



HAL
open science

Large population limit for a multilayer SIR model including households and workplaces

Madeleine Kubasch

► **To cite this version:**

Madeleine Kubasch. Large population limit for a multilayer SIR model including households and workplaces. 2023. hal-04106800v2

HAL Id: hal-04106800

<https://hal.science/hal-04106800v2>

Preprint submitted on 31 Jul 2023 (v2), last revised 13 Oct 2023 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Large population limit for a multilayer *SIR* model including households and workplaces

Madeleine Kubasch^{1,2}

¹Centre de mathématiques appliquées (CMAP), Ecole Polytechnique, 91128 Palaiseau, France

²MaIAGE, INRAE, Université Paris-Saclay, 78350 Jouy-en-Josas, France

July 31, 2023

Abstract

We study a multilayer *SIR* model with two levels of mixing, namely a global level which is uniformly mixing, and a local level with two layers distinguishing household and workplace contacts, respectively. We establish the large population convergence of the corresponding stochastic process. For this purpose, we use an individual-based model whose state space explicitly takes into account the duration of infectious periods. This allows to deal with the natural correlation of the epidemic states of individuals whose household and workplace share a common infected. In a general setting where a non-exponential distribution of infectious periods may be considered, convergence to the unique deterministic solution of a measure-valued equation is obtained. In the particular case of exponentially distributed infectious periods, we show that it is possible to further reduce the obtained deterministic limit, leading to a closed, finite dimensional dynamical system capturing the epidemic dynamics. This model reduction subsequently is studied from a numerical point of view. We illustrate that the dynamical system derived from the large population approximation is a pertinent model reduction when compared to simulations of the stochastic process or to an alternative edge-based compartmental model, both in terms of accuracy and computational cost.

Keywords. Epidemic process, household-workplace models, two layers of mixing, large population limit, agent-based model, model reduction.

Code availability. <https://github.com/m-kubasch/household-workplace-model>

Introduction

Among the possibilities of non-pharmaceutical epidemic control measures, social distancing may be used to slow down disease propagation. This has been emphasized by the recent COVID-19 epidemic, which lead to wide-spread implementation of non-pharmaceutical interventions. In particular, both empirical evidence [28] and simulation studies [36] have pointed out that school closure and teleworking are among the most effective of those measures. These results are not only limited to COVID-19, as it is equivalently known that particularly school closure mitigates influenza epidemics [25]. It thus seems relevant to take into account this kind of control measure in epidemic models, in order to study their impact on epidemic spread. A difficulty then arises from the fact that social distancing often disrupts particular contacts, such as workplace contacts only in the case of teleworking, while leaving intact other types of contact, for instance within households. Di Lauro *et al.* [11] have shown through the example of lockdowns that in order to model control measures impacting preferably specific types of contact, it is advantageous for the model to explicitly distinguish these contact structures. In addition, even though it is well known

that clustered contact patterns influence epidemic spread [13, 41], precisely understanding the impact on disease propagation of the way individuals are organized in households and workplaces is not straightforward [26, 5].

This motivates the study of models which explicitly distinguish different levels of contact, which as we shall see lead to interesting mathematical issues due to their multiscale population structure. A minimal model taking into account households and workplaces has been proposed by Pellis *et al.* [32], referred to as the household-workplace model. This model consists in a stochastic *SIR* model with two levels of mixing [8], a *global* and a *local* one, the latter being characterized by the presence of small contact structures. More precisely, contacts may take place within three different layers: households, workplaces and the general population. Households and workplaces correspond to the local level of mixing, while the general population constitutes the global one. In particular, it is possible to distinguish different rates at which acquaintances encounter one another within each of these three layers. This model has already been studied to some extent, establishing for instance different reproduction numbers, see *e.g.* Ball *et al.* [3] for an extensive overview. In particular, the reproduction number R_I proposed by Pellis *et al.* [32] is interesting, as it is intimately connected to the proportion of infections taking place within each layer of the contact network [5]. The R_0 used throughout this paper will thus refer to this particular reproduction number, and computations of R_0 and proportions of infections per layer will make use of the working package associated to [5].

One drawback of this model is its complexity, both mathematically and numerically. Indeed, it is not simple to analyse due to correlations arising as soon as an individual may belong to several small contact structures at once. Also, simulations require a significant amount of computation time, especially when considering larger population sizes. As a consequence, it is of interest to develop reduced models, which may be more prone to theoretical studies and/or numerical exploration. In particular, large population approximations of stochastic models have proven fruitful to achieve such model reductions in many contexts, among which epidemics on random graphs.

Historically, the standard *SIR* model developed by Kermack and McKendrick itself corresponds to the large population limit of the uniformly mixing stochastic *SIR* model. In the Markovian setting, the convergence of the stochastic model to its deterministic limit can be established using classical results on the convergence of finite type density-dependent Markov jump processes, *e.g.* [2]. In a more general setting, where infectious periods are not restricted to being exponentially distributed, the large population convergence of the stochastic model to the unique deterministic solution of a system of integral equations can also be obtained (see for instance [23] and references therein).

From the contact network point of view, uniform mixing is represented by complete graphs. One may notice that in order to describe an epidemic spreading on a graph, it is key to get a good understanding of the neighbourhoods of susceptible nodes, as they can only be infected through one of their neighbours. The complete graph hence is a favourable setting, as the neighbourhood of each node corresponds to the total population: the amount of information necessary to describe the infectious pressure on each susceptible thus reduces to the proportion of infected in the general population. For more general contact networks however, getting a picture of the neighbourhood of susceptible nodes over time is often difficult. As a consequence, it can be challenging to propose closed systems of equations correctly describing the epidemic dynamics. Several families of reduced models thus have evolved, including pairwise models (see [21] for a review), message passing [20], effective degree models [24] and edge-based compartmental models (EBCM) [40].

The latter have originally been developed in the particular case of an *SIR* model on configuration graphs, which have the property of being locally tree-like when the number of nodes grows to infinity. In other words, in the large population limit, there are no finite loops, and thus

the epidemic states of the neighbours of a susceptible node can be thought of as independent from one another. This is the key ingredient to the original EBCM which was first proposed by [40], and refined by [30]. Notice that this model reduction is especially parsimonious, as a handful of equations are enough to entirely describe the epidemic dynamics. Also, this model has been proven to be the large population limit of the underlying stochastic model [9, 17]. Since then, the equivalence with other reduced models has been established under appropriate assumptions [15, 42, 16, 22], which are thus equally exact in the large population limit thanks to the aforementioned convergence result. Note that the EBCM formalism has also been extended to include pertinent variations of the original model, such as non-Markovian settings [35], or an epidemic spreading on a dynamic multi-layer configuration model [16].

While the case of classical configuration models is thus well understood, taking into account the presence of clusters within the contact network raises a new challenge. In particular, some attention has been drawn to the case of small, highly connected groups referred to as *cliques*, which are well suited to represent structures such as households or workplaces. Several models have been proposed in the case where each individual belongs to a certain (random) number of such cliques which can be of random size. Such a network can for instance be obtained by extending the configuration model to include cliques, and an EBCM has been developed in this setting [41]. Alternative models are presented in [37], where contact rates are allowed to differ between cliques as random functions of their size, as well as [13], where cliques are not necessarily fully connected. Let us emphasize that these models share three major common points: they focus on the epidemic at the level of structures, as they keep track of the proportions of cliques containing a certain number of susceptibles and infected; the proposed reduced systems are high-dimensional for larger clique sizes; and for none of them, the convergence of the underlying stochastic model to the proposed reduced model has been established.

As already motivated in the case of our model, it further is of interest to distinguish several types of contacts, giving rise to multi-layer contact networks. Layers differ in terms of contact rates, and potentially also in terms of contact networks. To our knowledge, when considering small complete structures, this has first been achieved for household models. These models consist in distinguishing household contacts from contacts within the general population, each individual belonging to exactly one household. Such a model was proposed in [14], with a uniformly mixing general population, in which case the reduced model naturally appears to be the large population limit of its stochastic counterpart, which can again be formalized as a finite type density-dependent Markov jump process. Other possibilities include considering that general contacts form a configuration network [11, 26], or allowing individuals to belong to several cliques of the same type instead of only one household [6].

However, one may notice that the particular case of the household-workplace model is not covered by the previously introduced models, as it is necessary here to consider two layers of cliques. In this setting, the only reduced models proposed so far approach the epidemic dynamics using well calibrated uniformly mixing models [5, 10]. While these are capable of capturing some key characteristics of the epidemic, such as the epidemic peak size and final size, they do not allow for an accurate prediction of the epidemic dynamic over time.

As a consequence, in this paper, we will study the large population limit of the model with two levels of mixing which explicitly distinguishes households and workplaces. In order to do so, we will formalize the model in a finite population as an individual-based stochastic process, and establish that this sequence of processes converges in law when the size of the population grows to infinity. This allows, on the one hand, to identify our model reduction, and on the other hand establishes that this model reduction is legitimate since it is asymptotically exact. Besides, it paves the way for more quantitative estimates on this approximation. Notice here that each infected individual correlates the epidemic spread in his household and workplace, being infectious for exactly the same period of time in both structures. In order to deal with

this dependence, the duration of infectious periods will explicitly be taken into account in the model structure. This difficulty actually arises as soon as one considers the possibility of an individual belonging to several cliques at once, whether they are of different types or not, and we refer to [4] where a similar approach has been developed for branching approximations. This model formulation amounts to studying a measure-valued process mixing discrete and continuous components. More precisely, we establish the convergence of the individual-based process to the unique solution of an explicit measure-valued equation. In the particular case where this distribution is exponential, it is possible to go one step further and reduce the epidemic dynamics to a closed, finite dimensional dynamical system which is similar in spirit to reductions proposed in related settings [14, 37].

The present paper is structured as follows. Section 1 introduces the individual-based model, and Section 2.1 subsequently presents the convergence results in detail. Section 2.2 is devoted to numerical aspects. We first illustrate that the obtained dynamical system is in good accordance with stochastic simulations, discuss its implementation and examine its computational cost in terms of computation time compared to stochastic simulations. Next, we confront our reduced model to an alternative model reduction which we obtain using the EBCM formalism. Finally, Section 3 contains the proofs of our results.

Before proceeding, let us introduce some notations that will be used throughout the paper. For any integers $n \leq m$, we write $\llbracket n, m \rrbracket = \{n, \dots, m\}$. For a measurable space (E, \mathcal{E}) , let $\mathcal{M}_P(E)$ be the set of point measures, $\mathcal{M}_F(E)$ the set of finite measures and $\mathcal{M}_1(E)$ the set of probability measures on E . We define $\mathcal{M}_{P,1}(E) = \mathcal{M}_P(E) \cap \mathcal{M}_1(E)$ the set of punctual probability measures on E . For a measure μ on E and a suitable function f (either non-negative or belonging to $L^1(\mu)$), let $\langle \mu, f \rangle = \int_E f d\mu$. Also, for $x \in E$, δ_x designates the Dirac measure at point x . Further, for any metric space E and any integer m , let $\mathcal{C}(E, \mathbb{R}^m)$ be the set of continuous functions $f : E \rightarrow \mathbb{R}^m$. Similarly, $\mathcal{C}_b(E, \mathbb{R}^m)$ is defined as the subset of bounded functions $f \in \mathcal{C}(E, \mathbb{R}^m)$. Finally, the space $\mathcal{C}_b^1(E, \mathbb{R}^m)$ designates the set of bounded functions $f : E \rightarrow \mathbb{R}^m$ such that f is differentiable and its differential is continuous and bounded.

1 Presentation of the model

Let us begin by introducing the epidemic model of interest, in two successive steps. At first, a general model description is yielded, which corresponds to a more intuitive presentation of the model, before stating the mathematical model in detail using a measure-valued stochastic differential equation.

1.1 General presentation of the model

Let us start by describing the population structure of interest. Consider a population of K individuals. Each individual is part of exactly one household and one workplace, which are chosen independently from one another, and independently for each individual.

More precisely, such a population structure can be obtained as described in [5]. Suppose that households and workplaces are of size at least one and at most n_{\max} . Consider distributions (π_j^H) and (π_j^W) on $\llbracket 1, n_{\max} \rrbracket$. These distributions correspond to the large population limit of household and workplace size distributions, in the sense that in an infinite population, a proportion π_j^H of households would be of size j , while π_j^W would play a similar role for workplaces. In such an infinite population, the average household and workplace sizes, respectively m_H and m_W , would

be given for $X \in \{H, W\}$ by

$$m_X = \sum_{j=0}^{n_{\max}} j \pi_j^X.$$

In order to assign each individual of the finite-size population a household, we proceed as follows. Let $k \in \llbracket 0, K \rrbracket$ be the number of individuals who are not yet member of a household. While $k > 0$, choose a size \tilde{n} according to π^H , independently from the household sizes that were chosen during previous steps. The newly uncovered household is then of size $n = \min(\tilde{n}, k)$, and n individuals out of the k remaining ones are picked uniformly at random to assemble this new household. Consequently, it remains to update k to $k - n$. The process stops as soon as $k = 0$, as all individuals then belong to a household. Finally, this process is repeated independently for workplaces, using π^W instead of π^H . Throughout the paper, let us consider one particular sequence $(\mathbf{G}^K)_{K \geq 1}$ of realisations of this random population structure. The choice of this sequence being arbitrary among all sequences of such population structures, this does not lessen the generality of our results.

It remains to describe the way the disease spreads in the population. The epidemic model considered here is an extension of the standard *SIR* model. At each time, each individual is either *susceptible* if he has never encountered the disease and may be contaminated; *infected* if the individual is currently infectious, in which case he may transmit the disease to other susceptibles; or *recovered*, once the infectious period is over, in which case the individual has become immune against the disease.

The disease is transmitted among individuals as follows. Any contact between a susceptible and an infected individual within either level of mixing amounts in the contamination of the susceptible individual. Such an event is called an *infectious contact*. Within each household, each workplace and the general population, uniform mixing is assumed, but the parameterization differs slightly between the layers. Indeed, for households, we consider a *one-to-one* contact rate λ_H , meaning that whenever there are s susceptibles and i infected within a household, the next infectious contact occurs at rate $\lambda_H si$. Similarly, another one-to-one contact rate λ_W is associated to workplace contacts. Within the general population, on the other hand, a *one-to-all* contact rate β_G is considered: when there are s susceptibles and i infected within a population of size K , infectious contacts occur at rate $\frac{\beta_G}{K} si$. Here, contrary to the local level, frequency-dependence is considered for the global contact rate, as each given encounter between two individuals chosen at random within the general population becomes less likely when the population size K grows large.

Finally, infected individuals remain infectious for a period of time distributed according to a probability distribution ν on \mathbb{R}_+ , which we assume to be absolutely continuous with regard to the Lebesgue measure. Once they recover, they are supposed to be immune against the disease from there on. In particular, if ν is an exponential distribution, this corresponds to the Markovian *SIR* model.

1.2 The epidemic model at the level of households and workplaces

As we aim at investigating the large population limit of this model, we choose to enrich the population description as to obtain a closed Markov process. This corresponds to a favourable mathematical setting, as it allows us to use the associated martingale problem. In order to do so, we will represent the population in terms of particles which are described by a type. It seems natural to consider particles which correspond to entire structures, *i.e.* households and workplaces, which are characterized by their size and the number of susceptible and infected individuals they contain. Indeed, this point of view has already proven useful for deriving

reproduction numbers for related models [3], as well as the epidemic growth rate of the household-workplace model [33]. However, this is not enough to obtain a closed system of Markovian dynamics. The problem is that each infected individual correlates the spread of the epidemic within his household and his workplace, leading to an intricate correlation network. In order to circumvent this difficulty, similarly to [4], we will thus further characterize each structure by the infectious periods of its infected members.

More precisely, for a population of size $K \geq 1$, let K_H be the number of households and K_W the number of workplaces in \mathbf{G}^K . Label the K_H households in an arbitrary fashion $1, \dots, K_H$. Consider the set

$$E = \{(n, s, \tau) \in \llbracket 1, n_{\max} \rrbracket \times \llbracket 0, n_{\max} \rrbracket \times \mathbb{R}^{n_{\max}} : s \leq n; \forall j > n - s, \tau_j = 0\}.$$

Then for $k \in \llbracket 1, K_H \rrbracket$, the k -th household is characterized at time $t \geq 0$ by its type

$$x_k^H(t) = (n_k^H, s_k^H(t), \tau_k^H(t)) \in E.$$

The first two components of x_k^H correspond respectively to the size of the household (which is constant over time), and the number of susceptible members of the household at time t . The third component τ_k^H is a vector containing the *remaining infectious periods* of the members of the household and keeps track of infected and removed individuals. Indeed, at time t , there are $n_k^H - s_k^H(t)$ infected or removed individuals within the household. For $j \in \llbracket 1, n_k^H - s_k^H(t) \rrbracket$, $\tau_{k,j}^H(t)$ corresponds to the remaining duration of the infectious period at time t of the j -th member of the household who has contracted the disease. If $\tau_{k,j}^H(t) > 0$, the individual is still infectious and will remain so for $\tau_{k,j}^H(t)$ units of time. Otherwise, if $\tau_{k,j}^H(t) \leq 0$, the individual has recovered, and the recovery has occurred at time $t - |\tau_{k,j}^H(t)|$. For $j > n_k^H - s_k^H(t)$, $\tau_{k,j}^H(t)$ is set to zero and can be neglected for interpretation. Notice that, in principle, it would have been possible to define $\tau_k^H(t) \in \mathbb{R}^{n_k^H - s_k^H(t)}$, but letting $\tau_k^H \in \mathbb{R}^{n_{\max}}$ is convenient for computations.

Similarly, label the K_W workplaces in an arbitrary order $1, \dots, K_W$. For $\ell \in \llbracket 1, K_W \rrbracket$, the ℓ -th workplace is characterized by its type $x_\ell^W(t) = (n_\ell^W, s_\ell^W(t), \tau_\ell^W(t))$, which is defined analogously to household types.

By definition, these types evolve over time. On the one hand, for any $X \in \{H, W\}$, for any $k \in \llbracket 1, K_X \rrbracket$ and $j \in \llbracket 1, n_{\max} \rrbracket$, the j -th component of τ_k^X decreases linearly at unitary rate if it describes the remaining infectious period of an individual having contracted the disease at some previous time, and stays constant otherwise:

$$\forall j \in \llbracket 1, n_{\max} \rrbracket, \frac{d}{dt} \tau_{k,j}^X(t) = -\mathbf{1}_{\{j \leq n_k^X - s_k^X(t)\}}.$$

Let $(e_j)_{1 \leq j \leq n_{\max}}$ denote the canonical basis of $\mathbb{R}^{n_{\max}}$. Then for any $0 \leq t \leq T$, and $x = (n, s, \tau) \in E$, we may define $\Psi(x, T, t)$ as the type of a structure at time T given that it was in state x at time t , supposing that no infections occurred in the meantime:

$$\Psi(x, T, t) = \left(n, s, \tau - \sum_{j=1}^{n-s} (T-t)e_j \right).$$

On the other hand, infections within each level of mixing also cause the modification of the types of the household and the workplace of newly contaminated individuals. More precisely, consider a contamination occurring at time t . Suppose that the newly infected belongs to the k -th household and ℓ -th workplace. Let σ be the realisation of a random variable of distribution

ν , which is drawn independently for each new infected. Then x_k^H and x_ℓ^W jump from $x_k^H(t-)$ and $x_\ell^W(t-)$ to $j(x_k^H(t-), \sigma)$ and $j(x_\ell^W(t-), \sigma)$ respectively, where for any $x = (n, s, \tau) \in E$,

$$j(x, \sigma) = (n, s - 1, \tau + \sigma e_{n-s+1}).$$

It remains to describe how one identifies the household k and workplace ℓ the newly infected belongs to. Let $S(t-)$ be the number of susceptibles in the population previously to the infection event. If it takes place within the general population, any susceptible individual is chosen with uniform probability to be contaminated. The newly infected thus belongs to the k -th household with probability $s_k^H(t-)/S(t-)$, and independently to the ℓ -th workplace with probability $s_\ell^W(t-)/S(t-)$. Similarly, if the infection occurs within a household, only the workplace of the newly infected needs to be uncovered, and corresponds to the ℓ -th workplace with the same probability as previously. Within-workplace infections are treated analogously.

We are now ready to properly introduce the stochastic process which represents the previously described dynamics. We are interested in the stochastic process $(\zeta_t^K = (\zeta_t^{H|K}, \zeta_t^{W|K}))_{t \geq 0}$ taking values in $\mathfrak{M}_{P,1} = \mathcal{M}_{P,1}(E) \times \mathcal{M}_{P,1}(E)$, whose definition will be detailed hereafter. Here, $\zeta_t^{H|K}$ and $\zeta_t^{W|K}$ correspond respectively to the normalized counting measures associated to the distributions of household and workplace types at time t , *i.e.* for any time $t \geq 0$ and $X \in \{H, W\}$,

$$\zeta_t^{X|K} = \frac{1}{K_X} \sum_{k=1}^{K_X} \delta_{x_k^X(t)}.$$

Start by noticing that, as both household and workplace sizes are bounded, for $X \in \{H, W\}$, the following inequality holds:

$$\frac{K}{n_{\max}} \leq K_X \leq K.$$

Thus, studying the asymptotic $K \rightarrow \infty$ amounts to $(K_H, K_W) \rightarrow (\infty, \infty)$.

Recall that in the case of infections occurring within the general population, the rate of infections depends on the number of susceptibles $S(t)$ and infected $I(t)$ within the whole population at time $t \geq 0$. It is easy to compute these quantities using ζ^K as follows. For any $\tau \in \mathbb{R}^{n_{\max}}$, let $i(\tau) = \sum_{k=1}^{n_{\max}} \mathbf{1}_{\{\tau_k > 0\}}$. Then for any $t \geq 0$,

$$I^H(t) = \frac{1}{K_H} \sum_{k=1}^{K_H} i(\tau_k^H(t))$$

corresponds to the average number of infected individuals per household at time t . Similarly, one may define $S^H(t)$ as the average number of susceptibles per household at time t . Then

$$\forall X \in \{H, W\}, S(t) = K_X S^X(t) \text{ and } I(t) = K_X I^X(t). \quad (1)$$

Further, let N^H be the average household size, which is constant over time and always equal to K/K_H . This leads to $I(t)/K = I^X(t)/N^X$, which will be of use in computations. Notice that we will need to check that Equation (1) is well posed, as equalities of the type $K_H S^H(t) = K^W S^W(t)$ technically need to be proven for the stochastic process formalizing the model.

Finally, let us briefly emphasize that the partition of the population in households and workplaces is entirely conveyed by $\zeta_0^K \in \mathfrak{M}_{P,1}$, as it does not vary over time. In particular, the proportions of households and workplaces of each size are supposed to correspond to those observed in \mathbf{G}^K . Similarly, there are some natural constraints on ζ_0^K , as $\zeta^{H|K}$ and $\zeta^{W|K}$ describe the same population, once dispatched into households, and once dispatched into workplaces. As these assumptions are intuitive, they will not all be detailed here, and they will only be

emphasized when necessary. For instance, the total number of members in some epidemic state (susceptible, infected or recovered) within all households is equal to the total number of members in this state within all workplaces. Hence $K_H S^H(0) = K_W S^W(0)$ and $K_H I^H(0) = K_W I^W(0)$ almost surely. Further, at time 0, each infected needs to have the same remaining infectious period in both his household and workplace. In other words, almost surely,

$$\{\tau_{k,j}^H(0) : 1 \leq k \leq K_H, 1 \leq j \leq n_k^H - s_k^H(0)\} = \{\tau_{\ell,j}^W(0) : 1 \leq \ell \leq K_W, 1 \leq j \leq n_\ell^W - s_\ell^W(0)\}. \quad (2)$$

Forthcoming Lemma 1.2 shows that these conditions are enough to ensure that the previous characterization of $S(t)$ and $I(t)$ in terms of $S^X(t)$ and $I^X(t)$ is legitimate.

Before giving the proper definition of ζ^K , let us introduce some necessary notations. Let $U_K^G = (\mathbb{R}_+)^3 \times \llbracket 1, K_H \rrbracket \times \llbracket 1, K_W \rrbracket \times \mathbb{R}_+$, and consider the following measure on U_K^G :

$$\mu_K^G(du) = \mu_K^G(d\boldsymbol{\theta}, dk, d\ell, d\sigma) = d\boldsymbol{\theta} \otimes \mu_{\#}(dk) \otimes \mu_{\#}(d\ell) \otimes \nu(d\sigma),$$

where $d\boldsymbol{\theta}$ and $\mu_{\#}$ denote the Lebesgue measure on \mathbb{R}^3 and the standard counting measure, respectively. For $t \geq 0$ and $u = (\boldsymbol{\theta}, k, \ell, \sigma) \in U_K^G$ where $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$, let us define

$$\mathcal{I}_G(t, u) = \mathbf{1}_{\left\{ \theta_1 \leq \frac{\beta_G}{K} S(t) I(t), \theta_2 \leq \frac{s_k^H(t)}{S(t)}, \theta_3 \leq \frac{s_\ell^W(t)}{S(t)} \right\}}.$$

The idea is that \mathcal{I}_G will yield the correct rate for infection events in the general population. More precisely, the constraint on θ_1 corresponds to the rate of infectious contacts at that level of mixing, while the constraints on θ_2 and θ_3 are related to the probability that the newly infected belongs to the k -th household and ℓ -th workplace.

Similarly, let $U_K = (\mathbb{R}_+)^2 \times \llbracket 1, K_H \rrbracket \times \llbracket 1, K_W \rrbracket \times \mathbb{R}_+$, endowed with the measure

$$\mu_K(du) = \mu_K(d\boldsymbol{\theta}, dk, d\ell, d\sigma) = d\boldsymbol{\theta} \otimes \mu_{\#}(dk) \otimes \mu_{\#}(d\ell) \otimes \nu(d\sigma),$$

where with slight abuse of notation, $d\boldsymbol{\theta}$ designates the Lebesgue measure on \mathbb{R}^2 . Then for $t \geq 0$ and $u = (\boldsymbol{\theta}, k, \ell, \sigma) \in U_K$ where $\boldsymbol{\theta} = (\theta_1, \theta_2)$, we further introduce

$$\mathcal{I}_H(t, u) = \mathbf{1}_{\left\{ \theta_1 \leq \lambda_H s_k^H(t) i(\tau_k^H(t)), \theta_2 \leq \frac{s_\ell^W(t)}{S(t)} \right\}}.$$

This time, the constraints on θ_1 and θ_2 correspond respectively to the rate of infection within the k -th household and the probability of the newly infected belonging to the ℓ -th workplace. Also, for any $T \geq t \geq 0$ and $u = (\boldsymbol{\theta}, k, \ell, \sigma) \in U_K^G \cup U_K$, consider the following quantity, which will allow to keep track of the change in the household population due to an infection within the k -th household:

$$\Delta_H(u, T, t) = \delta(\Psi_{(j(x_k^H(t-), \sigma), T, t)}) - \delta(\Psi_{(x_k^H(t-), T, t)}).$$

Finally, define $\mathcal{I}_W(t, u)$ and $\Delta_W(u, T, t)$ analogously, by interchanging H and W as well as k and ℓ :

$$\mathcal{I}_W(t, u) = \mathbf{1}_{\left\{ \theta_1 \leq \lambda_W s_\ell^W(t) i(\tau_\ell^W(t)), \theta_2 \leq \frac{s_k^H(t)}{S(t)} \right\}} \text{ and } \Delta_W(u, T, t) = \delta(\Psi_{(j(x_\ell^W(t-), \sigma), T, t)}) - \delta(\Psi_{(x_\ell^W(t-), T, t)}).$$

In order to simplify notations in the following, for $Y \in \{H, W\}$, we let $U_K^Y = U_K$ and $\mu_K^Y = \mu_K$. We are now ready to yield the main characterization of ζ_t^K , as inspired by [38].

Proposition 1.1. Define on the same probability space as ζ_0^K , and independently from ζ_0^K , three independent Poisson point measures $Q^{Y|K}$ on $\mathbb{R}_+ \times U_K^Y$ with intensity $dt\mu_K^Y(du)$, for $Y \in \{H, W, G\}$.

Then $\zeta^K = (\zeta^{H|K}, \zeta^{W|K})$ is defined as the unique strong solution taking values in $\mathbb{D}(\mathbb{R}_+, \mathfrak{M}_{P,1})$ of the following equation. For $X \in \{H, W\}$,

$$\zeta_T^{X|K} = \frac{1}{K_X} \left(\sum_{j=1}^{K_X} \delta_{\Psi(x_j^X(0), T, 0)} + \sum_{Y \in \{H, W, G\}} \int_0^T \int_{U_K^Y} \mathcal{I}_Y(t-, u) \Delta_X(u, T, t) Q^{Y|K}(dt, du) \right). \quad (3)$$

The idea behind Equation (3) goes as follows. Let us focus for example on the distribution $\zeta_T^{H|K}$ of household types at time T . Each individual household's type contributes to the distribution at uniform weight $1/K_X$. If no infection event occurs between times 0 and T , then the state of the k -th household at time T is given by $\Psi(x_k^H(0), T, 0)$. However, suppose now that before time T , at least one initially susceptible member of the k -th household is infected, and let t be the first time at which such an event occurs. Then $x_k^H(t) = j(x_k^H(t-), \sigma)$ where σ is distributed according to ν . If no other infections affect this household up to time T , it will be in state $\Psi(j(x_k^H(t-), \sigma), T, t)$ instead of $\Psi(x_k^H(0), T, 0) = \Psi(x_k^H(t-), T, t)$. This reasoning is reflected in Δ_H , and can be iterated over the whole of $[0, T]$. Finally, the terms \mathcal{I}_Y for $Y \in \{G, H, W\}$ assure that all infection events occur at the corresponding rates.

Proof. The proof uses classical arguments, which will only be outlined here. Start by establishing existence of ζ^K . Consider the sequence $(T_n)_{n \geq 0}$ of successive jump times of ζ^K , where we define $T_0 = 0$. Then using a method similar to rejection sampling, $(T_n)_{n \geq 0}$ can be obtained as a subsequence of the jump times of a Poisson process with intensity $n_{\max}(\lambda_H n_{\max} + \lambda_W n_{\max} + \beta_G)K$, whose only limiting value is $+\infty$. Thus $\lim_{n \rightarrow +\infty} T_n = +\infty$ almost surely, ensuring that ζ^K takes values in $\mathbb{D}(\mathbb{R}_+, \mathfrak{M}_{P,1}(E))$.

Finally, uniqueness is obtained by an induction argument which proves that for any $n \geq 0$, $(T_n, \zeta_{T_n}^K)$ is uniquely determined by $(\zeta_0^K, (Q^{Y|K})_{Y \in \mathcal{S}})$ where $\mathcal{S} = \{G, H, W\}$. This obviously is true for $n = 0$. The induction step relies on the observation that T_{n+1} is uniquely determined by $(T_n, \zeta_{T_n}^K, (Q^{Y|K})_{Y \in \mathcal{S}})$ and $\zeta_{T_{n+1}}^K$ by $(T_{n+1}, T_n, \zeta_{T_n}^K, (Q^{Y|K})_{Y \in \mathcal{S}})$. The induction hypothesis allows to conclude. \square

Let us briefly show here that it follows directly from Proposition 1.1 that Equation (1) is well posed.

Lemma 1.2. Suppose that almost surely, $K_H S^H(0) = K_W S^W(0)$ and Equation (2) holds. Then for any $t \geq 0$, $K_H S^H(t) = K_W S^W(t)$ and $K_H I^H(t) = K_W I^W(t)$, almost surely.

Proof. Let $T \geq 0$, and for any $x = (n, s, \tau) \in E$, let $\mathbf{n}(x) = n$, $\mathbf{s}(x) = s$ and $\mathbf{i}(x) = i(\tau)$. Start by focusing on $S^X(T) = \langle \zeta_T^{X|K}, \mathbf{s} \rangle$ for $X \in \{H, W\}$. It follows from Equation (3) that

$$K_X S^X(T) = \sum_{j=1}^{K_X} \mathbf{s}(\Psi(x_j^X(0), T, 0)) + \sum_{Y \in \{H, W, G\}} \int_0^T \int_{U_K^Y} \mathcal{I}_Y(t-, u) \langle \Delta_X(u, T, t), \mathbf{s} \rangle Q^{Y|K}(dt, du).$$

Notice that on the one hand, for any $x \in E$ and $0 \leq t \leq T$, $\mathbf{s}(\Psi(x, T, t)) = \mathbf{s}(x)$. Hence the first term of the right-hand side equals $K_X S^X(0)$, and for any $u = (\boldsymbol{\theta}, k, \ell, \sigma)$,

$$\langle \Delta_H(u, T, t), \mathbf{s} \rangle = \mathbf{s}(\Psi(j(x_k^H(t-), \sigma), T, t)) - \mathbf{s}(\Psi(x_k^H(t-), T, t)) = -1.$$

The analogous computation yields $\langle \Delta_W(u, T, t), \mathbf{s} \rangle = -1$. Thus $K_H S^H(T) = K_W S^W(T)$ almost surely.

Let us now turn to $I^X(T) = \langle \zeta_T^{X|K}, \mathbf{i} \rangle$. This time, for any $u = (\boldsymbol{\theta}, k, \ell, \sigma)$ and $0 \leq t \leq T$, it holds that $\langle \Delta_X(u, T, t), \mathbf{i} \rangle = \mathbf{1}_{\{\sigma > (T-t)\}}$. Finally, Equation (2) ensures that

$$\sum_{k=1}^{K_H} \mathbf{i}(\Psi(x_k^H(0), T, 0)) = \sum_{k=1}^{K_H} \sum_{j=1}^{n_k^H - s_k^K(0)} \mathbf{1}_{\{\tau_{k,j}^H(0) > T\}} = \sum_{\ell=1}^{K_W} \mathbf{i}(\Psi(x_\ell^W(0), T, 0)).$$

The conclusion follows as previously from Equation (3). \square

2 Main results

In this section, we are going to present our main results on the convergence of $(\zeta^K)_{K \geq 1}$ in the Skorokhod space $\mathbb{D}(\mathbb{R}_+, \mathcal{M}_1(E))^2$. For $X \in \{H, W\}$, let \bar{X} be the complementary structure type, *i.e.* $\bar{X} = W$ if $X = H$ and vice-versa.

The following assumption on the sequence of initial conditions $(\zeta_0^K)_{K \geq 1}$ will be required from now on.

Assumption 2.1. *For any $X \in \{H, W\}$ and $T \geq 0$, suppose that:*

(i)

$$\lim_{N \rightarrow \infty} \sup_{K \geq 1} \mathbb{E} \left[\sup_{0 \leq t \leq T} \frac{1}{K_X} \sum_{k=1}^{K_X} \sum_{i=1}^{n_{\max}} \mathbf{1}_{\{n_k^X - s_k^X(0) \geq i, |\tau_{k,i}^X(0) - t| \geq N\}} \right] = 0.$$

(ii) *For any $c \in \mathbb{R}$, for any $i \in \llbracket 1, n_{\max} \rrbracket$,*

$$\lim_{\epsilon \rightarrow 0} \sup_{K \geq 1} \mathbb{E} \left[\frac{1}{K_X} \sum_{k=1}^{K_X} \mathbf{1}_{\{n_k^X - s_k^X(0) \geq i, |(\tau_{k,i}^X(0) - T) - c| \leq \epsilon\}} \right] = 0.$$

Briefly, the first assumption allows to control the impact of the initial condition on the queues of the distribution of remaining infectious periods at each time, while the second condition is related to aspects of absolute continuity. These conditions are for instance satisfied if for any $K \geq 1$, at time 0, the remaining infectious periods of infected individuals are i.i.d. of law ν , while those of recovered individuals are set to be equal to zero. Notice that this choice for recovered individuals does not represent a loss of generality, as it does not affect the epidemic spread and initially recovered individuals will remain recognizable at any time T as the only ones whose remaining infectious periods equal $-T$.

2.1 Large population approximation of $(\zeta^K)_{K \geq 1}$

For any $f \in C_b^1(\mathbb{R}_+ \times E, \mathbb{R})$, let $f_t(x) = f(t, x)$ and $f_t^{\mathcal{I}}(x) = \langle \nu, f_t(\mathbf{j}(x, \cdot)) \rangle$ for every $(t, x) \in \mathbb{R}_+ \times E$. Consider the differential operator \mathcal{A} defined as follows. For any $x = (n, s, \tau) \in E$,

$$\mathcal{A}f_t(x) = \partial_t f(t, x) - \sum_{k=1}^{n-s} \partial_{\tau_k} f(t, x).$$

Also, for any $x = (n, s, \tau) \in E$, let $\mathbf{n}(x) = n$, $\mathbf{s}(x) = s$ and $\mathbf{i}(x) = i(\tau)$ be the functions which to a structure in state x associate the corresponding structure size, number of susceptible and number of infected members, respectively. For instance, for any $X \in \{H, W\}$ the average rage of within-structure infections at time t is given by

$$\lambda_X \langle \zeta_t^{X|K}, \mathbf{si} \rangle = \frac{\lambda_X}{K_X} \sum_{k=1}^{K_X} \mathbf{s}(x_k^X(t)) \mathbf{i}(x_k^X(t)).$$

Notice also that as mentioned previously, the average size of structures is constant over time, hence $\langle \zeta_t^{X|K}, \mathbf{n} \rangle = \langle \zeta_0^{X|K}, \mathbf{n} \rangle$ for all $t \geq 0$. Finally, let $\mathfrak{M}_1 = \mathcal{M}_1(E)^2$. We are now ready to state our first result, whose proof is postponed to Section 3.1.

Theorem 2.2. *Suppose that $(\zeta_0^K)_{K \geq 1}$ satisfies Assumption 2.1 and converges in law to $\eta_0 \in \mathfrak{M}_1$. Then $(\zeta^K)_{K \geq 1}$ converges in $\mathbb{D}(\mathbb{R}_+, \mathcal{M}_1(E))^2$ to $\eta = (\eta^H, \eta^W)$ defined as the unique solution of the following system of Equations (4). For any $f \in \mathcal{C}_b^1(\mathbb{R}_+ \times E, \mathbb{R})$, for any $T \geq 0$,*

$$\begin{aligned} \langle \eta_T^X, f_T \rangle &= \langle \eta_0^X, f_0 \rangle + \int_0^T \langle \eta_t^X, \mathcal{A}f_t \rangle dt + \lambda_X \int_0^T \langle \eta_t^X, \mathbf{si}(f_t^I - f_t) \rangle dt \\ &+ \lambda_{\bar{X}} \int_0^T \frac{\langle \eta_t^{\bar{X}}, \mathbf{si} \rangle}{\langle \eta_t^{\bar{X}}, \mathbf{s} \rangle} \langle \eta_t^X, \mathbf{s}(f_t^I - f_t) \rangle dt + \beta_G \int_0^T \frac{\langle \eta_t^H, \mathbf{i} \rangle}{\langle \eta_0^H, \mathbf{n} \rangle} \langle \eta_t^X, \mathbf{s}(f_t^I - f_t) \rangle dt. \end{aligned} \quad (4)$$

From now on, let us assume that ν is the exponential distribution of parameter γ . As we shall see, it then is possible to deduce from Theorem 2.2 that the proportion of susceptible and infected individuals in the population converges to the solution of a dynamical system, when the size of the population grows large.

Let $s(t)$ and $i(t)$ be the proportions of susceptible and infectious individuals, respectively, in the population at time t according to distribution η_t . Further introduce the set

$$\mathbb{S} = \{(n - i, i) : 2 \leq n \leq n_{\max}, 0 \leq i \leq n - 1\}.$$

For $(S, I) \in \mathbb{S}$, let $n_{(S,I)}^H(t)$ be the proportion of households containing S susceptible and I infected individuals at time t , according to distribution η_t^H . Define $n_{(S,I)}^W(t)$ analogously for workplaces. Finally, consider

$$\tau_G(t) = \beta_G i(t), \text{ and } \tau_X(t) = \frac{\lambda_X}{m_X} \sum_{(S,I) \in \mathbb{S}} SI n_{(S,I)}^X(t) \text{ for } X \in \{H, W\}.$$

We assume that at time 0, a fraction ε of uniformly chosen individuals are infected amidst an otherwise susceptible population. Furthermore, at time 0, the remaining infectious period of each infected individual is supposed to be distributed according to ν , independently from one another. Let us emphasize here that actually, only this second assumption is crucial for the results to hold, while the original distribution of infected individuals does not need to be uniform (in which case forthcoming Equations (5) and (7) need to be adapted). We have chosen this particular initial condition as it has been previously considered in the literature, and refer to the Discussion for further comments.

In practice, this setting corresponds to the following probability distribution $\eta_{0,\varepsilon} = (\eta_{0,\varepsilon}^H, \eta_{0,\varepsilon}^W) \in \mathfrak{M}_1$ characterized for $X \in \{H, W\}$ as follows. For any $n \in \llbracket 1, n_{\max} \rrbracket$ and $s \in \llbracket 0, n \rrbracket$:

$$\eta_{0,\varepsilon}^X(n, s, d\tau) = \pi_n^X \binom{n}{s} (1 - \varepsilon)^s \varepsilon^{n-s} \left(\nu^{\otimes(n-s)} \otimes \delta_0^{\otimes(n_{\max}-n+s)} \right) (d\tau). \quad (5)$$

It then is possible to describe the epidemic dynamics by a finite, closed set of ordinary differential equations, as shown in the following result whose proof is postponed to Section 3.2.

Theorem 2.3. *Let $\varepsilon > 0$. Suppose that ν is the exponential distribution of parameter γ , and that $(\zeta_0^K)_{K \geq 1}$ satisfies Assumption 2.1 and converges in law to $\eta_{0,\varepsilon}$. Then the functions*

$(s, i, n_{(S,I)}^X : X \in \{H, W\}, (S, I) \in \mathbb{S})$ are characterized as being the unique solution of the following dynamical system: for any $t \geq 0$, for any $X \in \{H, W\}$, $(S, I) \in \mathbb{S}$,

$$\frac{d}{dt}s(t) = -(\tau_H(t) + \tau_W(t) + \tau_G(t)s(t)), \quad (6a)$$

$$\frac{d}{dt}i(t) = -\frac{d}{dt}s(t) - \gamma i(t), \quad (6b)$$

$$\begin{aligned} \frac{d}{dt}n_{(S,I)}^X(t) = & - \left(\lambda_X SI + \tau_{\bar{X}}(t) \frac{S}{s(t)} + \tau_G(t)S + \gamma I \right) n_{(S,I)}^X(t) \\ & + \gamma(I+1)n_{(S,I+1)}^X(t) \mathbf{1}_{\{S+I < n_{\max}\}} \\ & + \left(\lambda_X(S+1)(I-1) + \tau_{\bar{X}}(t) \frac{S+1}{s(t)} + \tau_G(t)(S+1) \right) n_{(S+1,I-1)}^X(t) \mathbf{1}_{\{I \geq 1\}}, \end{aligned} \quad (6c)$$

with initial conditions given by

$$s(0) = 1 - \varepsilon; \quad i(0) = \varepsilon; \quad n_{(S,I)}^X(0) = \binom{S+I}{I} \pi_{S+I}^X (1 - \varepsilon)^S \varepsilon^I. \quad (7)$$

This dynamical system may be understood as follows. Equation (6a) corresponds to the fact that the proportion of susceptibles decreases whenever a new infection occurs within the general population, or within a household or workplace. Similarly, Equation (6b) stems from the fact that newly contaminated individuals move from the susceptible to the infected state, which they in turn leave at rate γ since the infectious period is exponentially distributed. It remains to take an interest in Equation (6c). The first line indicates that a structure of type $(S, I) \in \mathbb{S}$ changes its composition upon either the infection of one of its susceptible members, be it an infection within the structure itself or outside of it (within a structure of the opposite type or in the general population), or upon the removal of one of its infected members. Simultaneously, a structure of type $(S, I+1)$ transforms into a structure of type (S, I) whenever one of its infected members recovers, while a structure of type $(S+1, I-1)$ becomes of type (S, I) upon infection of a susceptible member. In particular, this result shows that under the assumptions of Theorem 2.3, in the large population limit, we may neglect the natural correlation between structures caused by the fact that infected individuals belong to two structures at once. This allows to obtain a stronger model reduction than in Theorem 2.2, in the sense that the model reduces to a finite-dimensional ODE-system instead of a measure-valued equation.

Before detailing the proofs of Theorems 2.2 and 2.3, let us examine the latter from a numerical point of view.

2.2 Numerical assessment of the limiting dynamical system

The aim of this section is first to portray that the proposed large population limit, under the form of dynamical system (6a–c), is in good accordance with the original stochastic model for large population sizes. This secondarily leads to some practical comments on the implementation of the dynamical system. Finally, a comparison with another reduced model for epidemics with two layers of mixing will be established, namely with an Edge-Based Compartmental Model (EBCM) in the line of work of Volz *et al.* [41].

2.2.1 Implementation of the dynamical system and illustration of Theorem 2.3

Let us start by illustrating the result of Theorem 2.3 through numerical simulations. Using Gillespie's algorithm, we have performed fifty simulations of the epidemic within a population of $K = 10000$ individuals, where π^H and π^W are roughly inspired by the French household and workplace distributions as observed in 2018 by Insee [5]. These distributions are represented in Figure 1. Two sets of epidemic parameters have been considered, leading to either $R_0 = 2.5$ or

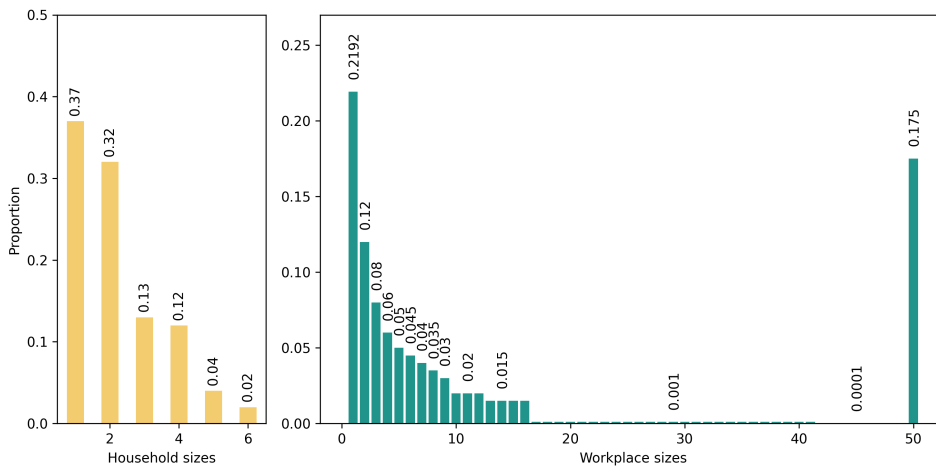


Figure 1: Household and workplace size distributions π^H (left) and π^W (right) used in simulations.

$R_0 = 1.2$. In both cases, the majority of contaminations take place at the local level. Indeed, for the first scenario with $R_0 = 2.5$, 42% and 18% of infections occur within households and workplaces, respectively. For the second scenario, mean-field infections are even less common, with 40% of within-household and 40% of within-workplace infections. Further, the epidemic is started by infecting either 10 or 100 individuals chosen uniformly at random at time 0 ($\varepsilon = 0.001$ or $\varepsilon = 0.01$, respectively). For each simulation, we have followed the evolution of the proportion of susceptible and infected individuals within the population over time. When $\varepsilon = 0.001$, time subsequently is shifted so that time 0 corresponds to the first moment when at least 5‰ of the population are infectious.

The simulation outcomes are presented in Figure 2. For each choice of parameters, the solutions s and i of dynamical system (6a-c) with initial condition given by (7) are plotted on the same graph. As previously for $\varepsilon = 0.001$, time is subsequently shifted to ensure that $i(0) = 0.005$, in order to synchronize the simulations among themselves and with their deterministic limit. As expected, one observes good accordance of the stochastic simulations and the deterministic functions (s, i) .

Before proceeding further, let us briefly emphasize a few aspects of the implementation of the proposed deterministic model. A potential drawback of dynamical system (6a-c) consists in its large dimension. Indeed, it holds that $\#\mathcal{S} = n_{\max}(n_{\max} + 1)/2 - 1$. The number of equations of dynamical system (6a-c) is hence given by $2 + 2\#\mathcal{S}$, and of order $O(n_{\max}^2)$. However, this fast-growing number of equations actually is manageable, as it is possible to implement the dynamical system in an automated way, in the sense that each equation does not need to be written one-by-one by the programmer. We refer to Appendix A.1 for details.

Nevertheless, the large dimension of the dynamical system of interest raises the question whether it is numerically speaking interesting to actually use it for numerical explorations when compared to stochastic simulations using Gillespie’s algorithm. Indeed, if the computation time needed to solve the dynamical system is not advantageous when compared to stochastic simulations, one may prefer to use the latter as it contains more information, such as the fluctuations around the large population limit. In order to address this question, we have compared the average time needed to either solve once dynamical system (6a-c), or to simulate one trajectory of the stochastic model using Gillespie’s algorithm, for different choices of epidemic parameters. Notice here that in practice, for stochastic simulations, it is often necessary to compute several individual trajectories in order to obtain the general behavior of the epidemic. However, as it is possible to execute these simulations in parallel, comparison to one individual

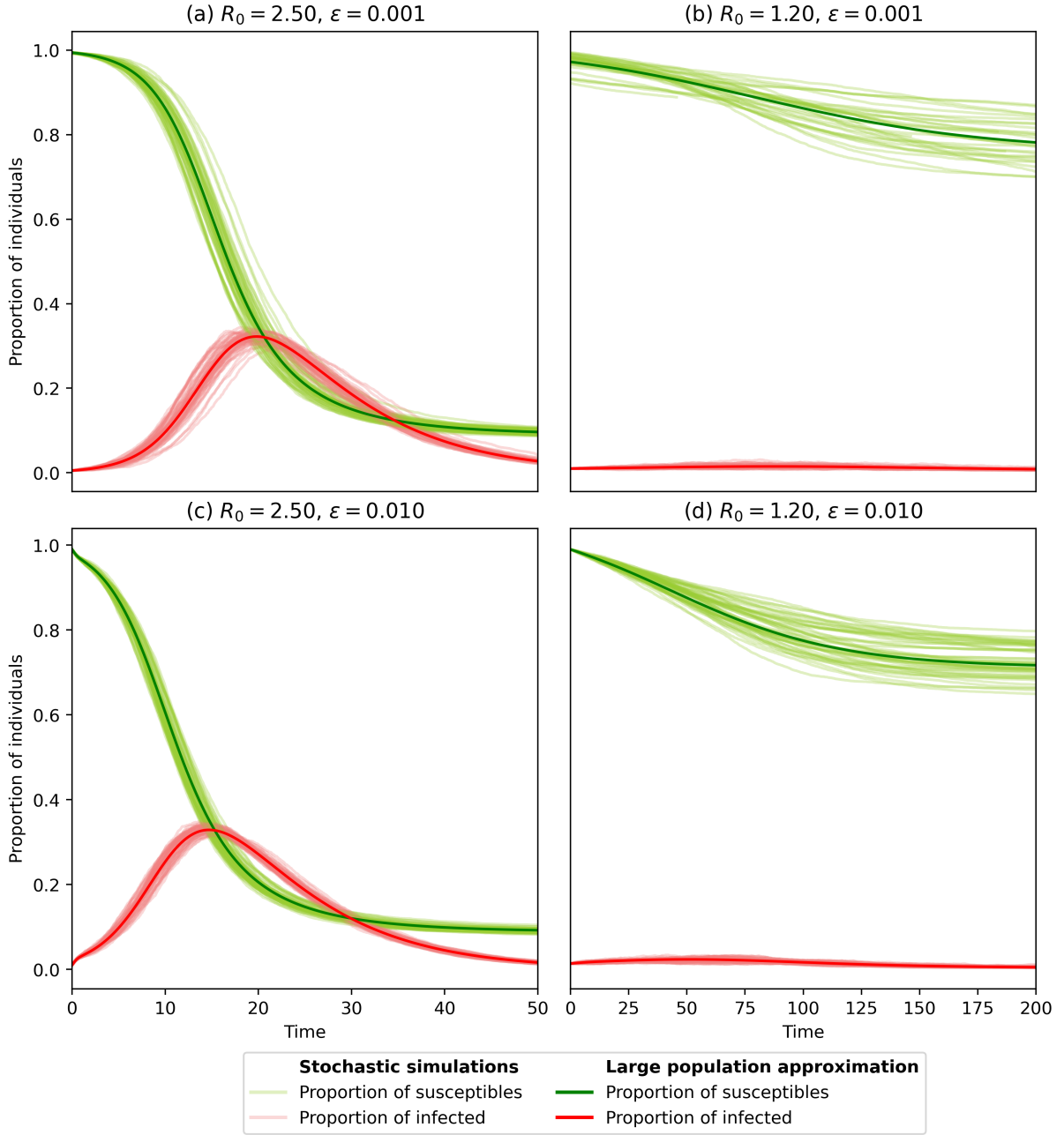


Figure 2: Comparison of the stochastic model with its large population approximation given by dynamical system (6a-c). Household and workplace distributions are those of Figure 1. Two sets of epidemic parameters were considered, namely $(\beta_G, \lambda_H, \lambda_W, \gamma) = (0.125, 1.5, 0.00115, 0.125)$ and $(\beta_G, \lambda_H, \lambda_W, \gamma) = (0.03, 0.05, 0.0015, 0.125)$ for the left and right column respectively ($R_0 = 2.5$ and $R_0 = 1.2$). The initial conditions are either $\epsilon = 0.001$ in Panels (a) and (b), or $\epsilon = 0.01$ in (c) and (d). For each of these scenarios, Gillespie's algorithm was used to simulate 50 trajectories of the stochastic model defined in Proposition 1.1 in a population of $K = 10000$ individuals (faint lines). For Panels (a) and (b), only trajectories reaching a threshold of 0.005 infected were kept, and time was shifted so that time 0 corresponds to the moment when this threshold is reached. Finally, the deterministic solution (s, i) of (6a-c) is represented for each scenario (thick lines). For Panels (a) and (b), the same time shifting procedure as for simulations is applied.

simulation seemed the most pertinent.

One hundred independent runs of each script were performed and their computation time assessed. For each run, the runtime of a reference function was also measured, in order to take into account fluctuations of computation time which occur among runs of the same script. The procedure and results are detailed in Appendix A.2. In summary, solving the reduced model is up to one order of magnitude faster than performing one stochastic simulation for values of $R_0 > 1$ that are not too close to the critical case $R_0 = 1$. This shows that the reduced model is pertinent for numerical exploration.

2.2.2 Comparison to Edge-Based Compartmental Models

One may notice that the population structure, as described in Section 1.1, can be regarded as a modification of the well-studied configuration model. Indeed, our network of household- and workplace-contacts may be seen as a two-layer graph, where each layer corresponds to a random graph generated as described in [29, 31], that we shall call clique configuration model (CCM) hereafter. This random graph model generalizes configuration models to include small, totally connected sub-graphs referred to as *cliques*. It then is possible to derive an EBCM for our household-workplace model by reasoning as in [41]. Details are provided in Appendix S1.1.

Edge-based compartmental models on CCM variants have been known to be in good accordance with simulations of the corresponding stochastic epidemic models in terms of proportions of susceptible and infected individuals within the population, under the assumption of a very small initial proportion of infected. In our case, we have confronted the EBCM with dynamical system (6a–c), as well as simulated trajectories of our stochastic model. Results are reported on Figure S1. The parameters considered are the same as for Panel (a) of Figure 2, except that several values of the initial proportion of infected ε have been explored. As expected, for very small values of ε , the EBCM and dynamical system (6a–c) both yield the correct epidemic dynamics, whereas for larger values of ε , the EBCM does not fit the simulated epidemic trajectories. We refer to Supplementary Material S1.1 for details.

Finally, one may notice that the EBCM is of higher dimension than dynamical system (6a–c). Proceeding like before, we obtain that the number of equations of the EBCM is of order $O(n_{\max}^3)$. This has a strong negative impact on computation time, as briefly illustrated in Supplementary Material S1.2, arguing against the applicability of this EBCM for numerical explorations.

To conclude, in the particular case of the household-workplace model studied in this article, the EBCM seems to be equivalent to the large population approximation described by dynamical system (6a–c), under the condition that the initial proportion of infected is very small. However, considering both the computational cost of its higher dimension and the loss of accuracy for more general initial conditions of the EBCM, the large population approximation given by dynamical system (6a–c) seems more pertinent in the case of the epidemic model under consideration.

3 Proofs

This section is devoted to establishing Theorems 2.2 and 2.3. As we will see, the proof of Theorem 2.2 naturally has the intrinsic difficulty of all convergence results for measure-valued processes, but it allows us to obtain a deterministic prediction of the dynamics of the structure type distributions during the course of an epidemic. At this level, the limiting object is rich, allowing it to convey detailed information on the distribution of remaining infectious periods within structures. This however comes at the cost of an infinite-dimensional limiting object, which motivates the interest in trying to further reduce its dimension by adopting a coarser population description. In the case where ν is the exponential distribution, Theorem 2.3 shows

that this actually is possible, the final reduced model taking the form of dynamical system (6a–c). As mentioned previously, the existence of an asymptotically exact, closed, finite-dimensional ODE-system capturing the epidemic dynamics was not obvious from the beginning. Indeed, in order to obtain this result, we need to eliminate from the state space E the continuous component τ , which needs to be transformed back into the simpler distinction between infected and removed states for each member of the structure. As we shall see in Subsection 3.2, this will demand some effort as removal events are not explicit in the model formalism adopted in Section 1.2. This difficulty will become apparent in the proof of forthcoming Proposition 3.15.

3.1 Proof of Theorem 2.2

Let us start with the proof of Theorem 2.2. It follows a classical scheme, establishing tightness of $(\zeta^K)_{K \geq 1}$, whose limiting values are shown to satisfy Equation (4). Uniqueness of the solutions of this equation given the initial condition then ensures the desired convergence result. In particular, the proof is inspired by [12] and [38].

3.1.1 Uniqueness and continuity of the solution of Equation (4).

We are first going to establish a uniqueness result for the solutions of Equation (4). Notice that we do not need to prove existence of solutions in this section, as forthcoming Proposition 3.8 constructs such solutions as limiting values of $(\zeta^K)_{K \geq 1}$.

Let us start with a technical lemma, whose proof we present for sake of completeness.

Lemma 3.1. *Let $f \in \mathcal{C}_b(E, \mathbb{R})$. There exists a sequence $(f_k)_{k \geq 1}$ taking values in $\mathcal{C}_b^1(E, \mathbb{R})$ such that f_k converges simply to f and $\sup_{k \geq 1} \|f_k\|_\infty \leq \|f\|_\infty$.*

Proof. Consider a mollifier ψ , i.e. $\psi \in \mathcal{C}^\infty(\mathbb{R}^{n_{\max}})$ is compactly supported, its mass $\int_{\mathbb{R}^{n_{\max}}} \psi(x) dx$ equals 1, and for $k \geq 1$, the function $\psi_k : x \mapsto k^{n_{\max}} \psi(kx)$ satisfies $\lim_{k \rightarrow \infty} \psi_k = \delta_0$ in the sense of distributions.

Consider $f \in \mathcal{C}_b(E, \mathbb{R})$. Define the sequence $(f_k)_{k \geq 1}$ as follows:

$$f_k : (n, s, \tau) \in E \mapsto f(n, s, \cdot) * \psi_k(\tau).$$

Then, for any $x \in E$, by definition of $(\psi_k)_{k \geq 1}$ it holds that $\lim_{k \rightarrow \infty} f_k(x) = f(x)$. Further, as ψ_k is of integral 1 for any k , it is obvious that for any k , $\|f_k\|_\infty \leq \|f\|_\infty$. Finally, it also follows from the usual properties of convolution that for any $k \geq 1$, f_k is smooth with respect to its last variable and all the corresponding partial derivatives are compactly supported, hence $f_k \in \mathcal{C}_b^1(E, \mathbb{R})$ for any $k \geq 1$. \square

We may now turn to the main result of this paragraph. With slight abuse of notation, for an element $\eta = (\eta_1, \eta_2) \in \mathfrak{M}_1$, we define its total variation norm by $\|\eta\|_{TV} = \|\eta_1\|_{TV} \vee \|\eta_2\|_{TV}$.

Proposition 3.2. *Let $\eta_\star \in \mathfrak{M}_1$. Then Equation (4) admits at most one measure-valued solution η which belongs to $\mathcal{C}(\mathbb{R}_+, (\mathfrak{M}_1, \|\cdot\|_{TV}))$, such that $\eta_0 = \eta_\star$.*

From now on, define $n^H = \langle \eta_0^H, \mathbf{n} \rangle$, $s_t^H = \langle \eta_t^H, \mathbf{s} \rangle$, $i_t^H = \langle \eta_t^H, \mathbf{i} \rangle$, and $s_t^W = \langle \eta_t^W, \mathbf{s} \rangle$. Let us establish the proposition.

Proof. First, notice that it follows immediately from Equation (4) that $\langle \eta_T^X, 1 \rangle = \langle \eta_0^X, 1 \rangle$ for $X \in \{H, W\}$, thus $\eta_0 \in \mathfrak{M}_1$ implies that for any $T \geq 0$, $\eta_T \in \mathfrak{M}_1$.

Let us show that any solution η of Equation (4) belongs to $\mathcal{C}(\mathbb{R}_+, (\mathfrak{M}_1, \|\cdot\|_{TV}))$. In order to do so, it is enough to show that $\eta^X \in \mathcal{C}(\mathbb{R}_+, (\mathcal{M}_1(E), \|\cdot\|_{TV}))$ for any $X \in \{H, W\}$. We are going to detail the proof for η^H only, as η^W can be handled in the same way.

Let $T \geq 0$ and $g \in \mathcal{C}_b^1(E, \mathbb{R})$ such that $\|g\|_\infty \leq 1$. Consider the function defined by

$$\forall (t, x) \in \mathbb{R} \times E, \quad f_t(x) = g(\Psi(x, T, t))$$

and recall that $f_t^{\mathcal{I}}(x) = \langle \nu, f_t(j(x, \cdot)) \rangle$. Then by definition, $f_T(x) = g(x)$. It follows from the assumption $g \in \mathcal{C}_b^1(E, \mathbb{R})$ that $f \in \mathcal{C}_b^1(\mathbb{R}_+ \times E, \mathbb{R})$. The advantage of this construction is that $t \mapsto \Psi(x, T, t)$ corresponds to a reversal of time, which cancels out the deterministic dynamics described by the differential operator \mathcal{A} . Indeed, letting $x = (n, s, \tau)$, a brief computation shows that

$$\partial_t f_t(x) = \sum_{k=1}^{n-s} \delta_{\tau_k} g(\Psi(x, T, t)) \text{ and } \partial_{\tau_k} f_t(x) = \delta_{\tau_k} g(\Psi(x, T, t)),$$

which yields that $\mathcal{A}f_t(x) = 0$ for all $(t, x) \in \mathbb{R} \times E$. Using the fact that $\langle \eta_T^H, g \rangle = \langle \eta_T^H, f_T \rangle$, it follows from Equation (4) that

$$\begin{aligned} \langle \eta_T^H, g \rangle &= \langle \eta_0^H, f_0 \rangle + \lambda_H \int_0^T \langle \eta_t^H, \mathbf{si}(f_t^{\mathcal{I}} - f_t) \rangle dt \\ &+ \lambda_W \int_0^T \frac{1}{s_t^W} \langle \eta_t^W, \mathbf{si} \rangle \langle \eta_t^H, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle dt + \beta_G \int_0^T \frac{i_t^H}{n^H} \langle \eta_t^H, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle dt. \end{aligned}$$

Recall that $i_t^H \leq n_{\max}$ and $\frac{i_t^H}{n^H} \leq 1$ since for any $x \in E$, $\mathbf{i}(x) \leq \mathbf{n}(x)$. We may notice that the following inequalities hold, as for any t , $\|f_t\|_\infty \leq 1$:

$$\begin{aligned} \langle \eta_t^H, \mathbf{si}(f_t^{\mathcal{I}} - f_t) \rangle &\leq 2(n_{\max})^2, & \langle \eta_t^H, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle &\leq 2n_{\max}, \\ \frac{1}{s_t^W} \langle \eta_t^W, \mathbf{si} \rangle &\leq n_{\max}, & \frac{i_t^H}{n^H} \langle \eta_t^H, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle &\leq 2n_{\max}. \end{aligned} \quad (8)$$

Let $C = 2n_{\max}(\lambda_H n_{\max} + \lambda_W n_{\max} + \beta_G)$ and let $\epsilon \in \mathbb{R}$. It then follows from inequalities (8) that

$$|\langle \eta_T^H - \eta_{T+\epsilon}^H, g \rangle| \leq C|T - (T + \epsilon)| = C|\epsilon|. \quad (9)$$

Consider now $h \in \mathcal{C}_b(E, \mathbb{R})$ such that $\|h\|_\infty \leq 1$. Lemma 3.1 ensures that there exists a sequence $(g_k)_{k \geq 1}$ taking values in $\mathcal{C}_b^1(E, \mathbb{R})$ which converges simply to h and such that $\|g_k\|_\infty \leq 1$. By dominated convergence, this implies that

$$|\langle \eta_T^H - \eta_{T+\epsilon}^H, h \rangle| \leq C|T - (T + \epsilon)| = C|\epsilon|.$$

As E is a Polish space, it follows from Proposition A.6.1 of [38] that

$$\|\eta_T^H - \eta_{T+\epsilon}^H\|_{TV} = \sup_{h \in \mathcal{C}_b(E, \mathbb{R}) : \|h\|_\infty \leq 1} |\langle \eta_T^H - \eta_{T+\epsilon}^H, h \rangle| \leq C|\epsilon|. \quad (10)$$

As this holds for any ϵ and any $T \geq 0$, the strong continuity of η^H is established.

It remains to establish uniqueness of the solution η of Equation (4) with initial condition $\eta_0 = \eta_\star$. We will once more establish uniqueness component-wise, and focus on η^H as η^W is treated in a similar fashion.

Let $\eta, \bar{\eta}$ be two solutions of Equation (4) with initial condition η_\star . Let $\bar{i}_t^H = \langle \bar{\eta}_t^H, \mathbf{i} \rangle$ and define \bar{s}_t^W in analogous manner. As before, let $T \geq 0$. Consider again $g \in \mathcal{C}_b^1(E, \mathbb{R})$ such that $\|g\|_\infty \leq 1$, and define f_t and $f_t^{\mathcal{I}}$ as previously. Then

$$\begin{aligned} |\langle \eta_T^H - \bar{\eta}_T^H, g \rangle| &\leq \lambda_H \int_0^T |\langle \eta_t^H - \bar{\eta}_t^H, \mathbf{si}(f_t^{\mathcal{I}} - f_t) \rangle| dt \\ &+ \lambda_W \int_0^T \left| \frac{1}{s_t^W} \langle \eta_t^W, \mathbf{si} \rangle \langle \eta_t^H, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle - \frac{1}{\bar{s}_t^W} \langle \bar{\eta}_t^W, \mathbf{si} \rangle \langle \bar{\eta}_t^H, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle \right| dt \\ &+ \frac{\beta_G}{m_H} \int_0^T \left| i_t^H \langle \eta_t^H, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle - \bar{i}_t^H \langle \bar{\eta}_t^H, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle \right| dt. \end{aligned}$$

Proceeding similarly as in Inequalities (8), we obtain that

$$|\langle \eta_T^H - \bar{\eta}_T^H, g \rangle| \leq \frac{C}{2} \int_0^T |\langle \eta_t^H - \bar{\eta}_t^H, f_t^I - f_t \rangle| dt \leq \frac{C}{2} \left(\int_0^T |\langle \eta_t^H - \bar{\eta}_t^H, f_t^I \rangle| dt + \int_0^T |\langle \eta_t^H - \bar{\eta}_t^H, f_t \rangle| dt \right).$$

On the one hand, by definition, $f_t \in \mathcal{C}_b(E, \mathbb{R})$ and $\|f_t\|_\infty \leq 1$. Thus, it is obvious that $|\langle \eta_t^H - \bar{\eta}_t^H, f_t \rangle| \leq \|\eta_t^H - \bar{\eta}_t^H\|_{TV}$. On the other hand, it follows from the usual criterion of continuity for parametric integrals that f_t^I is continuous on E and $\|f_t^I\|_\infty \leq \|f_t\|_\infty \leq 1$. As a consequence, $|\langle \eta_t^H - \bar{\eta}_t^H, f_t^I \rangle| \leq \|\eta_t^H - \bar{\eta}_t^H\|_{TV}$. Hence

$$|\langle \eta_T^H - \bar{\eta}_T^H, g \rangle| \leq C \int_0^T \|\eta_t^H - \bar{\eta}_t^H\|_{TV} dt.$$

We then may follow the same steps that allowed to establish Equation (10) from Equation (9), and obtain that

$$\|\eta_T^H - \bar{\eta}_T^H\|_{TV} \leq C \int_0^T \|\eta_t^H - \bar{\eta}_t^H\|_{TV} dt.$$

Gronwall's lemma then assures that

$$\forall t \in [0, T], \quad \|\eta_t^H - \bar{\eta}_t^H\|_{TV} = 0.$$

Obviously, one obtains the analogous result for η^W in the same manner. As $T \geq 0$ is arbitrary, this concludes the proof. \square

3.1.2 Tightness of $(\zeta^K)_{K \geq 1}$ in $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(E), w))^2$

Let us now turn to the tightness of $(\zeta^K)_{K \geq 1}$ in $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(E), w))^2$, where w designates the weak topology on $\mathcal{M}_F(E)$. We start by establishing the following preliminary result, whose proof simply relies on the chain rule.

Lemma 3.3. *Let $f \in \mathcal{C}_b^1(\mathbb{R}_+ \times E, \mathbb{R})$. Then for any $T \geq t_0 \geq 0$, for any $x \in E$,*

$$f(T, \Psi(x, T, t_0)) = f(t_0, x) + \int_{t_0}^T \mathcal{A}f(t, \Psi(x, t, t_0)) dt.$$

Proof. For $x \in E$ and $t_0 \in \mathbb{R}_+$, define $g_{t_0, x} : [t_0, +\infty) \rightarrow \mathbb{R}$, $T \mapsto f(T, \Psi(x, T, t_0))$. Let us start by noticing that for any $(t_0, x) \in \mathbb{R}_+ \times E$, $g_{t_0, x} \in \mathcal{C}^1(\mathbb{R}_+)$. Indeed, $g_{t_0, x} = f_{n, s} \circ h_{t_0, x}$, where

$$f_{n, s} : \mathbb{R}_+ \times \mathbb{R}^{n_{\max}} \rightarrow \mathbb{R}, \quad (u, \mathbf{v}) \mapsto f(u, (n, s, \mathbf{v}))$$

and $h_{t_0, x} : \mathbb{R}_+ \rightarrow \mathbb{R}^{1+n_{\max}}$, $t \mapsto \left(t, \tau - \sum_{k=1}^{n-s} (t - t_0) e_k \right)$.

The chain rule and a quick computation of the differentiable of $h_{t_0, x}$ yields that for every $t \geq t_0$,

$$g'_{t_0, x}(t) = \partial_1 f_{n, s}(h_{t_0, x}(t)) - \sum_{k=1}^{n-s} \partial_{k+1} f_{n, s}(h_{t_0, x}(t)).$$

Notice that $\partial_1 f_{n, s}(u, \mathbf{v}) = \partial_t f(u, (n, s, \mathbf{v}))$ and for $k \geq 2$, $\partial_k f_{n, s}(u, \mathbf{v}) = \partial_{\tau_{k-1}} f(u, (n, s, \mathbf{v}))$. As a consequence, we have shown that for any $(t_0, x) \in \mathbb{R}_+ \times E$, for every $t \geq t_0$, $g_{t_0, x}$ is differentiable at t and satisfies

$$g'_{t_0, x}(t) = \mathcal{A}f(t, \Psi(x, t, t_0)).$$

This concludes the proof. \square

Throughout the section, we use the notation $\mathcal{S} = \{H, W, G\}$. Also, for any $f \in \mathcal{C}_b^1(\mathbb{R}_+ \times E, \mathbb{R})$, let $f_t(x) = f(t, x)$ for any $(t, x) \in \mathbb{R}_+ \times E$. Finally, we define for any continuous bounded function $g : \mathbb{R}_+ \times E \rightarrow \mathbb{R}$, for any $t \geq 0$ and $u = (\boldsymbol{\theta}, k, \ell, \sigma) \in \bigcup_{Y \in \mathcal{S}} U_K^Y$:

$$g_{t,u}^H = g(t, j(x_k^H(t), \sigma)) - g(t, x_k^H(t)) \quad \text{and} \quad g_{t,u}^W = g(t, j(x_\ell^W(t), \sigma)) - g(t, x_\ell^W(t)).$$

Proposition 3.4. *Consider ζ^K as introduced in Proposition 1.1. For any $f \in \mathcal{C}_b^1(\mathbb{R}_+ \times E, \mathbb{R})$, $T \geq 0$ and $X \in \{H, W\}$,*

$$\langle \zeta_T^{X|K}, f_T \rangle = \langle \zeta_0^{X|K}, f_0 \rangle + \int_0^T \langle \zeta_t^{X|K}, \mathcal{A}f_t \rangle dt + \frac{1}{K_X} \sum_{Y \in \mathcal{S}} \int_0^T \int_{U_K^Y} \mathcal{I}_Y(t-, u) f_{t-,u}^X Q^{Y|K}(dt, du).$$

Proof. Let $f \in \mathcal{C}_b^1(\mathbb{R}_+ \times E, \mathbb{R})$ and $X \in \{H, W\}$. Recall that, by definition, for any bounded function $g : E \rightarrow \mathbb{R}$, for any $T \geq t \geq 0$ and $u = (\boldsymbol{\theta}, k, \ell, \sigma) \in \bigcup_{Y \in \mathcal{S}} U_K^Y$:

$$\langle \Delta_H(u, T, t), g \rangle = g(\Psi(j(x_k^H(t-), \sigma), T, t)) - g(\Psi(x_k^H(t-), T, t)),$$

and $\langle \Delta_W(u, T, t), g \rangle$ is defined analogously, by replacing H by W and k by ℓ .

From Equation (3), it follows that

$$\langle \zeta_T^{X|K}, f_T \rangle = \frac{1}{K_X} \sum_{j=1}^{K_X} f_T(\Psi(x_j^X(0), T, 0)) + \frac{1}{K_X} \sum_{Y \in \mathcal{S}} \int_0^T \int_{U_K^Y} \mathcal{I}_Y(t-, u) \langle \Delta_X(u, T, t), f_T \rangle Q^{Y|K}(dt, du).$$

Using the result from Lemma 3.3, this becomes:

$$\begin{aligned} \langle \zeta_T^{X|K}, f_T \rangle &= \frac{1}{K_X} \sum_{j=1}^{K_X} \left(f_0(x_j^X(0)) + \int_0^T \mathcal{A}f_t(\Psi(x_j^X(0), t, 0)) dt \right) \\ &\quad + \frac{1}{K_X} \sum_{Y \in \mathcal{S}} \int_0^T \int_{U_K^Y} \mathcal{I}_Y(t-, u) \left(\int_t^T \langle \Delta_X(u, z, t), \mathcal{A}f_z \rangle dz \right) Q^{Y|K}(dt, du) \\ &\quad + \frac{1}{K_X} \sum_{Y \in \mathcal{S}} \int_0^T \int_{U_K^Y} \mathcal{I}_Y(t-, u) f_{t-,u}^X Q^{Y|K}(dt, du). \end{aligned}$$

It follows from the definition of $\mathcal{C}_b^1(\mathbb{R}_+ \times E, \mathbb{R})$ that both f and $\mathcal{A}f$ are bounded, hence we may apply Fubini's theorem to obtain that

$$\begin{aligned} \langle \zeta_T^{X|K}, f_T \rangle &= \frac{1}{K_X} \sum_{j=1}^{K_X} f_0(x_j^X(0)) + \frac{1}{K_X} \sum_{Y \in \mathcal{S}} \int_0^T \int_{U_K^Y} \mathcal{I}_Y(t-, u) f_{t-,u}^X Q^{Y|K}(dt, du) \\ &\quad + \frac{1}{K_X} \int_0^T \left(\sum_{j=1}^{K_X} \mathcal{A}f_z(\Psi(x_j^X(0), z, 0)) + \sum_{Y \in \mathcal{S}} \int_0^z \int_{U_K^Y} \mathcal{I}_Y(t-, u) \langle \Delta_X(u, z, t), \mathcal{A}f_z \rangle Q^{Y|K}(dt, du) \right) dz. \end{aligned}$$

The first sum on the right-hand side equals $\langle \zeta_0^{X|K}, f_0 \rangle$. From the second line, one recognizes in the integrand the definition of $\langle \zeta_z^{X|K}, \mathcal{A}f_z \rangle$ from Equation (3). This yields the desired result. \square

For $Y \in \mathcal{S}$, let

$$\tilde{Q}^{Y|K}(dt, du) = Q^{Y|K}(dt, du) - dt \mu_K^Y(du)$$

be the compensated martingale-measure associated to $Q^{Y|K}$.

It follows that, for $f \in \mathcal{C}_b^1(\mathbb{R}_+ \times E, \mathbb{R})$ and $X \in \{H, W\}$,

$$\langle \zeta_T^{X|K}, f_T \rangle = M_T^{X|K}(f) + V_T^{X|K}(f),$$

where we define

$$M_T^{X|K}(f) = \frac{1}{K_X} \sum_{Y \in \mathcal{S}} \int_0^T \int_{U_K^Y} \mathcal{I}_Y(t-, u) f_{t-,u}^X \tilde{Q}^{Y|K}(dt, du)$$

and

$$V_T^{X|K}(f) = \langle \zeta_0^{X|K}, f_0 \rangle + \int_0^T \langle \zeta_t^{X|K}, \mathcal{A}f_t \rangle dt + \frac{1}{K_X} \sum_{Y \in \mathcal{S}} \int_0^T \int_{U_K^Y} \mathcal{I}_Y(t, u) f_{t,u}^X \mu_K^Y(du) dt.$$

Proposition 3.5. *Let $f \in C_b^1(\mathbb{R}_+ \times E, \mathbb{R})$ and $X \in \{H, W\}$. Then $(M_T^{X|K}(f))_{T \geq 0}$ is a square integrable martingale. Using the same notations as in Theorem 2.2, its quadratic variation is given by*

$$\langle M^{X|K}(f) \rangle_T = \frac{1}{K_X} \int_0^T \langle \zeta_t^{X|K}, \mathcal{H}_t^X((f_t^X)^2 - 2f_t^X f_t + f_t^2) \rangle dt,$$

where for any $t \geq 0$ and $x \in E$,

$$\mathcal{H}_t^X(x) = \beta_G \frac{I^H(t)}{N^H} \mathbf{s}(x) + \lambda_X \mathbf{s}(x) \mathbf{i}(x) + \lambda_{\bar{X}} \frac{\langle \zeta_t^{\bar{X}|K}, \mathbf{si} \rangle}{S_t^{\bar{X}}} \mathbf{s}(x).$$

Proof. Let $f \in C_b^1(E, \mathbb{R})$ and $X \in \{H, W\}$. Consider $M_T^{X|K}(f)$, which can be written as

$$M_T^{X|K}(f) = M_{X,T}^{X|K}(f) + M_{\bar{X},T}^{X|K}(f) + M_{G,T}^{X|K}(f),$$

where, for $Y \in \mathcal{S}$ and $T \geq 0$,

$$M_{Y,T}^{X|K}(f) = \int_0^T \int_{U_K^Y} \frac{1}{K_X} \mathcal{I}_Y(t-, u) f_{t-,u}^X \tilde{Q}^{Y|K}(dt, du).$$

Suppose that for any $Y \in \mathcal{S}$,

$$\mathbb{E} \left[\int_0^T \int_{U_K^Y} \left(\frac{1}{K_X} \mathcal{I}_Y(t, u) f_{t,u}^X \right)^2 \mu_K^Y(du) dt \right] < \infty,$$

then for all $Y \in \mathcal{S}$, $(M_{Y,T}^{X|K}(f))_{T \geq 0}$ is a square integrable martingale [27], implying that $(M_T^{X|K}(f))_{T \geq 0}$ also is a square integrable martingale. As $Q^{H|K}$, $Q^{W|K}$ and $Q^{G|K}$ are independent, it follows that

$$\langle M^{X|K}(f) \rangle_T = \sum_{Y \in \mathcal{S}} \langle M_Y^{X|K}(f) \rangle_T.$$

It thus is enough to study $(M_{Y,T}^{X|K}(f))_{T \geq 0}$ for all $Y \in \mathcal{S}$. In the following, we will detail the necessary computations in the case $X = H$, the case $X = W$ being similar.

Consider the case $Y = H$. Start by noticing that $\sum_{\ell=1}^{K_W} s_\ell^W(t) = K_W S^W(t)$, and that for any $k \in \llbracket 1, K_H \rrbracket$ and $t \in [0, T]$, $s_k^H(t)$ and $i_k^H(t)$ are less than n_{\max} . Hence, replacing $S(t)$ by $K_W S^W(t)$ in \mathcal{I}_H ,

$$\begin{aligned} \mathbb{E} \left[\langle M_H^{H|K}(f) \rangle_T \right] &= \mathbb{E} \left[\int_0^T \int_{U_K^H} \left(\frac{1}{K_H} \mathcal{I}_H(t, u) f_{t,u}^H \right)^2 \mu_K^H(du) dt \right] \\ &= \mathbb{E} \left[\int_0^T \frac{1}{K_H^2} \sum_{k=1}^{K_H} \lambda_H s_k^H(t) i(\tau_k^H(t)) \langle \nu, (f_t(j(x_k^H(t), \cdot)) - f_t(x_k^H(t)))^2 \rangle dt \right] \\ &\leq \frac{1}{K_H} \lambda_H (n_{\max})^2 4 \|f\|_\infty^2 T. \end{aligned}$$

Since further $K_H \geq K/n_{\max}$ and $\|f\|_\infty^2 < \infty$, we obtain that

$$\mathbb{E} \left[\langle M_H^{H|K}(f) \rangle_T \right] \leq \frac{4}{K} \lambda_X(n_{\max})^3 \|f\|_\infty^2 T < \infty. \quad (11)$$

Thus $(M_{H,t}^{H|K}(f))_{t \geq 0}$ is a square integrable martingale whose quadratic variations is given by

$$\langle M_H^{H|K}(f) \rangle_T = \int_0^T \int_{U_K^H} \left(\frac{1}{K_H} \mathcal{I}_H(t, u) f_{t,u}^H \right)^2 \mu_K^H(du) dt = \frac{\lambda_H}{K_H} \int_0^T \langle \zeta_t^{H|K}, \mathbf{si}((f_t^2)^{\mathcal{I}} - 2f_t^{\mathcal{I}} f_t + f_t^2) \rangle dt,$$

using the computations from Equation (3.1.2).

Similarly, $(M_{W,T}^{H|K}(f))_{T \geq 0}$ is a square integrable martingale of quadratic variation

$$\langle M_W^{H|K}(f) \rangle_T = \lambda_W \int_0^T \frac{\langle \zeta_t^{W|K}, \mathbf{si} \rangle}{S^W(t)} \langle \zeta_t^{H|K}, \mathbf{s}((f_t^2)^{\mathcal{I}} - 2f_t^{\mathcal{I}} f_t + f_t^2) \rangle dt.$$

Further, for the case $Y = G$, let us use the equalities $S(t) = K_H S^H(t)$ and $I(t) = K_H I^H(t)$. As $I(t)/K = I^H(t)/N^H \leq 1$ almost surely, we obtain that

$$\mathbb{E} \left[\langle M_G^{H|K}(f) \rangle_T \right] = \mathbb{E} \left[\int_0^T \int_{U_K^G} \left(\frac{1}{K_H} \mathcal{I}_G(t, u) f_{t,u}^H \right)^2 \mu_K^G(du) dt \right] \leq \frac{4}{K} \beta_G(n_{\max})^2 \|f\|_\infty^2 T.$$

As before, $(M_{G,T}^{H|K}(f))_{T \geq 0}$ thus is a square integrable martingale of quadratic variation given by

$$\langle M_G^{H|K}(f) \rangle_T = \frac{1}{K_H} \beta_G \int_0^T \frac{I^H(t)}{N^H} \langle \zeta_t^{H|K}, \mathbf{s}((f_t^2)^{\mathcal{I}} - 2f_t^{\mathcal{I}} f_t + f_t^2) \rangle dt.$$

This yields the desired result for $(M_T^{H|K}(f))_{T \geq 0}$, and proceeding similarly for $(M_T^{W|K}(f))_{T \geq 0}$ concludes the proof. \square

We are now ready to focus on the tightness of $(\zeta^K)_{K \geq 1}$, endowing $\mathcal{M}_F(E)$ with the vague topology v as a first step.

Proposition 3.6. *Under the assumptions of Theorem 2.2, the sequence $(\zeta^K)_{K \geq 1}$ is tight in $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(E), v))^2$.*

The proof relies on the fact that in order to establish tightness of $(\zeta^K)_{K \geq 1}$, it is enough to show that for any $X \in \{H, W\}$, $(\langle \zeta^{X|K}, f \rangle)_{K \geq 1}$ is tight for a large enough set of test functions f [34]. This in turn is ensured using the Aldous [1] and Rebolledo [18] criteria, whose application is straightforward thanks to the upper bounds established in the previous proof.

Proof. Once more, we will proceed component-wise and show that $(\zeta^{H|K})_{K \geq 1}$ and $(\zeta^{W|K})_{K \geq 1}$ are both tight in $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(E), v))$.

Let us focus on $(\zeta^{H|K})_{K \geq 1}$. According to Theorem 2.1 of [34], it is sufficient to show that for any function f belonging to a dense subset of

$$\mathcal{C}_0(E, \mathbb{R}) = \left\{ f : E \rightarrow \mathbb{R} \text{ continuous s.t. } \lim_{\|x\|_\infty \rightarrow \infty} |f(x)| = 0 \right\},$$

the sequence $(\langle \zeta^{H|K}, f \rangle)_{K \geq 1}$ is tight in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$. Notice that by density of $C_K^\infty(\mathbb{R}^{n_{\max}})$ in $\mathcal{C}_0(\mathbb{R}^{n_{\max}})$ endowed with the uniform norm, it follows that $\mathcal{C}_0(E, \mathbb{R}) \cap \mathcal{C}_b^1(E, \mathbb{R})$ is also dense in $\mathcal{C}_0(E, \mathbb{R})$ endowed with the uniform norm. Thus, let us consider $f \in \mathcal{C}_0(E, \mathbb{R}) \cap \mathcal{C}_b^1(E, \mathbb{R})$.

According to the Aldous [1] and Rebolledo [18] criteria, in order to prove the tightness of $(\langle \zeta^{H|K}, f \rangle)_{K \geq 1}$, it is enough to show that:

- (i) For any t belonging to a dense subset \mathcal{T} of \mathbb{R}^+ , both $(\langle M^{H|K}(f) \rangle_t)_{K \geq 0}$ and $(V_t^{H|K}(f))_{K \geq 0}$ are tight in \mathbb{R} .
- (ii) For any $T \geq 0$, for any $\epsilon, \alpha > 0$, there exist $\delta > 0$ and $K_0 \in \mathbb{N}$ such that for any two sequences of stopping times $(S_K)_{K \geq 1}$ and $(T_K)_{K \geq 1}$ satisfying $S_K \leq T_K \leq T$ for all integers K ,

$$\sup_{K \geq K_0} \mathbb{P} \left(|\langle M^{H|K}(f) \rangle_{S_K} - \langle M^{H|K}(f) \rangle_{T_K}| \geq \alpha, T_K \leq S_K + \delta \right) \leq \epsilon$$

$$\text{and } \sup_{K \geq K_0} \mathbb{P} \left(|V_{S_K}^{H|K}(f) - V_{T_K}^{H|K}(f)| \geq \alpha, T_K \leq S_K + \delta \right) \leq \epsilon.$$

Notice that, in order to establish (i), it is enough to show that for any $t \geq 0$,

$$\sup_{K \geq 1} \mathbb{E} \left[|\langle M^{H|K}(f) \rangle_t| \right] < \infty \text{ and } \sup_{K \geq 1} \mathbb{E} \left[|V_t^{H|K}(f)| \right] < \infty.$$

Recalling that $C = 2n_{\max}(\lambda_H n_{\max} + \lambda_W n_{\max} + \beta_G)$, it follows from Equation (11) that

$$\mathbb{E}[|\langle M^{H|K}(f) \rangle_t|] \leq \frac{1}{K} 2n_{\max} C \|f\|_{\infty}^2 t.$$

Similar computations yield that

$$\mathbb{E}[|V_t^{H|K}(f)|] \leq \|f\|_{\infty} + \|\mathcal{A}f\|_{\infty} + C \|f\|_{\infty} t.$$

As $f \in \mathcal{C}_b^1(E, \mathbb{R})$, this implies that (i) holds.

It remains to check (ii). Let $\epsilon, \alpha > 0$, and consider two sequences of stopping times $(S_K)_{K \geq 1}$ and $(T_K)_{K \geq 1}$ satisfying $S_K \leq T_K \leq T$ for all integers K . As previously, using Equation (11), we obtain the following upper bound:

$$\mathbb{E} \left[|\langle M^{H|K}(f) \rangle_{S_K} - \langle M^{H|K}(f) \rangle_{T_K}| \mathbb{1}_{T_K \leq S_K + \delta} \right] \leq \mathbb{E} \left[\int_{S_K}^{T_K} dt \mathbb{1}_{T_K \leq S_K + \delta} \right] \frac{1}{K} 2n_{\max} C \|f\|_{\infty}^2$$

$$\leq \frac{\delta}{K} 2n_{\max} C \|f\|_{\infty}^2.$$

Hence, using conditional Markov's inequality,

$$\mathbb{P} \left(|\langle M^{H|K}(f) \rangle_{S_K} - \langle M^{H|K}(f) \rangle_{T_K}| \geq \alpha, T_K \leq S_K + \delta \right) \leq \frac{\delta}{\alpha K} 2n_{\max} C \|f\|_{\infty}^2. \quad (12)$$

Proceeding similarly, we also obtain that

$$\mathbb{P} \left(|V_{S_K}^{H|K}(f) - V_{T_K}^{H|K}(f)| \geq \alpha, T_K \leq S_K + \delta \right) \leq \frac{\delta}{\alpha} (\|\mathcal{A}f\|_{\infty} + C \|f\|_{\infty}). \quad (13)$$

Equations (12) and (13) imply the existence of δ and K_0 such that (ii) is satisfied. Naturally, $\zeta^{W|K}$ can be handled analogously. This concludes the proof. \square

Finally, this result on the tightness of $(\zeta^K)_{K \geq 1}$ in $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(E), w))^2$ lets us establish the main result of this subsection.

Proposition 3.7. *Under the assumptions of Theorem 2.2, the sequence $(\zeta^K)_{K \geq 1}$ is tight in $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(E), w))^2$.*

Proof. Let $X \in \{H, W\}$. Tightness in $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(E), w))$ of $(\zeta^{X|K})_{K \geq 1}$ will be shown using Theorem 1.1.8 from [39], which we state in our setting for the sake of completeness. Let $\Phi : z \in \mathbb{R} \mapsto 6z^2 - 15z^4 + 10z^3$ and for $N \geq 1$, define smooth approximations of $x \in E \mapsto \mathbf{1}_{\{\|\tau(x)\|_\infty \geq N\}}$ by:

$$\forall x \in E, \forall N \geq 1, \phi_N(x) = \Phi(0 \vee (\|\tau(x)\|_\infty - (N - 1)) \wedge 1).$$

Then in order to ensure the tightness of $(\zeta^{X|K})_{K \geq 1}$ in $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(E), w))$, it is sufficient to show that for any $T \geq 0$, the following conditions hold:

(i) There exists a family of functions F which is dense in $\mathcal{C}_0(E, \mathbb{R})$ and stable under addition, such that for any $f \in F \cup \{x \in E \mapsto 1\}$, the sequence $(\langle \zeta^{X|K}, f \rangle)_{K \geq 1}$ is tight in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$.

(ii)

$$\lim_{N \rightarrow \infty} \limsup_{K \rightarrow \infty} \mathbb{E}[\sup_{t \leq T} \langle \zeta_t^{X|K}, \phi_N \rangle] = 0.$$

(iii) Any limiting value of $(\zeta^{X|K})_{K \geq 1}$, if it exists, belongs to $\mathcal{C}([0, T], (\mathcal{M}_F(E), w))$.

The proof hence consists in checking those assumptions. Let $T \geq 0$, and consider $X = H$, as the case $X = W$ can be treated similarly. We may see that (i) is satisfied, as we have shown in the proof of Proposition 3.6 that for any $f \in \mathcal{C}_0(E, \mathbb{R}) \cap \mathcal{C}_b^1(E, \mathbb{R})$, $(\langle \zeta^{H|K}, f \rangle)_{K \geq 1}$ is tight, and further for any $K \geq 1$, for any $T \geq 0$, $\langle \zeta_T^{H|K}, 1 \rangle = 1$ almost surely.

Let us now turn our attention to (ii). Start by noticing that for any $N \geq 1$ and $x \in E$,

$$\phi_N(x) \leq \mathbf{1}_{\{\|\tau(x)\|_\infty \geq N-1\}} \leq f_{N-1}(x) := \sum_{i=1}^{n_{\max}} f_{N-1,i}(x)$$

where $f_{N-1,i}(x) = \mathbf{1}_{\{\mathbf{n}(x) - \mathbf{s}(x) \geq i, |\tau_i(x)| \geq N-1\}}$.

Let $t \in [0, T]$. For any $N \geq t$, $x \in E$, $z \in [0, t]$ and $\sigma \geq 0$, it holds by definition that

$$f_{N-1,i}(\Psi(j(x_k^K(z-), \sigma), t, z)) - f_{N-1,i}(\Psi(x_k^H(z-), t, z)) = \mathbf{1}_{\{\mathbf{n}(x) - \mathbf{s}(x) = i-1, |\sigma - (t-z)| > N-1\}} \leq \mathbf{1}_{\{\sigma > N-1\}}.$$

Hence, using Proposition 1.1 and the above upper bounds, it follows that almost surely,

$$\langle \zeta_t^{H|K}, \phi_N \rangle \leq \frac{1}{K_H} \sum_{k=1}^{K_H} f_{N-1}(\Psi(x_k^H(0), t, 0)) + \frac{n_{\max}}{K_H} \sum_{Y \in \mathcal{S}} \int_0^t \int_{U_Y^K} \mathcal{I}_Y(z-, u) \mathbf{1}_{\{\sigma > N\}} Q_Y^K(dz, du).$$

Defining as previously $C = 2n_{\max}(\lambda_H n_{\max} + \lambda_W n_{\max} + \beta_G)$, this leads to the following upper bound:

$$\mathbb{E}[\sup_{t \leq T} \langle \zeta_t^{H|K}, \phi_N \rangle] \leq \mathbb{E}[\sup_{t \leq T} \frac{1}{K_H} \sum_{k=1}^{K_H} f_{N-1}(\Psi(x_k^H(0), t, 0))] + \frac{CT}{2} n_{\max} \nu((N - 1, +\infty)).$$

As a consequence, Assumption 2.1 (i) ensures that (ii) is satisfied. Notice that this assumption could actually be a little bit relaxed here, as it would be enough if the supremum over K were replaced by the limit superior over $K \rightarrow \infty$.

In order to check that condition (iii) holds, we will follow the arguments presented in [19]. Suppose that η^H is a limiting value of $(\zeta^{H|K})_{K \geq 1}$. By definition,

$$\sup_{t \in [0, T]} \sup_{f \in L^\infty, \|f\|_\infty \leq 1} |\langle \zeta_t^{H|K}, f \rangle - \langle \zeta_{t-}^{H|K}, f \rangle| \leq \frac{1}{K_H}.$$

As the application $\mu \mapsto \sup_{t \in [0, T]} |\langle \mu_t, f \rangle - \langle \mu_{t-}, f \rangle|$ is continuous on $\mathbb{D}([0, T], (\mathcal{M}_F(E), v))$ for any f in a measure-determining countable set, it follows that η^H belongs to $\mathcal{C}([0, T], (\mathcal{M}_F(E), v))$.

Let us now introduce $\phi_{N,M} = \phi_N(1 - \phi_M)$, which serves as a smooth and compactly supported approximation of $x \in E \mapsto \mathbf{1}_{\{N \leq \|\tau(x)\|_\infty \leq M\}}$, for $N \leq M$. As, on the one hand, $\mu \mapsto \sup_{t \in [0, T]} \langle \mu_t, \phi_{N,M} \rangle$ is continuous on $\mathbb{D}([0, T], (\mathcal{M}_F(E), v))$, and on the other hand, for any $K \geq 1$, $\sup_{t \in [0, T]} \langle \zeta_t^{X|K}, \phi_{N,M} \rangle \leq 1$, it follows that:

$$\mathbb{E}[\sup_{t \in [0, T]} \langle \eta_t^H, \phi_{N,M} \rangle] = \lim_{K \rightarrow \infty} \mathbb{E}[\sup_{t \in [0, T]} \langle \zeta_t^{X|K}, \phi_{N,M} \rangle] \leq \limsup_{K \rightarrow \infty} \mathbb{E}[\sup_{t \in [0, T]} \langle \zeta_t^{X|K}, \phi_N \rangle].$$

Letting M go to infinity in the left hand side, dominated convergence ensures that

$$\mathbb{E}[\sup_{t \in [0, T]} \langle \eta_t^H, \phi_N \rangle] \leq \limsup_{K \rightarrow \infty} \mathbb{E}[\sup_{t \in [0, T]} \langle \zeta_t^{X|K}, \phi_N \rangle] \xrightarrow{N \rightarrow \infty} 0,$$

where the convergence of the right hand side is achieved as in the proof of (ii). In particular, it thus is possible to extract a subsequence from $(\sup_{t \in [0, T]} \langle \eta_t^H, \phi_N \rangle)_N$ which converges almost surely to zero. This implies that for any ϵ , there exists N such that almost surely,

$$1 - \sup_{t \in [0, T]} \langle \eta_t^H, \mathbf{1}_{\{\|\tau(\cdot)\|_\infty \leq N\}} \rangle \leq \sup_{t \in [0, T]} \langle \eta_t^H, \phi_N \rangle < \epsilon$$

Thus $(\eta_t^H)_{t \in [0, T]}$ is almost surely tight.

Let $g \in \mathcal{C}_b(E)$, and let $g_N = g(1 - f_N)$. It then holds that for h small so that $t + h \in [0, T]$,

$$|\langle \eta_{t+h}^H, g \rangle - \langle \eta_t^H, g \rangle| \leq |\langle \eta_{t+h}^H, g - g_N \rangle| + |\langle \eta_{t+h}^H, g_N \rangle - \langle \eta_t^H, g_N \rangle| + |\langle \eta_t^H, g - g_N \rangle|.$$

Let $\epsilon > 0$. As $|g - g_N| \leq \|g\|_\infty f_N$, there exists N_0 such that $\sup_{t \in [0, T]} \langle \eta_t^H, g - g_{N_0} \rangle < \epsilon/3$. Further, as $\eta^H \in \mathcal{C}([0, T], (\mathcal{M}_F(E), v))$, for h small enough, $|\langle \eta_{t+h}^H, g_{N_0} \rangle - \langle \eta_t^H, g_{N_0} \rangle| < \epsilon/3$. This allows to conclude that $\eta^H \in \mathcal{C}([0, T], (\mathcal{M}_F(E), w))$, establishing (iii) and finally tightness of η^H in $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(E), w))$. \square

3.1.3 Identification of the limiting values of $(\zeta^K)_{K \geq 1}$

The tightness of $(\zeta^K)_{K \geq 1}$ in the space $\mathbb{D}(\mathbb{R}_+, \mathcal{M}_F(E), w)^2$ ensures that from any subsequence of $(\zeta^K)_{K \geq 1}$, one may extract a subsubsequence which converges in this space. The limits of these subsubsequences may be characterized as follows.

Proposition 3.8. *Under the assumptions of Theorem 2.2, all limiting values of $(\zeta^K)_{K \geq 1}$ in $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(E), w))^2$ are continuous with regard to the total variation norm, and solutions of Equation (4).*

Proof. Consider a subsequence $(\zeta^{\varphi(K)})_{K \geq 1}$ of $(\zeta^K)_{K \geq 1}$ which converges in law in $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(E), w))^2$, and let η be its limit.

Notice that it follows from the Proof of Proposition 3.7 that $\eta \in C(\mathbb{R}_+, (\mathcal{M}_F(E), w))^2$ almost surely. Hence, following Proposition A.6.1 of [38], for any $X \in \{H, W\}$,

$$\|\eta_T^X - \eta_{T-}^X\|_{TV} = \sup_{f \in \mathcal{C}_b(E, \mathbb{R}): \|f\|_\infty \leq 1} |\langle \eta_T^X, f \rangle - \langle \eta_{T-}^X, f \rangle| = 0 \text{ almost surely.}$$

It remains to show that η satisfies Equation (4). Let $T \geq 0$ and $f \in C_b^1(\mathbb{R}_+ \times E, \mathbb{R})$, and consider the application ψ_T^H defined by

$$\begin{aligned} \psi_T^H(\eta) &= \langle \eta_T^H, f_T \rangle - \langle \eta_0^H, f_0 \rangle - \int_0^T \langle \eta_t^H, \mathcal{A}f_t \rangle dt - \int_0^T \langle \eta_t^H, \lambda_H \mathbf{si}(f_t^{\mathcal{I}} - f_t) \rangle dt \\ &\quad - \int_0^T \frac{1}{\langle \eta_t^W, \mathbf{s} \rangle} \langle \eta_t^W, \lambda_W \mathbf{si} \rangle \langle \eta_t^H, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle dt - \beta_G \int_0^T \frac{\langle \eta_t^H, \mathbf{i} \rangle}{\langle \eta_0^H, \mathbf{n} \rangle} \langle \eta_t^H, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle dt. \end{aligned} \tag{14}$$

Start by noticing that $\psi_T^H(\zeta^{\varphi(K)}) = M_T^{H|\varphi(K)}(f)$, as $K_H S^H(t) = K_W S^W(t)$. Using Jensen's inequality, it follows from Equation (11) that

$$\mathbb{E}[|\psi_T^H(\zeta^{\varphi(K)})|^2] \leq \mathbb{E}[|\psi_T^H(\zeta^{\varphi(K)})|^2] = \mathbb{E}[\langle M^{H|K}(f) \rangle_t] \leq \frac{1}{K} 2n_{\max} C \|f\|_{\infty}^2 T \xrightarrow{K \rightarrow \infty} 0.$$

Suppose that $(\psi_T^H(\zeta^{\varphi(K)}))_{K \geq 1}$ converges in law to $\psi_T^H(\zeta)$. According to Theorem 3.5 of [7], it then is enough to prove that $(\psi_T^H(\zeta^{\varphi(K)}))_{K \geq 1}$ is uniformly integrable to obtain that its expectation converges to the expectation of $\psi_T^H(\zeta)$. In our case, uniform integrability is easily assured as the sequence $(\psi_T^H(\zeta^{\varphi(K)}))_{K \geq 1}$ is bounded. Indeed, using the fact that for all $T \geq 0$, $\zeta_T^{\varphi(K)} \in \mathfrak{M}_1$, we obtain from Equation (14) that

$$\left| \psi_T^H(\zeta^{\varphi(K)}) \right| \leq ((2 + C) \|f\|_{\infty} + \|\mathcal{A}f\|_{\infty}) T.$$

We may now conclude that

$$\mathbb{E}[\psi_T^H(\zeta)] = \lim_{K \rightarrow \infty} \mathbb{E}[\psi_T^H(\zeta^{\varphi(K)})] = 0,$$

which yields the desired result.

It thus suffices to show that $(\psi_T^H(\zeta^{\varphi(K)}))_{K \geq 1}$ converges in law to $\psi_T^H(\zeta)$. According to Skorokhod's representation theorem, there exists a probability space Ω on which one may define $(\tilde{\zeta}^K)_{K \geq 1}$ and $\tilde{\eta}$ equal in law to $(\zeta^{\varphi(K)})_{K \geq 1}$ and η , respectively, such that $(\tilde{\zeta}^K)_{K \geq 1}$ converges almost surely in $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(E), w))$ to $\tilde{\eta}$ on Ω . In particular, it holds that

$$\forall T \geq 0, \forall X \in \{H, W\}, \forall g \in \mathcal{C}_b(E), \quad \langle \tilde{\zeta}_T^{X|K}, g \rangle \xrightarrow{K \rightarrow \infty} \langle \tilde{\eta}_T^X, g \rangle \text{ almost surely.}$$

It follows immediately that for any $t \in [0, T]$ and $X \in \{H, W\}$, almost surely,

$$\langle \tilde{\zeta}_t^{X|K}, f_t \rangle, \langle \tilde{\zeta}_t^{X|K}, \mathcal{A}f_t \rangle \xrightarrow{K \rightarrow \infty} \langle \tilde{\eta}_t^X, f_t \rangle, \langle \tilde{\eta}_t^X, \mathcal{A}f_t \rangle. \quad (15)$$

Since $|\langle \tilde{\zeta}_t^{X|K}, \mathcal{A}f_t \rangle| \leq 2 \|Df\|_{\infty}$, where Df designates the differential of f , dominated convergence ensures that

$$\int_0^T \langle \tilde{\zeta}_t^{X|K}, \mathcal{A}f_t \rangle dt \xrightarrow{K \rightarrow \infty} \int_0^T \langle \tilde{\eta}_t^X, \mathcal{A}f_t \rangle dt \text{ almost surely.} \quad (16)$$

In order to establish the desired convergence of the last three terms of $\psi_T^H(\tilde{\zeta})$, we will make use of the following proposition, whose proof is postponed. In this context, a d -dimensional rectangle is a set A defined as the product of d intervals of $\mathbb{R} \cup \{-\infty, +\infty\}$.

Proposition 3.9. *For any $n \in \llbracket 1, n_{\max} \rrbracket$, for any $s \in \llbracket 0, n \rrbracket$, consider $m = m(n, s) < \infty$, a set $(A_k^{n,s})_{k \leq m}$ of $(n-s)$ -dimensional rectangles and a set $(\varphi_k^{n,s})_{k \leq m}$ of functions belonging to $C_b^1(\mathbb{R}^{n-s})$. For any $\tau \in \mathbb{R}^{n_{\max}}$, let $\tau_{1,n-s} = (\tau_1, \dots, \tau_{n-s})$. Define the function $\phi : E \rightarrow \mathbb{R}$ by*

$$\forall x = (n, s, \tau) \in E, \quad \phi(x) = \sum_{k=1}^{m(n,s)} \mathbf{1}_{A_k^{n,s}}(\tau_{1,n-s}) \varphi_k^{n-s}(\tau_{1,n-s}).$$

Then for any $X \in \{H, W\}$ and $T \geq 0$, it holds that

$$\langle \tilde{\zeta}_T^{X|K}, \phi \rangle \xrightarrow{K \rightarrow \infty} \langle \tilde{\eta}_T^X, \phi \rangle \text{ in } L^1.$$

Let us focus on the second-to-last term, representing infection events occurring within work-places, as the other two can be treated similarly.

The application $\phi(x) = \lambda_W \mathbf{s}(x) \mathbf{i}(x)$ is of the form described in Proposition 3.9, hence for any $t \in [0, T]$, $\langle \tilde{\zeta}_t^{W|K}, \lambda_W \mathbf{si} \rangle$ converges in L^1 to $\langle \tilde{\eta}_t^{W|K}, \lambda_W \mathbf{si} \rangle$ as K tends to infinity. Also, notice that as $f \in \mathcal{C}_b^1(\mathbb{R}_+ \times E, \mathbb{R})$, it follows that for any $t \in [0, T]$, $f_t^{\mathcal{I}} \in \mathcal{C}_b^1(E)$. Thus Proposition 3.9 ensures that for any $t \in [0, T]$, $\langle \tilde{\zeta}_t^{H|K}, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle$ converges in L^1 to $\langle \tilde{\eta}_t^{H|K}, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle$ as K tends to infinity.

In particular, the following convergence holds in probability:

$$X_t^K := \langle \tilde{\zeta}_t^{W|K}, \lambda_W \mathbf{si} \rangle \langle \tilde{\zeta}_t^{H|K}, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle \xrightarrow{K \rightarrow \infty} X_t := \langle \tilde{\eta}_t^{W|K}, \lambda_W \mathbf{si} \rangle \langle \tilde{\eta}_t^{H|K}, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle.$$

Letting $c = 2\lambda_W n_{\max}^2 \|f\|_{\infty}$ and $D = \{(x, y) : |x| \leq cy^2\}$, then for any $K \geq 1$, $(X_t^K, \langle \tilde{\zeta}_t^{W|K}, \mathbf{s} \rangle) \in D$ and $(X_t, \langle \tilde{\eta}_t^W, \mathbf{s} \rangle) \in D$, almost surely. As $\langle \tilde{\zeta}_t^{W|K}, \mathbf{s} \rangle$ converges almost surely to $\langle \tilde{\eta}_t^W, \mathbf{s} \rangle$, and the application $(x, y) \mapsto (x/y) \mathbf{1}_{\{y \neq 0\}}$ is continuous on D , we deduce the following convergence in probability:

$$Y_t^K := \frac{\langle \tilde{\zeta}_t^{W|K}, \lambda_W \mathbf{si} \rangle \langle \tilde{\zeta}_t^{H|K}, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle}{\langle \tilde{\zeta}_t^{W|K}, \mathbf{s} \rangle} \xrightarrow{K \rightarrow \infty} Y_t := \frac{\langle \tilde{\eta}_t^{W|K}, \lambda_W \mathbf{si} \rangle \langle \tilde{\eta}_t^{H|K}, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle}{\langle \tilde{\eta}_t^W, \mathbf{s} \rangle}.$$

In addition, for any $K \geq 1$ and any $t \in [0, T]$, $|Y_t^K| \leq 2\lambda_W n_{\max} \|f\|_{\infty}$. Thus using twice dominated convergence, we first obtain that the above convergence of $(Y_t^K)_{K \geq 1}$ to Y_t also holds in L^1 , and subsequently the following L^1 -convergence:

$$\int_0^T \frac{\langle \tilde{\zeta}_t^{W|K}, \lambda_W \mathbf{si} \rangle \langle \tilde{\zeta}_t^{H|K}, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle}{\langle \tilde{\zeta}_t^{W|K}, \mathbf{s} \rangle} dt \xrightarrow{K \rightarrow \infty} \int_0^T \frac{\langle \tilde{\eta}_t^{W|K}, \lambda_W \mathbf{si} \rangle \langle \tilde{\eta}_t^{H|K}, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle}{\langle \tilde{\eta}_t^W, \mathbf{s} \rangle} dt. \quad (17)$$

Reasoning in a similar manner, one also obtains:

$$\begin{aligned} \int_0^T \langle \tilde{\zeta}_t^H, \lambda_H \mathbf{si}(f_t^{\mathcal{I}} - f_t) \rangle dt &\xrightarrow{K \rightarrow \infty} \int_0^T \langle \tilde{\eta}_t^H, \lambda_H \mathbf{si}(f_t^{\mathcal{I}} - f_t) \rangle dt \text{ in } L^1, \\ \int_0^T \frac{\langle \tilde{\zeta}_t^H, \mathbf{i} \rangle \langle \tilde{\zeta}_t^H, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle}{\langle \tilde{\zeta}_t^H, \mathbf{n} \rangle} dt &\xrightarrow{K \rightarrow \infty} \int_0^T \frac{\langle \tilde{\eta}_t^H, \mathbf{i} \rangle \langle \tilde{\eta}_t^H, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle}{\langle \tilde{\eta}_t^H, \mathbf{n} \rangle} dt \text{ in } L^1. \end{aligned} \quad (18)$$

Thus, Equations (15-18) imply that all the terms on the right hand side of the definition of $\psi_T^H(\tilde{\zeta}^K)$ as stated in Equation (14) converge in probability, and thus their linear combination converges in probability to the linear combination of their limits. In other words, $\psi_T^H(\tilde{\zeta}^K)$ converges in probability to $\psi_T^H(\tilde{\eta})$, which ensures as desired that $\psi_T^H(\zeta^K)$ converges in law to $\psi_T^H(\eta)$. This concludes the proof. \square

In order to conclude, we only need to show that Proposition 3.9 holds.

Proof of Proposition 3.9. Step 1. Recall that a d -dimensional rectangle is a set A defined as the product of d intervals of $\mathbb{R} \cup \{-\infty, +\infty\}$. If all d intervals are included in \mathbb{R} , the rectangle will further be said finite.

Let $X \in \{H, W\}$, $n \in \llbracket 1, n_{\max} \rrbracket$ and $s \in \llbracket 0, n \rrbracket$. We start by showing that for any $T \geq 0$ and any finite $(n-s)$ -dimensional rectangle B ,

$$\langle \tilde{\zeta}_T^{X|K}, \mathbf{1}_{\{(n,s)\} \times B} \rangle \xrightarrow{K \rightarrow \infty} \langle \tilde{\eta}_T^{X|K}, \mathbf{1}_{\{(n,s)\} \times B} \rangle \text{ almost surely.} \quad (19)$$

As $\tilde{\eta}^X \in \mathcal{C}(\mathbb{R}_+, (\mathcal{M}_F(E), w))$, it follows that for any T , $\tilde{\zeta}_T^{X|K}$ converges in law to $\tilde{\eta}_T^X$ in $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(E), w))$. Thus, in order to establish the desired result, it is sufficient to show that $\mathbf{B} = \{(n, s)\} \times B$ is a $\tilde{\eta}_T^X$ -continuity set, in which case the Portmanteau theorem allows to conclude.

For any set A , let ∂A be the boundary of A . Then $\partial \mathbf{B} = \{(n, s)\} \times \partial B$. As B is a $(n - s)$ -dimensional rectangle, there exist $a_i < b_i \in \mathbb{R}$ for $i \in \llbracket 1, n - s \rrbracket$ such that B can be written as the product of intervals (potentially open, closed or half-open) delimited by $a_i < b_i$, for $i \in \llbracket 1, n - s \rrbracket$. Thus

$$\partial B = \bigcup_{i=1}^{n-s} \bigcup_{c \in \{a_i, b_i\}} \left(\prod_{j=1}^{i-1} [a_j, b_j] \times \{c\} \times \prod_{k=i+1}^{n-s} [a_k, b_k] \right).$$

Consider any $c \in \mathbb{R}$. We are going to prove that

$$\langle \tilde{\eta}_T^X, \mathbf{1}_{\{(n,s)\} \times (\prod_{j=1}^{i-1} [a_j, b_j] \times \{c\} \times \prod_{k=i+1}^{n-s} [a_k, b_k])} \rangle = 0, \quad (20)$$

which will be enough to conclude. In order to achieve this, let us introduce a mollifier $\psi \in \mathcal{C}^\infty(\mathbb{R})$ in the same sense as in the proof of Lemma 3.1, with compact support in $[-1, 1]$. For $\varepsilon > 0$, define the function $\varphi_\varepsilon : x \mapsto \varepsilon^{-1} \psi(x/\varepsilon)$, whose support lies in $[-\varepsilon, \varepsilon]$ and which converges to δ_0 in the sense of distributions, when ε goes to zero. For any $x = (n, s, \tau) \in E$, let

$$\phi_\varepsilon(x) = \mathbf{1}_{\{\mathbf{n}(x)=n, \mathbf{s}(x)=s\}} \left(\prod_{\substack{j=1 \\ j \neq i}}^{n-s} \mathbf{1}_{[a_j, b_j]} * \varphi_\varepsilon(\tau_j) \right) \mathbf{1}_c * \varphi_\varepsilon(\tau_i).$$

As $\phi_\varepsilon \in \mathcal{C}_b^1(E)$ and $\tilde{\zeta}_T^{X|K}$ converges almost surely to $\tilde{\eta}_T^X$ in $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(E), w))$, dominated convergence implies that

$$\mathbb{E}[\langle \tilde{\eta}_T^X, \phi_\varepsilon \rangle] = \lim_{K \rightarrow \infty} \mathbb{E}[\langle \tilde{\zeta}_T^{X|K}, \phi_\varepsilon \rangle].$$

Notice that

$$\phi_\varepsilon(x) \leq \mathbf{1}_{\{\mathbf{n}(x) - \mathbf{s}(x) > i\}} \mathbf{1}_c * \varphi_\varepsilon(\tau_i(x)).$$

Hence proceeding as in the Proof of Proposition 3.7, it follows that

$$\mathbb{E}[\langle \tilde{\zeta}_T^{X|K}, \phi_\varepsilon \rangle] \leq \mathbb{E} \left[\frac{1}{K_X} \sum_{k=1}^{K_X} \mathbf{1}_{\{n_k^X - s_k^X(0) \geq i, |(\tau_{k,i}^X(0) - T) - c| \leq \varepsilon\}} \right] + \frac{C}{2} \int_0^T \nu([c + (T-t) - \varepsilon, c + (T-t) + \varepsilon]) dt.$$

Absolute continuity of ν with regard to the Lebesgue measure and Assumption 2.1 ensure that the right hand side is dominated by a function $c(\varepsilon)$ which does not depend on K , and which goes to zero with ε . Thus $\mathbb{E}[\langle \tilde{\eta}_T^X, \phi_\varepsilon \rangle] \leq c(\varepsilon)$. In particular, one may construct a sequence $(\varepsilon_n)_{n \geq 1}$ which converges to 0 and satisfies $\sum_{n \geq 0} c(\varepsilon_n) < \infty$. Then on the one hand, the Borel-Cantelli lemma ensures that $\langle \tilde{\eta}_T^X, \phi_\varepsilon \rangle$ converges almost surely to 0 as n tends to infinity. On the other hand, by dominated convergence, $\langle \tilde{\eta}_T^X, \phi_\varepsilon \rangle$ converges almost surely to the left-hand side of Equation (20) as n tends to infinity, hence Equation (20) is proven to be true.

As a consequence, we conclude that $\langle \tilde{\eta}_T^X, \partial \mathbf{B} \rangle = 0$, and thus Equation (19) holds.

Step 2. Consider now a function ϕ as described in the proposition. For any N , let us introduce a partition of \mathbb{R}^{n-s} whose elements consist in (partially open) hypercubes of side length 2^{-N} . For every $k \leq m(n, s)$, a partition $(B_{k,j}^N)_{j \geq 1}$ of $A_k^{n,s}$ is obtained by taking the intersection of $A_k^{n,s}$ with those hypercubes. As $A_k^{n,s}$ is a rectangle itself, the family $(B_{k,j}^N)_{j \geq 1}$ consists of rectangles of side length at most 2^{-N} . For every j , consider a point $z_{k,j}^N$ belonging to

$B_{k,j}^N$. Finally, define the set $J_N(k) = \{j \geq 0 : \sup\{\|x\|_\infty : x \in B_{k,j}^N\} \leq N\}$, which contains only a finite number of elements. Then we can define the following approximation of ϕ :

$$\forall x \in E, \quad \phi_N(x) = \sum_{\substack{1 \leq n \leq n_{\max} \\ 0 \leq s \leq n}} \mathbf{1}_{\{n(x)=n, s(x)=s\}} \sum_{k=1}^{m(n,s)} \sum_{j=1}^{J_N(k)} \varphi_k^{n-s}(z_{k,j}^N) \mathbf{1}_{B_{k,j}^N}(\tau_{1,n-s}(x)).$$

Using our result from the first step, for every (n, s) such that $1 \leq n \leq n_{\max}$ and $0 \leq s \leq n$, for every $k \leq m(n, s)$ and $j \leq J_N(k)$, we obtain that

$$\lim_{K \rightarrow \infty} \langle \tilde{\zeta}_T^{X|K}, \phi_N \rangle = \langle \tilde{\eta}_T^X, \phi_N \rangle \text{ almost surely.} \quad (21)$$

On the one hand, notice that for any $x \in E$ such that $\|\tau(x)\|_\infty > N$, $\phi_N(x) = 0$. Hence for any $x = (n, s, \tau)$, one obtains the following inequality:

$$|\phi_N(x) - \phi(x)| \leq |\phi(x)| \mathbf{1}_{\{\|\tau\|_\infty > N\}} + \sum_{k=1}^{m(n,s)} \sum_{j=1}^{J_N(k)} |\varphi_k^{n-s}(z_{k,j}^N) - \varphi_k^{n-s}(\tau_{1,n-s})| \mathbf{1}_{B_{k,j}^N}(\tau_{1,n-s}). \quad (22)$$

On the one hand, for any such x , it is clear that $\lim_{N \rightarrow \infty} |\phi(x)| \mathbf{1}_{\{\|\tau\|_\infty > N\}} = 0$. On the other hand, as $\varphi_k^{n,s} \in \mathcal{C}_b^1(\mathbb{R}^{n-s})$, the mean value inequality implies that

$$\forall k \leq m(n, s), \forall j \leq J_N(k), \forall x \in B_{k,j}^N, \quad |\varphi_k^{n-s}(z_{k,j}^N) - \varphi_k^{n-s}(\tau_{1,n-s})| \leq \|D\phi\|_\infty d_N,$$

where d_N denotes the maximum of the diameters of d -dimensional hypercubes of side length 2^{-N} , for $d \leq n_{\max}$. In particular, this ensures that for any $x \in E$, $\lim_{N \rightarrow \infty} |\phi_N(x) - \phi(x)| = 0$. Thus, by dominated convergence,

$$\langle \tilde{\eta}_T^X, \phi_N \rangle \xrightarrow{N \rightarrow \infty} \langle \tilde{\eta}_T^X, \phi \rangle. \quad (23)$$

Furthermore, it follows from Equation (22) that for any $K \geq 1$,

$$\mathbb{E}[|\langle \tilde{\zeta}_T^{X|K}, \phi_N \rangle - \langle \tilde{\zeta}_T^{X|K}, \phi \rangle|] \leq \|\phi\|_\infty \sup_{K \geq 1} \mathbb{E}[\langle \tilde{\zeta}_T^{X|K}, \mathbf{1}_{\{\|\tau(\cdot)\|_\infty > N\}} \rangle] + \|D\phi\|_\infty d_N.$$

Reasoning as in the proof of Proposition 3.7, and using Assumption 2.1, we obtain that

$$\lim_{N \rightarrow \infty} \sup_{K \geq 1} \mathbb{E}[\langle \tilde{\zeta}_T^{X|K}, \mathbf{1}_{\{\|\tau(\cdot)\|_\infty > N\}} \rangle] = 0,$$

and as a consequence,

$$\lim_{N \rightarrow \infty} \sup_{K \geq 1} \mathbb{E}[|\langle \tilde{\zeta}_T^{X|K}, \phi_N \rangle - \langle \tilde{\zeta}_T^{X|K}, \phi \rangle|] = 0. \quad (24)$$

Noticing that

$$\mathbb{E}[|\langle \tilde{\zeta}_T^{X|K}, \phi \rangle - \langle \tilde{\eta}_T^X, \phi \rangle|] \leq \mathbb{E}[|\langle \tilde{\zeta}_T^{X|K}, \phi - \phi_N \rangle|] + \mathbb{E}[|\langle \tilde{\zeta}_T^{X|K}, \phi_N \rangle - \langle \tilde{\eta}_T^X, \phi_N \rangle|] + \mathbb{E}[|\langle \tilde{\eta}_T^X, \phi_N - \phi \rangle|]$$

and using Equation (24), as well as dominated convergence applied to Equations (21) and (23), finally yields that

$$\lim_{K \rightarrow \infty} \mathbb{E}[|\langle \tilde{\zeta}_T^{X|K}, \phi \rangle - \langle \tilde{\eta}_T^X, \phi \rangle|] = 0.$$

This concludes the proof. \square

3.1.4 Proof of Theorem 2.2

The previous results are sufficient to establish Theorem 2.2. Indeed, it follows from Propositions 3.7 and 3.8 that from every subsequence of $(\zeta^K)_{K \geq 1}$, one may extract a subsubsequence converging in $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(E), w))^2$ to a solution of Equation (4) which is continuous with respect to the total variation norm. As by assumption, ζ_0^K converges in law to $\eta_0 \in \mathfrak{M}_1$, Proposition 3.2 implies that all of these subsubsequences converge to the same limit η , which is the unique solution of Equation (4) with initial condition η_0 . As $\eta_0 \in \mathfrak{M}_1$, Proposition 3.2 further ensures that $\eta \in \mathbb{D}(\mathbb{R}_+, \mathfrak{M}_1)$. This establishes the convergence of $(\zeta^K)_{K \geq 1}$ to η in $\mathbb{D}(\mathbb{R}_+, \mathfrak{M}_1)$, as desired.

3.2 Proof of Theorem 2.3

This section is devoted to extracting dynamical system (6a-c) from the measure-valued integral equation (4), under the assumption that ν is the exponential distribution of parameter γ . From now on, we suppose that this assumption is satisfied, and further that $\eta_0 = \eta_{0,\varepsilon}$ as given by (5).

3.2.1 Preliminary study of the dynamical system

Before establishing Theorem 2.3 itself, let us start by showing that the solution of dynamical system (6a-c) endowed with initial condition (7) admits at most a unique solution. Existence will follow from the proofs of the forthcoming subsections, since they construct a solution to the dynamical system.

For this section, let us rewrite dynamical system (6a-c) as follows, in order to emphasize the associated Cauchy problem. Recall that the dynamical system is of dimension $d = 2 + 2\#\mathbb{S} = n_{\max}(n_{\max} + 1)$.

Let $y \in \mathcal{C}^1(\mathbb{R}_+, \mathbb{R}^d)$ and $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$ be defined such that dynamical system (6a-c) amounts to

$$y'(t) = f(y(t)) \quad \forall t \geq 0. \quad (25)$$

The components of y (and resp. f) will be called s , i and $n_{(S,I)}^X$ (resp. f_s , f_i and $f_{X;(S,I)}$) for $X \in \{H, W\}$ and $(S, I) \in \mathbb{S}$, in order to simplify their identification with the unknowns of the corresponding dynamical system. More precisely, consider the applications

$$\tau_X(y) = -\frac{\lambda_X}{m_X} \sum_{(S,I) \in \mathbb{S}} SI n_{(S,I)}^X \text{ for } X \in \{H, W\}, \text{ and } \tau_G(y) = \beta_G i.$$

Then $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is defined as follows, for any $y = (s, i, n_{(S,I)}^X : X \in \{H, W\}, (S, I) \in \mathbb{S}) \in \mathbb{R}^d$:

$$f_s(y) = -(\tau_H(y) + \tau_W(y) + \tau_G(y)s) \text{ and } f_i(y) = -f_s(y) - \gamma i,$$

while for all $X \in \{H, W\}$ and $(S, I) \in \mathbb{S}$

$$\begin{aligned} f_{X;(S,I)}(y) = & - \left[\left(\lambda_X I + \frac{\tau_{\bar{X}}(y)}{s} + \tau_G(y) \right) S - \gamma I \right] n_{(S,I)}^X + \gamma(I+1) n_{(S,I+1)}^X \mathbf{1}_{\{S+I < n_{\max}\}} \\ & + \left(\lambda_X(I-1) + \frac{\tau_{\bar{X}}(y)}{s} + \tau_G(y) \right) (S+1) n_{(S+1,I-1)}^X \mathbf{1}_{\{I \geq 1\}}. \end{aligned}$$

Also, notice that there are some natural constraints that we expect the solution of dynamical system (6a-c) to satisfy. Clearly, s , i and $n_{(S,I)}^X$ should belong to $[0, 1]$. Also, as the population is partitioned into susceptible, infected and removed individuals, it follows that $s+i \leq 1$. Similarly, as all individuals belong to exactly one household and one workplace, and as $n_{(S,I)}^X$ corresponds

to the proportion of structures of type X which contain S susceptible and I infected individuals, we expect that for $X \in \{H, W\}$,

$$\sum_{(S,I) \in \mathbb{S}} n_{(S,I)}^X \leq 1, \text{ and } \sum_{(S,I) \in \mathbb{S}} S n_{(S,I)}^X \leq m_X s. \quad (26)$$

We thus define the following set $V \subset \mathbb{R}^d$, which formalizes these constraints:

$$V = \left\{ y \in [0, 1]^d : s + i \leq 1, \sum_{(S,I) \in \mathbb{S}} n_{(S,I)}^X \leq 1 \text{ and } m_X s - \sum_{(S,I) \in \mathbb{S}} S n_{(S,I)}^X \geq 0 \quad \forall X \in \{H, W\} \right\}.$$

Proposition 3.10. *Let $y^* \in V$. Then the following assertions hold:*

- (i) *Suppose that there exists a solution y of the Cauchy problem (25) with initial condition $y(0) = y^*$. Then $y(t) \in V$ for any $t \geq 0$ for which y is well defined.*
- (ii) *At most, there exists a unique such solution y .*
- (iii) *In particular, for any $\varepsilon > 0$, the dynamical system (6a-c) endowed with initial condition (7) admits at most a unique solution.*

The proof of this proposition is available in Supplementary Material S2. It relies on establishing the Lipschitz continuity of f on V , from which uniqueness is deduced using Gronwall's lemma.

3.2.2 Absolute continuity

Consider any $n \in \llbracket 1, n_{\max} \rrbracket$ and $s \in \llbracket 0, n \rrbracket$. The aim of this section is to establish that conditionally on a structure being of size n and containing s susceptibles, the distribution of the remaining infectious periods of its $n - s$ members who have contracted the disease at some previous time, is absolutely continuous with respect to the Lebesgue measure on \mathbb{R}^{n-s} . More precisely, we will see that at any time $t \geq 0$, there exists a function $\rho_t^{X,n,s} : \mathbb{R}^{n-s} \rightarrow \mathbb{R}_+$ such that for any non-negative measurable function $f : E \rightarrow \mathbb{R}$, for any $X \in \{H, W\}$ and $t \geq 0$,

$$\langle \eta_t^X, \mathbf{1}_{\left\{ \begin{array}{l} \mathbf{n}(\bullet) = n \\ \mathbf{s}(\bullet) = s \end{array} \right\}} f(\bullet) \rangle = \int_{\mathbb{R}^{n-s}} f(n, s, \tau) \rho_t^{X,n,s}(\tau) d\tau.$$

In the following, for any $n \geq 1$, let 0_n be the zero of \mathbb{R}^n . Further, for $k \in \llbracket 1, n \rrbracket$ and $\tau \in \mathbb{R}^n$, let $\tau = (\tau_1, \dots, \tau_n)$ and $\tau_{1,k} = (\tau_1, \dots, \tau_k)$. Finally, $(e_k)_{1 \leq k \leq n_{\max}}$ is the canonical basis of $\mathbb{R}^{n_{\max}}$.

Lemma 3.11. *For each $X \in \{H, W\}$, there exists a family of measurable functions indexed by $n \in \llbracket 1, n_{\max} \rrbracket$ and $s \in \llbracket 0, n \rrbracket$ such that*

$$\forall s \in \llbracket 0, n - 1 \rrbracket, \rho^{X,n,s} : \mathbb{R}_+ \times \mathbb{R}^{n-s} \rightarrow \mathbb{R}_+ \\ (t, \tau) \mapsto \rho_t^{X,n,s}(\tau),$$

which verifies that for any $f : E \rightarrow \mathbb{R}$ non-negative measurable function,

$$\langle \eta_t^X, f \rangle = \sum_{n=1}^{n_{\max}} \left(f(n, n, 0_{n_{\max}}) \rho_t^{X,n,n} + \sum_{s=0}^{n-1} \int_{\mathbb{R}^{n-s}} f(n, s, \sum_{k=1}^{n-s} \tau_k e_k) \rho_t^{X,n,s}(\tau) d\tau \right).$$

The proof relies on studying the distribution η^X conditionally on structures being of given size and number of susceptibles. Absolute continuity with regard to the Lebesgue measure is then established by domination with another, absolutely continuous measure. We refer to Supplementary Material [S2](#) for details.

Notice that in the case $n = s$, $\rho^{X,n,n}$ depends on t alone, whereas for $s < n$, $\rho^{X,n,s}$ depends both on t and on $\tau \in \mathbb{R}^{n-s}$. Throughout the following, with some abuse of notation, this distinction will mostly be implicit. The adaptation to the case $s = n$ is generally straightforward. For instance, forthcoming Equation (27) reduces to

$$\frac{d}{dt}\rho_t^{X,n,n} = -\Lambda_t^X n \rho_t^{X,n,n},$$

where

$$\Lambda_t^X = \lambda_{\overline{X}} \frac{1}{s_{\overline{X}}} \langle \eta_{\overline{X}}, \mathbf{s}_i \rangle + \beta_G \frac{i_t^H}{nH}.$$

When integrating over τ , the case $n = s$ becomes

$$\int_{\mathbb{R}^0} \rho_t^{X,n,n}(\tau) d\tau = \rho_t^{X,n,n}.$$

Similarly, symbolically, for any $\tau \in \mathbb{R}^{n-s}$, $i(\tau) = 0$ whenever $n - s = 0$. This makes sense, as in a structure of size n containing n susceptibles, there are no infected. This leads for example to

$$\int_{\mathbb{R}^0} \mathbf{1}_{\{i(\tau)=0\}} \rho_t^{X,n,n}(\tau) d\tau = \rho_t^{X,n,n}.$$

When the adaptation is less clear, the case $n = s$ will be treated separately.

Proposition 3.12. *The family $(\rho^{X,n,s})_{1 \leq n \leq n_{\max}, 0 \leq s \leq n}$ is a weak solution to the following system of partial differential equations: for any $(n, s) \in \llbracket 1, n_{\max} \rrbracket \times \llbracket 0, n \rrbracket$, for any $\tau \in \mathbb{R}^{n-s}$ and $t \geq 0$,*

$$\begin{aligned} \partial_t \rho_t^{X,n,s}(\tau) - \sum_{k=1}^{n-s} \partial_{\tau_k} \rho_t^{X,n,s}(\tau) &= -s(\lambda_X i(\tau) + \Lambda_t^X) \rho_t^{X,n,s}(\tau) \\ &\quad + \mathbf{1}_{\{s+1 \leq n\}} \lambda_X (s+1) i(\tau_{1,n-s-1}) \rho_t^{X,n,s+1}(\tau_{1,n-s-1}) \mathbf{1}_{\{\tau_{n-s} > 0\}} \gamma e^{-\gamma \tau_{n-s}} \\ &\quad + \mathbf{1}_{\{s+1 \leq n\}} \Lambda_t^X (s+1) \rho_t^{X,n,s+1}(\tau_{1,n-s-1}) \mathbf{1}_{\{\tau_{n-s} > 0\}} \gamma e^{-\gamma \tau_{n-s}}, \end{aligned} \quad (27)$$

with initial condition given by

$$\rho_0^{X,n,n} = \pi_n^X \varepsilon^n \text{ and } \forall s < n, \tau \in \mathbb{R}^{n-s}, \rho_0^{X,n,s}(\tau) = \pi_n^X \binom{n}{s} \varepsilon^s (1 - \varepsilon)^{n-s} \mathbf{1}_{\{\tau \in (\mathbb{R}_+^*)^{n-s}\}} e^{-\gamma \sum_{k=1}^{n-s} \tau_k}.$$

Proof. Let $(n, s) \in \llbracket 1, n_{\max} \rrbracket \times \llbracket 0, n \rrbracket$. In the case $s = n$, Equation (27) follows directly from Equation (4) applied to $f(x) = \mathbf{1}_{\{\mathbf{n}(x)=n, \mathbf{s}(x)=s\}}$ for all $x \in E$ as $\mathcal{A}f = f^{\mathcal{I}} = 0$.

Suppose now that $s < n$. Consider f measurable on E such that $f(x) = 0$ if $(\mathbf{n}(x), \mathbf{s}(x)) \neq (n, s)$ and $f(x) = f_{n,s}(\tau_{1,n-s}(x))$ otherwise, where $f_{n,s} \in \mathcal{C}^\infty(\mathbb{R}^{n-s})$ is compactly supported. In particular, for any $\sigma > 0$ and $x \in E$ such that $\mathbf{s}(x) \geq 1$, notice that $f(j(x, \sigma))$ is zero, unless $s + 1 \leq n$ and $(\mathbf{n}(x), \mathbf{s}(x)) = (n, s + 1)$ in which case it is equal to $f_{n,s}(\tau_{1,n-s-1}(x) + \sigma e_{n-s})$. Recalling that $f^{\mathcal{I}}(x) = \langle \nu, f(j(x, \cdot)) \rangle$, injecting f into Equation (4) and differentiating with regard to t thus yields

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^{n-s}} f_{n,s}(\tau) \rho_t^{X,n,s}(\tau) d\tau &= - \sum_{k=1}^{n-s} \int_{\mathbb{R}^{n-s}} \partial_{\tau_k} f_{n,s}(\tau) \rho_t^{X,n,s}(\tau) d\tau \\ &\quad - \int_{\mathbb{R}^{n-s}} (\lambda_X s i(\tau) + \Lambda_t^X s) f_{n,s}(\tau) \rho_t^{X,n,s}(\tau) d\tau + \mathbf{1}_{\{s=n-1\}} \Lambda_t^X n \left(\int_0^{+\infty} f_{n,n-1}(\sigma e_1) \gamma e^{-\gamma \sigma} d\sigma \right) \rho_t^{X,n,n} \\ &\quad + \mathbf{1}_{\{s+1 < n\}} \int_{\mathbb{R}^{n-s-1}} (\lambda_X i(u) + \Lambda_t^X)(s+1) \left(\int_0^{+\infty} f_{n,s}(u + \sigma e_{n-s}) \gamma e^{-\gamma \sigma} d\sigma \right) \rho_t^{X,n,s+1}(u) du. \end{aligned}$$

An integration by parts leads to the following equality, as $f_{n,s}$ is of compact support which implies that the boundary terms of the integration by parts vanish:

$$-\sum_{k=1}^{n-s} \int_{\mathbb{R}^{n-s}} \partial_{\tau_k} f_{n,s}(\tau) \rho_t^{X,n,s}(\tau) d\tau = \sum_{k=1}^{n-s} \int_{\mathbb{R}^{n-s}} f_{n,s}(\tau) \partial_{\tau_k} \rho_t^{X,n,s}(\tau) d\tau.$$

Injecting this into the previous computation yields that $\rho^{X,n,s}$ is a weak solution to Equation (27).

Finally, initial conditions $\rho_0^{X,n,s}$ for $n \in \llbracket 1, n_{\max} \rrbracket$ and $s \in \llbracket 0, n \rrbracket$ follow directly from (5). \square

3.2.3 Proof of Theorem 2.3

Before proceeding, let us introduce the following notation. For $f \in \mathcal{C}_b^1(\mathbb{R}_+ \times E, \mathbb{R})$ and $T \geq t \geq 0$, define

$$f_{T,t} : x \in E \mapsto f(T, \Psi(x, T, t)) \text{ and } f_{T,t}^{\mathcal{I}} : x \in E \mapsto \langle \nu, f_{T,t}(\mathbf{j}(x, \cdot)) \rangle. \quad (28)$$

We can now state an intermediate result, which is similar in spirit to Proposition 3.4.

Proposition 3.13. *Let η be the unique solution in $\mathcal{C}(\mathbb{R}_+, \mathfrak{M}_1)$ of Equation (4). Then for any $f \in \mathcal{C}_b^1(\mathbb{R}_+ \times E, \mathbb{R})$, for $X \in \{H, W\}$,*

$$\begin{aligned} \langle \eta_T^X, f_T \rangle &= \langle \eta_0^X, f_{T,0} \rangle + \lambda_X \int_0^T \langle \eta_t^X, \mathbf{si}(f_{T,t}^{\mathcal{I}} - f_{T,t}) \rangle dt \\ &+ \lambda_{\bar{X}} \int_0^T \frac{1}{s_t^{\bar{X}}} \langle \eta_t^{\bar{X}}, \mathbf{si} \rangle \langle \eta_t^X, \mathbf{s}(f_{T,t}^{\mathcal{I}} - f_{T,t}) \rangle dt + \beta_G \int_0^T \frac{i_t^H}{n^H} \langle \eta_t^X, \mathbf{s}(f_{T,t}^{\mathcal{I}} - f_{T,t}) \rangle dt. \end{aligned} \quad (29)$$

Proof. The proof follows the exact same lines as the proof of Proposition 3.4, showing that for any $f \in \mathcal{C}_b^1(\mathbb{R}_+ \times E)$, Equation (29) leads to Equation (4) using Lemma 3.3. \square

Throughout this section, let h be the Heaviside step function, *i.e.* $h(z) = \mathbf{1}_{\{z > 0\}}$ for any real number z . Also, consider a mollifier ψ on \mathbb{R} in the same sense as in the Proof of Lemma 3.1. For $\alpha > 0$, define $\psi_\alpha(z) = \alpha^{-1} \psi(z/\alpha)$ for all $z \in \mathbb{R}$. Let us introduce, for any $\alpha > 0$, the function $h_\alpha : \mathbb{R} \rightarrow [0, 1]$ defined by

$$\forall z \in \mathbb{R}, \quad h_\alpha(z) = h * \psi_\alpha(z).$$

As ψ_α converges to δ_0 in the sense of distributions when α tends to zero, the families of functions $(h_\alpha)_{\alpha > 0}$ and $(\psi_\alpha)_{\alpha > 0}$ serve as smooth approximations of h and δ_0 respectively. This will be useful in the proofs of this section.

This allows us to establish the following proposition, which serves as a starting point of the proof of Theorem 2.3.

Proposition 3.14. *Under the assumptions of Theorem 2.3, it holds that for any $X \in \{H, W\}$, $(s_t^X)_{t \geq 0}$ and $(i_t^X)_{t \geq 0}$ satisfy*

$$\begin{aligned} \frac{d}{dt} s_t^X &= -n^X \left(\frac{\lambda_X}{n^{\bar{X}}} \langle \eta_t^X, \mathbf{si} \rangle + \frac{\lambda_{\bar{X}}}{n^{\bar{X}}} \langle \eta_t^{\bar{X}}, \mathbf{si} \rangle + \beta_G \frac{i_t^H}{n^H} \frac{s_t^X}{n^{\bar{X}}} \right), \\ \frac{d}{dt} i_t^X &= -\frac{d}{dt} s_t^X + \gamma i_t^X. \end{aligned} \quad (30)$$

Further,

$$s_0^X = (1 - \varepsilon) n^X \text{ and } i_0^X = \varepsilon n^X. \quad (31)$$

Proof. Notice that $\mathbf{s} \in \mathcal{C}_b^1(E, \mathbb{R})$ is such that $\mathcal{A}\mathbf{s} = 0$ and $\mathbf{s}^{\mathcal{I}}(x) - \mathbf{s}(x) = -1$ for all $x \in E$. It thus follows immediately from Equation (4) that, for any $X \in \{H, W\}$,

$$s_T^X = s_0^X - \int_0^T \left(\lambda_X \langle \eta_t^X, \mathbf{si} \rangle + \lambda_{\bar{X}} \frac{s_t^X}{s_t^{\bar{X}}} \langle \eta_t^{\bar{X}}, \mathbf{si} \rangle + \beta_G \frac{i_t^H}{n^H} s_t^X \right) dt.$$

Further, since $\eta \in \mathcal{C}(\mathbb{R}_+, (\mathfrak{M}_1, \|\cdot\|_{TV}))$, it follows that for any $t \geq 0$, $(\zeta_t^K)_{K \geq 1}$ converges in law to η_t . As \mathbf{s} is continuous and bounded on E , this implies that $\langle \zeta_t^{X|K}, \mathbf{s} \rangle \rightarrow \langle \eta_t^X, \mathbf{s} \rangle$ when K tends to infinity. The analogous result holds for \mathbf{n} . Hence Lemma 1.2 ensures that, for any $t \geq 0$, $s_t^X/s_t^{\bar{X}} = n^X/n^{\bar{X}}$. In other words,

$$s_T^X = s_0^X - n^X \int_0^T \left(\frac{\lambda_X}{n^X} \langle \eta_t^X, \mathbf{si} \rangle + \frac{\lambda_{\bar{X}}}{n^{\bar{X}}} \langle \eta_t^{\bar{X}}, \mathbf{si} \rangle + \beta_G \frac{i_t^H}{n^H} \frac{s_t^X}{n^X} \right) dt.$$

As $\eta \in \mathcal{C}(\mathbb{R}_+, (\mathfrak{M}_1, \|\cdot\|_{TV}))$, and \mathbf{s} and \mathbf{i} are bounded measurable functions, it follows that the integrand is continuous with regard to t . Thus, the first line of Equation (30) comes from the fundamental theorem of calculus.

Recall that $\eta_0 = \eta_{0,\varepsilon}$ as defined in Equation (5). In particular, we now have $n^X = m_X$, hence

$$s_0^X = \langle \eta_{0,\varepsilon}^X, \mathbf{s} \rangle = \sum_{n=1}^{n_{\max}} \pi_n^X \sum_{s=0}^n s \binom{n}{s} (1-\varepsilon)^s \varepsilon^{n-s} = (1-\varepsilon) \sum_{n=1}^{n_{\max}} n \pi_n^X = (1-\varepsilon) n^X. \quad (32)$$

This yields the first part of Equation (31).

It remains to take an interest in i_t^X . As \mathbf{i} does not belong to $\mathcal{C}_b^1(E, \mathbb{R})$ we cannot proceed in the same way. Remember that $\mathbf{i}(x) = \sum_{j=1}^{n_{\max}} h(\tau_j)$ for $x = (n, s, \tau) \in E$. Let us introduce, for any $\alpha > 0$, the function $\mathbf{i}_\alpha : E \rightarrow \llbracket 1, n_{\max} \rrbracket$ defined by

$$\forall x = (n, s, \tau) \in E, \quad \mathbf{i}_\alpha(x) = \sum_{j=1}^{n_{\max}} h_\alpha(\tau_j).$$

Fix $\alpha > 0$. As ψ'_α is bounded, it follows from the usual properties of convolution that $h_\alpha \in \mathcal{C}_b^1(\mathbb{R})$ and hence $\mathbf{i}_\alpha \in \mathcal{C}_b^1(E, \mathbb{R})$. Thus, we may apply Proposition 3.13 and obtain that

$$\begin{aligned} \langle \eta_T^X, \mathbf{i}_\alpha \rangle &= \langle \eta_0^X, (\mathbf{i}_\alpha)_{T,0} \rangle + \lambda_X \int_0^T \langle \eta_t^X, \mathbf{si}((\mathbf{i}_\alpha)_{T,t}^{\mathcal{I}} - (\mathbf{i}_\alpha)_{T,t}) \rangle dt \\ &+ \lambda_{\bar{X}} \int_0^T \frac{1}{s_t^{\bar{X}}} \langle \eta_t^{\bar{X}}, \mathbf{si} \rangle \langle \eta_t^X, \mathbf{s}((\mathbf{i}_\alpha)_{T,t}^{\mathcal{I}} - (\mathbf{i}_\alpha)_{T,t}) \rangle dt + \beta_G \int_0^T \frac{i_t^H}{n^H} \langle \eta_t^X, \mathbf{s}((\mathbf{i}_\alpha)_{T,t}^{\mathcal{I}} - (\mathbf{i}_\alpha)_{T,t}) \rangle dt. \end{aligned}$$

Notice that $\|\mathbf{i}_\alpha\|_\infty = n_{\max}$ for any $\alpha > 0$, as well as $\|\mathbf{s}\|_\infty = n_{\max}$, $\|\mathbf{i}\|_\infty = n_{\max}$, $\|\mathbf{si}\|_\infty \leq n_{\max}^2$ and $\langle \eta_t^{\bar{X}}, \mathbf{si} \rangle / s_t^{\bar{X}} \leq n_{\max}$ for any $t \geq 0$. As further $\eta_t \in \mathfrak{M}_1$ for any $t \geq 0$, we may let α go to zero and use dominated convergence to obtain that

$$\begin{aligned} \langle \eta_T^X, \mathbf{i} \rangle &= \langle \eta_0^X, \mathbf{i}_{T,0} \rangle + \lambda_X \int_0^T \langle \eta_t^X, \mathbf{si}(\mathbf{i}_{T,t}^{\mathcal{I}} - \mathbf{i}_{T,t}) \rangle dt \\ &+ \lambda_{\bar{X}} \int_0^T \frac{1}{s_t^{\bar{X}}} \langle \eta_t^{\bar{X}}, \mathbf{si} \rangle \langle \eta_t^X, \mathbf{s}(\mathbf{i}_{T,t}^{\mathcal{I}} - \mathbf{i}_{T,t}) \rangle dt + \beta_G \int_0^T \frac{i_t^H}{n^H} \langle \eta_t^X, \mathbf{s}(\mathbf{i}_{T,t}^{\mathcal{I}} - \mathbf{i}_{T,t}) \rangle dt. \end{aligned} \quad (33)$$

Further, for any $\sigma \geq 0$, $T \geq t \geq 0$ and $x \in E$,

$$i(\Psi(j(x, \sigma), T, t)) - i(\Psi(x, T, t)) = \mathbf{1}_{\{\sigma > (T-t)\}},$$

hence

$$\mathbf{i}_{T,t}^X(x) - \mathbf{i}_{T,t}(x) = \nu([T-t, \infty)) = e^{-\gamma(T-t)}.$$

Injecting this into Equation (33) and using as before that $s_t^X/s_t^{\bar{X}} = n^X/n^{\bar{X}}$ yields

$$i_T^X = \langle \eta_0^X, \mathbf{i}_{T,0} \rangle + n^X e^{-\gamma T} \int_0^T e^{\gamma t} \left(\frac{\lambda_X}{n^X} \langle \eta_t^X, \mathbf{si} \rangle + \frac{\lambda_{\bar{X}}}{n^{\bar{X}}} \langle \eta_t^{\bar{X}}, \mathbf{si} \rangle + \beta_G \frac{i_t^H}{n^H} \frac{s_t^X}{n^X} \right) dt. \quad (34)$$

As $\eta_0 = \eta_{0,\varepsilon}$, we may compute the first term of the right-hand side of this equation and obtain that

$$\langle \eta_0^X, \mathbf{i}_{T,0} \rangle = \langle \eta_{0,\varepsilon}^X, \mathbf{i}_{T,0} \rangle = e^{-\gamma T} \sum_{n=1}^{n_{\max}} \pi_n^X \sum_{s=0}^n \binom{n}{s} (1-\varepsilon)^s \varepsilon^{n-s} (n-s) = e^{-\gamma T} \varepsilon n^X. \quad (35)$$

Using the continuity of the integrand in Equation (34), we may now differentiate it with regard to T :

$$\frac{d}{dT} i_T^X = n^X \left(\frac{\lambda_X}{n^X} \langle \eta_T^X, \mathbf{si} \rangle + \frac{\lambda_{\bar{X}}}{n^{\bar{X}}} \langle \eta_T^{\bar{X}}, \mathbf{si} \rangle + \beta_G \frac{i_T^H}{n^H} \frac{s_T^X}{n^X} \right) - \gamma i_T^X = -\frac{d}{dT} s_T^X - \gamma i_T^X.$$

We thus have recovered the second line of Equation (30). Finally, the second half of Equation (31) is obtained by a computation analogous to (32). This concludes the proof. \square

For $(S, I) \in \mathbb{S}$, define the function $f^{(S,I)} : E \rightarrow \{0, 1\}$ by

$$f^{(S,I)}(x) = \mathbf{1}_{\{\mathbf{s}(x)=S, \mathbf{i}(x)=I\}}.$$

For $t \geq 0$, let $n_{(S,I)}^X(t) = \langle \eta_t^X, f^{(S,I)} \rangle$, which defines a continuous function on \mathbb{R}_+ as $f^{(S,I)}$ is bounded and measurable and $\eta \in \mathcal{C}(\mathbb{R}_+, (\mathfrak{M}_1, \|\cdot\|_{TV}))$. In words, this corresponds to the proportion of structures of type X which contain exactly S susceptible and I infected individuals. Notice that

$$\{x \in E : \mathbf{s}(x)\mathbf{i}(x) > 0\} = \{x \in E : \mathbf{s}(x)\mathbf{i}(x) > 0, (\mathbf{s}(x), \mathbf{i}(x)) \in \mathbb{S}\}.$$

We may thus rewrite the first line of Equation (30) as follows:

$$\frac{d}{dt} s_t^X = -n^X \left(\frac{\lambda_X}{n^X} \sum_{(S,I) \in \mathbb{S}} S I n_{(S,I)}^X(t) + \frac{\lambda_{\bar{X}}}{n^{\bar{X}}} \sum_{(S,I) \in \mathbb{S}} S I n_{(S,I)}^{\bar{X}}(t) + \beta_G \frac{i_t^H}{n^H} \frac{s_t^X}{n^X} \right).$$

Similarly, it holds that

$$\Lambda_t^X = \frac{\lambda_{\bar{X}}}{s_t^{\bar{X}}} \langle \eta_t^{\bar{X}}, \mathbf{si} \rangle + \beta_G \frac{i_t^H}{n^H} = \frac{\lambda_{\bar{X}}}{s_t^{\bar{X}}} \sum_{(S,I) \in \mathbb{S}} S I n_{(S,I)}^{\bar{X}}(t) + \beta_G \frac{i_t^H}{n^H},$$

which may also be written in terms of the notations of Equation (6a-c):

$$\Lambda_t^X = \left(\frac{s_t^{\bar{X}}}{n^{\bar{X}}} \right)^{-1} \tau_{\bar{X}}(t) + \beta_G \frac{i_t^H}{n^H}.$$

This motivates a closer study of the functions $n_{(S,I)}^X$ for $(S, I) \in \mathbb{S}$.

Proposition 3.15. *Let $X \in \{H, W\}$ and $(S, I) \in \mathbb{S}$. Under the assumptions of Theorem 2.3, it holds that*

$$\begin{aligned} \frac{d}{dt} n_{(S,I)}^X(t) &= \gamma \left((I+1)n_{(S,I+1)}^X(t) \mathbf{1}_{\{S+I < n_{\max}\}} - I n_{(S,I)}^X(t) \right) \\ &\quad + \lambda_X \left((S+1)(I-1)n_{(S+1,I-1)}^X(t) \mathbf{1}_{\{I \geq 1\}} - S I n_{(S,I)}^X(t) \right) \\ &\quad + \Lambda_t^X \left((S+1)n_{(S+1,I-1)}^X(t) \mathbf{1}_{\{I \geq 1\}} - S n_{(S,I)}^X(t) \right). \end{aligned} \quad (36)$$

Further

$$n_{(S,I)}^X(0) = \binom{S+I}{I} \pi_{S+I}^X (1-\varepsilon)^S \varepsilon^I. \quad (37)$$

Proof. Let $X \in \{H, W\}$ and $(S, I) \in \mathbb{S}$. It follows from Lemma 3.11 that

$$n_{(S,I)}^X(t) = \langle \eta_t^X, f^{(S,I)} \rangle = \sum_{n=S+I}^{n_{\max}} \int_{\mathbb{R}^{n-S}} \mathbf{1}_{\{i(\tau)=I\}} \rho_t^{X,n,S}(\tau) d\tau,$$

where in the second equality, the sum over n starts at $S+I$ as it is impossible to observe S susceptible and I infected members in a group of size $n < S+I$. Using Proposition 3.12, we may deduce that

$$\frac{d}{dt} n_{(S,I)}^X(t) = \sum_{n=S+I}^{n_{\max}} \int_{\mathbb{R}^{n-S}} \mathbf{1}_{\{i(\tau)=I\}} \partial_t \rho_t^{X,n,S}(\tau) d\tau.$$

Equation (27) yields that

$$\begin{aligned} \frac{d}{dt} n_{(S,I)}^X(t) &= \sum_{n=S+I}^{n_{\max}} \sum_{k=1}^{n-S} \int_{\mathbb{R}^{n-S}} \mathbf{1}_{\{i(\tau)=I\}} \partial_{\tau_k} \rho_t^{X,n,S}(\tau) d\tau \\ &\quad - \sum_{n=S+I}^{n_{\max}} \int_{\mathbb{R}^{n-S}} \mathbf{1}_{\{i(\tau)=I\}} (\lambda_X S I - \Lambda_t^X S) \rho_t^{X,n,S}(\tau) \\ &\quad + \sum_{n=S+I}^{n_{\max}} \mathbf{1}_{\{S+1 \leq n\}} \int_{\mathbb{R}^{n-S}} \mathbf{1}_{\{i(\tau)=I\}} \lambda_X (S+1) i(\tau_{1,n-S-1}) \rho_t^{X,n,S+1}(\tau_{1,n-S-1}) \mathbf{1}_{\{\tau_{n-S} > 0\}} \gamma e^{-\gamma \tau_{n-S}} d\tau \\ &\quad + \sum_{n=S+I}^{n_{\max}} \mathbf{1}_{\{S+1 \leq n\}} \int_{\mathbb{R}^{n-S}} \mathbf{1}_{\{i(\tau)=I\}} \Lambda_t^X (S+1) \rho_t^{X,n,S+1}(\tau_{1,n-S-1}) \mathbf{1}_{\{\tau_{n-S} > 0\}} \gamma e^{-\gamma \tau_{n-S}} d\tau. \end{aligned} \quad (38)$$

Let us start by noticing that the second line of Equation (38) may be written as

$$- (\lambda_X S I - \Lambda_t^X S) \sum_{n=S+I}^{n_{\max}} \int_{\mathbb{R}^{n-S}} \mathbf{1}_{\{i(\tau)=I\}} \rho_t^{X,n,S}(\tau) = - (\lambda_X S I - \Lambda_t^X S) n_{(S,I)}^X(t).$$

Next, let us pay attention to the third and fourth line of Equation (38). For any $\tau \in \mathbb{R}^{n-S}$, it appears that on the event $\tau_{n-S} > 0$, $i(\tau) = I$ if and only if $i(\tau_{1,n-S-1}) = I-1$. The third line thus becomes

$$\begin{aligned} &\sum_{n=S+I}^{n_{\max}} \mathbf{1}_{\{S+1 \leq n\}} \int_{\mathbb{R}^{n-S}} \mathbf{1}_{\{i(\tau_{1,n-S-1})=I-1\}} \lambda_X (S+1) (I-1) \rho_t^{X,n,S+1}(\tau_{1,n-S-1}) \mathbf{1}_{\{\tau_{n-S} > 0\}} \gamma e^{-\gamma \tau_{n-S}} d\tau \\ &= \lambda_X (S+1) (I-1) \sum_{n=S+I}^{n_{\max}} \mathbf{1}_{\{S+1 \leq n\}} \int_{\mathbb{R}^{n-S-1}} \mathbf{1}_{\{i(\tau')=I-1\}} \rho_t^{X,n,S+1}(\tau') d\tau' \\ &= \mathbf{1}_{\{I \geq 1\}} \lambda_X (S+1) (I-1) \sum_{n=S+I}^{n_{\max}} \int_{\mathbb{R}^{n-S-1}} \mathbf{1}_{\{i(\tau')=I-1\}} \rho_t^{X,n,S+1}(\tau') d\tau' \\ &= \mathbf{1}_{\{I \geq 1\}} \lambda_X (S+1) (I-1) n_{(S+1,I-1)}^X(t). \end{aligned}$$

In this computation, the indicator function of the event $\{I \geq 1\}$ appears as for any (n, s) , $\rho^{X,n,s}$ never charges the space $\{i(\tau) < 0\}$. As further $S + 1 \leq n$ whenever $S + I \leq n$ and $I \geq 1$, the indicator function of the set $\{S + 1 \leq n\}$ can be removed. Similarly, the fourth line of Equation (38) becomes

$$\mathbf{1}_{\{I \geq 1\}} \Lambda_t^X(S + 1) n_{(S+1, I-1)}^X(t).$$

It remains to focus on the first line of Equation (38). In order to do so, we will use another expression of $f^{(S,I)}$. For two integers $j \leq n$, let $\mathbb{B}(n, j)$ be the set of unordered subsets of j elements chosen in $\llbracket 1, n \rrbracket$. It then holds that for any $(n', s, \tau) \in E$,

$$f^{(S,I)}(n', s, \tau) = \sum_{n=S+I}^{n_{\max}} \mathbf{1}_{\left\{ \begin{smallmatrix} n'=n \\ s=S \end{smallmatrix} \right\}} \sum_{\mathbf{v} \in \mathbb{B}(n-S, I)} \prod_{j \in \mathbf{v}} h(\tau_j) \prod_{j \in \llbracket 1, n-S \rrbracket \setminus \mathbf{v}} (1 - h(\tau_j)),$$

where h designates as before the Heaviside step function. The idea behind this expression is that a structure of type (n, s, τ) contains S susceptible and I infected members if and only if $s = S$, and further exactly I out of the $n - S$ first components of τ (which describe the states of members of the structure having contracted the disease at some point, *i.e.* infected and removed members) are positive. In other words, there exists at most one element $\mathbf{v} \in \mathbb{B}(n - S, I)$ for which the term in the sum is not equal to zero, in which case the set \mathbf{v} corresponds to the indexes of infectious members, while $\llbracket 1, n - S \rrbracket \setminus \mathbf{v}$ is the set of removed members.

In the following, for any integers $1 \leq j \leq n$ and any $\mathbf{v} \in \mathbb{B}(n, j)$, consider the sets

$$A_{\mathbf{v}} = \{\tau \in \mathbb{R}^n : \tau_k > 0 \ \forall k \in \mathbf{v}, \tau_k \leq 0 \ \forall k \notin \mathbf{v}\}$$

and for any $k \in \mathbf{v}$,

$$A_{\mathbf{v};k} = \{(\tau_1, \dots, \tau_{k-1}, \tau_{k+1}, \dots, \tau_n) \in \mathbb{R}^{n-1} : \tau_j > 0 \text{ for } j \in \mathbf{v} \setminus \{k\} \text{ and } \tau_j \leq 0 \text{ for } j \notin \mathbf{v}\}.$$

With these notations, it holds that

$$\begin{aligned} J_t &:= \sum_{n=S+I}^{n_{\max}} \sum_{k=1}^{n-S} \int_{\mathbb{R}^{n-S}} \mathbf{1}_{\{i(\tau)=I\}} \partial_{\tau_k} \rho_t^{X,n,S}(\tau) d\tau = \sum_{n=S+I}^{n_{\max}} \sum_{k=1}^{n-S} \sum_{\mathbf{v} \in \mathbb{B}(n-S, I)} \int_{A_{\mathbf{v}}} \partial_{\tau_k} \rho_t^{X,n,S}(\tau) d\tau \\ &= \sum_{n=S+I}^{n_{\max}} \sum_{k=1}^{n-S} \sum_{\mathbf{v} \in \mathbb{B}(n-S, I)} (-1)^{\mathbf{1}_{\{k \in \mathbf{v}\}}} \int_{A_{\mathbf{v};k}} \rho_t^{X,n,S}(\tau_1, \dots, \tau_{k-1}, 0, \tau_{k+1}, \dots, \tau_n) \prod_{\substack{1 \leq j \leq n \\ j \neq k}} d\tau_j, \end{aligned}$$

where the second line follows from the first by integrating $\partial_{\tau_k} \rho_t^{X,n,S}$ over $A_{\mathbf{v}}$ with respect to τ_k . The goal is to show that

$$J_t = \gamma \left((I + 1) n_{(S, I+1)}^X(t) \mathbf{1}_{\{S+I < n_{\max}\}} - I n_{(S, I)}(t) \right), \quad (39)$$

which will conclude the proof of Equation (36).

In order to do so, let us define the following function for $x \in E$:

$$F^{(S,I)}(x) = \sum_{n=S+I}^{n_{\max}} \mathbf{1}_{\left\{ \begin{smallmatrix} \mathbf{n}(x)=n \\ \mathbf{s}(x)=S \end{smallmatrix} \right\}} \sum_{k=1}^{n-S} \sum_{\mathbf{v} \in \mathbb{B}(n-S, I)} (-1)^{\mathbf{1}_{\{k \in \mathbf{v}\}}} \mathbf{1}_{\{\tau_k(x)=0\}} H^{(n,S), \mathbf{v}, k}(\tau(x)),$$

where for any $\mathbf{v} \subseteq \llbracket 1, n - S \rrbracket$, $k \in \llbracket 1, n - S \rrbracket$ and $\tau \in \mathbb{R}^{n_{\max}}$,

$$H^{(n,S), \mathbf{v}, k}(\tau) = \prod_{\substack{j \in \mathbf{v} \\ j \neq k}} h(\tau_j) \prod_{\substack{j \in \llbracket 1, n-S \rrbracket \setminus \mathbf{v} \\ j \neq k}} (1 - h(\tau_j)).$$

It then follows that $J_t = \langle \eta_t^X, F^{(S,I)} \rangle$.

Furthermore, distinguishing the cases $k \in \mathbf{v}$ and $k \notin \mathbf{v}$ yields that

$$F^{(S,I)} = a^{(S,I)} - b^{(S,I)},$$

where for any $x \in E$,

$$a^{(S,I)}(x) = \sum_{n=S+I+1}^{n_{\max}} \mathbf{1}_{\left\{ \begin{smallmatrix} \mathbf{n}(x)=n \\ \mathbf{s}(x)=S \end{smallmatrix} \right\}} \sum_{k=1}^{n-S} \mathbf{1}_{\{\tau_k(x)=0\}} \sum_{\substack{\mathbf{v} \in \mathbb{B}(n-S,I) \\ k \notin \mathbf{v}}} H^{(n,S),\mathbf{v},k}(\tau(x)),$$

and

$$b^{(S,I)}(x) = \sum_{n=S+I}^{n_{\max}} \mathbf{1}_{\left\{ \begin{smallmatrix} \mathbf{n}(x)=n \\ \mathbf{s}(x)=S \end{smallmatrix} \right\}} \sum_{k=1}^{n-S} \mathbf{1}_{\{\tau_k(x)=0\}} \sum_{\substack{\mathbf{v} \in \mathbb{B}(n-S,I) \\ k \in \mathbf{v}}} H^{(n,S),\mathbf{v},k}(\tau(x)).$$

The heart of the proof then lies in the following result.

Lemma 3.16. *Under the assumptions of Theorem 2.3, for any $X \in \{H, W\}$ and $(S, I) \in \mathbb{S}$, for any $t \geq 0$,*

$$\langle \eta_t^X, a^{(S,I)} \rangle = \gamma(I+1)n_{(S,I+1)}^X(t) \mathbf{1}_{\{S+I < n_{\max}\}}, \quad (40)$$

and

$$\langle \eta_t^X, b^{(S,I)} \rangle = \gamma I n_{(S,I)}^X(t). \quad (41)$$

Before establishing Lemma 3.16, let us finish the proof of Proposition 3.15. It is clear that this lemma implies Equation (39), and thus Equation (36).

In order to obtain the initial condition of Equation (37), notice that by definition of $\eta_{0,\varepsilon}$,

$$\begin{aligned} \langle \eta_0^X, f^{(S,I)} \rangle &= \sum_{n=S+I}^{n_{\max}} \pi_n^X \binom{n}{S} (1-\varepsilon)^S \varepsilon^{n-S} \sum_{\mathbf{v} \in \mathbb{B}(n-S,I)} \int_{\mathbb{R}^{n-S}} \prod_{k \in \mathbf{v}} h(\tau_k) \prod_{k \in \llbracket 1, n-S \rrbracket \setminus \mathbf{v}} (1-h(\tau_k)) \nu(d\tau_1) \dots \nu(d\tau_{n-S}) \\ &= \sum_{n=S+I}^{n_{\max}} \pi_n^X \binom{n}{S} (1-\varepsilon)^S \varepsilon^{n-S} \sum_{\mathbf{v} \in \mathbb{B}(n-S,I)} \langle \nu, h \rangle^{\#\mathbf{v}} \langle \nu, 1-h \rangle^{(n-S)-\#\mathbf{v}}. \end{aligned}$$

Obviously, whenever $(n-S) - \#\mathbf{v} > 0$, the term vanishes as $\langle \nu, 1-h \rangle = 0$. Hence only the case $\mathbf{v} = \llbracket 1, n-S \rrbracket$ remains, which in turn corresponds to $n = S+I$. This concludes the proof. \square

Let us now turn to Lemma 3.16.

Proof of Lemma 3.16. Throughout the proof, we will continue using notations from the proof of Proposition 3.15. We will detail the arguments leading to Equation (40), as Equation (41) is obtained analogously.

Let $(S, I) \in \mathbb{S}$. We may start by noticing that in the case $S+I = n_{\max}$, it holds that $a^{(S,I)}(x) = 0$ for all $x \in E$. Indeed, the sum over n reduces to the case $n = n_{\max}$, and the set $\mathbb{B}(n_{\max}-S, I)$ contains exactly one element, namely $\llbracket 1, I \rrbracket$. Hence for any $k \in \llbracket 1, n_{\max}-S \rrbracket$, the set $\{\mathbf{v} \in \mathbb{B}(n_{\max}-S, I) : k \notin \mathbf{v}\}$ is empty, implying that $a^{(S,I)}$ is null on E . From now on, we will thus assume that $S+I < n_{\max}$.

Define the function $g^{(S,I)}$ on E as follows. For any $x \in E$,

$$g^{(S,I)}(x) = \sum_{n=S+I}^{n_{\max}} \mathbf{1}_{\left\{ \begin{smallmatrix} \mathbf{n}(x)=n \\ \mathbf{s}(x)=S \end{smallmatrix} \right\}} \sum_{k=1}^{n-S} h(\tau_k(x)) \sum_{\substack{\mathbf{v} \in \mathbb{B}(n-S,I-1) \\ k \notin \mathbf{v}}} H^{(n,S),\mathbf{v},k}(\tau(x)).$$

In particular, it holds by definition of $g^{(S,I)}$ that

$$\langle \eta_t^X, g^{(S,I)} \rangle = In_{(S,I)}^X(t).$$

Since we assume $S + I < n_{\max}$, the state $(S, I + 1)$ belongs to \mathbb{S} as well. Thus, we actually aim to show that for $G^{(S,I+1)} = \gamma g^{(S,I+1)} - a^{(S,I)}$,

$$\forall t \geq 0, \langle \eta_t^X, G^{(S,I+1)} \rangle = 0.$$

In order to achieve this, we will consider the following smooth approximations of $a^{(S,I)}$ and $g^{(S,I+1)}$ respectively, defined for $\alpha > 0$. For any $x \in E_X$,

$$\begin{aligned} a_\alpha^{(S,I)}(x) &= \sum_{n=S+I+1}^{n_{\max}} \mathbf{1}_{\left\{ \begin{smallmatrix} \mathbf{n}(x)=n, \\ \mathbf{s}(x)=S \end{smallmatrix} \right\}} \sum_{k=1}^{n-S} \psi_\alpha(\tau_k(x)) H_\alpha^{n-S,I,k}(\tau(x)), \\ g_\alpha^{(S,I+1)}(x) &= \sum_{n=S+I+1}^{n_{\max}} \mathbf{1}_{\left\{ \begin{smallmatrix} \mathbf{n}(x)=n, \\ \mathbf{s}(x)=S \end{smallmatrix} \right\}} \sum_{k=1}^{n-S} h_\alpha(\tau_k(x)) H_\alpha^{n-S,I,k}(\tau(x)), \end{aligned}$$

where

$$H_\alpha^{n-S,I,k}(\tau) = \sum_{\substack{\mathbf{v} \in \mathbb{B}(n-S,I) \\ k \notin \mathbf{v}}} \prod_{j \in \mathbf{v}} h_\alpha(\tau_j) \prod_{\substack{j \in \llbracket 1, n-S \rrbracket \\ j \neq k}} (1 - h_\alpha(\tau_j)).$$

Letting $G_\alpha^{(S,I+1)} = \gamma g_\alpha^{(S,I+1)} - a_\alpha^{(S,I)}$, dominated convergence ensures that

$$\langle \eta_t^X, G_\alpha^{(S,I+1)} \rangle \xrightarrow{\alpha \rightarrow 0} \langle \eta_t^X, G^{(S,I+1)} \rangle.$$

Classical properties of convolution imply that both $a_\alpha^{(S,I)}$ and $g_\alpha^{(S,I+1)}$ belong to $\mathcal{C}_b^1(E, \mathbb{R})$, allowing us to apply Proposition 3.13 to $G_\alpha^{(S,I+1)}(x)$. In order to do so, some preliminary computations are necessary. Indeed, we need to compute $f_{T,t}$ and $f_{T,t}^I$ as defined in Equation (28) for $f = a_\alpha^{(S,I)}$ and $f = g_\alpha^{(S,I+1)}$. Let $T > 0$, then for any $t \in [0, T]$, writing $e_{1,n-S}$ for the vector of dimension n_{\max} whose $(n-S)$ first components are set to 1 and the following all equal 0,

$$\left(a_\alpha^{(S,I)} \right)_{T,t}(x) = \sum_{n=S+I+1}^{n_{\max}} \mathbf{1}_{\left\{ \begin{smallmatrix} \mathbf{n}(x)=n, \\ \mathbf{s}(x)=S \end{smallmatrix} \right\}} \sum_{k=1}^{n-S} \psi_\alpha(\tau_k(x) - (T-t)) H_\alpha^{n-S,I,k}(\tau(x) - (T-t)e_{1,n-S}).$$

Similarly, for any $\sigma > 0$, distinguishing the cases where either $k = n - S$, or $k \neq n - S$ and $n - S \in \mathbf{v}$, or $k \neq n - S$ and $n - S \notin \mathbf{v}$ leads to

$$\begin{aligned} \left(a_\alpha^{(S,I)} \right)_{T,t}(j(x, \sigma)) &= \sum_{n=S+I+1}^{n_{\max}} \mathbf{1}_{\left\{ \begin{smallmatrix} \mathbf{n}(x)=n, \\ \mathbf{s}(x)-1=S \end{smallmatrix} \right\}} \left[\psi_\alpha(\sigma - (T-t)) H_\alpha^{n-S,I,n-S}(\tau(x) - (T-t)e_{1,n-S-1}) \right. \\ &\quad + h_\alpha(\sigma - (T-t)) \sum_{k=1}^{n-S-1} \psi_\alpha(\tau_k - (T-t)) H_\alpha^{n-S-1,I-1,k}(\tau(x) - (T-t)e_{1,n-S-1}) \\ &\quad \left. + (1 - h_\alpha(\sigma - (T-t))) \sum_{k=1}^{n-S-1} \psi_\alpha(\tau_k - (T-t)) H_\alpha^{n-S-1,I,k}(\tau(x) - (T-t)e_{1,n-S-1}) \right]. \end{aligned}$$

Define $m_{T,t}^h(\alpha) = \langle \nu, h_\alpha(\cdot - (T-t)) \rangle$ and $m_{T,t}^\psi(\alpha) = \langle \nu, \psi_\alpha(\cdot - (T-t)) \rangle$. Notice that the second line is equal to zero if $I = 0$, as the set $\mathbb{B}(n-S-I, -1)$ is empty. Similarly, in the case of the third line, when $n = S + 1 + I$ the set $\{\mathbf{v} \in \mathbb{B}(n-S-1, I)\}$ reduces to $\llbracket 1, I \rrbracket$ and hence for any

$k \in \llbracket 1, n-S \rrbracket$, the set $\{\mathbf{v} \in \mathbb{B}(n-S-1, I) : k \notin \mathbf{v}\}$ is empty. In particular, the third line vanishes if $S+I+1 = n_{\max}$. We may thus recognize that

$$\begin{aligned} \left(a_\alpha^{(S,I)}\right)_{T,t}^{\mathcal{I}}(x) &= m_{T,t}^\psi(\alpha) \sum_{n=S+I+1}^{n_{\max}} \mathbf{1}_{\left\{\begin{smallmatrix} \mathbf{n}(x)=n, \\ \mathbf{s}(x)-1=S \end{smallmatrix}\right\}} H_\alpha^{n-S,I,n-S}(\tau(x) - (T-t)e_{1,n-S-1}) \\ &\quad + m_{T,t}^h(\alpha) a_{\alpha;T,t}^{(S+1,I-1)} \mathbf{1}_{\{I \geq 1\}} + (1 - m_{T,t}^h(\alpha)) a_{\alpha;T,t}^{(S+1,I)} \mathbf{1}_{\{S+I+1 < n_{\max}\}}. \end{aligned}$$

Proceeding in the same way also yields

$$\begin{aligned} \left(g_\alpha^{(S,I+1)}\right)_{T,t}^{\mathcal{I}}(x) &= m_{T,t}^h(\alpha) \sum_{n=S+I+1}^{n_{\max}} \mathbf{1}_{\left\{\begin{smallmatrix} \mathbf{n}(x)=n, \\ \mathbf{s}(x)-1=S \end{smallmatrix}\right\}} H_\alpha^{n-S,I,n-S}(\tau(x) - (T-t)e_{1,n-S-1}) \\ &\quad + m_{T,t}^h(\alpha) g_{\alpha;T,t}^{(S+1,I)} \mathbf{1}_{\{I \geq 1\}} + (1 - m_{T,t}^h(\alpha)) g_{\alpha;T,t}^{(S+1,I+1)} \mathbf{1}_{\{S+I+1 < n_{\max}\}}. \end{aligned}$$

In particular, we thus obtain that

$$\begin{aligned} \left(G_\alpha^{(S,I+1)}\right)_{T,t}^{\mathcal{I}}(x) &= m_{T,t}^h(\alpha) G_{\alpha;T,t}^{(S+1,I)} \mathbf{1}_{\{I \geq 1\}} + (1 - m_{T,t}^h(\alpha)) G_{\alpha;T,t}^{(S+1,I+1)} \mathbf{1}_{\{S+I+1 < n_{\max}\}} \\ &\quad + (\gamma m_{T,t}^h(\alpha) - m_{T,t}^\psi(\alpha)) \sum_{n=S+I+1}^{n_{\max}} \mathbf{1}_{\left\{\begin{smallmatrix} \mathbf{n}(x)=n, \\ \mathbf{s}(x)-1=S \end{smallmatrix}\right\}} H_\alpha^{n-S,I,n-S}(\tau(x) - (T-t)e_{1,n-S-1}). \end{aligned}$$

Notice further that, on the one hand,

$$m_{T,t}^\psi(\alpha) = \int_0^{+\infty} \gamma e^{-\gamma x} \psi_\alpha(x - (T-t)) dx = \int_0^{+\infty} \gamma e^{-\gamma(x+(T-t))} \psi_\alpha(x) dx \xrightarrow{\alpha \rightarrow 0} \gamma e^{-\gamma(T-t)}.$$

On the other hand,

$$\begin{aligned} m_{T,t}^h(\alpha) &= \int_0^{+\infty} \gamma e^{-\gamma x} \int_0^{+\infty} \psi_\alpha(x - (T-t) - z) dz dx = \int_{\mathbb{R}} \int_y^{+\infty} \gamma e^{-\gamma x} dx \psi_\alpha(y - (T-t)) dy \\ &= \int_{\mathbb{R}} e^{-\gamma(y+(T-t))} \psi_\alpha(y) dy \xrightarrow{\alpha \rightarrow 0} e^{-\gamma(T-t)}. \end{aligned}$$

Thus, for any $T > 0$, we may apply Proposition 3.13 to $G_\alpha^{(S,I+1)}$, and let α go to zero in order to obtain by dominated convergence that

$$\begin{aligned} \langle \eta_T^X, G^{(S,I+1)} \rangle &= \langle \eta_0^X, G_{T,0}^{(S,I+1)} \rangle \\ &\quad + \lambda_X \int_0^T \langle \eta_t^X, \mathbf{si} \left(e^{-\gamma(T-t)} G_{T,t}^{(S+1,I)} \mathbf{1}_{\{I \geq 1\}} + (1 - e^{-\gamma(T-t)}) G_{T,t}^{(S+1,I+1)} \mathbf{1}_{\{S+I+1 < n_{\max}\}} - G_{T,t}^{(S,I+1)} \right) \rangle dt \\ &\quad + \int_0^T \Lambda_t^X \langle \eta_t^X, \mathbf{s} \left(e^{-\gamma(T-t)} G_{T,t}^{(S+1,I)} \mathbf{1}_{\{I \geq 1\}} + (1 - e^{-\gamma(T-t)}) G_{T,t}^{(S+1,I+1)} \mathbf{1}_{\{S+I+1 < n_{\max}\}} - G_{T,t}^{(S,I+1)} \right) \rangle dt. \end{aligned} \tag{42}$$

It is possible to compute $\langle \eta_0^X, G_{T,0}^{(S,I+1)} \rangle$. Indeed, by definition of $\eta_0 = \eta_{0,\varepsilon}$,

$$\begin{aligned} \langle \eta_0^X, a_{T,0}^{(S,I)} \rangle &= \sum_{n=S+I+1}^{n_{\max}} \pi_n^X \binom{n}{S} (1-\varepsilon)^S \varepsilon^{n-S} \\ &\quad \times \left[\sum_{k=1}^{n-S} \langle \nu, \delta_0(\cdot - T) \rangle \sum_{\substack{\mathbf{v} \in \mathbb{B}(n-S,I) \\ k \notin \mathbf{v}}} \prod_{j \in \mathbf{v}} \langle \nu, h(\cdot - T) \rangle \prod_{\substack{j \in \llbracket 1, n-S \rrbracket \\ j \neq k}} \langle \nu, 1 - h(\cdot - T) \rangle \right]. \end{aligned}$$

Of course, the computation of $\langle \eta_0^X, g_{T,0}^{(S,I+1)} \rangle$ is very similar. It suffices to replace the term $\langle \nu, \delta_0(\cdot - T) \rangle$ by $\langle \nu, h(\cdot - T) \rangle$. As $\langle \nu, \delta_0(\cdot - T) \rangle = \gamma \langle \nu, h(\cdot - T) \rangle$, it follows that

$$\langle \eta_0^X, G_{T,0}^{(S,I+1)} \rangle = 0. \tag{43}$$

Thus the term $\langle \eta_0^X, G_{T,0}^{(S,I+1)} \rangle$ of Equation (42) vanishes. Letting $C = n_{\max}(\lambda_X n_{\max} + \lambda_{\bar{X}} n_{\max} + \beta_G)$, we obtain that

$$|\langle \eta_T^X, G^{(S,I+1)} \rangle| \leq C \int_0^T \left(|\langle \eta_t^X, G_{T,t}^{(S+1,I)} \rangle| \mathbf{1}_{\{I \geq 1\}} + |\langle \eta_t^X, G_{T,t}^{(S+1,I+1)} \rangle| \mathbf{1}_{\{S+I+1 < n_{\max}\}} + |\langle \eta_t^X, G_{T,t}^{(S,I+1)} \rangle| \right) dt. \quad (44)$$

Notice that if $(S, I) \in \mathbb{S}$ is such that $S + I < n_{\max}$, then $(S', I') = (S, I + 1)$ is such that $(S', I') \in \mathbb{S}$, $S' + I' \geq 3$ and $I' \geq 1$. Similarly, on the event $I \geq 1$, $(S', I') = (S + 1, I)$ satisfies the same conditions, as well as $(S', I') = (S + 1, I + 1)$ if $S + I + 1 < n_{\max}$. In order to conclude, it thus suffices to establish the following lemma.

Lemma 3.17. *For any $(S, I) \in \mathbb{S}$ such that $I \geq 1$ and $S + I \geq 3$, for any $T \geq 0$,*

$$\forall t \in [0, T], \langle \eta_t^X, G_{T,t}^{(S,I)} \rangle = 0. \quad (45)$$

The desired result then follows from Equation (44). \square

Proof of Lemma 3.17. The key argument will be to apply Gronwall's lemma to $t \mapsto \langle \eta_t^X, G_{T,t}^{(S,I)} \rangle$. We thus need to show that for any $(S, I) \in \mathbb{S}$ such that $S + I \geq 3$ and $I \geq 1$, the function $t \mapsto \langle \eta_t^X, G_{T,t}^{(S,I)} \rangle$ is continuous on $[0, T]$. In order to do so, notice that we may apply Proposition 3.13 to $t \mapsto (G_\alpha^{(S,I)})_{T,t}$ which belongs to $\mathcal{C}_b^1(\mathbb{R}_+ \times E)$. As for any $z \in \mathbb{R}$ and $u \in [0, t]$ it holds that $z - (T - t) - (t - u) = z - (T - u)$, it follows that

$$\left((G_\alpha^{(S,I)})_{T,t} \right)_{t,u} = (G_\alpha^{(S,I)})_{T,u}.$$

Thus, proceeding like before and letting α go to zero, dominated convergence allows to conclude that for any $t \in [0, T]$,

$$\begin{aligned} \langle \eta_t^X, G_{T,t}^{(S,I)} \rangle &= \langle \eta_0^X, G_{T,0}^{(S,I)} \rangle \\ &+ \int_0^t \lambda_X \langle \eta_u^X, \mathbf{si} \left(e^{\gamma(T-u)} G_{T,u}^{(S+1,I-1)} \mathbf{1}_{\{I-1 \geq 1\}} + (1 - e^{\gamma(T-u)}) G_{T,u}^{(S+1,I)} \mathbf{1}_{\{S+I < n_{\max}\}} - G_{T,u}^{(S,I)} \right) \rangle du \\ &+ \int_0^t \Lambda_u^X \langle \eta_u^X, \mathbf{s} \left(e^{\gamma(T-u)} G_{T,u}^{(S+1,I-1)} \mathbf{1}_{\{I-1 \geq 1\}} + (1 - e^{\gamma(T-u)}) G_{T,u}^{(S+1,I)} \mathbf{1}_{\{S+I < n_{\max}\}} - G_{T,u}^{(S,I)} \right) \rangle du. \end{aligned}$$

The initial condition $t = 0$ vanishes according to Equation (43) as it holds for any $(S, I) \in \mathbb{S}$ satisfying $S + I < n_{\max}$, and indeed

$$\{(S, I + 1) : (S, I) \in \mathbb{S}, S + I < n_{\max}\} = \{(S, I) \in \mathbb{S} : S + I \geq 3, I \geq 1\}.$$

Further, both integrands are bounded by a constant which does not depend on t . The continuity of $t \mapsto \langle \eta_t^X, G_{T,t}^{(S,I)} \rangle$ on $[0, T]$ follows.

We are now ready to reason by induction. Notice that

$$\{(S, I) \in \mathbb{S} : S + I \geq 3, I \geq 1\} = \bigcup_{k=3}^{n_{\max}} \bigcup_{\ell=1}^{k-1} \{(k - \ell, \ell)\}.$$

For any $k \in [3, n_{\max}]$, for any $\ell \in [1, k - 1]$, Equation (44) becomes

$$|\langle \eta_T^X, G^{(k-\ell, \ell)} \rangle| \leq C \int_0^T \left(|\langle \eta_t^X, G_{T,t}^{(k-\ell+1, \ell-1)} \rangle| \mathbf{1}_{\{\ell-1 \geq 1\}} + |\langle \eta_t^X, G_{T,t}^{(k-\ell+1, \ell)} \rangle| \mathbf{1}_{\{k < n_{\max}\}} + |\langle \eta_t^X, G_{T,t}^{(k-\ell, \ell)} \rangle| \right) dt. \quad (46)$$

Let us start with the case $k = n_{\max}$. If further $\ell = 1$, Equation (46) reduces to

$$|\langle \eta_T^X, G^{(n_{\max}-1,1)} \rangle| \leq C \int_0^T |\langle \eta_t^X, G_{T,t}^{(n_{\max}-1,1)} \rangle| dt.$$

As $G^{(n_{\max}-1,1)} = G_{T,T}^{(n_{\max}-1,1)}$, the desired conclusion follows from Gronwall's lemma. Suppose now that Equation (45) holds for $(S, I) = (n_{\max} - \ell + 1, \ell - 1)$ with $\ell \in \llbracket 1, n_{\max} - 1 \rrbracket$. As further $k = n_{\max}$, we may notice that Equation (46) simplifies to

$$|\langle \eta_T^X, G^{(n_{\max}-\ell, \ell)} \rangle| \leq C \int_0^T |\langle \eta_t^X, G_{T,t}^{(n_{\max}-\ell, \ell)} \rangle| dt.$$

Again, Gronwall's lemma implies that Equation (45) holds for $(S, I) = (n_{\max} - \ell, \ell)$. We thus have established that Equation (45) is true for $(S, I) = (n_{\max} - \ell, \ell)$ for all $\ell \in \llbracket 0, n_{\max} - 1 \rrbracket$.

Let us now assume that Equation (45) is satisfied for $(S, I) = (k + 1 - \ell', \ell')$ for some $k \in \llbracket 3, n_{\max} - 1 \rrbracket$ and for all $\ell' \in \llbracket 1, k - 1 \rrbracket$. Then in particular, $\langle \eta_t^X, G_{T,t}^{(k,1)} \rangle = 0$ for all $t \in [0, T]$. Hence proceeding like before leads to the conclusion that $\langle \eta_t^X, G_{T,t}^{(k-1,1)} \rangle = 0$ for all $t \in [0, T]$ as the first two terms of the integrand of the associated inequality derived from Equation (46) are null. The same procedure allows to show that if Equation (45) holds for $(S, I) = (k - \ell + 1, \ell - 1)$ for some $\ell \in \llbracket 2, k - 1 \rrbracket$, then using the induction hypothesis, it also holds for $(S, I) = (k - \ell, \ell)$.

This concludes the proof. \square

We may finally focus on the main result of this section, namely Theorem 2.3.

Proof of Theorem 2.3. Before concluding, we need to emphasize that it would have been possible to choose $X = W$ when replacing S and I by $K_X S^X$ and $K_X I^X$ for \mathcal{I}_G in Proposition 3.5. All of the subsequent results still hold, simply replacing the household-related quantities in the definition of the rate for mean-field infections by their workplace-related counterparts.

As a consequence, Propositions 3.14 and 3.15 show that for any $X \in \{H, W\}$,

$$y^X = \left(\frac{s^X}{m_X}, \frac{i^X}{m_X}, n_{(S,I)}^X : (S, I) \in \mathbb{S}, n_{(S,I)}^{\bar{X}} : (S, I) \in \mathbb{S} \right)$$

satisfies the Cauchy problem (25) with initial condition (7). However, Proposition 3.10 ensures uniqueness of the solutions to this Cauchy problem. It hence is sensible to define, for $t \geq 0$,

$$s(t) = \frac{s^H(t)}{m_H} = \frac{s^W(t)}{m_W} \quad \text{and} \quad i(t) = \frac{i^H(t)}{m_H} = \frac{i^W(t)}{m_W}.$$

This leads to dynamical system (6a-c) with initial conditions (7), and concludes the proof. \square

Discussion

This paper has focused on proposing a new reduction for an *SIR* model with two levels of mixing, which explicitly includes households and workplaces. This reduced model was obtained as its large population limit, and the associated convergence of the stochastic model was established. In the general case of an arbitrary distribution ν of the length of infectious periods, the stochastic model converges under the large graph limit to the unique solution of a measure-valued equation. In the particular case where ν is an exponential distribution, we have shown that the epidemic dynamics can actually be reduced to a closed, finite-dimensional dynamical system. These results further are of interest since, to our knowledge, none of the previously proposed model

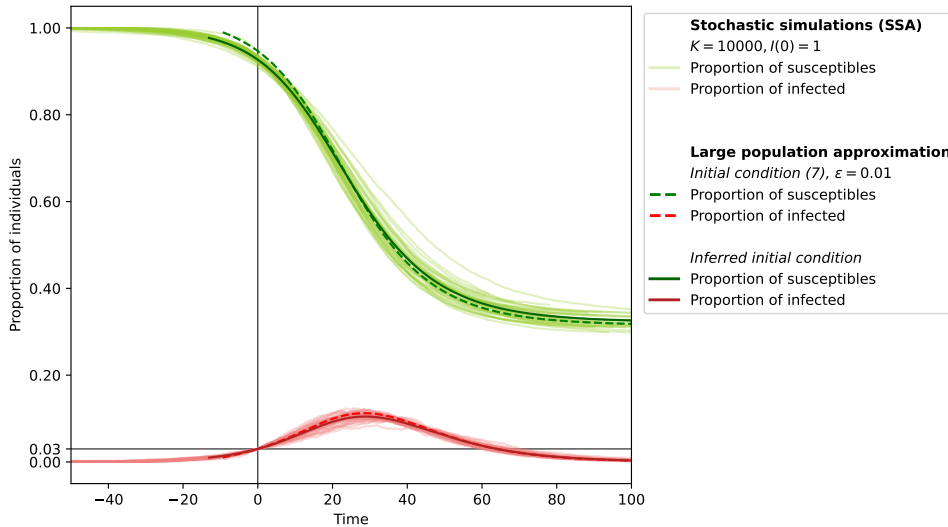


Figure 3: Comparison of the stochastic model starting from a single infected, with its large population approximation for two different choices of initial condition. The first initial condition is given by Equation (7) for $\varepsilon = 0.01$. The second initial condition is obtained by simulating a large number (> 2000) of stochastic epidemic trajectories, starting from a single infected until the proportion of infected reaches one percent. The initial condition corresponds to the average of the initial conditions observed in each simulation. Regarding the stochastic model, for this figure, 100 epidemics starting from a single infected were simulated. Similarly to Figure 2, only those reaching a threshold of 3% of infected are represented, and a time shift is applied to ease comparison between model outputs. Structure size distributions are those of Figure 1. Epidemic parameters: $(\beta_G, \lambda_H, \lambda_W, \gamma) = (0.085, 0.1, 0.001, 0.125)$, $R_0 = 1.7$.

reductions for epidemic models on networks whose nodes may belong to multiple cliques have been proven to coincide with the large graph limit of their stochastic counterparts.

Regarding the implementation of the reduced model, we have focused on the case where ν is the exponential distribution. Comparing the solution to the dynamical system with stochastic simulations of epidemic trajectories shows good accordance in practice, as expected. We also have emphasized that the dynamical system can be implemented in an automatic way, which implies that despite its high number of equations, it can be used even for larger structure sizes. Finally, some numerical evidence points to the reduced model being significantly less demanding than stochastic simulations in terms of computational cost for sufficiently large values of the reproduction number, which means that it is of interest for numerical explorations. A broader study of computational cost would be pertinent to refine these results.

Notice here that a possible model extension would be to consider a local level of mixing containing an arbitrary, yet finite, number of layers. As long as within each layer, each node is part of exactly one clique, and as long as cliques within each layer are constituted independently from one another as in the case for households and workplaces, the adaptation of the aforementioned results is expected to be straightforward. In particular, the dimension of the corresponding dynamical system should still be of order $O(n_{\max}^2)$, implying that the model should remain tractable.

Furthermore, we have compared the reduced model obtained in this work with the corresponding EBCM in the line of [41]. In the case of our household-workplace model with two levels of mixing, the EBCM seems the less appropriate choice, as it is less parsimonious and only approaches the epidemic well if the initial proportion of infected is very small. However, this may change if a more general contact structure within layers is considered, such as a configuration model for the global level, in which case it seems sensible to assume that EBCM-like equations will appear.

Finally, let us emphasize that by essence, the large population limit obtained here corresponds to a situation where the number of infected individuals is of the same order as the population size. In a realistic scenario, however, an epidemic is initiated by very few infected individuals. In the case of a large epidemic outbreak, the number of infected subsequently grows until it no longer is negligible when compared to the population size, at which point the large population limit correctly captures the dynamics of the outbreak. This raises the question: which initial condition is pertinent for the large population approximation? For uniformly mixing population, this is rather straightforward, for two main reasons. On the one hand, at each time, infected individuals are interchangeable in terms of infectious pressure exerted on susceptibles. On the other hand, the presence of recovered individuals at time t can be neglected in the study of the epidemic dynamics over the time interval $[t, \infty)$, simply by restricting the study to all other individuals, which still constitute a uniformly mixing population. As a consequence, in such a setting, it makes sense to suppose that at time zero, there are only infected and susceptible individuals, and that infected individuals are chosen uniformly at random in the population, with independent and identically distributed infectious period lengths.

This idea can of course be extended to our setting, and corresponds to the initial condition proposed in Theorem 2.3, while similar initial conditions have also been used in the literature in related settings [41, 11]. In our model however, one actually needs to know how infected and recovered individuals are distributed among households and workplaces, meaning that neither can recovered be ignored, nor is there any reason to believe that infected individuals are distributed uniformly at random in the population. Indeed, Figure 3 illustrates that when compared to stochastic simulations starting from a single infected, the large population approximation with initial condition given by Equation (7) fails to reproduce the epidemic dynamics, while they are correctly captured when using an initial condition which is inferred from stochastic simulations. As a consequence, it seems of interest to get a better understanding of this realistic initial condition, which arises from an epidemic started by a single infected. This may be achieved using a branching process approximation of the epidemic, which is designed to approach the initial, stochastic phase of the epidemic, and hence would represent a reduced model which complements the large population limit obtained in the present work.

Acknowledgment. I would like to thank Vincent Bansaye for his continuous support during the elaboration of this work and detailed feedback on the manuscript, as well as Viet Chi Tran and Frank Ball for pertinent discussions on related topics. Finally, I am very grateful to Elisabeta Vergu for her guidance and stimulating remarks. She will be missed.

References

- [1] David Aldous. Stopping Times and Tightness. *The Annals of Probability*, 6(2):335–340, April 1978.
- [2] Håkan Andersson and Tom Britton. Density dependent jump Markov processes. In *Stochastic Epidemic Models and Their Statistical Analysis*, Lecture Notes in Statistics. Springer, New York, NY, 2000.
- [3] Frank G. Ball, Lorenzo Pellis, and Pieter Trapman. Reproduction numbers for epidemic models with households and other social structures II: Comparisons and implications for vaccination. *Mathematical Biosciences*, 274:108–139, April 2016.
- [4] Frank G. Ball, David J. Sirl, and Pieter Trapman. Epidemics on random intersection graphs. *The Annals of Applied Probability*, 24(3):1081–1128, June 2014.

- [5] Vincent Bansaye, François Deslandes, Madeleine Kubasch, and Elisabeta Vergu. The epidemiological footprint of contact structures in models with two levels of mixing. Available on arXiv: <https://arxiv.org/abs/2303.05287>, 2023.
- [6] Rosanna C. Barnard, Istvan Z. Kiss, Luc Berthouze, and Joel C. Miller. Edge-Based Compartmental Modelling of an SIR Epidemic on a Dual-Layer Static–Dynamic Multiplex Network with Tunable Clustering. *Bulletin of Mathematical Biology*, 80(10):2698–2733, October 2018.
- [7] Patrick Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Statistics. Probability and Statistics Section. Wiley, New York, 2nd ed edition, 1999.
- [8] Tom Britton and Etienne Pardoux. Chapter 3 A General Two-Level Mixing Model. In Tom Britton and Etienne Pardoux, editors, *Stochastic Epidemic Models with Inference*, Lecture Notes in Mathematics, pages 159–213. Springer International Publishing, Cham, 2019.
- [9] Laurent Decreusefond, Jean-Stéphane Dhersin, Pascal Moyal, and Viet Chi Tran. Large graph limit for an SIR process in random network with heterogeneous connectivity. *The Annals of Applied Probability*, 22(2), April 2012.
- [10] María del Valle Rafo, Juan Pablo Di Mauro, and Juan Pablo Aparicio. Disease dynamics and mean field models for clustered networks. *Journal of Theoretical Biology*, 526:110554, October 2021.
- [11] Francesco Di Lauro, Luc Berthouze, Matthew D. Dorey, Joel C. Miller, and István Z. Kiss. The Impact of Contact Structure and Mixing on Control Measures and Disease-Induced Herd Immunity in Epidemic Models: A Mean-Field Model Perspective. *Bulletin of Mathematical Biology*, 83(11):117, November 2021.
- [12] Nicolas Fournier and Sylvie Méléard. A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *The Annals of Applied Probability*, 14(4):1880–1919, November 2004.
- [13] Laurent Hébert-Dufresne, Pierre-André Noël, Vincent Marceau, Antoine Allard, and Louis J. Dubé. Propagation dynamics on networks featuring complex topologies. *Physical Review E*, 82(3):036115, September 2010.
- [14] Thomas House and Matt J. Keeling. Deterministic epidemic models with explicit household structure. *Mathematical Biosciences*, 213(1):29–39, May 2008.
- [15] Thomas House and Matt J. Keeling. Insights from unifying modern approximations to infections on networks. *Journal of the Royal Society Interface*, 8(54):67–73, January 2011.
- [16] Karly A. Jacobsen, Mark G. Burch, Joseph H. Tien, and Grzegorz A. Rempała. The large graph limit of a stochastic epidemic model on a dynamic multilayer network. *Journal of Biological Dynamics*, 12(1):746–788, January 2018.
- [17] Svante Janson, Malwina Luczak, and Peter Windridge. Law of large numbers for the SIR epidemic on a random graph with given degrees. *Random Structures & Algorithms*, 45(4):726–763, 2014.
- [18] Anatole Joffe and Michel Metivier. Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Advances in Applied Probability*, 18(1):20–65, March 1986.

- [19] Benjamin Jourdain, Sylvie Méléard, and Wojbor A. Woyczynski. Lévy flights in evolutionary ecology. *Journal of Mathematical Biology*, 65(4):677–707, October 2012.
- [20] Brian Karrer and M. E. J. Newman. Message passing approach for general epidemic models. *Physical Review E*, 82(1):016101, July 2010.
- [21] Matt J. Keeling and Ken T. D. Eames. Networks and epidemic models. *Journal of the Royal Society, Interface*, 2(4):295–307, September 2005.
- [22] Istvan Z. Kiss, Eben Kenah, and Grzegorz A. Rempala. Necessary and sufficient conditions for exact closures of epidemic equations on configuration model networks, August 2022.
- [23] Thomas G. Kurtz. Epidemic models. In *Approximation of Population Processes*, number 36 in CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, Philadelphia, Pa, 1981.
- [24] Jennifer Lindquist, Junling Ma, P. van den Driessche, and Frederick H. Willeboordse. Effective degree network disease models. *Journal of Mathematical Biology*, 62(2):143–164, February 2011.
- [25] Giancarlo De Luca, Kim Van Kerckhove, Pietro Coletti, Chiara Poletto, Nathalie Bossuyt, Niel Hens, and Vittoria Colizza. The impact of regular school closure on seasonal influenza epidemics: A data-driven spatial transmission model for Belgium. *BMC Infectious Diseases*, 18(1):29, January 2018.
- [26] Junling Ma, Pauline van den Driessche, and Frederick H. Willeboordse. Effective degree household network disease model. *Journal of Mathematical Biology*, 66(1):75–94, January 2013.
- [27] Sylvie Meleard and Vincent Bansaye. *Stochastic Models for Structured Populations: Scaling Limits and Long Time Behavior*. Springer International Publishing, Cham, 2015.
- [28] Alba Mendez-Brito, Charbel El Bcheraoui, and Francisco Pozo-Martin. Systematic review of empirical studies comparing the effectiveness of non-pharmaceutical interventions against COVID-19. *Journal of Infection*, 83(3):281–293, September 2021.
- [29] Joel C. Miller. Percolation and epidemics in random clustered networks. *Physical Review E*, 80(2):020901, August 2009.
- [30] Joel C Miller. A note on a paper by Erik Volz: SIR dynamics in random networks. *Journal of Mathematical Biology*, 62:349–358, 2011.
- [31] Mark E. J. Newman. Random graphs with clustering. *Physical Review Letters*, 103(5):058701, July 2009.
- [32] Lorenzo Pellis, Neil M. Ferguson, and Christophe Fraser. Threshold parameters for a model of epidemic spread among households and workplaces. *Journal of The Royal Society Interface*, 6(40):979–987, November 2009.
- [33] Lorenzo Pellis, Neil M. Ferguson, and Christophe Fraser. Epidemic growth rate and household reproduction number in communities of households, schools and workplaces. *Journal of Mathematical Biology*, 63(4):691–734, October 2011.
- [34] Sylvie Roelly-Coppoletta. A criterion of convergence of measure-valued processes: Application to measure branching processes. *Stochastics: An International Journal of Probability and Stochastic Processes*, 17(1-2):43–65, 1986.

- [35] Neil Sherborne, Joel C. Miller, Konstantin B. Blyuss, and Istvan Z. Kiss. Mean-field models for non-Markovian epidemics on networks. *Journal of Mathematical Biology*, 76(3):755–778, February 2018.
- [36] Mario Ignacio Simoy and Juan Pablo Aparicio. Socially structured model for COVID-19 pandemic: Design and evaluation of control measures. *Computational and Applied Mathematics*, 41(1):14, December 2021.
- [37] Guillaume St-Onge, Laurent Hébert-Dufresne, and Antoine Allard. Heterogeneous transmission in groups induces a superlinear force of infection, February 2023.
- [38] Viet Chi Tran. *Modèles particuliers stochastiques pour des problèmes d'évolution adaptative et pour l'approximation de solutions statistiques*. PhD thesis, Université de Nanterre - Paris X, December 2006.
- [39] Viet Chi Tran. *Une ballade en forêts aléatoires*. Habilitation a Diriger les Recherches, Université Lille 1, November 2014.
- [40] Erik M. Volz. SIR dynamics in random networks with heterogeneous connectivity. *Journal of Mathematical Biology*, 56(3):293–310, March 2008.
- [41] Erik M. Volz, Joel C. Miller, Alison Galvani, and Lauren Ancel Meyers. Effects of Heterogeneous and Clustered Contact Patterns on Infectious Disease Dynamics. *PLoS Computational Biology*, 7(6):e1002042, June 2011.
- [42] Robert R. Wilkinson, Frank G. Ball, and Kieran J. Sharkey. The relationships between message passing, pairwise, Kermack–McKendrick and stochastic SIR epidemic models. *Journal of Mathematical Biology*, 75(6-7), April 2017.

Appendix A Implementation of the large population limit

A.1 Automatic implementation of the dynamical system

It is possible to implement dynamical system (6a-c) in an automated way, in the sense that equations do not need to be written individually. The key lies in the fact that the set \mathbb{S} can be constructed automatically, with an intrinsic organization of the states (S, I) it contains. For example, one may arrange them by growing number n of susceptible and infected members of the structure, and for each n , by growing number i of infected, leading to

$$\mathbb{S} = \{(2, 0), (1, 1), \dots, (n_{\max}, 0), (n_{\max} - 1, 1), \dots, (1, n_{\max} - 1)\}.$$

This in turn allows to make an explicit correspondence between any state $(S, I) \in \mathbb{S}$ and *e.g.* some position in a vector containing all functions of our dynamical system of interest. A similar idea was already employed in [33], for another purpose. With the previous structure of \mathbb{S} , one may for instance notice that for any $n \in \llbracket 2, n_{\max} \rrbracket$ and $i \in \llbracket 0, n - 1 \rrbracket$, the state $(n - i, i)$ is the $c(n - i, i)$ -th state enumerated in \mathbb{S} , where $c(n - i, i) = (n - 1)n/2 + i$. As a consequence, the general expression of Equation (6c) may be used to handle all the dynamics of the functions $n_{(S,I)}^X$, for $(S, I) \in \mathbb{S}$ and $X \in \{H, W\}$.

Also, notice that in practice, household sizes tend not to be as big as workplace sizes. It thus makes sense to distinguish explicitly a maximal size for each type of structure. This allows to avoid implementing unnecessary equations corresponding *e.g.* to household sizes that are not actually observed, and which thus artificially increase the dimension of the system.

A.2 Computational performance

The aim of this section is to numerically assess the computational cost associated to solving the large dimensional dynamic system (6a-c) in comparison to stochastic simulations using Gillespie's algorithm, also referred to as SSA (stochastic simulation algorithm). In order to do so, the average execution times of one stochastic simulation (SSA) and of one resolution of the associated dynamic system using the ODE solver `odeint` from the `scipy.integrate` library are compared.

Let us start by describing the general procedure. Each of the two scripts (stochastic simulation or reduced model) is executed one hundred times, all runs being independent from one another. For each run and each script, the computation time of the script of interest is measured, as well as the computation time of a reference function (summing all integers up to one billion with a simple for-loop). The ratios of the runtimes of both the script of interest and the reference function are computed. Comparison of the computation times for the stochastic and the reduced model is then based on the comparison of the averages of those normalised runtimes.

It remains to take an interest in the choice of the model parameters, namely the structure size distributions, the epidemic parameters *i.e.* the contact rates β_G , λ_H , λ_W and the removal rate γ , as well as the initial proportion of infected ε and the time interval $[0, T]$ on which the epidemic is simulated. For the stochastic model, the population size K will be fixed to ten thousand individuals. For all scenarios considered here, the structure size distributions will be those of Figure 1. Further, the initial proportion of individuals will be set to $\varepsilon = 0.005$, as Figure 2 indicates that for the stochastic model, the risk of extinction at the beginning of the epidemic due to stochastic fluctuations is small when starting from 50 infected at time 0, even for low values of R_0 . However, different values of the epidemic parameters will be considered, as to obtain scenarios that differ both in terms of R_0 and in terms of the proportions of infections occurring within the general population, within households or within workplaces, respectively referred to as p_G , p_H and p_W . The removal rate γ will be fixed at 0.125, and only the contact

rates will effectively vary. In total, ten different scenarios will be used, characterized by their values of $R_0 \in \{1.2, 1.4, 1.7, 2.0, 2.5\}$ and $(p_G, p_H, p_W) \in \{(0.2, 0.4, 0.4), (0.4, 0.4, 0.2)\}$.

Finally, parameter T will be chosen as follows. For each set of epidemic parameters detailed above, the reduced model is used to compute the time T_* at which the epidemic falls below one percent of infected individuals in the population, after the epidemic peak. T then is determined by rounding down T_* to the closest multiple of five.

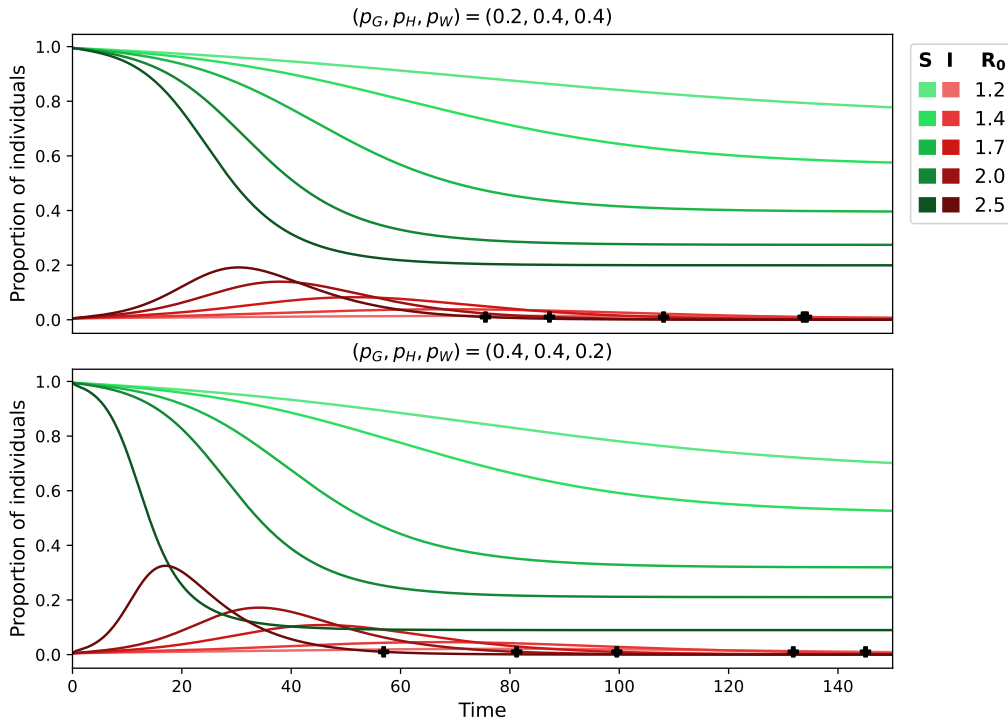


Figure A1: Proportion of susceptible (S) and infected (I) in the population, for each scenario detailed in Table A1, as given by dynamical system (6a-c). Scenarios are separated by values of (p_G, p_H, p_W) of infections per layer, namely $(0.2, 0.4, 0.4)$ and $(0.4, 0.4, 0.2)$ for the top and bottom panels, respectively. The corresponding values of R_0 are indicated by the color shades, as shown in the legend. The black crosses indicate for each curve that the proportion of infected falls below the threshold of one percent.

Figure A1 uses the reduced model to plot the trajectories of the proportion of susceptible and infected individuals in the population, for each scenario. The corresponding parameters are summarized in Table A1. Notice that in particular, this includes the parameters of Figure 2.

Let us now turn to the results. For each scenario of Table A1, measurement of average normalised computation times was repeated three times. The results are shown in Figure A2, which indicates for each scenario the ratio of the average normalised runtime for one resolution of dynamical system (6a-c) over the average normalised runtime of one stochastic simulation. Let us first take an interest in the datasets labeled $(p_G, p_H, p_W) = (0.2, 0.4, 0.4)$ and $(p_G, p_H, p_W) = (0.4, 0.4, 0.2)$. One may notice first that for each scenario, the results of all three repeats are close to one another, indicating that the results are reproducible. Further, for both possible values of (p_G, p_H, p_W) , the results indicate a shared general trend. Indeed, for values of R_0 close to the critical case $R_0 = 1$, the ratio exceeds one, and diminishes subsequently, falling below one between $R_0 = 1.4$ and $R_0 = 1.7$ and attaining values of order 10^{-1} . This behavior suggests that solving dynamic system (6a-c) is advantageous in terms of computation time for intermediate or high values of R_0 , being up to one order of magnitude faster than one stochastic simulation. As the time interval $[0, T]$ on which the epidemic is studied originally depends on the scenario and is significantly shorter for larger values of R_0 , one may wonder whether this

	$R_0 = 1.2$	$R_0 = 1.4$	$R_0 = 1.7$	$R_0 = 2.0$	$R_0 = 2.5$	(p_G, p_H, p_W)
β_G	0.03	0.035	0.045	0.05	0.06	
λ_H	0.05	0.07	0.09	0.15	0.2	(0.2, 0.4, 0.4)
λ_W	0.0015	0.0016	0.0018	0.002	0.0022	
T	130	130	105	85	75	
β_G	0.06	0.07	0.085	0.1	0.125	(0.4, 0.4, 0.2)
λ_H	0.06	0.07	0.1	0.15	1.5	
λ_W	0.00075	0.0008	0.001	0.0011	0.00115	
T	145	130	95	80	55	

Table A1: Values of the contact rates β_G , λ_H and λ_W considered, grouped by value of R_0 and proportions of infections per layer (p_G, p_H, p_W) characterizing the scenarios.

difference influences the results. As a consequence, we have repeated the same procedure for all of the scenarios characterised by $(p_G, p_H, p_W) = (0.2, 0.4, 0.4)$, with fixed $T = 75$. Figure A2 shows that the associated results are very similar to those obtained previously, pleading against this hypothesis.

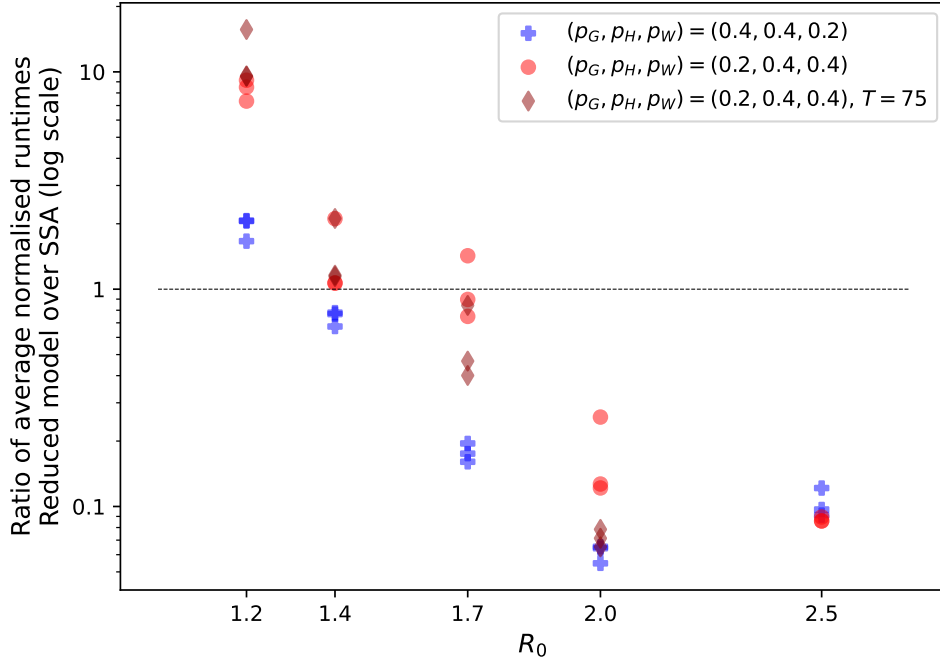


Figure A2: Ratio of the average normalised computation time for solving once dynamic system (6a-c) over the average normalised computation time for one stochastic simulation (SSA). This ratio was computed three times for each scenario of Table A1. The results are presented as a function of R_0 , while colors indicate the value of (p_G, p_H, p_W) . Unless stated otherwise, the parameter T from Table A1 was used. The dotted line indicates the threshold of one.

Of course, this comparison could be pushed further. For instance, the most basic version of the SSA algorithm was used, and more advanced methods such as τ -leaping are expected to accelerate stochastic simulations. Also, a more thorough exploration of the parameter space would be pertinent, assessing for instance the influence of the structure size distributions.

Supplementary Material

S1 Edge-based compartmental model

S1.1 Presentation of the EBCM

Let us start by describing how to obtain the population structure of the local level of mixture described in Section 1.1 using a clique configuration model (CCM). In our case, each node belongs to exactly one clique within each layer (one household and one workplace, respectively). Let us briefly notice that whenever a node is picked uniformly at random, the probability of it belonging to a structure of type X and size n is given by $\hat{\pi}_n^X = n\pi_n^X/n^X$, for any $n \in \llbracket 1, n_{\max} \rrbracket$ and $X \in \{H, W\}$. As a consequence, the layer corresponding to structures of type $X \in \{H, W\}$ is obtained by the following two steps. First, associate to each node a structure size distributed according to the size-biased law $\hat{\pi}^X$. This is done independently for each node. Second, for $k \in \llbracket 1, n_{\max} \rrbracket$, form cliques of size k by drawing uniformly without replacement k -tuples in the set of nodes of associated structure size k . This step stops when all nodes of associated clique size k belong to a clique. This procedure is repeated independently for each layer, allowing to assemble households and workplaces.

Let us now turn to deriving the EBCM. Consider s and i the proportions of susceptible and infected individuals in the population, respectively. Let $\theta_n^X(t)$ for $X \in \{H, W\}$ and $n \in \llbracket 1, n_{\max} \rrbracket$ be the chance of a susceptible belonging to a structure of type X and size n to escape infection within this structure, and $\theta^G(t)$ the chance of escaping infection through the mean-field level, up to time $t \geq 0$. The key idea is that a node is susceptible at time t if and only if it has escaped infection up to time t , and the risks of infection within each layer are independent from one another. This makes use of properties of the CCM, which heuristically explain the decoupling of the risk of infection in the two local layers from one another. Further, the fact that in an infinite population, each individual structure has a negligible impact on the proportion of infected yields the intuition behind the decorrelation of the risks of infection at the local and global level. This leads to

$$s = \theta^G \prod_{X \in \{H, W\}} \left(\sum_{n=1}^{n_{\max}} \hat{\pi}_n^X \theta_n^X \right).$$

As we are considering an *SIR* model, it follows that $i'(t) = -s'(t) + \gamma i(t)$, so that the difficulty resides in understanding the dynamics of θ^G and $\theta_n^X(t)$, for $X \in \{H, W\}$ and $n \in \llbracket 1, n_{\max} \rrbracket$.

Define for $X \in \{H, W\}$ and $n \in \llbracket 2, n_{\max} \rrbracket$:

$$m_n^X = \theta^G \hat{\pi}_n^X \theta_n^X \sum_{k=1}^{n_{\max}} \hat{\pi}_k^{\bar{X}} \theta_k^{\bar{X}},$$

which corresponds to the proportion of individuals who are susceptible and belong to a structure of type X and size n . Also, let $n_{(S,I,R)}^X$ be the proportion of susceptibles belonging to a structure of type X containing exactly S susceptibles, I infected and R removed individuals. This allows us to introduce the following quantities, which participate in the rates at which a member of a structure of type X and size n is infected, either within the considered structure or outside of it, respectively:

$$T_n^X = \lambda_X \sum_{\substack{(S,I,R) \in \mathbb{N}^3 \\ S+I+R=n}} SI n_{(S,I,R)}^X \quad \text{and} \quad \tau_n^X = \left(\beta_G i + \sum_{k=1}^{n_{\max}} T_k^{\bar{X}} \right) \frac{\hat{\pi}_n^X \theta_n^X}{\sum_{k=1}^{n_{\max}} \hat{\pi}_k^X \theta_k^X}.$$

One obtains the following dynamics:

$$\frac{d}{dt} \theta^G = -\beta_G i \theta^G, \quad \text{and} \quad \forall X \in \{H, W\}, \forall n \in \llbracket 2, n_{\max} \rrbracket, \quad \frac{d}{dt} \theta_n^X = -\frac{T_n^X}{m_n^X} \theta_n^X.$$

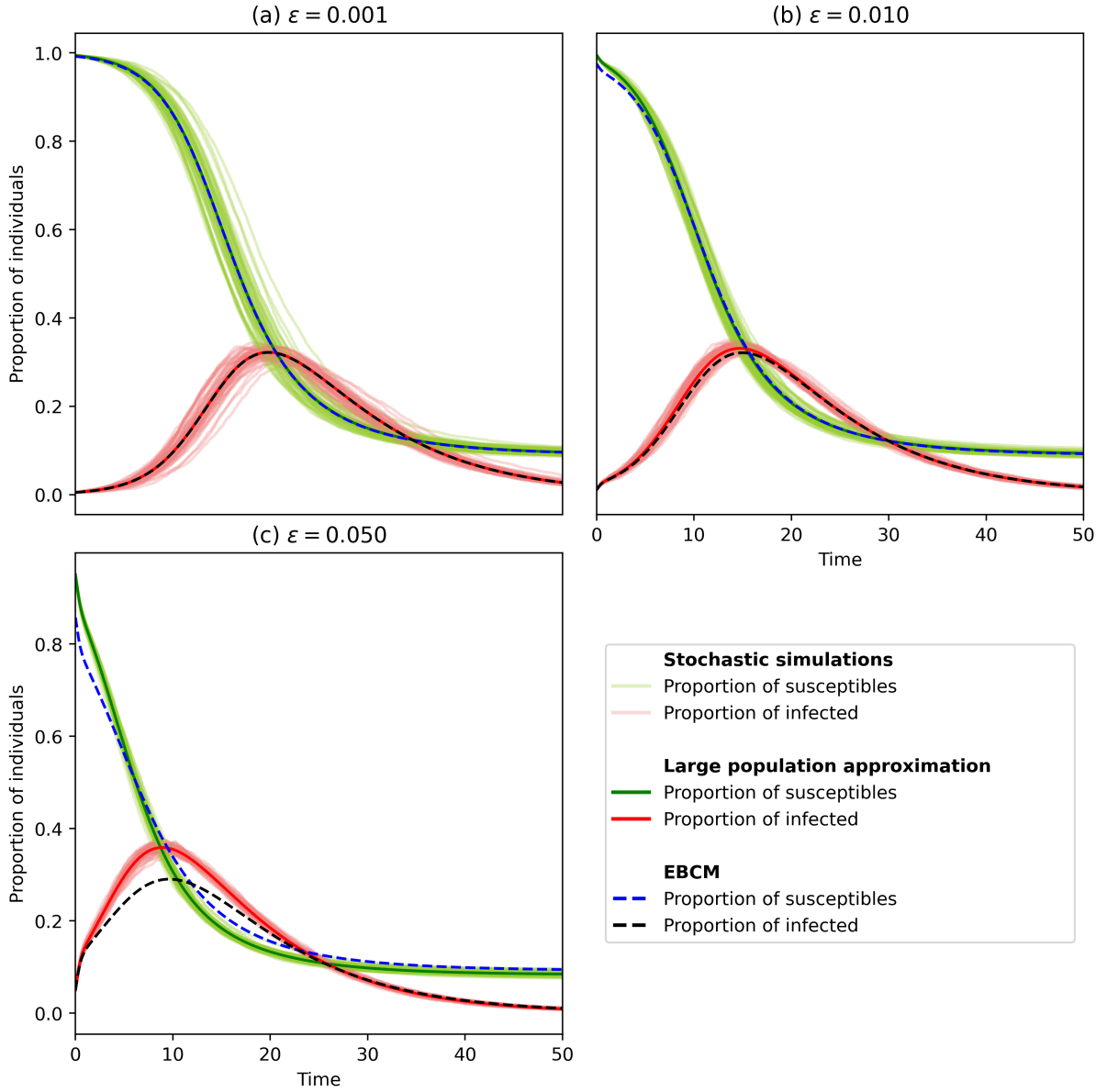


Figure S1: Comparison of the stochastic model with the large population approximation given by dynamical system (6a–c) and the corresponding EBCM. Household and workplace distributions are those of Figure 1. Epidemic parameters were set to $(\beta_G, \lambda_H, \lambda_W, \gamma) = (0.125, 1.5, 0.00115, 0.125)$. Initial conditions correspond to $\varepsilon \in \{0.001, 0.01, 0.05\}$ as indicated for each panel. For each of these scenarios, Gillespie’s algorithm was used to simulate 50 trajectories of the stochastic model defined in Proposition 1.1 in a population of $K = 10000$ individuals (faint lines). For Panel (a), only trajectories reaching a threshold of 0.005 infected were kept, and time was shifted so that time 0 corresponds to the moment when this threshold is reached. Finally, the deterministic solution (s, i) of both dynamical system (6a–c) (thick lines) and the EBCM (dashed lines) are represented for each scenario. For Panel (a), the same time shifting procedure as for simulations is applied.

Further, for any $X \in \{H, W\}$, $n \in \llbracket 2, n_{\max} \rrbracket$ and $(S, I, R) \in \mathbb{N}^3$ such that $S + I + R = n$ and either $S \geq 2$ or $SI \geq 1$:

$$\begin{aligned} \frac{d}{dt} n_{(S,I,R)}^X &= - \left(\lambda_X SI + \frac{\tau_n^X}{m_n^X} S + \gamma I \right) n_{(S,I,R)}^X \\ &\quad + \gamma (I + 1) n_{(S,I+1,R-1)}^X \mathbf{1}_{\{R \geq 1\}} \\ &\quad + \left(\lambda_X (S + 1)(I - 1) + \frac{\tau_n^X}{m_n^X} (S + 1) \right) n_{(S+1,I-1,R)}^X \mathbf{1}_{\{I > 1\}}. \end{aligned}$$

Additionally, as in a structure of size one, no infection may occur within the structure itself, θ_1^H and θ_1^W are constant over time. Finally, it remains to define the initial conditions. Following Volz *et al.* [41], we consider the case $\varepsilon \ll 1$. Then the only quantities which are not null at time zero are: for any $X \in \{H, W\}$, $n \in \llbracket 1, n_{\max} \rrbracket$ and $I \in \llbracket 1, n - 1 \rrbracket$,

$$\begin{aligned} i(0) &= \varepsilon, \\ \theta^G(0) &= \theta_n^X(0) = 1 - \varepsilon, \\ n_{(n,0,0)}^X &= \frac{1}{n} \hat{\pi}_n^X (1 - \varepsilon)^n, \\ n_{(n-I,I,0)}^X &= \hat{\pi}_n^X (1 - \varepsilon)^{n-I} \varepsilon^I. \end{aligned} \tag{S1}$$

The proportions of susceptible and infected as predicted by both the EBCM and dynamical system (6a-c) are shown in Figure S1, for different values of ε . Let us first notice that in the case $\varepsilon = 0.001$, corresponding to Panel (a) of Figure S1, the solutions (s, i) of both the EBCM and dynamical system (6a-c) are in perfect accordance, emphasizing the fact that for very small values of ε , the EBCM seems to yield the correct asymptotic population dynamics. However, for larger values of ε , the EBCM struggles to reproduce these dynamics. At first, as for $\varepsilon = 0.01$ in Panel (b) of Figure S1, the difference is mainly visible at the beginning of the epidemic and seems to dampen afterwards. For even larger values of ε , the solution to the EBCM does not succeed in reproducing the dynamics over time, as shown in Panel (c) of Figure S1. The problem for capturing the epidemic dynamics for higher values of ε lies in the fact that defining the proper initial condition for the EBCM is not straightforward, leading to initial conditions consisting in an approximation which is only sensible whenever ε is very small.

S1.2 Computational performance

In order to compare the computation times needed to solve either dynamical system (6a-c) or the dynamical system associated to the EBCM which has been introduced above, we will proceed similarly as in Appendix A.2, making use of the ODE solver `odeint` from the `scipy.integrate` library in both cases. However, this time, only one parameter set will be used, corresponding to the parameters chosen for Panel (a) of Figure 2. Further, the average normalised computation time is only computed once, instead of having three repeats as in Appendix A.2. Considering the relatively small fluctuations between repeats for all scenarios in Figure A2, this is not expected to significantly affect the qualitative result.

The model parameters and the associated average runtimes are shown in Table S1. Due to the excessive computation needed to solve the EBCM, only 10 runs of this script were performed. However, considering that the average normalised runtime for solving the EBCM is several orders of magnitude higher than the average normalised runtime for solving dynamical system (6a-c), this again is not expected to significantly alter the results. Finally, the computation times necessary for solving the EBCM are relatively homogeneous over all runs, indicating that the average computation time is not biased by an outlier.

	Runs	Normalised runtimes		
		Average	Minimum	Maximum
Dynamical system (6a-c)	50	0.15	0.14	0.17
EBCM	10	2076	1887	2254

Table S1: Numerical assessment of the computation time needed to solve either dynamical system (6a-c) or the EBCM introduced in Appendix S1.1. Model parameters: household and workplace size distribution from Figure 1; $(\beta_G, \lambda_H, \lambda_W, \gamma) = (0.125, 1.5, 0.00115, 0.125)$; initial proportion of infected $\varepsilon = 0.005$; resolution of the numerical system over the time interval $[0, 30]$.

S2 Proofs

S2.1 Proof of Proposition 3.10

Let us start with the following lemma, which will be needed afterwards.

Lemma A.1. *Consider a solution y of the dynamical system (6a-c) and let $\Delta(t) = m_X s(t) - \sum_{(S,I) \in \mathbb{S}} S n_{(S,I)}^X(t)$. Then*

$$\frac{d}{dt} \Delta(t) = \gamma n_{(1,1)}^X(t) - \left(\tau_G(t) + \frac{1}{s(t)} \tau_{\bar{X}}(t) \right) \Delta(t).$$

Proof of Lemma A.1. Let $X \in \{H, W\}$. First, notice that

$$\{(S+1, I-1) : (S, I) \in \mathbb{S}, I \geq 1\} = \{(S, I) \in \mathbb{S} : S > 1\}.$$

As $S-1=0$ whenever $S=1$, we thus obtain that

$$\begin{aligned} & \sum_{(S,I) \in \mathbb{S}} S^2 I n_{(S,I)}^X - \sum_{(S,I) \in \mathbb{S}} S(S+1)(I-1) n_{(S+1, I-1)}^X \mathbf{1}_{\{I \geq 1\}} \\ &= \sum_{(S,I) \in \mathbb{S}} S^2 I n_{(S,I)}^X - \sum_{(S,I) \in \mathbb{S}} (S-1) S I n_{(S,I)}^X = \sum_{(S,I) \in \mathbb{S}} S I n_{(S,I)}^X. \end{aligned} \quad (\text{S2})$$

Similarly, $\{(S, I+1) : (S, I) \in \mathbb{S}, S+I < n_{\max}\} = \mathbb{S} \setminus \{(S, I) \in \mathbb{S} : I=0\} \cup \{(1, 1)\}$. As $SI=0$ whenever $I=0$, it follows that

$$\sum_{(S,I) \in \mathbb{S}} \gamma S I n_{(S,I)}^X - \sum_{(S,I) \in \mathbb{S}} \gamma S(I+1) n_{(S, I+1)}^X \mathbf{1}_{\{S+I < n_{\max}\}} = \gamma n_{(1,1)}^X. \quad (\text{S3})$$

The desired conclusion then results directly from Equations (6a-c), regrouping the terms of the form of Equations (S2) and (S3) in order to simplify the expression. \square

We are now ready to focus on the desired result.

Proof of proposition 3.10. (i) By assumption, $y(0) \in V$. Let us start by checking that all components of y , as well as Δ , stay non-negative over time.

Let $t_0 \geq 0$ be such that $y(t_0) \in V$. If $i(t_0) = 0$, then $i'(t_0) = -s'(t_0) \geq 0$ by assumption, which ensures that i will not become negative on a neighbourhood of t_0 . Similar arguments hold for the lower bounds of Δ and s , using Lemma A.1 and inequality (26), respectively.

Let us now turn our attention to $n_{(S,I)}^X$ for $X \in \{H, W\}$ and $(S, I) \in \mathbb{S}$. Recall from Equation (6c) that its derivative may be ill defined, due to the division by $s(t)$. However, inequality (26) ensures that the ratios $\tau_{\bar{X}} n_{(S,I)}^X / s$ are well defined at all time, for any $(S, I) \in \mathbb{S}$. As a

consequence, we may now notice as previously that, if $n_{(S,I)}^X(t_0) = 0$, Equation (6c) ensures that $\frac{d}{dt}n_{(S,I)}^X(t_0) \geq 0$.

The desired conclusion follows: whenever either of the quantities of interest reach zero, their derivatives are non-negative which ensures that they do not become negative shortly thereafter.

Next, let us have a look at the upper bounds. For $X \in \{H, W\}$, a brief computation yields

$$\frac{d}{dt} \sum_{(S,I) \in \mathbb{S}} n_{(S,I)}^X(t) = -\gamma n_{(1,1)}^X(t) \leq 0.$$

This assures that starting from $y^* \in V$, the inequality $\sum_{(S,I) \in \mathbb{S}} n_{(S,I)}^X(t) \leq 1$ holds. For $X \in \{H, W\}$ and $(S, I) \in \mathbb{S}$, it follows that if $n_{(S,I)}^X(t_0) = 1$, then for any $(S', I') \in \mathbb{S} \setminus (S, I)$, $n_{(S',I')}^X(t_0) = 0$. Thus

$$\frac{d}{dt} n_{(S,I)}^X(t_0) = - \left(\lambda_X SI + \tau_{\bar{X}}(t_0) \frac{S}{s(t_0)} + \tau_G(t_0)S + \gamma I \right) \leq 0,$$

from which one may deduce that $n_{(S,I)}^X$ remains less than or equal to one. The remaining upper bounds on s , i and $s + i$ may be obtained using similar arguments.

We thus have established that if $y^* \in V$, then $y(t) \in V$ for all t such that y is well defined.

(ii) In order to prove that there exists at most a unique solution y for any initial condition $y^* \in V$, let us start by showing that f is Lipschitz continuous on V .

First, consider f_s . Let $y = (s, i, n_{(S,I)}^X : X \in \{H, W\}, (S, I) \in \mathbb{S})$ and $\hat{y} = (\hat{s}, \hat{i}, \hat{n}_{(S,I)}^X : X \in \{H, W\}, (S, I) \in \mathbb{S})$ be two elements of V , and let $X \in \{H, W\}$. Letting $c_X = \lambda_X \# \mathbb{S} (n_{\max})^2 / m_X$, we obtain that

$$|\tau_X(y) - \tau_X(\hat{y})| \leq \frac{\lambda_X}{m_X} \sum_{(S,I) \in \mathbb{S}} SI |n_{(S,I)}^X - \hat{n}_{(S,I)}^X| \leq c_X \|y - \hat{y}\|_\infty.$$

Further, recall that $0 \leq \tau_G(y) \leq \beta_G$ for any $y \in V$.

Letting $c_s = c_H + c_W + \beta_G$, it follows that

$$|f_s(y) - f_s(\hat{y})| \leq |\tau_H(y) - \tau_H(\hat{y})| + |\tau_W(y) - \tau_W(\hat{y})| + |\tau_G(y)s - \tau_G(\hat{y})\hat{s}| \leq c_s \|y - \hat{y}\|_\infty.$$

Similarly, letting $c_i = c_s + \gamma$,

$$|f_i(y) - f_i(\hat{y})| \leq |f_s(y) - f_s(\hat{y})| + \gamma |i - \hat{i}| \leq c_i \|y - \hat{y}\|_\infty.$$

Finally, notice that inequality (26) implies that $\max(Sn_{(S,I)}^X/s, S\hat{n}_{(S,I)}^X/\hat{s}) \leq m_X$ for any $X \in \{H, W\}$ and $(S, I) \in \mathbb{S}$. As a consequence, we obtain that

$$|\tau_{\bar{X}}(y) \frac{Sn_{(S,I)}^X}{s} - \tau_{\bar{X}}(\hat{y}) \frac{S\hat{n}_{(S,I)}^X}{\hat{s}}| \leq m_X |\tau_{\bar{X}}(y) - \tau_{\bar{X}}(\hat{y})| \leq m_X c_{\bar{X}} \|y - \hat{y}\|_\infty,$$

from which we may deduce, letting $C_X = 2(\lambda_X (n_{\max})^2 + m_X c_{\bar{X}} + \beta_G + \gamma)$, that

$$|f_{X;(S,I)}(y) - f_{X;(S,I)}(\hat{y})| \leq C_X \|y - \hat{y}\|_\infty.$$

This establishes the desired Lipschitz continuity of f on V , with associated Lipschitz constant $c_L = \max(C_H, C_W)$.

Suppose now that there are two solutions y and \bar{y} of Equations (6a-c) and such that $y(0) = \bar{y}(0)$. It then holds that for any $T \geq 0$,

$$\|y(T) - \bar{y}(T)\|_\infty \leq \int_0^T \|f(y(t)) - f(\bar{y}(t))\|_\infty dt \leq c_L \int_0^T \|y(t) - \bar{y}(t)\|_\infty dt,$$

Thus Gronwall's lemma ensures that $\|y(t) - \bar{y}(t)\|_\infty = 0$ for any $t \leq T$. T being arbitrary, the desired conclusion on uniqueness follows.

(iii) In order to establish (iii), it simply remains to show that the initial condition y^* defined by Equation (7) belongs to V . Let us start by noticing that, following Equation (7), for any $X \in \{H, W\}$,

$$\sum_{(S,I) \in \mathbb{S}} n_{(S,I)}^X(0) = \sum_{n=2}^{n_{\max}} \pi_n^X \sum_{I=0}^{n-1} \binom{n}{I} \varepsilon^I (1-\varepsilon)^{n-I} = \sum_{n=2}^{n_{\max}} \pi_n^X (1-\varepsilon^n) \leq 1.$$

Similarly,

$$\begin{aligned} \sum_{(S,I) \in \mathbb{S}} S n_{(S,I)}^X(0) &= \sum_{n=2}^{n_{\max}} \pi_n^X \sum_{I=0}^{n-1} (n-I) \binom{n}{I} \varepsilon^I (1-\varepsilon)^{n-I} = \sum_{n=2}^{n_{\max}} \pi_n^X n (1-\varepsilon) \\ &= (m_X - \pi_1^X) (1-\varepsilon) \leq m_X s(0), \end{aligned}$$

where we have used the fact that $s(0) = 1 - \varepsilon$, and recognizing $\mathbb{E}[n - B]$ for $B \sim \mathcal{B}(n, \varepsilon)$ to deduce the second equality. The other conditions following immediately from (7), we conclude that $y^* \in V$. \square

S2.2 Proof of Lemma 3.11

Let $X \in \{H, W\}$. To begin with, let us introduce the family of sets

$$E_{(n,s)} = \{x \in E : \mathbf{n}(x) = n, \mathbf{s}(x) = s\}, \quad \forall (n, s) \in \llbracket 1, n_{\max} \rrbracket \times \llbracket 0, n \rrbracket$$

which constitutes a partition of E . In the following, let (n_t, s_t, τ_t) be distributed according to η_t^X , for any $t \geq 0$. Let f be a non-negative measurable function on E . Then

$$\begin{aligned} \langle \eta_t^X, f \rangle &= \mathbb{E}[f(n_t, s_t, \tau_t)] \\ &= \sum_{n=1}^{n_{\max}} \sum_{s=0}^n \mathbb{E}[f(n_t, s_t, \tau_t) | n_t = n, s_t = s] \mathbb{P}(n_t = n, s_t = s). \end{aligned}$$

Conditionally on $\{n_t = n, s_t = s\}$, almost surely $(n_t, s_t, \tau_t) \in E_{(n,s)}$ and $(\tau_t)_k = 0$ for any $k > n - s$. Thus

$$\begin{aligned} \langle \eta_t^X, f \rangle &= \sum_{n=1}^{n_{\max}} f(n, n, 0_{n_{\max}}) \mathbb{P}(n_t = n, s_t = s) \\ &\quad + \sum_{n=1}^{n_{\max}} \sum_{s=0}^{n-1} \mathbb{E} \left[f \left(n, s, \sum_{k=1}^{n-s} (\tau_t)_k e_k \right) \middle| n_t = n, s_t = s \right] \mathbb{P}(n_t = n, s_t = s). \end{aligned} \tag{S4}$$

Let $p_t^X(n, s) = \mathbb{P}(n_t = n, s_t = s)$. Then on the one hand, define for any $n \in \llbracket 1, n_{\max} \rrbracket$,

$$\rho_t^{X,n,n} = p_t^X(n, n).$$

On the other hand, for $s < n$, let $\tilde{\eta}_t^{X,n,s}$ be the distribution of $((\tau_t)_1, \dots, (\tau_t)_{n-s})$ conditionally on $\{n_t = n, s_t = s\}$. If we manage to show that $\tilde{\eta}_t^{X,n,s}$ is absolutely continuous with regard to the Lebesgue measure on \mathbb{R}^{n-s} , then letting $\tilde{\rho}_t^{X,n,s}$ be its density, the result follows with

$$\forall \tau \in \mathbb{R}^{n-s}, \quad \rho_t^{X,n,s}(\tau) = p_t^X(n, s) \tilde{\rho}_t^{X,n,s}(\tau).$$

In order to show that for any $n \in \llbracket 1, n_{\max} \rrbracket$ and $s \in \llbracket 1, n-1 \rrbracket$, $\tilde{\eta}_t^{X,n,s}$ is absolutely continuous with regard to the Lebesgue measure on \mathbb{R}^{n-s} , we will proceed by induction.

Let us first consider the case $s = n-1$. We want to apply Equation (4) to test functions $\varphi^* \in \mathcal{C}_b^1(E, \mathbb{R}_+)$ which are equal to zero outside of the set $\{x \in E : \mathbf{n}(x) = n, \mathbf{s}(x) = n-1\}$. Hence, let $\varphi \in \mathcal{C}_b^1(\mathbb{R})$ be non-negative, and define $\varphi^* \in \mathcal{C}_b^1(E, \mathbb{R}_+)$ by

$$\forall x \in E, \quad \varphi^*(x) = \varphi(\tau_1(x)) \mathbf{1}_{\{\mathbf{n}(x), \mathbf{s}(x) = (n, s)\}}.$$

Let $T > 0$ and for $t \in [0, T]$, define $f_t(x) = \varphi^*(\Psi(x, T, t))$ for $x \in E$. Using Equation (S4) and applying Equation (4) to f_t after noticing that $\mathcal{A}f_t(x) = 0$ for any $(t, x) \in \mathbb{R}_+ \times E$, we obtain that

$$\begin{aligned} p_t^X(n, n-1) \langle \tilde{\eta}_T^{X,n,n-1}, \varphi \rangle &= \langle \eta_T^X, \varphi^* \rangle = \langle \eta_T^X, f_T \rangle \\ &= \langle \eta_0^X, f_0 \rangle + \lambda_H \int_0^T \langle \eta_t^X, \mathbf{si}(f_t^{\mathcal{I}} - f_t) \rangle dt + \int_0^T \Lambda_t^X \langle \eta_t^X, \mathbf{s}(f_t^{\mathcal{I}} - f_t) \rangle dt. \end{aligned}$$

As $\varphi \geq 0$ and letting $C = n_{\max}(\lambda_H n_{\max} + \lambda_W n_{\max} + \beta_G)$, it follows that

$$p_t^X(n, n-1) \langle \tilde{\eta}_T^{X,n,n-1}, \varphi \rangle \leq \langle \eta_0^X, f_0 \rangle + C \int_0^T \langle \eta_t^X, f_t^{\mathcal{I}} \rangle dt. \quad (\text{S5})$$

Notice that Equation (5) with a change of variables $z = \sigma - T$ implies that

$$\langle \eta_0^X, f_0 \rangle = n\pi_n^X (1-\varepsilon)^{n-1} \varepsilon \int_{\mathbb{R}} \varphi(z) \mathbf{1}_{\{z > -T\}} \gamma e^{-\gamma(z+T)} dz.$$

Furthermore, by definition, for any $x \in E$,

$$f_t^{\mathcal{I}}(x) = \mathbf{1}_{\{\mathbf{n}(x)=n, \mathbf{s}(x)=n\}} \int_0^{+\infty} \varphi(\sigma - (T-t)) \gamma e^{-\gamma\sigma} d\sigma.$$

Thus Equation (S4), Fubini's theorem and a change of variables $z = \sigma - (T-t)$ lead to

$$\int_0^T \langle \eta_t^X, f_t^{\mathcal{I}} \rangle dt = \int_{\mathbb{R}} \varphi(z) \int_0^T p_t^X(n, n) \gamma e^{-\gamma(z+(T-t))} dt dz.$$

Let

$$\begin{aligned} H_T^{X,n,n-1} : z \in \mathbb{R} &\mapsto \left(n\pi_n^X (1-\varepsilon)^{n-1} \varepsilon \mathbf{1}_{\{z > -T\}} \gamma e^{-\gamma(z+T)} \right) \\ &+ \int_0^T p_t^X(n, n) \gamma e^{-\gamma(z+(T-t))} dt. \end{aligned}$$

Then $H_T^{X,n,n-1}$ is integrable on \mathbb{R} with respect to the Lebesgue measure, as its integral is equal to the right-hand side of Equation (S5) which is finite. It further satisfies

$$p_t^X(n, s) \langle \tilde{\eta}_T^{X,n,n-1}, \varphi \rangle \leq \int_{\mathbb{R}} H_T^{X,n,n-1}(z) \varphi(z) dz.$$

In particular, let B be a Borel set which is null for the Lebesgue measure. Consider a mollifier ψ on \mathbb{R} in the same sense as for the Proof of Lemma 3.1. For $\alpha > 0$, let $\psi_\alpha : z \in \mathbb{R} \mapsto \alpha^{-1} \psi(z/\alpha)$. Then for any $\alpha > 0$, we may define $\varphi_\alpha = \mathbf{1}_B * \psi_\alpha$, which is an element of $\mathcal{C}^1(\mathbb{R})$ with compact support. Thus for every $\alpha > 0$,

$$p_T^X(n, s) \langle \tilde{\eta}_T^{X,n,n-1}, \varphi_\alpha \rangle \leq \int_{\mathbb{R}} H_T^{X,n,n-1}(z) \varphi_\alpha(z) dz. \quad (\text{S6})$$

As φ_α further is bounded by 1 for every $\alpha > 0$, and as $\tilde{\eta}_T^{X,n,n-1}$ is a finite measure and $H_T^{X,n,n-1}$ is integrable on \mathbb{R} with respect to the Lebesgue measure, dominated convergence allows to let α go to 0 on both sides of inequality (S6). As B is a null set for the Lebesgue measure, the right-hand side goes to 0. This demonstrates that $\tilde{\eta}_T^{X,n,n-1}$ is absolutely continuous with respect to the Lebesgue measure on \mathbb{R} .

Let us now suppose that $\tilde{\eta}_T^{X,n,s+1}$ is absolutely continuous with respect to the Lebesgue measure on \mathbb{R}^{n-s-1} for some $n' \in \llbracket 1, n_{\max} \rrbracket$ and some $s \in \llbracket 0, n-2 \rrbracket$. This time, let $\varphi \in \mathcal{C}_b^1(\mathbb{R}^{n-s})$ be non-negative and define $\varphi^* \in \mathcal{C}_b^1(E, \mathbb{R}_+)$ by

$$\forall x \in E, \quad \varphi^*(x) = \mathbf{1}_{\{\mathbf{n}(x)=n, \mathbf{s}(x)=s\}} \varphi(\tau_1(x), \dots, \tau_{n-s}(x)).$$

For any $T \geq 0$ and $t \in [0, T]$, let $f_t : x \in E \rightarrow \varphi^*(\Psi(x, T, t))$. Proceeding like before, we obtain that

$$p_T^X(n, s) \langle \tilde{\eta}_T^{X,n,s}, \varphi \rangle \leq \langle \eta_0^X, f_0 \rangle + C \int_0^T \langle \eta_t^X, f_t^I \rangle dt. \quad (\text{S7})$$

In order to simplify notations in the following, define for $T \geq t \geq 0$ and $(\tau_1, \dots, \tau_{n-s}) \in \mathbb{R}^{n-s}$ the function

$$\varphi_{T,t}(\tau_1, \dots, \tau_{n-s}) = \varphi(\tau_1 - (T-t), \dots, \tau_{n-s} - (T-t)).$$

Then similarly to the case $(n, n-1)$, Equation (5) leads to

$$\langle \eta_0^X, f_0 \rangle = \pi_n^X \binom{n}{s} (1-\varepsilon)^s \varepsilon^{n-s} \int_{\mathbb{R}^{n-s}} \varphi_{T,0}(\sigma_1, \dots, \sigma_{n-s}) \prod_{j=1}^{n-s} (\mathbf{1}_{\{\sigma_j > 0\}} \gamma e^{-\gamma \sigma_j}) d\sigma_1 \dots d\sigma_{n-s}. \quad (\text{S8})$$

Notice here that for any $T > 0$ and $t \in [0, T]$, the application $z \in \mathbb{R}^{n-s} \mapsto z - \sum_{k=1}^{n-s} (T-t) e_k$, where e_k is the k -th vector of the canonical basis of \mathbb{R}^{n-s} , defines a \mathcal{C}^1 -diffeomorphism from \mathbb{R}^{n-s} into itself, whose Jacobian matrix is the identity matrix of dimension $n-s$. Using the case $t=0$, Equation (S8) becomes by change of variables:

$$\langle \eta_0^X, f_0 \rangle = \pi_n^X \binom{n}{s} (1-\varepsilon)^s \varepsilon^{n-s} \int_{\mathbb{R}^{n-s}} \varphi(z_1, \dots, z_{n-s}) \prod_{j=1}^{n-s} (\mathbf{1}_{\{z_j > -T\}} \gamma e^{-\gamma(z_j+T)}) dz_1 \dots dz_{n-s}.$$

Moreover, notice that this time, for any $x \in E$ and $\sigma > 0$,

$$f_t^I(x) = \mathbf{1}_{\{\mathbf{n}(x)=n, \mathbf{s}(x)-1=s\}} \varphi_{T,t}(\tau_1(x), \dots, \tau_{n-s-1}(x), \sigma).$$

Thus, for $t \in [0, T]$,

$$\langle \eta_t^X, f_t^I \rangle = \int_{\mathbb{R}^{n-s-1}} \left(\int_0^{+\infty} \varphi_{T,t}(\tau_1, \dots, \tau_{n-s-1}, \sigma) \gamma e^{-\gamma \sigma} d\sigma \right) p_t^X(n, s+1) \tilde{\eta}_t^{X,n,s+1}(d\tau_1, \dots, d\tau_{n-s-1}).$$

Fubini's theorem together with the induction hypothesis yield

$$\langle \eta_t^X, f_t^I \rangle = \int_{\mathbb{R}_+} \int_{\mathbb{R}^{n-s-1}} \varphi_{T,t}(\tau_1, \dots, \tau_{n-s-1}, \sigma) p_t^X(n, s+1) \tilde{\rho}_t^{X,(n,s+1)}(\tau_1, \dots, \tau_{n-s-1}) d\tau_1 \dots d\tau_{n-s-1} \gamma e^{-\gamma \sigma} d\sigma.$$

The previously introduced family of \mathcal{C}^1 -diffeomorphisms may serve again for a change of variables, allowing to obtain that

$$\begin{aligned} \langle \eta_t^X, f_t^I \rangle &= p_t^X(n, s+1) \int_{\mathbb{R}^{n-s}} \left(\varphi(z_1, \dots, z_{n-s-1}, z_{n-s}) \mathbf{1}_{\{z_{n-s} > -(T-t)\}} \gamma e^{-\gamma(z_{n-s} + (T-t))} \right. \\ &\quad \left. \times \rho_t^{X,n,s+1}(z_1 + (T-t), \dots, z_{n-s-1} + (T-t)) \right) dz_1 \dots dz_{n-s}. \end{aligned}$$

For $z = (z_1, \dots, z_{n-s}) \in \mathbb{R}^{n-s}$, define

$$H_T^{X,n,s}(z) = \pi_n^X \binom{n}{s} (1 - \varepsilon)^s \varepsilon^{n-s} \prod_{j=1}^{n-s} (\mathbf{1}_{\{z_j > -T\}} \gamma e^{-\gamma(z_j + T)}) \\ + \int_0^T \rho_t^{X,n,s+1}(z_1 + (T-t), \dots, z_{n-s-1} + (T-t)) \mathbf{1}_{\{z_{n-s} > -(T-t)\}} \gamma e^{-\gamma(z_{n-s} + (T-t))} dt.$$

It then follows like before from Equation (S7) that $H_T^{X,n,s}$ is integrable on \mathbb{R}^{n-s} with respect to the Lebesgue measure, and

$$p_T^X(n, s) \langle \tilde{\eta}_T^{X,n,s}, \varphi \rangle \leq \int_{\mathbb{R}^{n-s}} H_T^{X,n,s}(z_1, \dots, z_{n-s}) dz_1 \dots dz_{n-s}.$$

The absolute continuity of $\tilde{\eta}_T^{X,n,s}$ with respect to the Lebesgue measure on \mathbb{R}^{n-s} follows using the same arguments as previously. This concludes the proof. \square