *Network Science*, 7(4):556–580, 2019.
[51] A.-S. Sznitman. Topics in propagation of chaos. In *Ecole d'été de probabilités de Saint-Flour XIX—1989*, pages 165–251. Springer, 1991.
[52] E. Volz. SIR dynamics in random networks with heterogeneous connectivity. *Journal of mathematical biology*, 56:293–310, 2008.
[53] F. J. Wang. Limit theorems for age and density dependent stochastic population models. *Journal of Mathematical Biology*, 2(4):373–400, 1975.
[54] F. J. Wang. Gaussian approximation of some closed stochastic epidemic models. *Journal of Applied Probability*, 14(2):221–231, 1977.

Department of Computational Applied Mathematics and Operations Research, George R. Brown School of Engineering and Computing, Rice University, Houston, TX 77005
Email address: gdpang@rice.edu

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

Université Paris-Saclay, Université d'Evry Val d'Essonne, CNRS, LaMME, UMR 8071 91037 Evry, France
Email address: aurelien.velleret@nsup.org