INFECTION MODELS ON DENSE DYNAMIC RANDOM GRAPHS

SIMONE BALDASSARRI, PETER BRAUNSTEINS, FRANK DEN HOLLANDER,
AND MICHEL MANDJES

ABSTRACT. We consider Susceptible-Infected-Recovered (SIR) models on dense dynamic random graphs, in which the joint dynamics of vertices and edges are *co-evolutionary*, i.e., they influence each other bidirectionally. In particular, edges appear and disappear over time depending on the states of the two connected vertices, on how long they have been infected, and on the total density of susceptible and infected vertices. Our main results establish functional laws of large numbers for the densities of susceptible, infected, and recovered vertices, jointly with the underlying evolving random graphs in the graphon space. Our results are supported by simulations, which characterize the limiting size of the epidemics, i.e., the limiting density of susceptible vertices, and how the *peak* of the epidemics depends on the rate of the evolution of the underlying graph.

The proofs of our main results rely on the careful construction of a *mimicking process*, obtained by approximating the two-way feedback interaction between vertex and edge dynamics with a mean-field type interaction, acting only as one-way feedback, that remains sufficiently close to the original co-evolutionary process. To treat the more general setting in which edge dynamics are affected by the proportions of susceptible and infected individuals, we introduce a methodological extension of existing techniques. We thus show that our model exhibits multiple epidemic peaks – a phenomenon observed in real-world epidemics – which can emerge in models that incorporate mutual feedback between vertex and edge dynamics.

MSC 2020 *subject classifications.* 60F17, 60K35, 60K37.

Key words and phrases. SIR dynamics, graph dynamics, co-evolution.

Acknowledgment. The work in this paper was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 101034253, and by the NWO Gravitation project NETWORKS under grant no. 024.002.003. SB was further supported through "Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni" (GNAMPA-INdAM).

Date: February 16, 2026.

Contents

1. Introduction 2
2. Model and main results 5
2.1. Model definition 5
2.2. Main results 9
3. Illustrations 10
3.1. Dynamics without global feedback 10
3.2. Multiple peaks with global feedback 13
4. Road map 16
5. Proofs 19
5.1. Proof of Theorem 2.1 19
5.2. Coupling 21
5.3. Proof of Theorem 2.2 26
Appendix A. Graphons, Lévy metric and an additional lemma 26
References 27

1. Introduction

Mathematical models play a crucial role in understanding and managing the spread of infectious diseases. These models offer insight into how various parameters influence the dynamics of an epidemic, which can be instrumental in forecasting outbreaks and designing effective control and mitigation strategies.

The base epidemic model is the *Susceptible-Infected-Recovered* (SIR) compartmental model. The deterministic version of this model is characterized by a system of differential equations that captures the evolution of the proportion of susceptible, infected, and recovered individuals in the population over time. The same differential equations also appear in the stochastic version of the model, as a large population limit established through a *functional law of large numbers* (FLLN). In the present paper, we establish such a FLLN for a considerably more general version of the base SIR model.

The basic version of the SIR model relies on simplifying assumptions that are typically not satisfied in practice. One key assumption is *homogeneous mixing*, which implies that every individual in the population is equally likely to come into contact with every other individual. If we represent each individual in the population as a vertex in a graph and each social connection through which the epidemic can spread as an edge, then the assumption of homogeneous mixing

is equivalent to assuming that *every edge is present in the graph*. However, this assumption overlooks several crucial aspects of real-world social networks during a pandemic. In reality, the network is *heterogeneous*, reflecting that some individuals have more social connections than others. It is also *dynamic*, as connections can appear and disappear over time. Moreover, these dynamics are *co-evolutionary*: the individual epidemiological states (i.e., susceptible, infected, or recovered) are influenced by the structure of the network (since the epidemic spreads through existing connections) and the network itself evolves in response to the epidemic, as individuals may choose to alter their connections to prevent further spread of the disease.

At a more precise level, a realistic model should incorporate two types of co-evolution:

(a) First, the dependence of the network on the individual epidemiological states can be *local*. For example, susceptible individuals may choose to break ties with other individuals they suspect to be infected.
(b) Second, there can be *global* feedback mechanisms. For instance, quarantine measures may be introduced once the proportion of infected individuals exceeds a certain threshold.

Ideally, we would work with a model in which the underlying graph is heterogeneous (i.e., not complete), is dynamic (i.e., evolves over time), and exhibits co-evolution (i.e., the network influences the epidemic dynamics and is, in turn, shaped by them – both at local and global scales).

In this paper we introduce a stochastic SIR model that explicitly captures the realistic network features discussed above. We analyze this model in the *dense regime* in which the number of edges scales roughly as n^2, with n denoting the number of individuals. Our first goal is to establish a FLLN for the proportion of susceptible, infected, and recovered individuals over time. While our model is inherently more complex, this result is in a similar spirit as the FLLN for the base SIR model. In view of the general co-evolutionary dynamics incorporated in the model, we are also interested in how the network evolves during the pandemic. Our second objective is therefore to establish a FLLN for the dynamic network. The proposed model is capable of reproducing complex real-world phenomena, such as the emergence of multiple infection peaks — patterns that cannot be captured in SIR models without incorporating the co-evolutionary feedback between individual states and network structure. We use our results to gain insight into the evolution of the epidemic and the state of the network as the epidemic evolves.

Our results should be viewed within the context of SIR models on random graphs. We begin by providing a brief overview of this subarea. Broadly, the literature can be categorized into four main groups:

- A FLLN for the SIR model on a *sparse static* configuration model was proposed in [37]. This result was subsequently rigorously established in [8], [10], [19], and [26] under progressively weaker assumptions on the distribution of the vertex degrees in the

configuration model. A similar result for an SIR model on a sparse static stochastic block model appeared more recently in [13].

- FLLNs for SIR models on *dense static* random graphs have been considered in [20, 28, 34]. In these papers, the random graph through which the epidemic spreads is constructed by sampling from a reference graphon. The results in [34] apply when the process exhibits non-Markovian dynamics but hold only when then number of edges scale as n^2, whereas the results in [20, 28] consider Markovian dynamics but also hold when the number of edges scale as n^{1+a} for $a \in (0, 1]$.
- FLLNs for the SIR model on *sparse dynamic* random graphs with co-evolutionary feedback are established in [5, 6, 16, 27] (see also [21] for the SI model). In these papers, susceptible individuals can break edges with infected neighbors, potentially rewiring the edge to another individual. A primary focus of these papers is to establish a discontinuity in the final proportion of susceptible individuals at the critical value of the contact rate. In these papers, the co-evolutionary dynamics are local, in the sense that susceptible individuals break the connection with their infected neighbors in a manner that does not depend on the overall state of the epidemic (i.e., the total number of infected individuals).
- A FLLN for the SIR model on *dense dynamic* random graphs is established in [24]. In this paper the underlying graph is a dynamic stochastic block model that allows for graph degrees that scale at least as $\log n$ (so at least $n \log n$ edges). However, there is no co-evolutionary feedback in the graph dynamics considered, i.e., the state of the graph evolves independently of the epidemic states of the individuals.

Currently, the literature lacks a FLLN for an SIR model on a dense dynamic random graph that incorporates both local and global co-evolutionary feedback. Moreover, no efforts have been made to establish such a law for the evolution of the network itself (i.e., jointly with the proportion of susceptible, infected, and recovered individuals). The present paper aims to fill these gaps. More broadly, our contributions belong to the growing literature for related models on dynamic random graphs. See for instance [2, 9, 18, 23, 25, 29, 33, 36] and references therein.

Methodologically, our paper builds on two earlier works [15, 3] concerning graph-valued stochastic processes. A central element of these papers is the concept of a *graphon* (developed in [11, 12, 30, 31, 32]), which is used to describe the limit of the network. Graphons have been extensively used to characterize the limit of other dense graph-valued processes; see for instance [1, 4, 14, 17, 22, 35]. However, to the best of our knowledge, the present paper is the first to establish a limit in the space of graphons for a co-evolutionary model with a global feedback mechanism.

In [15] limits are established for a graph-valued processes with one-way dependence (i.e., not co-evolutionary), in the setting where the edge dynamics can depend on the collective states of all vertices across the network (i.e., global dependence). On the other hand, in [3] limits are established for graph-valued processes that are co-evolutionary, but where the co-evolutionary

feedback is local only (i.e., no global feedback). In the present paper, we use the mechanism in [15] to allow the edge dynamics to depend, for example, on the total number of infected individuals (i.e., global feedback), and we adapt the proof technique in [3] to establish our results in this more general setting. This technique builds on the notion of a *mimicking process* introduced in [3], which approximates the bidirectional interaction between vertex and edge dynamics by a mean-field-type, *single-directional* interaction that remains sufficiently close to the original co-evolutionary process. Our analysis suggests that incorporating global feedback mechanisms can be achieved with surprising ease within the framework established in [3]. We expect that this approach is applicable more broadly to other dense co-evolutionary processes.

The paper is organized as follows. In Section 2 we detail the system dynamics, present our main theoretical findings, and provide some discussion, concluding remarks and outlook. The simulations reported in Section 3 reveal that our theory is capable of reproducing patterns observed in practice, most notably trajectories with multiple peaks. A road map of the proof is provided in Section 4, while the actual proofs are provided in Section 5.

2. Model and main results

In this section we first define the stochastic dynamics underlying our SIR model on a dense dynamic random graph. This description specifies the way that we incorporate both local and global co-evolutionary feedback. We then state and briefly discuss our main results.

2.1. **Model definition.** In this paper we analyze an SIR model on a dynamic random graph. Given a time horizon $T > 0$, we denote by $(G_n(t))_{t\in[0,T]}$ an evolving random graph with vertex set $[n] := \{1, 2, ..., n\}$. For $i, j \in [n]$, we let $e_{ij}(t) = 1$ if edge ij is active at time t and 0 otherwise. The graph $G_n(t)$ is undirected, so $e_{ij}(t) = e_{ji}(t)$, and contains no self-loops, so $e_{ii}(t) = 0$. Let $x_i(t) \in \{\mathrm{S}, \mathrm{I}, \mathrm{R}\}$ be the state of vertex $i \in [n]$ at time $t \in [0, T]$, where 'S' means that it is susceptible, 'I' means that it is infected, and 'R' means that it is recovered.

Initialization. At time $t = 0$ we initialize the graph as an Erdős-Rényi random graph with connection probability $p_0 \in (0, 1)$. Each vertex is infected with probability $q_0 \in (0, 1)$ and susceptible with probability $1 - q_0$, independently of the states of the other vertices and of the initial random graph. Thus, no vertex is in the recovered state at $t = 0$.

Definition of vertex types. For any infected vertex $i \in [n]$, we define its *infection time* as

$$t_i^{\mathrm{I}} := \inf\{s \in [0, T] : x_i(s) = \mathrm{I}\}.$$

We define *type* of vertex i at time t as

$$y_i(t) := \begin{cases} -1 & \text{if } x_i(t) = \mathrm{S}, \\ t - t_i^{\mathrm{I}} & \text{if } x_i(t) = \mathrm{I}, \\ T + 1 & \text{if } x_i(t) = \mathrm{R}, \end{cases} \tag{2.1}$$

where $t - t_i^{\mathrm{I}}$ is the length of time that vertex i has been infected for. We then define the *empirical type process*, $(F_n(t;\cdot))_{t\in[0,T]}$, through

$$F_n(t;y) := \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{y_i(t) \leqslant y\}, \quad y \in \mathbb{R}. \tag{2.2}$$

Vertex and edge dynamics. For each infected vertex i, we denote by $\mathscr{I}(\cdot)$ its *infectivity*, where $\mathscr{I} : [0,\infty) \to [0,1]$ is a continuous deterministic function.

- *Vertex dynamics.* Let $\mathcal{N}_i^{\mathrm{I}}(t)$ be the set of infected neighbors of vertex i at time t. At any time $t \in [0,T]$, $x_i(t)$ transitions from S to I at rate

$$\frac{\lambda}{n}\sum_{j\in\mathcal{N}_i^{\mathrm{I}}(t)} \mathscr{I}(y_j(t)),$$

  and $x_i(t)$ makes a transition from I to R at a (normalised) rate of 1.
- *Edge dynamics.* At any time $t \in [0,T]$, $e_{ij}(t)$ transitions from 0 to 1 (i.e., from inactive to active) at rate

$$\begin{cases} \gamma\,\pi_{\mathrm{SS}}(F_n(t;\cdot)), & \text{if } x_i(t) = x_j(t) = \mathrm{S}, \\ \gamma\,\pi_{\mathrm{SI}}(y_i(t), F_n(t;\cdot)), & \text{if } x_i(t) = \mathrm{I},\ x_j(t) = \mathrm{S}, \\ \gamma\,\pi_{\mathrm{SI}}(y_j(t), F_n(t;\cdot)), & \text{if } x_i(t) = \mathrm{S},\ x_j(t) = \mathrm{I}, \\ \gamma\,\pi_{\mathrm{II}}(y_i(t), y_j(t), F_n(t;\cdot)), & \text{if } x_i(t) = x_j(t) = \mathrm{I}, \end{cases}$$

  and $e_{ij}(t)$ transitions from 1 to 0 (i.e., from active to inactive) at the same rate with $\gamma\,\pi_s(\cdot)$ replaced by $\gamma\,(1-\pi_s(\cdot))$, with $s \in \{\mathrm{SS},\mathrm{SI},\mathrm{II}\}$. When $x_i(t)$ transitions from I to R we suppose that vertex i resamples adjacent edges with the initial probability p_0, after which point these edges are static (i.e., they no longer turn on and off).

Note that in our setup the functions $\pi_s(\cdot)$ can depend on $F_n(t;\cdot)$, and therefore on the proportion of susceptible, infected and recovered vertices, and also on the distribution of the times which infected vertices have been infected for. We impose a Lipschitz continuity condition on the functions $\pi_{\mathrm{II}}(\cdot)$. Specifically, we suppose that there exists $L_{\mathrm{II}} < \infty$ such that

$$|\pi_{\mathrm{II}}(a_1,b_1,F_1) - \pi_{\mathrm{II}}(a_2,b_2,F_2)| \leqslant L_{\mathrm{II}}[|a_1-a_2| + |b_1-b_2| + d_L(F_1,F_2)]$$

for all $a_1,a_2,b_1,b_2 > 0$ and all distribution functions F_1 and F_2, where d_L is the Lévy metric defined in (A.4). We also impose analogous Lipschitz continuity assumptions on π_{SS} and π_{SI}.

Graphons. For any $t \in [0,T]$, let $h^{G_n(t)}$ be a function from $[0,1]^2$ to $[0,1]$ which is characterized by

$$h^{G_n(t)}(x,y) := e_{\lceil nx\rceil,\lceil ny\rceil}(t). \tag{2.3}$$

The function $h^{G_n(t)}$ is referred to as *empirical graphon* associated with $G_n(t)$ and encodes the adjacency matrix of $G_n(t)$ (see Figure 2.1).

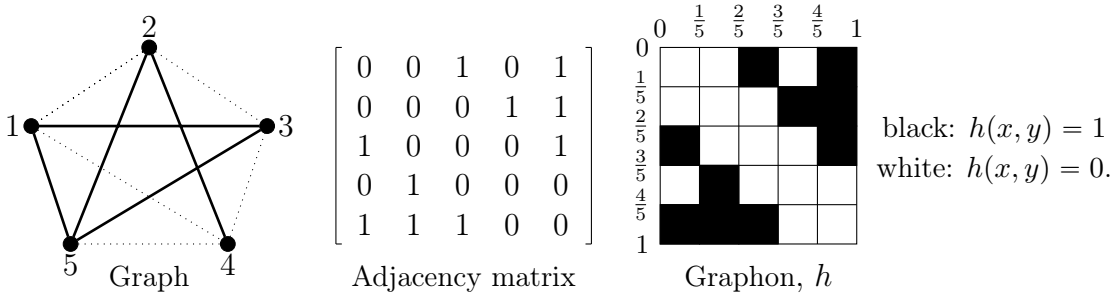

Figure 2.1: Illustration of the relation between graph, adjacency matrix, and graphon.

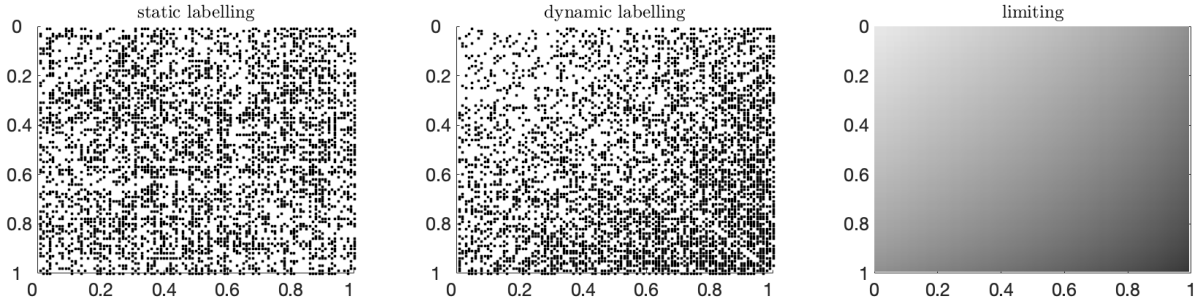

Figure 2.2: Two graphons with different labellings (left and middle panels), plus the limiting graphon when the labelling of the middle panel has been used (right panel).

Empirical graphons belong to the space of *graphons* $\mathscr{W}$, containing all functions $h : [0,1]^2 \mapsto [0,1]$ such that $h(x,y) = h(y,x)$ for all $(x,y) \in [0,1]^2$. We endow $\mathscr{W}$ with the *cut distance*:

$$d_{\square}(h_1, h_2) := \sup_{S,T \subseteq [0,1]} \left| \int_{S \times T} \mathrm{d}x \, \mathrm{d}y \, [h_1(x,y) - h_2(x,y)] \right|, \quad h_1, h_2 \in \mathscr{W}.$$

The space $\mathscr{W}$ is used to establish limiting results for random graphs; we establish our main result in the space of $\mathscr{W}$-valued paths.

We need to carefully choose how to label the vertices, since $d_{\square}(h_1, h_2)$ can depend on the labeling convention. For example, in Figure 2.2 the left and center empirical graphons represent the same graph up to a relabeling of vertices, but only the center graphon is close to the right limiting graphon in the cut distance. Throughout this paper we choose to label the vertices such that

$$y_1(t) \leqslant y_2(t) \leqslant \cdots \leqslant y_n(t), \quad \text{for all } t \in [0,T],$$

with ties broken arbitrarily. This means that susceptible vertices have a lower label than infected vertices, infected vertices have a lower label than recovered vertices, and infected vertices that have been infected for a short time have a lower label than those that have been infected for a long time. This also implies that the labels of the vertices change over time.

Candidate limit. In FLLNs, the limiting object can be informally understood as capturing the system's dynamics in the limit of large system size. This deterministic limit, often referred to as the *fluid limit*, is typically characterized by a system of differential equations. Below, we outline the procedure for identifying this candidate limit in the context of our system.

We let $(F(t;\cdot))_{t\in[0,T]}$ be a time-varying distribution function which can be thought of as a candidate limit of $(F_n(t;\cdot))_{t\in[0,T]}$ as $n\to\infty$. For any $t\in[0,T]$, $F(t;y)$ is absolutely continuous on $y\in[0,t)$ and has point masses at -1, t, and $T+1$, with no probability mass elsewhere. We characterize F through differential equations; however, in order to write them in a compact form we require additional notation. First, let $f_{\mathrm{I}}(t;y) := \frac{\mathrm{d}}{\mathrm{d}y}F(t;y)$ for $y\in[0,t)$, which can be interpreted as the limiting density of the amount of time that infected individuals have been infected for at time t, and

$$\begin{aligned} p_{\mathrm{S}}(t) &:= F(t;-1) - \lim_{s\uparrow -1} F(t;s), \\ p_{\mathrm{I}}(t) &:= q_0 e^{-t} + \int_0^{t-} \mathrm{d}y\, f_{\mathrm{I}}(t;y), \\ p_{\mathrm{R}}(t) &:= F(t;T+1) - \lim_{s\uparrow T+1} F(t;s), \end{aligned} \tag{2.4}$$

which can be interpreted as the limiting proportion of susceptible, infected, and recovered individuals at time t.

Next, let H be an edge connection probability function, defined by

$$\begin{aligned} &H(t;u,v,(F(s;\cdot))_{s\in[0,t]}) \\ &\quad := \mathbb{P}(e_{ij}(t)=1 \mid y_i(t)=u, y_j(t)=v, (F_n(s;\cdot))_{s\in[0,t]} = (F(s;\cdot))_{s\in[0,t]}); \end{aligned}$$

we remark that this probability is defined more rigorously as an integral of the objects π_{SS}, π_{SI}, and π_{II} that we introduced in Equation (4.2). Finally let

$$\mathcal{J}(t) := \mathcal{I}(t;F(t;\cdot)) := \int_0^t \mathrm{d}F(t;u)\, H(t;-1,u,F(t;\cdot))\, \mathscr{I}(u), \tag{2.5}$$

where $\lambda\mathcal{J}(t)$ can be interpreted as the limiting rate (as $n\to\infty$) at which susceptible individuals become infected at time t. The function F is characterized by the initial conditions

$$p_{\mathrm{S}}(0) = 1-q_0, \qquad p_{\mathrm{I}}(0) = q_0, \tag{2.6}$$

with the initial infection-age distribution concentrated at $u=0$ (that is, $f_{\mathrm{I}}(0;u) = q_0\,\delta_0(u)$), while the updating equations are

$$\begin{aligned} \frac{\partial}{\partial t}p_{\mathrm{S}}(t) &= -\lambda\mathcal{J}(t)\, p_{\mathrm{S}}(t), \\ \frac{\partial}{\partial t}f_{\mathrm{I}}(t;u) + \frac{\partial}{\partial u}f_{\mathrm{I}}(t;u) &= -f_{\mathrm{I}}(t;u), \qquad u>0, \\ f_{\mathrm{I}}(t;0) &= \lambda\mathcal{J}(t)\, p_{\mathrm{S}}(t), \\ \frac{\partial}{\partial t}p_{\mathrm{R}}(t) &= p_{\mathrm{I}}(t), \end{aligned} \tag{2.7}$$

where $f_{\mathrm{I}}(t;u)$ denotes the density of infected individuals with infection age u at time t, and $p_{\mathrm{I}}(t) = \int_0^\infty f_{\mathrm{I}}(t;u)\,\mathrm{d}u$. Due to the initially infected individuals, there is a probability mass at type t for all $t \in [0,T]$, and hence $f_{\mathrm{I}}(t;u)$ does not exist when $u = t$. This technical issue can be resolved in a standard way by introducing the measure-valued notation $F_{\mathrm{I}}(t;\,\mathrm{d}u)$ and treating the point mass at t separately. Since this modification is entirely straightforward and does not affect any of the subsequent arguments, we suppress it in the notation to streamline the presentation.

Note that (2.7) can be written (in a more involved form) in terms of F and the model parameters defined above, i.e., p_0, q_0, λ, γ, $\mathscr{I}(\cdot)$, $\pi_{\mathrm{SS}}(\cdot)$, $\pi_{\mathrm{SI}}(\cdot,\cdot)$, and $\pi_{\mathrm{II}}(\cdot,\cdot,\cdot)$. Moreover, F can be recovered from the solution to (2.7) by letting, for $t \in [0,T]$,

$$F(t;y) = \begin{cases} 0, & y < -1, \\ p_{\mathrm{S}}(t), & y \in [-1,0), \\ p_{\mathrm{S}}(t) + \int_0^y \mathrm{d}u\, f_{\mathrm{I}}(t;u), & y \in [0,t), \\ p_{\mathrm{S}}(t) + p_{\mathrm{I}}(t), & y \in [t,T+1), \\ 1, & y \geqslant T+1. \end{cases} \tag{2.8}$$

We also need to construct a candidate limit for the graphon valued process $(h^{G_n(t)})_{t\in[0,T]}$ as $n \to \infty$. To this end, we let $\bar{F}$ denote the generalized inverse of F, which is formally defined as

$$\bar{F}(t;x) := \inf\{u : F(t;u) > x\}.$$

For $t \in [0,T]$, let the candidate limit $g^{[F]}(t;\cdot) \in \mathscr{W}$ be given via

$$g^{[F]}(t;x,y) := H(t;\bar{F}(t;x),\bar{F}(t;y),(F(t;\cdot)_{t\in[0,T]})). \tag{2.9}$$

2.2. **Main results.** First we establish that the candidate limit $(F(t;\cdot))_{t\in[0,T]}$ is well-defined. Note that, given (2.9), this also implies that $(g^{[F]}(t;\cdot))_{t\in[0,T]}$ is well-defined.

Theorem 2.1. *There is a unique solution $(F(t;\cdot))_{t\in[0,T]}$ to the system of differential equations characterized by* (2.7) *with initial condition* (2.6). *In particular, for all $t \in [0,T]$, $F(t;\cdot)$ in* (2.8) *follows from (i)*

$$p_{\mathrm{S}}(t) = (1-q_0)\exp\left(-\lambda\int_0^t \mathrm{d}s\,\mathcal{J}(s)\right)$$

and (ii) $f_I(t;u)$, with $u \in [0,t)$, being the unique solution of the fixed-point equation

$$f_{\mathrm{I}}(t;u) = \mathrm{e}^{-t} f_{\mathrm{I}}(0;u-t) + \lambda\int_0^t \mathrm{d}s\,\mathrm{e}^{-(t-s)}\,\mathcal{J}(s)\,p_{\mathrm{S}}(s).$$

Next we establish convergence of the random processes $(F_n(t;\cdot))_{t\in[0,T]}$ and $(h^{G_n(t)})_{t\in[0,T]}$ to their respective candidate limits $(F(t;\cdot))_{t\in[0,T]}$ and $(g^{[F]}(t;\cdot,\cdot))_{t\in[0,T]}$, as $n \to \infty$. To state this result, we let $(\mathcal{M}, d_L)$ denote the space of probability distributions equipped with the Lévy metric (which induces the weak topology; see (A.4)). In the sequel, denote by $D((\mathcal{M},d_L),[0,T])$

(resp. $D((\mathscr{W}, d_{\square}), [0,T])$) the Skorokhod space of $\mathcal{M}$-valued (resp. $\mathscr{W}$-valued) paths on $[0,T]$. Here and in the sequel, '$\Rightarrow$' denotes convergence in distribution.

Theorem 2.2. *As $n \to \infty$, on $D((\mathcal{M}, d_L), [0,T])$,*

$$(F_n(t;\cdot))_{t\in[0,T]} \Rightarrow (F(t;\cdot))_{t\in[0,T]}$$

and, on $D((\mathscr{W}, d_{\square}), [0,T])$,

$$(h^{G_n(t)}(\cdot,\cdot))_{t\in[0,T]} \Rightarrow (g^{[F]}(t;\cdot,\cdot))_{t\in[0,T]}.$$

3. Illustrations

In this section, we illustrate our findings. In the first subsection, we focus on the case with only local feedback, presenting expressions for the basic reproduction number R_0 and the final fraction of susceptible individuals, and discussing the epidemic peak. The second subsection presents a model with global feedback, which shows that multiple infection peaks can arise.

3.1. Dynamics without global feedback.

In this subsection, we consider a simpler setting in which global feedback is absent. More specifically, we make the assumption that

$$\pi_{\mathrm{SS}}(\cdot) = p_0, \quad \text{and} \quad \pi_{\mathrm{SI}}(u,F) = \pi_{\mathrm{SI}}(u). \tag{3.1}$$

Note that $\pi_{\mathrm{SS}}(\cdot) = p_0$ implies that susceptible individuals are connected with probability p_0 *at any time* t, and $\pi_{\mathrm{SI}}(u,F) = \pi_{\mathrm{SI}}(u)$ entails that there is no global co-evolutionary feedback mechanism. Under these assumptions, we can follow standard intuitive calculations to derive compact expressions for the basic reproduction number R_0 and the final proportion of susceptible individuals $p_{\mathrm{S}}(\infty)$ as n becomes large (cf. [8, Section 1]).

Note that, under (3.1), an individual who was infected u time units ago is connected to any susceptible individual with probability

$$p_{\mathrm{SI}}(u) := p_0 \mathrm{e}^{-\gamma u} + \int_0^u \mathrm{d}s\, \gamma \mathrm{e}^{-\gamma s}\, \pi_{\mathrm{SI}}(u-s).$$

The probability that the infected individual infects any susceptible individual throughout their lifetime is then

$$\bar{p}_{\mathrm{SI},n}(u) := \int_0^\infty \mathrm{d}u\, p_{\mathrm{SI}}(u)\, \frac{\lambda}{n} \mathscr{I}(u)\, \mathrm{e}^{-u}.$$

Considering the expected number of infections caused by a single infected individual among a large number (i.e., $n \to \infty$) of susceptible individuals, we then obtain the basic reproduction number

$$R_0 = \int_0^\infty \mathrm{d}u\, p_{\mathrm{SI}}(u)\, \lambda \mathscr{I}(u)\, \mathrm{e}^{-u}. \tag{3.2}$$

To write down a formula for the final value $p_{\mathrm{S}}(\infty)$, note that an initially susceptible individual does not ever get infected with probability $p_{\mathrm{S}}(\infty)/(1-q_0)$, and if it remains susceptible, then it

does not get infected by any of the $n(1-s(\infty))$ individuals that get infected at some point. Consequently

$$\begin{aligned}\frac{p_{\mathrm{S}}(\infty)}{1-q_0} &= \lim_{n\to\infty} (1-\bar{p}_{\mathrm{SI},n}(u))^{n(1-p_s S(\infty))} \\ &= \lim_{n\to\infty} \left(1 - \int_0^\infty \mathrm{d}u\, p_{\mathrm{SI}}(u)\, \frac{\lambda}{n}\, \mathscr{I}(u)\, \mathrm{e}^{-u}\right)^{n(1-p_{\mathrm{S}}(\infty))} \\ &= \lim_{n\to\infty} \left(1 - \frac{R_0(1-p_{\mathrm{S}}(\infty))}{n(1-p_{\mathrm{S}}(\infty))}\right)^{n(1-p_{\mathrm{S}}(\infty))} = \mathrm{e}^{-R_0(1-p_{\mathrm{S}}(\infty))},\end{aligned}$$

which implies that $p_{\mathrm{S}}(\infty)$ is the unique solution in $[0,1]$ of the fixed-point equation

$$p_{\mathrm{S}}(\infty) = (1-q_0)\, \mathrm{e}^{-R_0(1-p_{\mathrm{S}}(\infty))}. \tag{3.3}$$

Let $p_{\mathrm{S},n}(t)$ denote the proportion of susceptible individuals at time t in a system with n individuals. As a consequence of Theorem 2.2, it holds that if $p_{\mathrm{S}}(\infty) = \lim_{t\to\infty} \lim_{n\to\infty} p_{\mathrm{S},n}(t)$, then $p_{\mathrm{S}}(\infty)$ satisfies (3.3). It should be noted, however, that the limiting proportion of susceptible individuals is most naturally expressed as $p_{\mathrm{S}}(\infty) = \lim_{n\to\infty} \lim_{t\to\infty} p_{\mathrm{S},n}(t)$. The required interchange of limits is generally justified by dominating the epidemic process with a subcritical branching process after some large time t (see, for instance, [13, Section 6.3]). Here, we encounter two challenges: (1) the vertex types are continuous, which would likely require working with a dominating branching process with a continuous type space, and (2) the dynamic nature of the edges complicates the construction of such a dominating branching process. Addressing these technical challenges is beyond the scope of the present work.

We are also interested in the *peak* of the epidemic, defined as

$$i_{\max} := \max_{t\geqslant 0} p_{\mathrm{I}}(t).$$

A classical result for the conventional SIR model is that

$$i_{\max} = 1 - R_0^{-1} + R_0^{-1} \log R_0^{-1}. \tag{3.4}$$

In our setting, deriving an analogous formula is challenging (see also [24, Section 3.2] for a related analysis). Nevertheless, the findings below illustrate how γ, which defines the timescale of the edge dynamics, influences $i_{\max}$.

To this end, we first analyze the effect of γ on the basic reproduction number $R_0 \equiv R_0(\gamma)$, assuming $\mathscr{I}(\cdot) \equiv 1$. From (3.2), by direct computations we find

$$R_0(\gamma) = \frac{\lambda}{\gamma+1} p_0 + \lambda\gamma \int_0^\infty \mathrm{d}u\, \mathrm{e}^{-u} \left(\int_0^u \mathrm{d}s\, \mathrm{e}^{-\gamma s}\, \pi_{\mathrm{SI}}(u-s)\right) =: A(\gamma) + B(\gamma);$$

observe that in the static setting (in which $\gamma = 0$) this reduces to $R_0(0) = \lambda p_0$. After applying Fubini–Tonelli and the change of variables $v = u - s$ in the inner integral, we can write

$$B(\gamma) = \lambda\gamma \int_0^\infty \mathrm{d}s\, \mathrm{e}^{-\gamma s} \left(\int_0^\infty \mathrm{d}v\, \mathrm{e}^{-(s+v)}\, \pi_{\mathrm{SI}}(v)\right) = \lambda\gamma\, C \int_0^\infty \mathrm{d}s\, \mathrm{e}^{-(\gamma+1)s} = \frac{\lambda\gamma\, C}{\gamma+1},$$

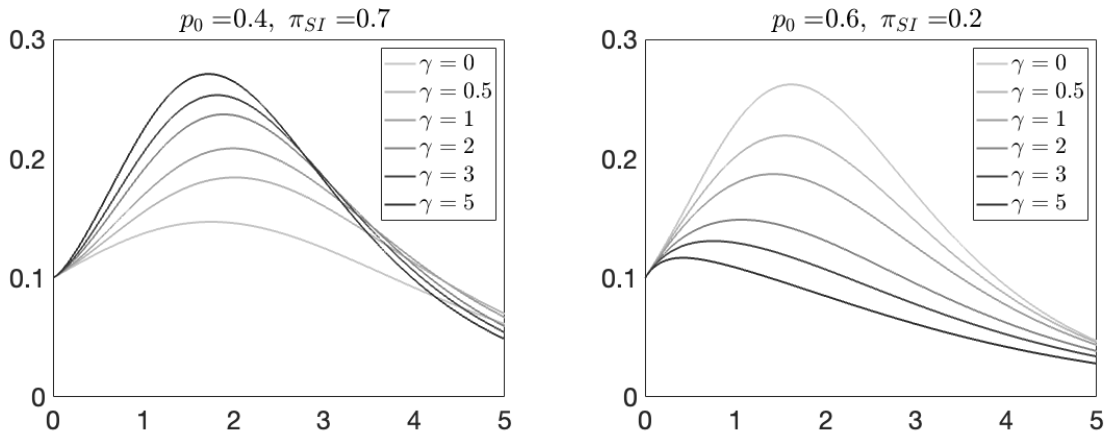

Figure 3.1: Each solid curve represents the numerical solution of the system of PDEs for $\pi_{\mathrm{SS}} = p_0$, $\lambda = 4$, $q_0 = 0.1$, and varying $p_0, \pi_{\mathrm{SI}}, \gamma$.

where

$$C := \int_0^\infty \mathrm{d}v\, \mathrm{e}^{-v} \pi_{\mathrm{SI}}(v).$$

Note that $C < \infty$ because $\pi_{\mathrm{SI}}(\cdot)$ was assumed to be Lipschitz continuous. It thus follows that

$$R_0(\gamma) = \frac{\lambda(p_0 + \gamma C)}{\gamma + 1},$$

and as a consequence

$$R_0'(\gamma) = \frac{\lambda(C - p_0)}{(\gamma + 1)^2}.$$

We therefore conclude that R_0 is increasing in γ if and only if $C > p_0$. Note that through Equation (3.3) we see that $p_{\mathrm{S}}(\infty)$ is monotonically increasing in R_0, and hence $p_{\mathrm{S}}(\infty)$ is monotonically increasing in γ if and only if $C > p_0$. The condition $C > p_0$ is simplest to understand when $\pi_{\mathrm{SI}}(\cdot)$ is constant (say π_{SI}), where it reduces to $\pi_{\mathrm{SI}} > p_0$. In this case, once an individual becomes infected, the probability that they are connected to a susceptible individual, $p_{\mathrm{SI}}(u)$, increases with the time since infection u. If the edge re-sampling rate γ increases, then the increase in $p_{\mathrm{SI}}(u)$ is sped up, leading to an increase in R_0 (see (3.2)). Equivalent reasoning also explains why R_0 is decreasing in γ when $C \equiv \pi_{\mathrm{SI}} < p_0$.

We now turn to the peak of the epidemic, $i_{\max}$. First observe that if the classical result (3.4) were to hold for our model, then $i_{\max}$ would be increasing in R_0. Second, since the number of infected individuals at any given time cannot exceed the total fraction that will eventually be infected, we obtain the upper bound $i_{\max} \leqslant 1 - p_{\mathrm{S}}(\infty)$, where the right-hand side is likewise increasing in R_0. Taken together with the preceding discussion, these observations suggest that $i_{\max}$ is monotonically increasing in γ if and only if $C > p_0$. As we are unable to derive a counterpart of (3.4) for the present model, we cannot establish this claim rigorously. Nevertheless, the conclusion is strongly supported by the numerical results shown in Figure 3.1.

3.2. **Multiple peaks with global feedback.** In this subsection, we present an example that illustrates how global feedback between the epidemic state and network connectivity can produce multiple infection peaks. To capture this behavior, we consider a model in which network connectivity evolves in response to the overall infection level. For $a > 0$, we define

$$\phi(t) := \int_0^a \mathrm{d}s\, f_{\mathrm{I}}(t; s),$$

which can be equivalently written as $\phi(t) = F_n(t; a) - F_n(t; 0)$. Here, $\phi(t)$ denotes the proportion of individuals who were infected at most a time units ago and who are still infectious at time t. It may be interpreted as the population's perceived overall 'threat level'. Since $\phi(t)$ depends on infections that occurred in the past (namely, within the preceding a time units), the perceived threat level generally lags behind the actual threat level. The latter is most naturally quantified by the current rate of infection (see (2.5) for the limiting rate as $n \to \infty$). For $0 < \phi_1 < \phi_2 < \infty$, define the piecewise linear function

$$d(\phi) = \begin{cases} 0.1 & \text{if } \phi \leqslant \phi_1, \\ 0.1 + 0.8\dfrac{\phi - \phi_1}{\phi_2 - \phi_1} & \text{if } \phi_1 < \phi < \phi_2, \\ 0.9 & \text{if } \phi \geqslant \phi_2, \end{cases}$$

which is a behavioral control function, representing how strongly the population is distancing: $d(\phi) \approx 0.9$ (resp. $d(\phi) \approx 0.1$) means that the system is in 'distancing mode' (resp. 'normal mode'). We then put

$$\begin{aligned} \pi_{\mathrm{SS}}(\phi) &:= (1 - d(\phi))\, p_{\mathrm{SS}}^{\mathrm{norm}} + d(\phi)\, p_{\mathrm{SS}}^{\mathrm{dist}}, \\ \pi_{\mathrm{SI}}(\phi) &:= (1 - d(\phi))\, p_{\mathrm{SI}}^{\mathrm{norm}} + d(\phi)\, p_{\mathrm{SI}}^{\mathrm{dist}}, \end{aligned} \tag{3.5}$$

for given numbers $p_{\mathrm{SS}}^{\mathrm{norm}}$, $p_{\mathrm{SS}}^{\mathrm{dist}}$, $p_{\mathrm{SI}}^{\mathrm{norm}}$, and $p_{\mathrm{SI}}^{\mathrm{dist}}$, where it is envisaged that $p_{\mathrm{SS}}^{\mathrm{norm}} > p_{\mathrm{SS}}^{\mathrm{dist}}$ and $p_{\mathrm{SI}}^{\mathrm{norm}} > p_{\mathrm{SI}}^{\mathrm{dist}}$. We have thus constructed a simplified model of 'behavioural response' (e.g., by government interventions or voluntary distancing), enforcing that contact rates drop when people perceive high infection.

In the numerical illustration in Figures 3.2 and 3.3, we have chosen the parameters:

$$\begin{gathered} p_0 = 0.1, \quad q_0 = 0.05, \quad \phi_1 = 0.24, \quad \phi_2 = 0.28, \quad \gamma = 20, \quad \lambda = 10, \\ p_{\mathrm{SS}}^{\mathrm{norm}} = 0.9, \quad p_{\mathrm{SS}}^{\mathrm{dist}} = 0.3, \quad p_{\mathrm{SI}}^{\mathrm{norm}} = 0.6, \quad p_{\mathrm{SI}}^{\mathrm{dist}} = 0.01, \quad \pi_{\mathrm{II}} = 0.3, \quad a = 1. \end{gathered}$$

These figures display the functional law of large numbers for the epidemic states and network that was established in Theorem 2.2, and simulations with finite n. The explanation below focuses on the large population limit, and describes how to interpret the figures. The following phases can be distinguished:

- *Initial rise:* In the top row of Figure 3.2 we observe that the proportion of infected individuals is increasing for $t \in [0, 0.69]$. As displayed in Figure 3.3, the process is in normal mode for the vast majority of this time interval. Consequently, there is a high proportion of active SI edges through which the epidemic can spread, and still a relatively high proportion of susceptible individuals which can be infected. Indeed,

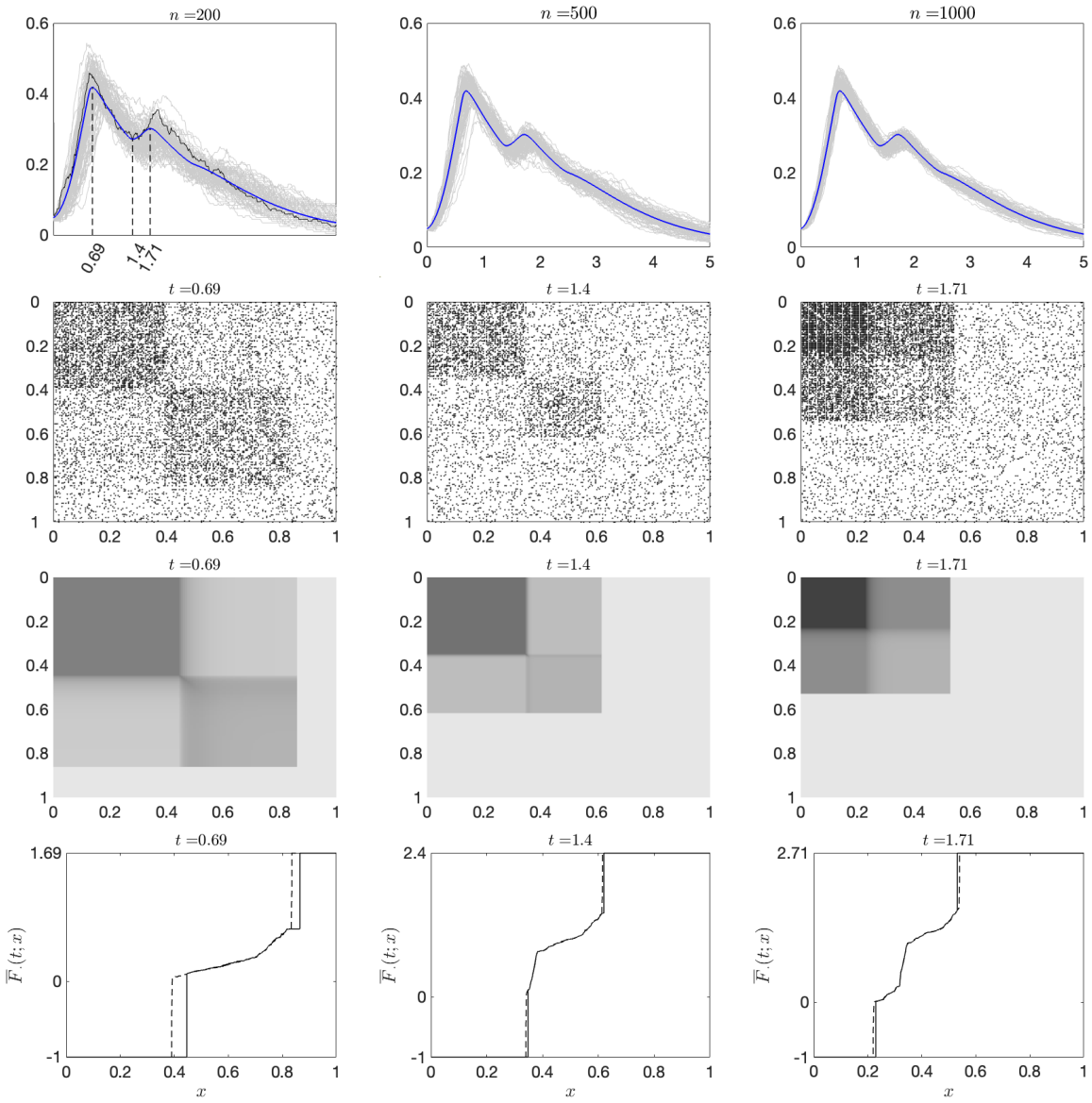

Figure 3.2: The grey trajectories in the top row, from left to right, correspond to simulations of 100 stochastic SIR trajectories of the fraction of infected individuals for $n = 200, 500, 1000$, respectively. The solid blue curves are the numerical solution of the system of PDEs with the same parameters. Observe that the 'cloud' of simulated trajectories shrinks with n, reflecting the convergence to the deterministic limiting path. The second row displays the empirical graphons corresponding to the black trajectory ($n = 200$) when $t = 0.69$ (first peak), $t = 1.4$ (valley between the first two peaks) and $t = 1.71$ (second peak); a dot represents an edge. The labels of the vertices are updated dynamically so that they are ordered lexicographically, i.e., the vertices with state S have lower labels than the vertices with state I, which in turn have lower labels than the vertices with state R, and then by increasing type. The third row displays the corresponding FLLN. The bottom row displays $\bar{F}_{200}(t;x)$ (dashed line) corresponding to the black trajectory and $\bar{F}(t;x)$ (solid line) when $t = 0.69, 1.4, 1.71$ for $x \in [0,1]$.

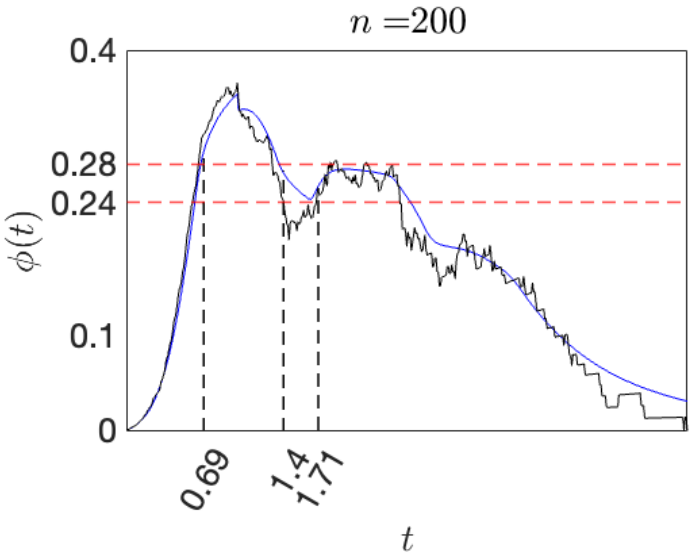

Figure 3.3: An illustration of $\phi(t)$ for $t \in [0, 5]$: the black and the blue lines correspond to the empirical black trajectory and the smooth blue curve, respectively, of the top left panel in Figure 3.2.

the initial proportion of susceptible individuals is $1 - p_0 = 0.9$ and, by observing the bottom left plot of Figure 3.2, we see that at $t = 0.69$ the proportion of susceptible individuals is just over 0.4 (i.e., the proportion of type -1 vertices).

- *First peak:* Just before $t = 0.69$ the process enters distancing mode, and after $t = 0.69$ the proportion of infected individuals starts to decline. In distancing mode, when resampled, the edge probability between SI individuals is $\pi_{\text{SI}}(\phi) = 0.069$, compared to $\pi_{\text{SI}}(\phi) = 0.541$ in normal mode (see Equation (3.5)). The adjustment to distancing mode is not instantaneous, however, since edges are resampled at the finite rate $\gamma = 20$. Consequently, in the left middle two panels in Figure 3.2, where we observe the empirical graphon and limiting graphon at $t = 0.69$, the graph is losing SI edges rapidly. From the bottom left plot of Figure 3.2, we see that the proportions of susceptible and infected individuals are approximately 0.43 and $0.4 = 0.83 - 0.43$ respectively (refer to x-axis); hence these SI edges correspond to $(x, y) \in [0, 0.43] \times [0.43, 0.83] \cup [0.43, 0.83] \times [0, 0.43]$ in the left second and third row of Figure 3.2.
- *Decline from first peak and first dip:* Figure 3.3 demonstrates that even once the proportion of infected individuals starts to decline, the perceived danger level $\phi(t)$ is increasing (i.e., for $t \in [0.69, 1]$). This is because $\phi(t)$ is calculated using past infections, and hence the perceived danger level lags behind the actual threat level, i.e., the current rate of infection. This can also be observed after the first dip ($t = 1.4$), where the perceived danger level continues to decline even though the proportion of infected individuals is increasing, and after the second peak, where the perceived threat level continues to increase even though the proportion of infected individuals is declining. It

is these delayed dynamics that lead to the two infection peaks that we observe in the top row of Figure 3.2.

- *Second peak:* At $t = 1.71$, the proportion of infected individuals reaches its second peak. In the bottom right of Figure 3.2, for $x \in [0.23, 0.56]$, we can observe the types (i.e., the time since infection) of the currently infected individuals. Here we see that there have been two waves of infection: of those currently infected, a proportion of approximately $0.1 = 0.33 - 0.23$ (refer to x-axis) of the total population were infected in the second wave which occurred between 0 and 0.31 time units ago (refer to y-axis), approximately 0.04 were infected during the dip which occurred between 0.31 and 1.02 time units ago, and the remaining $0.19 = 0.56 - 0.37$ were infected during the first wave which occurred between 1.02 and 1.71 time units ago.
- *Final decline:* After $t = 1.71$ the proportion of infected individuals declines monotonically. This is because there are now very few susceptible individuals left to infect. Indeed, the bottom right of Figure 3.2 shows that the proportion of susceptible individuals is just over 0.2 at $t = 1.71$. Moreover, at $t = 1.71$, the proportion of infected individuals is no longer increasing despite the relatively high proportion of active SI edges in comparison to $t = 0.69$ and $t = 1.4$, which can be observed by comparing the plots in the third row of Figure 3.2. The continued decline in the proportion of infected individuals after $t = 1.71$ is therefore because the density of SI edges is never high enough to compensate for the low number of susceptible individuals to infect.

4. Road map

In this section, we outline a four-step procedure leading to the proof of our main results. Each step is presented systematically, as this approach may also serve as a useful framework for analyzing related co-evolutionary systems.

The main idea is that we construct a process, referred to as the *mimicking process*, that exhibits *one-way dependence*: the edge dynamics depend on the states of the vertices, but not vice versa. This structural simplification makes the mimicking process significantly more tractable than the original co-evolutionary model, which features two-way dependence. In the first three steps of the procedure outlined below we prove our FLLN claims for the mimicking process. In the fourth step we show that the two processes remain 'sufficiently close', allowing the FLLN that we establish for the mimicking process to carry over to the co-evolutionary process.

Step 1: Define the mimicking process, and establish convergence of its empirical type process. A key role in our proofs is played by the mimicking process, a process with one-way dependence that shares the same edge dynamics as the co-evolutionary model but differs in its vertex dynamics. It is expressed via the evolving random graph process $(G_n^*(t))_{t\in[0,T]}$ that corresponds to the empirical type process $(F_n^*(t;\cdot))_{t\in[0,T]}$. For this process, we define the generalized type analogously to the original model, now denoted by $X_i^*(t) = (x_i^*(t), y_i^*(t))$. In the mimicking

process at any time $t \in [0,T]$, $x_i^*(t)$ transitions from state S to state I at rate $\lambda \mathcal{J}(t)$, as given in (2.5). The infected and recovered vertices behave as in the original model. We define the mimicking process in full detail in Section 5.2.

We then focus on establishing the convergence of the empirical type process of the mimicking process, $(F_n^*(t;\cdot))_{t\in[0,T]}$ as $n \to \infty$. The following property plays a key role.

Property 4.1. *The empirical type process satisfies a stochastic-process limit: $F_n^* \Rightarrow F$ as $n \to \infty$ on $D(\mathcal{M}([0,T] \cup \{T+1\}), [0,T])$.*

To establish this property in our setting, we characterize F via the generator of the process $(x_i^*(t), y_i^*(t))_{t\in[0,T]}$. It takes the form

$$(\mathcal{L}_t f)(x,y) = \big(\lambda \mathcal{J}(t)\, \mathbf{1}_{\{x=\mathrm{S}\}} + \mathbf{1}_{\{x=\mathrm{I}\}}\big)[f(x',y) - f(x,y)] + b(x,y)\, \frac{\partial}{\partial y} f(x,y), \tag{4.1}$$

where x' is the state of the selected vertex after changing its state x at time t. Here $b(x,y)$ is characterized as follows: (i) $b(x,y) = 0$ for $x \in \{\mathrm{S}, \mathrm{R}\}$, because the type of a susceptible or recovered vertex does not change in time if its state does not, and (ii) $b(\mathrm{I}, y) = 1$. Indeed, if vertex i has state I during the entire time interval $(t, t + \mathrm{d}t)$, then

$$y_i^*(t + \mathrm{d}t) = y_i^*(0) + t + \mathrm{d}t - t_i^{\mathrm{I}} = y_i^*(t) + \mathrm{d}t,$$

so that the time derivative of $y_i^*(t)$ equals 1. This gives rise to the Kolmogorov forward equations in (2.7) that characterizes F, and that verify Property 4.1 above.

Step 2: Express the edge-connection probability, corresponding to the mimicking process, in terms of the types. The goal is develop expressions (i) for the probability that in the mimicking process there is an active edge at time t, in terms of the types of vertices i and j, and (ii) for the path of the empirical distribution $(F_n^*(s;\cdot))_{s\in[0,t]}$ up to time t. In this context, the key property is the following.

Property 4.2. *At any time t, edge ij is active with probability*

$$H(t; y_i^*(t), y_j^*(t), (F_n^*(s;\cdot))_{s\in[0,t]}),$$

conditionally independently of all the other edges given $(y_i^*(s), y_j^*(s), F_n^*(s;\cdot))_{s\in[0,t]}$.

Given the paths of $x_i^*(\cdot)$ and $x_j^*(\cdot)$, the probability that edge ij is active at time t is

$$p_{ij}(t) = \mathrm{e}^{-\gamma t} p_0 + \gamma \int_0^t \mathrm{d}s\, \mathrm{e}^{-\gamma s} \pi_{x_i^*(t-s), x_j^*(t-s)}(\cdot) = \mathrm{e}^{-\gamma t} p_0 + \gamma \int_0^t \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)} \pi_{x_i^*(s), x_j^*(s)}(\cdot).$$

To verify Property 4.2, we need to show that this probability can be expressed only in terms of the types of vertices i and j at time t (i.e., $x_i^*(t)$ and $x_j^*(t) = S$) and $(F_n^*(s;\cdot))_{s\in[0,t]}$. To this end, we first observe that if $x_i^*(t) = x_j^*(t) = \mathrm{S}$, then

$$p_{ij}(t) = \mathrm{e}^{-\gamma t} p_0 + \gamma \int_0^t \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)} \pi_{\mathrm{SS}}(F_n^*(s;\cdot)).$$

Denote $y^*_{ij,\vee}(t) := y^*_i(t) \vee y^*_j(t)$ and $y^*_{ij,\wedge}(t) := y^*_i(t) \wedge y^*_j(t)$. If $x^*_i(t) = x^*_j(t) = \mathrm{I}$, then, using the identities $t^{\mathrm{I}}_i \wedge t^{\mathrm{I}}_j = t - y^*_{ij,\vee}(t)$ and $t^{\mathrm{I}}_i \vee t^{\mathrm{I}}_j = t - y^*_{ij,\wedge}(t)$, we can write

$$\begin{aligned}
p_{ij}(t) &= \mathrm{e}^{-\gamma t} p_0 + \gamma \int_0^{t^{\mathrm{I}}_i \wedge t^{\mathrm{I}}_j} \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)}\, \pi_{\mathrm{SS}}(F^*_n(s;\cdot)) \\
&\quad + \gamma \int_{t^{\mathrm{I}}_i \wedge t^{\mathrm{I}}_j}^{t^{\mathrm{I}}_i \vee t^{\mathrm{I}}_j} \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)}\, \pi_{\mathrm{SI}}(s - (t^{\mathrm{I}}_i \wedge t^{\mathrm{I}}_j), F^*_n(s;\cdot)) \\
&\quad + \gamma \int_{t^{\mathrm{I}}_i \vee t^{\mathrm{I}}_j}^{t} \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)}\, \pi_{\mathrm{II}}(s - (t^{\mathrm{I}}_i \wedge t^{\mathrm{I}}_j), s - (t^{\mathrm{I}}_i \vee t^{\mathrm{I}}_j), F^*_n(s;\cdot)) \\
&= \mathrm{e}^{-\gamma t} p_0 + \gamma \int_0^{t - y^*_{ij,\vee}(t)} \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)}\, \pi_{\mathrm{SS}}(F^*_n(s;\cdot)) \\
&\quad + \gamma \int_{t - y^*_{ij,\vee}(t)}^{t - y^*_{ij,\wedge}(t)} \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)}\, \pi_{\mathrm{SI}}(s - (t - y^*_{ij,\vee}(t)), F^*_n(s;\cdot)) \\
&\quad + \gamma \int_{t - t - y^*_{ij,\wedge}(t)}^{t} \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)}\, \pi_{\mathrm{II}}(s - (t - y^*_{ij,\vee}(t)), s - (t - y^*_{ij,\wedge}(t)), F^*_n(s;\cdot)).
\end{aligned}$$

Similarly, if $x^*_i(t) = \mathrm{I}$ and $x^*_j(t) = \mathrm{S}$, then

$$\begin{aligned}
p_{ij}(t) \quad &= \mathrm{e}^{-\gamma t} p_0 + \gamma \int_0^{t - y^*_i(t)} \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)} \pi_{\mathrm{SS}}(F^*_n(s;\cdot)) \\
&\qquad + \gamma \int_{t - y^*_i(t)}^{t} \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)} \pi_{\mathrm{SI}}(s - (t - y^*_i(t)), F^*_n(s;\cdot)).
\end{aligned}$$

The case $x^*_i(t) = \mathrm{S}$ and $x^*_j(t) = \mathrm{I}$ is fully analogous. Finally, if $x_i(t) = \mathrm{R}$ or $x_j(t) = \mathrm{R}$, then evidently $p_{ij}(t) = p_0$.

In view of the above computations, we can write the probability $p_{ij}(t)$, in terms of $y^*_i(t)$, $y^*_j(t)$ and $(F^*_n(s;\cdot))_{s\in[0,t]}$ only:

$$\begin{aligned}
p_{ij}(t) = {} & \mathbf{1}\{y^*_{ij,\vee}(t) = T+1\} p_0 + \mathbf{1}\{y^*_{ij,\vee}(t) \neq T+1\}\, \mathrm{e}^{-\gamma t} p_0 + \\
& \mathbf{1}\{y^*_{ij,\vee}(t) = -1\}\, \gamma \int_0^t \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)} \pi_{\mathrm{SS}}(F^*_n(s;\cdot)) + \\
& \mathbf{1}\{y^*_{ij,\vee}(t) \in [0, T+1)\}\, \gamma \int_0^{t - y^*_{ij,\vee}(t)} \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)} \pi_{\mathrm{SS}}(F^*_n(s;\cdot)) + \\
& \mathbf{1}\{y^*_{ij,\wedge}(t) = -1\} \mathbf{1}\{y^*_{ij,\vee}(t) \in [0, T+1)\} \\
& \qquad \times \gamma \int_{t - y^*_{ij,\vee}(t)}^{t} \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)} \pi_{\mathrm{SI}}(s - (t - y^*_{ij,\vee}(t)), F^*_n(s;\cdot)) + \\
& \mathbf{1}\{y^*_{ij,\wedge}(t) \in [0, T+1)\} \mathbf{1}\{y^*_{ij,\vee}(t) \in [0, T+1)\} \\
& \qquad \times \gamma \Bigg(\int_{t - (y^*_{ij,\vee}(t))}^{t - (y^*_{ij,\wedge}(t))} \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)} \pi_{\mathrm{SI}}(s - (t - y^*_{ij,\vee}(t)), F^*_n(s;\cdot)) + \\
& \qquad\qquad \int_{t - (y^*_{ij,\vee}(t))}^{t} \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)} \pi_{\mathrm{II}}(s - (t - y^*_{ij,\vee}(t)), s - (t - y^*_{ij,\wedge}(t)), F^*_n(s;\cdot)) \Bigg).
\end{aligned}$$

$$\int_{t-y^*_{ij,\wedge}(t)}^{t} \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)}\pi_{\mathrm{II}}(s-(t-y^*_{ij,\vee}(t)), s-(t-y^*_{ij,\wedge}(t)), F^*_n(s;\cdot))\Bigg).$$

Consequently, the probability of having an active edge between two vertices with type u and v at time t is given by, with $(u,v)^- := u \wedge v$ and $(u,v)^+ := u \vee v$,

$$\begin{aligned}
H(t;u,v,F^*_n(t;\cdot)) = {} & \mathbf{1}\{(u,v)^+ = T+1\}p_0 + \mathbf{1}\{(u,v)^+ \neq T+1\}\,\mathrm{e}^{-\gamma t}p_0 + \\
& \mathbf{1}\{(u,v)^+ = -1\}\,\gamma \int_0^t \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)}\pi_{\mathrm{SS}}(F^*_n(s;\cdot)) + \\
& \mathbf{1}\{(u,v)^+) \in [0,T+1)\}\,\gamma \int_0^{t-(u,v)^+} \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)}\pi_{\mathrm{SS}}(F^*_n(s;\cdot)) \\
& \mathbf{1}\{(u,v)^- = -1\}\mathbf{1}\{(u,v)^+ \in [0,T+1)\} + \\
& \quad \times \gamma \int_{t-(u,v)^+}^{t} \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)}\pi_{\mathrm{SI}}(s-t+(u,v)^+, F^*_n(s;\cdot)) + \\
& \mathbf{1}\{(u,v)^- \in [0,T+1)\}\mathbf{1}\{(u,v)^+ \in [0,T+1)\} \\
& \quad \times \gamma\Bigg(\int_{t-(u,v)^+}^{t-(u,v)^-} \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)}\pi_{\mathrm{SI}}(s-t+(u,v)^+, F^*_n(s;\cdot)) + \\
& \qquad \int_{t-(u,v)^-}^{t} \mathrm{d}s\, \mathrm{e}^{-\gamma(t-s)}\pi_{\mathrm{II}}(s-t+(u,v)^+, s-t+(u,v)^-, F^*_n(s;\cdot))\Bigg). \quad (4.2)
\end{aligned}$$

We have thus verified Property 4.2.

Step 3: Establish convergence of the graphon-valued process that corresponds to the mimicking process. If F is the limit described in Property 4.1, then the induced reference graphon process (see (2.9)) is our candidate limit for the graph-valued process $G^*_n(t)$ in the space of graphons. To show this, we use the continuous mapping theorem, for which we need the following property.

Property 4.3. *The map $F \mapsto g^{[F]}$ from $D(\mathcal{M}([0,T]\cup\{T+1\},[0,T]))$ to $D((\mathcal{W},d_\square),[0,T])$ is continuous, where $\mathcal{W}$ and $d_\square$ are defined in Appendix A.*

Step 4: Extend to the original co-evolutionary model. We apply the framework above to the mimicking process, and couple the mimicking process with the co-evolutionary process in such a way that discrepancies during the time interval $[0,T]$ have a sufficiently small probability. This will imply that $h^{G_n(t)}(\cdot,\cdot) \implies g^{[F]}(t;\cdot,\cdot)$ as $n\to\infty$ in the space $D((\mathscr{W},d_\square),[0,T])$

5. Proofs

5.1. **Proof of Theorem 2.1.** The expression for $p_{\mathrm{S}}(t)$ directly follows from the integration of the first equation in (2.7). We just need to prove that, for any $t\in[0,T]$ and $u\in[-1,T+1]$, there exists a unique solution $f_{\mathrm{I}}(t,u)$ of the second equation in (2.7). To this end, we first derive the fixed-point equation $f_{\mathrm{I}}(t,u)$ must satisfy by using the method of characteristics, and then we use the Banach fixed-point theorem to prove existence and uniqueness of the solution.

Method of characteristics. The characteristics satisfy $\mathrm{d}u/\mathrm{d}t = 1$, meaning that $u(t) = u(0) + t$. Along this curve the PDE becomes an ODE for the profile $F(t) = f_{\mathrm{I}}(t, u(t))$, which reads as

$$F'(t) + F(t) = \lambda\, \mathcal{J}(t)\, p_{\mathrm{S}}(t),$$

which is equivalent to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{t} F(t)\right) = \lambda \mathrm{e}^{t} \mathcal{J}(t)\, p_{\mathrm{S}}(t).$$

This thereby leads to

$$F(t) = \mathrm{e}^{-t} F(0) + \lambda \int_0^t \mathrm{d}s\, \mathrm{e}^{-(t-s)} \mathcal{J}(s)\, p_{\mathrm{S}}(s),$$

which reads as

$$f_{\mathrm{I}}(t, u) = \mathrm{e}^{-t} f_{\mathrm{I}}(0, u - t) + \lambda \int_0^t \mathrm{d}s\, \mathrm{e}^{-(t-s)} \mathcal{J}(s)\, p_{\mathrm{S}}(s).$$

Note that the solution is not explicit, but it depends on solving a coupled system along all the characteristics. Indeed, the term $\mathcal{J}(s)$ depends on the function $f_{\mathrm{I}}(s, \cdot)$ itself.

Contraction property. Define the space $X = C([0, T] \times [-1, T + 1], \mathbb{R})$ with the norm

$$\|f\|_\infty := \sup_{t \in [0,T],\, u \in [-1, T+1]} |f(t, u)|, \qquad f \in X.$$

Define now the operator $\mathcal{T} : X \to X$ as

$$(\mathcal{T} f)(t, u) := \mathrm{e}^{-t} f(0, u - t) + \lambda \int_0^t \mathrm{d}s\, \mathrm{e}^{-(t-s)} \mathcal{J}_f(s)\, p_{\mathrm{S}}(s),$$

where $\mathcal{J}_f(s) := \int_0^s \mathrm{d}F_f(s; u)\, H(s; -1, u, F_f(s; \cdot))\, \mathscr{I}(u)$, with F_f defined as in (2.8) after replacing f_{I} by f (note that also $p_{\mathrm{I}}(t)$ is obtained after this replacement in (2.4)). A solution f_{I} of the starting PDE corresponds then to a fixed point of $f = \mathcal{T} f$. Letting $f, g \in X$, we find, by using that $\mathcal{J}(\cdot)$ is Lipschitz continuous with Lipschitz constant L,

$$\begin{aligned}
\|(\mathcal{T} f)(t, u) - (\mathcal{T} g)(t, u)\|_\infty &= \sup_{t \in [0,T]} \left| \lambda \int_0^t \mathrm{d}s\, \mathrm{e}^{-(t-s)} p_{\mathrm{S}}(s)\, (\mathcal{J}_f(s) - \mathcal{J}_g(s)) \right| \\
&\leqslant \lambda L \|f - g\|_\infty \sup_{t \in [0,T]} \int_0^t \mathrm{d}s\, s \mathrm{e}^{-(t-s)} \\
&\leqslant \lambda L \|f - g\|_\infty \sup_{t \in [0,T]} \int_0^t \mathrm{d}s\, s, \\
&= K(T) \|f - g\|_\infty,
\end{aligned}$$

with $K(T) := \frac{1}{2} \lambda L T^2$, showing that $\mathcal{T}$ is a contraction for small enough T.

Local existence and uniqueness. First, note that $K(T) > 0$. Moreover, if $K(T) < 1$, then $\mathcal{T}$ is a contraction for any $t \in [0, T]$, therefore leading to global existence and uniqueness by the Banach fixed-point theorem. Otherwise, by the same theorem we deduce that there exists a unique solution on $[0, T_0] \times [-1, T + 1]$, with T_0 the unique positive value such that $K(T_0) = 1$.

Global existence. After obtaining the solution on $[0, T_0]$, we use the value at $t = T_0$ as new initial data and repeat the same argument. Since $T < \infty$ and the norm of the solution can be bounded uniformly in t and u as

$$\sup_{t\in[0,T]} \sup_{u\in[-1,T+1]} |f_{\mathrm{I}}(t,u)| \leqslant \sup_{u\in[-1,T+1]} |f_{\mathrm{I}}(0,u)| + \lambda K(T) < \infty,$$

we can iterate finitely many times to cover all $t \in [0, T]$. This yields global existence and uniqueness for (t, u) in the whole rectangle $[0, T] \times [-1, T + 1]$.

5.2. **Coupling.** To prove convergence of the process in the space of graphons, we construct a mimicking process following Steps 1–3 in Section 4 and that is close to the original process in L_1 (Step 4).

Mimicking process. Suppose that the process $(G_n^*(t))_{t\in[0,T]}$ is characterised by the following dynamics:

- $G_n^*(0)$ is an ERRG with connection probability p_0.
- Each vertex i being susceptible (resp. infected) is equipped with an independent rate-$\lambda\mathcal{J}^*(t)$ (resp. rate-1) Poisson clock, with $\mathcal{J}^*(t) := \mathcal{I}(t; f_{\mathrm{I}}^*(t,\cdot), (F^*(t,\cdot))_{t\in[0,T]})$, where $f_{\mathrm{I}}^*(t,\cdot) = \mathbb{P}(X_i^*(t) \in (\mathrm{I},\cdot))$ and $(F^*(t,\cdot))_{t\in[0,T]}$ is the associated limiting empirical type process. When the clock associated to vertex i rings, the following happens:
  - if $x_i(t) = \mathrm{S}$, then vertex i gets infected;
  - if $x_i(t) = \mathrm{I}$, then vertex i recovers and all the adjacent edges to vertex i are resampled according to the initial edge probability p_0.
- Each edge (not involving any vertex having state R) is equipped with an independent rate-γ Poisson clock. When the clock associated to edge ij rings, then the edge is active with a probability that depends on the state of the connecting vertices. This probability is defined as follows:
  - $\pi_{\mathrm{SS}}(F^*(t,\cdot))$ if both adjacent vertices are susceptible in $G_n^*(t)$;
  - $\pi_{\mathrm{SI}}(y_i^*(t), F^*(t,\cdot))$ (resp. $\pi_{\mathrm{SI}}(y_j^*(t), F^*(t,\cdot))$) if vertex i (resp. vertex j) is infected and vertex j (resp. vertex i) is susceptible, where $y_i^*(t)$ denotes the type of vertex i in $G_n^*(t)$ for any $i \in [n]$;
  - $\pi_{\mathrm{II}}(y_i^*(t), y_j^*(t), F^*(t,\cdot))$ if both adjacent vertices are infected in $G_n^*(t)$.

Lemma 5.1. *The process* $\{G_n^*(t))_{t\geqslant 0}\}_{n\in\mathbb{N}}$ *satisfies the assumptions of* [15, Theorem 3.10]*, and therefore* $h^{G_n^*} \Rightarrow g^{[F^*]}$ *as* $n \to \infty$ *in the space* $D((\mathcal{W}, d_{\square}), [0, T])$.

Proof. To prove the claim, it suffices to verify that [15, Assumptions 3.1, 3.6–3.7, 3.9] hold. The dynamics of each vertex in the mimicking process are independent of those of the other vertices. Consequently, the first three properties discussed in Section 4 are satisfied. We may therefore follow the same reasoning as in [3, Section 2.4.1] to establish a FLLN for the empirical type process as $n \to \infty$. □

Lemma 5.2. *There exists a coupling of* $\{(G_n(t))_{t\in[0,T]}\}_{n\in\mathbb{N}}$ *and* $\{(G_n^*(t))_{t\in[0,T]}\}_{n\in\mathbb{N}}$ *such that, for any* $\delta > 0$*,*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(||\tilde{h}^{G_n(t)} - \tilde{h}^{G_n^*(t)}||_{L_1} > \delta \text{ for some } t \in [0,T] \right) = 0. \tag{5.1}$$

Proof. The claim is proved in three steps. In Step I, we construct a coupling between the original and the mimicking processes. The key idea is to associate susceptible vertices with Poisson clocks whose rates differ between the two processes, and to couple these clocks by simulating a Poisson process of intensity λ up to time t and retaining each point in $G_n(t)$ and $G_n^*(t)$ with appropriately chosen probabilities, specified below. The first retained point determines the infection time of the vertex in the corresponding process; if no point is retained up to time t, the vertex remains susceptible. The same construction applies to the edge dynamics. A *difference* arises when a vertex (resp. an edge) is infected (resp. active) in one process but susceptible (resp. inactive) in the other. Step II characterizes the probability of such differences, and Step III provides precise estimates.

Step I: description of the coupling. Suppose that $(G_n(t))_{t\geqslant 0}$ and $(G_n^*(t))_{t\geqslant 0}$ are generated in the following manner. Recall that $\mathcal{N}_i^{\mathrm{I}}(t)$ is the set of infected neighbors of vertex i at time t.

- Each initially infected vertex $i \in [n]$ is assigned the same (coupled) rate-1 Poisson clock in both processes. When the clock associated with vertex i rings, vertex i recovers and all the adjacent edges to vertex i are resampled according to the initial edge probability p_0 in both $G_n(t)$ and $G_n^*(t)$.
- Each initially susceptible vertex $i \in [n]$ is assigned the same (coupled) rate-λ Poisson clock in both processes. When the clock associated to vertex i rings, generate an outcome U drawn from Unif$(0,1)$ distribution.
  - If $U \leqslant \sum_{j\in\mathcal{N}_i^{\mathrm{I}}(t)} \mathscr{I}(y_j)/n$, then vertex i becomes infected, otherwise remains susceptible, in $G_n(t)$.
  - If $U \leqslant \mathcal{J}^*(t)$, then vertex i becomes infected, otherwise remains susceptible, in $G_n^*(t)$.
- Assign each edge the same (coupled) rate-γ Poisson clock in both processes. When the Poisson clock associated with edge ij rings, generate an outcome U drawn from Unif$(0,1)$ distribution.
  - When $x_i(t) = x_j(t) = \mathrm{S}$, if $U \leqslant \pi_{\mathrm{SS}}(F_n(t,\cdot))$, then edge ij is active, otherwise it is inactive, in $G_n(t)$.
  - When either $x_i(t) = \mathrm{S}$ and $x_j(t) = \mathrm{I}$, or $x_i(t) = \mathrm{I}$ and $x_j(t) = \mathrm{S}$, if $U \leqslant \pi_{\mathrm{SI}}(y_j(t), F_n(t,\cdot))$ or $U \leqslant \pi_{\mathrm{SI}}(y_i(t), F_n(t,\cdot))$, respectively, then edge ij is active, otherwise it is inactive, in $G_n(t)$.
  - When $x_i(t) = x_j(t) = \mathrm{I}$, if $U \leqslant \pi_{\mathrm{II}}(y_i(t), y_j(t), F_n(t,\cdot))$, then edge ij is active, otherwise it is inactive, in $G_n(t)$.
  - When $x_i^*(t) = x_j^*(t) = \mathrm{S}$, if $U \leqslant \pi_{\mathrm{SS}}(F^*(t,\cdot))$, then edge ij is active, otherwise it is inactive, in $G_n^*(t)$.

– When either $x_i^*(t) = \mathrm{S}$ and $x_j^*(t) = \mathrm{I}$, or $x_i^*(t) = \mathrm{I}$ and $x_j^*(t) = \mathrm{S}$, if $U \leqslant \pi_{\mathrm{SI}}(y_j^*(t), F^*(t,\cdot))$ or $U \leqslant \pi_{\mathrm{SI}}(y_i^*(t), F^*(t,\cdot))$, respectively, then edge ij is active, otherwise it is inactive, in $G_n^*(t)$.
– When $x_i^*(t) = x_j^*(t) = \mathrm{I}$, if $U \leqslant \pi_{\mathrm{II}}(y_i^*(t), y_j^*(t), F^*(t,\cdot))$, then edge ij is active, otherwise it is inactive, in $G_n^*(t)$.

Step II: majorization of the L_1 distance. If vertex i is susceptible in both models and the clock associated with vertex i rings, then a *difference* is formed when vertex i is susceptible in one process and infected in the other. This happens with probability

$$\left| \sum_{j \in \mathcal{N}_i^{\mathrm{I}}(t)} \frac{\mathscr{I}(y_j)}{n} - \mathcal{J}^*(t) \right|.$$

If edge ij has the same state (active or inactive) in both models and the clock associated with edge ij rings, then a difference is formed with probability depending on the generalised types $X_i(t), X_j(t), X_i^*(t), X_j^*(t)$, and on $F_n(t,\cdot), F^*(t,\cdot)$. For instance, if $x_i(t) = x_i^*(t) = \mathrm{S}$ and $x_j(t) = x_j^*(t) = \mathrm{I}$, this probability is equal to

$$|\pi_{\mathrm{SI}}(y_j(t), F_n(t,\cdot)) - \pi_{\mathrm{SI}}(y_j^*(t), F^*(t,\cdot))|.$$

Step III: bounding the dominating process. Our approach is to show that, over a small time window $[t, t+\Delta)$ with $\Delta > 0$, the number of differences is stochastically dominated by a random variable with suitable properties. Let

- $t_\Delta = [t, t+\Delta)$ be the time window,
- $N_E \sim \mathrm{Bin}(\binom{n}{2}, 1 - \mathrm{e}^{-\gamma\Delta})$ be the total number of edge clocks that ring in t_Δ,
- $N_V \sim \mathrm{Bin}(\mathcal{N}_{\mathrm{S}}(t), 1 - \mathrm{e}^{-\lambda\Delta})$ be the total number of clocks associated with susceptible vertices in both processes at time t that ring in t_Δ, where $\mathcal{N}_{\mathrm{S}}(t)$ is the number of susceptible vertices at time t,
- $d_E(t)$ be the total number of differences in the edges between $G_n(t)$ and $G_n^*(t)$,
- $d_V(t)$ be the total number of differences in vertex states between $F_n(t;\cdot)$ and $F_n^*(t;\cdot)$, where F_n is defined in (2.2) and F_n^* is defined analogously using the graph $G_n^*(t)$. Here, by $F_n(t;\cdot)$ and $F_n^*(t;\cdot)$ we mean the complete information on vertex types and states.

Given $\Omega \equiv (N_E, N_V, d_E(t), d_V(t), F_n(t;\cdot), F_n^*(t;\cdot))$, we want to bound the probability $\Pi_n(\Delta)$ that, when an edge clock (of an edge chosen uniformly at random) rings, a difference is formed during any time $s \in t_\Delta$. Note that if the clock of a vertex pair ij rings, then a difference can only be formed when either $x_i(s) \neq x_i^*(s)$ or $x_j(s) \neq x_j^*(s)$, or there is a difference in the types or in the empirical type distributions $F_n(s;\cdot)$ and $F^*(s;\cdot)$. At any time t, we can then partition the set of all the vertex pairs as those connecting at least one vertex with a difference of state between the two models, and those which connect both vertices having the same state in the two models, which we refer to as $E_1(t)$ and $E_2(t)$, respectively. Since each vertex i is part of n vertex pairs, note that

$$|E_1(t)| \leqslant n\,(d_V(t) + N_V).$$

Observing that $d_V(t) = \sum_{i=1}^{n} \mathbf{1}\{x_i(t) \neq x_i^*(t)\}$, considering the worst case scenario that every time a vertex clock rings during t_Δ a new difference is formed, and observing that the type of two infected vertices can differ of at most Δ during t_Δ, we have the following upper bound:

$$\begin{aligned}\Pi_n(\Delta) &\leqslant \frac{|E_1(t)|}{\binom{n}{2}} + L\left(2\Delta + d_L(F_n(s;\cdot), F^*(s;\cdot))\right) \\ &\leqslant \frac{n(d_V(t)+N_V)}{\binom{n}{2}} + L\left(2\Delta + d_L(F_n(s;\cdot), F^*(s;\cdot))\right).\end{aligned} \tag{5.2}$$

where we have used the Lipschitz continuity of the functions $\pi_{\mathrm{SS}}, \pi_{\mathrm{SI}}, \pi_{\mathrm{II}}$, with the constant L defined as $\max\{L_{\mathrm{SS}}, L_{\mathrm{SI}}, L_{\mathrm{II}}\} < \infty$. To control this probability, using the triangular inequality for the Lévy metric (see Appendix A), we can write

$$d_L(F_n(s;\cdot), F^*(s;\cdot)) \leqslant \underbrace{d_L(F_n(s;\cdot), F_n^*(s;\cdot))}_{(a)} + \underbrace{d_L(F_n^*(s;\cdot), F^*(s;\cdot))}_{(b)}.$$

Considering the worst case scenario as explained above, for any $s \in t_\Delta$, the term (a) can be bounded as

$$d_L(F_n(s,\cdot), F_n^*(s,\cdot)) \leqslant \frac{d_V(t)+N_V}{n},$$

while the term (b) tends to zero as $n \to \infty$ directly from the Glivenko-Cantelli theorem, because in the mimicking process vertices behave independently and therefore $F_n^*(s;\cdot)$ converges in distribution to $F^*(s;\cdot)$. Thus, (5.2) becomes

$$\begin{aligned}\Pi_n(\Delta) &\leqslant \frac{n(d_V(t)+N_V)}{\binom{n}{2}} + L\left(2\Delta + \frac{d_V(t)+N_V}{n} + o(n)\right) \\ &\leqslant \left(1+\frac{L}{2}\right)\frac{n(d_V(t)+N_V)}{\binom{n}{2}} + 2L\Delta + o(n).\end{aligned}$$

Because this bound is uniform in t_Δ and does not depend on N_E, given Ω, and the term $o(n)$ can be ignored in the limit as $n \to \infty$, we therefore have

$$d_E(t+\Delta) - d_E(t) \overset{\mathrm{st}}{\leqslant} \mathrm{Bin}\left(N_E, \left(1+\frac{L}{2}\right)\frac{n(d_V(t)+N_V)}{\binom{n}{2}} + 2L\Delta\right),$$

where $\overset{\mathrm{st}}{\leqslant}$ means 'stochastically dominated by'.

Given Ω, we next want to bound the probability that, when a vertex clock (of a susceptible vertex in both processes at time t, chosen uniformly at random, that is) rings, a difference is formed during any time $s \in t_\Delta$. In other words, we wish to establish an upper bound on

$$\frac{1}{n}\sum_{i=1}^{n}\left|\sum_{j\in\mathcal{N}_i^{\mathrm{I}}(s)}\frac{\mathscr{I}(y_j)}{n} - \mathcal{J}^*(s)\right|,$$

which applies for all $s \in t_\Delta$. Applying the triangular inequality, we have

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\Bigg|\sum_{j\in\mathcal{N}_i^{\mathrm{I}}(s)}\frac{\mathscr{I}(y_j)}{n}-\mathcal{J}^*(s)\Bigg| \\ \leqslant \underbrace{\frac{1}{n}\sum_{i=1}^{n}\Bigg|\sum_{j\in\mathcal{N}_i^{\mathrm{I}}(s)}\frac{\mathscr{I}(y_j)}{n}-\sum_{j\in\mathcal{N}_i^{*\mathrm{I}}(s)}\frac{\mathscr{I}(y_j^*)}{n}\Bigg|}_{(i)}+\underbrace{\frac{1}{n}\sum_{i=1}^{n}\Bigg|\sum_{j\in\mathcal{N}_i^{*\mathrm{I}}(s)}\frac{\mathscr{I}(y_j^*)}{n}-\mathcal{J}^*(s)\Bigg|}_{(ii)},\end{aligned}$$

where $\mathcal{N}_i^{*\mathrm{I}}(s)$ (evidently) denotes the set of infected neighbours of vertex i in $G_n^*(s)$. We first bound (i) as follows:

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\Bigg|\sum_{j\in\mathcal{N}_i^{\mathrm{I}}(s)}\frac{\mathscr{I}(y_j)}{n}-\sum_{j\in\mathcal{N}_i^{*\mathrm{I}}(s)}\frac{\mathscr{I}(y_j^*)}{n}\Bigg| &= \frac{1}{n^2}\sum_{i=1}^{n}\Bigg|\sum_{j\in\mathcal{N}_i^{\mathrm{I}}(s)}\mathscr{I}(y_j)-\sum_{j\in\mathcal{N}_i^{*\mathrm{I}}(s)}\mathscr{I}(y_j^*)\Bigg| \\ &\leqslant \frac{1}{n^2}(d_E(t)+N_E),\end{aligned}$$

where we used that $\sup_{u\in[0,t+\Delta]}\mathscr{I}(u)=1$.

We next bound (ii). To this end, we use the convergence of the mimicking process in the space of graphons. Let $g_n^*(s;u,v)=h^{G_n^*}(s;u,v)$ denote the empirical graphon at time s. In addition, let $r_n^{\mathrm{S+I}}(s)$ and $r^{\mathrm{S+I}}(s)$ denote the sum of the proportions of susceptible and infected vertices in the empirical and limiting graphon, respectively, at time s. Similarly, $r_n^{\mathrm{S}}(s)$ and $r^{\mathrm{S}}(s)$ denote the corresponding quantities for susceptible vertices. Thus, we can write

$$\frac{1}{n}\sum_{j\in\mathcal{N}_i^{*\mathrm{I}}(s)}\mathscr{I}(y_j^*)=\int_{r_n^{\mathrm{S}}(s)}^{r_n^{\mathrm{S+I}}(s)}\mathrm{d}u\,g_n^*\Big(s;\frac{i}{n},u\Big)\mathscr{I}(F^*(s;u)),$$

using the definition of the empirical graphon. In addition, by Lemma A.1 we have

$$\mathcal{J}^*(s)=\int_{r^{\mathrm{S}}(s)}^{r^{\mathrm{S+I}}(s)}\mathrm{d}y\,g^{[F^*]}(s;y,0)\mathscr{I}(F^*(s;y)).$$

We can therefore write

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\Bigg|\sum_{j\in\mathcal{N}_i^{*\mathrm{I}}(s)}\frac{\mathscr{I}(y_j^*)}{n}-\mathcal{J}^*(s)\Bigg| &= \frac{1}{n}\sum_{i=1}^{n}\Bigg|\int_{r_n^{\mathrm{S}}(s)}^{r_n^{\mathrm{S+I}}(s)}\mathrm{d}u\,g_n^*\Big(s;\frac{i}{n},u\Big)\mathscr{I}(F^*(s;u)) \\ &\qquad -\int_{r^{\mathrm{S}}(s)}^{r^{\mathrm{S+I}}(s)}\mathrm{d}y\,g^{[F^*]}(s;y,0)\mathscr{I}(F^*(s;y))\Bigg| \\ &= \Bigg|\int_{r_n^{\mathrm{S}}(s)}^{r_n^{\mathrm{S+I}}(s)}\mathrm{d}u\,g_n^*(s;0,u)\mathscr{I}(F^*(s;u)) \\ &\qquad -\int_{r^{\mathrm{S}}(s)}^{r^{\mathrm{S+I}}(s)}\mathrm{d}y\,g^{[F^*]}(s;y,0)\mathscr{I}(F^*(s;y))\Bigg| \\ &\leqslant \big|(r_n^{\mathrm{S+I}}(s)-r^{\mathrm{S+I}}(s))+(r_n^{\mathrm{S}}(s)-r^{\mathrm{S}}(s))\big|\end{aligned}$$

$$+ d_{\square}\left(g_n^*(s;0,\cdot), g^{[F^*]}(s;0,\cdot)\right).$$

Note that, by the FLLN claimed in Lemma 5.1), as $n \to \infty$ each of these terms tends to zero uniformly in s with high probability. Thus, we conclude that

$$d_V(t+\Delta) - d_V(t) \overset{\text{st}}{\leqslant} \text{Bin}\left(N_V, \frac{1}{n^2}(d_E(t) + N_E) + \epsilon(n)\right),$$

where $\epsilon(n)$ is a function that can be chosen arbitrarily small with high probability, and can therefore effectively ignored in the limit as $n \to \infty$.

Our plan is now to bound the number of differences formed in the vertices and edges between the two models by considering the time intervals $[0,\Delta], [\Delta, 2\Delta], ..., [T-\Delta, T]$ in sequence, and then use the union bound. Without loss of generality, we are tacitly assuming that T is a multiple of Δ. We use the results above which, for n large, can be summarised as follows:

$$\begin{aligned}
N_E &\overset{\text{st}}{\leqslant} \text{Bin}\left(\binom{n}{2}, \lambda\Delta\right),\\
N_V &\overset{\text{st}}{\leqslant} \text{Bin}\left(n, \gamma\Delta\right),\\
d_E(t+\Delta) - d_E(t) &\overset{\text{st}}{\leqslant} \text{Bin}\left(N_E, \left(1+\frac{L}{2}\right)\frac{n(d_V(t)+N_V)}{\binom{n}{2}} + 2L\Delta\right),\\
d_V(t+\Delta) - d_V(t) &\overset{\text{st}}{\leqslant} \text{Bin}\left(N_V, \frac{d_E(t)+N_E}{n^2}\right).
\end{aligned}$$

The proof is then completed by following arguments analogous to those used in the proof of [3, Lemma 3.6]. □

5.3. **Proof of Theorem 2.2.** Theorem 2.2 now follows from Lemmas 5.1–5.2 and [15, Theorem 3.10].

Appendix A. Graphons, Lévy metric and an additional lemma

Graphons. Let $\mathcal{W}$ be the space of functions $h\colon [0,1]^2 \to [0,1]$ such that $h(x,y) = h(y,x)$ for all $(x,y) \in [0,1]^2$, endowed with the *cut distance*

$$d_{\square}(h_1,h_2) := \sup_{S,T\subseteq[0,1]} \left|\int_{S\times T} \mathrm{d}x\,\mathrm{d}y\,[h_1(x,y) - h_2(x,y)]\right|, \quad h_1,h_2 \in \mathcal{W}. \tag{A.1}$$

On $\mathcal{W}$, called the space of graphons, there is a natural equivalence relation $\sim$. Let Σ be the space of measure-preserving bijections $\sigma\colon [0,1] \to [0,1]$. Then $h_1(x,y) \sim h_2(x,y)$ if $\delta_{\square}(h_1,h_2) = 0$, where $\delta_{\square}$ is the *cut metric* defined by

$$\delta_{\square}(\tilde{h}_1,\tilde{h}_2) := \inf_{\sigma_1,\sigma_2\in\Sigma} d_{\square}(h_1^{\sigma_1}, h_2^{\sigma_2}), \qquad \tilde{h}_1,\tilde{h}_2 \in \widetilde{\mathcal{W}}, \tag{A.2}$$

with $h^{\sigma}(x,y) = h(\sigma x, \sigma y)$. This equivalence relation yields the quotient space $(\widetilde{\mathcal{W}}, \delta_{\square})$, which is compact.

A finite simple undirected graph G on n vertices can be represented as a graphon $h^G \in \mathcal{W}$ by setting

$$h^G(x,y) := \begin{cases} 1 & \text{if there is an edge between vertex } \lceil nx \rceil \text{ and vertex } \lceil ny \rceil, \\ 0 & \text{otherwise,} \end{cases} \tag{A.3}$$

which referred to as the *empirical graphon* associated with G, and has a block structure.

Lévy metric. We equip the space $\mathcal{M}(\mathbb{R}_+)$ with the *Lévy metric*, so that for two distribution functions F, G we have

$$d_L(F,G) := \inf\{\epsilon > 0 : F(x-\epsilon) - \epsilon \leqslant G(x) \leqslant F(x+\epsilon) + \epsilon, \forall\ x \in \mathbb{R}\} \tag{A.4}$$

An important property of the Lévy metric is that it satisfies the triangle inequality.

Lemma A.1. *For any $s \in t_\Delta$,*

$$\mathcal{J}^*(s) = \int_{r^{\mathrm{S}}(s)}^{r^{\mathrm{S+I}}(s)} \mathrm{d}y\, g^{[F^*]}(s;y,0)\mathscr{I}(F^*(s;y)).$$

Proof. By definition, we have

$$\mathcal{J}^*(s) = \mathcal{I}(s; f_{\mathrm{I}}^*(s,\cdot), F^*(s,\cdot)) = \int_0^T \mathrm{d}u\, f_{\mathrm{I}}^*(s,u) H(s;-1,u,F^*(s,\cdot))\mathscr{I}(u).$$

By using the change of variable $u = F^*(s;y)$, we obtain

$$\begin{aligned} \mathcal{J}^*(s) &= \int_{r^{\mathrm{S}}(s)}^{r^{\mathrm{S+I}}(s)} \mathrm{d}y\, H\big(s;-1,F^*(s;y),F^*(s,\cdot)\big)\mathscr{I}(F^*(s;y)) \\ &= \int_{r^{\mathrm{S}}(s)}^{r^{\mathrm{S+I}}(s)} \mathrm{d}y\, g^{[F^*]}(s;y,0)\mathscr{I}(F^*(s;y)). \end{aligned}$$

where in the first equality we used the fact that $\mathrm{d}y = f_{\mathrm{I}}^*(s, F^*(s;y))\,\mathrm{d}u$. This follows from the derivative of the inverse function, which is well defined because the function $F^*(s;\cdot)$ is right-continuous and non-decreasing. Thus, the derivative is zero only in the intervals in which the function is constant, meaning that there are no infected individuals with those types, and therefore we may think the integral not running on these values. Finally, note that the integral goes from $r^{\mathrm{S}}(s)$ to $r^{\mathrm{S+I}}(s)$ because otherwise $\mathscr{I}(\cdot) \equiv 0$ and therefore the integrand is zero. □

References

[1] S. Athreya, F. den Hollander, A. Röllin. Co-evolving vertex and edge dynamics in dense graphs. Preprint at arXiv:2504.06493 (2025).

[2] L. Avena, R. Baldasso, R.S. Hazra, F. den Hollander, M. Quattropani. The voter model on random regular graphs with random rewiring. Preprint at arXiv:2501.08703 (2025).

[3] S. Baldassarri, P. Braunsteins, F. den Hollander, M. Mandjes. Opinion dynamics on dense dynamic random graphs. Preprint at arXiv:2410.14618 (2024).

[4] S. Baldassarri, N. Kriukov. Functional central limit theorem for the subgraph count of the voter model on dynamic random graphs. To appear in *Bernoulli*. Preprint at arXiv:2503.11541 (2025).

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

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

[7] F. Ball, T. Britton, P. Trapman. An epidemic in a dynamic population with importation of infectives. *The Annals of Applied Probability* 27 (2017) 242–274.

[8] A. Barbour, G. Reinert. Approximating the epidemic curve. *Electronic Journal of Probability* 18 (2013) 1–30.

[9] R. Basu, A. Sly. Evolving voter model on dense random graphs. *The Annals of Applied Probability* 27 (2017) 1235–1288.

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

[11] C. Borgs, J. Chayes, L. Lovász, V. Sós, K. Vesztergombi, K. Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing. *Advances in Mathematics* 219 (2008) 1801–1851.

[12] C. Borgs, J. Chayes, L. Lovász, V. Sós, K. Vesztergombi, K. Convergent sequences of dense graphs II: Multiway cuts and statistical physics. *Annals of Mathematics* 176 (2012) 151–219.

[13] C. Borgs, K. Huang, C. Ikeokwu. A Law of Large Numbers for SIR on the Stochastic Block Model: A Proof via Herd Immunity. Preprint at arXiv:2410.07097 (2024).

[14] P. Braunsteins, F. den Hollander, M. Mandjes. A sample-path large deviation principle for dynamic Erdős–Rényi random graphs. *The Annals of Applied Probability* 33 (2023) 3278–3320.

[15] P. Braunsteins, F. den Hollander, M. Mandjes. Graphon-valued processes with vertex-level fluctuations. Preprint at arXiv:2209.01544 (2022).

[16] W. Chen, Y. Hou, D. Yao. SIR epidemics on evolving Erdős–Rényi graphs. *ALEA* 22 (2025) 245–282.

[17] H. Crane. Dynamic random networks and their graph limits. *The Annals of Applied Probability,* 26 (2016), 691–721.

[18] G.L.B. da Silva, R. Oliveira, D. Valesin. The contact process over a dynamical d-regular graph. *Annales de l'Institut Henri Poincaré: Probabilités et Statistiques* 60 (2024) 2849–2877.

[19] L. Decreusefond, J. S. Dhersin, P. Moyal, V. C. Tran. Large graph limit for an SIR process in random network with heterogeneous connectivity. *Annals of Applied Probability* 22(2) (2012) 541–575.

[20] J. F. Delmas, P. Frasca, F. Garin, V. C. Tran, A. Velleret, P. A. Zitt. Individual-based SIS models on (not so) dense large random networks. *ALEA* 21 (2024) 1375–1405.

[21] R. Durrett, D. Yao. Susceptible–infected epidemics on evolving graphs. *Electronic Journal of Probability* 27 (2022) 1–66.

[22] F. Garbe, J. Hladký, M. Sileikis, F. Skerman (2022). From flip processes to dynamical systems on graphons. *Annales de l'Institut Henri Poincare (B) Probabilites et statistiques.* 60 (2024) 2878–2922.

[23] M. Hilário, D. Ungaretti, D. Valesin, M.E. Vares. Results on the contact process with dynamic edges or under renewals. *Electronic Journal of Probability* 27 (2022) 1–31.

[24] Y. Huang, A. Rollin. The SIR epidemic on a dynamic Erdős–Rényi random graph. Preprint at arXiv:2404.12566, 2024.

[25] E. Jacob, A. Linker, P. Mörters. The contact process on dynamical scale-free networks. *Annales de l'Institut Henri Poincaré: Probabilités et Statistiques* 61 (2025) 1279–1318.

[26] S. Janson, M. Luczak, P. Windridge. Law of large numbers for the SIR epidemic on a random graph with given degrees. *Random Structures & Algorithms* 45(4) (2014) 726–763.

[27] Y. Jiang, R. Kassem, G. York, M. Junge, R. Durrett. SIR epidemics on evolving graphs. Preprint at arXiv:1901.06568, 2019.

[28] D. Keliger, I. Horváth, B. Takács. Local-density dependent Markov processes on graphons with epidemiological applications. *Stochastic Processes and their Applications* 148 (2022) 324–352.

[29] A.A. Lashari, P. Trapman. Branching process approach for epidemics in dynamic partnership network. *Journal of Mathematical Biology* 76 (2018) 265–294.
[30] L. Lovász. Large Networks and Graph Limits. American Mathematical Society, Providence RI, USA, 2012.
[31] L. Lovász, B. Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B* 96 (2006) 933–957.
[32] L. Lovász, B. Szegedy. Szemerédi's lemma for the analyst. *Geometric and Functional Analysis* 17 (2007) 252–270.
[33] M. Milewska, R. van der Hofstad, B. Zwart. SIR on locally converging dynamic random graphs. Preprint at arXiv:2501.09623 (2025).
[34] G. Pang, É. Pardoux, A. Velleret. Stochastic SIR model with individual heterogeneity and infection-age dependent infectivity on large non-homogeneous random graphs. Preprint at arXiv:2502.04225, 2025.
[35] B. Ráth. Time evolution of dense multigraph limits under edge-conservative preferential attachment dynamics. *Random Structures & Algorithms* 41 (2012), 365–390.
[36] D. Valesin. The contact process on random graphs. *Ensaios Matemáticos* 40 (2024) 1–115.
[37] E. Volz. SIR dynamics in random networks with heterogeneous connectivity. *Journal of Mathematical Biology* 56 (2008) 293–310.

Gran Sasso Science Institute, Viale Francesco Crispi 7, 67100 L'Aquila, Italy

Email address: simone.baldassarri@gssi.it

School of Mathematics and Statistics, Anita B. Lawrence Centre, UNSW Sydney, Sydney NSW 2052, Australia

Email address: p.braunsteins@unsw.edu.au

Mathematisch Instituut, Universiteit Leiden, Einsteinweg 55, 2333 CC Leiden, The Netherlands

Email address: denholla@math.leidenuniv.nl

Mathematisch Instituut, Universiteit Leiden, Einsteinweg 55, 2333 CC Leiden, The Netherlands

Email address: m.r.h.mandjes@math.leidenuniv.nl