

On the continuum limit of epidemiological models on graphs: Convergence and approximation results

*Original*

On the continuum limit of epidemiological models on graphs: Convergence and approximation results / Ayuso De Dios, B.; Dovetta, S.; Spinolo, L. V.. - In: MATHEMATICAL MODELS AND METHODS IN APPLIED SCIENCES. - ISSN 0218-2025. - 34:8(2024), pp. 1483-1532. [10.1142/S0218202524500271]

*Availability:*

This version is available at: 11583/2990431 since: 2024-07-10T13:27:12Z

*Publisher:*

World Scientific

*Published*

DOI:10.1142/S0218202524500271

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

(Article begins on next page)

# ON THE CONTINUUM LIMIT OF EPIDEMIOLOGICAL MODELS ON GRAPHS: CONVERGENCE AND APPROXIMATION RESULTS

BLANCA AYUSO DE DIOS<sup>1,3</sup>, SIMONE DOVETTA<sup>2</sup>, AND LAURA V. SPINOLO<sup>3</sup>

ABSTRACT. We focus on an epidemiological model (the archetypical SIR system) defined on graphs and study the asymptotic behavior of the solutions as the number of vertices in the graph diverges. By relying on the theory of *graphons* we provide a characterization of the limit and establish convergence results. We also provide approximation results for both deterministic and random discretizations.

AMS Subject Classification (2020): 35Q92, 92D30, 35R02, 05C99

Keywords: epidemiological models, SIR, continuum limit, sampling, graph limit, graphon

## 1. INTRODUCTION

In the present work we consider epidemiological models (that is, models describing the spreading of an infectious disease) defined on graphs. In real-world applications, each vertex of the graph can represent a single individual, but also a group of people living together or sharing a specific trait (for instance, people in the same age group), or a geographical entity like a neighbourhood, a town, a region. We are specifically interested in describing the behaviour of the model as the number  $n$  of vertices in the graph goes to  $+\infty$  and we establish convergence and approximation results, based on both deterministic and random algorithms. Our analysis borrows tools from the theory of *graphons*, which plays a prominent role in modern graph theory. In particular, our investigation was inspired by the available literature on dynamical systems on graphons, largely developed e.g. in the series of works [KM17, KM22, Med14a, Med14b, Med19C, Med19, MT20] (see also references therein) with a prominent focus on Kuramoto models and nonlinear/nonlocal diffusion equations. To the best of our knowledge, the present work is among the very first ones using the theory of graphons to study the asymptotic properties of epidemiological models on graphs as the number of vertices diverges (see also [DDZ22, GC19, VFG20] for papers with a completely different focus).

Although the origin of mathematical epidemiology can be traced back to Bernoulli [Ber60], nowadays the archetype of epidemiological models is the celebrated SIR (Susceptible-Infected-Recovered or Removed) system introduced by Kermack and McKendrick in [McK26, KMK27]. The dimensionless form of the SIR system is

$$\frac{ds}{dt} = -\beta si, \quad \frac{di}{dt} = \beta si - \gamma i, \quad \frac{dr}{dt} = \gamma i.$$

Here, the unknowns  $s, i, r : \mathbb{R}_+ \rightarrow \mathbb{R}$  represent the percentage of susceptible, infected and recovered individuals, whereas  $\beta \geq 0$  and  $\gamma \geq 0$  denote the (possibly time-dependent) infecting and recovering coefficients, respectively. Note furthermore that the sum  $s + i + r \equiv 1$  remains constant in time. The analysis of SIR models, and of more refined variants (SEIR, MSEIR, ...), has grown enormously through the decades, resulting in a huge amount of literature (we refer e.g. to the monographs and lecture notes [BCC01, BDW08, DG05, Mur02, Mur03], to the review [Het00] and to the references therein for extended discussions on this subject). Note that, in its standard version above, the SIR model implicitly assumes that contacts among individuals are uniform, i.e. that every individual has the same probability of interacting with any other. Since this is far from being true in real life, where social interactions are governed by complex patterns, several authors have investigated epidemiological models defined on networks (see for instance [BBPSV05, DGM08, MPSV02, New02, PSV01] and the reviews [NPP16, PSCMV15]). The network approach is by now regarded as a feasible tool to describe epidemic spreads through heterogeneous populations. The importance of taking into account the role of population heterogeneity in the modelling of an epidemics has been stressed once more with the COVID19 pandemic, which fuelled in the last years an upsurge of research on epidemiological models from different mathematical communities (recent works in these directions are for instance [ABNPV21, ABS21, BRCDF20]).

In what follows, even though we are confident that the results discussed in the present work could be applied to refined models with more compartments (SEIR, MSEIR, ...), to ease the exposition we

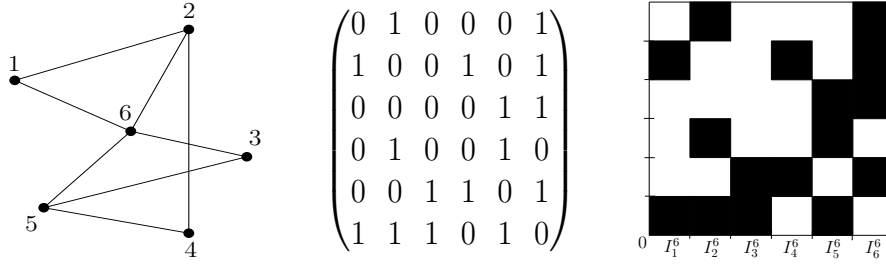


FIGURE 1. A simple graph  $\mathcal{G}$  with  $n = 6$  (leftmost), its adjacency matrix (center) and the pixel representation (rightmost) of the graphon  $W_{\mathcal{G}}$  associated to  $\mathcal{G}$  as in (1.7).

focus on the SIR model only. To discuss our main contributions we now fix a graph  $\mathcal{G}_n$  with  $n$  vertices labeled  $\{1, \dots, n\}$ , and we consider the SIR model

$$\begin{cases} \frac{ds_j^n}{dt} = -s_j^n(t) \frac{1}{n} \sum_{k=1}^n \beta_k^n(t) A_{jk}^n(t) i_k^n(t) \\ \frac{di_j^n}{dt} = s_j^n(t) \frac{1}{n} \sum_{k=1}^n \beta_k^n(t) A_{jk}^n(t) i_k^n(t) - \gamma_j^n(t) i_j^n(t) \\ \frac{dr_j^n}{dt} = \gamma_j^n(t) i_j^n(t). \end{cases} \quad j = 1, \dots, n \quad (1.1)$$

In the above system,  $s_j^n, i_j^n, r_j^n$  represent the percentage of susceptible, infected and recovered individuals at the vertex  $j$ , whereas  $\beta_j^n$  and  $\gamma_j^n$  are the (possibly time dependent) infecting and recovering coefficients.  $(A_{jk}^n)_{j,k=1,\dots,n}$  is the adjacency matrix which describes the connections (*edges*) between the vertices: the (possibly time dependent) coefficient  $A_{jk}^n$  indicates the weight given to the edge between the vertices  $j$  and  $k$ . If  $j$  and  $k$  are not connected, then  $A_{jk}^n = 0$ . We consider graphs which contain no loops and no multiple edges, undirected (edges are not oriented) and with nonnegative weights. These characteristics of  $\mathcal{G}_n$  translate into the conditions

$$A_{jj}^n = 0 \text{ for every } j, \quad A_{jk}^n = A_{kj}^n \geq 0 \text{ for every } j, k = 1, \dots, n. \quad (1.2)$$

Note however that in the present work we actually never use the no-loops condition  $A_{jj}^n = 0$ . System (1.1) is augmented with the initial conditions

$$s_j^n(0) = s_{j,0}^n, \quad i_j^n(0) = i_{j,0}^n, \quad r_j^n(0) = r_{j,0}^n \quad \text{for every } j = 1, \dots, n \quad (1.3)$$

and in view of modelling considerations they are such that

$$0 \leq s_{j,0}^n, i_{j,0}^n, r_{j,0}^n \leq 1, \quad s_{j,0}^n + i_{j,0}^n + r_{j,0}^n = 1 \quad \text{for every } j = 1, \dots, n, \quad (1.4)$$

which as we will show (see Lemma 2.2 in §2) implies that the solution satisfies

$$0 \leq s_j^n(t), i_j^n(t), r_j^n(t) \leq 1, \quad s_j^n(t) + i_j^n(t) + r_j^n(t) \equiv 1 \quad \text{for every } j = 1, \dots, n \text{ and every } t \in \mathbb{R}_+. \quad (1.5)$$

In a nutshell, the present paper aims at discussing the  $n \rightarrow +\infty$  limit of system (1.1). To do so, we rely on the celebrated theory of *graphons*, which has recently flourished with a series of fundamental works like [BCCG21, BCCZ18, BCCZ19, BCL10, BCLSV06, BCLSV08, BCLSV12, BS02, Lov12, LS07]. In general, a function  $W \in L^1([0, 1]^2; \mathbb{R})$  is called a *graphon* if  $W(x, y) = W(y, x)$  (see Definition B.1 in Appendix B). We refer to Appendix B for some rigorous definitions and a brief overview, mainly based on [BCCZ19], of the main results concerning graphons related to the present paper. Here we just mention that the basic idea of the theory of graphons is to identify a given graph  $\mathcal{G}_n$  with  $n$  vertices and adjacency matrix  $(A_{jk}^n)_{j,k=1,\dots,n}$  with a piecewise constant function  $W_{\mathcal{G}_n} : [0, 1]^2 \rightarrow \mathbb{R}$  (the *step-graphon*) defined on the unit square. By considering a uniform partition of  $]0, 1[$  into  $n$  intervals

$$I_j^n := \left] \frac{j-1}{n}, \frac{j}{n} \right], \quad j = 1, \dots, n, \quad (1.6)$$

and setting

$$W_{\mathcal{G}_n}(x, y) = A_{jk}^n \quad \text{if } (x, y) \in I_j^n \times I_k^n, \quad (1.7)$$

for  $\mathcal{G}_n$  satisfying (1.2), the piecewise constant function  $W_{\mathcal{G}_n}$  fullfills  $W_{\mathcal{G}_n}(x, x) = 0$ ,  $W_{\mathcal{G}_n}(x, y) = W_{\mathcal{G}_n}(y, x)$  and  $W_{\mathcal{G}_n} \geq 0$ . In Figure 1 is given the so called pixel picture of the graph. On the left

there is a simple graph<sup>1</sup>  $\mathcal{G}_n$  with  $n = 6$ , on the middle its adjacency matrix and on the right the pixel picture of the piecewise constant step-graphon  $W_{\mathcal{G}_n}$ . Here, the little squares have all edge length  $1/n$ . The 0's entries in  $A$  correspond to white squares, the 1's to black squares. The origin  $(0, 0)$  is placed at the upper left corner, in analogy with indexing matrix elements.

To apply the theory of graphons to system (1.1), we need to identify functions defined on the set of vertices of  $\mathcal{G}_n$  with functions defined on the unit interval. To this end, for a graph  $\mathcal{G}_n$  with  $n$  vertices we consider a vector valued function  $(u_i^n)_{i=1, \dots, n}$  defined say on  $\mathbb{R}_+$ , where each component  $u_j^n$  represents the value attained at a given vertex  $j$  of  $\mathcal{G}_n$ . In this work, we will always identify the vector valued function  $(u_i^n)_{i=1, \dots, n}$  with the *scalar* piecewise constant function  $u^n$  defined on  $\mathbb{R}_+ \times [0, 1]$  given by

$$u^n(t, x) := \sum_{j=1}^n u_j^n(t) \mathbb{1}_{I_j^n}(x), \quad (1.8)$$

with  $\mathbb{1}_{I_j^n}$  denoting the characteristic function of the interval  $I_j^n$ , defined in (1.6).

In this framework, the (formal) limit of (1.1) is thus the system

$$\begin{cases} \partial_t s(t, x) = -s(t, x) \int_0^1 \beta(t, y) W(t, x, y) i(t, y) dy \\ \partial_t i(t, x) = s(t, x) \int_0^1 \beta(t, y) W(t, x, y) i(t, y) dy - \gamma(t, x) i(t, x) \\ \partial_t r(t, x) = \gamma(t, x) i(t, x). \end{cases} \quad (t, x) \in \mathbb{R}_+ \times [0, 1] \quad (1.9)$$

We augment (1.9) with the initial condition

$$s(0, x) = s_0(x), \quad i(0, x) = i_0(x), \quad r(0, x) = r_0(x) \quad (1.10)$$

and in virtue of (1.5) we assume

$$0 \leq s_0, i_0, r_0 \leq 1, \quad s_0 + i_0 + r_0 = 1 \quad \text{a.e. on } [0, 1]. \quad (1.11)$$

The notion of solution for system (1.9) we will consider in this paper, which we refer to as *distributional solution*, is made precise in the next definition.

**Definition 1.1.** Fix  $T \in (0, \infty]$  and  $W \in L_{\text{loc}}^1([0, T] \times [0, 1]^2; \mathbb{R})$ ,  $\beta, \gamma \in L^\infty([0, T] \times [0, 1]; \mathbb{R})$  and  $s_0, i_0, r_0 \in L^\infty([0, 1]; \mathbb{R})$ . A triple  $(s, i, r) \in L^\infty([0, T] \times [0, 1]; \mathbb{R}^3)$  is a distributional solution of the Cauchy problem (1.9), (1.10) if for every  $(\varphi, \psi, \eta) \in C_c^\infty([0, T] \times [0, 1]; \mathbb{R}^3)$  we have

$$\begin{aligned} \int_{\mathbb{R}_+} \int_0^1 s(t, x) \left( \partial_t \varphi(t, x) - \varphi(t, x) \int_0^1 \beta(t, y) W(t, x, y) i(t, y) dy \right) dx dt &= \int_0^1 s_0(x) \varphi(0, x) dx \\ \int_{\mathbb{R}_+} \int_0^1 i(t, x) \left( \partial_t \psi(t, x) - \gamma(t, x) \psi(t, x) + s(t, x) \psi(t, x) \int_0^1 \beta(t, y) W(t, x, y) i(t, y) dy \right) dx dt & \\ &= \int_0^1 i_0(x) \psi(0, x) dx \\ \int_{\mathbb{R}_+} \int_0^1 r(t, x) \partial_t \eta(t, x) + \gamma(t, x) \eta(t, x) i(t, x) dx dt &= \int_0^1 r_0(x) \eta(0, x) dx. \end{aligned} \quad (1.12)$$

Clearly, the approximation and convergence to the continuum model (1.9) of the corresponding discrete models (1.1) hinge on suitable properties of the graphon under exam. We anticipate here that our analysis will exploit two different sets of assumptions on graphons: on the one side we take  $L^2$  graphons (cf. Case 1), on the other side we take  $L^1$  graphons with almost everywhere bounded degree function (cf. Case 2). For this reason, each of our main results will report in a single statement what happens in both situations, distinguishing between Case 1 and Case 2. Since it is readily seen that none of these conditions implies the other one, we believe it is interesting to treat both cases, thus obtaining results that apply to a wider class of graphons (see also (i) right after Theorem 1.5 below for a further comparison between the two assumptions).

Our first result, Proposition 1.2, guarantees the existence and uniqueness of a distributional solution (in the sense of Definition 1.1) to the Cauchy problem (1.9), (1.10).

<sup>1</sup>Hence  $A_{jk} \in \{0, 1\}$  for all  $j, k$

**Proposition 1.2.** Fix  $T > 0$  and  $W \in L^1([0, T] \times [0, 1]^2; \mathbb{R})$  satisfying  $W \geq 0$  and  $s_0, i_0, r_0 \in L^\infty([0, 1]; \mathbb{R})$  satisfying (1.11). Assume furthermore that  $\gamma, \beta \in L^\infty([0, T] \times [0, 1]; \mathbb{R})$  satisfy  $\gamma, \beta \geq 0$  almost everywhere on  $[0, T] \times [0, 1]$ . Then if either of the following conditions hold:

- CASE 1:  $W \in L^2([0, T] \times [0, 1]^2; \mathbb{R})$ ; or
- CASE 2: there is a constant  $K_d > 0$  such that

$$\operatorname{ess\,sup}_{t \in [0, T], x \in [0, 1]} \int_0^1 W(t, x, y) dy \leq K_d; \quad (1.13)$$

then the Cauchy problem (1.9), (1.10) admits exactly one distributional solution satisfying

$$0 \leq s, i, r \leq 1, \quad s + i + r = 1, \quad \text{a.e. on } [0, T] \times [0, 1]. \quad (1.14)$$

Prior to state our first main result, we recall the meaning of the so called *cut norm*  $\|\cdot\|_\square$ , which for a graphon  $W$  is defined as

$$\|W\|_\square := \sup_{S, T \subseteq [0, 1]} \left| \iint_{S \times T} W(x, y) dx dy \right|, \quad (1.15)$$

where the supremum is taken over all pairs of measurable subsets  $S, T$  in  $[0, 1]$ . This norm is the prototypical one in the theory of graphons (see the whole discussion in Appendix B). We can now state our first main convergence result.

**Theorem 1.3.** Fix  $T > 0$  and let  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  be a sequence of time-dependent weighted undirected graphs with adjacency matrix satisfying (1.2) and

$$\lim_{n \rightarrow +\infty} \int_0^T \|W_{\mathcal{G}_n}(t, \cdot, \cdot) - W(t, \cdot, \cdot)\|_\square dt = 0 \quad (1.16)$$

for some graphon  $W \in L^1([0, T] \times [0, 1]^2; \mathbb{R})$ . Assume furthermore that either of the following conditions holds:

- CASE 1: there is a constant  $K_0 > 0$  such that

$$\|W_{\mathcal{G}_n}\|_{L^2([0, T] \times [0, 1]^2; \mathbb{R})} \leq K_0 \quad \text{for every } n \in \mathbb{N}; \text{ or} \quad (1.17)$$

- CASE 2: there is a constant  $K_1 > 0$  such that

$$\operatorname{ess\,sup}_{t \in [0, T], x \in [0, 1]} \int_0^1 W_{\mathcal{G}_n}(t, x, y) dy \leq K_1 \quad \text{for every } n \in \mathbb{N}. \quad (1.18)$$

Recalling the notation (1.8), assume also that the coefficients  $\beta^n, \gamma^n$  satisfy

$$0 \leq \beta^n, \gamma^n \leq M \text{ a.e. on } [0, T] \times [0, 1], \quad \beta^n \rightarrow \beta, \gamma^n \rightarrow \gamma \text{ strongly in } L^1([0, T] \times [0, 1]; \mathbb{R}), \quad (1.19)$$

for some constant  $M > 0$  and some functions  $\beta, \gamma \in L^\infty([0, T] \times [0, 1]; \mathbb{R})$ .

If  $(s^n, i^n, r^n)$  is the solution of (1.1), (1.3) for given initial data  $s_0^n, i_0^n, r_0^n$  satisfying (1.4) for every  $n$ , then we have up to subsequences

$$(s^n, i^n, r^n) \xrightarrow{*} (s, i, r) \text{ weakly* in } L^\infty([0, T] \times [0, 1]; \mathbb{R}^3) \text{ as } n \rightarrow +\infty, \quad (1.20)$$

where  $(s, i, r) : [0, T] \times [0, 1] \rightarrow [0, 1]^3$  is the distributional solution of (1.9), (1.10) satisfying (1.14) and attaining the initial condition (1.10) for some data  $(s_0, i_0, r_0)$  satisfying (1.11).

Some remarks are in order. First, assumption (1.16) is entirely natural from the point of view of graphon theory in view of the analysis in [BCCZ19, BCLSV08]. We refer to Appendix B and in particular to Remark B.2 for a more technical discussion, here we just mention that assumption (1.16) is satisfied (up to subsequences and vertices relabelling) in two very relevant cases: (i) if  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  is a dense<sup>2</sup> sequence of graphs such that  $\|W_{\mathcal{G}_n}\|_{L^p([0, T] \times [0, 1]^2; \mathbb{R})}$  is uniformly bounded for some  $p > 1$ ; (ii) if  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  is a sparse sequence of graphs satisfying suitable topological assumptions (see Appendix B for the precise conditions), then (1.16) holds true up to renormalization<sup>3</sup>, i.e. provided we replace  $\mathcal{G}_n$  by  $\mathcal{G}_n / \|W_{\mathcal{G}_n}\|_{L^1([0, T] \times [0, 1]^2; \mathbb{R})}$ .

Second, both conditions (1.17) and (1.18) pass to the limit, i.e. they are satisfied by the limit graphon  $W$  in (1.16), see respectively Lemmas 2.4-2.5 below. This implies that the limit system (1.9)

<sup>2</sup>We refer to Appendix B for the definitions of dense and sparse sequence of graphs

<sup>3</sup>Actually, condition (1.16) is directly satisfied with no renormalization by sparse graph sequences too, but this is of limited interest since the limit would be the trivial function  $W \equiv 0$

satisfies the assumptions of Proposition 1.2 and in particular that there is a unique solution of the Cauchy problem (1.9), (1.10) satisfying (1.14). This in turn ensures that, by fixing the initial data (1.10) of the limit problem, then there is no need to pass to subsequences in (1.20), as the whole sequence converges.

Third, by looking at the proof of Theorem 1.3 one realizes that, if  $\mathcal{G}_n$  is a sequence of random graphs satisfying (1.16) and either (1.17) or (1.18) almost surely, then (1.20) holds almost surely.

Finally, Theorem 1.3 can be regarded as a *convergence* result: given a sequence of graphs  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  satisfying (1.16) for some  $W$ , we characterize the asymptotic behaviour of the corresponding sequence of SIR models (1.1).

In the next results, we somehow adopt the opposite viewpoint: given a graphon  $W \in L^1([0, T] \times [0, 1]^2; \mathbb{R})$  satisfying suitable conditions and the corresponding system (1.9), we look at discrete approximations to the solution of (1.9). Our approach here focuses on the approximation of  $W$  by sequences of graphs constructed in such a way that the corresponding solutions of (1.1) converge to the solution of (1.9). Obviously, different samplings of  $W$  provide different approximating graphs and the actual convergence of the specific approximation to the limit problem depends on the considered discretization. In this work, we study both deterministic and random approximations. However, before discussing in detail our results in this direction, we stress immediately that all the sampling procedures we take into account retain much more information on the limit graphon  $W$  than the rather mild assumption (1.16) required by Theorem 1.3. As a consequence, all the convergence results we obtain (see e.g. (1.24)-(1.25) and (1.29)-(1.31) below) hold in a much stronger topology than the one in (1.20).

We now introduce the deterministic approximation we consider. Let  $n \in \mathbb{N}$ , recall (1.6) and set

$$\begin{aligned} \langle u \rangle_{I_j^n} &:= n \int_{I_j^n} u(x) dx \quad \text{for every } j = 1, \dots, n, \quad \forall u \in L^1([0, 1]; \mathbb{R}), \\ \langle U \rangle_{I_j^n \times I_k^n} &:= n^2 \iint_{I_j^n \times I_k^n} U(x) dx \quad \text{for every } j, k = 1, \dots, n, \quad \forall U \in L^1([0, 1]^2; \mathbb{R}). \end{aligned} \quad (1.21)$$

Given a possibly time-dependent graphon  $W$ , for every  $n \in \mathbb{N}$  we then let  $\mathcal{G}_n$  be the possibly time-dependent graph with  $n$  vertices and adjacency matrix

$$A_{jk}^n(t) := \langle W(t, \cdot, \cdot) \rangle_{I_j^n \times I_k^n}. \quad (1.22)$$

We can now state our first approximation result.

**Theorem 1.4.** *Fix  $T > 0$  and assume  $W \in L^1([0, T] \times [0, 1]^2; \mathbb{R}_+)$  is a non-negative time-dependent graphon. Let  $\beta, \gamma \in L^\infty([0, T] \times [0, 1]; \mathbb{R})$  and  $s_0, i_0, r_0 \in L^\infty([0, 1]; \mathbb{R})$  satisfy  $\beta, \gamma \geq 0$  and (1.11), respectively. Set*

$$\begin{aligned} \beta_j^n(t) &:= \langle \beta(t, \cdot) \rangle_{I_j^n}, \quad \gamma_j^n(t) := \langle \gamma(t, \cdot) \rangle_{I_j^n} \quad \text{for a.e. } t \in \mathbb{R}_+ \text{ and every } j = 1, \dots, n \\ s_{j,0}^n &:= \langle s_0 \rangle_{I_j^n}, \quad i_{j,0}^n := \langle i_0 \rangle_{I_j^n}, \quad r_{j,0}^n := \langle r_0 \rangle_{I_j^n} \quad \text{for every } j = 1, \dots, n \end{aligned} \quad (1.23)$$

and let  $\mathcal{G}_n$  be the graph with adjacency matrix  $(A_{jk}^n)_{j,k=1,\dots,n}$  defined in (1.22). Let  $(s^n, i^n, r^n)$  be the solution of the Cauchy problem (1.1), (1.3) written using the notation (1.8). Then, provided  $(s, i, r)$  is the distributional solution of (1.9), (1.10) satisfying (1.14), the following results hold:

- CASE 1: if  $W \in L^2([0, T] \times [0, 1]^2; \mathbb{R}_+)$  then

$$\lim_{n \rightarrow +\infty} \operatorname{ess\,sup}_{t \in [0, T]} (\|s(t, \cdot) - s^n(t, \cdot)\|_{L^2([0, 1]; \mathbb{R})} + \|i(t, \cdot) - i^n(t, \cdot)\|_{L^2([0, 1]; \mathbb{R})} + \|r(t, \cdot) - r^n(t, \cdot)\|_{L^2([0, 1]; \mathbb{R})}) = 0; \quad (1.24)$$

- CASE 2: if  $W$  satisfies (1.13) then

$$\lim_{n \rightarrow +\infty} \operatorname{ess\,sup}_{t \in [0, T]} (\|s(t, \cdot) - s^n(t, \cdot)\|_{L^1([0, 1]; \mathbb{R})} + \|i(t, \cdot) - i^n(t, \cdot)\|_{L^1([0, 1]; \mathbb{R})} + \|r(t, \cdot) - r^n(t, \cdot)\|_{L^1([0, 1]; \mathbb{R})}) = 0. \quad (1.25)$$

Note that with the sampling procedure defined by (1.22) and (1.23) the resulting discretization coincides with a discontinuous piecewise constant Galerkin approximation to (1.9). In particular, if for instance  $\beta_j^n$  is as in (1.23) and we define  $\beta^n$  using the notation (1.8), then  $\beta^n$  is exactly the  $L^2$ -orthogonal projection of  $\beta$  onto the space of piecewise constant functions on the uniform partition of  $[0, 1]$  of granularity  $1/n$ . Similarly, if  $W(t, \cdot) \in L^2([0, 1]^2; \mathbb{R})$  the matrix  $(A_{jk}^n(t))_{1 \leq j, k \leq n}$  in (1.22)

coincides with the  $L^2$ -orthogonal projection of the graphon  $W(t, \cdot)$  onto the space of piecewise constant functions on the uniform cartesian grid of  $[0, 1]^2$ .

We now introduce our second sampling procedure, which belongs to the family of the so called  $W$ -random graphs introduced in [DF81], popularized by [LS06] and then widely used in the context of graphon theory. However, let us stress that we will discuss such procedure only for time-independent graphons, so that the remainder of our analysis here will apply only to this kind of graphons (see (ii) right after Theorem 1.5 below for further details on this point).

Inspired by the construction given in [KM22, Med19], for any graphon  $W \in L^1([0, 1]^2; \mathbb{R}_+)$  and for  $\alpha \in ]0, 1[$  fixed, we introduce the *scaled sparse  $W$ -random graph*  $n^\alpha \mathcal{G}(n, W, n^{-\alpha})$  as the graph with  $n$  nodes and adjacency matrix given by

$$\mathbb{P}(A_{jk}^n = n^\alpha) = \langle \widehat{W}^n \rangle_{I_j^n \times I_k^n}, \quad \text{where } \widehat{W}^n := \min\{1, n^{-\alpha} W\}, \quad \mathbb{P}(A_{jk}^n = 0) = 1 - \mathbb{P}(A_{jk}^n = n^\alpha). \quad (1.26)$$

The scaling factor  $n^{-\alpha}$  is the so called target edge density of the graph and it does enter in the definition (1.26) of the matrix probability distribution. On the one hand, it dictates the level to trim the unbounded graphon  $W$ . On the other hand, it allows to ensure that the cut-off in the minimum acts only on part of the edges.

*Remark 1.1.* The construction considered in [KM22, Med19] (and more generally the definition of the so-called *sparse  $W$ -random graph* in the theory of graphons) is slightly different from (1.26). In [KM22, Med19] the *sparse  $W$ -random graph*  $\mathcal{G}(n, W, n^{-\alpha})$  has adjacency matrix

$$\mathbb{P}(A_{jk}^n = 1) = \langle \widehat{W}^n \rangle_{I_j^n \times I_k^n}, \quad \text{where } \widehat{W}^n := \min\{1, n^{-\alpha} W\}, \quad \mathbb{P}(A_{jk}^n = 0) = 1 - \langle \widehat{W}^n \rangle_{I_j^n \times I_k^n}. \quad (1.27)$$

Here, we consider directly the *scaled sparse  $W$ -random graph*  $n^\alpha \mathcal{G}(n, W, n^{-\alpha})$  as defined in (1.28) since that is the graph capable of approximating the graphon  $W$ .

In the case of bounded graphons  $W \in L^\infty([0, 1]^2; [0, 1])$  it is standard to choose  $\alpha = 0$ , so that the adjacency matrix of the resulting graph boils down to a random matrix with entries following the Bernoulli distribution

$$\mathbb{P}(A_{jk}^n = 1) = \langle W \rangle_{I_j^n \times I_k^n}, \quad \mathbb{P}(A_{jk}^n = 0) = 1 - \langle W \rangle_{I_j^n \times I_k^n}. \quad (1.28)$$

Such sparse  $W$ -random graph is denoted here by  $\mathcal{G}(n, W)$ .

We now state the main results for the random approximation to (1.9) given by the scaled sparse  $W$ -random graphs  $n^\alpha \mathcal{G}(n, W, n^{-\alpha})$  as random samplings of  $W$ .

**Theorem 1.5.** *Fix  $T > 0$  and let  $W \in L^1([0, 1]^2; \mathbb{R}_+)$  be a graphon and  $\gamma, \beta \in L^\infty([0, T] \times [0, 1]; \mathbb{R})$  satisfy  $\beta, \gamma \geq 0$ . Let  $\alpha \in ]0, 1[$  be fixed. For every  $n \in \mathbb{N}$  and  $j = 1, \dots, n$ , let  $\beta_j^n, \gamma_j^n, s_{j,0}^n, i_{j,0}^n, r_{j,0}^n$  be as in (1.23) and  $\mathcal{G}_n = n^\alpha \mathcal{G}(n, W, n^{-\alpha})$  be the scaled sparse  $W$ -random graph given by (1.26). Let  $(s^n, i^n, r^n)$  be the solution of the Cauchy problem (1.1), (1.3) written by using the notation (1.8). Then the following results hold:*

- **CASE 1:** *if  $W \in L^2([0, 1]^2; \mathbb{R}_+)$  and  $\alpha \in ]0, 1/2[$  then*

$$\begin{aligned} \lim_{n \rightarrow +\infty} \operatorname{ess\,sup}_{t \in [0, T]} (\|s(t, \cdot) - s^n(t, \cdot)\|_{L^2([0, 1]; \mathbb{R})} + \|i(t, \cdot) - i^n(t, \cdot)\|_{L^2([0, 1]; \mathbb{R})} \\ + \|r(t, \cdot) - r^n(t, \cdot)\|_{L^2([0, 1]; \mathbb{R})}) = 0 \quad \text{almost surely;} \end{aligned} \quad (1.29)$$

- **CASE 2:** *if there is a constant  $K_a > 0$  such that*

$$\operatorname{ess\,sup}_{x \in [0, 1]} \int_0^1 W(x, y) dy \leq K_a, \quad (1.30)$$

*then, for  $\alpha \in ]0, 1[$ ,*

$$\begin{aligned} \lim_{n \rightarrow +\infty} \operatorname{ess\,sup}_{t \in [0, T]} (\|s(t, \cdot) - s^n(t, \cdot)\|_{L^1([0, 1]; \mathbb{R})} + \|i(t, \cdot) - i^n(t, \cdot)\|_{L^1([0, 1]; \mathbb{R})} \\ + \|r(t, \cdot) - r^n(t, \cdot)\|_{L^1([0, 1]; \mathbb{R})}) = 0 \quad \text{almost surely.} \end{aligned} \quad (1.31)$$

We wish to note (as will be evident from the proof of Theorem 1.5) that if  $W \in L^\infty([0, 1]^2; [0, 1])$ , the above theorem applies with  $\mathcal{G}_n$  given by the sparse random graph  $\mathcal{G}(n, W)$  as defined in (1.28).

Prior to close this introduction we comment on the main novelties of our approximation results. The proof of Theorems 1.4–1.5 borrow some of the techniques from the analysis in [Med19] concerning the limit of Kuramoto systems and nonlinear/nonlocal diffusion defined on graphs (see also [Med14a, Med14b, Med19C] and [BCD20] for a different but related problem). However, with respect to previous analyses [KM17, Med19, MT20, KM22], the specific features of the SIR system allows here to obtain the following improvements:

- (i) we relax the regularity assumptions on  $W$ , namely we require only  $W \in L^2([0, 1]^2; \mathbb{R}_+)$  (cf. CASE 1). This in particular implies that Theorems 1.4–1.5 (contrary to [Med19, Theorem 3.1] and the main results in [MT20, KM22]) directly apply to the power law case  $W(x, y) = (xy)^{-\mu}$ ,  $\mu \in ]0, 2^{-1}[$ . As pointed out for instance in the introduction to [BCCG21, BCCZ19] (see also the one in [MT20]), the power law case is usually considered technically challenging but also fascinating and compelling in view of applications, as several real-world networks have a power law structure. From the technical standpoint, when  $W$  is an  $L^2$  graphon we circumvent condition (1.13) by a careful manipulation of the particular structure of the SIR system. Furthermore, unlike in [Med19, MT20, KM22], we can also go below the  $L^2$  summability framework (cf. CASE 2), provided we adopt (1.13) (resp. (1.30) in the random case). Note that (1.13) (resp. (1.30) in the random case) is for instance satisfied by the graphons  $W(x, y) = (x + y)^{-\mu}$ ,  $\mu \in ]0, 1[$ , which do not belong to  $L^2([0, 1]^2; \mathbb{R})$  if  $\mu \geq 2^{-1}$ ;
- (ii) for the deterministic approximation in Theorem 1.4, we allow for time-dependent graphons. This might be particularly relevant in view of applications because it allows to consider changes in the network structure (e.g. due to the implementation of lockdown measures). However, with the random approximation in Theorem 1.5 our results are valid only for time independent graphons. This limitation is technical, mainly due to measurability issues on the integrals with respect to time of certain stochastic processes arising in the computations. Hence, at present we are not able to adapt the techniques used in the proof of Theorem 1.5 to deal with time dependent graphons  $W$ .

**Paper outline.** The remainder of the paper is organized as follows. In §2 we collect some preliminary results. In §3 we discuss the proof of Proposition 1.2 and of Theorem 1.3. §4 contains the proof of Theorem 1.4 for the deterministic approximation, while in §5 we present the proof of Theorem 1.5 for the random approximation. The paper is completed with two appendices. Appendix A collects the main notation used in the paper, whereas Appendix B comprises a brief overview of some definitions and results on graphons (providing in particular sufficient conditions for (1.16) to hold).

**Norms notation.** In the remainder of the paper, symbols like  $\|u\|_p$  will be used to denote the  $L^p$  norm of  $u$  when computed with respect to all the variables  $u$  depends on. Conversely, when the  $L^p$  norm of  $u$  is computed only with respect to some of its variables, we will use the full notation indicating explicitly the space of integration. Moreover, for the sake of brevity we will often omit the value space when it is equal to  $\mathbb{R}$ , writing e.g.  $L^p(0, 1), L^p([0, T] \times [0, 1])$  to denote  $L^p([0, 1]; \mathbb{R}), L^p([0, T] \times [0, 1]; \mathbb{R})$ .

## 2. PRELIMINARY RESULTS

In this section we collect some subsidiary results that will be useful for the subsequent analysis. Precisely, in §2.1 we prove existence and uniqueness of the solution for the discrete SIR system (1.1), whereas in §2.2 we establish a priori estimates for distributional solutions of the continuum SIR system (1.9). In §2.3 we show how conditions (1.17) and (1.18) for a sequence of graphons  $W_n$  imply corresponding conditions on the limit graphon  $W$ . In §2.4 we recall a standard convergence results for integral means that will be frequently used in the following.

**2.1. The discrete SIR model on graphs.** The following lemma is a direct consequence of the Cauchy Lipschitz Picard Lindelöf Theorem on existence and uniqueness of solutions of ODEs. In the statement we use the notation (1.8).

**Lemma 2.1.** *Fix  $n \in \mathbb{N}$  and assume  $\beta_j^n, \gamma_j^n, A_{jk}^n \in L^\infty(\mathbb{R}_+)$  satisfy  $\beta_j^n, \gamma_j^n, A_{jk}^n \geq 0$  a.e. on  $\mathbb{R}_+$ , for every  $j, k = 1, \dots, n$ . For every given initial condition (1.3) satisfying (1.4) there are  $\nu > 0$  and a unique solution  $s^n, i^n, r^n : [0, \nu] \times [0, 1] \rightarrow \mathbb{R}$  of the Cauchy problem (1.1), (1.3).*

Note that the above lemma only provides local-in-time existence and uniqueness. However, the local-in-time solution can be extended for every  $t \geq 0$  and hence it is actually a global in-time-solution  $s^n, i^n, r^n : \mathbb{R}_+ \times [0, 1] \rightarrow \mathbb{R}$ . This is a consequence of the following lemma. The proof is standard, but we provide it for the sake of completeness.

**Lemma 2.2.** *Under the assumptions of Lemma 2.1, the solution of the Cauchy problem (1.1), (1.3) satisfies (1.5).*

*Proof.* The equality in (1.5) follows from the fact that by adding the three lines of (1.1) we get that the derivative of  $s_j^n(t) + i_j^n(t) + r_j^n(t)$  vanishes.

Next, we point out that  $s_j^n(t) = 0$  for every  $t \in \mathbb{R}$  is a solution of the equation at the first line of (1.1). By the uniqueness part of the Cauchy Lipschitz Picard Lindelöf Theorem, this implies that, since  $s_{j,0}^n \geq 0$ , then  $s_j^n(t) \geq 0$  for every  $t$ . We now prove that  $i_j^n(t) \geq 0$ , for every  $j = 1, \dots, n$ . As a first step, we establish the proof under the further assumption

$$A_{jk}^n > 0, \quad \text{for every } j, k = 1, \dots, n. \quad (2.1)$$

We separately consider the two possible instances:

INSTANCE (A):  $i_{j,0}^n > 0$  for every  $j = 1, \dots, n$ . We set

$$\bar{t} := \sup\{t \in \mathbb{R}_+ : i_j^n(\tau) > 0 \text{ for every } j = 1, \dots, n \text{ and every } \tau \in [0, t]\}.$$

We point out that  $\bar{t} > 0$  and we now show that  $\bar{t} = +\infty$ . Assume by contradiction that  $\bar{t} < +\infty$  and, just to fix the ideas, assume that  $i_1^n(\bar{t}) = 0$ . Note that, in this case, it must be  $s_1^n(\bar{t}) > 0$ . Indeed, if it were  $s_1^n(\bar{t}) = i_1^n(\bar{t}) = 0$ , then  $r_1^n(\bar{t}) = 1$ . However,  $(0, 0, 1)$  is an equilibrium for the component  $(s_1^n, i_1^n, r_1^n)$  (as for any other component), so that it would follow  $s_{1,0}^n = i_{1,0}^n = 0$ ,  $r_{1,0}^n = 1$ , contradicting the assumption  $i_{1,0}^n > 0$ . Hence,  $s_1^n(\bar{t}) > 0$ . Then there are two possibilities:

- (i)  $i_2^n(\bar{t}) = \dots = i_n^n(\bar{t}) = 0$ . Since  $(s_1^n, \dots, s_n^n, 0, \dots, 0, r_1^n, \dots, r_n^n)$  is an equilibrium for (1.1) for every  $(s_1^n, \dots, s_n^n, r_1^n, \dots, r_n^n)$ , this implies  $i_1^n(0) = \dots = i_n^n(0) = 0$ , which contradicts the definition of INSTANCE (A).
- (ii) there is  $k$  such that  $i_k^n(\bar{t}) > 0$ . Since  $A_{jk}^n > 0$  and  $s_1^n(\bar{t}) > 0$ , this implies that  $di_1^n/dt > 0$  at  $t = \bar{t}$  and again contradicts the definition of  $\bar{t}$ .

INSTANCE (B): there is  $j = 1, \dots, n$  such that  $i_{j,0}^n = 0$ . Note that, if  $s_{j,0}^n = i_{j,0}^n = 0$ , then  $r_{j,0}^n = 1$ ,  $(s_j^n(t), i_j^n(t), r_j^n(t)) = (0, 0, 1)$  for every  $t$  and the  $j$ -th component does not interact with any other component in the system. Therefore, without loss of generality we can restrict our attention only to the components  $j$  for which  $s_{j,0}^n > 0$ . Just to fix the ideas, we assume  $i_{1,0}^n = 0$ . If  $i_{2,0}^n = \dots = i_{n,0}^n = 0$  then by arguing as in item (i) above we conclude that  $i_1^n(t) = \dots = i_n^n(t) = 0$  for every  $t$ , which in particular yields our claim. If there is  $k$  such that  $i_{k,0}^n > 0$ , then  $di_j^n/dt > 0$  at  $t = 0$  for every  $j$  for which  $i_{j,0}^n = 0$  (since  $A_{jk}^n > 0$  and  $s_{j,0}^n > 0$  by assumption). This implies that  $i_j^n(t) > 0$  for every  $j = 1, \dots, n$  and  $t \in ]0, \sigma[$  for some sufficiently small  $\sigma > 0$ . We can then consider the Cauchy problem obtained by coupling (1.1) with the datum assigned at  $\sigma/2$  and apply the same argument as in the previous case.

To conclude the proof of the inequality  $i_j^n \geq 0$  we are left to remove the assumption (2.1). To this end, we rely on an approximation argument: we fix  $\varepsilon > 0$ , replace  $A_{jk}^n$  with  $A_{jk}^n + \varepsilon$ , for every  $j, k = 1, \dots, n$ , and term  $s_j^{n\varepsilon}(t), i_j^{n\varepsilon}(t), r_j^{n\varepsilon}(t)$  the corresponding solution of the Cauchy problem (1.1), (1.3). Since  $A_{jk}^n + \varepsilon > 0$ , by the previous step  $i_j^{n\varepsilon}(t) \geq 0$  for every  $t$  and every  $\varepsilon > 0$ . By the continuous dependence of the solution of ODEs on parameters,  $i_j^{n\varepsilon}(t)$  converges to  $i_j^n(t)$  as  $\varepsilon \rightarrow 0^+$  and this implies  $i_j^n(t) \geq 0$ .

Finally, since for every  $j = 1, \dots, n$  we have  $\gamma_j^n \geq 0$  and  $i_j^n \geq 0$ , then  $dr_j^n/dt \geq 0$  and this yields  $r_j^n \geq 0$ .  $\square$

**2.2. A priori estimates for distributional solutions of (1.9).** We now derive a preliminary result which will be useful in the next section to establish Proposition 1.2.

**Lemma 2.3.** *Fix  $T > 0$ ,  $W \in L^1([0, T] \times [0, 1]^2; \mathbb{R})$  and  $\beta, \gamma \in L^\infty([0, T] \times [0, 1])$ . Let  $(s, i, r) \in L^\infty([0, T] \times [0, 1]; \mathbb{R}^3)$  be a distributional solution of (1.9) satisfying (1.14). Then, either:*

- CASE 1: *if  $W \in L^2([0, T] \times [0, 1]^2; \mathbb{R})$ , then  $\partial_t s, \partial_t i, \partial_t r \in L^2([0, T] \times [0, 1])$  and*

$$\|\partial_t s\|_2, \|\partial_t i\|_2, \|\partial_t r\|_2 \leq C(\|W\|_2, \|\beta\|_\infty, \|\gamma\|_\infty), \quad (2.2)$$

where  $C(\|W\|_2, \|\beta\|_\infty, \|\gamma\|_\infty)$  is a suitable constant only depending on  $\|W\|_2, \|\beta\|_\infty$  and  $\|\gamma\|_\infty$ ; or

- CASE 2: if (1.13) holds true, then  $\partial_t s, \partial_t i, \partial_t r \in L^\infty([0, T] \times [0, 1])$  and

$$\|\partial_t s\|_\infty, \|\partial_t i\|_\infty, \|\partial_t r\|_\infty \leq C(K_d, \|\beta\|_\infty, \|\gamma\|_\infty), \quad (2.3)$$

where  $C(K_d, \|\beta\|_\infty, \|\gamma\|_\infty)$  is a suitable constant only depending on  $K_d, \|\beta\|_\infty$  and  $\|\gamma\|_\infty$ .

*Proof.* Let  $(s, i, r)$  be a distributional solution of (1.9) satisfying (1.14). Owing to (1.12), the distributional derivatives  $(\partial_t s, \partial_t i, \partial_t r)$  are given by the right hand side of (1.9) and hence are locally summable functions. We now separately consider the following two cases.

CASE 1:  $W \in L^2([0, T] \times [0, 1]^2; \mathbb{R})$ . By using Jensen's inequality we get

$$\begin{aligned} \int_0^T \int_0^1 [\partial_t s(t, x)]^2 dx dt &= \int_0^T \int_0^1 \left( -s(t, x) \int_0^1 \beta(t, y) W(t, x, y) i(t, y) dy \right)^2 dx dt \\ &\leq \|s\|_\infty^2 \|\beta\|_\infty^2 \|i\|_\infty^2 \int_0^T \int_0^1 \left( \int_0^1 W(t, x, y) dy \right)^2 dx dt \\ &\stackrel{\text{Jensen}}{\leq} \|s\|_\infty^2 \|\beta\|_\infty^2 \|i\|_\infty^2 \int_0^T \int_0^1 \int_0^1 W^2(t, x, y) dy dx dt \stackrel{(1.14)}{\leq} \|\beta\|_\infty^2 \int_0^T \int_0^1 \int_0^1 W^2(t, x, y) dy dx dt, \end{aligned}$$

which yields a control on the first term in (2.2), and relying on a similar argument we control also the other terms in (2.2).

CASE 2:  $W$  satisfies (1.13). We get

$$|\partial_t s(t, x)| = \left| -s(t, x) \int_0^1 \beta(t, y) W(t, x, y) i(t, y) dy \right| \stackrel{(1.14)}{\leq} \|\beta\|_\infty \left| \int_0^1 W(t, x, y) dy \right| \stackrel{(1.13)}{\leq} K_d \|\beta\|_\infty.$$

By using an analogous argument we control the other terms in (2.3).  $\square$

### 2.3. Limit conditions coming from (1.17) and (1.18).

**Lemma 2.4.** CASE 1. Assume (1.16) and (1.17). Then the limit graphon satisfies  $\|W\|_{L^2([0, T] \times [0, 1]^2; \mathbb{R})} \leq K_0$ .

*Proof.* Owing to (1.17), we have that, as  $n \rightarrow +\infty$  and up to subsequences,  $W_{\mathcal{G}_n} \rightharpoonup V$  weakly in  $L^2([0, T] \times [0, 1]^2)$ , for some limit function  $V \in L^2([0, T] \times [0, 1]^2)$  with  $\|V\|_2 \leq K_0$  by weak lower semicontinuity. To conclude, we have to show that  $V \equiv W$ . By the uniqueness of the distributional limit, it suffices to show that

$$\lim_{n \rightarrow +\infty} \int_0^T \iint_{[0, 1]^2} \xi(t, x, y) [W_{\mathcal{G}_n} - W](t, x, y) dx dy dt = 0 \quad (2.4)$$

for every test function  $\xi \in C_c^\infty([0, T] \times [0, 1]^2)$ . By the Stone-Weierstrass Theorem, to verify (2.4) it suffices to show that for every  $\eta \in C^\infty([0, T])$ ,  $\varphi, \psi \in C^\infty([0, 1])$  we have

$$\lim_{n \rightarrow +\infty} \int_0^T \eta(t) \iint_{[0, 1]^2} \varphi(x) \psi(y) [W_{\mathcal{G}_n} - W](t, x, y) dx dy dt = 0. \quad (2.5)$$

To this end, we point out that

$$\begin{aligned} &\left| \int_0^T \eta(t) \iint_{[0, 1]^2} \varphi(x) \psi(y) [W_{\mathcal{G}_n} - W](t, x, y) dx dy dt \right| \leq \int_0^T \left| \eta(t) \iint_{[0, 1]^2} \varphi(x) \psi(y) [W_{\mathcal{G}_n} - W](t, x, y) dx dy \right| dt \\ &\leq \|\eta\|_{C^0} \int_0^T \left| \iint_{[0, 1]^2} \varphi(x) \psi(y) [W_{\mathcal{G}_n} - W](t, x, y) dx dy \right| dt \\ &= \|\eta\|_{C^0} \|\varphi\|_{C^0} \|\psi\|_{C^0} \int_0^T \left| \iint_{[0, 1]^2} \frac{\varphi(x)}{\|\varphi\|_{C^0}} \frac{\psi(y)}{\|\psi\|_{C^0}} [W_{\mathcal{G}_n} - W](t, x, y) dx dy \right| dt \\ &\leq 4 \|\eta\|_{C^0} \|\varphi\|_{C^0} \|\psi\|_{C^0} \int_0^T \| [W_{\mathcal{G}_n} - W](t, \cdot, \cdot) \|_{\square} dt, \end{aligned}$$

where the last step is a consequence of the properties (B.2) of the cut norm recalled in Appendix B. By (1.16), this yields (2.5) and concludes the proof of the lemma.  $\square$

**Lemma 2.5.** CASE 2. Assume (1.16) and (1.18). Then the limit graphon  $W$  satisfies (1.13) with  $K_d = K_1$ .

*Proof.* For a.e.  $t \in [0, T]$ , we have  $W(t, \cdot, \cdot) \in L^1([0, 1]^2)$  and, owing to (1.16) and up to subsequences,

$$\lim_{n \rightarrow +\infty} \|W_{\mathcal{G}_n}(t, \cdot, \cdot) - W(t, \cdot, \cdot)\|_{\square} = 0. \quad (2.6)$$

For any such  $t$ , we fix  $\hat{x} \in ]0, 1[$  Lebesgue point for the map  $x \mapsto \int_0^1 W(t, x, y) dy$ , which belongs to  $L^1(0, 1)$ . We fix  $\varepsilon < \min\{\hat{x}, 1 - \hat{x}\}$  and point out that, by combining (2.6) with the definition of cut norm (1.15) we have

$$\int_{\hat{x}-\varepsilon}^{\hat{x}+\varepsilon} \int_0^1 W(t, x, y) dy dx = \lim_{n \rightarrow +\infty} \int_{\hat{x}-\varepsilon}^{\hat{x}+\varepsilon} \int_0^1 W_{\mathcal{G}_n}(t, x, y) dy dx \stackrel{(1.18)}{\leq} 2K_1\varepsilon, \quad (2.7)$$

which in turn implies owing to Lebesgue's Differentiation Theorem

$$\int_0^1 W(t, \hat{x}, y) dy = \lim_{\varepsilon \rightarrow 0^+} \frac{1}{2\varepsilon} \int_{\hat{x}-\varepsilon}^{\hat{x}+\varepsilon} \int_0^1 W(t, x, y) dy dx \stackrel{(2.7)}{\leq} K_1$$

and this yields (1.13) with  $K_d = K_1$ .  $\square$

**2.4. A classical approximation result.** We now recall a standard approximation result that will be often used in the next sections. For the sake of completeness, we briefly sketch its proof.

**Lemma 2.6.** Assume  $u \in L^1(0, 1)$ ,  $U \in L^1([0, 1]^2)$ , recall (1.6) and (1.21), set

$$u_j^n = \langle u \rangle_{I_j^n}, \quad A_{jk}^n = \langle U \rangle_{I_j^n \times I_k^n}$$

and use the notation (1.7) and (1.8). Then

$$u^n \rightarrow u \text{ strongly in } L^1(0, 1), \quad W_{\mathcal{G}_n} \rightarrow U \text{ strongly in } L^1([0, 1]^2). \quad (2.8)$$

If furthermore  $u \in L^2(0, 1)$ ,  $U \in L^2([0, 1]^2)$ , then

$$u^n \rightarrow u \text{ strongly in } L^2(0, 1), \quad W_{\mathcal{G}_n} \rightarrow U \text{ strongly in } L^2([0, 1]^2). \quad (2.9)$$

*Proof.* We only provide the proof of (2.9), the proof of (2.8) being entirely analogous and slightly easier.

STEP 1: we first establish (2.9) for  $u \in C^0([0, 1])$ . For fixed  $\varepsilon > 0$ , by the uniform continuity of  $u$  there is  $n_\varepsilon$  such that, for every  $n \geq n_\varepsilon$ , if  $|x - y| \leq n^{-1}$  then  $|u(x) - u(y)| \leq \varepsilon$ . This implies that, for every  $j = 1, \dots, n$  and  $x \in I_j^n$ , we have

$$|u(x) - u_j^n| = \left| u(x) - n \int_{I_j^n} u(y) dy \right| = \left| n \int_{I_j^n} [u(x) - u(y)] dy \right| \leq n \int_{I_j^n} |u(x) - u(y)| dy \stackrel{n \geq n_\varepsilon}{\leq} n \int_{I_j^n} \varepsilon dy = \varepsilon.$$

This in turn yields  $\|u - u^n\|_2 \leq \varepsilon$  and by the arbitrariness of  $\varepsilon$  we get the desired convergence result.

STEP 2: we consider the general case. For any  $u, v \in L^2(0, 1)$  we term  $u^n$  and  $v^n$  the corresponding piecewise constant approximation, so that by using Jensen's inequality we get

$$\begin{aligned} \|u^n - v^n\|_2^2 &= \sum_{j=1}^n \int_{I_j^n} |u_j^n - v_j^n|^2 dx = \sum_{j=1}^n \int_{I_j^n} \left( n \int_{I_j^n} [u - v](y) dy \right)^2 dx \\ &\stackrel{\text{Jensen}}{\leq} \sum_{j=1}^n \int_{I_j^n} n \int_{I_j^n} [u - v]^2(y) dy dx = \sum_{j=1}^n \int_{I_j^n} [u - v]^2(y) dy = \|u - v\|_2^2. \end{aligned} \quad (2.10)$$

We now fix  $\varepsilon > 0$  and choose  $v \in C^0([0, 1])$  in such a way that  $\|u - v\|_2 \leq \varepsilon$ . We then have

$$\|u - u^n\|_2 \leq \|u - v\|_2 + \|v - v^n\|_2 + \|v^n - u^n\|_2 \stackrel{(2.10)}{\leq} 2\|u - v\|_2 + \|v - v^n\|_2 \leq 2\varepsilon + \|v - v^n\|_2.$$

By using STEP 1, we can find  $n_\varepsilon$  such that if  $n \geq n_\varepsilon$  then  $\|v - v^n\|_2 \leq \varepsilon$  and by plugging this inequality into the above expression and using the arbitrariness of  $\varepsilon$  we obtain the desired convergence result.  $\square$

## 3. CONVERGENCE RESULTS FOR SIR MODEL ON GRAPHS

This section contains the proofs of Proposition 1.2 and Theorem 1.3. The exposition is organized as follows:

- we first establish in §3.1 the uniqueness part of Proposition 1.2, namely the uniqueness of the distributional solution to (1.9), (1.10);
- in §3.2 we prove Theorem 1.3. As a byproduct, the proof provides a constructive argument that yields the existence part of Proposition 1.2, contingent on showing a sequence of graphs satisfying (1.16) and either (1.17) or (1.18);
- in §3.3, we prove the existence part of Proposition 1.2 by exhibiting a family of graphs satisfying the above requirements. This concludes the proof Proposition 1.2.

**3.1. Proof of the uniqueness part of Proposition 1.2.** We fix two distributional solutions  $(s_1, i_1, r_1)$  and  $(s_2, i_2, r_2)$  satisfying (1.14) and we separately consider CASES 1 and 2 below.

CASE 1:  $W \in L^2([0, T] \times [0, 1]^2; \mathbb{R}_+)$ . We set  $u_1 := s_1 + i_1$ ,  $u_2 := s_2 + i_2$  and we recall that, owing to Lemma 2.3,  $(s_1, i_1, r_1)$  and  $(s_2, i_2, r_2)$  have Sobolev regularity and the equalities in (1.9) are satisfied pointwise almost everywhere. By using Hölder's and Young's inequalities we get

$$\begin{aligned}
\frac{d}{dt} \int_0^1 [s_1 - s_2]^2(t, x) dx &= 2 \int_0^1 [s_1 - s_2] [\partial_t s_1 - \partial_t s_2](t, x) dx \\
&\stackrel{(1.9)}{=} - 2 \underbrace{\int_0^1 [s_1 - s_2]^2(t, x) \int_0^1 \beta(t, y) W(t, x, y) i_1(t, y) dy dx}_{\geq 0} \\
&\quad - 2 \int_0^1 s_2(t, x) [s_1 - s_2](t, x) \int_0^1 \beta(t, y) W(t, x, y) [i_1 - i_2](t, y) dy dx \\
&\stackrel{\text{Hölder}}{\leq} 2 \|s_2\|_\infty \|\beta\|_\infty \|W(t, \cdot, \cdot)\|_{L^2([0, 1]^2)} \|s_1(t, \cdot) - s_2(t, \cdot)\|_{L^2(0, 1)} \|i_1(t, \cdot) - i_2(t, \cdot)\|_{L^2(0, 1)} \\
&\stackrel{(1.14), i_i = u_i - s_i, \text{Young}}{\leq} 2 \|\beta\|_\infty \|W(t, \cdot, \cdot)\|_{L^2([0, 1]^2)} \left[ \frac{3}{2} \|s_1(t, \cdot) - s_2(t, \cdot)\|_{L^2(0, 1)}^2 + \frac{1}{2} \|u_1(t, \cdot) - u_2(t, \cdot)\|_{L^2(0, 1)}^2 \right].
\end{aligned} \tag{3.1}$$

Next, we point out that

$$\partial_t [u_1 - u_2] = \gamma [i_1 - i_2] = \gamma [u_1 - u_2] - \gamma [s_1 - s_2] \tag{3.2}$$

and by using again Hölder's and Young's inequalities this yields

$$\begin{aligned}
\frac{d}{dt} \int_0^1 [u_1 - u_2]^2(t, x) dx &= \int_0^1 \gamma [u_1(t, x) - u_2(t, x)]^2 dx - \int_0^1 \gamma [s_1(t, x) - s_2(t, x)] [u_1(t, x) - u_2(t, x)] dx \\
&\stackrel{\text{Hölder, Young}}{\leq} \|\gamma\|_\infty \left[ \frac{3}{2} \|u_1(t, \cdot) - u_2(t, \cdot)\|_{L^2(0, 1)}^2 + \frac{1}{2} \|s_1(t, \cdot) - s_2(t, \cdot)\|_{L^2(0, 1)}^2 \right].
\end{aligned} \tag{3.3}$$

Combining (3.1) and (3.3) we get

$$\begin{aligned}
&\frac{d}{dt} \left[ \|u_1(t, \cdot) - u_2(t, \cdot)\|_{L^2(0, 1)}^2 + \|s_1(t, \cdot) - s_2(t, \cdot)\|_{L^2(0, 1)}^2 \right] \\
&\leq \frac{7}{2} \left[ \|\beta\|_\infty \|W(t, \cdot, \cdot)\|_{L^2([0, 1]^2)} + \|\gamma\|_\infty \right] \left[ \|s_1(t, \cdot) - s_2(t, \cdot)\|_{L^2(0, 1)}^2 + \|u_1(t, \cdot) - u_2(t, \cdot)\|_{L^2(0, 1)}^2 \right]
\end{aligned}$$

and by Grönwall Lemma this implies  $\|u_1(t, \cdot) - u_2(t, \cdot)\|_{L^2(0, 1)}^2 + \|s_1(t, \cdot) - s_2(t, \cdot)\|_{L^2(0, 1)}^2 = 0$  for every  $t \in [0, T]$ . Since  $i_1 = u_1 - s_1$ ,  $i_2 = u_2 - s_2$  this in turn implies  $s_1 = s_2$ ,  $i_1 = i_2$  a.e. on  $[0, T] \times [0, 1]$  and, since  $r_1 = 1 - s_1 - i_1$ ,  $r_2 = 1 - s_2 - i_2$ , it also implies  $r_1 = r_2$  a.e. on  $[0, T] \times [0, 1]$ .

CASE 2:  $W \in L^1([0, T] \times [0, 1]^2; \mathbb{R}_+)$  satisfies (1.13). As before we set  $u_1 := s_1 + i_1$ ,  $u_2 := s_2 + i_2$  and we get

$$\begin{aligned}
\frac{d}{dt}|s_1 - s_2|(t, x) &= \text{sign}[s_1 - s_2][\partial_t s_1 - \partial_t s_2](t, x) \\
&= \underbrace{-\text{sign}[s_1 - s_2][s_1 - s_2](t, x) \int_0^1 \beta(t, y)W(t, x, y)i_1(t, y)dy}_{\geq 0} \\
&\quad + \text{sign}[s_1 - s_2]s_2(t, x) \int_0^1 \beta(t, y)W(t, x, y)[i_1 - i_2](t, y)dy \\
&\stackrel{(1.13), (1.14), W \geq 0}{\leq} K_d \|\beta\|_\infty \| [i_1 - i_2](t, \cdot) \|_{L^\infty(0,1)} \\
&\stackrel{(1.13), (1.14)}{\leq} K_d \|\beta\|_\infty \left[ \| [s_1 - s_2](t, \cdot) \|_{L^\infty(0,1)} + \| [u_1 - u_2](t, \cdot) \|_{L^\infty(0,1)} \right].
\end{aligned} \tag{3.4}$$

By (3.2) we get

$$\frac{d}{dt} \| [u_1 - u_2](t, \cdot) \|_{L^\infty(0,1)} \leq \|\gamma\|_\infty \left[ \| [s_1 - s_2](t, \cdot) \|_{L^\infty(0,1)} + \| [u_1 - u_2](t, \cdot) \|_{L^\infty(0,1)} \right] \tag{3.5}$$

and combining (3.4) and (3.5) with Grönwall Lemma we get  $\| [s_1 - s_2](t, \cdot) \|_{L^\infty(0,1)} = 0$  and  $\| [u_1 - u_2](t, \cdot) \|_{L^\infty(0,1)} = 0$  for every  $t \in [0, T]$ . Arguing as in the previous case this in turn implies  $s_1 = s_2$ ,  $i_1 = i_2$ ,  $r_1 = r_2$  a.e. on  $[0, T] \times [0, 1]$ .

**3.2. Proof of Theorem 1.3.** We now provide the proof of Theorem 1.3. Let  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$ ,  $\{\beta^n\}_{n \in \mathbb{N}}$  and  $\{\gamma^n\}_{n \in \mathbb{N}}$  be fixed as in the statement of the theorem. Recalling the notation (1.8) we observe that  $(s^n, i^n, r^n)$  is a distributional solution of

$$\begin{cases} \partial_t s^n(t, x) = -s^n(t, x) \int_0^1 \beta^n(t, y)W_{\mathcal{G}_n}(t, x, y)i^n(t, y) dy \\ \partial_t i^n(t, x) = s^n(t, x) \int_0^1 \beta^n(t, y)W_{\mathcal{G}_n}(t, x, y)i^n(t, y) dy - \gamma^n(t, x)i^n(t, x) & x \in [0, 1] \\ \partial_t r^n(t, x) = \gamma^n(t, x)i^n(t, x) \end{cases} \tag{3.6}$$

and satisfies the initial condition  $s^n(0, \cdot) = s_0^n$ ,  $i^n(0, \cdot) = i_0^n$ ,  $r^n(0, \cdot) = r_0^n$ . This implies that for every  $(\varphi, \psi, \eta) \in (C_c^\infty([0, T] \times [0, 1]))^3$  we have

$$\begin{aligned}
&\int_0^T \int_0^1 s^n(t, x) \left( \partial_t \varphi(t, x) - \varphi(t, x) \int_0^1 \beta^n(t, y)W_{\mathcal{G}_n}(t, x, y)i^n(t, y) dy \right) dx dt = \int_0^1 s_0^n(x) \varphi(0, x) dx \\
&\int_0^T \int_0^1 i^n(t, x) (\partial_t \psi(t, x) - \gamma^n(t, x) \psi(t, x)) + s^n(t, x) \psi(t, x) \int_0^1 \beta^n(t, y)W_{\mathcal{G}_n}(t, x, y)i^n(t, y) dy dx dt \\
&\quad = \int_0^1 i_0^n(x) \psi(0, x) dx \\
&\int_0^T \int_0^1 r^n(t, x) \partial_t \eta(t, x) + \gamma^n(t, x) \eta(t, x) i^n(t, x) dx dt = \int_0^1 r_0^n(x) \eta(0, x) dx.
\end{aligned} \tag{3.7}$$

We recall the bounds (1.5), which owing to the notation (1.8) imply

$$\|s^n\|_\infty, \|i^n\|_\infty, \|r^n\|_\infty \leq 1 \tag{3.8}$$

and we conclude that, up to subsequences,

$$s^n \rightharpoonup^* s, \quad i^n \rightharpoonup^* i, \quad r^n \rightharpoonup^* r \quad \text{weakly}^* \text{ in } L^\infty([0, T] \times [0, 1]), \tag{3.9}$$

for some limit function  $(s, i, r) \in L^\infty([0, T] \times [0, 1]; \mathbb{R}^3)$ . Also,

$$s_0^n \rightharpoonup^* s_0, \quad i_0^n \rightharpoonup^* i_0, \quad r_0^n \rightharpoonup^* r_0 \quad \text{weakly}^* \text{ in } L^\infty(0, 1), \tag{3.10}$$

for some limit function  $(s_0, i_0, r_0) \in L^\infty([0, 1]; \mathbb{R}^3)$ . We now pass to the limit in (3.7) by separately considering the two cases.

CASE 1: we assume (1.17) and proceed according to the following steps.

STEP 1A: since  $(s^n, i^n, r^n)$  is a distributional solution of (3.6), by (2.2) and recalling (1.17) and (1.19) we conclude that

$$\|\partial_t s^n\|_2, \|\partial_t i^n\|_2, \|\partial_t r^n\|_2 \leq C(K_0, M). \quad (3.11)$$

We now show that this implies that, up to subsequences, for every  $\psi \in L^2([0, T] \times [0, 1]^2)$

$$\int_0^T \int_0^1 \int_0^1 \psi(t, x, y) s^n(t, x) i^n(t, y) dx dy dt \rightarrow \int_0^T \int_0^1 \int_0^1 \psi(t, x, y) s(t, x) i(t, y) dx dy dt, \quad (3.12)$$

that is the product  $s^n(t, x) i^n(t, y)$  weakly converges to  $s(t, x) i(t, y)$  in  $L^2([0, T] \times [0, 1]^2)$ . To this end, we point out that for every fixed  $n \in \mathbb{N}$ , since  $\partial_t s^n, \partial_t i^n, \partial_t r^n$  are all  $L^2$  functions owing to (3.11), we can select a representative of  $(s^n, i^n, r^n)$  (which a priori as  $L^\infty$  functions are only defined up to negligible sets) such that the map  $t \mapsto (s^n, i^n, r^n)(t, \cdot)$  is continuous from  $[0, T]$  to  $L^2(0, 1)$  endowed with the strong topology. In the following we will always work with this representative, which allows us to give a meaning to the value  $(s^n, i^n, r^n)(t, \cdot) \in L^2(0, 1)$  for every  $t \in [0, T]$ .

Using the  $L^\infty$  bounds and a standard diagonal argument we can extract a further subsequence such that

$$s^n(q, \cdot) \rightharpoonup^* s(q, \cdot), \quad i^n(q, \cdot) \rightharpoonup^* i(q, \cdot), \quad r^n(q, \cdot) \rightharpoonup^* r(q, \cdot) \text{ weakly}^* \text{ in } L^\infty(0, 1), \text{ for every } q \in [0, T] \cap \mathbb{Q}. \quad (3.13)$$

Next, we fix  $t \in [0, T]$  and a test function  $\psi \in L^2([0, T] \times [0, 1]^2)$ . We also fix a sequence of rational numbers  $q_k \rightarrow t$  as  $k \rightarrow +\infty$ . We then have

$$\begin{aligned} & \left| \iint_{[0,1]^2} \psi(t, x, y) s^n(t, x) i^n(t, y) dx dy - \iint_{[0,1]^2} \psi(t, x, y) s(t, x) i(t, y) dx dy \right| \\ & \leq \underbrace{\left| \iint_{[0,1]^2} \psi(t, x, y) s^n(t, x) i^n(t, y) dx dy - \iint_{[0,1]^2} \psi(t, x, y) s^n(q_k, x) i^n(t, y) dx dy \right|}_{E_k^n} \\ & \quad + \underbrace{\left| \iint_{[0,1]^2} \psi(t, x, y) s^n(q_k, x) i^n(t, y) dx dy - \iint_{[0,1]^2} \psi(t, x, y) s^n(q_k, x) i^n(q_k, y) dx dy \right|}_{B_k^n} \\ & \quad + \underbrace{\left| \iint_{[0,1]^2} \psi(t, x, y) s^n(q_k, x) i^n(q_k, y) dx dy - \iint_{[0,1]^2} \psi(t, x, y) s(q_k, x) i(q_k, y) dx dy \right|}_{C_k^n} \\ & \quad + \underbrace{\left| \iint_{[0,1]^2} \psi(t, x, y) s(q_k, x) i(q_k, y) dx dy - \iint_{[0,1]^2} \psi(t, x, y) s(t, x) i(t, y) dx dy \right|}_{D_k^n}. \end{aligned} \quad (3.14)$$

We now estimate each of the above terms separately. Note that

$$E_k^n \leq \|\psi(t, \cdot, \cdot)\|_{L^2([0,1]^2)} \underbrace{\|i^n\|_\infty}_{\leq 1} \|s^n(t, \cdot) - s^n(q_k, \cdot)\|_{L^2(0,1)} \quad (3.15)$$

and, owing to Jensen's inequality,

$$\begin{aligned} \|s^n(t, \cdot) - s^n(q_k, \cdot)\|_{L^2(0,1)}^2 &= \int_0^1 [s^n(t, \cdot) - s^n(q_k, \cdot)]^2 dx = \int_0^1 \left[ \int_{q_k}^t \partial_\tau s^n(\tau, \cdot) d\tau \right]^2 dx \\ &\stackrel{\text{Jensen}}{\leq} |q_k - t| \int_0^1 \int_{q_k}^t (\partial_\tau s^n(\tau, \cdot))^2 d\tau dx \stackrel{(3.11)}{\leq} C(K_0, M) |q_k - t|, \end{aligned} \quad (3.16)$$

so that plugging the above inequality into (3.15) we get

$$E_k^n \leq C(K_0, M) \|\psi(t, \cdot, \cdot)\|_{L^2([0,1]^2)} \sqrt{|q_k - t|}, \quad \text{for every } n.$$

By an analogous argument, we get  $B_k^n \leq C(K_0, M) \|\psi(t, \cdot, \cdot)\|_{L^2([0,1]^2)} \sqrt{|q_k - t|}$ .

Owing to (3.11), we have

$$\partial_t s^n \rightharpoonup \partial_t s, \quad \partial_t i^n \rightharpoonup \partial_t i \quad \text{weakly in } L^2([0, T] \times [0, 1]),$$

and by recalling (3.16) and using the lower semicontinuity of the norm with respect to the weak convergence we get

$$\int_0^1 \int_{q_k}^t (\partial_\tau s(\tau, \cdot))^2 d\tau dx \leq C(K_0, M), \quad \int_0^1 \int_{q_k}^t (\partial_\tau i(\tau, \cdot))^2 d\tau dx \leq C(K_0, M).$$

Hence, arguing as before this yields  $D_k^n \leq C(K_0, M) \|\psi(t, \cdot, \cdot)\|_{L^2([0,1]^2)} \sqrt{|q_k - t|}$ , for every  $n \in \mathbb{N}$ . To conclude, we fix  $\varepsilon > 0$  and choose  $k$  in such a way that  $E_k^n + B_k^n + D_k^n \leq \varepsilon$ , for every  $n$ . Next, we recall (3.13) and choose  $n_\varepsilon$  in such a way that  $C_k^n \leq \varepsilon$  for every  $n \geq n_\varepsilon$ . By the arbitrariness of  $\varepsilon$ , this implies that the left hand side of (3.14) vanishes in the  $n \rightarrow +\infty$  limit for a.e.  $t \in [0, T]$ . By Lebesgue's Dominated Convergence Theorem this yields (3.12).

STEP 1B: we show that the limit  $(s, i, r)$  in (3.9) is a distributional solution of (1.9), (1.10). We subtract (1.12) from (3.7) and in particular we get

$$\begin{aligned} & \left| \int_0^T \int_0^1 \left( s^n(t, x) \int_0^1 \beta^n(t, y) W_{\mathcal{G}_n}(t, x, y) i^n(t, y) dy - s(t, x) \int_0^1 \beta(t, y) W(t, x, y) i(t, y) dy \right) \varphi(t, x) dx dt \right| \\ & \leq \underbrace{\left| \int_0^T \int_0^1 \int_0^1 (W_{\mathcal{G}_n}(t, x, y) - W(t, x, y)) \beta^n(t, y) s^n(t, x) i^n(t, y) \varphi(t, x) dy dx dt \right|}_{I_1} \\ & + \underbrace{\left| \int_0^T \int_0^1 \int_0^1 W(t, x, y) (\beta^n(t, y) s^n(t, x) i^n(t, y) - \beta(t, y) s(t, x) i(t, y)) \varphi(t, x) dy dx dt \right|}_{I_2} \end{aligned} \quad (3.17)$$

for every  $\varphi \in C_c^\infty([0, T] \times [0, 1]^2)$ . We first control  $I_2$ :

$$\begin{aligned} I_2 & \leq \underbrace{\left| \int_0^T \int_0^1 \int_0^1 W(t, x, y) [\beta^n(t, y) - \beta(t, y)] s^n(t, x) i^n(t, y) \varphi(t, x) dy dx dt \right|}_{:=I_{21}} \\ & + \underbrace{\left| \int_0^T \int_0^1 \int_0^1 W(t, x, y) \beta(t, y) [s^n(t, x) i^n(t, y) - s(t, x) i(t, y)] \varphi(t, x) dy dx dt \right|}_{:=I_{22}}. \end{aligned}$$

To see that  $I_{22}$  vanishes in the  $n \rightarrow +\infty$  limit it suffices to point out that  $W\beta \in L^2([0, T] \times [0, 1]^2)$  and then use (3.12). To deal with  $I_{21}$  we observe that

$$\begin{aligned} I_{21} & \stackrel{\|s_n\|_\infty, \|i_n\|_\infty \leq 1}{\leq} \int_0^T \int_0^1 \int_0^1 W(t, x, y) |\beta^n(t, y) - \beta(t, y)| dy dx dt \\ & \stackrel{\text{H\"older}}{\leq} \|W\|_{L^2([0, T] \times [0, 1]^2)} \left( \int_0^T \int_0^1 |\beta^n(t, y) - \beta(t, y)|^2 dy dt \right)^{1/2} \\ & \stackrel{(1.17)}{\leq} K_0 \left( \int_0^T \int_0^1 |\beta^n(t, y) - \beta(t, y)|^2 dy dt \right)^{1/2} \stackrel{(1.19)}{\leq} \sqrt{2MK_0} \left( \int_0^T \int_0^1 |\beta^n(t, y) - \beta(t, y)| dy dt \right)^{1/2} \end{aligned}$$

and by (1.19) the right-hand side of the above equation vanishes in the  $n \rightarrow +\infty$  limit. Going back to (3.17), this shows that  $I_2$  converges to 0. Let us focus now on  $I_1$ . We set

$$L := \|\beta^n\|_\infty \|s^n\|_\infty \|i^n\|_\infty \|\varphi\|_\infty \stackrel{\|s^n\|_\infty, \|i^n\|_\infty \leq 1}{\leq} \|\beta\|_\infty \|\varphi\|_\infty \stackrel{(1.23)}{\leq} M \|\varphi\|_\infty.$$

By multiplying and dividing  $I_1$  by  $L$  and making use of (B.2) in Appendix B, we get

$$\begin{aligned} I_1 & = L \left| \int_0^T \int_0^1 \int_0^1 (W_{\mathcal{G}_n}(t, x, y) - W(t, x, y)) \cdot \frac{s^n(t, x)}{\|s^n\|_\infty} \frac{\varphi(t, x)}{\|\varphi\|_\infty} \frac{\beta^n(t, y)}{\|\beta^n\|_\infty} \frac{i^n(t, y)}{\|i^n\|_\infty} dy dx dt \right| \\ & \leq L \int_0^T \sup_{\substack{\|f(t, \cdot)\|_\infty \leq 1 \\ \|g(t, \cdot)\|_\infty \leq 1}} \left| \int_{[0, 1]^2} (W_{\mathcal{G}_n}(t, x, y) - W(t, x, y)) f(t, x) g(t, y) dy dx \right| dt \\ & \stackrel{(B.2)}{\leq} 4L \int_0^T \|W_{\mathcal{G}_n} - W\|_{\square} dt \stackrel{(1.16)}{\rightarrow} 0 \quad \text{as } n \rightarrow +\infty. \end{aligned}$$

This implies that  $I_1 \rightarrow 0$  as  $n \rightarrow +\infty$  and hence that the left hand side of (3.17) vanishes in the  $n \rightarrow +\infty$  limit. By relying on similar arguments, one can pass to the limit in all the other terms in (3.7) and show that  $(s, i, r)$  is a distributional solution of (1.9), (1.10).

CASE 2: we assume the family of graphs satisfy (1.18). The proof follows the same argument as in the previous case, so we only provide a sketch and highlight the points where there is some difference. By using the proof of Lemma 2.3 and recalling (1.19) we conclude that

$$\|\partial_t s^n\|_\infty, \|\partial_t i^n\|_\infty, \|\partial_t r^n\|_\infty \leq C(K_1, M),$$

which implies that there is a representative of  $(s, i, r)$  such that the maps  $t \mapsto s(t, \cdot)$ ,  $t \mapsto i(t, \cdot)$ ,  $t \mapsto r(t, \cdot)$  are continuous from  $[0, T]$  in  $L^\infty(0, 1)$  endowed with the ess sup (strong) topology. The goal is now to show that this implies that, up to subsequences, for every  $\psi \in L^1([0, T] \times [0, 1]^2)$

$$\int_0^T \int_0^1 \int_0^1 \psi(t, x, y) s^n(t, x) i^n(t, y) dx dy dt \rightarrow \int_0^T \int_0^1 \int_0^1 \psi(t, x, y) s(t, x) i(t, y) dx dy dt. \quad (3.18)$$

The proof follows the same lines as the proof of (3.12): up to subsequences, we have (3.13) and hence, for a given  $t \in [0, T]$ , we have the splitting as in (3.14), where now

$$E_k^n \leq \|\psi(t, \cdot, \cdot)\|_{L^1([0, 1]^2)} \|s^n(t, \cdot) - s^n(q_k, \cdot)\|_{L^\infty(0, 1)}$$

and by using the continuity of the map  $t \mapsto s(t, \cdot)$  from  $[0, T] \rightarrow L^\infty(0, 1)$  we conclude that  $\|s^n(t, \cdot) - s^n(q_k, \cdot)\|_{L^\infty(0, 1)}$  is arbitrarily small, provided  $q_k$  is sufficiently close to  $t$ . By reasoning similarly we can show that the other terms on the right hand side of (3.14) are arbitrarily small and hence establish (3.18). The rest of the proof in CASE 2 works as in CASE 1.

*Remark 3.1.* In system (1.1) the term  $s_j^n$  only interacts with  $n^{-1} \sum_{k=1}^n \beta_k^n(t) A_{jk}^n i_k^n$ . This allows us to pass to the limit in the integral term in the first equation of (1.9). From the point of view of applications it would be also meaningful to consider the following SIR system:

$$\begin{cases} \frac{ds_j(t)}{dt} = -s_j(t) \frac{1}{n} \sum_{k=1}^n \beta_k^n(t) A_{jk}^n i_k(t) - \beta_j^n(t) s_j^n(t) i_j^n(t) \\ \frac{di_j(t)}{dt} = s_j(t) \frac{1}{n} \sum_{k=1}^n \beta_k^n(t) A_{jk}^n i_k(t) + \beta_j^n(t) s_j^n(t) i_j^n(t) - \gamma_j^n(t) i_j(t) \\ \frac{dr_j(t)}{dt} = \gamma_j^n(t) i_j(t) \end{cases} \quad j = 1, \dots, n \quad (3.19)$$

where the interaction of the node  $j$  with itself is not scaled by the factor  $n^{-1}$  since it is somehow independent from the size of the network. The proof of Theorem 1.3 does not directly extend to (3.19) because from the weak\* convergence of  $s^n$  and  $i^n$  alone we cannot infer that the product  $s^n(t, x) i^n(t, x)$  converges to  $s(t, x) i(t, x)$ . As a matter of fact, the argument in the proof of Theorem 1.3 shows that there is  $\pi \in L^\infty([0, T] \times [0, 1])$ ,  $0 \leq \pi \leq 1$ , such that  $(s^n, i^n, r^n)$  converge weakly\* in  $L^\infty([0, T] \times [0, 1])$  to a distributional solution of

$$\begin{cases} \partial_t s(t, x) = -s(t, x) \int_0^1 \beta(t, y) W(t, x, y) i(t, y) dy - \beta(t, x) \pi(t, x) \\ \partial_t i(t, x) = s(t, x) \int_0^1 \beta(t, y) W(t, x, y) i(t, y) dy + \beta(t, x) \pi(t, x) - \gamma(t, x) i(t, x) \\ \partial_t r(t, x) = \gamma(t, x) i(t, x) \end{cases} \quad x \in [0, 1]$$

satisfying the initial condition (1.3).

**3.3. Existence of distributional solution of (1.9).** We can now provide the proof of the existence part of Proposition 1.2. Let  $W, T, \gamma, \beta, s_0, i_0, r_0$  be as in the statement of Proposition 1.2. Owing to the proof of Theorem 1.3, to establish the existence part of Proposition 1.2 it suffices to exhibit sequences of graphs  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$ , coefficients  $\{\beta^n\}_{n \in \mathbb{N}}$  and  $\{\gamma^n\}_{n \in \mathbb{N}}$  and initial data  $\{(s_0^n, i_0^n, r_0^n)\}_{n \in \mathbb{N}}$  such that (1.16), (1.19), (3.10) and either (1.17) or (1.18) are satisfied.

We define  $\beta^n, \gamma^n$  and  $(s_0^n, i_0^n, r_0^n)$  by using (1.21), (1.23) and the notation (1.8). Note that, if  $\beta_k^n$  and  $\gamma_k^n$  are as in (1.23), then

$$|\beta_k^n(t)| \leq \|\beta\|_\infty, \quad |\gamma_k^n(t)| \leq \|\gamma\|_\infty, \quad \text{for every } k = 1, \dots, n \text{ and a.e. } t \in [0, T] \quad (3.20)$$

and, by using (2.8),

$$\lim_{n \rightarrow +\infty} \|\gamma(t, \cdot) - \gamma^n(t, \cdot)\|_{L^1(0,1)} = 0, \quad \lim_{n \rightarrow +\infty} \|\beta(t, \cdot) - \beta^n(t, \cdot)\|_{L^1(0,1)} = 0 \quad \text{for a.e. } t \in [0, T].$$

Owing to Lebesgue's Dominated Convergence Theorem and to (3.20) this yields

$$\lim_{n \rightarrow +\infty} \int_0^T \|\beta(s, \cdot) - \beta^n(s, \cdot)\|_{L^1(0,1)} ds = 0, \quad \lim_{n \rightarrow +\infty} \int_0^T \|\gamma(s, \cdot) - \gamma^n(s, \cdot)\|_{L^1(0,1)} ds = 0. \quad (3.21)$$

This implies that both conditions in (1.19) are satisfied. To establish (3.10) we use again (2.8). We now verify the assumptions on  $W_{\mathcal{G}_n}$ . If  $W \in L^2([0, T] \times [0, 1]^2; \mathbb{R}_+)$ , then by Jensen's inequality

$$\begin{aligned} \|W_{\mathcal{G}_n}(t, \cdot, \cdot)\|_{L^2([0,1]^2)}^2 &= \sum_{j,k=1}^n \iint_{I_j^n \times I_k^n} (A_{jk}^n)^2 dx dy \stackrel{(1.22)}{=} \sum_{j,k=1}^n \iint_{I_j^n \times I_k^n} \langle W(t, \cdot, \cdot) \rangle_{I_j^n \times I_k^n}^2 dx dy \\ &\stackrel{\text{Jensen}}{\leq} \sum_{j,k=1}^n \iint_{I_j^n \times I_k^n} \langle W^2(t, \cdot, \cdot) \rangle_{I_j^n \times I_k^n} dx dy = \sum_{j,k=1}^n \iint_{I_j^n \times I_k^n} W^2(t, x, y) dx dy = \|W(t, \cdot, \cdot)\|_{L^2([0,1]^2)}^2 \end{aligned} \quad (3.22)$$

for a.e.  $t \in [0, T]$ , and this yields (1.17) with  $K_0 = \|W\|_2$ . If  $W \in L^1([0, T] \times [0, 1]^2; \mathbb{R}_+)$  satisfies (1.13), we have

$$\begin{aligned} \operatorname{ess\,sup}_{t \in [0, T], x \in [0, 1]} \int_0^1 W_{\mathcal{G}_n}(t, x, y) dy &\stackrel{(1.22)}{=} \operatorname{ess\,sup}_{t \in [0, T], j=1, \dots, n} \sum_{k=1}^n \int_{I_k^n} \langle W(t, \cdot, \cdot) \rangle_{I_j^n \times I_k^n} dy \\ &= \operatorname{ess\,sup}_{t \in [0, T], j=1, \dots, n} \sum_{k=1}^n n \iint_{I_j^n \times I_k^n} W(t, x, y) dx dy = \operatorname{ess\,sup}_{t \in [0, T], j=1, \dots, n} n \int_j^{j+1} \int_0^1 W(t, x, y) dy dx \\ &\stackrel{(1.13)}{\leq} \operatorname{ess\,sup}_{t \in [0, T], j=1, \dots, n} n \int_{I_j^n} K_d dx = K_d, \end{aligned} \quad (3.23)$$

which yields (1.18) with  $K_1 = K_d$ . We are left to verify (1.16). To this end, we recall that, owing to (B.2) in Appendix B,

$$\|U(t, \cdot, \cdot)\|_{\square} \leq \|U(t, \cdot, \cdot)\|_{L^1([0,1]^2)} \quad \text{for a.e. } t \in [0, T] \quad (3.24)$$

for every time-dependent graphon  $U \in L^1([0, T] \times [0, 1]^2)$ . Next, we combine (2.8) and either (3.22) or (3.23) with Lebesgue's Dominated Convergence Theorem to get

$$\lim_{n \rightarrow +\infty} \int_0^T \|W_{\mathcal{G}_n}(s, \cdot, \cdot) - W(s, \cdot, \cdot)\|_{L^1([0,1]^2)} ds = 0,$$

which owing to (3.24) yields (1.16).

#### 4. DETERMINISTIC APPROXIMATION TO SIR ON GRAPHONS

This section presents the convergence analysis for the deterministic approximations to (1.9), i.e. the proof of Theorem 1.4. We consider each case of the theorem separately.

**4.1. Proof of Theorem 1.4 (CASE 1).** We first prove the following a priori estimate on the  $L^2$ -error of the approximation at each fixed time.

**Lemma 4.1.** *Under the same assumptions as in the statement of Theorem 1.4 (CASE 1) we have*

$$\begin{aligned} &\|s(t, \cdot) - s^n(t, \cdot)\|_{L^2(0,1)}^2 + \|i(t, \cdot) - i^n(t, \cdot)\|_{L^2(0,1)}^2 + \|r(t, \cdot) - r^n(t, \cdot)\|_{L^2(0,1)}^2 \\ &\leq B \left( \|s_0 - s_0^n\|_2^2 + \|i_0 - i_0^n\|_2^2 + \|r_0 - r_0^n\|_2^2 \right) \\ &\quad + B \int_0^t \left( \|\beta W - \beta^n W_{\mathcal{G}_n}\|_{L^2([0,1]^2)}^2 + \|\gamma(s, \cdot) - \gamma^n(s, \cdot)\|_{L^2(0,1)}^2 \right) ds \end{aligned} \quad (4.1)$$

for every  $t \in [0, T]$ , where  $B$  is a suitable constant only depending on  $T, \|\beta\|_\infty, \|\gamma\|_\infty$  and  $\|W\|_2$ .

*Proof.* We rely on a Grönwall type argument and proceed according to the following steps.

STEP 1: we subtract the first line of (3.6) from the first line of (1.9) and multiply the result by  $2[s - s^n]$ . This yields

$$\begin{aligned} \partial_t [s - s^n]^2(t, x) &= \underbrace{-2[s - s^n]^2(t, x) \int_0^1 \beta(t, y) W(t, x, y) i(t, y) dy}_{J_1(t, x)} \\ &\quad - \underbrace{2s^n(t, x)[s - s^n](t, x) \int_0^1 (\beta(t, y) W(t, x, y) i(t, y) - \beta^n(t, y) W_{\mathcal{G}_n}(t, x, y) i^n(t, y)) dy}_{J_2(t, x)}. \end{aligned} \quad (4.2)$$

Recalling that  $\beta, W \geq 0$  by assumption and that  $i \geq 0$  owing to (1.14) we have

$$J_1(t, x) \leq 0 \quad \text{for a.e. } (t, x) \in [0, T] \times [0, 1]. \quad (4.3)$$

To control  $J_2$ , we split it as

$$\begin{aligned} |J_2(t, x)| &\leq 2 \underbrace{\left| s^n(t, x)[s - s^n](t, x) \int_0^1 \beta(t, y) W(t, x, y) [i - i^n](t, y) dy \right|}_{J_{21}} \\ &\quad + \underbrace{\left| 2s^n(t, x)[s - s^n](t, x) \int_0^1 i^n(t, y) [\beta(t, y) W(t, x, y) - \beta^n(t, y) W_{\mathcal{G}_n}(t, x, y)] dy \right|}_{J_{22}}. \end{aligned} \quad (4.4)$$

By using Hölder's and Young's inequalities we get

$$\begin{aligned} \int_0^1 J_{21}(t, x) dx &\stackrel{(3.8)}{\leq} 2 \int_0^1 |s - s^n|(t, x) \int_0^1 \beta(t, y) W(t, x, y) |i - i^n|(t, y) dy dx \\ &\stackrel{\text{Hölder}}{\leq} 2 \|\beta\|_\infty \| [s - s^n](t, \cdot) \|_{L^2(0,1)} \| [i - i^n](t, \cdot) \|_{L^2(0,1)} \| W(t, \cdot, \cdot) \|_{L^2([0,1]^2)} \\ &\stackrel{\text{Young}}{\leq} \|\beta\|_\infty \| W(t, \cdot, \cdot) \|_{L^2([0,1]^2)} \left[ \| [s - s^n](t, \cdot) \|_{L^2(0,1)}^2 + \| [i - i^n](t, \cdot) \|_{L^2(0,1)}^2 \right]. \end{aligned} \quad (4.5)$$

To control  $J_{22}$ , we point out that

$$\begin{aligned} &\int_0^1 i^n(t, y) \left| \beta(t, y) W(t, x, y) - \beta^n(t, y) W_{\mathcal{G}_n}(t, x, y) \right| dy \\ &\leq \| i^n(t, \cdot) \|_{L^2(0,1)} \| [\beta W - \beta^n W_{\mathcal{G}_n}](t, x, \cdot) \|_{L^2(0,1)} \stackrel{(3.8)}{\leq} \| [\beta W - \beta^n W_{\mathcal{G}_n}](t, x, \cdot) \|_{L^2(0,1)}, \end{aligned}$$

and combining (3.8) with Young's Inequality we get

$$\int_0^1 J_{22}(t, x) dx \leq \| s - s^n \|_{L^2(0,1)}^2 + \| [\beta W - \beta^n W_{\mathcal{G}_n}](t, \cdot, \cdot) \|_{L^2([0,1]^2)}^2. \quad (4.6)$$

Coupling (4.2), (4.3), (4.4), (4.5), (4.6) we get

$$\begin{aligned} \frac{d}{dt} \int_0^1 |s - s^n|^2(t, x) dx &\leq \|\beta\|_\infty \| W(t, \cdot, \cdot) \|_{L^2([0,1]^2)} \left[ \| [s - s^n](t, \cdot) \|_{L^2(0,1)}^2 + \| [i - i^n](t, \cdot) \|_{L^2(0,1)}^2 \right] \\ &\quad + \| [s - s^n](t, \cdot) \|_{L^2(0,1)}^2 + \| [\beta W - \beta^n W_{\mathcal{G}_n}](t, \cdot, \cdot) \|_{L^2([0,1]^2)}^2. \end{aligned} \quad (4.7)$$

STEP 2: we set  $v := s + i$ ,  $v^n := s^n + i^n$  and point out that owing to (3.6) we have

$$\partial_t [v - v^n] = -\gamma i + \gamma^n i^n = -\gamma [v - s] + \gamma^n [v^n - s^n] = -\gamma [v - v^n] + \gamma [s - s^n] - v^n [\gamma - \gamma^n] + s^n [\gamma - \gamma^n].$$

This yields

$$\partial_t [v - v^n]^2 = \underbrace{-2\gamma [v - v^n]^2}_{\leq 0} + 2\gamma [v - v^n][s - s^n] - 2v^n [v - v^n][\gamma - \gamma^n] + 2s^n [v - v^n][\gamma - \gamma^n]$$

and by integrating, recalling (3.8), noting that  $0 \leq v, v^n \leq 1$  and using Hölder's and Young's inequalities we get

$$\begin{aligned} \frac{d}{dt} \int_0^1 [v - v^n]^2(t, x) dx &\leq \|\gamma\|_\infty \left[ \|[v - v^n](t, \cdot)\|_{L^2(0,1)}^2 + \|[s - s^n](t, \cdot)\|_{L^2(0,1)}^2 \right] \\ &\quad + 2\|[v - v^n](t, \cdot)\|_{L^2(0,1)}^2 + 2\|[\gamma - \gamma^n](t, \cdot)\|_{L^2(0,1)}^2. \end{aligned} \quad (4.8)$$

STEP 3: by combining (4.7) and (4.8) and pointing out that  $\|[i - i^n](t, \cdot)\|_{L^2(0,1)}^2 \leq 2(\|[v - v^n](t, \cdot)\|_{L^2(0,1)}^2 + \|[s - s^n](t, \cdot)\|_{L^2(0,1)}^2)$  we get

$$\begin{aligned} \frac{d}{dt} &\left[ \|[s - s^n](t, \cdot)\|_{L^2(0,1)}^2 + \|[v - v^n](t, \cdot)\|_{L^2(0,1)}^2 \right] \\ &\leq \|[\beta W - \beta^n W_{\mathcal{G}_n}](t, \cdot, \cdot)\|_{L^2([0,1]^2)}^2 + 2\|[\gamma - \gamma^n](t, \cdot)\|_{L^2(0,1)}^2 \\ &\quad + \|[s - s^n](t, \cdot)\|_{L^2(0,1)}^2 \left[ 3\|\beta\|_\infty \|W(t, \cdot, \cdot)\|_{L^2([0,1]^2)} + \|\gamma\|_\infty + 1 \right] \\ &\quad + \|[v - v^n](t, \cdot)\|_{L^2(0,1)}^2 \left[ 2\|\beta\|_\infty \|W(t, \cdot, \cdot)\|_{L^2([0,1]^2)} + \|\gamma\|_\infty + 2 \right]. \end{aligned}$$

Owing to Grönwall Lemma and to the identities  $i = v - s$ ,  $i^n = v^n - s^n$ ,  $r = 1 - v$ ,  $r^n = 1 - v^n$  we eventually arrive at (4.1).  $\square$

4.1.1. *Conclusion of the proof of Theorem 1.4 (CASE 1).* In view of Lemma 4.1, to establish Theorem 1.4 (CASE 1) we are left to show that the right hand side of (4.1) converges to 0 as  $n \rightarrow +\infty$  uniformly in  $t \in [0, T]$ . Owing to Lemma 2.6, we get for the initial data

$$\lim_{n \rightarrow +\infty} \|s_0 - s_0^n\|_2^2 = 0, \quad \lim_{n \rightarrow +\infty} \|i_0 - i_0^n\|_2^2 = 0, \quad \lim_{n \rightarrow +\infty} \|r_0 - r_0^n\|_2^2 = 0. \quad (4.9)$$

Arguing analogously as in (3.22), we obtain

$$\|\beta^n(t, \cdot)\|_{L^2(0,1)} \leq \|\beta(t, \cdot)\|_{L^2(0,1)}, \quad \|\gamma^n(t, \cdot)\|_{L^2(0,1)} \leq \|\gamma(t, \cdot)\|_{L^2(0,1)} \quad \text{for a.e. } t \in [0, T]. \quad (4.10)$$

Using again (2.9) from Lemma 2.6, we get

$$\lim_{n \rightarrow +\infty} \|\beta(t, \cdot) - \beta^n(t, \cdot)\|_{L^2(0,1)}^2 = 0, \quad \lim_{n \rightarrow +\infty} \|\gamma(t, \cdot) - \gamma^n(t, \cdot)\|_{L^2(0,1)}^2 = 0 \quad \text{for a.e. } t \in [0, T]$$

and by virtue of the Lebesgue's Dominated Convergence Theorem together with (4.10) this yields

$$\lim_{n \rightarrow +\infty} \int_0^T \|\beta(t, \cdot) - \beta^n(t, \cdot)\|_{L^2(0,1)}^2 dt = 0, \quad \lim_{n \rightarrow +\infty} \int_0^T \|\gamma(s, \cdot) - \gamma^n(s, \cdot)\|_{L^2(0,1)}^2 ds = 0. \quad (4.11)$$

Next, we use Hölder's inequality and point out that

$$\begin{aligned} \|[\beta W - \beta^n W_{\mathcal{G}_n}](t, \cdot, \cdot)\|_{L^2([0,1]^2)}^2 &\leq \left( \|[\beta W - \beta^n W](t, \cdot, \cdot)\|_{L^2([0,1]^2)} + \|[\beta^n W - \beta^n W_{\mathcal{G}_n}](t, \cdot, \cdot)\|_{L^2([0,1]^2)} \right)^2 \\ &\leq \left( \|[\beta W - \beta^n W](t, \cdot, \cdot)\|_{L^2([0,1]^2)} + \|\beta^n(t, \cdot)\|_{L^\infty(0,1)} \| [W - W_{\mathcal{G}_n}](t, \cdot, \cdot) \|_{L^2([0,1]^2)} \right)^2 \\ &\leq 2\|[\beta W - \beta^n W](t, \cdot, \cdot)\|_{L^2([0,1]^2)}^2 + 2\|\beta^n(t, \cdot)\|_{L^\infty(0,1)}^2 \| [W - W_{\mathcal{G}_n}](t, \cdot, \cdot) \|_{L^2([0,1]^2)}^2. \end{aligned} \quad (4.12)$$

We now want to show that

$$\lim_{n \rightarrow +\infty} \underbrace{\int_0^T \|\beta W(s, \cdot, \cdot) - \beta^n W(s, \cdot, \cdot)\|_{L^2([0,1]^2)}^2 ds}_{:= H_n} = 0. \quad (4.13)$$

To this end, it suffices to show that, for every subsequence  $\{H_{n_j}\}_{j \in \mathbb{N}}$ , there is a further subsubsequence  $\{H_{n_{j_h}}\}_{h \in \mathbb{N}}$  such that  $\lim_{h \rightarrow +\infty} H_{n_{j_h}} = 0$ . We therefore fix an arbitrary subsubsequence  $\{H_{n_j}\}_{j \in \mathbb{N}}$  and notice that from (3.21) it follows that there is a subsubsequence  $\beta^{n_{j_h}}(t, x)$  that converges to  $\beta(t, x)$  for a.e.  $(t, x) \in [0, T] \times [0, 1]$ . This implies that  $W^{\beta^{n_{j_h}}}(t, x)$  converges to  $W\beta(t, x)$  for a.e.  $(t, x) \in [0, T] \times [0, 1]$ . Hence, the bound

$$|W[\beta^{n_{j_h}} - \beta]|^2 \leq |W|^2 [\|\beta^{n_{j_h}}\|_\infty + \|\beta\|_\infty]^2 \leq 4|W|^2 \|\beta\|_\infty^2 \quad \text{a.e. on } [0, T] \times [0, 1]^2$$

together with the Lebesgue's Dominated Convergence Theorem yields  $\lim_{h \rightarrow +\infty} H_{n_{j_h}} = 0$  and concludes the proof of (4.13). Owing to (2.9), (3.22) and to the Lebesgue's Dominated Convergence Theorem we also have

$$\lim_{n \rightarrow +\infty} \int_0^T \|W(s, \cdot, \cdot) - W_{\mathcal{G}_n}(s, \cdot, \cdot)\|_{L^2([0,1]^2)}^2 ds = 0. \quad (4.14)$$

Substituting (4.13) and (4.14) into (4.12), using (4.11) and (4.9) and recalling (4.1) we eventually arrive at (1.24).

**4.2. Proof of Theorem 1.4 (CASE 2).** The proof follows the same argument as for (CASE 1), so we only sketch and highlight the points where there are differences. We subtract the first line of (3.6) from the first line of (1.9) and multiply the result by  $\text{sign}[s - s^n]$ . Next, we integrate in space and arguing as in STEP 1 of the proof of Lemma 4.1 we get

$$\frac{d}{dt} \int_0^1 |s - s^n|(t, x) dx \leq \int_0^1 \int_0^1 \beta(t, y) W(t, x, y) |i - i^n|(t, y) dy dx + \int_0^1 \int_0^1 |\beta W - \beta^n W_{\mathcal{G}_n}|(t, x, y) dy dx. \quad (4.15)$$

From the symmetry property  $W(t, x, y) = W(t, y, x)$  we immediately get

$$\begin{aligned} \int_0^1 \int_0^1 \beta(t, y) W(t, x, y) |i - i^n|(t, y) dx dy &\leq \|\beta\|_\infty \int_0^1 |i - i^n|(t, y) \int_0^1 W(t, x, y) dx dy \\ &\stackrel{(1.13)}{\leq} K_d \|\beta\|_\infty \| [i - i^n](t, \cdot) \|_{L^1(0,1)}. \end{aligned}$$

By plugging the above inequality in (4.15) and then arguing as in the proof of Lemma 4.1 we arrive at

$$\begin{aligned} &\| [s - s^n](t, \cdot) \|_{L^1(0,1)} + \| [i - i^n](t, \cdot) \|_{L^1(0,1)} + \| [r - r^n](t, \cdot) \|_{L^1(0,1)} \\ &\leq D \left( \|s_0 - s_0^n\|_1 + \|i_0 - i_0^n\|_1 + \|r_0 - r_0^n\|_1 \right) \\ &\quad + D \int_0^t \left( \| [\beta W - \beta^n W_{\mathcal{G}_n}](s, \cdot, \cdot) \|_{L^1([0,1]^2)} + \| [\gamma - \gamma^n](s, \cdot) \|_{L^1(0,1)} \right) ds \quad \text{for every } t \in [0, T], \end{aligned}$$

for a suitable constant  $D$  only depending on  $T, K_d, \|\beta\|_\infty$  and  $\|\gamma\|_\infty$ . Hence, we need to ensure convergence of the right hand side of the above expression. We first point out that

$$\| [\beta W - \beta^n W_{\mathcal{G}_n}](s, \cdot, \cdot) \|_{L^1([0,1]^2)} \leq \| [\beta - \beta^n] W(s, \cdot, \cdot) \|_{L^1([0,1]^2)} + \| [\beta^n [W - W_{\mathcal{G}_n}]](s, \cdot, \cdot) \|_{L^1([0,1]^2)}.$$

The first term in the above sum can be further estimated by

$$\| [\beta - \beta^n] W(s, \cdot, \cdot) \|_{L^1([0,1]^2)} = \int_0^1 |\beta - \beta^n|(t, y) \int_0^1 W(s, x, y) dx dy \stackrel{(1.13)}{\leq} K_d \| [\beta - \beta^n](s, \cdot) \|_{L^1(0,1)},$$

while to control the second term it suffices to note that, by construction,  $\|\beta^n\|_\infty \leq \|\beta\|_\infty$  and this implies  $\| [\beta^n (W - W_{\mathcal{G}_n})](s, \cdot, \cdot) \|_{L^1([0,1]^2)} \leq \|\beta\|_\infty \| [W - W_{\mathcal{G}_n}](s, \cdot, \cdot) \|_{L^1([0,1]^2)}$ .

The rest of the argument is basically the same as for CASE 1 in §4.1.1, the main difference is that to control  $\| [W - W_{\mathcal{G}_n}](s, \cdot, \cdot) \|_{L^1([0,1]^2)}$  we use (2.8) from Lemma 2.6 instead of (2.9).

*Remark 4.1.* As pointed out in Remark 3.1, from the point of view of applications it would be also reasonable to consider the SIR model (3.19). As a matter of fact, the proof of Theorem 1.4 does extend to (3.19). More precisely, under the same assumptions as in Theorem 1.4, the solution  $(s^n, i^n, r^n)$  of (3.19) satisfies (1.24), provided  $s, i, r : [0, T] \times [0, 1] \rightarrow [0, 1]$  is the distributional solution of

$$\begin{cases} \partial_t s(t, x) = -s(t, x) \int_0^1 \beta(t, y) W(t, x, y) i(t, y) dy - \beta(t, x) s(t, x) i(t, x) \\ \partial_t i(t, x) = s(t, x) \int_0^1 \beta(t, y) W(t, x, y) i(t, y) dy + \beta(t, x) s(t, x) i(t, x) - \gamma(t, x) i(t, x) \\ \partial_t r(t, x) = \gamma(t, x) i(t, x) \end{cases} \quad x \in [0, 1]$$

coupled with the initial condition (1.10).

## 5. RANDOM APPROXIMATION: CONVERGENCE ANALYSIS OF THE RANDOM SAMPLINGS

In this section we present the convergence analysis for the random approximation as stated in Theorem 1.5.

The proof is inspired by the argument used in [Med19] for Kuramoto systems. In short, at each step of the approximation, we will introduce an auxiliary deterministic SIR system on an averaged graph and split the total error of the approximation in two components: a random error (between the original random model and the auxiliary averaged one) and a deterministic error (between the averaged model and the limit one).

The exposition in §5 is organized as follows. In §5.1 we review some preliminary probability results. In §5.2 we pass from the random to the averaged model. The proof of Theorem 1.5 is carried out in §5.3 for CASE 1 and in §5.4 for CASE 2.

**5.1. Preliminary results.** We begin by stating (without proof) three classical results in Probability Theory. The first one is the Borel–Cantelli Lemma (see for instance [Cin11, Chapter 3]).

**Lemma 5.1** (Borel-Cantelli). *Let  $\{\mathcal{E}_n\}_{n \in \mathbb{N}}$  be a sequence of events in a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and set  $\mathcal{E} := \limsup_{n \rightarrow +\infty} \mathcal{E}_n = \bigcap_{k=1}^{\infty} \bigcup_{n \geq k} \mathcal{E}_n$ . If  $\sum_{n \in \mathbb{N}} \mathbb{P}(\mathcal{E}_n) < \infty$ , then  $\mathbb{P}(\mathcal{E}) = 0$ .*

The second tool is Chebyshev’s inequality: let  $X$  be a random variable on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with bounded second moment. Then,

$$\mathbb{P}(|X - \mathbb{E}X| \geq t) \leq \frac{\text{Var}(X)}{t^2} \quad \text{for every } t > 0, \quad (5.1)$$

where  $\text{Var}(X) := \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$ .

The third tool is Hoeffding’s inequality: let  $X_1, \dots, X_m$  be independent random variables on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  such that  $a \leq X_i \leq b$  almost surely for every  $i = 1, \dots, m$ , for some  $a, b \in \mathbb{R}$ . Then, for every  $t > 0$  it holds

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^m X_i - \sum_{i=1}^m \mathbb{E}(X_i) \geq t\right) &\leq \exp\left[-\frac{2t^2}{m(b-a)^2}\right], \quad \text{for every } t > 0, \\ \mathbb{P}\left(\left|\frac{1}{m} \sum_{i=1}^m X_i - \frac{1}{m} \sum_{i=1}^m \mathbb{E}(X_i)\right| \geq t\right) &\leq \exp\left[-\frac{2mt^2}{(b-a)^2}\right] \quad \text{for every } t > 0. \end{aligned} \quad (5.2)$$

We now recall that, for  $u^n$  as in (1.8), the  $L^2$  and  $L^1$  discrete norms are defined by

$$\|u^n\|_2 = \left(\frac{1}{n} \sum_{j=1}^n |u_j^n|^2\right)^{1/2}, \quad \|u^n\|_1 = \frac{1}{n} \sum_{j=1}^n |u_j^n|, \quad (5.3)$$

and the discrete Hölder’s inequality reads

$$\sum_{i=1}^n x_i y_i \leq \left(\sum_{i=1}^n x_i^2\right)^{1/2} \left(\sum_{i=1}^n y_i^2\right)^{1/2} \quad \text{for every } x_1, \dots, x_n, y_1, \dots, y_n \in \mathbb{R}. \quad (5.4)$$

Furthermore, note that, adapting the argument given in Lemma 2.2 to prove (1.5), if  $A_{jk}^n$  is the random variable (1.26) and  $(s^n, i^n, r^n)$  is the solution of the Cauchy problem (1.1), (1.3), written according to the notation (1.8), then

$$0 \leq s_j^n, i_j^n, r_j^n \leq 1, \quad s_j^n + i_j^n + r_j^n = 1 \quad \text{almost surely and for a.e. } (t, x) \in [0, T] \times [0, 1]. \quad (5.5)$$

**5.2. From a random to a deterministic model: the SIR system on the averaged graph.**

For any given  $n \in \mathbb{N}$  and starting from the sparse  $W$ -random graph  $\mathcal{G}_n = n^\alpha \mathcal{G}(n, W, n^{-\alpha})$  (or  $\mathcal{G}_n = \mathcal{G}(n, W)$ ) we consider the averaged graph, denoted by  $\bar{\mathcal{G}}_n$ , defined by setting

$$\bar{A}_{jk}^n := \mathbb{E}(A_{jk}^n) \stackrel{(1.26)}{=} n^\alpha < \widehat{W}^n >_{I_j^n \times I_k^n}, \quad \text{where } \widehat{W}^n = \min\{1, n^{-\alpha} W\}. \quad (5.6)$$

Using always the notation (1.8) we denote by  $(\bar{s}^n, \bar{i}^n, \bar{r}^n)$  the solution of the SIR system on  $\bar{\mathcal{G}}_n$

$$\begin{cases} \frac{d\bar{s}_j^n(t)}{dt} = -\bar{s}_j^n(t) \frac{1}{n} \sum_{k=1}^n \beta_k^n(t) \bar{A}_{jk}^n \bar{i}_k^n(t), \\ \frac{d\bar{i}_j^n(t)}{dt} = \bar{s}_j^n(t) \frac{1}{n} \sum_{k=1}^n \beta_k^n(t) \bar{A}_{jk}^n \bar{i}_k^n(t) - \gamma_j^n(t) \bar{i}_j^n(t) \\ \frac{d\bar{r}_j^n(t)}{dt} = \gamma_j^n(t) \bar{i}_j^n(t), \end{cases} \quad j = 1, \dots, n, \quad (5.7)$$

coupled with the initial conditions (1.3). Note that the coefficients  $\beta_j^n$  and  $\gamma_j^n$  and the initial data  $s_{j,0}^n, i_{j,0}^n, r_{j,0}^n$  are as in (1.23). Note that the above system (5.7) is indeed a *deterministic* approximation to (1.9).

From the definition of the averaged graph  $\bar{\mathcal{G}}_n$ , and depending on the regularity of  $W$ , we can derive the following estimates that will be useful for our analysis:

- CASE 1: : if  $W \in L^2([0, 1]^2; \mathbb{R}_+)$  then

$$\begin{aligned} \sum_{j,k=1}^n \frac{[\bar{A}_{jk}^n]^2}{n^2} &\stackrel{(5.6)}{=} \sum_{j,k=1}^n \frac{1}{n^2} \left( n^2 \int_{I_j^n} \int_{I_k^n} n^\alpha \widehat{W}^n(x, y) dx dy \right)^2 \\ &\leq \sum_{j,k=1}^n \frac{1}{n^2} n^2 \int_{I_j^n} \int_{I_k^n} [n^\alpha \widehat{W}^n(x, y)]^2 dx dy \leq \sum_{j,k=1}^n \int_{I_j^n} \int_{I_k^n} [W(x, y)]^2 dx dy = \|W\|_2^2; \end{aligned} \quad (5.8)$$

- CASE 2: : if  $W \in L^1([0, 1]^2; \mathbb{R}_+)$  satisfies (1.30) then

$$\begin{aligned} \frac{1}{n} \sum_{j=1}^n \bar{A}_{jk}^n &\stackrel{(5.6)}{=} \frac{1}{n} \sum_{j=1}^n n^\alpha \langle \widehat{W}^n \rangle_{I_j^n \times I_k^n} \leq \frac{1}{n} \sum_{j=1}^n \langle W \rangle_{I_j^n \times I_k^n} = \frac{1}{n} \sum_{j=1}^n n^2 \int_{I_j^n} \int_{I_k^n} W(x, y) dx dy \\ &= n \int_{I_j^n} \int_0^1 W(x, y) dx dy \stackrel{(1.30)}{\leq} K_a. \end{aligned} \quad (5.9)$$

Note furthermore that, by construction,

$$\mathbb{E}(\bar{A}_{jk}^n - A_{jk}^n) \stackrel{(5.6)}{=} 0. \quad (5.10)$$

**5.3. Proof of Theorem 1.5 (CASE 1).** In what follows we always assume  $W \in L^2([0, 1]^2; \mathbb{R}_+)$ . We start by proving the following lemma, which provides an estimate (with probability one) on the difference between the random approximation  $(s^n, i^n, r^n)$  and the deterministic approximation  $(\bar{s}^n, \bar{i}^n, \bar{r}^n)$  solution of the associated averaged model.

**Lemma 5.2.** *Under the same assumptions as in the statement of Theorem 1.5 (CASE 1), let  $\bar{s}^n, \bar{i}^n, \bar{r}^n : [0, T] \times [0, 1] \rightarrow \mathbb{R}$  be the solution of the Cauchy problem (5.7), (1.3). Then*

$$\lim_{n \rightarrow +\infty} \left[ \|[s^n - \bar{s}^n](t, \cdot)\|_{L^2(0,1)} + \|[i^n - \bar{i}^n](t, \cdot)\|_{L^2(0,1)} + \|[r^n - \bar{r}^n](t, \cdot)\|_{L^2(0,1)} \right] = 0 \quad (5.11)$$

$\mathbb{P}$ -almost surely and for a.e.  $t \in [0, T]$ .

*Proof.* We proceed according to the following steps.

STEP 1: we subtract (1.1) from (5.7), multiply the result by  $2[\bar{s}_j^n - s_j^n]$ , use discrete Hölder (5.4) to obtain

$$\begin{aligned} \frac{d}{dt} [\bar{s}_j^n - s_j^n]^2 &= - \underbrace{\frac{2[\bar{s}_j^n - s_j^n]^2}{n} \sum_{k=1}^n \beta_k^n \bar{A}_{jk}^n \bar{i}_k^n}_{\geq 0} - \frac{2s_j^n [\bar{s}_j^n - s_j^n]}{n} \sum_{k=1}^n \beta_k^n [\bar{A}_{jk}^n \bar{i}_k^n - A_{jk}^n i_k^n] \\ &\stackrel{(3.20), (5.5)}{\leq} 2\|\beta\|_\infty \frac{|\bar{s}_j^n - s_j^n|}{n} \sum_{k=1}^n \bar{A}_{jk}^n |\bar{i}_k^n - i_k^n| + \frac{2s_j^n |\bar{s}_j^n - s_j^n|}{n} \sum_{k=1}^n \beta_k^n i_k^n [\bar{A}_{jk}^n - A_{jk}^n] \\ &\stackrel{(5.4), (5.5)}{\leq} 2\|\beta\|_\infty \frac{|\bar{s}_j^n - s_j^n|}{n} \left( \sum_{k=1}^n [\bar{A}_{jk}^n]^2 \right)^{1/2} \left( \sum_{k=1}^n |\bar{i}_k^n - i_k^n|^2 \right)^{1/2} + 2[\bar{s}_j^n - s_j^n] Z_j^n(t), \end{aligned}$$

where we have set

$$Z_j^n(t) := \frac{s_j^n(t)}{n} \sum_{k=1}^n \beta_k^n(t) i_k^n(t) [\bar{A}_{jk}^n - A_{jk}^n]. \quad (5.12)$$

Hence, summing over all nodes and using the discrete Hölder Inequality (5.4), we have

$$\begin{aligned} \frac{d}{dt} \sum_{j=1}^n [\bar{s}_j^n - s_j^n]^2 &\leq 2\|\beta\|_\infty \left( \sum_{k=1}^n |\bar{i}_k^n - i_k^n|^2 \right)^{1/2} \sum_{j=1}^n |\bar{s}_j^n - s_j^n| \left( \sum_{k=1}^n \frac{[\bar{A}_{jk}^n]^2}{n^2} \right)^{1/2} + 2 \sum_{j=1}^n [\bar{s}_j^n - s_j^n] Z_j^n \\ &\stackrel{(5.4)}{\leq} 2\|\beta\|_\infty \left( \sum_{k=1}^n |\bar{i}_k^n - i_k^n|^2 \right)^{1/2} \left( \sum_{j=1}^n |\bar{s}_j^n - s_j^n|^2 \right)^{1/2} \underbrace{\left( \sum_{j,k=1}^n \frac{[\bar{A}_{jk}^n]^2}{n^2} \right)^{1/2} + 2 \left( \sum_{j=1}^n [\bar{s}_j^n - s_j^n]^2 \right)^{1/2}}_{\leq \|W\|_2 \text{ by (5.8)}} \left( \sum_{j=1}^n (Z_j^n)^2 \right)^{1/2} \\ &\stackrel{\text{Young}}{\leq} \|\beta\|_\infty \|W\|_2 \sum_{j=1}^n |\bar{i}_j^n - i_j^n|^2 + \|\beta\|_\infty \left[ \|W\|_2 + 1 \right] \sum_{j=1}^n |\bar{s}_j^n - s_j^n|^2 + \sum_{j=1}^n (Z_j^n)^2. \end{aligned} \quad (5.13)$$

Next, we use an argument similar to the one in the proof of Proposition 1.2 in §3.1. We sum the first two equations in (1.1), subtract it from the sum of the first two equations in (5.7) and set  $\bar{v}_j^n := \bar{s}_j^n + \bar{i}_j^n$ ,  $v_j^n := s_j^n + i_j^n$  to get

$$\frac{d}{dt} [\bar{v}_j^n - v_j^n] = -\gamma_j^n [\bar{v}_j^n - v_j^n] + \gamma_j^n [\bar{s}_j^n - s_j^n]. \quad (5.14)$$

Multiplying the above equation by  $2[\bar{v}_j^n - v_j^n]$  we find

$$\frac{d}{dt} [\bar{v}_j^n - v_j^n]^2 = - \underbrace{2\gamma_j^n [\bar{v}_j^n - v_j^n]^2}_{\geq 0} + 2\gamma_j^n [\bar{s}_j^n - s_j^n] [\bar{v}_j^n - v_j^n] \stackrel{(3.20), \text{Young}}{\leq} \|\gamma\|_\infty \left[ [\bar{s}_j^n - s_j^n]^2 + [\bar{v}_j^n - v_j^n]^2 \right],$$

which in turn entails

$$\frac{d}{dt} \sum_{j=1}^n [\bar{v}_j^n - v_j^n]^2 \leq \|\gamma\|_\infty \left[ \sum_{j=1}^n [\bar{s}_j^n - s_j^n]^2 + \sum_{j=1}^n [\bar{v}_j^n - v_j^n]^2 \right]. \quad (5.15)$$

STEP 2: we combine the estimates (5.13) and (5.15), and note that

$$\sum_{k=1}^n |\bar{i}_k^n - i_k^n|^2 \leq 2 \sum_{k=1}^n |\bar{v}_k^n - v_k^n|^2 + 2 \sum_{k=1}^n |\bar{s}_k^n - s_k^n|^2.$$

Recalling now the expression of the  $L^2$  norm (5.3) and applying Grönwall Lemma, we eventually arrive at

$$\begin{aligned} &\|[\bar{v}^n - v^n](t, \cdot)\|_{L^2(0,1)}^2 + \|[\bar{s}^n - s^n](t, \cdot)\|_{L^2(0,1)}^2 \\ &\leq \exp[\bar{B}t] \frac{1}{n} \int_0^t \exp[-\bar{B}\tau] \sum_{j=1}^n (Z_j^n)^2(\tau) d\tau \quad \text{almost surely and for every } t \in [0, T], \end{aligned} \quad (5.16)$$

for a suitable constant  $\bar{B}$  only depending on  $\|\beta\|_\infty$ ,  $\|\gamma\|_\infty$  and  $\|W\|_2$ . Owing to the equalities  $\bar{v}_j^n := \bar{s}_j^n + \bar{i}_j^n$ ,  $v_j^n := s_j^n + i_j^n$ ,  $\bar{r}_j^n = 1 - \bar{v}_j^n$ ,  $r_j^n = 1 - v_j^n$ , to establish (5.11) it suffices to show that the right hand side of (5.16) vanishes in the  $n \rightarrow +\infty$  limit, almost surely. To this end we notice that

$$\begin{aligned} &\frac{1}{n} \int_0^T \exp[-\bar{B}s] \sum_{j=1}^n (Z_j^n)^2(s) ds \stackrel{(5.12)}{=} \frac{1}{n^3} \sum_{j=1}^n \sum_{k=1}^n |\bar{A}_{jk}^n - A_{jk}^n|^2 \int_0^T \exp[-\bar{B}\tau] [s_j^n(\tau)]^2 [\beta_k^n i_k^n]^2(\tau) d\tau \\ &+ \frac{2}{n^3} \sum_{j=1}^n \sum_{\substack{k,\ell=1 \\ k \neq \ell}}^n (\bar{A}_{jk}^n - A_{jk}^n) (\bar{A}_{j\ell}^n - A_{j\ell}^n) \int_0^T [s_j^n(\tau)]^2 [\beta_k^n i_k^n \beta_\ell^n i_\ell^n](\tau) \exp[-\bar{B}\tau] d\tau. \end{aligned} \quad (5.17)$$

To control the first term in the above sum, we point out that almost surely we have

$$\begin{aligned} 0 &\leq \frac{1}{n^3} \int_0^T \exp[-\bar{B}\tau] \sum_{j=1}^n [s_j^n(\tau)]^2 \left[ \sum_{k=1}^n [\beta_k^n i_k^n]^2(\tau) |\bar{A}_{jk}^n - A_{jk}^n|^2 \right] d\tau \\ &\stackrel{(5.5), (3.20)}{\leq} \|\beta\|_\infty \frac{1}{n^3} \sum_{j,k=1}^n |\bar{A}_{jk}^n - A_{jk}^n|^2 \int_0^T \exp[-\bar{B}\tau] d\tau. \end{aligned}$$

Note that  $\{|\bar{A}_{jk}^n - A_{jk}^n|^2\}_{j,k=1,\dots,n}$  are independent random variables and, thanks to (1.26) and (5.6), satisfy

$$0 \leq |\bar{A}_{jk}^n - A_{jk}^n|^2 \leq n^{2\alpha} \quad \text{almost surely}$$

together with

$$\mathbb{E}(|\bar{A}_{jk}^n - A_{jk}^n|^2) = \mathbb{E}([A_{jk}^n]^2) - [\mathbb{E}(A_{jk}^n)]^2 \stackrel{(1.26)}{=} n^{2\alpha} < \widehat{W}^n >_{I_j^n \times I_j^n} [1 - < \widehat{W}^n >_{I_j^n \times I_j^n}] \leq n^{2\alpha}. \quad (5.18)$$

Hoeffding's inequality (5.2) then yields

$$\mathbb{P} \left( \sum_{j,k=1}^n |\bar{A}_{jk}^n - A_{jk}^n|^2 - \sum_{j,k=1}^n \mathbb{E}(|\bar{A}_{jk}^n - A_{jk}^n|^2) \geq t \right) \leq \exp \left( -2 \left[ \frac{t}{n^{2\alpha+1}} \right]^2 \right). \quad (5.19)$$

Setting  $t := n^\beta$  with  $\beta \in ]1 + 2\alpha, 3[$  in (5.19), we apply Borel-Cantelli Lemma to the event

$$\mathcal{E}_n := \left\{ \sum_{j,k=1}^n |\bar{A}_{jk}^n - A_{jk}^n|^2 \geq \sum_{j,k=1}^n \mathbb{E}(|\bar{A}_{jk}^n - A_{jk}^n|^2) + n^\beta \right\},$$

to obtain

$$\begin{aligned} \lim_{n \rightarrow +\infty} n^{-3} \sum_{j,k=1}^n |\bar{A}_{jk}^n - A_{jk}^n|^2 &\leq \lim_{n \rightarrow +\infty} n^{-3} \sum_{j,k=1}^n \mathbb{E}(|\bar{A}_{jk}^n - A_{jk}^n|^2) + \lim_{n \rightarrow +\infty} n^{\beta-3} \\ &\stackrel{(5.18)}{\leq} \lim_{n \rightarrow +\infty} n^{2\alpha-1} + \lim_{n \rightarrow +\infty} n^{\beta-3} = 0 \quad \text{almost surely.} \end{aligned}$$

We now turn our attention to the second term in (5.17). We first observe that, owing to (5.5) and (3.20), we have

$$0 \leq \int_0^T [s_j^n(\tau)]^2 [\beta_k^n i_k^n \beta_\ell^n i_\ell^n](\tau) \exp[-\bar{B}\tau] d\tau \leq \|\beta\|_\infty \int_0^T \exp[-\bar{B}\tau] d\tau \quad \text{almost surely.}$$

Next, notice that for  $k \neq \ell$  the random variables  $\bar{A}_{jk}^n - A_{jk}^n$  and  $\bar{A}_{j\ell}^n - A_{j\ell}^n$  are independent, which together with the linearity of the expectation implies

$$\mathbb{E} \left( \sum_{j=1}^n \sum_{k,\ell=1,k \neq \ell}^n (\bar{A}_{jk}^n - A_{jk}^n)(\bar{A}_{j\ell}^n - A_{j\ell}^n) \right) = \sum_{j=1}^n \sum_{k,\ell=1,k \neq \ell}^n \mathbb{E}(\bar{A}_{jk}^n - A_{jk}^n) \mathbb{E}(\bar{A}_{j\ell}^n - A_{j\ell}^n) \stackrel{(5.10)}{=} 0.$$

Hence, Chebyshev's inequality (5.1) yields

$$\mathbb{P} \left( \sum_{j=1}^n \sum_{k,\ell=1,k \neq \ell}^n (\bar{A}_{jk}^n - A_{jk}^n)(\bar{A}_{j\ell}^n - A_{j\ell}^n) \geq t \right) \leq \frac{1}{t^2} \mathbb{E} \left( \left[ \sum_{j=1}^n \sum_{k,\ell=1,k \neq \ell}^n (\bar{A}_{jk}^n - A_{jk}^n)(\bar{A}_{j\ell}^n - A_{j\ell}^n) \right]^2 \right). \quad (5.20)$$

The linearity of the expectation allows us to write

$$\begin{aligned} &\mathbb{E} \left( \left[ \sum_{j=1}^n \sum_{k,\ell=1,k \neq \ell}^n [\bar{A}_{jk}^n - A_{jk}^n][\bar{A}_{j\ell}^n - A_{j\ell}^n] \right]^2 \right) = \sum_{j=1}^n \sum_{\substack{k,\ell=1 \\ k \neq \ell}}^n \mathbb{E}([\bar{A}_{jk}^n - A_{jk}^n]^2 [\bar{A}_{j\ell}^n - A_{j\ell}^n]^2) \\ &+ 2 \sum_{i,j=1}^n \sum_{\substack{k,\ell,h,m=1 \\ k,\ell,h,m \in \mathcal{I}}}^n \mathbb{E}([\bar{A}_{jk}^n - A_{jk}^n][\bar{A}_{j\ell}^n - A_{j\ell}^n][\bar{A}_{ih}^n - A_{ih}^n][\bar{A}_{im}^n - A_{im}^n]) \end{aligned}$$

with the set of indices  $\mathcal{I} := \{k, \ell, h, m = 1, \dots, n : k \neq \ell, h \neq m, [k \neq h \vee \ell \neq m] \wedge [k \neq m \vee \ell \neq h]\}$ . Observe that the independence of  $\bar{A}_{jk}^n - A_{jk}^n$  and (5.10) guarantee that all the terms in the second sum above vanish. Precisely, if  $k \neq \ell, h \neq m, k \neq h, k \neq m$

$$\mathbb{E}((\bar{A}_{jk}^n - A_{jk}^n)(\bar{A}_{j\ell}^n - A_{j\ell}^n)(\bar{A}_{ih}^n - A_{ih}^n)(\bar{A}_{im}^n - A_{im}^n)) = \mathbb{E}(\bar{A}_{jk}^n - A_{jk}^n)\mathbb{E}(\bar{A}_{j\ell}^n - A_{j\ell}^n)\mathbb{E}(\bar{A}_{ih}^n - A_{ih}^n)\mathbb{E}(\bar{A}_{im}^n - A_{im}^n) \stackrel{(5.10)}{=} 0,$$

and arguing similarly one finds

$$\mathbb{E}((\bar{A}_{jk}^n - A_{jk}^n)(\bar{A}_{j\ell}^n - A_{j\ell}^n)(\bar{A}_{ih}^n - A_{ih}^n)(\bar{A}_{im}^n - A_{im}^n)) = 0 \quad \text{for every } (k, \ell, h, m) \in \mathcal{I}. \quad (5.21)$$

As for the terms in the first sum, since  $k \neq \ell$ , the independence together with (5.18) gives

$$\mathbb{E}([\bar{A}_{jk}^n - A_{jk}^n]^2[\bar{A}_{j\ell}^n - A_{j\ell}^n]^2) = \mathbb{E}([\bar{A}_{jk}^n - A_{jk}^n]^2)\mathbb{E}([\bar{A}_{j\ell}^n - A_{j\ell}^n]^2) \stackrel{(5.18)}{\leq} n^{4\alpha}. \quad (5.22)$$

Plugging (5.22) and (5.21) into (5.20) we arrive at

$$\mathbb{P}\left(n^{-3} \sum_{j=1}^n \sum_{k, \ell=1, k \neq \ell}^n |\bar{A}_{jk}^n - A_{jk}^n| |\bar{A}_{j\ell}^n - A_{j\ell}^n| \geq t\right) \leq \frac{1}{n^6 t^2} n^{4\alpha+3} = \frac{1}{t^2} n^{4\alpha-3}.$$

By setting  $t := n^\sigma$ , with  $\sigma \in ]2(\alpha - 1), 0[$ , the right hand side of the above expression boils down to  $n^{4\alpha-3-2\sigma}$ . Then, since  $4\alpha - 3 - 2\sigma < -1$  we can apply the Borel-Cantelli Lemma and conclude that

$$\lim_{n \rightarrow +\infty} n^{-3} \sum_{j=1}^n \sum_{k, \ell=1, k \neq \ell}^n |\bar{A}_{jk}^n - A_{jk}^n| |\bar{A}_{j\ell}^n - A_{j\ell}^n| \leq \lim_{n \rightarrow +\infty} n^\sigma = 0 \quad \text{almost surely.}$$

This eventually implies that the right hand side of (5.16) vanishes almost surely in the  $n \rightarrow +\infty$  limit and completes the proof of Lemma 5.2.  $\square$

We consider the splitting of the error

$$s - s^n = (s - \bar{s}^n) + (\bar{s}^n - s^n) \quad i - i^n = (i - \bar{i}^n) + (\bar{i}^n - i^n) \quad r - r^n = (r - \bar{r}^n) + (\bar{r}^n - r^n).$$

Hence, by virtue of triangle inequality and Lemma 5.2, to conclude the proof of Theorem 1.5 it suffices to show that

$$\lim_{n \rightarrow +\infty} \left[ \| [s - \bar{s}^n](t, \cdot) \|_{L^2(0,1)} + \| [i - \bar{i}^n](t, \cdot) \|_{L^2(0,1)} + \| [r - \bar{r}^n](t, \cdot) \|_{L^2(0,1)} \right] = 0 \quad \text{for every } t \in [0, T].$$

Observe that  $(\bar{s}^n, \bar{i}^n, \bar{r}^n)$  is a deterministic approximation (with a particular sampling of the graphon  $W$ ) of (1.9) and in particular the proof of Lemma 4.1 works for  $(\bar{s}^n, \bar{i}^n, \bar{r}^n)$  provided we replace  $W_{\mathcal{G}_n}$  by the step graphon  $W_{\bar{\mathcal{G}}_n}$  of the averaged graph  $\bar{\mathcal{G}}_n$  (with  $(\bar{A}_{jk})_{j,k=1,\dots,n}$  defined in (5.6)) and guarantee that  $W_{\bar{\mathcal{G}}_n}$  converges to  $W$  in the  $L^2$ -norm (corresponding to (4.14) in §4.1.1). Therefore, arguing exactly as in §4.1.1, the only point we need to show is that

$$\lim_{n \rightarrow +\infty} \|W - W_{\bar{\mathcal{G}}_n}\|_2 = 0. \quad (5.23)$$

Now recall  $\bar{\mathcal{G}}_n$  is the averaged graph of the scaled sparse  $W$ -random graph defined through (1.26) and based on the trimmed graphon with target edge density  $n^\alpha$ . We now introduce an auxiliary graph, denoted by  $\tilde{\mathcal{G}}_n$ , defined through the adjacency matrix as in (1.22), i.e., the  $L^2$ -orthogonal projection of  $W$  on the space of piecewise constant functions on  $[0, 1]^2$ . Then, considering the associated step-graphons, the standard triangle inequality gives

$$\|W - W_{\bar{\mathcal{G}}_n}\|_2 \leq \|W - W_{\tilde{\mathcal{G}}_n}\|_2 + \|W_{\tilde{\mathcal{G}}_n} - W_{\bar{\mathcal{G}}_n}\|_2.$$

Owing to estimate (2.9) from Lemma 2.6, the first term in the above sum vanishes in the  $n \rightarrow +\infty$  limit. To control the limit of the second term, Jensen's inequality directly gives

$$\begin{aligned} \|W_{\tilde{\mathcal{G}}_n} - W_{\bar{\mathcal{G}}_n}\|_2 &= \sum_{j,k=1}^n \frac{|\langle W \rangle_{I_j^n \times I_k^n} - n^\alpha \langle \widehat{W}^n \rangle_{I_j^n \times I_k^n}|^2}{n^2} \\ &\leq \sum_{j,k=1}^n \iint_{I_j^n \times I_k^n} [W - \min\{n^\alpha, W\}]^2(x, y) dx dy = \iint_{[0,1]^2} [W - \min\{n^\alpha, W\}]^2(x, y) dx dy. \end{aligned}$$

As  $\lim_{n \rightarrow +\infty} [W - \min\{n^\alpha, W\}]^2(x, y) = 0$  for a.e.  $(x, y) \in [0, 1]^2$  and  $|W - \min\{n^\alpha, W\}| \leq 2W$  on  $[0, 1]^2$ , then Lebesgue's Dominated Convergence Theorem yields (5.23) and this concludes the proof of Theorem 1.5 (CASE 1).

**5.4. Proof of Theorem 1.5 (CASE 2).** In this subsection we assume  $W \in L^1([0, 1]; \mathbb{R}_+)$  to satisfy (1.30). The proof in this case follows the same lines as that of CASE 1 above, so we only provide a sketch focusing on the points where the arguments are different. We first show that

$$\lim_{n \rightarrow +\infty} \left[ \| [s^n - \bar{s}^n](t, \cdot) \|_{L^1(0,1)} + \| [i^n - \bar{i}^n](t, \cdot) \|_{L^1(0,1)} + \| [r^n - \bar{r}^n](t, \cdot) \|_{L^1(0,1)} \right] = 0 \quad (5.24)$$

$\mathbb{P}$ -almost surely and for a.e.  $t \in [0, T]$ . To this end, we subtract (1.1) from (5.7), multiply the result by  $\text{sign}(\bar{s}_j^n - s_j^n)$  and get

$$\begin{aligned} \frac{d}{dt} |\bar{s}_j^n - s_j^n| &\leq - \underbrace{\frac{|\bar{s}_j^n - s_j^n|}{n} \sum_{k=1}^n \beta_k^n \bar{A}_{jk}^n \bar{i}_k^n}_{\geq 0} + \frac{|s_j^n|}{n} \sum_{k=1}^n \beta_k^n \bar{A}_{jk}^n |\bar{i}_k^n - i_k^n| + \frac{\text{sign}[\bar{s}_j^n - s_j^n] s_j^n}{n} \sum_{k=1}^n \beta_k^n i_k^n (\bar{A}_{jk}^n - A_{jk}^n) \\ &\stackrel{(5.5), (3.20)}{\leq} \frac{\|\beta\|_\infty}{n} \sum_{k=1}^n \bar{A}_{jk}^n |\bar{i}_k^n - i_k^n| + H_j^n, \end{aligned}$$

where we have set

$$H_j^n(t) := \frac{\text{sign}[\bar{s}_j^n - s_j^n] s_j^n}{n} \sum_{k=1}^n \beta_k^n i_k^n (\bar{A}_{jk}^n - A_{jk}^n).$$

This implies

$$\frac{d}{dt} \sum_{j=1}^n |\bar{s}_j^n - s_j^n| \leq \frac{\|\beta\|_\infty}{n} \sum_{k=1}^n |\bar{i}_k^n - i_k^n| \sum_{j=1}^n \bar{A}_{jk}^n + \sum_{j=1}^n H_j^n \stackrel{(5.9)}{\leq} K_a \|\beta\|_\infty \sum_{k=1}^n |\bar{i}_k^n - i_k^n| + \sum_{j=1}^n H_j^n. \quad (5.25)$$

We then set  $\bar{v}_j^n := \bar{s}_j^n + \bar{i}_j^n$ ,  $v_j^n := s_j^n + i_j^n$  and point out that owing to (5.14) we have

$$\frac{d}{dt} |\bar{v}_j^n - v_j^n| \leq - \underbrace{\gamma_j^n |\bar{v}_j^n - v_j^n|}_{\geq 0} + \gamma_j^n |\bar{s}_j^n - s_j^n|,$$

so that summing the above expression over  $j$ , recalling (5.25), using the inequality  $|\bar{i}_j^n - i_j^n| \leq |\bar{v}_j^n - v_j^n| + |\bar{s}_j^n - s_j^n|$ , the definition of the  $L^1$  norm as in (5.3) and applying Grönwall's Lemma we arrive at

$$\begin{aligned} \| [s^n - \bar{s}^n](t, \cdot) \|_{L^1(0,1)} + \| [i^n - \bar{i}^n](t, \cdot) \|_{L^1(0,1)} &\leq \frac{1}{n} \exp[\bar{D}t] \int_0^T \exp[-\bar{D}\tau] \sum_{j=1}^n H_j^n(\tau) d\tau \\ &= \frac{1}{n^2} \sum_{j,k=1}^n (\bar{A}_{jk}^n - A_{jk}^n) \exp[\bar{D}t] \int_0^T \exp[-\bar{D}\tau] \text{sign}[\bar{s}_j^n - s_j^n] [s_j^n \beta_k^n i_k^n](\tau) d\tau \quad \text{almost surely.} \end{aligned} \quad (5.26)$$

In the above expression  $\bar{D}$  is a suitable constant depending only on  $\|\beta\|_\infty$ ,  $\|\gamma\|_\infty$  and  $K_a$ . To control the right hand side of (5.26), note that

$$\left| \int_0^T \exp[-\bar{D}\tau] \text{sign}[\bar{s}_j^n - s_j^n] [s_j^n \beta_k^n i_k^n](\tau) d\tau \right| \stackrel{(5.5), (3.20)}{\leq} \|\beta\|_\infty \int_0^T \exp[-\bar{D}\tau] d\tau \quad \text{almost surely.} \quad (5.27)$$

Also,

$$\mathbb{E} \left( \sum_{j,k=1}^n (\bar{A}_{jk}^n - A_{jk}^n) \right) = \sum_{j,k=1}^n \mathbb{E}(\bar{A}_{jk}^n - A_{jk}^n) \stackrel{(5.10)}{=} 0$$

and

$$n^\alpha [ \langle \widehat{W}^n \rangle_{I_j^n \times I_k^n} - 1 ] \leq \bar{A}_{jk}^n - A_{jk}^n \leq n^\alpha [ \widehat{W}^n \rangle_{I_j^n \times I_k^n} ].$$

Hence, Hoeffding's inequality (5.2) yields

$$\mathbb{P} \left( \frac{1}{n^2} \left| \sum_{j,k=1}^n (\bar{A}_{jk}^n - A_{jk}^n) \right| \geq t \right) \leq \exp \left( -2 \left[ \frac{n^2 t}{n^{1+\alpha}} \right]^2 \right).$$

By choosing now  $t = n^\eta$  with  $\eta \in ]\alpha - 1, 0[$  and applying Borel-Cantelli Lemma we obtain

$$\lim_{n \rightarrow +\infty} \frac{1}{n^2} \left| \sum_{j,k=1}^n (\bar{A}_{jk}^n - A_{jk}^n) \right| \leq \lim_{n \rightarrow +\infty} n^\eta = 0 \quad \text{almost surely.}$$

Coupling with (5.27) implies that the right hand side of (5.26) vanishes almost surely in the  $n \rightarrow +\infty$  limit, establishing (5.24). The rest of the proof of Theorem 1.5 (CASE 2) follows mutatis mutandis as in the proof of Theorem 1.5 (CASE 1) and is therefore omitted.

#### ACKNOWLEDGMENTS

This work was motivated by a question posed by Giovanni Naldi, whom the authors wish to thank. They also wish to thank Giuseppe Patanè and all the colleagues of the IMATI workgroup on COVID-19 for interesting discussions, and Enrico Priola and Giuseppe Savaré for clarifying remarks on time-dependent random variables. S.D. is supported by the INdAM GNAMPA 2023 Project *Modelli non-lineari in presenza di interazioni puntuali*. S.D. acknowledges that this study was carried out within the project E53D23005450006 “Nonlinear dispersive equations in presence of singularities” – funded by European Union – Next Generation EU within the PRIN 2022 program (D.D. 104 - 02/02/2022 Ministero dell’Università e della Ricerca). This manuscript reflects only the authors’ views and opinions and the Ministry cannot be considered responsible for them. L.V.S. is a member of the GNAMPA group of INDAM, of the PRIN 2020 Project 20204NT8W4, PRIN 2022 Project 2022YXWSLR, PRIN 2022 PNRR Project P2022XJ9SX, and of the CNR FOE 2022 Project STRIVE.

#### APPENDIX A. NOTATION

For the reader’s convenience we collect here the main notation used in the present paper.

- $\mathbb{R}_+ := [0, +\infty[$
- a.e., for a.e.  $x$ : almost everywhere, for almost every  $x$ , with respect to the standard Lebesgue measure
- $\mathbb{1}_E$ : the characteristic function of the set  $E$
- $\|\cdot\|_{\square}$ : the cut norm defined by (1.15)
- $\delta_{\square}$ : the cut metric defined by (B.1)
- $W_{\mathcal{G}}$ : a step-graphon associated to the undirected graph  $\mathcal{G}$  as in (1.7)
- $I_j^n$ : see (1.6)
- $\langle u \rangle_{I_j^n}, \langle U \rangle_{I_j^n \times I_k^n}$ : see (1.21)
- $\mathcal{G}(n, W)$ : the sparse  $W$ -random graph as in (1.28)
- $\mathcal{G}(n, W, n^{-\alpha})$ : the sparse  $W$ -random graph as in (1.27)
- $n^{\alpha} \mathcal{G}(n, W, n^{-\alpha})$ : the scaled sparse  $W$ -random graph as in (1.26)
- $K_d$ : the constant in (1.13)
- $K_0$ : the constant in (1.17)
- $K_1$ : the constant in (1.18)
- $K_a$ : the constant in (1.30)
- $C_c^{\infty}(\Omega)$ : the set of infinitely differentiable, compactly supported functions on the open set  $\Omega$
- $C^0([0, T])$ : the Banach space of continuous functions defined on the interval  $[0, T]$ , endowed with the norm

$$\|u\|_{C^0} := \max_{t \in [0, T]} |u(t)|.$$

#### APPENDIX B. GRAPHONS AND CONVERGING GRAPH SEQUENCES

In this appendix we briefly overview some notions and results in the theory of graphs and graphons that we need in the present paper. The exposition mainly follows the work [BCCZ19].

##### B.1. Graphons and cut distance.

**Definition B.1.** A *graphon* is a summable function  $W : [0, 1]^2 \rightarrow \mathbb{R}$  satisfying  $W(x, y) = W(y, x)$  for a.e.  $x, y \in [0, 1]$ . In the following we denote by  $\mathcal{W}$  the set of graphons.

Given a weighted undirected graph  $\mathcal{G}_n$  with  $n$  vertices and adjacency matrix  $(A_{jk})_{j,k=1,\dots,n}$ , we can define a step-graphon  $W_{\mathcal{G}_n}$  associated to  $\mathcal{G}_n$  by using formula (1.7) (see also Figure 1).

As anticipated in the Introduction, the theory of graphons makes crucial use of the cut norm defined by (1.15). The notion of cut norm was first introduced by Frieze and Kannan [FK99], and its key importance for graphons has been recently highlighted in a series of papers (see [BCCZ18, BCCZ19, BCLSV06, BCLSV08, BCLSV12] and the references therein). In particular, as pointed out in [BCLSV08] the cut norm is intimately connected with the notion of *left convergence* for graph

sequences (see Remark B.1 below for a very brief discussion of this point and §B.2 for some comments about the analytic properties of the cut norm).

From the graph theory viewpoint, a serious drawback of the notion of cut norm is the following. Assume  $\mathcal{G}$  and  $\mathcal{G}'$  are the very same graph, but have different vertex labeling. Then in general the adjacency matrices of  $\mathcal{G}$  and  $\mathcal{G}'$  are not the same and hence owing to (1.7) and (1.15) we have  $\|W_{\mathcal{G}} - W_{\mathcal{G}'}\|_{\square} \neq 0$ . A way out this issue is the notion of *cut distance*. To define it, we first introduce some notation. First, we recall that a map  $\phi : [0, 1] \rightarrow [0, 1]$  is measure-preserving if, for every measurable set  $A \subseteq [0, 1]$ , the set  $\phi^{-1}(A)$  is also measurable and furthermore  $|\phi^{-1}(A)| = |\phi(A)|$ . Given a graphon  $W$  and a measure preserving map  $\phi$ , we set  $W^{\phi}(x, y) := W(\phi(x), \phi(y))$ , for every  $x, y \in [0, 1]$ .

**Definition B.2.** For every  $U, W \in \mathcal{W}$ , the cut distance  $\delta_{\square}$  between  $U$  and  $W$  is defined as

$$\delta_{\square}(U, W) := \inf_{\phi} \|U - W^{\phi}\|_{\square} \quad (\text{B.1})$$

where the infimum is taken over all measure preserving maps  $\phi : [0, 1] \rightarrow [0, 1]$ .

Note that, strictly speaking,  $\delta_{\square}$  is only a pseudometric and not a distance since the equality  $\delta_{\square}(U, W) = 0$  does not imply that  $U \equiv W$ . As a matter of fact,  $\delta_{\square}(W, W^{\phi}) = 0$ , for every fixed  $W \in \mathcal{W}$  and any given measure preserving map  $\phi$ . For this reason, we will always tacitly identify two graphons with zero cut distance.

*Remark B.1* (Cut distance and left convergence of graph sequences). Let us briefly recall the link between the cut metric and the so called left convergence of graph sequences, redirecting to [BCLSV08] for an extended discussion. Very loosely speaking, a sequence of graphs  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  is left convergent if the *local structure* of the graphs somehow stabilizes in the  $n \rightarrow +\infty$  limit, in the sense that the (suitably normalized) number of copies in  $\mathcal{G}_n$  of any given finite subgraph tends to a limit value as  $n \rightarrow +\infty$  (see [BCLSV08] for the rigorous definition). The connection between left convergence and cut distance is unveiled by [BCLSV08, Theorem 3.8]: if  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  is a sequence of undirected graphs with uniformly bounded adjacency matrices, then  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  is left convergent if and only if  $\lim_{n \rightarrow +\infty} \delta_{\square}(W_{\mathcal{G}_n}, W) = 0$ , for some bounded graphon  $W$ .

**B.2. Analytic properties of the cut distance.** We refer to [Jan14] for a detailed discussion of the cut norm from the point of view of mathematical analysis. Since we have used it in the paper, we recall the relation from [Jan14, equation (4.1)]

$$\|W\|_{\square} \leq \sup_{\|f\|_{\infty}, \|g\|_{\infty} \leq 1} \left| \int_{[0, 1]^2} W(x, y) f(x) g(y) dx dy \right| \leq 4 \|W\|_{\square}. \quad (\text{B.2})$$

where the supremum is taken over all real-valued, measurable functions  $f, g$  defined on  $[0, 1]$ .

It is also interesting to point out that convergence in the cut norm is in between strong and weak convergence in  $L^1([0, 1]^2)$ . Indeed, the cut norm induces a weaker topology than the strong one, see [Lov12, Section 8.3]. On the other hand, one can show that convergence in the cut norm is strictly stronger than the weak convergence, see [Jan14, Appendix F]. Loosely speaking, the reason why convergence in the cut norm is strictly stronger than weak convergence is the following. Weak convergence in  $L^1([0, 1]^2)$  can be characterized as convergence tested against characteristic functions in the form  $\mathbb{1}_{S \times T}$ , with  $S, T \subseteq [0, 1]$ . Convergence in the cut norm requires that this convergence is uniform with respect to  $S$  and  $T$ .

**B.3. Compactness results for dense and sparse graph sequences.** We now quote [BCCZ19, Theorem 2.13] (see also [LS07, Theorem 5.1] for the case  $p = \infty$ ).

**Theorem B.3.** *Assume  $1 < p \leq \infty$  and  $C > 0$  and consider a sequence of graphons  $\{W_n\}_{n \in \mathbb{N}} \subset \mathcal{W}$  such that  $\|W_n\|_p \leq C$  for every  $n \in \mathbb{N}$ . Then there is  $W \in \mathcal{W}$  such that, up to subsequences,*

$$\lim_{n \rightarrow +\infty} \delta_{\square}(W_n, W) = 0.$$

The analogous of Theorem B.3 for  $p = 1$  is false in general, but compactness of bounded  $L^1$  balls can be recovered by adding a uniform integrability assumption, see [BCCZ19, Theorem C.7]. We now recall the following fundamental definitions.

**Definition B.4.** A sequence of undirected graphs  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  is said to be dense or sparse if

$$\liminf_{n \rightarrow +\infty} \|W_{\mathcal{G}_n}\|_1 > 0 \quad \text{or} \quad \lim_{n \rightarrow +\infty} \|W_{\mathcal{G}_n}\|_1 = 0,$$

respectively. In the previous expression,  $W_{\mathcal{G}_n}$  is the step graphon defined as in (1.7).

The quantity  $\|W_{\mathcal{G}_n}\|_1$  essentially represents the edge density of the graph. According to the previous definition, if  $\mathcal{G}_n$  is a sequence of dense simple graphs then the number of edges is  $O(n^2)$  as  $n$  is large enough, whereas if  $\mathcal{G}_n$  is a sequence of sparse simple graphs then the number of edges is  $o(n^2)$  as  $n \rightarrow +\infty$ . Given a sequence of dense graphs  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$ , by applying Theorem B.3 and its analogue in the case  $p = 1$  [BCCZ19, Theorem C.7] to the sequence  $\{W_{\mathcal{G}_n}\}_{n \in \mathbb{N}}$  we obtain useful compactness results. The same results apply to sequences of sparse graphs, but in this case the associated sequence of graphons converges to the uninformative limit  $W \equiv 0$ . As a matter of fact, however, most of the graph sequences relevant for real-world applications are sparse, see the introduction to [BCCZ19]. As pointed out in [BCCZ19], a way to circumvent this obstruction and extract information from the asymptotic behavior of sparse sequences is to normalize  $\mathcal{G}_n$  and consider the sequence  $\{\mathcal{G}_n / \|W_{\mathcal{G}_n}\|_1\}_{n \in \mathbb{N}}$ . Before discussing a selection of the main results in [BCCZ19] we have to introduce some further notation.

**Definition B.5** ( $(C, \eta) - L^p$  upper regular graph). Fix  $C, \eta > 0$ . Given  $p \in ]1, +\infty[$  and a graph  $\mathcal{G}$  with  $|V(\mathcal{G})| = n \geq \eta^{-1}$  vertices and weighted adjacency matrix  $A = (A_{k\ell})_{k, \ell=1}^n$ , we say that  $\mathcal{G}$  is a  $(C, \eta) - L^p$  upper regular graph if the following holds. Let  $\mathcal{P} := \{V_1, \dots, V_m\}$  be any partition of  $V(\mathcal{G})$  into disjoint sets such that  $|V_j| \geq \eta n$  for every  $j = 1, \dots, m$ . Then

$$\sum_{i,j=1}^m \frac{|V_i||V_j|}{n^2} \left| \sum_{k \in V_i, \ell \in V_j} \frac{A_{k\ell}}{|V_i||V_j|} \right|^p \leq C^p \|W_{\mathcal{G}}\|_1^p. \quad (\text{B.3})$$

For  $p = \infty$ , the definition is analogous: it suffices to replace (B.3) with

$$\max_{1 \leq i, j \leq m} \left| \sum_{k \in V_i, \ell \in V_j} \frac{A_{k\ell}}{|V_i||V_j|} \right| \leq C \|W_{\mathcal{G}}\|_1.$$

Very loosely speaking, a graph  $\mathcal{G}$  is  $(C, \eta) - L^p$  upper regular if, whenever its vertices are divided into a certain number of groups none of which is too small, then averaging the edge weights with respect to the partition gives a weighted graph with bounded  $L^p$  norm. Note furthermore that if  $p = 1$  then equation (B.3) is always satisfied with  $C = 1$ , so the definition is only meaningful if  $p > 1$ . We now quote [BCCZ19, Definition 2.7].

**Definition B.6.** Fix  $p \in ]1, +\infty[$  and assume that  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  is a sequence of sparse graphs such that  $|V(\mathcal{G}_n)| = n$ . We say that  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  is a  $C$ -upper  $L^p$  regular sequence if, for every  $\eta > 0$ , there is  $n_0 = n_0(\eta)$  so that  $\mathcal{G}_n$  is  $(C + \eta, \eta)$ -upper  $L^p$  regular, for every  $n \geq n_0$ .

Note that [BCCZ19, Proposition A.1] implies that, if  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  is a  $C$ -upper  $L^p$  regular sequence of simple graphs for some  $p > 1$ , then the average degree of  $\mathcal{G}_n$  blows up as  $n \rightarrow +\infty$ . The importance of Definition B.6 is given by the next result [BCCZ19, Theorem 2.8, Proposition 2.10].

**Theorem B.7.** Fix  $p \in ]1, +\infty[$  and  $C > 0$ . Assume that  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  is a  $C$ -upper  $L^p$  regular sequence of graphs. Then there is  $W \in \mathcal{W}$  such that  $\|W\|_p \leq C$  and up to subsequences

$$\lim_{n \rightarrow +\infty} \delta_{\square} \left( \frac{W_{\mathcal{G}_n}}{\|W_{\mathcal{G}_n}\|_1}, W \right) = 0. \quad (\text{B.4})$$

Conversely, if  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  is a sequence of graphs such that for some  $W \in \mathcal{W}$  with  $\|W\|_p < +\infty$  equation (B.4) holds true, then  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  is a  $\|W\|_p$ -upper  $L^p$  regular sequence.

In the case  $p = 1$  the compactness result is still true provided we replace the assumption that  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  is a  $C$ -upper  $L^p$  regular sequence of graphs with the assumption that  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  is a uniformly upper regular sequence, see [BCCZ19, Theorem C.13].

**B.3.1. Convergence in the cut metric and convergence in the cut norm.** Summing up, cut-metric convergence provides a convenient framework to discuss the asymptotic properties of sequences of both dense and sparse large graphs. As a matter of fact, however, one often consider sequences converging in the cut norm. Although this might sound like a stronger assumption, the following proposition states that the two conditions are equivalent, up to a relabeling of the vertices. For the proof we refer to [BCLSV08, Lemma 5.3] in the case  $p = +\infty$  and [BCCZ19, Proposition 5.2] in the case  $1 < p < \infty$ .

**Proposition B.8.** *If  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  is a sequence of graphs satisfying the assumptions of either Theorem B.3 or Theorem B.7, then there are  $W \in \mathcal{W}$  and, for every  $n$ , a labeling of the vertices of  $\mathcal{G}_n$  such that*

$$\|W_{\mathcal{G}_n} - W\|_{\square} \rightarrow 0 \quad \text{or} \quad \left\| \frac{W_{\mathcal{G}_n}}{\|W_{\mathcal{G}_n}\|_1} - W \right\|_{\square} \rightarrow 0$$

as  $n \rightarrow +\infty$ , respectively.

*Remark B.2.* By combining Theorems B.3–B.7 we can conclude that assumption (1.16) is satisfied by a large class of sequences of undirected graphs. In particular, owing to Theorem B.3 it is satisfied up to subsequences and vertex relabelings by any sequence  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  such that  $\|W_{\mathcal{G}_n}\|_p$  is uniformly bounded. If the sequence  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  is dense, then the limit  $W$  we obtain this way is already meaningful. If the sequence  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  is sparse, then the limit is the uninformative trivial graphon  $W \equiv 0$ . However, Theorem B.7 implies that, in the sparse case, assumption (1.16) is satisfied (up to subsequences and vertex relabelings) by the normalized sequence  $\{\mathcal{G}_n/\|W_{\mathcal{G}_n}\|_1\}_{n \in \mathbb{N}}$  provided  $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$  is  $C$ -upper  $L^p$  regular for some  $C > 0$  and  $p > 1$ .

## REFERENCES

- [ABS21] Alimohammadi Y., Borgs C., Saberi A., *Algorithms using local graph features to predict epidemics*, Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA) (2022), 3430–3451.
- [ABNPV21] Almeida L., Bliman P.-A., Nadin G., Perthame B., Vauchelet N., *Final size and convergence rate for an epidemic in heterogeneous population*, Math. Models Methods Appl. Sci. **31**(5) (2021), 1021–1051.
- [BBPSV05] Barthelemy M., Barrat A., Pastor-Satorras R., Vespignani A., *Dynamical patterns of epidemic outbreaks in complex heterogeneous networks*, J. Theoretical Biology **235** (2005), 275–288.
- [BRCDF20] Bellomo N., Bingham R., Chaplain M.A.J., Dosi G., Forni G., Knopoff D.A., Lowengrub J., Twarock R., Virgillito M.E., *A multiscale model of virus pandemic: heterogeneous interactive entities in a globally connected world*, Math. Models Methods Appl. Sci. **30** (2020), 1591–1651.
- [BS02] Benjamini I., Schramm O., *Recurrence of distributional limits of finite planar graphs*, Electr. J. Probab. **6** (2001), 1–13.
- [Ber60] Bernoulli D., *Essai d’une nouvelle analyse de la mortalité causée par la petite vérole et des avantages de l’inoculation pour la prévenir*, Mémoires de Mathématiques et de Physique, Académie Royale des Sciences, Paris, 1760, 1–45.
- [BCCG21] Borgs C., Chayes J.T., Cohn H., Ganguly S., *Consistent nonparametric estimation for heavy-tailed sparse graphs*, Ann. Statist. **49**(4) (2021), 1904–1930.
- [BCCZ18] Borgs C., Chayes J.T., Cohn H., Zhao Y., *An  $L^p$  theory of sparse graph convergence II: LD convergence, quotients, and right convergence*, Annals of Prob. **45** (2018), 337–396.
- [BCCZ19] Borgs C., Chayes J.T., Cohn H., Zhao Y., *An  $L^p$  theory of sparse graph convergence I: Limits, sparse random graph models, and power law distributions*, Trans. Am. Math. Soc. **372** (2019), 3019–3062.
- [BCL10] Borgs C., Chayes J.T., Lovász L., *Moments of two-variable functions and the uniqueness of graph limits*, Geometr. Funct. Anal. **19** (2010), 1597–1619.
- [BCLSV06] Borgs C., Chayes J.T., Lovász L., Sós V.T., Vesztergombi K., *Counting graph homomorphisms*, Top. Discr. Math. Ser. Algor. Combin. **26** (2006), 315–371.
- [BCLSV08] Borgs C., Chayes J.T., Lovász L., Sós V.T., Vesztergombi K., *Convergent sequences of dense graphs I. Subgraph frequencies, metric properties and testing*, Adv. Math. **219** (2008), 1801–1851.
- [BCLSV12] Borgs C., Chayes J.T., Lovász L., Sós V.T., Vesztergombi K., *Convergent sequences of dense graphs II. Multiway cuts and statistical physics*, Ann. Math. **176** (2012), 151–219.
- [BCD20] Braides A., Cermelli P., Dovetta S.,  $\Gamma$ -limit of the cut functional on dense graph sequences, ESAIM: COCV **26** (2020) 26.
- [BCC01] Brauer F., Castillo-Chávez C., *Mathematical Models in Population Biology and Epidemiology*, Springer-Verlag, New York, 2001.
- [BDW08] Brauer F., van den Driessche P., Wu J. (Editors), *Lecture Notes in Mathematical Epidemiology*, Springer-Verlag, Berlin, Heidelberg, 2008.
- [Cin11] Çinlar E., *Probability and Stochastics*, Graduate Texts in Mathematics **261**, Springer-Verlag New York, 2011.
- [DDZ22] Delmas J.-F., Dronnier D., Zitt P.-A., *An infinite-dimensional metapopulation SIS model*, J. Diff. Eq. **313** (2022), 1–53.

- [DG05] Daley D.J., Gani J., *Epidemic Modeling: An Introduction*, Cambridge University Press, Cambridge, 2005.
- [DF81] Diaconis P., Freedman D., *On the statistics of vision: the Julesz Conjecture*, *J. Math. Psychology* **2** (1981), 112–138.
- [DGM08] Dorogovtsev S.N., Goltsev A.V., Mendes J.F.F., *Critical phenomena in complex networks*, *Rev. Mod. Phys.* **80**(4), (2008), 1275–1335.
- [FK99] Frieze A., Kannan R., *Quick approximation to matrices and applications*, *Combinatorica* **19** (1999), 175–220.
- [GC19] Gao S., Caines P.E., *Spectral representations of graphons in very large network systems control*, *IEEE Conference on Decision and Control* (2019), 5068–5075.
- [Het00] Hethcote H.W., *The mathematics of infectious diseases*, *SIAM Review* **42**(4) (2000), 599–653.
- [Jan14] Janson S., *Graphons, cut norm and distance, couplings and rearrangements*, Vol. 4 of *New York J. Math. Monogr.* (2013).
- [KM17] Kaliuzhnyi–Verbovetskyi D., Medvedev G.S., *The semilinear heat equation on sparse random graphs*, *SIAM J. Math. Anal.* **49**(2) (2017), 1333–1355.
- [KM22] Kaliuzhnyi–Verbovetskyi D., Medvedev G.S., *Sparse Monte Carlo method for nonlocal diffusion problems*, *SIAM J. Num. Anal.* **60**(6) (2022), 3001–3028.
- [KMK27] Kermack W.O., McKendrick A.G., *Contribution to the mathematical theory of epidemics*, *Proc. Roy. Soc. London A* **115** (1927), 700–721.
- [Lov12] Lovász L., Vol. 60 of *Large networks and graph limits*, American Mathematical Society Colloquium Publications (2012).
- [LS06] Lovász L., Szegedy B., *Limits of dense graph sequences*, *J. Combin. Theory Ser. B* **96**(6) (2006), 933–957.
- [LS07] Lovász L., Szegedy B., *Szemerédi lemma’s for analyst*, *Geom. Funct. Anal.* **17** (2007), 252–270.
- [Med14a] Medvedev G.S., *The nonlinear heat equation on dense graphs and graph limits*, *SIAM J. Math. Anal.* **46**(4) (2014), 2743–2766.
- [Med14b] Medvedev G.S., *The nonlinear heat equation on  $W$ -random graphs*, *Arch. Rational Mech. Anal.* **212**(3) (2014), 781–803.
- [Med19C] Medvedev G.S., *Correction to: The nonlinear heat equation on  $W$ -random graphs*, *Arch. Rational Mech. Anal.* **231**(3) (2019), 1305–1308.
- [Med19] Medvedev G.S., *The continuum limit of the Kuramoto model on sparse random graphs*, *Comm. Math. Sciences* **17**(4) (2019), 883–898.
- [MT20] Medvedev G.S., Tang X., *The Kuramoto model on power law graphs: synchronization and contrast states*, *J. Nonlinear Sci.* **30** (2020), 2405–2427.
- [McK26] McKendrick A.G., *Applications of mathematics to medical problems*, *Proc. Edinburgh Math. Soc.*, **44** (1926), 98–130.
- [MPSV02] Moreno Y., Pastor–Satorras R., Vespignani A., *Epidemic outbreaks in complex heterogeneous networks*, *European Phys. J. B* **26**(4) (2002), 521–529.
- [New02] Newman M.E.J., *Spread of epidemic disease on networks*, *Phys. Rev. E* **66** (2002), 016128.
- [Mur02] Murray J.D., *Mathematical Biology: I. An Introduction*, volume 17 of *Interdisciplinary Applied Mathematics*, Springer–Verlag, New York, third edition, (2002).
- [Mur03] Murray J.D., *Mathematical biology: II. Spatial models and biomedical applications*, volume 18 of *Interdisciplinary Applied Mathematics*. Springer–Verlag, New York, third edition, (2003).
- [NPP16] Nowzari C., Preciado V.M., Pappas G.J., *Analysis and control of epidemics: a survey of spreading processes on complex networks*, *IEEE Control Syst. Mag.* **36**(1) (2016), 26–46.
- [PSCMV15] Pastor–Satorras R., Castellano C., Van Mieghem P., Vespignani A., *Epidemic processes in complex networks*, *Rev. Mod. Phys.* **87**(3) (2015), 925–979.
- [PSV01] Pastor–Satorras R., Vespignani A., *Epidemic dynamics and endemic states in complex networks*, *Phys. Rev. E* **63** (2001), 066117.
- [VFG20] Vizuete R., Frasca P., Garin F., *Graphon–based sensitivity analysis of SIS epidemics*, *IEEE Control Systems Letters* **4**(3) (2020), 542–547.

<sup>1</sup>UNIVERSITÀ DEGLI STUDI MILANO–BICOCCA, DIPARTIMENTO DI MATEMATICA E APPLICAZIONI, VIA ROBERTO COZZI, 55, 20126 MILANO, ITALY.

*Email address:* blanca.ayuso@unimib.it

<sup>2</sup>POLITECNICO DI TORINO, DIPARTIMENTO DI SCIENZE MATEMATICHE, CORSO DUCA DEGLI ABRUZZI 24, 10125 TORINO, ITALY.

*Email address:* simone.dovetta@polito.it

<sup>3</sup>IMATI-CNR, VIA FERRATA 5, 27100 PAVIA, ITALY.

*Email address:* spinolo@imati.cnr.it