Stochastic heterogeneous SIR model with infection-age dependent infectivity on large 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 is 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. To each individual is associated a random infectivity function of its infection age, which is allowed to depend upon 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, and reflects the heterogeneities in individual characteristics and social connectivity.

1. Introduction

In the present paper, we are interested in the large population limit of the 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 contribute to the individual heterogeneity. Another source of heterogeneity arises from the individual connectivities (such as households, workplaces, 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, 54, 52, 9, 5, 34, 25, 41, 42, 15], on weighted (configuration) graphs [12, 18, 13, 53], on dynamic (evolving) graphs [3, 2, 35, 22, 14, 24, 30]. (See also the relevant models on random networks with household structures in [7, 4, 6, 40], and on multilayer networks in [31, 40].)

Much relevant to our work are the large population limits for epidemic models on random graphs. In particular, [17, 33] proved a functional law of large numbers (FLLN) for a Markovian SIR model on a configuration model graph with specified degree distributions

Date: July 9, 2026.

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.

and edges being randomly matched, and established the measure-valued limit as a system of nonlinear differential equations, which verifies the conjecture in [54]. In [39], a functional 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, 33, 39, 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 [38], 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 latter two references ([38, 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 survey [28]. 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 [44] (see also [56, 55] for Gaussian approximations and [51] for measure-valued state descriptors and PDE limit for the FLLN), and epidemic models with varying infectivity are studied in [26, 45, 27], where deterministic or stochastic integral equations are derived for the FLLNs and FCLTs, respectively. When adding the age of infection into the variables of the model, measure-valued processes have been 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 [46] and [47], respectively. This is extended to spatially dense SIR models in [48], where the FLLN is established with a PDE limit on a graphon. 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 [40], an individual-based multi-layer SIR model with households and workplaces to take into account the 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 behaviors. So they may be distributed on a compact subset of $\mathbb{R}^d$, as in the spatial SIR model considered in [48, 37], which we generalize in the simplest setting of pairwise interactions. Concerning 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. 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 [38], there is no heterogeneity in the contact rate and the dependency of the graph on individual types is more restrictive. In [10], the study of SIR dynamics on large graphs only focuses on stochastic block models.

Each individual is moreover associated with a random varying infectivity function, which reflects the force of infection of each individual at any time after 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 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 the force of infection that is exerted from its infectious neighbors upon each individual.

To describe the evolution dynamics, we use three measure-valued processes for the three compartments, where at each time t 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 characterize the measure-valued process tracking the infection ages of infected individuals as the unique solution of a PDE. To note, the dynamics described from the set of measure-valued equations and the PDE may be seen as evolving on a graphon.

The PDE model is linear, with a boundary condition given by the product of the measure-valued susceptible process and the aggregate force of infection. A specific form of Picard iteration for these quantities is used to deduce the existence of solutions. Solution uniqueness follows from classical methods based on the characteristics of the equation, as in [23, Section 3.2], e.g., adapted for measure-valued solutions.

The novelty of our paper is in the combination of two aspects: heterogeneity of the individuals and of the graph of interactions on the one hand, and infection-age dependent infectivity on the other hand. The second aspect has been considered recently by some of the authors of the present paper in a homogeneous situation. We borrow from these works, see notably [27], the effective techniques consisting in the construction of an intermediate model for which the infection rates are derived from the deterministic limiting model. Whereas the limiting infectivity function was previously described as acting on total count processes, it is acting in the present paper at the individual level, with a dependence on the characteristics of each individual through the limiting connectivity kernel and infectivity function.

The variance-based approach to approximating random sums by their expectations, previously used in [27] for treating the heterogeneity in the choice of the infectivity function, is effective in our framework with additional heterogeneity, see Lemma 5.7. 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 be optimal.

Nonetheless, the combination of both heterogeneities induces some complications for the control of the error induced by the replacement of the original model by that simpler one, see in particular the proof of Lemma 5.6.

As in [20], we have relaxed as much as possible the regularity assumptions of the interaction kernel in the limit, which we only assume to be almost everywhere continuous. In addition, we allow for the individual characteristics to take values in a state space as general as possible, which implies that the state descriptors evolve as measure-valued processes. Whereas the initial condition is generally assumed to be given through an independent and identically distributed (i.i.d.) sample, we have merely assumed weak convergence in probability in terms of distributions. The resulting difficulty concerns particularly the discrepancies in the initial conditions, see the proof of Lemma 5.10, which is based on a new convergence result of independent interest for weakly continuous kernels on general Polish spaces, see Proposition 5.9.

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 our assumptions in Section 2. In Section 3, we state 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 of the convergence in the FLLN are given in Section 5, with additional technical supporting results proved in Appendix A. Concluding remarks are provided in Section 6.

1.2. **Notation.** We denote by $\mathbb{N} = \{1, 2, \ldots\}$ the set of positive integers. 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

$$\|f\|_\infty = \sup_{x \in \mathcal{X}} |f(x)|.$$

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

In what follows $(\mathbb{X}, \mathrm{d})$ will denote a Polish metric space, i.e., complete and separable. We denote by $\mathcal{M}_1(\mathbb{X})$ and $\mathcal{M}(\mathbb{X})$ the sets of probability and finite measures on $\mathbb{X}$, respectively. For $\mu \in \mathcal{M}(\mathbb{X})$ and a real-valued measurable function f defined on $\mathbb{X}$, we will sometimes denote the integral of f with respect to the measure μ, if well-defined, by $\langle \mu, f \rangle = \int_{\mathbb{X}} f(x)\, \mu(\mathrm{d}x) = \int f \,\mathrm{d}\mu$. We endow $\mathcal{M}_1(\mathbb{X})$ and $\mathcal{M}(\mathbb{X})$ with the topology of weak convergence. $\mathcal{M}_1(\mathbb{X})$ and $\mathcal{M}(\mathbb{X})$ are Polish metric spaces with the Lévy-Prokhorov metric, cf. [49, Section II.4]. We denote by $\mathbb{D}(\mathbb{R}_+, \mathcal{M}(\mathbb{X}))$ the space of paths from $\mathbb{R}_+$ to $\mathcal{M}(\mathbb{X})$ that are right-continuous with left limits (cád-lág). This space is endowed with the Skorokhod topology (see, *e.g.*, [8, Chapter 3]). We shall also use the two following Skorokhod spaces: $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ and $\mathbb{D}(\mathbb{R}_+, \mathcal{M}(\mathbb{X} \times \mathbb{R}_+))$.

2. Model description

We consider a population consisting of N individuals, who are in one 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, nor 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 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 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 infected has an infectivity function depending randomly on its infection age:

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

where λ_i^N is assumed not to be identically 0. In particular, for the initially infected individuals $j \in \mathcal{I}^N(0)$, their infectivity level at a given time $t > 0$ is:

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

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 $\mathcal{Y}^N = \{Y_i^N\}_{i \le N}$ be an i.i.d. sequence of uniform random variables on $[0,1]$, independent from $\mathcal{F}_0^N$. Let $\widehat{\lambda}$ be a deterministic measurable function from $\mathbb{X} \times [0,1] \times \mathbb{R}_+$ to $\mathbb{R}_+$ that is càd-làg in its last component. 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.

There is no further generality in considering $\mathcal{Y}^N = \{Y_i^N\}_{i\le N}$ be an i.i.d. sequence of random variables taking values in $\mathbb{Y}$, where $\mathbb{Y}$ is a Polish space, by adjusting the definition of $\widehat{\lambda}$, see [36, Theorem 3.19].

Assumption 1. *There exists a global upper-bound $\lambda^* < \infty$ on the random functions λ_i^N taking values in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)$, in the sense that the following holds almost surely:*

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

Remark 2.2. *For Assumption 1 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 $A^N(0)$:

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

while assuming that there is no pathological situation where the above set would be empty. We specify the individual distribution of infection duration through its cumulative distribution function. 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. *There exists a measurable function $(F_x^c(a))_{x\in\mathbb{X}, a\in\mathbb{R}_+}$ with values in $[0,1]$ which is right-continuous, non-increasing in the variable 'a' such that $\lim_{a\to\infty} F_x^c(a) = 0$ for all $x \in \mathbb{X}$ and such that the two following identities hold a.s. for any $N \ge 1$, $t > 0$, $i \in \mathcal{S}^N(0)$ and $j \in \mathcal{I}^N(0)$:*

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

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 $\mathbb{P}_0^N(\eta_j^{N,0} > t) = \mathbb{P}\big(\eta_j^{N,0} > t \,\big|\, \mathcal{F}_0^N\big)$. Note that the condition on $\eta_j^{N,0}$ in Assumption 2 excludes the case where the denominator in the above expression would be zero.

Let $F_x := 1 - F_x^c$, which is interpreted as the c.d.f. of the infection durations η_i^N given that $X_i^N = x$.

Remark 2.3. *This assumption appears more explicitly stated under the formalism presented in Remark 2.1. By definition, there exists a deterministic measurable function $g : \mathbb{X} \times [0,1] \to \mathbb{R}_+$ such that $\eta_i^N = g(X_i^N, Y_i^N)$ a.s. For each $x \in \mathbb{X}$, we define the cumulative distribution:*

$$F_x(t) = \mathbb{P}(g(x, Y_1^N) \le t) = \int_{[0,1]} \mathbb{1}_{\{g(x,y)\le t\}} \mathrm{d}y$$

and set $F^c_x(\cdot) = 1 - F_x(\cdot)$. F^c_x *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, for any* $N \geq 1$*, any* $i \in \mathcal{S}^N(0)$*, and any* $t > 0$*:*

$$\mathbb{P}^N_0(\eta^N_i > t) = \mathbb{P}\big(g(X^N_i, Y^N_i) > t \,\big|\, X^N_i\big) = F^c_{X^N_i}(t).$$

For any N *and any initially infected individual* $j \in \mathcal{I}^N(0)$*, the constraint that* $\eta^{N,0}_j$ *is a well-defined positive quantity, which translates into* $g(X^N_j, Y^N_j) > A^N_j(0)$*, explains the form of the second identity in Assumption 2, where for any* $t > 0$*,*

$$\begin{aligned}\mathbb{P}^N_0(\eta^{N,0}_j > t) &= \mathbb{P}\big(g(X^N_j, Y^N_j) > t + A^N_j(0) \,\big|\, X^N_j, A^N_j(0), \mathbb{1}_{\{g(X^N_j, Y^N_j) > A^N_j(0)\}}\big)\\ &= \frac{F^c_{X^N_j}(A^N_j(0) + t)}{F^c_{X^N_j}(A^N_j(0))}.\end{aligned}$$

Assumption 3. *There exists a deterministic measurable function* $\bar{\lambda}$ *from* $\mathbb{X} \times \mathbb{R}_+$ *to* $\mathbb{R}_+$ *such that the two following identities hold for any* $N \geq 1$*, any* $a \geq 0$*, any* $t \geq 0$*, any* $i \in \mathcal{S}^N(0)$ *and any* $j \in \mathcal{I}^N(0)$*:*

$$\begin{aligned}\mathbb{E}^N_0\big[\lambda^N_i(a)\big] &= \bar{\lambda}(X^N_i, a),\\ \mathbb{E}^N_0\big[\lambda^{N,0}_j(t)\big] &= \frac{\bar{\lambda}(X^N_j, A^N_j(0) + t)}{F^c_{X^N_j}(A^N_j(0))},\end{aligned}$$

where $F^c_{X^N_j}(A^N_j(0)) > 0$ *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^**.*

Additional conditions on the regularity of $\bar{\lambda}$ will be stated in Assumption 6, once the limiting measures on $\mathbb{X}$ will be defined. Recall that $\mathbb{E}^N_0\big[\lambda^N_i(a)\big]$ equals by definition $\mathbb{E}\big[\lambda^N_i(a) \,\big|\, \mathcal{F}^N_0\big]$ and similarly for $\mathbb{E}^N_0\big[\lambda^{N,0}_j(t)\big]$.

Remark 2.4. *Under the formalism presented in Remark 2.1,* $\bar{\lambda}$ *is directly expressed as follows:*

$$\bar{\lambda}(x,a) = \int_{[0,1]} \widehat{\lambda}(x,y,a)\,\mathrm{d}y\,,$$

and the first identity in Assumption 3 is automatic. We can also check the second identity:

$$\begin{aligned}\mathbb{E}^N_0\big[\lambda^{N,0}_j(t)\big] &= \mathbb{E}\big[\widehat{\lambda}(X^N_j, Y^N_j, t + A^N_j(0)) \,\big|\, X^N_j, A^N_j(0), \mathbb{1}_{\{g(X^N_j, Y^N_j) > A^N_j(0)\}}\big]\\ &= \frac{\int_{[0,1]} \widehat{\lambda}(X^N_j, y, t + A^N_j(0))\,\mathrm{d}y}{F^c_{X^N_j}(A^N_j(0))} = \frac{\bar{\lambda}(X^N_j, t + A^N_j(0))}{F^c_{X^N_j}(A^N_j(0))},\end{aligned}$$

where we have exploited in the second line the fact that $\widehat{\lambda}(X^N_j, y, t + A^N_j(0)) > 0$ *entails* $g(X^N_j, Y^N_j) > t + A^N_j(0)$ *and a fortiori* $g(X^N_j, Y^N_j) > A^N_j(0)$*. In addition, for any* $x \in \mathbb{X}$ *and any* $a \geq 0$ *such that* $F^c_x(a) > 0$*, since* $\widehat{\lambda}(x,y,a) = 0$ *on the set* $\{g(x,y) \leq a\}$*,*

$$\frac{\bar{\lambda}(x,a)}{F^c_x(a)} = \frac{\int_{[0,1]} \widehat{\lambda}(x,y,a)\,\mathbb{1}_{\{g(x,y)>a\}}\,\mathrm{d}y}{\int_{[0,1]} \mathbb{1}_{\{g(x,y)>a\}}\,\mathrm{d}y}\,,$$

and is actually upper-bounded by λ^* as soon as $\widehat{\lambda}$ is itself upper-bounded by λ^*.

Remark 2.5. *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 positive 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:*

$$\bar{\lambda}(X_i^N, a) = \widetilde{\lambda}(X_i^N, a) \cdot F^c_{X_i^N}(a). \tag{2.3}$$

Remark 2.6. *As an instance of the case presented in the above Remark 2.5, 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 $\widetilde{\lambda}(x,a) \equiv \check{\lambda}$ and $g(x,y) = \check{g}(x)\cdot y$. The expression for $\bar{\lambda}$ further simplifies:*

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

Remark 2.7. *In the previous assumptions, we did not allow the infectious period to extend indefinitely, in line with the SIR formalism. It does however not appear crucial in the following proofs. Some SI formalism without recovery or even a mix between SI and SIR formalism could be treated very similarly. We mean situations where $\mathbb{P}(\eta_i^N = +\infty) > 0$ for some i, with possibly F_x upper-bounded by a constant strictly smaller than 1 for some non-negligible set of $x \in \mathbb{X}$ and the function g in Remark 2.3 with possibly infinite values.*

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 $\kappa^N : \mathbb{X}\times\mathbb{X} \to [0,1]$:

$$\mathbb{P}_0^N\big(i \overset{N}{\sim} j\big) = \mathbb{P}\big(i \overset{N}{\sim} j \,\big|\, \mathcal{F}_0^N\big) = \kappa^N(X_i^N, X_j^N), \tag{2.4}$$

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 $\mathcal{F}_0^N$.

Remark 2.8. *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 $i,j = 1,\dots,N$, where the X_i^N are uniformly and independently distributed in $[0,1]$. Another very natural choice of X_i^N is also $X_i^N = i/N$ for $i = 1,\dots,N$, which also leads to the limiting empirical measure $\bar{\mu}_X$, as defined in the formula (2.20) below, being the Lebesgue measure on $[0,1]$.*

Remark 2.9. *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* [38]*. Following* [20]*, this form of scaling is however not assumed in order 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 Remark 2.12, and in Section 6 after the proofs.*

2.3. **Force of infection.** 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.

Remark 2.10. *Heterogeneity in the susceptibility to infection can be included in the model through these 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.,* [43]*, we allow for the interplay between characteristics that are continuous and for an additional degree of independent randomness.*

Assumption 4. *The values* $(\omega^N(i,j))_{i,j\in[\![1,N]\!]}$ *are mutually independent between different edges and independent of the* $(\lambda_i^N)_{i\le N}$ *conditionally on* $\mathcal{F}_0^N$*. There exists a deterministic measurable function* $\gamma^N : \mathbb{X}\times\mathbb{X}\to\mathbb{R}_+$ *such that:*

$$\mathbb{E}\big[\omega^N(i,j)\,\big|\,\mathcal{F}_0^N,\mathcal{E}^N\big] = \begin{cases}\gamma^N(X_i^N,X_j^N) & \text{if } (i,j)\in\mathcal{E}^N,\\ 0 & \text{otherwise.}\end{cases} \tag{2.5}$$

Remark 2.11. *Under the formalism given in Remark 2.1, the values* $(\omega^N(i,j))_{i,j\in[\![1,N]\!]}$ *are independent of the* $(\lambda_i^N)_{i\le N}$ *if they are independent of the* (Y_i^N)*, conditionally on* $\mathcal{F}_0^N$*.*

The function γ^N captures the expected infectious rate of interaction for active contacts. As a consequence of these definitions, we have

$$N\cdot\mathbb{E}_0^N\big[\omega^N(i,j)\big] = \bar{\omega}^N(X_i^N,X_j^N), \quad \text{where } \ \bar{\omega}^N := N\cdot\kappa^N\cdot\gamma^N\,. \tag{2.6}$$

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.

We 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 $(i,j)\in\mathcal{E}^N$:

$$\upsilon^N(i,j) = \mathrm{Var}\big[\omega^N(i,j)\,\big|\,\mathcal{F}_0^N,\mathcal{E}^N\big]\,. \tag{2.7}$$

Our upcoming Assumption 8 entails that the quantity $\upsilon^N(i,j)$ is finite, for any $i,j\in[\![1,N]\!]$. Given the above definition of γ^N, this definition of υ^N shall be interpreted as follows:

$$\mathbb{E}\big[\omega^N(i,j)^2\,\big|\,\mathcal{F}_0^N,\mathcal{E}^N\big] = \begin{cases}\upsilon^N(i,j)+\gamma^N(X_i^N,X_j^N)^2 & \text{if } (i,j)\in\mathcal{E}^N,\\ 0 & \text{otherwise.}\end{cases} \tag{2.8}$$

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

$$\overline{\mathfrak{F}}_i^N(t) = \sum_{j\in\mathcal{I}^N(t)} \omega^N(i,j)\cdot\lambda_j^N(A_j^N(t))\,. \tag{2.9}$$

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

$$D_i^N(t) = \int_0^t\int_0^\infty \mathbb{1}_{\{D_i^N(s^-)=0\}}\mathbb{1}_{\left\{u\le\overline{\mathfrak{F}}_i^N(s^-)\right\}}Q_i(\mathrm{d}s,\mathrm{d}u)\,, \tag{2.10}$$

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 mutually 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

$$\tau_i^N = \inf\{t\ge 0;\, D_i^N(t)=1\};\quad \mathcal{D}^N(t) := \{i\in\mathcal{S}^N(0);\, \tau_i^N\le t\}, \tag{2.11}$$

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 $D_i^N(t)=1$. 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 (2.10), we also obtain the following alternative expression for $\mathfrak{F}_i^N(t)$ for each $i=1,\dots,N$:

$$\begin{aligned}\overline{\mathfrak{F}}_i^N(t) &= \sum_{k\in\mathcal{I}^N(0)}\omega^N(i,k)\lambda_k^N(A_k^N(0)+t)\\ &\quad+\sum_{j\in\mathcal{S}^N(0)}\omega^N(i,j)\int_0^t\int_0^\infty \lambda_j^N(t-s)\mathbb{1}_{\{D_j^N(s^-)=0\}}\mathbb{1}_{\left\{u\le\overline{\mathfrak{F}}_j^N(s^-)\right\}}Q_j(\mathrm{d}s,\mathrm{d}u)\,.\end{aligned} \tag{2.12}$$

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

$$\bar\mu_t^{S,N}(\mathrm{d}x) = \frac{1}{N}\sum_{i\in\mathcal{S}^N(t)}\delta_{X_i^N}(\mathrm{d}x)\,, \tag{2.13}$$

$$\bar\mu_t^{I,N}(\mathrm{d}x,\mathrm{d}a) = \frac{1}{N}\sum_{i\in\mathcal{I}^N(t)}\delta_{X_i^N}(\mathrm{d}x)\delta_{A_i^N(t)}(\mathrm{d}a)\,, \tag{2.14}$$

$$\bar\mu_t^{R,N}(\mathrm{d}x) = \frac{1}{N}\sum_{i\in\mathcal{R}^N(t)}\delta_{X_i^N}(\mathrm{d}x)\,. \tag{2.15}$$

For any $t>0$, $\bar\mu_t^{S,N}$ and $\bar\mu_t^{R,N}$ 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 $\bar\mu_t^{I,N}$ 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:

$$\langle \bar{\mu}_t^{S,N}, \varphi \rangle = \langle \bar{\mu}_0^{S,N}, \varphi \rangle - \frac{1}{N} \sum_{i \in \mathcal{D}^N(t)} \varphi(X_i^N)\,, \tag{2.16}$$

  with

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

- for the infected process:

$$\langle \bar{\mu}_t^{I,N}, \psi \rangle = \frac{1}{N} \sum_{j \in \mathcal{I}^N(0)} \mathbb{1}_{\left\{\eta_j^{N,0} > t\right\}} \psi(X_j^N, A_j^N(0) + t) + \frac{1}{N} \sum_{i \in \mathcal{D}^N(t)} \mathbb{1}_{\left\{\tau_i^N + \eta_i^N > t\right\}} \psi(X_i^N, t - \tau_i^N)\,, \tag{2.17}$$

  with

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

- for the recovered process:

$$\langle \bar{\mu}_t^{R,N}, \varphi \rangle = \langle \bar{\mu}_0^{R,N}, \varphi \rangle + \frac{1}{N} \sum_{j \in \mathcal{I}^N(0)} \mathbb{1}_{\left\{\eta_j^{N,0} \le t\right\}} \varphi(X_j^N) + \frac{1}{N} \sum_{i \in \mathcal{D}^N(t)} \mathbb{1}_{\left\{\tau_i^N + \eta_i^N \le t\right\}} \varphi(X_i^N)\,. \tag{2.18}$$

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: $S^N(t) = N\langle \bar{\mu}_t^{S,N}, \mathbb{1}\rangle$, $I^N(t) = N\langle \bar{\mu}_t^{I,N}, \mathbb{1}\rangle$ and $R^N(t) = N\langle \bar{\mu}_t^{R,N}, \mathbb{1}\rangle$ with $\mathbb{1}(x) \equiv 1$ and $\mathbb{1}(a,x) \equiv 1$.

We make the following assumptions on the initial quantities.

Assumption 5. *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 $\bar{\mu}_0^{R,N}$.*

We consider also the two distributions $\bar{\mu}_X^N$ and $\bar{\mu}_X$ of characteristics on $\mathbb{X}$ for the whole population:

$$\bar{\mu}_X^N(\mathrm{d}x) = \bar{\mu}_0^{S,N}(\mathrm{d}x) + \langle \bar{\mu}_0^{I,N}(\mathrm{d}x, .), \mathbf{1}\rangle + \bar{\mu}_0^{R,N}(\mathrm{d}x) = \frac{1}{N} \sum_{i \le N} \delta_{X_i^N}(\mathrm{d}x), \tag{2.19}$$

$$\bar{\mu}_X(\mathrm{d}x) = \bar{\mu}_0^S(\mathrm{d}x) + \langle \bar{\mu}_0^I(\mathrm{d}x, .), \mathbf{1}\rangle + \bar{\mu}_0^R(\mathrm{d}x). \tag{2.20}$$

Here $\mathbf{1}(a) \equiv 1$. A direct consequence of Assumption 5 is that $\bar{\mu}_X$ is a probability distribution, and $\bar{\mu}_X^N$ converges weakly to $\bar{\mu}_X$.

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. We include additional regularity conditions together with the first assumption, on the quantities introduced in Assumptions 2, 3 and 5.

Assumption 6. *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.

$\bar{\mu}_0^I$ is absolutely continuous with respect to $(\bar{\mu}_X \otimes Leb)$, where Leb denotes the Lebesgue measure on $\mathbb{R}_+$. The function : $(x, a) \mapsto F_x(a)$ is $\bar{\mu}_0^I$-a.e. continuous while the function $\bar{\lambda}$ is $(\bar{\mu}_X \otimes Leb)$-a.e. continuous.

Assumption 6 is stated in terms of the convergence of the product $N\gamma^N\kappa^N$ without requiring separate convergence properties of γ^N and κ^N. 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 instance with more denseness but less frequent contacts for the infections during travels than at the workplaces.

Regarding our next convergence results in probability, there is no loss of generality in considering the following assumption.

Assumption 7. *There exists $\omega^* > 0$ that is an upper-bound of $\bar{\omega}^N$ uniformly on $\mathbb{X}^2$ and on 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 8. *The two following quantities $\bar{\gamma}^N$ and Υ^N converge in probability to zero as $N \to \infty$:*

$$\begin{aligned}
\bar{\gamma}^N &:= \frac{1}{N^2} \sum_{i,j} \gamma^N(X_i^N, X_j^N), \\
\Upsilon^N &:= \frac{1}{N} \sum_{i,j} \mathbb{E}_0^N\left[\upsilon^N(i,j); (i,j) \in \mathcal{E}^N\right] \\
&= \frac{1}{N} \sum_{i,j} \kappa^N(X_i^N, X_j^N) \cdot \mathbb{E}_0^N\left[\upsilon^N(i,j) \,\middle|\, (i,j) \in \mathcal{E}^N\right].
\end{aligned} \tag{2.21}$$

By $\mathbb{E}_0^N\left[\upsilon^N(i,j) \middle| (i,j) \in \mathcal{E}^N\right]$ is meant the ratio of $\mathbb{E}\left[\upsilon^N(i,j); (i,j) \in \mathcal{E}^N \,\middle|\, \mathcal{F}_0^N\right]$ divided by $\mathbb{P}\left[(i,j) \in \mathcal{E}^N \,\middle|\, \mathcal{F}_0^N\right] = \kappa^N(X_i^N, X_j^N)$.

Remark 2.12. *To better understand the variance term in Assumption 8, 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 (2.4) *and* (2.5) *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 6 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 8 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 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 8. So we do not have any additional restriction on* $\bar{\sigma}_N$ *in this model.*

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.*

3. Functional law of large numbers

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

Theorem 3.1. *Under Assumptions 1–8,*

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

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

$$\bar{\mu}^S_t(\mathrm{d}x) = \bar{\mu}^S_0(\mathrm{d}x) - \int_0^t \overline{\mathfrak{F}}(s,x)\bar{\mu}^S_s(\mathrm{d}x)\mathrm{d}s\,, \tag{3.1}$$

and

$$\begin{aligned}\overline{\mathfrak{F}}(t,x) = &\int_0^\infty\!\!\int_{\mathbb{X}} \bar{\omega}(x,x')\frac{\bar{\lambda}(x',a'+t)}{F^c_{x'}(a')}\bar{\mu}^I_0(\mathrm{d}x',da') \\ &+ \int_0^t\!\!\int_{\mathbb{X}} \bar{\omega}(x,x')\bar{\lambda}(x',t-s)\overline{\mathfrak{F}}(s,x')\bar{\mu}^S_s(\mathrm{d}x')\mathrm{d}s\,,\end{aligned} \tag{3.2}$$

with $\bar{\lambda}$, $\bar{\omega}$ *being given respectively in Assumptions 3 and 6. The uniqueness of the above system is stated among the potential candidates* $(\bar{\mu}^S_t, \overline{\mathfrak{F}}(t))_{t\geq 0}$ *in which* $\bar{\mu}^S_\cdot$ *is a* $\mathcal{M}(\mathbb{X})$*-valued càd-làg process such that* $\mu_t(\mathbb{X}) \leq 1$ *for any* $t \geq 0$, *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 $(\bar{\mu}^S_\cdot, \overline{\mathfrak{F}}(\cdot))$ *and the initial conditions* $\bar{\mu}^I_0$ *and* $\bar{\mu}^R_0$, *the explicit formulas for* $\langle \bar{\mu}^I_t, \psi\rangle$ *and* $\langle \bar{\mu}^R_t, \phi\rangle$ *with the test functions* $\psi \in C_b(\mathbb{X}\times\mathbb{R}_+)$ *and* $\varphi \in C_b(\mathbb{X})$ *are as follows:*

$$\begin{aligned}\langle \bar{\mu}^I_t, \psi\rangle = &\int_0^\infty\!\!\int_{\mathbb{X}} \psi(x,a+t)\frac{F^c_x(a+t)}{F^c_x(a)}\bar{\mu}^I_0(\mathrm{d}x,\mathrm{d}a) \\ &+ \int_0^t\!\!\int_{\mathbb{X}} \psi(x,t-s)F^c_x(t-s)\overline{\mathfrak{F}}(s,x)\bar{\mu}^S_s(\mathrm{d}x)\mathrm{d}s\,,\end{aligned} \tag{3.3}$$

and

$$\langle \bar\mu_t^R, \varphi\rangle = \langle \bar\mu_0^R, \varphi\rangle + \int_0^\infty\int_{\mathbb{X}} \varphi(x)\Big(1-\frac{F_x^c(a+t)}{F_x^c(a)}\Big)\bar\mu_0^I(\mathrm{d}x,\mathrm{d}a) + \int_0^t\int_{\mathbb{X}} \varphi(x)F_x(t-s)\overline{\mathfrak{F}}(s,x)\bar\mu_s^S(\mathrm{d}x)\mathrm{d}s\,. \tag{3.4}$$

Note that the unique solution to equation (3.1) is given by

$$\bar\mu_t^S(\mathrm{d}x) = \exp\Big(-\int_0^t \overline{\mathfrak{F}}(s,x)\mathrm{d}s\Big)\bar\mu_0^S(\mathrm{d}x)\,, \tag{3.5}$$

and that (3.3) results, evaluated for any $x\in\mathbb{X}$ with the function $\psi(x',a') = \omega(x,x')\frac{\bar\lambda(x',a')}{F_{x'}^c(a')}$ and after comparison with (3.2), in the following expression for $\overline{\mathfrak{F}}$:

$$\overline{\mathfrak{F}}(t,x) = \int_0^\infty\int_{\mathbb{X}} \bar\omega(x,x')\frac{\bar\lambda(x',a')}{F_{x'}^c(a')}\bar\mu_t^I(\mathrm{d}x', da')\,. \tag{3.6}$$

Remark 3.2. *Our conditions on the limiting kernel of interaction $\bar\omega$ encompass most, if not all, of the (known) pairwise kernel interactions found in applications. 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 dynamics on a graphon.*

Consider the special case: $\bar\omega(x,x') = \bar\gamma\cdot\bar\kappa(x,x') = \bar\gamma\cdot x\cdot x'$ for some value $\bar\gamma>0$. This multiplicative form for $\bar\omega(x,x')$ is a typical example of a graphon. Here we assume that the (X_i^N) are uniformly distributed in $[0,1]$, and $\bar\kappa^N=\bar\kappa$ for any N and $\bar\kappa(x,x') = x\cdot x'$ as in Remark 2.8, and $\bar\gamma^N = \bar\gamma/N$ in Assumption 6 (hence, γ^N does not depend on the types of the individuals i and j in contact and $\Upsilon^N\equiv 0$). This kernel structure displays a convenient property of separation between three contributions: 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:

$$\overline{\mathfrak{F}}(t) = \int_0^\infty\int_{[0,1]} x'\cdot\frac{\bar\lambda(x',a')}{F_{x'}^c(a')}\bar\mu_t^I(\mathrm{d}x', da')\,.$$

A reformulation of this system of integro-delay equations (3.1)-(3.4) is proposed in Section 4.3 as a system of partial differential equations.

We also remark that in the special case of $\lambda_i^N(a)$ discussed in Remark 2.5, given the expression of $\bar\lambda(x,a)$ in (2.3), we obtain the following:

$$\overline{\mathfrak{F}}(t,x) = \int_0^\infty\int_{\mathbb{X}} \bar\omega(x,x')\widetilde\lambda(x',a')\bar\mu_t^I(\mathrm{d}x', da')\,.$$

4. Properties of the limiting system of equations

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

Proposition 4.1. *Under Assumptions 1 and 6, the set of equations (3.1)–(3.2) has a unique solution $(\bar\mu_\cdot^S, \overline{\mathfrak{F}}_\cdot)$.*

Proof. We first prove the uniqueness. Before proceeding, we establish some useful bounds. Suppose that $(\bar{\mu}^S_., \overline{\mathfrak{F}}_.)$ is a solution. By Assumptions 1 and 6, we have $\bar{\lambda}(x,t) \leq \lambda^*$ and $\bar{\omega}(x,x') \leq \omega^*$ for some $\omega^* > 0$. Recalling (3.2) and Assumption 3, we derive

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

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 (2.20) is a probability measure.

On the other hand by (3.1),

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

Suppose now that there are two solutions $(\bar{\mu}^{S,\ell}_., \overline{\mathfrak{F}}^\ell_.)$, $\ell = 1, 2$. Let $D_t = \|\overline{\mathfrak{F}}^1(t,.) - \overline{\mathfrak{F}}^2(t,.)\|_\infty$. By (4.5):

$$\begin{aligned}
\overline{\mathfrak{F}}^1(t,x) - \overline{\mathfrak{F}}^2(t,x) &= \int_0^t\!\!\int_{\mathbb{X}} \bar{\omega}(x,x')\bar{\lambda}(x',t-s)\overline{\mathfrak{F}}^1(s,x')(\bar{\mu}^{S,1}_s(\mathrm{d}x') - \bar{\mu}^{S,2}_s(\mathrm{d}x'))\mathrm{d}s \\
&\quad + \int_0^t\!\!\int_{\mathbb{X}} \bar{\omega}(x,x')\bar{\lambda}(x',t-s)(\overline{\mathfrak{F}}^1(s,x') - \overline{\mathfrak{F}}^2(s,x'))\bar{\mu}^{S,2}_s(\mathrm{d}x')\mathrm{d}s\,.
\end{aligned} \tag{4.3}$$

Since : $a \in \mathbb{R}_+ \mapsto e^{-a}$ is 1-Lipschitz continuous:

$$\left\| \exp\left(- \int_0^s \overline{\mathfrak{F}}^1(r,\,.)\mathrm{d}s\right) - \exp\left(- \int_0^s \overline{\mathfrak{F}}^2(r,\,.)\mathrm{d}r\right)\right\|_\infty \leq \int_0^s D_r \,\mathrm{d}r\,.$$

We then deduce from (3.5) that: $\left\|\bar{\mu}^{S,1}_s - \bar{\mu}^{S,2}_s\right\|_{TV} \leq \int_0^s D_r\,\mathrm{d}r$. By injecting this inequality in (4.3) with (4.1) and (4.2), we deduce for any $T > 0$ and $t \in [0,T]$ that $D_t \leq C\int_0^t D_s \mathrm{d}s$ where $C = \lambda^*\omega^* + (\lambda^*\omega^*)^2 T$. Thanks to Gronwall's inequality (for any $T > 0$), the proof of the uniqueness is concluded.

Finally, existence can be proved by the following modified Picard iteration. We initialize with $\bar{\mu}^{S,0}_s(\mathrm{d}x') = \bar{\mu}^S_0(\mathrm{d}x')$, $\overline{\mathfrak{F}}^0(t,x) = 0$, then for any $k \geq 0$:

$$\bar{\mu}^{S,k+1}_t(\mathrm{d}x) = \bar{\mu}^S_0(\mathrm{d}x) - \int_0^t \overline{\mathfrak{F}}^k(s,x)\bar{\mu}^{S,k+1}_s(\mathrm{d}x)\mathrm{d}s\,, \tag{4.4}$$

and

$$\begin{aligned}
\overline{\mathfrak{F}}^{k+1}(t,x) &= \int_0^\infty\!\!\int_{\mathbb{X}} \bar{\omega}(x,x')\frac{\bar{\lambda}(x',a'+t)}{F^c_{x'}(a')}\bar{\mu}^I_0(\mathrm{d}x', da') \\
&\quad + \int_0^t\!\!\int_{\mathbb{X}} \bar{\omega}(x,x')\bar{\lambda}(x',t-s)\overline{\mathfrak{F}}^k(s,x')\bar{\mu}^{S,k+1}_s(\mathrm{d}x')\mathrm{d}s\,.
\end{aligned} \tag{4.5}$$

In these two definitions, we make the choice of the product between $\overline{\mathfrak{F}}^k$ and $\bar{\mu}^{S,k+1}$ because then we have the following explicit expression of $\bar{\mu}^{S,k+1}$ in terms of $\overline{\mathfrak{F}}^k$:

$$\bar{\mu}_t^{S,k+1}(\mathrm{d}x) = \exp\Big(- \int_0^t \overline{\mathfrak{F}}^k(r,x)\mathrm{d}r \Big) \bar{\mu}_0^S(\mathrm{d}x)\,, \tag{4.6}$$

similarly to (3.5). We can thus justify similarly the a priori estimates (4.2) and (4.1) for the above approximating sequence. Then, the arguments used for proving uniqueness yield the convergence of our Picard iteration, exactly as in the usual case of globally Lipschitz coefficients. ☐

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 $\bar{\mu}_X$-a.e. continuous.*

Proof. Recalling (3.2) and Assumption 3, we first deduce the following inequality for any $t \geq 0$ and any $x_1, x_2 \in \mathbb{X}$:

$$\begin{aligned}\Big|\overline{\mathfrak{F}}(t,x_1) - \overline{\mathfrak{F}}(t,x_2)\Big| \leq \lambda^* \int_{\mathbb{X}} |\bar{\omega}(x_1,x') - \bar{\omega}(x_2,x')| \Big\{ \Big(\frac{F^c_{x'}(a'+t)}{F^c_{x'}(a')} \Big) \bar{\mu}_0^I(\mathrm{d}x', \mathrm{d}a') \\ + \int_0^t \overline{\mathfrak{F}}(s,x') \exp\Big(- \int_0^s \overline{\mathfrak{F}}(r,x')\mathrm{d}r \Big) \mathrm{d}s \, \bar{\mu}_0^S(\mathrm{d}x') \Big\}.\end{aligned}$$

With similar arguments as in the proof of Proposition 4.1 to deduce (4.1), notably with Assumption 5, we obtain

$$\|\overline{\mathfrak{F}}(.,x_1) - \overline{\mathfrak{F}}(.,x_2)\|_\infty \leq \lambda^* \int_{\mathbb{X}} |\bar{\omega}(x_1,x') - \bar{\omega}(x_2,x')| \, \bar{\mu}_X(\mathrm{d}x'). \tag{4.7}$$

As stated in Appendix A in Lemma 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, since $\bar{\omega}$ is $\bar{\mu}_X^{\otimes 2}$-a.e. continuous by virtue of Assumption 6. 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 3.1 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 $\bar{\mu}_0^I(\mathbb{X} \times \{0\}) = 0$. Then, $\bar{\mu}_t^I(\mathrm{d}x, \mathrm{d}a)$ in (3.3) is the unique solution to the following equation: for any $\psi \in C^{0,1}(\mathbb{X} \times \mathbb{R}_+)$ and $t > 0$,*

$$\frac{d}{dt}\langle \bar{\mu}_t^I, \psi \rangle = \langle \bar{\mu}_t^I, \partial_a \psi - h\psi \rangle + \int_{\mathbb{X}} \psi(x,0) \overline{\mathfrak{F}}(t,x) \bar{\mu}_t^S(\mathrm{d}x)\,. \tag{4.8}$$

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

$$\langle \partial_t \bar{\mu}_t^I, \psi \rangle + \langle \partial_a \bar{\mu}_t^I, \psi \rangle = -\langle h\, \bar{\mu}_t^I, \psi \rangle \tag{4.9}$$

with the initial condition $\bar\mu_0^I$ given in Assumption 5 and the boundary condition at $a=0$: $\tilde\mu_t^I(\mathrm{d}x,0)=\overline{\mathfrak{F}}(t,x)\bar\mu_t^S(\mathrm{d}x)$, where $\tilde\mu_t^I(\mathrm{d}x,a)$ is the "density" of $\bar\mu_t^I(\mathrm{d}x,\mathrm{d}a)$, that is, $\bar\mu_t^I(\mathrm{d}x,\mathrm{d}a)=\tilde\mu_t^I(\mathrm{d}x,a)\mathrm{d}a$ and $a\mapsto\tilde\mu_t^I(\mathrm{d}x,a)$ is continuous at 0^+.

Recalling (3.5), remark that we can also write the last term in (4.8) as

$$\int_{\mathbb{X}}\psi(x,0)\overline{\mathfrak{F}}(t,x)\exp\Big(-\int_0^t\overline{\mathfrak{F}}(s,x)\mathrm{d}s\Big)\bar\mu_0^S(\mathrm{d}x)\,.$$

which only involves the initial $\bar\mu_0^S(\mathrm{d}x)$ and $\overline{\mathfrak{F}}(t,x)$.

We directly identify from Proposition 4.3 and (3.6) that the system of equations under consideration is a weak formulation of the following system. Here, $u_t^I(x,a)$ (resp. $u_t^S(x)$) represent the density of $\bar\mu_t^I$ (resp. $\bar\mu_t^S$) with respect to $\bar\mu_X(\mathrm{d}x)\,\mathrm{d}a$ (resp. $\bar\mu_X(\mathrm{d}x)$):

$$\begin{cases}\partial_t u_t^I(x,a)+\partial_a u_t^I(x,a)=-h(x,a)\,u_t^I(x,a)\,,\\ u_t^I(x,0)=-\partial_t u_t^S(x)=\overline{\mathfrak{F}}(t,x)\,u_t^S(x)\,,\\ \overline{\mathfrak{F}}(t,x)=\displaystyle\int_0^\infty\!\!\int_{\mathbb{X}}\bar\omega(x,x')\frac{\bar\lambda(x',a')}{F^c_{x'}(a')}u_t^I(x',a')\,\bar\mu_X(\mathrm{d}x')\,\mathrm{d}a'\,.\end{cases}\tag{4.10}$$

We can associate to this system the density process u^R of recovered individual, defined by $u_t^R(x)=1-u_t^S(x)-\int_0^\infty u_t^I(x,a)\,\mathrm{d}a$ and solution to $\partial_t u_t^R(x)=\int_0^\infty h(x,a)\,u_t^I(x,a)\mathrm{d}a$. While the uniqueness of the system in Proposition 4.3 entails the one of the above system (by defining $\bar\mu_t^S(\mathrm{d}x)=u_t^I(x,a)\,\bar\mu_X(\mathrm{d}x)$ and $\bar\mu_t^I(\mathrm{d}x,\mathrm{d}a)=u_t^I(x,a)\,\bar\mu_X(\mathrm{d}x)\,\mathrm{d}a$), the regularity of such solutions and in what stronger sense the system (4.10) is satisfied is left open.

Proof. By taking derivative with respect to t in the expression of $\langle\bar\mu_t^I,\psi\rangle$ in (3.3), we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle\bar\mu_t^I,\psi\rangle&=\int_0^\infty\!\!\int_{\mathbb{X}}\partial_a\psi(x,a+t)\frac{F_x^c(a+t)}{F_x^c(a)}\bar\mu_0^I(\mathrm{d}x,\mathrm{d}a)\\&\quad-\int_0^\infty\!\!\int_{\mathbb{X}}\psi(x,a+t)\frac{f_x(a+t)}{F_x^c(a)}\bar\mu_0^I(\mathrm{d}x,\mathrm{d}a)\\&\quad+\int_{\mathbb{X}}\psi(x,0)\overline{\mathfrak{F}}(t,x)\bar\mu_t^S(\mathrm{d}x)\\&\quad-\int_0^t\!\!\int_{\mathbb{X}}\psi(x,t-s)f_x(t-s)\overline{\mathfrak{F}}(s,x)\bar\mu_s^S(\mathrm{d}x)\mathrm{d}s\\&\quad+\int_0^t\!\!\int_{\mathbb{X}}\partial_a\psi(x,t-s)F_x^c(t-s)\overline{\mathfrak{F}}(s,x)\bar\mu_s^S(\mathrm{d}x)\mathrm{d}s\,.\end{aligned}$$

Next, from (3.3), we observe that

$$\begin{aligned}\langle\bar\mu_t^I,\partial_a\psi\rangle&=\int_0^\infty\!\!\int_{\mathbb{X}}\partial_a\psi(x,a+t)\frac{F_x^c(a+t)}{F_x^c(a)}\bar\mu_0^I(\mathrm{d}x,\mathrm{d}a)\\&\quad+\int_0^t\!\!\int_{\mathbb{X}}\partial_a\psi(x,t-s)F_x^c(t-s)\overline{\mathfrak{F}}(s,x)\bar\mu_s^S(\mathrm{d}x)\mathrm{d}s\,,\end{aligned}$$

and

$$\langle\bar\mu_t^I,h\psi\rangle=\int_0^\infty\!\!\int_{\mathbb{X}}\psi(x,a+t)\frac{f_x(a+t)}{F_x^c(a)}\bar\mu_0^I(\mathrm{d}x,\mathrm{d}a)$$

$$+\int_0^t\int_{\mathbb{X}}\psi(x,t-s)f_x(t-s)\overline{\mathfrak{F}}(s,x)\bar{\mu}_s^S(\mathrm{d}x)\mathrm{d}s\,.$$

Hence, the last three identities lead to the expression in (4.8).

Next, (3.3) entails the following formula

$$\begin{aligned}\langle\bar{\mu}_t^I,\psi\rangle=\int_t^\infty\int_{\mathbb{X}}\psi(x,a)\frac{F_x^c(a)}{F_x^c(a-t)}\bar{\mu}_0^I(\mathrm{d}x,\mathrm{d}a-t)\\+\int_0^t\int_{\mathbb{X}}\psi(x,a)F_x^c(a)\overline{\mathfrak{F}}(t-a,x)\bar{\mu}_{t-a}^S(\mathrm{d}x)\mathrm{d}a\,.\end{aligned}\tag{4.11}$$

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

Now integrating (4.8) over the interval $[0,t]$, we obtain

$$\langle\bar{\mu}_t^I,\psi\rangle=\langle\bar{\mu}_0^I,\psi\rangle+\int_0^t\langle\bar{\mu}_s^I,\partial_a\psi-h\psi\rangle ds+\int_0^t\int_{\mathbb{X}}\psi(x,0)\overline{\mathfrak{F}}(s,x)\bar{\mu}_s^S(\mathrm{d}x)\mathrm{d}s\,.\tag{4.12}$$

We now choose in (4.12) $\psi_n(x,a)=\varsigma(x)(1-na)^+$ for some $\varsigma\in C^1(\mathbb{R}_+)$. Since $\bar{\mu}_0^I(\mathbb{X}\times\{0\})=0$, we deduce from (4.11) that $\bar{\mu}_t^I(\mathbb{X}\times\{0\})=0$ holds for any $t\geq0$. We have

$$\langle\bar{\mu}_t^I,\psi_n\rangle\to0\quad\text{and}\quad\int_0^t\langle\bar{\mu}_s^I,h\psi_n\rangle\mathrm{d}s\to0,\quad\text{as }n\to\infty.$$

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

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

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

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

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

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

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

$$\int_0^t\langle\bar{\mu}_s^I,\partial_a\psi\rangle\mathrm{d}s+\int_0^t\int_{\mathbb{X}}\psi(x,0)\overline{\mathfrak{F}}(s,x)\bar{\mu}_s^S(\mathrm{d}x)\mathrm{d}s=-\int_0^t\langle\partial_a\bar{\mu}_s^I,\psi\rangle\mathrm{d}s.\tag{4.13}$$

Hence, from (4.8) and (4.13), we obtain the PDE model in (4.9) with the boundary condition at $a=0$: $\tilde{\mu}_s^I(\mathrm{d}x,0)=\overline{\mathfrak{F}}(s,x)\bar{\mu}_s^S(\mathrm{d}x)$.

We now prove uniqueness. The argument which follows is based on the characteristics of the equation, as in [23, Section 3.2], yet adapted for measure-valued solutions in a way reminiscent of duality approaches as in [16, Section 5.5].

Let us consider in addition to $\bar{\mu}^I$ any arbitrary solution $(\widetilde{\mu}^I_t)$ to the PDE in (4.8) 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 $\varphi_0 \in C^{0,1}(\mathbb{X} \times [0,t])$:

$$\Psi_t(x,a) = \begin{cases} \psi_0(x, a-t) \cdot \exp\Big[\int_{a-t}^a h(x,a')\mathrm{d}a'\Big] & \text{for any } x \in \mathbb{X}, a \in (t,\infty), \\ \varphi_0(x, t-a) \cdot \exp\Big[\int_0^a h(x,a')\mathrm{d}a'\Big] & \text{for any } x \in \mathbb{X}, a \in [0,t]. \end{cases} \tag{4.14}$$

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 (4.17).

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 $a \in (t,\infty)$,

$$\begin{aligned} \partial_t\Psi_t(x,a) &= -\partial_a\psi_0(x,a-t)\exp\Big[\int_{a-t}^a h(x,a')\mathrm{d}a'\Big] + \Psi_t(x,a)\cdot h(x,a-t), \\ \partial_a\Psi_t(x,a) &= \partial_a\psi_0(x,a-t)\ \exp\Big[\int_{a-t}^a h(x,a')\mathrm{d}a'\Big] + \Psi_t(x,a)\cdot\big[h(x,a) - h(x,a-t)\big]. \end{aligned} \tag{4.15}$$

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

$$\begin{aligned} \partial_t\Psi_t(x,a) &= \partial_a\varphi_0(x,t-a)\exp\Big[\int_0^a h(x,a')\mathrm{d}a'\Big], \\ \partial_a\Psi_t(x,a) &= -\partial_a\psi_0(x,t-a)\exp\Big[\int_0^a h(x,a')\mathrm{d}a'\Big] + \Psi_t(x,a)\cdot h(x,a). \end{aligned} \tag{4.16}$$

Thanks to (4.15) and (4.16), 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:

$$\langle \Delta\mu^I_t, \Psi_t\rangle = \langle \Delta\mu^I_0, \psi_0\rangle = 0, \tag{4.17}$$

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 (4.14) satisfies $\Psi_T(x,a) = \psi(x,a)$. Since $h(x,a) = -\partial_a \log F^c_x(a)$, we know that $\exp\Big[-\int_0^a h(x,a')\mathrm{d}a'\Big] = F^c_x(a) > 0$. 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 $\psi_0(x,a-T)\cdot\exp\Big[\int_{a-T}^a h(x,a')\mathrm{d}a'\Big]$ agrees with $\psi(x,a)$ and similarly for any $a \in [0,T]$ with $\varphi_0(x,T-a)\cdot\exp\Big[\int_0^a h(x,a')\mathrm{d}a'\Big]$ instead. Thanks to (4.17), 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 (4.9), with the boundary condition at $a = 0$ as specified in Proposition 4.3. Then, the integration by parts formula in (4.13) holds with $\widetilde{\mu}^I$ instead of $\bar{\mu}^I$, and so (4.12) 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. $\square$

5. Proof of Theorem 3.1, 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 $D_i^N(t)$ in (2.10). Define

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

and

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

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 $\mathbb{R}$:*

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

where the limit $\bar{\mu}^S$ is the first term of the unique solution of (3.1)-(3.2).

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

$$\mathbb{P}_0^N(\widetilde{D}_i^N(t) = 0) = \exp\left(- \int_0^t \overline{\mathfrak{F}}(s, X_i^N) \mathrm{d}s \right),$$

we deduce

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

Recall from Assumption 5 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

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

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

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

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:

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

We relate this convergence to the differences $\overline{\mathfrak{F}}_i^N(s) - \overline{\mathfrak{F}}(s, X_i^N)$ for $s \geq 0$ and $i \in \mathcal{S}^N(0)$. We deduce from (2.9) and (3.2) that those differences can then be decomposed into seven terms, by exploiting the following definitions.

We first define

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

where $\widetilde{\tau}_j^N$ is the jump time of 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 jump of (D_j^N) and that of $(\widetilde{D}_j^N)$.

Remark 5.2. *Similarly to $\mathcal{D}^N(t)$ in (2.11), we can define $\widetilde{\mathcal{D}}^N(t)$ accordingly to $(\widetilde{\tau}_j^N)$:*

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

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

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

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

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

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 $\mathrm{d}s\mathrm{d}u$:

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

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 $t \geq 0$,

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

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

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

in a similar way as $\overline{\mathfrak{V}}_i^{N,1}(t)$, now for the initially infected individuals. We recall from Assumption 3 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 $\overline{\mathfrak{E}}^{N,1}$, now 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*

$$\begin{aligned}
\overline{\mathfrak{F}}_i^N(t) - \overline{\mathfrak{F}}(t, X_i^N) = {}& \overline{\mathfrak{A}}_i^N(t) && \textit{(discrepancies between the jumps)}\\
&+ \overline{\mathfrak{V}}_i^{N,0}(t) + \overline{\mathfrak{V}}_i^{N,1}(t) && \textit{(Law of Large Number)}\\
&+ \overline{\mathfrak{L}}^{N,0}(t, X_i^N) + \overline{\mathfrak{L}}^{N,1}(t, X_i^N) && \textit{(discrepancies between the kernels)}\\
&+ \overline{\mathfrak{E}}^{N,0}(t, X_i^N) + \overline{\mathfrak{E}}^{N,1}(t, X_i^N) && \textit{(discrepancies between the initial conditions).}
\end{aligned}$$

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 (3.5) 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 (3.2), 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 (2.11) 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 (2.12), 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 (2.17), 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

$$\begin{aligned}\overline{\mathfrak{F}}_i^{N,0}(t) - \overline{\mathfrak{V}}_i^{N,0}(t) - \overline{\mathfrak{L}}^{N,0}(t,X_i^N) &= \int_{\mathbb{X}}\!\int_0^\infty \bar\omega(X_i^N,x')\cdot\frac{\bar\lambda(x',a'+t)}{F^c_{x'}(a')}\bar\mu_0^{I,N}(\mathrm{d}x',\mathrm{d}a')\\ &= \overline{\mathfrak{E}}^{N,0}(t,X_i^N) + \overline{\mathfrak{F}}^0(t,X_i^N)\,,\end{aligned} \tag{5.21}$$

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 (2.16) 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:

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

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 1, 5, 6 and 8, there exist a constant $C > 0$ such that*

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

holds a.s. for any $t \geq 0$, where $\overline{\Delta}_D^N(t)$ 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 in Assumption 7.

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

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

We observe for any $j \in \mathcal{S}^N(0)$, the following identity by virtue of (2.6):

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

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

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

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

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

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

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

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

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

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 (2.8) and (2.4) to deduce the following inequality:

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

Remark that from (2.6), for any $i,j\in\mathcal{S}^N(0)$, $\kappa^N(X_i^N,X_j^N)\cdot\gamma^N(X_i^N,X_j^N)=\bar{\omega}^N(X_i^N,X_j^N)/N\le \omega^*/N$. Recalling the definitions given in Assumption 8, we thus deduce

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

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

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 $(\bar{\mathfrak{V}}^N)_{N\ge1}$ defined in (5.27):*

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

Proof. We treat this component in the same way as $\mathfrak{U}^N$, starting with the Cauchy-Schwartz inequality:

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

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^*:

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

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

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

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 3:

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

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

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

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

$\overline{\mathfrak{L}}^{N,1}$ 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, 3 and 5:*

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

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 3, 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, 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 (2.13)-(2.14), their added masses is upper-bounded by 1, which concludes the proof of Lemma 5.8. $\square$

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 $\mathcal{Y}$ be a Polish space. Let $\mu \in \mathcal{M}(\mathbb{X})$, $\nu \in \mathcal{M}(\mathcal{Y})$ and $k : \mathbb{X}\times\mathcal{Y} \mapsto \mathbb{R}$ 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 $\mathcal{M}(\mathbb{X})$ 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:*

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

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 $\mathcal{Y} = \mathbb{X} \times \mathbb{R}_+$ and

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

$$\nu^{N,0}(\mathrm{d}x', \mathrm{d}a') := \bar{\mu}_0^{I,N}(\mathrm{d}x', \mathrm{d}a'), \quad \nu^0(\mathrm{d}x', \mathrm{d}a') := \bar{\mu}_0^I(\mathrm{d}x', \mathrm{d}a').$$

Assumption 5 states that μ^N and $\nu^{N,0}$ converge weakly to respectively μ and ν^0. We recall the upper-bound of $\bar{\lambda}(x', a'+t)/F_{x'}^c(a')$ by λ^* as a consequence of Assumptions 2 and 3. We recall also the uniform upper-bound ω^* of $\bar{\omega}^N$, as defined in Assumption 7. $\omega^* \cdot \lambda^*$ is thus a global upper-bound of k_t^0.

By virtue of Assumption 6 and since $t > 0$, the following property holds for $(\mu \otimes \nu^0)$ almost every triplet (x, x', a') such that $F_{x'}^c(a') = 0$: $F_{\check{x}'}^c(\check{a}' + t) = 0$ holds for any $(\check{x}, \check{x}', \check{a}')$ in a small enough neighborhood of (x, x', a'), thus also $\bar{\lambda}(\check{x}', \check{a}' + t) = 0$ and $k(\check{x}, \check{x}', \check{a}') = 0$.

The following property holds on the other hand, for $(\mu \otimes \nu^0)$ almost every triplet (x, x', a') such that $F_{x'}^c(a') > 0$: $F_{\check{x}'}^c(\check{a}') > 0$ holds for any $(\check{x}, \check{x}', \check{a}')$ in a small enough neighborhood of (x, x', a'). Still by virtue of Assumption 6 on $\bar{\lambda}$ and $\bar{\mu}_0^I$, the function : $(x', a') \mapsto \bar{\lambda}(x', a'+t)$ is ν^0-a.e. continuous. We thus conclude that k_t^0 is $(\mu \otimes \nu^0)$-a.e. continuous.

Proposition 5.9 does therefore entail that $\langle \bar{\mu}_0^{S,N}, |\overline{\mathfrak{E}}^{N,0}(t,.)|\rangle$ tends to zero.

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 $\mathcal{Y} = \mathbb{X}$ and

$$\mu^N = \nu^{N,1} := \bar{\mu}_0^{S,N}, \quad \mu = \nu^1 = \bar{\mu}_0^S, \quad k_t^1(x,x') = \bar{\omega}(x,x') \cdot \mathbb{E}[\bar{\lambda}(x', t - \widetilde{\tau}_{x'})].$$

The fact that μ^N and $\nu^{N,1}$ converge weakly in probability to $\mu = \nu^1$ follows from Assumption 5. k_t^1 is bounded under Assumptions 1 and 6. To check that k_t^1 is $(\mu \otimes \nu^1)$ a.e. continuous, we exploit the following alternative expression:

$$k_t^1(x,x') = \bar{\omega}(x,x') \int_0^t \bar{\lambda}(x', t-s) \cdot \overline{\mathfrak{F}}(s,x') \cdot \exp\Big[- \int_0^s \overline{\mathfrak{F}}(r,x') \mathrm{d}r \Big] \mathrm{d}s,$$

derived with the same argument as for (5.24). Thanks to Assumptions 1 and 6 and Lemma 4.2, we can conclude that k^1 is $(\mu \otimes \nu^1)$ a.e. continuous, as stated in Lemma A.3 in Appendix A.

Proposition 5.9 does therefore entail that $\langle \bar{\mu}_0^{S,N}, |\overline{\mathfrak{E}}^{N,1}(t,.)|\rangle$ tends to zero. □

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 $N \to \infty$, $\overline{\Delta}_D^N(t)$ 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 : $t \mapsto \overline{\Delta}_D^N(t)$ 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:

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

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

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

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 8 and 6, such that

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

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 Assumptions 3 and 7, for the defining properties of λ^* and ω^*, we obtain that

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

holds for any $s \geq 0$ and any $x \in \mathbb{X}$. Since $\bar{\mu}_0^{S,N}(\mathbb{X}) \leq 1$, 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 $\overline{\Delta}_D^N(t)$ converges to 0 in probability for any t as $N \to \infty$. □

Before proceeding, let us state a result which is exactly Theorem II.4.1 in [49] and will be needed in the next proof. Let us generally consider a Polish space $\mathcal{Y}$. We say that a subset $\mathfrak{M}$ of $\mathcal{C}_b(\mathcal{Y})$ is separating if (i) it includes the constant function : $x \mapsto 1$ and (ii) it discriminates elements of $\mathcal{M}_1(\mathcal{Y})$, in that for any $(\nu,\nu') \in \mathcal{M}_1(\mathcal{Y})^2$, $\nu = \nu'$ is equivalent to the property that $\langle\nu,\phi\rangle = \langle\nu,\phi\rangle$ for any $\phi \in \mathfrak{M}$. It is well known that $\mathcal{C}_b(\mathcal{Y})$ itself is separating. A sequence of random elements of $\mathbb{D}(\mathbb{R}_+, \mathcal{M}(\mathcal{Y}))$ is said to be C-tight if it is tight and any limit of a converging subsequence is a.s. continuous.

Proposition 5.12. *A sequence of processes $(\nu^N)_{n\in\mathbb{N}^*}$ is C-tight in $\mathbb{D}(\mathbb{R}_+, \mathcal{M}_1(\mathcal{Y}))$ if and only if:*

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

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

(b) ***Tightness of the projections.*** *The sequence* $(\langle \nu^N_\cdot, \varphi\rangle)_{N\in\mathbb{N}^*}$ *is C-tight in* $\mathbb{D}(\mathbb{R}_+,\mathbb{R})$ *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*

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

Proof. We apply Proposition 5.12 with $\mathcal{Y}=\mathbb{X}$, $\nu^N=\bar\mu^{(S,N)}$ 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 $\bar\mu_0^{(S,N)}\Rightarrow\bar\mu_0^S$ in probability (exploiting for instance the Lévy-Prokhorov metric).

Concerning Point (b), let us consider any $\varphi\in\mathcal{C}_{b,+}(\mathbb{X})$. 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 $\overline\Delta_D^N$ in respectively (5.2) and (5.6):

$$\big|\langle\bar\mu_t^{(S,N)},\varphi\rangle-\langle\widetilde\mu_t^{(S,N)},\varphi\rangle\big|\le\|\varphi\|_\infty\cdot\overline\Delta_D^N(t), \tag{5.40}$$

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. 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 $\bar\mu^{(S)}_\cdot$, which concludes that $\bar\mu^{(S,N)}$ converges to $\bar\mu^{(S)}$ in $\mathbb{D}(\mathbb{R}_+,\mathcal{M}(\mathbb{X}))$. □

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

Proposition 5.14. *The following convergence to 0 holds a.s. 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

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

where $C,\lambda^*<\infty$ while $\overline\Delta_D^N(t)$, $\overline{\mathfrak{V}}^N$, $\|\bar\omega^N-\bar\omega\|_\infty$ and $\overline{\mathfrak{E}}^N(t)$ all tend to 0 thanks respectively to Proposition 5.11, Assumption 8, Assumption 6, 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 $\varphi\in C_b(\mathbb{X})$:

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

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 $\bar{\mu}^{N,S}$ 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*

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

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 $t \geq 0$:

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

and

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

We recall that $\mathcal{D}^N(t)$ is defined in (2.11) 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 $\langle \bar{\mu}_0^{R}, \varphi \rangle$ is part of Assumption 5.

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 5, 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

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

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):

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

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 $(0,t]$.

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

Thanks to Assumption 5 and (3.5), the above conditional expectation converges in probability to:

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

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

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

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 $\{D_i^N(r)=\widetilde{D}_i^N(r),\forall r\in[0,t]\}$:

$$|\langle\widetilde{\mu}_t^{R,N,1}-\bar{\mu}_t^{R,N,1},\varphi\rangle|\le\|\varphi\|_\infty\cdot\overline{\Delta}_D^N(t). \quad (5.52)$$

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 (3.4), 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 $\mathcal{Y}=\mathbb{X}$ 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 (2.19). Since the latter converges in probability to $\bar{\mu}_X$, recalling (2.19), we can find for any $\varepsilon>0$ some compact set K_ε such that

$$\sup_{N\ge1}\mathbb{P}\Big(\bar{\mu}_X^N(K_\varepsilon^c)>\varepsilon\Big)<\varepsilon.$$

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 $\mathbb{D}_1$, which concludes the proof of Proposition 5.15. □

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

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

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:

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

$$= \frac{1}{N} \sum_{i \in \mathcal{S}^N(0)} \left(\mathbb{1}_{\{D_i^N(t)=0\}} \psi(X_i^N, -1) + \mathbb{1}_{\{D_i^N(t)=1\}} \mathbb{1}_{\{\eta_i^N > t - \tau_i^N\}} \psi(X_i^N, t - \tau_i^N) \right)$$
$$+ \frac{1}{N} \sum_{j \in \mathcal{I}^N(0)} \mathbb{1}_{\{\eta_i^{N,0} > t\}} \psi(X_j^N, A_j^N(0) + t) \,.$$

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 $\mathcal{Y} = \mathbb{X} \times [-1, \infty)$ 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 [36, 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 $t \geq 0$:

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

and

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

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

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

Thanks to Assumption 5, the above conditional expectation converges in probability to:

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

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

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

So we conclude to the convergence in probability of $\langle \bar{\mu}_t^{I,N,0}, \psi\rangle$ to $\langle \bar{\mu}_t^{I,N,0}, \psi\rangle$ 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):

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

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

Thanks to Assumption 5 and (3.5), the above conditional expectation converges in probability to

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

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

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

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

$$|\langle \widetilde{\mu}_t^{I,N,1} - \bar{\mu}_t^{I,N,1}, \psi\rangle| \le \|\psi\|_\infty \cdot \overline{\Delta}_D^N(t). \tag{5.60}$$

The right-hand side converges in probability to 0 as N tends to infinity thanks to Proposition 5.11. 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

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

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

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

Recalling (5.62), this entails Point (a), for any $T > 0$ with $K^{(\mathcal{Y})} = K \times [-1, A+T]$. 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 $\mathbb{D}(\mathbb{R}_+, \mathcal{M}(\mathbb{X}\times\mathbb{R}_+))$. $\square$

6. Concluding remarks

Typical examples of pairwise kernel interactions found in applications are contact matrices over a discrete space as in [11], e.g., as well as spatial interactions as in [37] with characteristics that are distributed on a compact subset of $\mathbb{R}^d$. Our conditions encompass both settings, and any combination of those.

They also allow for a scaling factor $\varepsilon^N > 0$ depending on N such that $\bar{\kappa}^N = \varepsilon^N \cdot \bar{\kappa}$ where $\bar{\kappa}$ can be a constant but also a non-degenerate bounded measurable function. A typical choice is for instance $\varepsilon^N = N^{-\alpha}$ with $\alpha \in (0, 1)$. Provided we assume as in [38] that the contact rate is fixed at some value $\bar{\gamma}^N > 0$, $N\varepsilon^N \to \infty$ is required (and sufficient) for Assumption 8 to hold. We do not need to require the convergence to infinity of $N\varepsilon^N/\log(N)$ as in [38]. We remark that $N\varepsilon^N \to \infty$ ensures that the node degrees diverge, with no condition on their relation to N. 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 case where $\varepsilon^N = N^{-1}$ leads to limiting equations of a different nature, as hinted by the exploratory simulations in [20] and clarified in the results of [10] for specific cases of stochastic block models.

Besides, our conditions guarantee the robustness of the convergence against perturbations that concern a negligible subset of edges, since the two statistics introduced in Assumption 8 are averages.

Moreover, it can be helpful to introduce different kinds of interactions with their corresponding scaling factors. For example, we may consider the (X_i^N) as spatial coordinates, say uniformly distributed in $[0, 1]$ and with the norm-distance $d(x, y) = |x - y|$ on the circle so that all the locations are equivalent (and where 0 is merged with 1). 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:

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

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: $(\kappa^{N,\mathcal{L}} \cdot \gamma^{N,\mathcal{L}})/(\kappa^{N,\mathcal{G}} \cdot \gamma^{N,\mathcal{G}}) \geq (2\delta)/(1-2\delta)$.

A possible extension of our result would be to consider individual types that evolve in time, typically seasonal or when the individual locations move in space. Another natural objective would be 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 [37]. The inclusion of superspreading events, defined as the occurrence of highly heterogeneous transmission, would require to relax the boundedness conditions (see [19] in this direction regarding $\bar{\omega}$ or [29] regarding λ_i^N), which also leads to additional difficulties.

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. Lemma A.2 is used to deduce the a.e. continuity of sections of a.e. continuous kernels and a related result is stated in the next Lemma A.3. 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 [36, 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 [50]), 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 [36, Lemma 4.2] that $\psi^{\widetilde{N}_n}(t)$ converges in probability to $\psi(t)$ locally uniformly in t. □

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

Lemma A.2. *Let $\mathcal{Y}$ be a Polish space. Let $\mu \in \mathcal{M}(\mathbb{X})$, $\nu \in \mathcal{M}(\mathcal{Y})$ and $k : \mathbb{X} \times \mathcal{Y} \mapsto \mathbb{R}$ be a bounded measurable function that is $(\mu \otimes \nu)$-a.e. continuous. Then the function : $x \mapsto \bar{k}(x, .)$ is μ a.e. continuous from $\mathbb{X}$ with values in $L^1(\nu)$.*

Lemma A.2 is firstly exploited in the proof of Lemma 4.2 with $\mathcal{Y} = \mathbb{X}$, $\mu = \nu = \bar{\mu}_X$ and $k = \bar{\omega}$. It is also involved in the proof of the following Lemma A.3.

Proof. For any $\delta, \eta > 0$, $x \in \mathbb{X}$ and $x' \in B(x, \eta)$:

$$\int_{\mathcal{Y}} |k(x,y) - k(x',y)|\, \nu(\mathrm{d}y) \leq \delta + \|k\|_\infty\, \nu\big(\mathcal{G}_x[\delta, \eta]\big), \tag{A.1}$$

where $\mathcal{G}_x[\delta, \eta] = \{y \in \mathcal{Y},\ \exists x'' \in B(x, \eta),\ |k(x,y) - k(x'',y)| > \delta\}$.

By assumption, there exists a measurable subset N of $\mathbb{X} \times \mathcal{Y}$ such that $\mu \otimes \nu(N) = 0$ and such that k is continuous in any $(x,y) \in \mathbb{X}^2 \setminus N$. By the Fubini-Tonelli theorem, for any $x \in \mathbb{X}$, the section N_x is measurable, where $N_x := \{y \in \mathbb{X}, (x,y) \in N\}$. In addition, $\int_{\mathbb{X}} \nu(N_x)\, \mu(\mathrm{d}x) = \mu \otimes \nu(N) = 0$, so that $\nu(N_x) = 0$ for μ almost every x. For any $y \notin N_x$, k is continuous in (x,y), thus $k(x,.)$ is continuous in y, which entails $\cap_{n\geq 1} \mathcal{G}_x[\delta, 2^{-n}] \in N_x$. For any x such that $\nu(N_x) = 0$, $\nu\big(\mathcal{G}_x[\delta, \eta]\big)$ tends to zero as η tends to zero. Recalling (A.1), it concludes the proof of Lemma A.2. □

Lemma A.3. *Let $t > 0$, $\mathcal{Y}$ be a Polish space and $\mu \in \mathcal{M}(\mathcal{Y})$. Let the measurable function $k : \mathcal{Y} \times [0,t] \to \mathbb{R}$ be bounded and $\mu \otimes$ Leb-a.e. continuous and the measurable function $F : \mathcal{Y} \times [0,t] \to \mathbb{R}$ be bounded and such that : $y \mapsto F(y,.)$ is μ-a.e. continuous with respect to the uniform norm in $[0,t]$. Then, : $y \mapsto \int_0^t k(y,a)\, F(y,a)\, \mathrm{d}a$ is μ-a.e. continuous.*

Leb denotes the Lebesgue measure on $[0,t]$. This lemma is to be applied in the proof of Lemma 5.10 with $\mathcal{Y} = \mathbb{X} \times \mathbb{X}$, $\mu(\mathrm{d}y) = \bar{\mu}_X \otimes \bar{\mu}_X(\mathrm{d}x, \mathrm{d}x')$, $k(y,a) = \bar{\omega}(x,x')\, \bar{\lambda}(x', t-a)$ and $F(y,a) = \overline{\mathcal{F}}(x',a) \cdot \exp[-\int_0^a \overline{\mathcal{F}}(x',r)\, \mathrm{d}r]$.

It is standard that k is bounded and $\mu \otimes$ *Leb*-a.e. continuous under Assumption 3, 7 and 6. Thanks to (4.3), F is bounded and for any $y_1 = (x_1, x_1') \in \mathcal{Y}$ and $y_2 = (x_2, x_2') \in \mathcal{Y}$:

$$\|F(y_1,.) - F(y_2,.)\|_{L^\infty([0,t])} \leq (1 + \lambda^*\, \omega^*\, t)\, \|\overline{\mathcal{F}}(x_1',.) - \overline{\mathcal{F}}(x_2',.)\|_{L^\infty([0,t])}\,,$$

which entails the required regularity of F thanks to Lemma 4.2.

Proof. For any $y_1, y_2 \in \mathcal{Y}$:

$$\begin{aligned}
&|\int_0^t k(y_1,a)\, F(y_1,a)\, \mathrm{d}a - \int_0^t k(y_2,a)\, F(y_2,a)\, \mathrm{d}a| \\
&\quad \leq \|F\|_\infty\, \|k(y_1,.) - k(y_2,.)\|_{L^1(Leb)} + t\, \|k\|_\infty\, \|F(y_1,.) - F(y_2,.)\|_{L^\infty([0,t])}\,.
\end{aligned}$$

It concludes the proof of Lemma A.3 by virtue of the assumed regularity of k and F, thanks to Lemma A.2 for the first term in k. □

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 $\varepsilon^N(x)$:

$$\varepsilon^N(x) = \Big| \int_{\mathcal{Y}} k(x,y)[\nu^N - \nu](dy) \Big|\,. \tag{A.2}$$

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 $\mu^N(\mathbb{X})$ then tends to zero. In the following, we can thus assume that both $\nu(\mathcal{Y}) > 0$ and $\mu(\mathbb{X}) > 0$.

The irregularities of k will be located through the following subset $\mathcal{G}[\delta,\eta]$ of $\mathbb{X}\times\mathcal{Y}$, defined for any $\delta,\eta>0$:

$$\mathcal{G}[\delta,\eta] := \{(x,y)\in\mathbb{X}\times\mathcal{Y};\ Diam_k(B[(x,y);\eta])>\delta\}, \tag{A.3}$$

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

$$Diam_k(A) := \sup\{|k(z)-k(z')|; z,z'\in A\}, \tag{A.4}$$

while $B[(x,y);\eta]$ denotes the open ball centered in (x,y) of radius η. The size η of the vicinities in (A.3) shall be considered sufficiently small to ensure that discrepancies of order δ are exceptional. The measurability of the set $\mathcal{G}[\delta,\eta]$ can be more directly verified through the following definition, which happens to be equivalent to the one given in (A.3):

$$\mathcal{G}[\delta,\eta] = \cup_{\{q\in\mathbb{Q}_+,m\geq 1\}} k^{-1}([0,q])^\eta \cap k^{-1}([q+\delta+2^{-m},\|k\|_\infty])^\eta,$$

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.

Step 1: Convergence of $\langle\mu,\varepsilon^N\rangle$ to zero.
Since k is $(\mu\otimes\nu)$-a.e. continuous, in particular, $k(x,.)$ is ν-a.e. continuous for x on a measurable set $\mathcal{A}\subset\mathbb{X}$ such that $\mu(\mathcal{A})=1$. For any $x\in\mathcal{A}$, thanks to the Portmanteau theorem, see e.g. [32], 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

$$\lim_{N\to\infty}\langle\mu,\varepsilon^N\rangle = 0. \tag{A.6}$$

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 $(\mu\otimes\nu)$-a.e. continuous, the following convergence to zero holds for any δ:

$$\lim_{\eta\to 0}[\mu\otimes\nu](\mathcal{G}[\delta,\eta]) = 0. \tag{A.7}$$

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:

$$[\mu\otimes\nu^N](\mathcal{G}[\delta,\eta]) \leq [\mu\otimes\nu](\mathcal{G}[\delta,\eta]^\eta)+\eta, \tag{A.8}$$

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.3) that $\mathcal{G}[\delta,\eta]^\eta\subset\mathcal{G}[\delta,2\eta]$. Recalling (A.7) and coming back to (A.8), we have proved the following convergence to zero for any δ:

$$\lim_{\eta\to 0}[\mu\otimes\nu^N](\mathcal{G}[\delta,\eta]) = 0. \tag{A.9}$$

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:

$$\mathcal{H}^N[\delta] := \{x \in \mathbb{X};\ \varepsilon^N(x) \geq \delta\}\,. \tag{A.10}$$

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 $\mathcal{Y}$:

$$\mathcal{G}_x[\delta,\eta] := \{y \in \mathcal{Y};\ (x,y) \in \mathcal{G}[\delta,\eta]\}, \tag{A.11}$$

namely the restriction of $\mathcal{G}[\delta,\eta]$, recall (A.3), with x as the first coordinate.

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

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

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.11), for N sufficiently large, we obtain

$$\begin{aligned}\Big|\int_{\mathcal{G}_x[\delta,\eta]^c} |k(x,y) - k(x',y)|\,[\nu^N - \nu](dy)\Big| &\leq [\nu^N \vee \nu](\mathcal{G}_x[\delta,\eta]^c)\cdot\delta \\ &\leq 2\delta\cdot\nu(\mathcal{Y})\,,\end{aligned} \tag{A.12}$$

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,

$$\Big|\int_{\mathcal{G}_x[\delta,\eta]} |k(x,y) - k(x',y)|\,[\nu^N - \nu](dy)\Big| \leq [\nu^N \vee \nu](\mathcal{G}_x[\delta,\eta]), \tag{A.13}$$

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.2) to combine this result with (A.12) and (A.13), 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), (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

$$\langle \mu^N, \varepsilon^N\rangle \leq 2\theta\,\delta\cdot\mu(\mathbb{X}) + \mu^N(\mathcal{H}^N[\theta\,\delta]), \tag{A.14}$$

where we recall (A.10).

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

$$\mu\Big(\Big\{x \in \mathbb{X};\ \nu(\mathcal{G}_x[\delta,\eta]) \geq \frac{\delta}{2}\Big\}\Big) \leq \frac{2[\mu\otimes\nu](\mathcal{G}[\delta,\eta])}{\delta} \leq \delta. \tag{A.15}$$

Similarly,

$$\mu\Big(\Big\{x \in \mathbb{X};\ \nu^N(\mathcal{G}_x[\delta,\eta]) \geq \frac{\delta}{2}\Big\}\Big) \leq \frac{2[\mu\otimes\nu^N](\mathcal{G}[\delta,\eta])}{\delta} \leq \delta. \tag{A.16}$$

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

$$\mu^N(\mathcal{H}^N[\theta\,\delta]) \le \mu(\mathcal{H}^N[\theta\,\delta]^\eta) + \eta, \tag{A.17}$$

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

$$\mu(\mathcal{H}^N[\theta\,\delta]^\eta) \le \mu(\mathcal{H}^N[\delta]) + \mu(\mathcal{H}^N[\delta]^c \cap \mathcal{H}^N[\theta\,\delta]^\eta). \tag{A.18}$$

Thanks to the Markov inequality, we obtain

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

which converges to 0 as N tends to infinity as stated in (A.6). We thus restrict to N sufficiently large in order to ensure that

$$\mu(\mathcal{H}^N[\delta]) \le \delta. \tag{A.19}$$

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

$$\begin{aligned}&\mu(\mathcal{H}^N[\delta]^c \cap \mathcal{H}^N[\theta\,\delta]^\eta)\\ &\qquad \le \mu\Big(\Big\{x \in \mathbb{X};\ \nu^N(\mathcal{G}_x[\delta,\eta]) \ge \frac{\delta}{2}\Big\}\Big) + \mu\Big(\Big\{x \in \mathbb{X};\ \nu(\mathcal{G}_x[\delta,\eta]) \ge \frac{\delta}{2}\Big\}\Big).\end{aligned} \tag{A.20}$$

For the next upper-bound, valid for η sufficiently small then N sufficiently large, we recall (A.15), (A.16), (A.18),(A.19), (A.20) and get

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

It remains to combine this result with (A.14) and (A.17) to conclude the proof that $(\langle \mu^N, \varepsilon^N \rangle)$ tends to zero 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 [36, Lemma 4.2] to relate the convergence of probability to a.s. convergence of sequence extractions. Let $(N[k])_{k\ge1} \in \mathbb{N}^{\mathbb{N}}$ be an increasing sequence. Since μ^N and ν^N converge in probability, we can extract a subsequence $(\check{N}[\ell])_{\ell\ge1} = (N[K[\ell]])_{\ell\ge1}$ 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 zero 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. □

Conflict of interest The authors declare that they have no conflict of interest.
Funding declaration This research was not supported by any specific funding.

Acknowledgement

We thank the reviewer for the helpful comments that have led to substantial improvements of the exposition of the paper.

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 erdős-rényi graphs. *Latin American Journal of Probability and Mathematical Statistics*, 22(1):245–245, 2025.

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

[16] D. Dawson. Measure-Valued Processes, Stochastic Partial Differential Equations, and Interacting Systems. http://www.ams.org/crmp/005, Mar. 1994.

[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] L. C. Evans. *Partial Differential Equations.* American Mathematical Society, Providence, R.I, 2nd ed. edition, 2010.

[24] J. Fernley, P. Mörters, and M. Ortgiese. The contact process on a graph adapting to the infection. *Stochastic Processes and their Applications*, 183:104596, 2025.

[25] 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.

[26] R. Forien, G. Pang, and É. Pardoux. Epidemic models with varying infectivity. *SIAM Journal on Applied Mathematics*, 81(5):1893–1930, 2021.

[27] 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.
[28] 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.
[29] R. Forien and É. Pardoux. Stochastic epidemic model with varying infectivity and waning immunity: the law of large numbers with unbounded infectivity. *arXiv preprint arXiv:2606.11845*, 2026.
[30] Y. Huang and A. Röllin. The SIR epidemic on a dynamic Erdös-Rényi random graph. *arXiv preprint arXiv:2404.12566*, 2024.
[31] 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.
[32] J. Jacod and A. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer-Verlag, Berlin, 2 edition, 2003.
[33] 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.
[34] 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.
[35] Y. Jiang, R. Kassem, G. York, M. Junge, and R. Durrett. SIR epidemics on evolving graphs. *arXiv preprint arXiv:1901.06568*, 2019.
[36] O. Kallenberg. *Foundations of Modern Probability*. Springer, New York, 2002.
[37] A. Kanga and É. Pardoux. Spatial SIR epidemic model with varying infectivity without movement of individuals: Law of Large Numbers. *Pure and Applied Functional Analysis*, to appear, 2026.
[38] D. Keliger, I. Horváth, and B. Takács. Local-density dependent Markov processes on graphons with epidemiological applications. *Stochastic Processes and their Applications*, 148:324–352, 2022.
[39] 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.
[40] M. Kubasch. Large population limit for a multilayer SIR model including households and workplaces. *arXiv preprint arXiv:2305.17064*, 2023.
[41] 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.
[42] 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.
[43] 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.
[44] G. Pang and É. Pardoux. Functional limit theorems for non-Markovian epidemic models. *The Annals of Applied Probability*, 32(3):1615–1665, 2022.
[45] G. Pang and É. Pardoux. Functional central limit theorems for epidemic models with varying infectivity. *Stochastics*, 95(5):819–866, 2023.
[46] 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.
[47] G. Pang and É. Pardoux. SPDE for stochastic SIR epidemic models with infection-age dependent infectivity. *Stochastics and Partial Differential Equations: Analysis and Computations*, 2025. doi.org/10.1007/s40072-025-00403-x.
[48] G. Pang and É. Pardoux. Spatially dense stochastic epidemic models with infection-age dependent infectivity. *Stochastic Analysis and Applications*, 44(1):65–116, 2026.
[49] E. Perkins. Dawson-Watanabe superprocesses and measure-valued diffusions. In *Lectures on probability theory and statistics (Saint-Flour, 1999)*, volume 1781 of *Lecture Notes in Math.*, pages 125–324. Springer, Berlin, 2002.
[50] G. Pólya and G. Szegö. *Problems and Theorems in Analysis I: Series, integral calculus, theory of functions*. Springer Verlag, Berlin, Heidelberg, 1978.
[51] G. Reinert. The asymptotic evolution of the general stochastic epidemic. *The Annals of Applied Probability*, 5(4):1061–1086, 1995.
[52] Y. Shang. Mixed SI (R) epidemic dynamics in random graphs with general degree distributions. *Applied Mathematics and Computation*, 219(10):5042–5048, 2013.

[53] K. Spricer and T. Britton. An SIR epidemic on a weighted network. *Network Science*, 7(4):556–580, 2019.
[54] E. Volz. SIR dynamics in random networks with heterogeneous connectivity. *Journal of Mathematical Biology*, 56:293–310, 2008.
[55] F. J. Wang. Limit theorems for age and density dependent stochastic population models. *Journal of Mathematical Biology*, 2(4):373–400, 1975.
[56] 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