CrossMark

*Adv. Appl. Probab.* **54,** 880–912 (2022)
*doi:*10.1017/apr.2022.52

# A FUNCTIONAL CENTRAL LIMIT THEOREM FOR SI PROCESSES ON CONFIGURATION MODEL GRAPHS

WASIUR R. KHUDABUKHSH,* *University of Nottingham*
CASPER WOROSZYLO,** *BHP Billiton*
GRZEGORZ A. REMPAŁA,*** *The Ohio State University*
HEINZ KOEPPL,**** *Technische Universität Darmstadt*

**Abstract**

We study a stochastic compartmental susceptible–infected (SI) epidemic process on a configuration model random graph with a given degree distribution over a finite time interval. We split the population of graph vertices into two compartments, namely, S and I, denoting susceptible and infected vertices, respectively. In addition to the sizes of these two compartments, we keep track of the counts of SI-edges (those connecting a susceptible and an infected vertex) and SS-edges (those connecting two susceptible vertices). We describe the dynamical process in terms of these counts and present a functional central limit theorem (FCLT) for them as the number of vertices in the random graph grows to infinity. The FCLT asserts that the counts, when appropriately scaled, converge weakly to a continuous Gaussian vector semimartingale process in the space of vector-valued càdlàg functions endowed with the Skorokhod topology. We discuss applications of the FCLT in percolation theory and in modelling the spread of computer viruses. We also provide simulation results illustrating the FCLT for some common degree distributions.

*Keywords:* SI process; functional CLT; configuration model; random graphs; scaling limit

2020 Mathematics Subject Classification: Primary 60F17
Secondary 60F05; 92D30

## 1. Introduction

Large-graph scaling limits, such as the functional laws of large numbers (FLLNs) and the functional central limit theorems (FCLTs), of dynamical processes on random graphs have received much attention of late. Although dynamical processes on random graphs themselves have long been studied by mathematicians, physicists, epidemiologists, computer scientists, and engineers, a comprehensive and mathematically rigorous body of work on various scaling

Received 6 February 2018; revision received 25 September 2021.
* Postal address: School of Mathematical Sciences, The University of Nottingham, University Park, Nottingham, NG7 2RD, UK.
** Postal address: 480 Queen Street, Level 12, Brisbane QLD 4000, Australia.
*** Postal address: Mathematical Biosciences Institute, The Ohio State University, Jennings Hall 3rd Floor, 1735 Neil Ave., Columbus, OH 43210, United States of America.
**** Postal address: Bioinspired Communication Systems, Technische Universität Darmstadt, Rundeturmstrasse 12, 64283 Darmstadt, Germany.

© The Author(s), 2022. Published by Cambridge University Press on behalf of Applied Probability Trust.



https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

limits under general settings remains elusive. Such scaling limits have been derived rigorously only for a handful of special cases to date. Notable breakthroughs in the context of epidemiological processes include [2, 3, 21, 31, 30], appearing primarily in the probability theory literature. They provide FLLNs under various sets of technical assumptions. However, scaling limits in the form of FCLTs have not been well investigated to the best of our knowledge. We are not aware of any rigorously derived FCLT for dynamical processes on random graphs, except for diffusion-type approximations attempted in special cases like an early-stage epidemic in [27] and normal approximation theorems for graph statistics in configuration model (CM) random graphs [7]. In this paper, we provide an FCLT for a particular type of binary dynamics on CM random graphs (see [47, Chapter 7], [16]) as $n$, the number of vertices in the graph, grows to infinity.

### 1.1. Our contribution

In the current paper we study a stochastic compartmental susceptible–infected (SI) epidemic process on a CM random graph with a given degree distribution over a finite time interval $\mathcal{T}_0 := [0, T]$, for some $T > 0$. In this setting, we segregate the population into two compartments, namely, S and I, containing the susceptible and the infected individuals, respectively. In addition to the sizes of these two compartments, we keep track of the counts of SI-edges (those connecting a susceptible and an infected individual) and SS-edges (those connecting two susceptible individuals). We describe the dynamical process in terms of these counts and present an FCLT for them as $n$ grows to infinity. To be precise, let $X_{\mathrm{SI},i}(t)$ and $X_{\mathrm{SS},i}(t)$ denote the numbers of infected and susceptible neighbours of a susceptible vertex $i$ at time $t$. Based on these local processes, define

$$X_{\mathrm{SS}}(t) := \sum_{i \in S(t)} X_{\mathrm{SS},i}(t) \quad \text{and} \quad X_{\mathrm{SI}}(t) := \sum_{i \in S(t)} X_{\mathrm{SI},i}(t),$$

where $S(t)$ denotes the set of susceptible vertices at time $t$. We will denote the set of infected vertices at time $t$ by $I(t)$. While $X_{\mathrm{SI}}$ is the count of SI-type edges, the process $X_{\mathrm{SS}}$ counts each SS-type edge twice. Let

$$X(t) := (X_{\mathrm{S}}(t), X_{\mathrm{SI}}(t), X_{\mathrm{SS}}(t))$$

denote the aggregated state vector of the system at time $t \geq 0$, where $X_{\mathrm{S}}(t) := |\, S(t) \,|$ keeps track of the number of susceptible vertices. Also, let

$$X_{\mathrm{S}\bullet}(t) := X_{\mathrm{SI}}(t) + X_{\mathrm{SS}}(t).$$

An FLLN for the SI process approximates the scaled counts $n^{-1}X$ by the solution to a system of ordinary differential equations (ODEs). That is, $n^{-1}X \approx x$ for sufficiently large $n$, where $x$ is the solution to the FLLN limit ODE. In this paper, we prove an FCLT for the fluctuations of the scaled process $n^{-1}X$ around the FLLN limit $x$. Our main result (Theorem 2 in Section 4.3) asserts that, under appropriate technical assumptions, the fluctuation process $Y(t) := \sqrt{n}\big(n^{-1}X(t) - x(t)\big)$ converges weakly to a Gaussian vector semimartingale in $D^{(3)}$, the space of real 3-dimensional vector-valued càdlàg functions on $\mathcal{T}_0$ endowed with the Skorokhod topology.

The technical assumptions needed for the result to hold will be made precise later in Section 3. A precise statement of the FCLT result with additional discussion and further details is presented in Section 4.

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

### 1.2. Proof strategy and paper outline

Our derivation relies on the application of an FCLT for local martingales due to Rebolledo, referred to as the *Rebolledo theorem* hereinafter. In [44], Rebolledo provided sufficient conditions for the convergence of local martingales to a continuous Gaussian (vector) martingale in terms of the associated optional and predictable quadratic variation processes, and the martingale process containing 'big' jumps of the original process. Helland later provided a simpler proof of the Rebolledo theorem in [28]. For our purpose, we do not need the Rebolledo theorem in its full generality; a version of it tailored to the setting of square-integrable martingales suffices, and for the sake of completeness we provide the statement of such a version in Appendix C.

The first step towards the FCLT is to perform a Doob–Meyer decomposition ([36]) of the semimartingale of the vector of counts into a zero-mean martingale and a compensator. Then one can show that the predictable quadratic variation of the appropriately scaled martingale process converges in probability to a deterministic quantity. Furthermore, in the limit, its sample paths turn out to be close to continuous in the sense that their 'big' jumps vanish. Weak convergence in the sense of [14] is then established by applying the Rebolledo theorem.

The rest of the paper is structured as follows. The construction of the CM random graph and the epidemic process on it are described in detail in Section 2. In Section 3, we make our technical assumptions precise and provide a law of large numbers, before presenting our FCLT (Theorem 2) in Section 4. Necessary technical lemmas are also discussed leading up to the FCLT. In Section 5, we discuss applications of our FCLT in percolation theory (from a non-equilibrium statistical mechanics point of view), and in computer science in the context of spread of computer viruses. In Section 5, we also discuss the peculiar case of Poisson-type degree distributions under which the correlation-equations-based mean-field pair approximation approach correctly estimates the limiting process. For illustration, we provide simulation results for some common degree distributions.

## 2. The SI process on CM random graphs

### 2.1. Notational conventions

We use the symbols $\mathbb{N}$ and $\mathbb{R}$ to denote the set of natural numbers and the set of real numbers, respectively. Also, we use $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ and $\mathbb{R}_+ := \mathbb{R} \setminus (-\infty, 0]$. Given $R \subseteq \mathbb{R}$, we denote the $\sigma$-field of Borel subsets of $R$ by $\mathcal{B}(R)$. Recall $\mathcal{T}_0 := [0, T]$ for some $t > 0$. We denote by $D = D(\mathcal{T}_0)$ the space of real functions on $\mathcal{T}_0$ that are right continuous and have left-hand limits. Functions in $D$ are called càdlàg. Unless otherwise mentioned, the space $D$ is assumed to be endowed with the Skorokhod topology [14, Chapter 3], which turns $D$ into a Polish space. We call $D$ the Skorokhod space. Let the triplet $(\Omega, \mathcal{F}, \mathbb{P})$ denote our probability space. For an event $A$, we use $\mathbf{1}(A)$ to denote the indicator (or characteristic) function of $A$. We shall use the shorthand notation $(a)_b = a(a-1)(a-2)\cdots(a-b+1)$ for $a > b$ and $a, b \in \mathbb{N}$. The symbols $O(.)$ and $o(.)$ are the big-O and small-o notation, respectively, and they carry their usual meanings. For a differentiable function $f$ defined on some set $E \subseteq \mathbb{R}^d$, we denote its partial derivative with respect to the $i$th variable by $\partial_i f$, for $i = 1, 2, \ldots, d$. With some abuse of notation, we use $\partial f(x)$ to denote the derivative of a differentiable function of a single variable at $x$. For a stochastic process $Z(t)$ with paths in $D$, we denote the associated jump process $Z(t) - Z(t-)$ by $\delta Z(t)$. We also define $\mathcal{T} := (0, T] \subset \mathcal{T}_0$.

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

### 2.2. Model

We begin with the class of all CM random graphs [47, Chapter 7] with $n$ vertices, for $n \in \mathbb{N}$. The main advantage of the CM is that it allows one to fix the degrees before constructing the graph itself. There are numerous real-life situations where random graphs with a prescribed degree sequence (or a probability distribution for the degree sequence) are reasonable and intuitive. See [47, Chapter 7] for some examples.

Given degrees $d_1, d_2, \ldots, d_n$ for $n$ vertices, we first assign $d_i$ half-edges to vertex $i$. The CM random graph is then obtained by matching (or pairing) all available half-edges uniformly at random. Two paired half-edges form an edge. We assume the degrees $d_1, d_2, \ldots, d_n$ are drawn from a probability distribution, which we call the degree distribution. If the sum of the drawn degrees is not even, we add a parity half-edge to the last vertex. That is, we first draw $d_1, d_2, \ldots, d_{n-1}$ and $d'_n$ as an independent and identically distributed random sample from the degree distribution and then set $d_n := d'_n + 1\Big(d'_n + \sum_{j=1}^{n-1} d_j \text{ is odd}\Big)$. However, for large $n$ the contribution of the added indicator function is negligible with probability one (see, e.g., [47, Section 7.6, p. 239]), and we will thus ignore it in our asymptotic calculations.

Additionally, we assume the first three moments of the degree distribution are finite. Note that, by matching half-edges uniformly at random, we obtain a *multigraph*, because the resulting graph may have self-loops and multiple edges. Interestingly, when the degrees $d_1, d_2, \ldots, d_n$ are drawn independently, [23, Theorem 3.1.2] states that as the number of vertices grows to infinity, the numbers of self-loops and parallel edges have independent Poisson limits whose means depend only on the first two moments of the degree distribution. Since the focus of this paper is on the functional limit of the fluctuations of the scaled stochastic process $n^{-1}X$, the contributions of self-loops and parallel edges are negligible in the limit. Therefore, we will ignore self-loops and parallels edges in our calculations of the asymptotic terms.

Let us denote the probability generating function of the underlying degree distribution by $\psi$, i.e.,

$$\psi(x) := \sum_k x^k p_k, \tag{2.1}$$

where $p_k$ is the probability that a randomly chosen vertex has degree $k$, for $k \in \mathbb{N}_0$. We denote the class of all CM random graphs with $n$ vertices by $\mathcal{G}(\psi, n)$.

A half-edge is referred to as SI-type (resp., IS-type) if it originates at a susceptible (resp., infected) vertex and is a part of an SI-type edge. Similarly, an SS-type half-edge is a part of an SS-type edge and an II-type half-edge is a part of an II-type edge (connecting two infected individuals).

We consider the Markovian SI model on CM random graphs. Each infected individual (represented by a vertex of the graph) infects one of its neighbours at rate $\beta > 0$, independently of the other neighbours. We split the population into two compartments consisting respectively of the susceptible and the infected individuals. In order to prove the FCLT result outlined in the introduction, we need to approximate the probability distributions of the counts $X_{\mathrm{S}}$, $X_{\mathrm{SI}}$, and $X_{\mathrm{SS}}$, which requires us to calculate the probability that a susceptible vertex of degree $k$ has $l$ infected neighbours, for $l \le k$. As we explain below, obtaining this probability is particularly easy if we allow the graph to be constructed dynamically as the infection spreads. Consequently, we adopt the following dynamic construction of the graph first proposed in [21].

**Dynamic construction of the graph.** Suppose we are given $n$ vertices with degrees (or numbers of half-edges) $d_1, d_2, \ldots, d_n$. A subset of size $X_{\mathrm{S}}(0)$ of $n$ vertices is chosen uniformly at

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

random. Those vertices are designated as susceptible, and the remaining vertices as infected. Note that since we can write

$$X_{\mathrm{SI}}(t) = \sum_{i \in I(t)} X_{\mathrm{IS},i}(t),$$

where $X_{\mathrm{IS},i}(t)$ denotes the number of IS-type half-edges originating at an infected vertex $i$ at time $t$, it suffices to follow the infected vertices in the dynamic construction.

We start at time $t = 0$ by infecting a random subset of vertices and revealing their connections by matching their half-edges uniformly at random with the available half-edges (see below). We then associate with each IS-type half-edge an independent exponential clock with parameter $\beta$. The first of all these clocks that rings determines the next event. Therefore, at time $t$ we know that the next infection will take place in a time exponentially distributed with parameter $\beta X_{\mathrm{SI}}(t)$. Let $t^* > t$ denote the time of the next infection, with its left limit denoted by $t^{*-}$. At $t^*$, we update the state of the epidemic as follows.

Step 1. We randomly match the IS-type half-edge that has rung to a half-edge originating from a susceptible vertex. This susceptible vertex is newly infected with degree, say, $k$.

Step 2. We choose uniformly $k-1$ half-edges among all the available half-edges (they are either of type IS or of type SS). Let $n_{\mathrm{IS}}$ (resp., $n_{\mathrm{SS}}$) be the number of IS-type (resp., of SS-type) half-edges drawn among these $k-1$ half-edges. Note that the probability of drawing a specific pair $(n_{\mathrm{IS}}, n_{\mathrm{SS}})$ is given by the hypergeometric distribution

$$\frac{\binom{X_{\mathrm{SI}}(t^{*-})}{n_{\mathrm{IS}}}\binom{X_{\mathrm{SS}}(t^{*-})}{n_{\mathrm{SS}}}}{\binom{X_{\mathrm{S}\bullet}(t^{*-})}{k-1}}.$$

The chosen $n_{\mathrm{IS}}$ half-edges of type IS determine the infected neighbours of the newly infected individual. The remaining $n_{\mathrm{SS}}$ edges of type SS remain open in the sense that the susceptible neighbour is not fixed.

Step 3. We change the status of the $n_{\mathrm{IS}}$ (resp., $n_{\mathrm{SS}}$) IS-type (resp., SS-type) edges created to II-type (resp., SI-type).

Step 4. We change the status of the newly infected from S to I and wait for the next clock to ring at a new time $t^{**} > t^*$ to repeat Steps 1–4. The process stops when there are no more IS-type half-edges.

The above algorithm (see Figure 1 for a pictorial representation) gives rise to the natural filtration $\mathcal{F}_t$ as the $\sigma$-field generated by the process history up to and including time $t > 0$ [22, Chapter 1, p. 14]. The filtration $\mathcal{F}_t$ includes information about the status of all vertices (susceptible or infected), and all paired and unpaired half-edges. To be precise, we define

$$\mathcal{F}_t := \sigma\left(\{S(s), I(s), \{X_{\mathrm{IS},i}, X_{\mathrm{II},i}\}_{i \in I(s)} \mid s \le t\}\right).$$

We let $\mathcal{F}_0$ contain *all* $\mathbb{P}$-null sets in $\mathcal{F}$. We include all $\mathbb{P}$-null sets in $\mathcal{F}_0$ so that the filtration family $\{\mathcal{F}_t\}$ is complete. It is also right continuous (i.e., $\mathcal{F}_{t+} = \mathcal{F}_t$ for every $t \ge 0$, where $\mathcal{F}_{t+} := \cap_{s>0}\mathcal{F}_{t+s}$ is the $\sigma$*-field of events immediately after* $t$), because it is generated by a right continuous jump process [1, Chapter II, p. 61]. Therefore, the usual Dellacherie conditions on $\{\mathcal{F}_t\}$ are satisfied [44, 33, 25]. We denote the $\sigma$-field of events strictly prior to $t \in \mathbb{R}_+$ by $\mathcal{F}_{t-}$ and write $\mathcal{F}_{0-} := \mathcal{F}_0$, by convention.

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

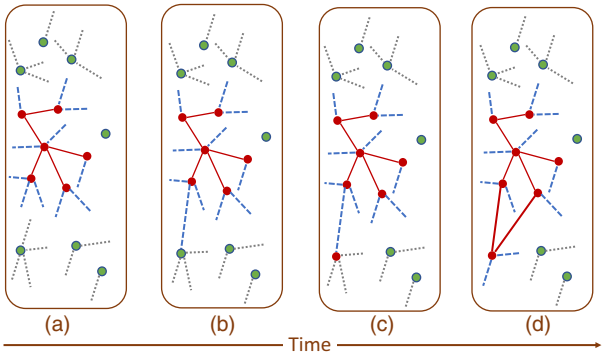



FIGURE 1. Dynamics of the stochastic SI model over a finite time interval $\mathcal{T}_0$. Susceptible vertices are shown in yellowish green and infected vertices in dark red. Different types of edges are shown in different colours. (a) State of the system at a given time. We know the degree of each vertex. For each infected vertex, we know the number of infected neighbours it has (the number of II-type edges, which are shown in solid red) and the number of IS-type half-edges emanating from it, which are shown in dashed blue. The green dotted half-edges emanating from the susceptible vertices are of undetermined types. (b) An IS-type half-edge rings and is randomly matched with a half-edge originating from a susceptible vertex. A full edge is thus formed. (c) We change the status of the chosen susceptible vertex to infected (and change its colour to dark red). (d) We determine the types of the remaining half-edges of the newly infected vertex.

Denote the collection of vertices of degree $k$ that remain susceptible at time $t$ by $S_k(t)$. Recall that $X(t) := \big(X_{\mathrm{S}}(t), X_{\mathrm{SI}}(t), X_{\mathrm{SS}}(t)\big)$ denotes the aggregated state vector of the system at time $t \geq 0$. Note that, since we may calculate $X_{\mathrm{SS}}(t)$ if we know $X_{\mathrm{SI}}(t)$ and $X_{\mathrm{II}}(t)$, the process $X$ is adapted to $\mathcal{F}_t$. However, the process $X$ itself is not necessarily Markovian. Let $\tilde{X} := (X_{\mathrm{SI},i}, X_{\mathrm{SS},i})_{i\in S}$. By the Doob–Meyer decomposition theorem [36], we decompose $X$ as

$$X(t) = X(0) + \int_0^t \mathbb{F}_{\mathrm{X}}\big(\tilde{X}(s)\big)\ \mathrm{d}s + M'(t), \tag{2.2}$$

where $M'(t) := (M'_{\mathrm{S}}(t), M'_{\mathrm{SI}}(t), M'_{\mathrm{SS}}(t))$ is a zero-mean martingale adapted to the filtration $\mathcal{F}_t$, and $\mathbb{F}_{\mathrm{X}}\big(\tilde{X}\big) := \big(\mathbb{F}_{\mathrm{S}}\big(\tilde{X}\big), \mathbb{F}_{\mathrm{SI}}\big(\tilde{X}\big), \mathbb{F}_{\mathrm{SS}}\big(\tilde{X}\big)\big)$ is an integrable function given by

$$\begin{aligned}
\mathbb{F}_{\mathrm{S}}(\tilde{x}) &:= -\beta x_{\mathrm{SI}},\\
\mathbb{F}_{\mathrm{SI}}(\tilde{x}) &:= \sum_{i\in S} \beta x_{\mathrm{SI},i}\big(x_{\mathrm{SS},i} - x_{\mathrm{SI},i}\big),\\
\mathbb{F}_{\mathrm{SS}}(\tilde{x}) &:= -2\sum_{i\in S} \beta x_{\mathrm{SI},i} x_{\mathrm{SS},i},
\end{aligned} \tag{2.3}$$

for $\tilde{x} := (x_{\mathrm{SI},i}, x_{\mathrm{SS},i})_{i\in S}$ and $x_{\mathrm{SI}} := \sum_{i\in S} x_{\mathrm{SI},i}$, $x_{\mathrm{SS}} := \sum_{i\in S} x_{\mathrm{SS},i}$. The first equation follows from the fact that the number of susceptible vertices decreases by 1 at rate $\beta X_{\mathrm{SI}}$ because of the Markovian assumption on the SI process and the independence of the infection transmissions along the edges. Similarly, when a susceptible vertex $i$ gets infected, which happens at rate $\beta X_{\mathrm{SI},i}$ (again because of the independence of infection transmissions along each of its $X_{\mathrm{SI},i}$ infectious edges), the SI edges connected to that susceptible vertex turn into II-type edges, whereas its SS-type edges turn into SI-type edges. Therefore, the net change in the total count of SI-edges is $\big(X_{\mathrm{SS},i} - X_{\mathrm{SI},i}\big)$ whenever a susceptible vertex $i$ gets infected. Because

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

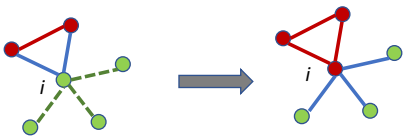



FIGURE 2. Change to the edge counts. When a susceptible vertex $i$ gets infected (turns red from green; see Step 2 of the dynamic construction), the SI edges attached to that vertex turns into II edges, whereas SS edges become new SI edges. This visualization is helpful in understanding the Doob–Meyer decomposition in Equation (2.2).

of the independence assumption, we sum over all susceptible vertices to arrive at the second equation. See Figure 2 for a visualization. The third equation can be explained similarly by computing the rates and the corresponding net changes in the total count of SS-edges when a susceptible vertex gets infected.

We note that as a consequence of the uniformly-at-random matching of half edges in the dynamic construction, similarly to Step 2, we obtain also the following hypergeometric distribution (see also [30]) for the number of infected and susceptible neighbours of a susceptible vertex of degree $k$ at time $t$.

**Lemma 1.** *For $k \in \mathbb{N}_0$ and $i \in S_k(t)$, conditionally on the process history until time $t$, the vector $(X_{\mathrm{SI},i}(t), X_{\mathrm{SS},i}(t))$ follows the hypergeometric distribution*

$$\mathbb{P}\big(X_{\mathrm{SI},i}(t) = n_{\mathrm{SI}}, X_{\mathrm{SS},i}(t) = n_{\mathrm{SS}} \mid \mathcal{F}_t\big) = \frac{\binom{X_{\mathrm{SI}}(t)}{n_{\mathrm{SI}}}\binom{X_{\mathrm{SS}}(t)}{n_{\mathrm{SS}}}}{\binom{X_{\mathrm{S}\bullet}(t)}{k}}, \tag{2.4}$$

*supported on $n_{\mathrm{SI}} + n_{\mathrm{SS}} = k$, where $n_{\mathrm{SI}}, n_{\mathrm{SS}} \in \mathbb{N}_0$.*

*Proof.* Consider $X_{\mathrm{SI}}(t)$, the number of half-edges of type IS emanating from the set of infected vertices. By construction, they all connect to the susceptible vertices uniformly at random. Thus the probability that $n_{\mathrm{SI}}$ of them connect to a given vertex $i \in S_k$ when a total of $X_{\mathrm{S}\bullet}(t)$ half-edges are available to choose from is

$$\frac{\binom{k}{n_{\mathrm{SI}}}\binom{X_{\mathrm{S}\bullet}(t)-k}{X_{\mathrm{SI}}(t)-n_{\mathrm{SI}}}}{\binom{X_{\mathrm{S}\bullet}(t)}{X_{\mathrm{SI}}(t)}},$$

which works out to be the same as the hypergeometric probability (2.4). □

Since we partition the collection of susceptible vertices $S(t)$ by their degree $k \in \mathbb{N}_0$, we have $S(t) = \cup_k S_k(t)$. Therefore, $X_{\mathrm{S}}(t) = \sum_k X_{S_k}(t)$, where $X_{S_k}(t)$ is the size of $S_k(t)$. Note that $X_{\mathrm{S}\bullet}(t) = \sum_k k X_{S_k}(t)$. To study the large graph limit of the system, we also define the quantity

$$\theta(t) := \exp\left(-\beta \int_0^t \frac{X_{\mathrm{SI}}(s)}{X_{\mathrm{S}\bullet}(s)}\,\mathrm{d}s\right), \tag{2.5}$$

which can be intuitively described as the probability that a degree-1 vertex that was susceptible at time zero remains susceptible till time $t > 0$ ([48, 37, 38]). It may be described equivalently as

$$\theta(t) = \theta(0) + \int_0^t \mathbb{F}_0(X_{\mathrm{SI}}(s), X_{\mathrm{S}\bullet}(s), \theta(s))\ \mathrm{d}s,$$

where $\theta(0) = 1$ and $\mathbb{F}_0(X_{\mathrm{SI}}, X_{\mathrm{S}\bullet}, \theta) := -\beta\theta\dfrac{X_{\mathrm{SI}}}{X_{\mathrm{S}\bullet}}$.

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

## 3. The law of large numbers

As done in [30], we make the following technical assumptions. Unless otherwise stated, all limits below and elsewhere in the paper are taken in the large graph limit, i.e. as $n \to \infty$.

**A1** There exists a positive constant $c_{S\bullet}$ such that $\mathbb{P}\big(n^{-1}X_{S\bullet}(T) > c_{S\bullet}\big) \to 1$.

**A2** The fraction of initially susceptible vertices is non-random and converges to some $\alpha_S$, i.e.,

$$n^{-1}X_S(0) \to \alpha_S. \tag{3.1}$$

We also assume that the initially infected and susceptible vertices are selected uniformly at random and $0 < \alpha_S < 1$.

**A3** The degree distribution $\{p_k\}$ satisfies $\sum_k k^3 p_k < \infty$.

We note that by virtue of the uniformly-at-random selection of the infected vertices at time zero, A2 also implies (see [30])

$$\begin{aligned} n^{-1}X_I(0) &\to \alpha_I = 1 - \alpha_S > 0, \\ n^{-1}X_{SI}(0) &\to \alpha_{SI} = \alpha_S(1-\alpha_S)\partial\psi(1), \\ n^{-1}X_{SS}(0) &\to \alpha_{SS} = \alpha_S^2\partial\psi(1). \end{aligned} \tag{3.2}$$

Below, for simplicity, we use the vector notation $\alpha := (\alpha_S, \alpha_{SI}, \alpha_{SS})$. The process $X_I$ captures the number of infected individuals. Although it appears that the assumption A1 may be unnecessary, we retain it for the sake of consistency with the work [30] since in our FCLT derivations we rely on some of the results obtained there.

We note that simulating the stochastic system satisfying A1, A2, and A3 is straightforward with the help of the dynamic construction described above. Indeed, for a given choice of degree distribution satisfying A3 with mean degree $\partial\psi(1)$, we may draw degrees $d_1, d_2, \ldots, d_n$ and assign $d_i$ half-edges to vertex $i$. Then we initialize the number of infected and susceptible nodes as in the dynamic graph construction above, assigning the initial proportions according to $\alpha$. We then run the epidemic process following the dynamic construction up until a finite time $T$, which is independent of $n$, when there are still susceptible vertices left. This simulation algorithm also constructs the stochastic process $X$ satisfying A1, A2, and A3.

For $f : \mathbb{R} \to \mathbb{R}$ and $r \in \mathbb{N}$, the quantity $f^r$ is understood as $(r-1)$-times multiplication of $f$ with itself, for $r \in \mathbb{N}_0$, with the convention $f^0 := 1$. The symbol $\partial^r f$ denotes the $r$th derivative of the function $f$, and by convention we write $\partial f := \partial^1 f$. Now, define the operator $\mathbb{D}^r f$ as

$$\mathbb{D}^r f := f^{r-1}\frac{\partial^r f}{(\partial f)^r}, \tag{3.3}$$

for $f : \mathbb{R} \to \mathbb{R}$ and $r \in \mathbb{N}$, whenever the division of $\partial^r f$ by $(\partial f)^r$ is permissible, i.e., $\partial f \neq 0$. The operator $\mathbb{D}^r$ is used to capture the impact of the graph structure on the limiting dynamics through the degree distribution. Let $x := (x_S, x_{SI}, x_{SS})$, and let $\theta : \mathbb{R}_+ \to [0, 1]$ be a function. Then define $\kappa(\vartheta)$ as

$$\kappa(\vartheta) := \mathbb{D}^2\psi \circ \vartheta. \tag{3.4}$$

Note that

$$\kappa(\vartheta)(x) = \frac{\psi(\vartheta(x))\partial^2(\psi\circ\vartheta)(x))}{(\partial(\psi\circ\vartheta)(x))^2}, \qquad x \in \mathbb{R}. \tag{3.5}$$

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

If $(x, \vartheta)$ is the large graph FLLN limit of $(n^{-1}X, \theta)$, then, following [30, Section 3.3.3], we interpret $\kappa(\vartheta)$ as the limiting ratio of the average excess degree of a susceptible vertex chosen randomly as a neighbour of an infectious individual, to the average degree of a susceptible vertex $\mu_S$. The quantity $\kappa(\vartheta)$ allows us to count various pairs accurately. In general, the operator $\mathbb{D}^{r+1}\psi \circ \vartheta$ recursively compares the excess degree of a susceptible vertex randomly chosen as a neighbour of $r$ infected individuals with that of a randomly chosen susceptible vertex. Therefore, it allows us to count various $r$-configurations (different subgraph structures on $r$ vertices, e.g., triples, quadruples, etc.) accurately in the limit. To be precise, in Lemma 5 in Appendix B, we explicitly show

$$\mathbb{D}^{r+1}\psi \circ \theta = \frac{\mu_S^{(r)}(\theta)}{\mu_S(\theta)}\mathbb{D}^{r}\psi \circ \theta \xrightarrow{\mathrm{P}} \mathbb{D}^{r+1}\psi \circ \vartheta, \tag{3.6}$$

where $\mu_S^{(r)}$ is the average excess degree of a susceptible vertex randomly chosen as a neighbour of $r$ infected individuals. In Appendix B, we calculate these quantities explicitly.

Let us also define the operator $\mathbb{H}(x, \vartheta) := (\mathbb{H}_x(x, \vartheta)\,, \mathbb{H}_0(x, \vartheta)\,)$, where $\mathbb{H}_x(x, \vartheta) := (\mathbb{H}_S(x, \vartheta)\,, \mathbb{H}_{SI}(x, \vartheta)\,, \mathbb{H}_{SS}(x, \vartheta))$ and $\mathbb{H}_0(x, \vartheta)$ are given by

$$\begin{aligned}
\mathbb{H}_S(x, \vartheta) &:= -\beta x_{SI},\\
\mathbb{H}_{SI}(x, \vartheta) &:= \beta\kappa(\vartheta)\frac{x_{SI}}{x_S}(x_{SS} - x_{SI}) - \beta x_{SI},\\
\mathbb{H}_{SS}(x, \vartheta) &:= -2\beta\kappa(\vartheta)\frac{x_{SI}x_{SS}}{x_S},\\
\mathbb{H}_0(x, \vartheta) &:= -\beta\frac{x_{SI}}{\alpha_S\partial\psi(\vartheta)}.
\end{aligned} \tag{3.7}$$

Now, noting that A1 implies $P(X_{S\bullet} > k) \to 1$ for any $k > 0$ and 3 implies $\sum_k k^2 p_k < \infty$, recall the strong law on large graphs due to [30]. In the following we take the norm $\|.\|$ to be the standard Euclidean norm.

**Theorem 1.** *Assume A1, A2, and A3 for a CM graph $\mathcal{G}(\psi, n)$. Then, for any $T > 0$, the following holds:*

$$\sup_{0<t\le T} \|(X(t)/n, \theta(t)) - (x(t), \vartheta(t))\| \xrightarrow{\mathrm{P}} 0,$$

*where $(x, \vartheta) := ((x_S, x_{SI}, x_{SS}), \vartheta)$ is the solution of*

$$(x(t), \vartheta(t)) = (x(0), \vartheta(0)) + \int_0^t \mathbb{H}(x(s), \vartheta(s))\ \mathrm{d}s, \tag{3.8}$$

*with the initial condition $x(0) = \alpha$ and $\vartheta(0) = 1$.*

Observe that in the absence of recovery, the numbers of susceptible and infected individuals are linearly related as $X_S + X_I = n$ in the standard susceptible–infected–recovered (SIR) model. The proof therefore follows immediately from setting the recovery rate in the SIR model to zero and assuming that there is only one layer in [30]. The crucial observation is that the neighbourhood distribution of a susceptible vertex, conditional on the process history, can be expressed as a hypergeometric distribution (see Lemma 1) whose mixed moments can be approximated by the corresponding multinomial ones. This allows us to 'average out' the individual-based quantities such as $X_{SI,i}$ for $i \in S$. The convergence is then established by calculating several quadratic variations. The proof of our FCLT, presented in Section 4, exploits similar calculations.

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

## 4. Functional central limit theorem

Having obtained the FLLN, we now derive an FCLT for $X$ after an appropriate scaling. To this end, we begin by first defining the scaled martingale process

$$M(t) = (M_{\mathrm{S}}(t), M_{\mathrm{SI}}(t), M_{\mathrm{SS}}(t)) := n^{-1/2} M'(t), \tag{4.1}$$

which is square-integrable. We study the quadratic variation of the scaled martingale $M(t)$. The idea is to check whether either the optional or the predictable quadratic variation of the scaled process $M$ converges in probability to a deterministic limit. If either of them does, and if the paths of $M$ become approximately continuous in the limit ('big' jumps disappear), we can make use of the Rebolledo theorem (see Appendix C, as well as [44, 28]) to establish the asymptotic limit.

For each $\epsilon > 0$, define

$$M^{\epsilon}(t) := \left(M^{\epsilon}_{\mathrm{S}}(t), M^{\epsilon}_{\mathrm{SI}}(t), M^{\epsilon}_{\mathrm{SS}}(t)\right) \tag{4.2}$$

to be a vector of square-integrable martingales (only) containing all jumps of $M(t)$ larger in absolute value than $\epsilon$.

We use the shorthand notation $\langle M\rangle(t)$ for the $3 \times 3$ matrix of predictable covariation processes of the components of $M(t)$. That is,

$$\langle M\rangle(t) := \begin{pmatrix} \langle M_{\mathrm{S}}\rangle(t) & \langle M_{\mathrm{S}}, M_{\mathrm{SI}}\rangle(t) & \langle M_{\mathrm{S}}, M_{\mathrm{SS}}\rangle(t) \\ \langle M_{\mathrm{SI}}, M_{\mathrm{S}}\rangle(t) & \langle M_{\mathrm{SI}}\rangle(t) & \langle M_{\mathrm{SI}}, M_{\mathrm{SS}}\rangle(t) \\ \langle M_{\mathrm{SS}}, M_{\mathrm{S}}\rangle(t) & \langle M_{\mathrm{SS}}, M_{\mathrm{SI}}\rangle(t) & \langle M_{\mathrm{SS}}\rangle(t) \end{pmatrix}. \tag{4.3}$$

Here $\langle M_{\mathrm{S}}\rangle(t) := \langle M_{\mathrm{S}}, M_{\mathrm{S}}\rangle(t)$ etc., by convention. See Appendix C for definitions of quadratic variation processes. Define $\langle M^{\epsilon}\rangle$ similarly. We shall study the large graph limits of $\langle M\rangle(t)$ and $\langle M^{\epsilon}\rangle(t)$ as $n \to \infty$ for each $t \in \mathcal{T}_0$. For this purpose, we need the neighbourhood distribution of a susceptible vertex $i$ of degree $k$, i.e., the distribution of $(X_{\mathrm{SI},i}, X_{\mathrm{SS},i})$ for a vertex $i \in S_k$, for all $k \in \mathbb{N}$.

We quote an important remark from [30] that will come in handy for the derivations.

**Remark 1.** Note that the total number of edges in the graph is $2^{-1}\sum_i d_i$. With $0 < \alpha_{\mathrm{S}} < 1$, it immediately follows that $n^{-1}X_{\mathrm{SI}} \le n^{-1}X_{\mathrm{S}\bullet} \le \partial\psi(1)$ and $n^{-1}X_{\mathrm{SS}} \le n^{-1}X_{\mathrm{S}\bullet} \le \partial\psi(1)$ for sufficiently large $n \in \mathbb{N}$. Also note that $0 \le \theta \le 1$ and $\partial\psi(\theta) \le \partial\psi(1)$, so that $\alpha_{\mathrm{S}}\theta\partial\psi(\theta) \le \partial\psi(1)$. By virtue of A1, $n^{-1}X_{\mathrm{S}\bullet}$ is bounded away from 0 on $\mathcal{T}_0$, and hence so is $\theta$ (see [30, Section 3.1]). As a consequence of [30, Lemma 1(b)], we can take the same lower bound for $\alpha_{\mathrm{S}}\theta\partial\psi(\theta)$. Let us denote by $\xi > 0$ the uniform lower bound for $n^{-1}X_{\mathrm{S}\bullet}$ and $\alpha_{\mathrm{S}}\theta\partial\psi(\theta)$, so that we can write $n^{-1}X_{\mathrm{S}\bullet} \in [\xi, \partial\psi(1)] \subset \mathbb{R}_+$. Note that all these bounds hold with probability 1.

### 4.1. Deterministic limit of $\langle M\rangle(t)$

Recall that $(x, \vartheta) := ((x_{\mathrm{S}}, x_{\mathrm{SI}}, x_{\mathrm{SS}}), \vartheta)$. Let us begin by defining the following operators acting on the function $(x, \vartheta)$:

$$\begin{aligned} v_{\mathrm{S}} &:= \beta x_{\mathrm{SI}}, \\ v_{\mathrm{SI}} &:= \beta\left(\frac{x_{\mathrm{SI}}(x_{\mathrm{SS}} - x_{\mathrm{SI}})^2}{x_{\mathrm{S}}^2}\mathbb{D}^3\psi(\vartheta) - \frac{x_{\mathrm{SI}}(x_{\mathrm{SS}} - 3x_{\mathrm{SI}})}{x_{\mathrm{S}}}\mathbb{D}^2\psi(\vartheta) + x_{\mathrm{SI}}\right), \end{aligned}$$

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

$$\begin{aligned} v_{\mathrm{SS}} &:= 4\beta \frac{x_{\mathrm{SI}} x_{\mathrm{SS}}}{x_{\mathrm{S}}} \left( \frac{x_{\mathrm{SS}}}{x_{\mathrm{S}}} \mathbb{D}^3 \psi(\vartheta) + \mathbb{D}^2 \psi(\vartheta) \right), \\ v_{\mathrm{S,SI}} &:= -\beta \left( \frac{x_{\mathrm{SI}}(x_{\mathrm{SS}} - x_{\mathrm{SI}})}{x_{\mathrm{S}}} \mathbb{D}^2 \psi(\vartheta) - x_{\mathrm{SI}} \right), \\ v_{\mathrm{S,SS}} &:= 2\beta \frac{x_{\mathrm{SI}} x_{\mathrm{SS}}}{x_{\mathrm{S}}} \mathbb{D}^2 \psi(\vartheta), \\ v_{\mathrm{SI,SS}} &:= -2\beta \frac{x_{\mathrm{SI}} x_{\mathrm{SS}}(x_{\mathrm{SS}} - x_{\mathrm{SI}})}{x_{\mathrm{S}}^2} \mathbb{D}^3 \psi(\vartheta). \end{aligned} \tag{4.4}$$

The intuition behind the operators in Equation (4.4) will be clear when we compute the predictable quadratic variation $\langle M \rangle$ of the martingale process $M$ and seek its limit as $n \to \infty$. In particular, we shall see that each of the terms on the right-hand side of Equation (4.4) is a function of various multinomial moments, which we find as limits of the corresponding hypergeometric ones in order to approximate the expected jump sizes conditional on the process history.

Now, define a $\mathcal{T}$-indexed family of matrices $\{V(t)\}$ as follows

$$V(t) := \begin{pmatrix} V_{\mathrm{S}}(t) & V_{\mathrm{S,SI}}(t) & V_{\mathrm{S,SS}}(t) \\ V_{\mathrm{SI,S}}(t) & V_{\mathrm{SI}}(t) & V_{\mathrm{SI,SS}}(t) \\ V_{\mathrm{SS,S}}(t) & V_{\mathrm{SS,SI}}(t) & V_{\mathrm{SS}}(t) \end{pmatrix}, \tag{4.5}$$

where, given $v_{id_1,id_2}(x, \vartheta)$ for $id_1, id_2 \in \{\mathrm{S}, \mathrm{SI}, \mathrm{SS}\}$ in Equation (4.4),

$$V_{id_1,id_2}(t) := \int_0^t v_{id_1,id_2}(x(s), \vartheta(s))\, \mathrm{d}s, \tag{4.6}$$

with the convention $v_{id_1,id_2} := v_{id_2,id_1}$ for $id_1, id_2 \in \{\mathrm{S}, \mathrm{SI}, \mathrm{SS}\}$ and $v_{id_1,id_2} := v_{id_1}$ whenever $id_1 = id_2 \in \{\mathrm{S}, \mathrm{SI}, \mathrm{SS}\}$. Note that this also sets the convention $V_{id_1,id_2}(t) := V_{id_2,id_1}(t)$ for $id_1, id_2 \in \{\mathrm{S}, \mathrm{SI}, \mathrm{SS}\}$ and $V_{id_1,id_2}(t) := V_{id_1}(t)$ whenever $id_1 = id_2 \in \{\mathrm{S}, \mathrm{SI}, \mathrm{SS}\}$ for each $t \in \mathcal{T}$.

Let us now present our first result providing the deterministic limit of $\langle M \rangle$ in the following lemma. Recall that proving a deterministic limit of $\langle M \rangle$ will satisfy one of the conditions of the Rebolledo theorem. The key strategy in establishing the result will be to approximate various hypergeometric moments by the corresponding multinomial ones.

**Lemma 2.** *Consider the stochastic SI model described in Section 2.2. Assume A1, A2, and A3 for a CM graph $\mathcal{G}(\psi, n)$. Then*

$$\langle M \rangle\,(t) \xrightarrow{\mathrm{P}} V(t),$$

*for each $t \in \mathcal{T}$, as $n \to \infty$, where $V(t)$ is as defined in Equation (4.5), and $(x, \vartheta)$ is the solution of Equation (3.8) with $x(0) = \alpha$ and $\vartheta(0) = 1$.*

*Proof of Lemma 2.* To show convergence of the matrix random process $\langle M \rangle\,(t)$ to $V(t)$, we show componentwise convergence of the respective components. The general strategy for proving convergence for these components remains the same. To conserve space, we only demonstrate here the strategy for establishing $M_{\mathrm{SI}}(t) \xrightarrow{\mathrm{P}} V_{\mathrm{SI}}(t)$. The remaining assertions follow similarly.

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

**Computation of $\langle M_{\mathrm{SI}}\rangle$.** The process $M_{\mathrm{SI}}$ jumps only if a susceptible vertex gets infected. Therefore, the predictable quadratic variation is computed as follows:

$$\langle M_{\mathrm{SI}}\rangle(t) = \left\langle n^{-1/2} M'_{\mathrm{SI}} \right\rangle(t) = \int_0^t \sum_k \frac{1}{n} \sum_{i\in S_k} \beta X_{\mathrm{SI},i}\big(X_{\mathrm{SS},i} - X_{\mathrm{SI},i}\big)^2 \,\mathrm{d}s.$$

Now, for a randomly selected $i \in S_k$, we seek the (conditional) moments $\mathbb{E}\Big[X_{\mathrm{SI},i}\big(X_{\mathrm{SS},i} - X_{\mathrm{SI},i}\big)^2 \big| \mathcal{F}_{t-}\Big]$. Define the function $C_h^k : \mathcal{T}_0 \to \mathbb{R}$ as

$$C_h^k(t) := \mathbb{E}\big[X_{\mathrm{SI},i}(t)(X_{\mathrm{SS},i}(t) - X_{\mathrm{SI},i}(t))^2 \big| \mathcal{F}_{t-}\big].$$

Following the computations in Appendix A, we get

$$\begin{aligned} C_h^k(t) = &\frac{(k)_3 X_{\mathrm{SI}}}{(X_{\mathrm{S}\bullet})_3}\left[(X_{\mathrm{SS}})_2 - 2(X_{\mathrm{SI}}-1)X_{\mathrm{SS}} + (X_{\mathrm{SI}}-1)_2\right] \\ &- \frac{(k)_2 X_{\mathrm{SI}}}{(X_{\mathrm{S}\bullet})_2}[X_{\mathrm{SS}} - 3(X_{\mathrm{SI}}-1)] + k\frac{X_{\mathrm{SI}}}{X_{\mathrm{S}\bullet}}. \end{aligned}$$

To be precise, the processes on the right-hand side of the above equation are evaluated at $t-$. To approximate the hypergeometric moments by corresponding multinomial ones, define the multinomial compensator $C_m^k : \mathcal{T}_0 \times [\xi, \partial\psi(1)] \to \mathbb{R}$ as

$$\begin{aligned} C_m^k(t,z) := &\frac{(k)_3 n^{-3} X_{\mathrm{SI}}}{z^3}\left(X_{\mathrm{SS}}{}^2 - 2X_{\mathrm{SI}}X_{\mathrm{SS}} + X_{\mathrm{SI}}^2\right) \\ &- \frac{(k)_2 n^{-2} X_{\mathrm{SI}}}{z^2}(X_{\mathrm{SS}} - 3X_{\mathrm{SI}}) + \frac{k n^{-1} X_{\mathrm{SI}}}{z} \\ = &\frac{(k)_3 n^{-3} X_{\mathrm{SI}}(X_{\mathrm{SS}} - X_{\mathrm{SI}})^2}{z^3} - \frac{(k)_2 n^{-2} X_{\mathrm{SI}}(X_{\mathrm{SS}} - 3X_{\mathrm{SI}})}{z^2} + \frac{k n^{-1} X_{\mathrm{SI}}}{z}. \end{aligned}$$

Again, to be precise, the processes on the right-hand side of the above equation are evaluated at $t-$. Observe that there exists an $L > 0$ such that

$$C_m^k(t, z(t)) \le Lk^3, \tag{4.7}$$

uniformly in $n$. This holds because $n^{-1}X_{\mathrm{SI}}$ and $n^{-1}X_{\mathrm{SS}}$ are uniformly bounded above by virtue of Remark 1 and $z$ is bounded away from zero, by definition. The function $C_m^k(t, z(t))$ is also Lipschitz continuous in $z$. Now recall the definition of $\nu_{\mathrm{SI}}$ from Equation (4.4) and define

$$\begin{aligned} \Delta(t) := &\sum_k \frac{1}{n} \sum_{i\in S_k} \beta X_{\mathrm{SI},i}(t)\big(X_{\mathrm{SS},i}(t) - X_{\mathrm{SI},i}(t)\big)^2 - \nu_{\mathrm{SI}}(x(t), \vartheta(t)) \\ = &\sum_k \frac{1}{n} \sum_{i\in S_k} \beta X_{\mathrm{SI},i}(t)\big(X_{\mathrm{SS},i}(t) - X_{\mathrm{SI},i}(t)\big)^2 - \nu_{\mathrm{SI}}\big(n^{-1}X(t), \theta(t)\big) \\ &+ \nu_{\mathrm{SI}}(n^{-1}X(t), \theta(t)) - \nu_{\mathrm{SI}}(x(t), \vartheta(t)) \\ = &\,\Delta_1(t) + \Delta_2(t), \end{aligned}$$

where

$$\Delta_1(t) := \sum_k \frac{1}{n} \sum_{i\in S_k} \beta X_{\mathrm{SI},i}(t)(X_{\mathrm{SS},i}(t) - X_{\mathrm{SI},i}(t))^2 - \nu_{\mathrm{SI}}\big(n^{-1}X(t), \theta(t)\big),$$

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

and $\Delta_2(t) := v_{\mathrm{SI}}(n^{-1}X(t), \theta(t)) - v_{\mathrm{SI}}(x(t), \vartheta(t))$. To show $\langle M_{\mathrm{SI}} \rangle \xrightarrow{\mathrm{P}} V_{\mathrm{SI}}$, it suffices to show $\sup_{t\in\mathcal{T}} |\Delta(t)| \xrightarrow{\mathrm{P}} 0$. We achieve this by separately showing $\sup_{t\in\mathcal{T}} |\Delta_1(t)| \xrightarrow{\mathrm{P}} 0$ and $\sup_{t\in\mathcal{T}} |\Delta_2(t)| \xrightarrow{\mathrm{P}} 0$.

**Convergence of** $\Delta_1(t)$. Observe that

$$\begin{aligned}\Delta_1(t) &= \sum_k \frac{1}{n} \sum_{i\in S_k} \beta X_{\mathrm{SI},i}\big(X_{\mathrm{SS},i} - X_{\mathrm{SI},i}\big)^2 - \beta\Bigg[\frac{n^{-3}X_{\mathrm{SI}}(X_{\mathrm{SS}} - X_{\mathrm{SI}})^2}{\alpha_{\mathrm{S}}^2} \frac{\partial^3\psi(\theta)}{(\partial\psi(\theta))^3} \\ &\quad - \frac{n^{-2}X_{\mathrm{SI}}(X_{\mathrm{SS}} - 3X_{\mathrm{SI}})}{\alpha_{\mathrm{S}}} \frac{\partial^2\psi(\theta)}{(\partial\psi(\theta))^2} + n^{-1}X_{\mathrm{SI}}\Bigg] \\ &= \sum_k \Bigg[\frac{1}{n} \sum_{i\in S_k} \beta X_{\mathrm{SI},i}\big(X_{\mathrm{SS},i} - X_{\mathrm{SI},i}\big)^2 - \beta\Bigg\{\frac{n^{-3}X_{\mathrm{SI}}(X_{\mathrm{SS}} - X_{\mathrm{SI}})^2}{\alpha_{\mathrm{S}}^2} \frac{(k)_3\theta^k p_k}{(\theta\partial\psi(\theta))^3} \\ &\quad - \frac{n^{-2}X_{\mathrm{SI}}(X_{\mathrm{SS}} - 3X_{\mathrm{SI}})}{\alpha_{\mathrm{S}}} \frac{(k)_2\theta^k p_k}{(\theta\partial\psi(\theta))^2} + n^{-1}X_{\mathrm{SI}} \frac{k\theta p_k}{\theta\partial\psi(\theta)}\Bigg\}\Bigg] \\ &= \sum_k \Bigg[\frac{1}{n} \sum_{i\in S_k} \beta X_{\mathrm{SI},i}\big(X_{\mathrm{SS},i} - X_{\mathrm{SI},i}\big)^2 - \beta\alpha_{\mathrm{S}} p_k \theta^k C_m^k(t, \alpha_{\mathrm{S}}\theta\partial\psi(\theta))\Bigg].\end{aligned}$$

The second equality follows from expressing the derivatives of the probability generating functions as sums over all possible degrees $k$ and then collecting terms involving degree $k$. Define

$$\Delta_1^{(k)}(t) := \frac{1}{n} \sum_{i\in S_k} \beta X_{\mathrm{SI},i}\big(X_{\mathrm{SS},i} - X_{\mathrm{SI},i}\big)^2 - \beta\alpha_{\mathrm{S}} p_k \theta^k C_m^k(t, \alpha_{\mathrm{S}}\theta\partial\psi(\theta)).$$

Our task boils down to showing that $\sup_{t\in\mathcal{T}} \left|\sum_k \Delta_1^{(k)}(t)\right| \xrightarrow{\mathrm{P}} 0$ as $n \to \infty$. We achieve this in two steps. First we show that the tails of $\sum_k \Delta_1^{(k)}(t)$ are negligible. Second, we show that each term $\Delta_1^{(k)}(t)$ converges to zero uniformly in probability for a fixed $k \in \mathbb{N}$.

**Step I: Tails are negligible.** Let us begin by showing that as $N \to \infty$,

$$\sup_{n\in\mathbb{N}} \sup_{t\in\mathcal{T}} \left|\sum_{k>N} \Delta_1^{(k)}(t)\right| \xrightarrow{\mathrm{P}} 0.$$

Observe that, for sufficiently large $n$,

$$\left|\frac{1}{n} \sum_{k>N} \sum_{i\in S_k} \beta X_{\mathrm{SI},i}\big(X_{\mathrm{SS},i} - X_{\mathrm{SI},i}\big)^2\right| \le \frac{\beta}{n} \sum_{k>N} k^3 X_{S_k} \le 2\beta \sum_{k>N} k^3 p_k, \tag{4.8}$$

because $n^{-1}X_{S_k} \le 2p_k$ for sufficiently large $n$ in light of Remark 1. Following Remark 1 and the bound on $C_m^k$ from Equation (4.7), we get

$$\left|\sum_{k>N} \beta\alpha_{\mathrm{S}} p_k \theta^k C_m^k(t, \alpha_{\mathrm{S}}\theta\partial\psi(\theta))\right| \le \beta L \sum_{k>N} k^3 p_k. \tag{4.9}$$

Therefore, we get $\sup_{n\in\mathbb{N}} \sup_{t\in\mathcal{T}} \left|\sum_{k>N} \Delta_1^{(k)}(t)\right| \xrightarrow{\mathrm{P}} 0$ by combining the inequalities (4.8) and (4.9), in view of A3.

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

**Step II: Uniform convergence in probability for a fixed k.** In addition to Step I, it is sufficient to show $\sup_{t\in\mathcal{T}}|\Delta_1^{(k)}(t)| \xrightarrow{\mathrm{P}} 0$ for an arbitrarily fixed $k\in\mathbb{N}$ to justify $\sup_{t\in\mathcal{T}}|\Delta_1(t)| \xrightarrow{\mathrm{P}} 0$. Observe that

$$
\begin{aligned}
|\Delta_1^{(k)}(t)| &= \left|\frac{1}{n}\sum_{i\in S_k}\beta X_{\mathrm{SI},i}\big(X_{\mathrm{SS},i}-X_{\mathrm{SI},i}\big)^2-\beta\alpha_{\mathrm{S}}p_k\theta^k C_m^k(t,\alpha_{\mathrm{S}}\theta\partial\psi(\theta))\right| \\
&\le \beta n^{-1}\left|\sum_{i\in S_k}X_{\mathrm{SI},i}\big(X_{\mathrm{SS},i}-X_{\mathrm{SI},i}\big)^2-X_{S_k}C_h^k(t)\right| && (4.10)\\
&\quad+\beta n^{-1}X_{S_k}\big|C_h^k(t)-C_m^k\big(t,n^{-1}X_{\mathrm{S}\bullet}\big)\big| && (4.11)\\
&\quad+\beta\big|n^{-1}X_{S_k}C_m^k\big(t,n^{-1}X_{\mathrm{S}\bullet}\big)-\alpha_{\mathrm{S}}p_k\theta^kC_m^k\big(t,n^{-1}X_{\mathrm{S}\bullet}\big)\big| && (4.12)\\
&\quad+\beta\alpha_{\mathrm{S}}p_k\theta^k\big|C_m^k\big(t,n^{-1}X_{\mathrm{S}\bullet}\big)-C_m^k(t,\alpha_{\mathrm{S}}\theta\partial\psi(\theta))\big|. && (4.13)
\end{aligned}
$$

We show that each of the above summands converges uniformly in probability to zero.

Define the process $\Delta_{1,1}^{(k)}(t):=\sum_{i\in S_k}X_{\mathrm{SI},i}\big(X_{\mathrm{SS},i}-X_{\mathrm{SI},i}\big)^2-X_{S_k}C_h^k(t)$. Observe that $\Delta_{1,1}^{(k)}(t)$ is a zero-mean, piecewise constant, càdlàg martingale with paths in $D$. The jumps of $\Delta_{1,1}^{(k)}(t)$ take place when a vertex of degree $k$ gets infected. The quadratic variation of $\Delta_{1,1}^{(k)}(t)$ is therefore the sum of its squared jumps,

$$
\Big[\Delta_{1,1}^{(k)}\Big](t)=\sum_{s\le t}\Big(\delta\Delta_{1,1}^{(k)}(s)\Big)^2\le k^6n,
$$

because the number of jumps cannot exceed $n$. Therefore, by Doob's martingale inequality we get $\sup_{t\in\mathcal{T}}\big|n^{-1}\Delta_{1,1}^{(k)}(t)\big| \xrightarrow{\mathrm{P}} 0$, since

$$
\mathbb{E}\Big[\Big[\Delta_{1,1}^{(k)}\Big](t)\Big]=\mathbb{E}\Big[\Big(\Delta_{1,1}^{(k)}(t)\Big)^2\Big]=O(n).
$$

That is, the quantity in (4.10) converges uniformly in probability to zero.

For the term in (4.11), take into account $n^{-1}X_{S_k}\le 1$ and observe that

$$
\sup_{t\in\mathcal{T}}\big|C_h^k(t)-C_m^k\big(t,n^{-1}X_{\mathrm{S}\bullet}\big)\big|\le\frac{c_1k^3}{X_{S_\bullet}(T)-2},
$$

for some $c_1>0$, because $X_{\mathrm{S}\bullet}$ is non-increasing on $\mathcal{T}_0$. Therefore, by A1, the quantity in (4.11) converges to zero uniformly in probability.

Now observe that

$$
\begin{aligned}
&\sup_{t\in\mathcal{T}}\big|n^{-1}X_{S_k}C_m^k\big(t,n^{-1}X_{\mathrm{S}\bullet}\big)-\alpha_{\mathrm{S}}p_k\theta^kC_m^k\big(t,n^{-1}X_{\mathrm{S}\bullet}\big)\big| \\
&\qquad\le Lk^3\sup_{t\in\mathcal{T}}\big|n^{-1}X_{S_k}-\alpha_{\mathrm{S}}p_k\theta^k\big|\xrightarrow{\mathrm{P}}0,
\end{aligned}
$$

by virtue of the bound on $C_m^k$ in Equation (4.7) and [30, Lemma 1(a)]. Therefore, the term in (4.12) also converges to zero uniformly in probability.

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

Finally, by virtue of the Lipschitz continuity of $C_m^k(t, z)$ in $z$, we get

$$\sup_{t\in\mathcal{T}} \left|C_m^k\left(t, n^{-1}X_{S\bullet}\right) - C_m^k(t, \alpha_S\theta\partial\psi(\theta))\right| \le c_2 \sup_{t\in\mathcal{T}} \left|n^{-1}X_{S\bullet} - \alpha_S\theta\partial\psi(\theta)\right|,$$

for some $c_2 > 0$. Because $\sup_{t\in\mathcal{T}} \left|n^{-1}X_{S\bullet} - \alpha_S\theta\partial\psi(\theta)\right| \xrightarrow{P} 0$ as shown in [30], we conclude that the term in (4.13) converges to zero uniformly in probability.

Having shown that the terms in (4.10), (4.11), (4.12), and (4.13) converge to zero uniformly in probability, we establish that $\sup_{t\in\mathcal{T}} \left|\Delta_1^{(k)}(t)\right| \xrightarrow{P} 0$ uniformly in probability for any fixed $k \in \mathbb{N}$. Finally, by virtue of Step I and Step II, we obtain $\sup_{t\in\mathcal{T}} |\Delta_1(t)| \xrightarrow{P} 0$.

**Convergence of** $\Delta_2(t)$. Note that $v_{SI}(n^{-1}X, \theta)$ is Lipschitz continuous on its domain, which we can take as $(0, 1] \times [\xi, \partial\psi(1)]^2 \times [\xi, 1]$, by Remark 1. Therefore,

$$\sup_{t\in\mathcal{T}} \left|v_{SI}(n^{-1}X, \theta) - v_{SI}(x, \vartheta)\right| \le c_3 \sup_{t\in\mathcal{T}} \left\|(n^{-1}X, \theta) - (x, \vartheta)\right\|,$$

for some Lipschitz constant $c_3 > 0$. Since $(x, \vartheta)$ is the solution of Equation (3.8) with initial condition $x(0) = \alpha$ and $\vartheta(0) = 1$, by virtue of Theorem 1 we obtain $\sup_{t\in\mathcal{T}} |\Delta_2(t)| \xrightarrow{P} 0$.

**Final conclusion.** Since $\sup_{t\in\mathcal{T}} |\Delta_1(t)| \xrightarrow{P} 0$ and $\sup_{t\in\mathcal{T}} |\Delta_2(t)| \xrightarrow{P} 0$, we conclude $\sup_{t\in\mathcal{T}} |\Delta(t)| \xrightarrow{P} 0$, which is a sufficient condition for

$$\langle M_{SI}\rangle(t) \xrightarrow{P} V_{SI}(t) = \int_0^t v_{SI}(x(s), \vartheta(s))\, ds. \qquad \square$$

We remark that the various moment estimates used in the proof above (and elsewhere in the paper) ignore the contributions of self-loops and parallel edges. Their contributions are asymptotically negligible as discussed earlier in Section 2.2. Also, note that we may need to add a parity edge to the last vertex if the sum of the drawn degrees is not even. This is relevant for the calculations in Steps I and II. However, as discussed earlier in Section 2.2, the contribution due to this minor adjustment is negligible and hence is not shown explicitly.

### 4.2. Asymptotic rarefaction of jumps

Recall that $M^\epsilon := \left(M_S^\epsilon, M_{SI}^\epsilon, M_{SS}^\epsilon\right)$ is the vector of square-integrable martingales containing all jumps of components of $M$ larger than $\epsilon$ in absolute value, for $\epsilon > 0$; i.e., $M_{id}(t) - M_{id}^\epsilon(t)$ is a locally square-integrable martingale and $|\delta M_{id}(t) - \delta M_{id}^\epsilon(t)| \le \epsilon$ for all $id \in \{S, SI, SS\}$ and $t \in \mathcal{T}$. We wish to show $\left\langle M_{id}^\epsilon\right\rangle(t) \xrightarrow{P} 0$ for all $id \in \{S, SI, SS\}$ and $t \in \mathcal{T}$, as $n \to \infty$. We would like to point out that this condition is essentially the *strong asymptotic rarefaction of jumps condition of the second type* (strong ARJ(2)) as described in [44, 1]. Intuitively this ensures that the sample paths of the martingale $M(t)$ are close to continuous in the limit. Before proceeding further, we offer the following remark.

**Lemma 3.** *For the CM graph $\mathcal{G}(\psi, n)$ along with A3, the following holds true:*

$$n^{-\frac{1}{2}}\, d_{\max} \xrightarrow{a.s.} 0, \tag{4.14}$$

*where $d_{\max}$ is the maximum degree observed in a realization of $\mathcal{G}(\psi, n)$.*

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

*Proof of Lemma 3.* The result follows by a direct application of the result in [8, Theorem 5.2] along with A3. □

Let us now compute the predictable quadratic variation of $M^\epsilon$ and establish its asymptotic limit.

**Lemma 4.** *Consider the stochastic SI model described in Section 2.2. Assume A1, A2, and A3 for a CM graph $\mathcal{G}(\psi, n)$. Consider the vector $M^\epsilon$ of square-integrable martingales containing all jumps of components of $M(t)$ larger than $\epsilon$ in absolute value for $\epsilon > 0$, as defined in Equation (4.2). Then, as $n \to \infty$, for all $id \in \{\mathrm{S}, \mathrm{SI}, \mathrm{SS}\}$, for each $t \in \mathcal{T}$,*

$$\langle M^\epsilon_{id} \rangle (t) \xrightarrow{\mathrm{P}} 0. \tag{4.15}$$

*Proof of Lemma 4.* We proceed in the following two steps.

**Computation of $\langle M^\epsilon_{\mathrm{S}} \rangle$.** Note that the original process $M'_{\mathrm{S}}$ makes only unit jumps. Then, for arbitrary $\epsilon > 0$,

$$\langle M^\epsilon_{\mathrm{S}} \rangle (t) \leq \int_0^t \mathbb{E}\Big[ \big(\delta M^\epsilon_{\mathrm{S}}(s)\big)^2 \mathbf{1}\big(|\delta M'_{\mathrm{S}}(s)| > n^{1/2}\epsilon\big) \big| \mathcal{F}_{s-} \Big] \, \mathrm{d}s = 0 \quad \forall\, n > \frac{1}{\epsilon^2}$$

$$\implies \langle M^\epsilon_{\mathrm{S}} \rangle (t) \xrightarrow{\mathrm{P}} 0 \quad \text{for all } 0 < t \leq T \text{ and for all } \epsilon > 0 \text{ as } n \to \infty.$$

**Computation of $\langle M^\epsilon_{\mathrm{SI}} \rangle$ and $\langle M^\epsilon_{\mathrm{SS}} \rangle$.** Note that both $M'_{\mathrm{SI}}$ and $M'_{\mathrm{SS}}$ jump only if infection of a vertex occurs. This in particular implies that the jump sizes of $M'_{\mathrm{SI}}$ and $M'_{\mathrm{SS}}$ are bounded above by the degree of the vertex getting infected. Therefore, they are also bounded above by the maximum degree $d_{\max}$. For an arbitrary $\epsilon > 0$, and for $id \in \{\mathrm{SI}, \mathrm{SS}\}$,

$$\begin{aligned} \langle M^\epsilon_{id} \rangle (t) &\leq \int_0^t \mathbb{E}\Big[ (\delta M^\epsilon_{id}(s))^2 \mathbf{1}\Big( |n^{-1/2} d_{\max}| > \epsilon \Big) \mid \mathcal{F}_{s-} \Big] \, \mathrm{d}s \\ &\leq t n^{-1} d^2_{\max} \mathbf{1}\Big( |n^{-1/2} d_{\max}| > \epsilon \Big). \end{aligned}$$

By Lemma 3, along with the continuous mapping theorem and the fact that almost sure convergence implies convergence in probability, we can claim that the right-hand side of the above inequality $t n^{-1} d^2_{\max} \mathbf{1}\big(|n^{-1/2} d_{\max}| > \epsilon\big) \xrightarrow{\mathrm{P}} 0$ for each $0 < t \leq T$ and $\epsilon > 0$. Therefore, for all $h > 0$,

$$\mathbb{P}\big( \langle M^\epsilon_{id} \rangle (t) > h \big) \leq \mathbb{P}\Big( t n^{-1} d^2_{\max} \mathbf{1}\Big( |n^{-1/2} d_{\max}| > \epsilon \Big) > h \Big) \to 0$$

as $n \to \infty$, establishing $\langle M^\epsilon_{id} \rangle (t) \xrightarrow{\mathrm{P}} 0$ as $n \to \infty$ for all $0 < t \leq T$ and $\epsilon > 0$. This completes the proof. □

### 4.3. Statement and proof of the FCLT

Having shown the convergence of all relevant quadratic variation processes, we are now ready to present the FCLT. First we state that the function $V$ found in Lemma 2 is a positive semidefinite matrix-valued function on $\mathcal{T}$, with positive semidefinite increments. Set $V(0) := \mathbf{0}$, the $3 \times 3$ null matrix, so that we can treat $V(t)$ as a positive semidefinite matrix-valued function on the entirety of $\mathcal{T}_0$. Let us denote by $\mathcal{V}$ the collection of all such positive

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

semidefinite $3 \times 3$ matrix-valued functions on $\mathcal{T}_0$ that have positive semidefinite increments and that are $\mathbf{0}$ at time zero. Given such a matrix-valued function $V \in \mathcal{V}$, let $G$ be a continuous Gaussian vector martingale such that $\langle G \rangle = [G] = V$. Such a process always exists [1, Chapter II, p. 83]. In particular, $G(t) - G(s) \sim N(\mathbf{0}, V(t) - V(s))$, the multivariate normal distribution for $0 \leq s \leq t$.

**Theorem 2.** (Functional central limit theorem.) *Consider the stochastic SI model described in Section 2.2. Assume A1, A2, and A3 for a CM graph $\mathcal{G}(\psi, n)$. Consider, for $t \in \mathcal{T}_0$, the fluctuation process*

$$Y(t) := \sqrt{n}\big(n^{-1}X(t) - x(t)\big). \tag{4.16}$$

*Assume $\lim_{n\to\infty} Y(0) = U(0)$, for some non-random $U(0)$. Then there exists a matrix-valued function $V \in \mathcal{V}$ on $\mathcal{T}_0$ such that*

$$Y \overset{\mathcal{D}}{\Longrightarrow} U \text{ in } D^{(3)} \text{as } n \to \infty, \tag{4.17}$$

*where $U$ is a continuous Gaussian vector semimartingale satisfying*

$$U(t) = U(0) + G(t) + \int_0^t \nabla \mathbb{H}_{\mathrm{x}}(x(s), \vartheta(s))\, U(s)\, \mathrm{d}s, \tag{4.18}$$

*where $\nabla \mathbb{H}_{\mathrm{x}}(x, \vartheta) := ((\partial_j \mathbb{H}_{\mathrm{i}}(x, \vartheta)))$ for $i, j \in \{\mathrm{S}, \mathrm{SI}, \mathrm{SS}\}$ and $G$ is a continuous Gaussian vector martingale such that $\langle G \rangle = [G] = V$, provided $V$ remains finite on the entirety of $\mathcal{T}_0$ and $\nabla \mathbb{H}_{\mathrm{x}}(x(s), \vartheta(s))$ is continuous.*

*Proof of Theorem 2.* We first prove an FCLT for the martingale process $M$ defined in Equation (4.1). We wish to apply Rebolledo's FCLT for local martingales on $M$. A version of the Rebolledo theorem adequate for our purpose is provided in Appendix C. Note that, in light of the Doob–Meyer decomposition given in Equation (2.2), $M$ is indeed a pure-jump, zero-mean, locally square-integrable, càdlàg martingale. After having established an FCLT for the martingale process $M$, we prove convergence of the fluctuation process $Y$. It suffices to carry out the following three steps.

**Step I: Deterministic limit of $\langle M \rangle$.** Let $(x, \vartheta)$ be the solution of Equation (3.8) with initial condition $x(0) = \alpha$ and $\vartheta(0) = 1$, as given in Theorem 1. Then, by virtue of Lemma 2, we conclude, for each $t \in \mathcal{T}$, $\langle M \rangle (t) \xrightarrow{\mathrm{P}} V(t)$, where the matrix-valued function $V$ is defined in Equation (4.5), and we set $V(0) := \mathbf{0}$, the $3 \times 3$ null matrix.

**Step II: Asymptotic rarefaction of jumps.** Let $\epsilon > 0$ be arbitrary. Consider the vector $M^\epsilon$ of square-integrable martingales containing all jumps of components of $M(t)$ larger than $\epsilon$ in absolute value for $\epsilon > 0$, as defined in Equation (4.2). Then, by means of Lemma 4, we conclude $\left\langle M^\epsilon_{id} \right\rangle (t) \xrightarrow{\mathrm{P}} 0$, for each $t \in \mathcal{T}$ and $id \in \{\mathrm{S}, \mathrm{SI}, \mathrm{SS}\}$.

Now let $G$ be the continuous Gaussian vector martingale such that $\langle G \rangle = [G] = V$. In light of Rebolledo's theorem for locally square-integrable martingales (see Appendix C and also [1, Chapter II, p. 83]), Step I and Step II are sufficient to establish

$$(M(t_1), M(t_2), \ldots, M(t_l)) \overset{\mathcal{D}}{\Longrightarrow} (G(t_1), G(t_2), \ldots, G(t_l)) \quad \text{as} n \to \infty$$

for all $t_1, t_2, \ldots, t_l \in \mathcal{T}$. Furthermore, since $\mathcal{T}$ is dense in $\mathcal{T}_0$, we conclude $M \overset{\mathcal{D}}{\Longrightarrow} G$ in $D^{(3)}$ as $n \to \infty$, and $\langle M \rangle$ and $[M]$ converge uniformly on compact subsets of $\mathcal{T}_0$, in probability, to $V$.

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

**Step III: Convergence of the fluctuation process.** In keeping with the Doob–Meyer decomposition given in Equation (2.2),

$$Y(t) = Y(0) + M(t) + \int_0^t \sqrt{n}\left(\frac{1}{n}\mathbb{F}_{\mathrm{X}}\big(\tilde{X}(s)\big) - \mathbb{H}_{\mathrm{x}}(x(s), \vartheta(s))\right) \mathrm{d}s,$$

we expect the following limit process:

$$U(t) = U(0) + G(t) + \int_0^t \nabla\mathbb{H}_{\mathrm{x}}(x(s), \vartheta(s))\, U(s)\, \mathrm{d}s. \tag{4.19}$$

Indeed, define

$$\begin{aligned}\Delta(t) :=& \int_0^t \sqrt{n}\left(\frac{1}{n}\mathbb{F}_{\mathrm{X}}\big(\tilde{X}(s)\big) - \mathbb{H}_{\mathrm{x}}(x(s), \vartheta(s)) - \frac{1}{\sqrt{n}}\nabla\mathbb{H}_{\mathrm{x}}(x(s), \vartheta(s))\ Y(s)\right) \mathrm{d}s \\ =& \int_0^t \sqrt{n}\left(\frac{1}{n}\mathbb{F}_{\mathrm{X}}\big(\tilde{X}(s)\big) - \mathbb{H}_{\mathrm{x}}\left(\frac{1}{n}X(s), \theta(s)\right) + \mathbb{H}_{\mathrm{x}}\left(\frac{1}{n}X(s), \theta(s)\right)\right. \\ & \left. - \mathbb{H}_{\mathrm{x}}(x(s), \vartheta(s)) - \frac{1}{\sqrt{n}}\nabla\mathbb{H}_{\mathrm{x}}(x(s), \vartheta(s))\ Y(s)\right) \mathrm{d}s.\end{aligned}$$

Note that the strong law of large numbers in Theorem 1 establishes uniform convergence (in probability) of the operators $n^{-1}\mathbb{F}_{\mathrm{X}}\big(\tilde{X}(s)\big)$ and $\mathbb{H}_{\mathrm{x}}\left(\frac{1}{n}X(s), \theta(s)\right)$, and the latter operator is Lipschitz continuous on its domain (see [30]). In light of Theorem 1 and A3, it follows from the Lipschitz continuity of various multinomial compensators $C_m^k$ introduced in the proof of Lemma 2 that

$$\lim_{n\to\infty} \sqrt{n}\left(\frac{1}{n}\mathbb{F}_{\mathrm{X}}\big(\tilde{X}(s)\big) - \mathbb{H}_{\mathrm{x}}\left(\frac{1}{n}X(s), \theta(s)\right)\right) = 0.$$

Moreover, we have just shown $M \overset{\mathcal{D}}{\Longrightarrow} G$ in $D^{(3)}$. If $V$ remains finite on the entirety of $\mathcal{T}_0$, the matrix-valued function $\nabla\mathbb{H}_{\mathrm{x}}(x(s), \vartheta(s))$ is continuous, and $\lim_{n\to\infty} Y(0) = U(0)$, for some non-random $U(0)$, then we have $\sup_{t\in\mathcal{T}_0} |\Delta(t)| \xrightarrow{\mathrm{P}} 0$ following Theorem 1, and by an application of the continuous mapping theorem we conclude

$$Y \overset{\mathcal{D}}{\Longrightarrow} U \text{ in } D^{(3)} \text{ as } n \to \infty,$$

where the Gaussian semimartingale $U$ satisfies Equation (4.19) with the Gaussian martingale $G$ being such that $\langle G\rangle = [G] = V$. This completes the proof. □

## 5. Applications

Here we consider some applications of our results. As we discuss these applications, we shall also present some numerical and simulation results that are intended not only to provide insights into the dynamics of the process, but also to serve as a verification of our results.

### 5.1. Percolation

There is a connection between the stochastic SI model and the percolation theory known from statistical physics and developed to study the process of liquid filtering ('percolating') through a porous medium. Classical equilibrium mechanics studies its stationary behaviour

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

and is premised upon the axiom that the underlying quantum-mechanical laws are designed so as to maximize the entropy. The stationary distribution of such a stochastic system is given by the Boltzmann ensemble. This classical treatment of the subject, however, does not explain the non-equilibrium behaviour of the dynamical system, i.e., when it is still in a transient phase. Consequently, the non-equilibrium behaviour of percolation has aroused much interest in recent times. Some notable contributions include [29, 5]. The standard treatment of percolation, both equilibrium and non-equilibrium, has been extended in another important direction concerning the structure of the porous medium. Traditionally it has been studied on lattices and grids. Of late, however, percolation on random graphs has also been considered ([9, 19, 46]). Continuing in this direction, we shall treat (non-equilibrium) percolation as a dynamical process on a CM random graph and study its behaviour over a finite time interval.

One of the key quantities of interest in the study of non-equilibrium percolation is the time evolution of the number of wetted sites (also called 'active' vertices in the literature). The correspondence between our stochastic SI model as described in Section 2.2 and non-equilibrium percolation is visible if we treat the infected vertices as the ones wetted during the process of percolation. Accordingly, in this context, we give the process $X(t)$ an appropriate new interpretation. The process $X_{\mathrm{S}}(t)$, for example, captures the number of unwetted sites until time $t$, and the process $X_{\mathrm{SI}}$ captures the number of channels (bonds) through which the liquid can percolate. In Figure 1, the percolated component up to a given time (the wetted part of the graph) is shown in red. Having made the correspondence precise, we can apply Theorem 2 to approximate these quantities in the large graph limit.

**Numerical illustration.** In Figure 3, we show some simulation results to check the accuracy of our scaling limit. We compare the expected sample paths of $X_{\mathrm{S}}$ and $X_{\mathrm{SI}}$ provided by Theorems 1 and 2 with estimates obtained using simulations of Gillespie's algorithm on a CM graph. In particular, we consider a Poisson degree distribution in Figure 3a and a 3-regular random graph in Figure 3b (obtained by the CM construction with degree distribution $p_k = \mathsf{1}(k = 3)$). In Figure 4, we compare the simulated sample paths of the true Gillespie dynamics and those corresponding to the diffusion approximation for a negative binomial degree distribution. Figures 3a, 3b, and 4 show convincing accuracy of the diffusion approximation. In Figure 5, we show the time evolution of the correlation coefficient between the jumps of $X_{\mathrm{S}}$ and $X_{\mathrm{SI}}$, and also the expected sample path coupled with 95%-confidence ellipses in the space of $X_{\mathrm{S}}$ and $X_{\mathrm{SI}}$. The orientation of the confidence ellipses is calculated as the angle of the eigenvector corresponding to the largest eigenvalue of the covariance matrix towards the $x$-axis. To be specific, the orientation is given by $\omega := \arctan e_2/e_1$, where $e := (e_1, e_2)$ is the eigenvector corresponding to the largest eigenvalue of the covariance matrix. The lengths of the major and minor axes are determined using the eigenvalues and by looking up the probability table of the chi-squared distribution (recall that squared normal variates follow a chi-squared distribution). The MATLAB script used to draw the ellipses is based on a script provided by [45].

The existence of a giant component and the proportion of vertices on the giant component play an important role in percolation theory, especially from an equilibrium point of view in statistical mechanics. The case of a degree distribution $\{p_k\}_{k\in\mathbb{N}_0}$ such that $\sum_{k\in\mathbb{N}_0} k^2 p_k = 2\sum_{k\in\mathbb{N}_0} k p_k$ and $p_1 = 0$ (or, equivalently, $p_0 + p_2 = 1$) is a curious one, in that quite different behaviours of the giant component are observable for such a degree distribution. (The reader should refer to [32] for examples of such behaviours.) Barring this exceptional case, in light of A3, the condition for existence of a giant component is satisfied (see [39]) for our stochastic SI

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

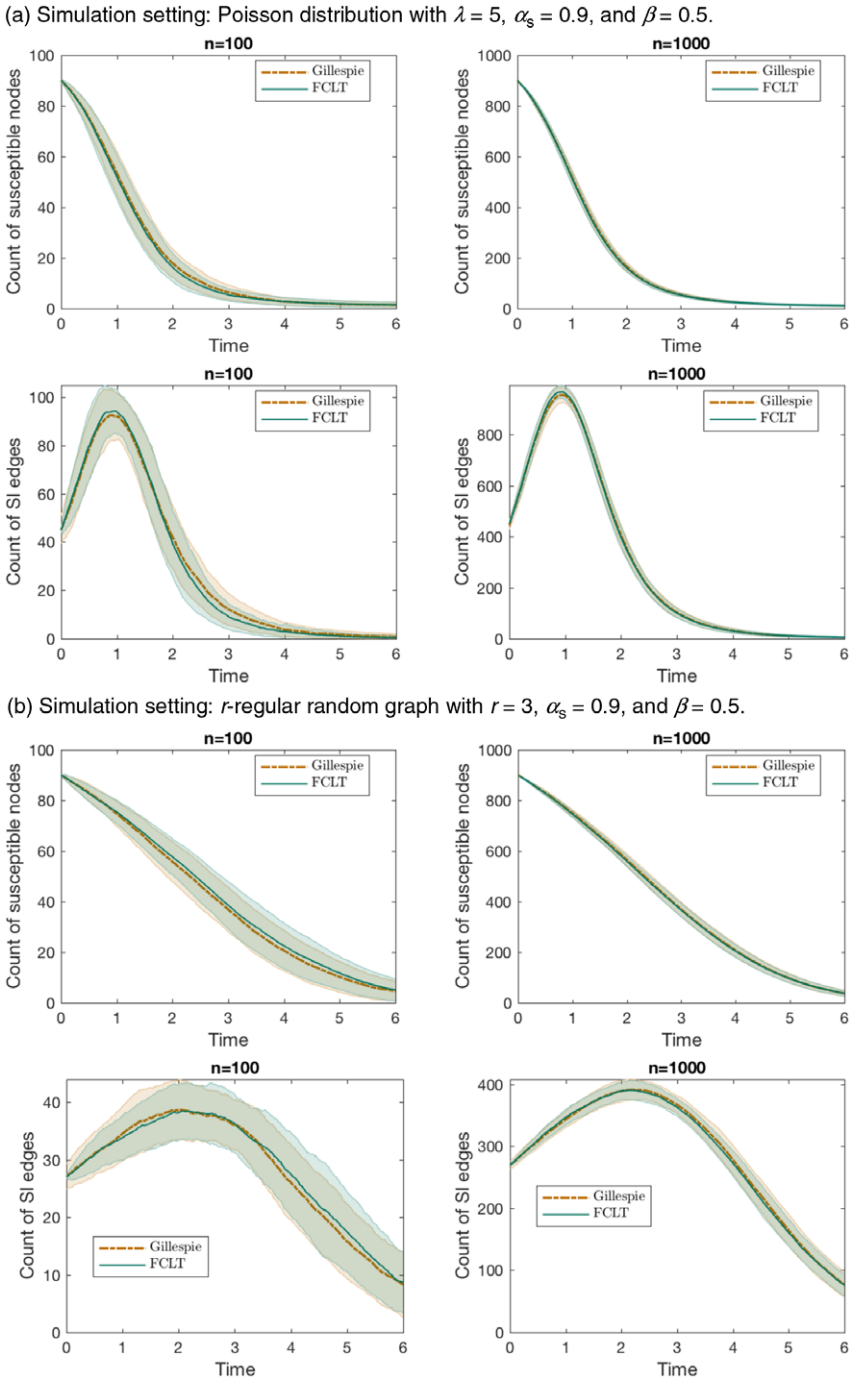



FIGURE 3. Comparison of our diffusion approximation with simulation results obtained by Gillespie's algorithm.

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

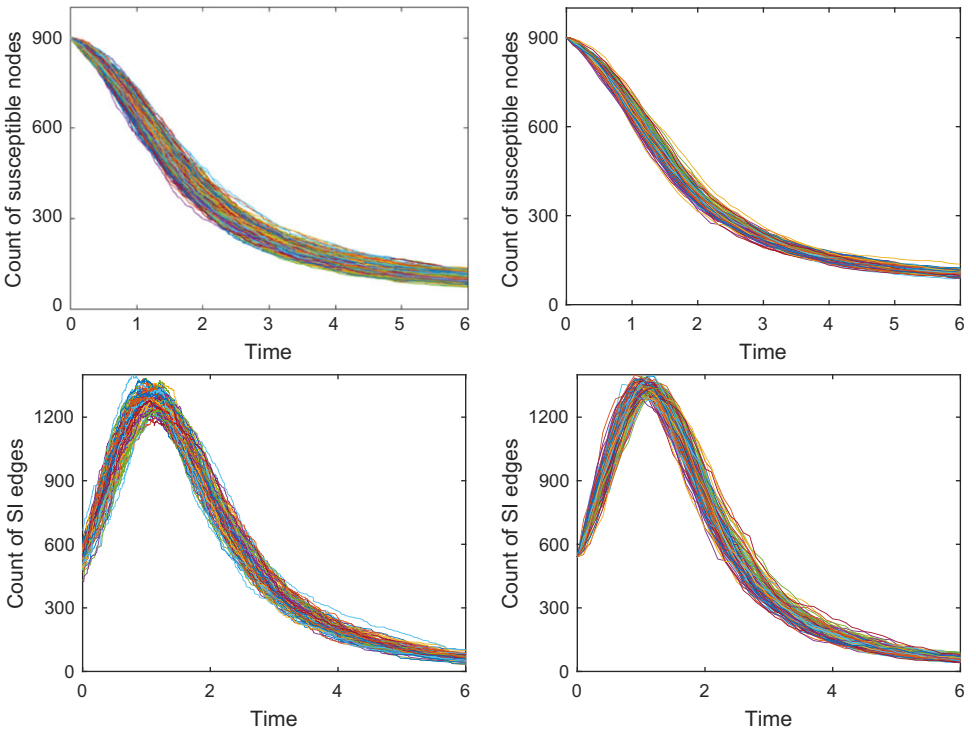



FIGURE 4. Comparison of simulated sample paths. *Left:* Sample paths obtained through Gillespie's algorithm. *Right:* Sample paths obtained through diffusion approximation. Simulation setting: $n = 1000$, negative binomial distribution with $r = 2$, $p = 3/4$.

model in the traditional sense. To be precise, setting $\alpha_S = 1$ and taking the asymptotic limit in time, one finds the fraction of vertices on the giant component to be $1 - \psi(\theta_\infty)$, where $\theta_\infty > 0$ is the solution of $\partial\psi(1)\theta_\infty = \partial\psi(\theta_\infty)$ (see [31, 40]). However, as mentioned earlier, we take a non-equilibrium point of view and concern ourselves with the time evolution of the fraction of vertices on the infected part of the graph, the 'percolated component'. As a by-product of the scaling limits in Theorem 1 and Theorem 2, the variable $\theta$ defined in Equation (2.5) gives us a tool to approximate the proportion of susceptible individuals in the population (and hence the proportion of infected vertices as well). We expect the fraction of infected individuals to converge in probability to $1 - \alpha_S\psi(\vartheta)$ as $n \to \infty$, $\vartheta$ being the scaling limit of $\theta$. A fixed time interval $\mathcal{T}_0$ enables us to look for critical values in the space of the infection rate $\beta > 0$. This allows us to decide whether the system 'percolates' in the sense that the fraction of vertices on the percolated component achieves a value greater than a pre-specified one (usually close to unity) by time $T$. Using different colours in Figure 6, we depict the fraction of vertices on the percolated component as a function of both time and the infection rate $\beta$ (let us call such a figure a percolation profile) for three degree distributions with the same mean. Questions such as whether the system with an infection rate $\beta$ 'percolates' are immediately settled by drawing a horizontal line and checking whether the line passes through the colour corresponding to the pre-specified level. See Figure 7 for another comparative view, highlighting the need to take into account higher moments of the degree distribution.

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

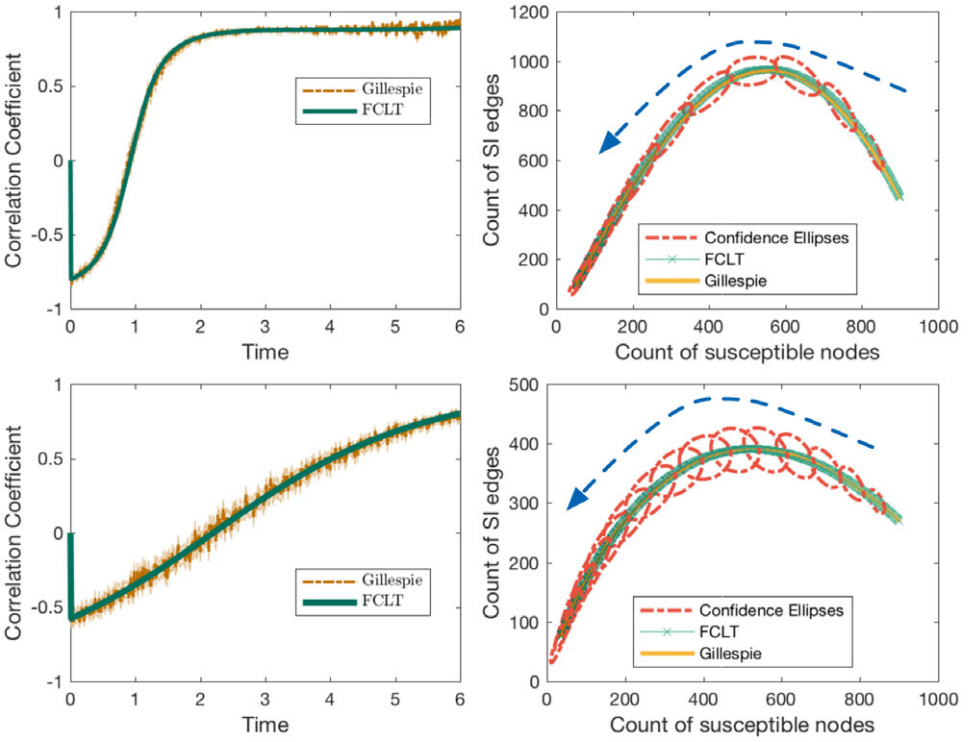



FIGURE 5. The figures on the left depict the time evolution of the correlation coefficient between jumps of $X_S$ and $X_{SI}$ as estimated from numerical simulations (via Gillespie's algorithm), pitted against theoretical values computed from the FCLT (Theorem 2). The figures on the right show the expected sample path in the space of $X_S$ and $X_{SI}$. The two lines correspond to numerical simulation and theoretical values. The dotted ellipses are the 95%-confidence ellipses based on estimates of covariances between $X_S$ and $X_{SI}$ from the diffusion approximation created using a MATLAB script provided by [45]. The arrows indicate the time direction. *Above:* Poisson distribution with mean 5. *Below:* $r$-regular random graph with $r = 3$. In both cases, $n = 1000$, $\alpha_S = 0.9$, and $\beta = 0.5$.

### 5.2. The strange case of Poisson-type degree distributions and the exactness of the pair approximation

We consider a particular class of degree distributions called 'Poisson-type' (PT) by [30]. A degree distribution with probability generating function $\psi$ is called PT if $\kappa(\vartheta)$, defined in Equation (3.4), is a constant, i.e., $\kappa(\vartheta) = \kappa$ for some $\kappa \in \mathbb{R}$ (or, equivalently, $\partial\psi(\vartheta) = \partial\psi(1)(\psi(\vartheta))^{\kappa}$). As a consequence, the operators defined in Equation (3.3) are also constants, and satisfy

$$\mathbb{D}^r \psi \circ \vartheta = \prod_{i=1}^{r-1} (i\kappa - i + 1) = ((r-1)\kappa - r + 2)\ \mathbb{D}^{r-1} \psi \circ \vartheta,$$

with $\mathbb{D}^0 \psi \circ \vartheta = 1$. The PT class includes Poisson ($\kappa = 1$, irrespective of the mean of the distribution), degenerate $\left(r\text{-regular random graphs, } \kappa = \frac{r-1}{r} < 1\right)$, binomial $\left(\kappa = \frac{N-1}{N} < 1,\right.$ independent of $p$ for $\left.\text{Bin}(N, p)\right)$, and negative binomial $\left(\kappa = \frac{r+1}{r} > 1,\right.$ independent of $p$

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

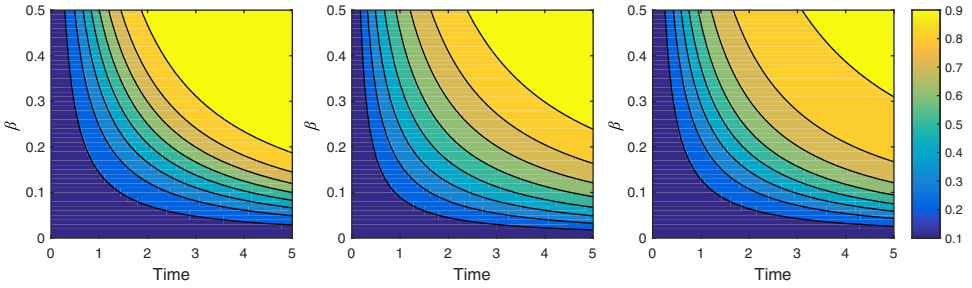



FIGURE 6. Comparison of percolation profiles of three degree distributions having the same mean. *Left:* Poisson distribution with mean 6. *Middle:* A heterogeneous population with degree distribution $p_k := 0.7 \times \mathbf{1}(k=1) + 0.2 \times \mathbf{1}(k=4) + 0.1 \times \mathbf{1}(k=45)$. Such a degree distribution represents a population segregated into three classes. Weak vertices constitute the biggest class, followed by medium-strength vertices and then strong vertices. *Right:* Negative binomial distribution with parameters $r=2$, $p=3/4$. The figures show the time evolution of the fraction of vertices on the percolated component for varying infection rates $\beta$. We assume the initial fraction of infected vertices is 0.1 in all three cases. The yellow region in each of the plots corresponds to the terminal state. Questions such as whether the system with an infection rate $\beta$ 'percolates' are immediately settled by drawing a horizontal line and checking whether the line passes through the colour corresponding to a pre-specified level. It is worth noting that the limiting proportions of vertices in the giant components eventually getting infected are also quite different for the three degree distributions. While the proportion is around 0.91 (approximately) for the negative binomial distribution and the hand-picked degree distribution in the middle, it is relatively high for the Poisson distribution (approximately 0.99).

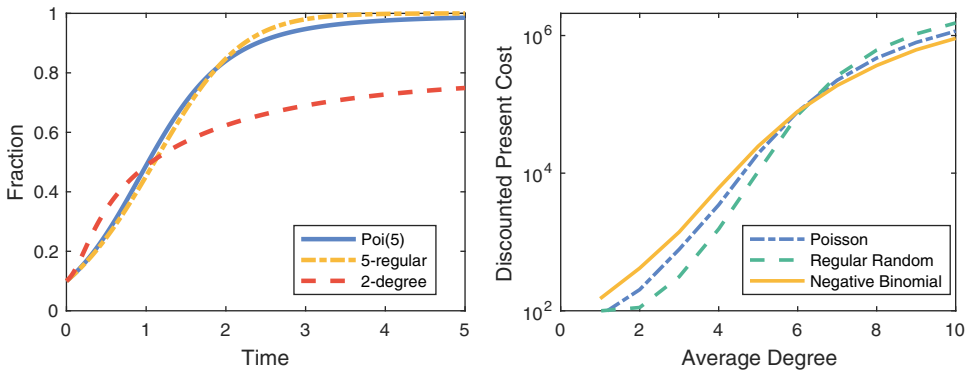



FIGURE 7. *Left:* Comparison of the time evolution of the fraction of vertices on the percolated component for different degree distributions with the same mean. The 2-degree distribution in the plot refers to the degree distribution $p_k = 0.5 \times \mathbf{1}(k=1) + 0.5 \times \mathbf{1}(k=9)$, where none of the vertices have degree 5 yet the average degree is 5. This presents a pathological case and highlights the need to take into account higher moments of the degree distribution. *Right:* Comparison of discounted cost against increasing average degree for different degree distributions. With increasing average degree the graphs lose sparsity and facilitate the spread of a computer virus. Therefore, they incur higher cost. When the average degree is very small, regular random graphs seem favourable compared to random graphs with negative-binomially distributed degrees. The costs are computed with $n = 1000$, $\gamma = 1$.

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

for $\mathrm{NB}(r, p)\big)$ degree distributions. The PT class is particularly peculiar in that it totally decouples the vector $x = (x_{\mathrm{S}}, x_{\mathrm{SI}}, x_{\mathrm{SS}})$ and the matrix-valued function $V$ from the auxiliary variable $\vartheta$, so that an autonomous system of ODEs can be obtained for $x$ and $V$, rendering $\vartheta$ redundant. This allows for great simplification in the limiting equations. Define $\mathbb{G}(x) := (\mathbb{G}_{\mathrm{S}}(x)\,, \mathbb{G}_{\mathrm{SI}}(x)\,, \mathbb{G}_{\mathrm{SS}}(x))$ as

$$\begin{aligned} \mathbb{G}_{\mathrm{S}}(x) &:= -\beta x_{\mathrm{SI}}, \\ \mathbb{G}_{\mathrm{SI}}(x) &:= \beta\kappa \frac{x_{\mathrm{SI}}}{x_{\mathrm{S}}}(x_{\mathrm{SS}} - x_{\mathrm{SI}}) - \beta x_{\mathrm{SI}}, \\ \mathbb{G}_{\mathrm{SS}}(x) &:= -2\beta\kappa \frac{x_{\mathrm{SI}} x_{\mathrm{SS}}}{x_{\mathrm{S}}}. \end{aligned} \tag{5.1}$$

Plugging $\mathbb{D}^2\psi \circ \vartheta = \kappa$ and $\mathbb{D}^3\psi \circ \vartheta = \kappa(2\kappa - 1)$ into Equation (4.4), we find that the matrix-valued function $V$ is entirely determined by $x$. The following is immediate.

**Corollary 1.** (Scaling limit under PT distributions.) *Assume* A1, A2, *and* A3 *for a CM graph* $\mathcal{G}(\psi, n)$ *with* $\partial\psi(\vartheta) = \partial\psi(1)(\psi(\vartheta))^{\kappa}$ *for some* $\kappa \in \mathbb{R}$. *Then the following law of large numbers holds:*

$$\sup_{0<t\leq T} \left\| n^{-1}X(t) - x(t) \right\| \xrightarrow{\mathrm{P}} 0,$$

*where* $x$ *is the solution of* $x(t) = x(0) + \int_0^t \mathbb{G}(x(s))\ \mathrm{d}s$ *with* $x(0) = \alpha$. *Moreover, the fluctuation process* $Y$ *defined in Equation* (4.16) *converges weakly to a continuous Gaussian vector semimartingale* $U$ *satisfying*

$$U(t) = U(0) + G(t) + \int_0^t \nabla\mathbb{G}(x(s))\ U(s)\,\mathrm{d}s,$$

*where* $G$ *is a Gaussian vector martingale such that* $\langle G\rangle = [G] = V$.

In fact, one can obtain a smaller system by expressing $x_{\mathrm{SI}}$ and $x_{\mathrm{SS}}$ explicitly as functions of $x_{\mathrm{S}}$ (see [30]). This is remarkable because, under the PT class, the graph structure affects the scaling limits *only* through two summary statistics of $\psi$, namely the mean $\partial\psi(1)$ and $\kappa = \mathbb{D}^2\psi(1)$. Recall that $\kappa$, as defined in Equation 3.4, is the limiting ratio of the average excess degree of a susceptible vertex chosen at random as a neighbour of an infected vertex, to the average degree of a susceptible vertex. Therefore it is, in general, dependent on time through $\vartheta$. Under the PT class, this ratio remains constant throughout the entire course of time $\mathcal{T}_0$. Moreover, the mean $\partial\psi(1)$ only affects the initial condition $x(0) = \alpha$ through Equation (3.2). The dynamics of the limiting process are then dictated by the constant $\kappa$ under the PT class.

Now we revisit the correlation equations approach of [43], from the ecology literature, to study the dynamics of counts of singles, pairs, triples, and quadruples of the form $A$, $AB$, $ABC$, $ABCD$, where $A, B, C, D \in \{\mathrm{S}, \mathrm{I}\}$. Following [43], we use the notation $[\cdot]$ to denote the count. In this mean-field approach, the dynamics of singles are described by those of pairs, the dynamics of pairs by those of triples, and so on. In this context, pair approximation refers to approximating the count of triples by pairs as

$$[ABC] \approx \kappa \frac{[AB]\,[BC]}{[B]},$$

and closing the system at the level of pairs (also known as pair closure). In order to draw an analogy, we divide the counts by $n$ and use the same notation for the scaled counts. We also

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

set the same initial condition ([S], [SI], [SS]) $= \alpha$ at $t = 0$. The pair approximation then yields a system of ODEs for ([S], [SI], [SS]) that exactly matches the limiting ODEs for $n^{-1}X$, i.e.,

$$\frac{\mathrm{d}}{\mathrm{d}t}([\mathrm{S}], [\mathrm{SI}], [\mathrm{SS}]) = \mathbb{G}\left(([\mathrm{S}], [\mathrm{SI}], [\mathrm{SS}])\right). \tag{5.2}$$

Therefore, under the PT class, the pair approximation is *exact* in the sense that it correctly estimates the limiting means of various counts. By virtue of Corollary 1, our FCLT further enables it to correctly estimate *all* other higher limiting moments, because $V$ is now entirely determined by the solution of Equation (5.2). As the PT class is quite big, our FCLT thus greatly enhances the usefulness of the pair approximation.

### 5.3. Spread of computer viruses

Epidemic models have been used in the context of the spread of computer viruses for some years now. The correspondence between our model and the application area under consideration is apparent without requiring much change in nomenclature. Early works in this direction did not take into account the inherent graph structure and assumed 'homogeneous mixing' in some sense. Recent works, however, have duly studied it on more realistic computer networks, which are often modelled as random graphs, without the assumption of 'homogeneous mixing'. Lelarge, for example, based much of his work on classical Erdős–Rényi random graphs and CMs (see, e.g., [35]). Interested readers are referred to [34, 35, 49] for an overview of relevant literature. Applying our results, we can approximate the number of virus-affected computers over time and the numbers of edges of different types. Additionally, one might be interested in estimating some 'cost' involving the count variables in a linear or nonlinear fashion. For instance, if the cost function is polynomial in the count variables, the mixed moments of various orders can be approximated by means of Theorem 2. To illustrate the concept using a simple example, we assume an exponentiated form for the incurred cost to emphasize the severity of a computer being virus-affected. We can then compute time-discounted expected incurred cost and study how it behaves with decreasing sparsity of the underlying graph. To this end, define

$$\begin{aligned} I(t) &:= \exp\bigl(cX_I(t)\bigr), \\ C_\psi &:= \mathbb{E}\left[\int_{\mathcal{T}_0} \exp\bigl(-\gamma t\bigr) I(t)\,\mathrm{d}t\right] = \int_{\mathcal{T}_0} \exp\bigl(-\gamma t\bigr)\mathbb{E}[I(t)]\,\mathrm{d}t, \end{aligned} \tag{5.3}$$

where $c > 0$ and $\gamma > 0$ are constants. In Figure 7, we plot the discounted cost $C_\psi$ against an increasing average degree of the underlying graph, engendering decreasing sparsity. When the average degree is very small, regular random graphs seem favourable compared to random graphs with negative-binomially distributed degrees.

## 6. Conclusion and future work

We conclude the paper with a brief literature review followed by a short discussion. In summary, we have studied the susceptible–infected (SI) model by formulating a stochastic process on CM random graphs. Even though this is a simple infection model, the Markovian process on the entirety of the random graph suffers state space explosion as $n$ grows to infinity. Analysis of the non-Markovian aggregate process also becomes complicated. Therefore, scaling limits retaining key features of the network are generally of interest.

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

### 6.1. Related works

In the recent scientific literature one comes across a host of dynamical processes arising from epidemiology ([18, 41, 27]), statistical physics ([17]), and computer science ([35, 34, 49]). These dynamical processes are often similar and hence lend themselves to application across disciplines ([42]). In pursuit of scaling limits, much of the research has been inspired by the mean-field approach from statistical physics. For instance, the authors of [42] study epidemic dynamics on the scale-free networks of [4]. The majority of work in this direction aims to obtain limiting ordinary differential equations (ODEs) for the proportions of individuals in different compartments of the population. Notwithstanding the simplicity of these methods, the scaling limits presented are approximate and lack mathematical rigour by design. See [23, Chapter 1] for a critique. The standard mean-field method was further improved by the use of pair approximation in [26]. Several other improvements yielding less approximate results have been proposed since. A detailed account is presented in [10]. Some of these approximate results have been followed up by probabilists and improved upon ([23, 24, 20]).

A related line of research concerns the first-passage percolation (FPP) on random graphs ([9, 19, 46, 13, 11, 12]). Given a graph $G = (V, E)$ with $n$ vertices, we assign random weights $W_e$ to edges $e \in E$. We assume the weights are independent and identically distributed. For two vertices $i, j \in V$, the *passage time* from $i$ to $j$ is defined as

$$\mathsf{c}_n(i,j) := \min_{\mathsf{p}(i,j)} \sum_{e \in \mathsf{p}(i,j)} W_e,$$

where the minimum is taken over all paths $\mathsf{p}(i, j)$ from the vertex $i$ to the vertex $j$. By convention, $\mathsf{c}_n(i, i) := 0$, and $\mathsf{c}_n(i, j) := \infty$ if there is no path from the vertex $i$ to the vertex $j$. For two typical vertices $i$ and $j$, chosen uniformly at random, one looks for a sequence of reals $\tilde{\mathsf{c}}_n$ such that

$$\left(\mathsf{c}_n - \tilde{\mathsf{c}}_n\right) \stackrel{\mathcal{D}}{\Longrightarrow} Q,$$

for some limiting random variable $Q$ with usually continuous support. Sometimes the scaling limit of the number of edges on the shortest path between two typical vertices is studied along with the passage times. That is, if $\mathsf{h}_n$ denotes the number of edges on the shortest path between two typical vertices (chosen uniformly at random), we seek a sequence $\tilde{\mathsf{h}}_n$ of scaling constants such that

$$\left(\mathsf{h}_n - \tilde{\mathsf{h}}_n\right) \stackrel{\mathcal{D}}{\Longrightarrow} \tilde{Q},$$

for some limiting random variable $\tilde{Q}$. The random variable $\mathsf{h}_n$ is often called the typical hop-count. One of the important questions in the study of FPP concerns the growth rate of the sequence of scaling constants $\tilde{\mathsf{c}}_n$ and $\tilde{\mathsf{h}}_n$, e.g., whether they are of order $\log(n)$. Universality results are also important in the study of FPPs and form a substantial body of literature. While the hop-count $\mathsf{h}_n$ measures typical distances in the random graph, the passage time $\mathsf{c}_n$ can be interpreted as the typical (minimum) amount of time required for an infectious disease to be transmitted from an infected vertex to a susceptible vertex. As a result, limit theorems for the passage time $\mathsf{c}_n$ provide a complementary view to our FCLT for the SI process on CM random graphs as demonstrated by the numerical results in Section 5.

In an epidemiological context, limit theorems for a discrete-time random graph epidemic model were derived in [2] under rather restrictive assumptions, such as finiteness of a $(4 + \delta)$th

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

moment of the degree distribution for some $\delta > 0$. The work of Volz in [48] presented scaling limits for susceptible–infected–recovered (SIR) models on random graphs in the form of ODEs. The authors in [21] later proved Volz's results rigorously by summarizing the epidemic process on CM random graphs into some measure-valued equations. Several similar law-of-large-numbers-type scaling limits under varying sets of technical assumptions surfaced afterwards. For example, uniformly bounded degrees were assumed in [15, 6]. The authors of [31] assume the degree of a randomly chosen susceptible vertex to be uniformly integrable and the maximum degree of initially infected vertices to be $o(n)$. The work in [3] studies a variant of the standard compartmental SIR with notions of local (within households, for example) and global contacts, and uses a branching process approximation to derive threshold behaviour and final outcome in the event of a global epidemic. Recently a law of large numbers for the stochastic SIR process on a multilayer CM was derived in [30] assuming finiteness of the second moment of the underlying degree distribution.

Although a number of scaling limits in the form of laws of large numbers have surfaced over the years, appropriate diffusion approximations have not yet been fully explored. Our FCLT attempts to complement the laws of large numbers already available in the literature. In particular, our FCLT lends itself as an approximating tool in applications where laws of large numbers are inadequate, e.g., in situations involving higher-order moments.

In our present work, we have disregarded 'recovery' of the infected vertices. The reason behind this exclusion is our inability to evaluate the neighbourhood distribution of an infected vertex in the presence of spontaneous recovery of its neighbours. One difficulty is that, unlike the susceptible vertices (of a given degree) that have remained untouched by the process of infection and hence receive identically distributed neighbourhoods upon uniformly-at-random matching of half-edges, the infected vertices are not identically distributed, because they already possess partially formed neighbourhoods consisting of infected and recovered neighbours. This corresponds to the part of the graph that has already been revealed up to a given time. Recall the construction of the CM random graph, where the graph is dynamically revealed as infection spreads (see Section 4). As a result, the hypergeometric argument as mentioned in Lemma 1 seems inadequate. For the purpose of obtaining a law of large numbers, we can circumvent this difficulty by suitably bounding the jump sizes of different martingales arising in the proof by the degrees of the vertices concerned. Therefore, we actually do not need the exact neighbourhood distribution of an infected individual for deriving laws of large numbers. However, to establish an FCLT, one needs to find the limit of the quadratic covariation process, which would involve the task of approximating quantities such as $\sum_{k\in\mathbb{N}_0}\sum_{i\in I_k} X_{\mathrm{IS},i}^2$, where $I_k$ is the collection of degree-$k$ vertices that are infected and $X_{\mathrm{IS},i}$ is the number of susceptible neighbours of an infected individual of degree $k$. We suspect that elaborate bookkeeping of the infection-spreading process would be necessary to approximate such quantities. We have not been able to find a simple workaround so far and intend to pursue this problem in the near future.

## Appendix A. Hypergeometric moments

Here, we compute various (conditional) moments that are useful for our derivations. Let us use the shorthand notation

$$\varsigma(n_{\mathrm{SI}}; X_{\mathrm{SI}}, X_{\mathrm{S}\bullet}, k) := \mathbb{P}\big(X_{\mathrm{SI},i} = n_{\mathrm{SI}}, X_{\mathrm{SS},i} = k - n_{\mathrm{SI}} \mid \mathcal{F}_t\big)\,,$$

conditional on the process history as given in Lemma 1. The following moments are then computed keeping Lemma 1 in mind.

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

In a straightforward fashion, we get for $i \in S_k$

$$\begin{aligned}\mathbb{E}\big[(X_{\mathrm{SI},i})_3 \mid \mathcal{F}_t\big] &= \sum_{n_{\mathrm{SI}}} (n_{\mathrm{SI}})_3 \frac{\binom{X_{\mathrm{SI}}}{n_{\mathrm{SI}}}\binom{X_{\mathrm{S}\bullet}-X_{\mathrm{SI}}}{k-n_{\mathrm{SI}}}}{\binom{X_{\mathrm{S}\bullet}}{k}} \\ &= \sum_{n_{\mathrm{SI}}} (n_{\mathrm{SI}})_3 \frac{\frac{(X_{\mathrm{SI}})_3}{(n_{\mathrm{SI}})_3}\binom{X_{\mathrm{SI}}-3}{n_{\mathrm{SI}}-3}\binom{(X_{\mathrm{S}\bullet}-3)-(X_{\mathrm{SI}}-3)}{(k-3)-(n_{\mathrm{SI}}-3)}}{\frac{(X_{\mathrm{S}\bullet})_3}{(k)_3}\binom{X_{\mathrm{S}\bullet}-3}{k-3}} \\ &= \frac{(k)_3(X_{\mathrm{SI}})_3}{(X_{\mathrm{S}\bullet})_3} \sum_{n_{\mathrm{SI}}} \varsigma(n_{\mathrm{SI}}; X_{\mathrm{SI}}-3, X_{\mathrm{S}\bullet}-3, k-3) \\ &= \frac{(k)_3(X_{\mathrm{SI}})_3}{(X_{\mathrm{S}\bullet})_3}.\end{aligned}$$

Similarly, we can derive for $i \in S_k$

$$\mathbb{E}\big[(X_{\mathrm{SI},i})_2 \mid \mathcal{F}_t\big] = \sum_{n_{\mathrm{SI}}} (n_{\mathrm{SI}})_2 \frac{\binom{X_{\mathrm{SI}}}{n_{\mathrm{SI}}}\binom{X_{\mathrm{S}\bullet}-X_{\mathrm{SI}}}{k-n_{\mathrm{SI}}}}{\binom{X_{\mathrm{S}\bullet}(t)}{k}} = \frac{(k)_2(X_{\mathrm{SI}})_2}{(X_{\mathrm{S}\bullet})_2},$$

whence we get

$$\begin{aligned}\mathbb{E}\Big[X_{\mathrm{SI},i}^3 \mid \mathcal{F}_t\Big] &= \mathbb{E}\big[(X_{\mathrm{SI},i})_3 \mid \mathcal{F}_t\big] + 3\mathbb{E}\big[(X_{\mathrm{SI},i})_2 \mid \mathcal{F}_t\big] + \mathbb{E}\big[X_{\mathrm{SI},i} \mid \mathcal{F}_t\big] \\ &= \frac{(k)_3(X_{\mathrm{SI}})_3}{(X_{\mathrm{S}\bullet})_3} + 3\frac{(k)_2(X_{\mathrm{SI}})_2}{(X_{\mathrm{S}\bullet})_2} + k\frac{X_{\mathrm{SI}}}{X_{\mathrm{S}\bullet}}.\end{aligned}$$

Proceeding in a similar fashion, for $i \in S_k$,

$$\begin{aligned}\mathbb{E}\big[X_{\mathrm{SI},i}(X_{\mathrm{SS},i})_2 \mid \mathcal{F}_t\big] &= \sum_{0\le n_{\mathrm{SI}}+n_{\mathrm{SS}}=k} n_{\mathrm{SI}}(n_{\mathrm{SS}})_2 \frac{\binom{X_{\mathrm{SI}}}{n_{\mathrm{SI}}}\binom{X_{\mathrm{S}\bullet}-X_{\mathrm{SI}}}{n_{\mathrm{SS}}}}{\binom{X_{\mathrm{S}\bullet}}{k}} \\ &= \sum_{0\le n_{\mathrm{SI}}+n_{\mathrm{SS}}=k} n_{\mathrm{SI}}(n_{\mathrm{SS}})_2 \frac{\frac{X_{\mathrm{SI}}}{n_{\mathrm{SI}}}\binom{X_{\mathrm{SI}}-1}{n_{\mathrm{SI}}-1}\frac{(X_{\mathrm{SS}})_2}{(n_{\mathrm{SS}})_2}\binom{X_{\mathrm{S}\bullet}-X_{\mathrm{SI}}-2}{k-n_{\mathrm{SI}}-2}}{\frac{(X_{\mathrm{S}\bullet})_3}{(k)_3}\binom{X_{\mathrm{S}\bullet}-3}{k-3}} \\ &= \frac{(k)_3 X_{\mathrm{SI}}(X_{\mathrm{SS}})_2}{(X_{\mathrm{S}\bullet})_3} \sum_{n_{\mathrm{SI}}} \varsigma(n_{\mathrm{SI}}, ; X_{\mathrm{SI}}-1, X_{\mathrm{S}\bullet}-3, k-3) \\ &= \frac{(k)_3 X_{\mathrm{SI}}(X_{\mathrm{SS}})_2}{(X_{\mathrm{S}\bullet})_3}.\end{aligned}$$

Similarly, for $i \in S_k$,

$$\begin{aligned}\mathbb{E}\big[(X_{\mathrm{SI},i})_2 X_{\mathrm{SS},i} \mid \mathcal{F}_t\big] &= \frac{(k)_3(X_{\mathrm{SI}})_2 X_{\mathrm{SS}}}{(X_{\mathrm{S}\bullet})_3}, \\ \mathbb{E}\big[X_{\mathrm{SI},i} X_{\mathrm{SS},i} \mid \mathcal{F}_t\big] &= \frac{(k)_2 X_{\mathrm{SI}} X_{\mathrm{SS}}}{(X_{\mathrm{S}\bullet})_2}.\end{aligned}$$

## Appendix B. Interpretation of the $\mathbb{D}$ operator

Here we provide an intuitive explanation for the $\mathbb{D}$ operator defined in Equation (3.3) in the context of SI processes on CM random graphs. Recall that $\mu_{\mathrm{S}}$ and $\mu_{\mathrm{S}}^{(r)}$ respectively denote the

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

average degree of a randomly chosen susceptible vertex and the average excess degree of a susceptible vertex randomly chosen as a neighbour of $r$ infected individuals. In Section 3 we mentioned that the operator $\mathbb{D}^{r+1}\psi \circ \vartheta$ recursively compared a susceptible vertex randomly chosen as a neighbour of $r$ infected individuals with a randomly chosen susceptible vertex. We now make this notion of comparison precise.

**Lemma 5.** *Assume* A1, A2, *and* A3 *for the stochastic SI model on the CM graph* $\mathcal{G}(\psi, n)$. *Then* $\mathbb{D}^r\psi \circ \theta \xrightarrow{\mathrm{P}} \mathbb{D}^r\psi \circ \vartheta$ *uniformly on* $\mathcal{T}_0$, *and the following recurrence relation for* $\mathbb{D}^r$ *holds:*

$$\mathbb{D}^{r+1}\psi \circ \theta = \frac{\mu_{\mathrm{S}}^{(r)}(\theta)}{\mu_{\mathrm{S}}(\theta)}\mathbb{D}^r\psi \circ \theta. \tag{B.1}$$

*Proof of Lemma 5.* The probability that a randomly chosen vertex $i$ is susceptible and is of degree $k$ is given by $\mathbb{P}(i \in S_k(t)) = n^{-1}X_{\mathrm{S}}(0)\theta^k(t)p_k$. The following is then immediate:

$$\mu_{\mathrm{S}}(\theta(t)) = \sum_k k\mathbb{P}(i \in S_k(t) \mid i \in S(t)) = \frac{\sum_k k\theta^k(t)p_k}{\sum_k \theta^k(t)p_k} = \frac{\theta(t)\partial\psi(\theta(t))}{\psi(\theta(t))}.$$

In order to explicitly calculate $\mu_{\mathrm{S}}^{(r)}$, it will be helpful to keep the dynamic construction of the graph in mind. In particular, we make use of the neighbourhood distribution of a susceptible vertex given in Lemma 1. Therefore,

$$\begin{aligned}\mu_{\mathrm{S}}^{(r)}(\theta(t)) &= \frac{\sum_k (k-r)\mathbb{P}(i \in S_k(t))\,\mathbb{E}\big[(X_{\mathrm{SI},i})_r \mid \mathcal{F}_{t-}\big]}{\sum_k \mathbb{P}(i \in S_k(t))\,\mathbb{E}\big[(X_{\mathrm{SI},i})_r \mid \mathcal{F}_{t-}\big]} \\ &= \frac{\sum_k (k)_{r+1}\theta^k(t)p_k}{\sum_k (k)_r\theta^k(t)p_k} \\ &= \frac{\theta(t)\,\partial^{r+1}\psi(\theta(t))}{\partial^r\psi(\theta(t))}.\end{aligned}$$

The recurrence relation then follows in a straightforward manner:

$$\mathbb{D}^{r+1}\psi \circ \theta = \frac{\theta\,\partial^{r+1}\psi(\theta)}{\partial^r\psi(\theta)} \times \frac{\psi(\theta)}{\theta\partial\psi(\theta)} \times \frac{\psi^{r-1}(\theta)\partial^r\psi(\theta)}{(\partial\psi(\theta))^r} = \frac{\mu_{\mathrm{S}}^{(r)}(\theta)}{\mu_{\mathrm{S}}(\theta)}\mathbb{D}^r\psi \circ \theta.$$

The convergence $\mathbb{D}^r\psi \circ \theta \xrightarrow{\mathrm{P}} \mathbb{D}^r\psi \circ \vartheta$, uniformly on $\mathcal{T}_0$, follows by virtue of Theorem 1. This completes the proof. □

For our purposes, we only need

$$\frac{\mu_{\mathrm{S}}^{(1)}(\theta)}{\mu_{\mathrm{S}}(\theta)} \xrightarrow{\mathrm{P}} \kappa(\vartheta) = \mathbb{D}^2\psi \circ \vartheta, \qquad \frac{\mu_{\mathrm{S}}^{(2)}(\theta)}{\mu_{\mathrm{S}}(\theta)}\kappa(\theta) \xrightarrow{\mathrm{P}} \mathbb{D}^3\psi \circ \vartheta,$$

and hence the interpretation in Section 3 as a limiting ratio follows. The two operators $\mathbb{D}^2\psi \circ \vartheta$ and $\mathbb{D}^3\psi \circ \vartheta$ essentially allow us to correctly estimate various pair and triple counts in the large graph limit.

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

## Appendix C. Rebolledo's theorem

Here we furnish a statement of Rebolledo's theorem, as our derivation of the FCLT relies heavily on it. The version presented here differs slightly from the original one in [44]. In our application, we only need a central limit theorem for locally square-integrable martingales. Therefore, we borrow the following version from [1].

Let $M_1, M_2, \ldots, M_n, \ldots$ be a sequence of vector-valued locally square-integrable martingales. We allow the possibility of their being defined on different sample spaces for each $n$. Let us denote the components of $M_n$ by $M_n := (M_{n,1}, M_{n,2}, \ldots, M_{n,k})$ for some $k \in \mathbb{N}$. Now, for each $\epsilon > 0$, define $M_n^{(\epsilon)}$ to be a vector-valued locally square-integrable martingale that contains all jumps of $M_n$ larger in absolute value than $\epsilon$ as its only jumps. Write $M_n^{(\epsilon)} := \left(M_{n,1}^{(\epsilon)}, M_{n,2}^{(\epsilon)}, \ldots, M_{n,k}^{(\epsilon)}\right)$. Then the absolute value of the difference between jump sizes of $M_{n,i}$ and $M_{n,i}^{(\epsilon)}$ is necessarily smaller than $\epsilon$.

The optional quadratic (co-)variation process $\left[M_{n,i}, M_{n,j}\right]$ of the processes $M_{n,i}$ and $M_{n,j}$ is defined as the sum of the products of the jump sizes of the processes. That is, for $i, j = 1, 2, \ldots, k$, we define

$$\left[M_{n,i}, M_{n,j}\right] := \sum_{s \le t} \delta M_{n,i}(s) \delta M_{n,j}(s).$$

By convention, we define $\left[M_{n,i}\right] = \left[M_{n,i}, M_{n,j}\right]$. For $i, j = 1, 2, \ldots, k$, the predictable quadratic (co-)variation process between $M_{n,i}$ and $M_{n,j}$ is given by

$$\left\langle M_{n,i}, M_{n,j} \right\rangle (t) := \int_0^t \mathrm{Cov}(\mathrm{d}M_{n,i}(s),\ \mathrm{d}M_{n,j}(s) \mid \mathcal{F}_{s-})\, \mathrm{d}s,$$

with the convention $\left\langle M_{n,i} \right\rangle = \left\langle M_{n,i}, M_{n,i} \right\rangle$. Here, $\mathrm{d}M_{n,i}(s)$ and $\mathrm{d}M_{n,j}(s)$ are the increments of $M_{n,i}$, and $M_{n,j}$ respectively. Denote the corresponding $k \times k$ matrix-valued quadratic variation processes by $\langle M_n \rangle := \left(\left( \left\langle M_{n,i}, M_{n,j} \right\rangle \right)\right)$. Similarly, for the process $M_n^{(\epsilon)}$, denote the corresponding predictable quadratic variation process by $\left\langle M_n^{(\epsilon)} \right\rangle := \left(\left( \left\langle M_{n,i}^{(\epsilon)}, M_{n,j}^{(\epsilon)} \right\rangle \right)\right)$. Similarly, let $[M_n]$ denote the matrix of optional quadratic variations of the components of $M_n$. That is, $[M_n] := \left(\left( \left[M_{n,i}, M_{n,j}\right] \right)\right)$. Finally, let $\mathcal{T}_0 := [0, T]$ be as before and let $\mathcal{T} \subset \mathcal{T}_0$.

**Theorem 3.** (Rebolledo's theorem for locally square-integrable martingales.) *Consider the following three conditions:*

$$\langle M_n \rangle (t) \xrightarrow{\mathrm{P}} V(t), \forall t \in \mathcal{T}, \tag{C.1}$$

$$[M_n] (t) \xrightarrow{\mathrm{P}} V(t), \forall t \in \mathcal{T}, \tag{C.2}$$

$$\left\langle M_{n,i}^{(\epsilon)} \right\rangle (t) \xrightarrow{\mathrm{P}} 0, \forall t \in \mathcal{T}, \epsilon > 0, i = 1, 2, \ldots, k, \tag{C.3}$$

*as $n \to \infty$, where $V$ is a $k \times k$ semidefinite matrix-valued, continuous deterministic function on $\mathcal{T}_0$ with positive semidefinite increments and $V(0) = \mathbf{0}$, the zero matrix.*

*Then either of Equation (C.1) or Equation (C.2), together with Equation (C.3), implies the following finite-dimensional convergence:*

$$(M_n(t_1), M_n(t_2), \ldots, M_n(t_l)) \xRightarrow{\mathcal{D}} (W(t_1), W(t_2), \ldots, W(t_l))$$

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

*as $n \to \infty$, where $W$ is a Gaussian vector martingale with $[W] = \langle W \rangle = V$ and $t_1, t_2, \ldots, t_l \in \mathcal{T}$. Moreover, both Equation (C.1) and Equation (C.2) then hold.*

*If, in addition, $\mathcal{T}$ is dense in $\mathcal{T}_0$, then either of Equation (C.1) or Equation (C.2) together with Equation (C.3) implies the following weak convergence:*

$$M_n \overset{\mathcal{D}}{\Longrightarrow} W$$

*as $n \to \infty$ in $D^{(k)}$, the space of $\mathbb{R}^k$-valued càdlàg functions on $\mathcal{T}_0$ endowed with the Skorokhod topology. Furthermore, $\langle M_n \rangle$ and $[M_n]$ converge uniformly on compact subsets of $\mathcal{T}_0$ to $V$ in probability.*

## Acknowledgements

First of all, the authors would like to thank the anonymous reviewers for their constructive feedback, which significantly improved the quality of the paper. Most of the work was conducted while the first author was a PhD student at the Technische Universität Darmstadt in Germany. It was initiated when the first author was visiting the Mathematical Biosciences Institute (MBI) at the Ohio State University in the summer of 2016. The first author acknowledges the hospitality of MBI during his visit to the institute.

## Funding information

The work was co-funded by the German Research Foundation (DFG) as part of project C3 within the Collaborative Research Center (CRC) 1053—MAKI, and by the National Science Foundation under the grants RAPID DMS-1513489 and DMS-1853587. MBI is receiving major funding from the National Science Foundation under the grant DMS-1440386. The first author was supported by the President's Postdoctoral Scholars Program of the Ohio State University.

## Competing interests

There were no competing interests to declare which arose during the preparation or publication process for this article.

## References

[1] Andersen, P. K., Borgan, Ø., Gill, R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer, New York.
[2] Andersson, H. (1998). Limit theorems for a random graph epidemic model. *Ann. Appl. Prob.* **8**, 1331–1349.
[3] Ball, F. and Neal, P. (2002). A general model for stochastic SIR epidemics with two levels of mixing. *Math. Biosci.* **180**, 73–102.
[4] Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science* **286**, 509–512.
[5] Barato, A. C. and Hinrichsen, H. (2009). Nonequilibrium phase transition in a spreading process on a timeline. *J. Statist. Mech.* **2009**, article no. P02020.
[6] Barbour, A. D. and Reinert, G. (2013). Approximating the epidemic curve. *Electron. J. Prob.* **18**, article no. 54, 30 pp.
[7] Barbour, A. D. and Röllin, A. (2019). Central limit theorems in the configuration model. *Ann. Appl. Prob.* **29**, 1046–1069.
[8] Barndorff-Nielsen, O. (1963). On the limit behaviour of extreme order statistics. *Ann. Math. Statist.* **34**, 992–1002.
[9] Baroni, E., van der Hofstad, R. and Komjáthy, J. (2015). *First passage percolation on random graphs with infinite variance degrees*. Preprint. Available at https://arxiv.org/abs/1506.01255.
[10] Barrat, A., Barthélemy, M. and Vespignani, A. (2008). *Dynamical Processes on Complex Networks*. Cambridge University Press.

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

[11] BHAMIDI, S., VAN DER HOFSTAD, R. AND HOOGHIEMSTRA, G. (2010). First passage percolation on random graphs with finite mean degrees. *Ann. Appl. Prob.* **20**, 1907–1965.
[12] BHAMIDI, S., VAN DER HOFSTAD, R. AND HOOGHIEMSTRA, G. (2011). First passage percolation on the Erdős–Rényi random graph. *Combin. Prob. Comput.* **20**, 683–707.
[13] BHAMIDI, S., VAN DER HOFSTAD, R. AND HOOGHIEMSTRA, G. (2017). Universality for first passage percolation on sparse random graphs. *Ann. Prob.* **45**, 2568–2630.
[14] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd edn. John Wiley, New York.
[15] BOHMAN, T. AND PICOLLELLI, M. (2012). SIR epidemics on random graphs with a fixed degree sequence. *Random Structures Algorithms* **41**, 179–214.
[16] BOLLOBÁS, B. (2001). *Random Graphs*, 2nd edn. Cambridge University Press.
[17] BOLLOBÁS, B. AND RIORDAN, O. (2006). *Percolation*. Cambridge University Press.
[18] BRAUER, F. AND VAN DEN DRIESSCHE, P. (2003). Some directions for mathematical epidemiology. In *Dynamical Systems and Their Applications in Biology* (Cape Breton Island, NS, **2001**) (Fields Inst. Commun., Vol. 36), American Mathematical Society, Providence, RI, pp. 95–112.
[19] CALLAWAY, D. S., NEWMAN, M. E., STROGATZ, S. H. AND WATTS, D. J. (2000). Network robustness and fragility: percolation on random graphs. *Phys. Rev. Lett.* **85**, article no. 5468.
[20] CHATTERJEE, S. AND DURRETT, R. (2009). Contact processes on random graphs with power law degree distributions have critical value 0. *Ann. Prob.* **37**, 2332–2356.
[21] DECREUSEFOND, L., DHERSIN, J.-S., MOYAL, P. AND TRAN, V. C. (2012). Large graph limit for an SIR process in random network with heterogeneous connectivity. *Ann. Appl. Prob.* **22**, 541–575.
[22] DURRETT, R. (2010). *Probability: Theory and Examples*, 4th edn. Cambridge University Press.
[23] DURRETT, R. (2010). *Random Graph Dynamics*. Cambridge University Press.
[24] DURRETT, R. (2010). Some features of the spread of epidemics and information on a random graph. *Proc. Nat. Acad. Sci. USA* **107**, 4491–4498.
[25] FLEMING, T. R. AND HARRINGTON, D. P. (1991). *Counting Processes and Survival Analysis*. John Wiley, New York.
[26] GLEESON, J. P. (2011). High-accuracy approximation of binary-state dynamics on networks. *Phys. Rev. Lett.* **107**, article no. 068701.
[27] GRAHAM, M. AND HOUSE, T. (2014). Dynamics of stochastic epidemics on heterogeneous networks. *J. Math. Biol.* **68**, 1583–1605.
[28] HELLAND, I. S. (1982). Central limit theorems for martingales with discrete or continuous time. *Scand. J. Statist.* **9**, 79–94.
[29] HINRICHSEN, H. (2006). Non-equilibrium phase transitions. *Physica A* **369**, 1–28.
[30] JACOBSEN, K. A., BURCH, M. G., TIEN, J. H. AND REMPAŁA, G. A. (2018). The large graph limit of a stochastic epidemic model on a dynamic multilayer network. *J. Biol. Dynamics* **12**, 746–788.
[31] JANSON, S., LUCZAK, M. AND WINDRIDGE, P. (2014). Law of large numbers for the SIR epidemic on a random graph with given degrees. *Random Structures Algorithms* **45**, 726–763.
[32] JANSON, S. AND LUCZAK, M. J. (2009). A new approach to the giant component problem. *Random Structures Algorithms* **34**, 197–216.
[33] KARATZAS, I. AND SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd edn. Springer, New York.
[34] KEPHART, J. O. AND WHITE, S. R. (1993). Measuring and modeling computer virus prevalence. In *Proc. 1993 IEEE Computer Society Symposium on Research in Security and Privacy*, Institute of Electrical and Electronics Engineers, Piscataway, NJ, pp. 2–15.
[35] LELARGE, M. (2012). Diffusion and cascading behavior in random networks. *Games Econom. Behavior* **75**, 752–775.
[36] MEYER, P. A. (1962). A decomposition theorem for supermartingales. *Illinois J. Math.* **6**, 193–205.
[37] MILLER, J. C. (2011). A note on a paper by Erik Volz: SIR dynamics in random networks. *J. Math. Biol.* **62**, 349–358.
[38] MILLER, J. C., SLIM, A. C. AND VOLZ, E. M. (2012). Edge-based compartmental modelling for infectious disease spread. *J. R. Soc. Interface* **9**, 890–906.
[39] MOLLOY, M. AND REED, B. (1995). A critical point for random graphs with a given degree sequence. *Random Structures Algorithms* **6**, 161–179.
[40] MOLLOY, M. AND REED, B. (1998). The size of the giant component of a random graph with a given degree sequence. *Combinatorics Prob. Comput.* **7**, 295–305.
[41] NEWMAN, M. E. J. (2002). Spread of epidemic disease on networks. *Phys. Rev. E* **66**, article no. 016128, 11 pp.
[42] PASTOR-SATORRAS, R. AND VESPIGNANI, A. (2001). Epidemic dynamics and endemic states in complex networks. *Phys. Rev. E* **63**, article no. 066117.

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press

[43] RAND, D. A. (1999). Correlation equations and pair approximations for spatial ecologies. In *Advanced Ecological Theory: Principles and Applications*, Blackwell Science, Oxford, pp. 100–142.
[44] REBOLLEDO, R. (1980). Central limit theorems for local martingales. *Z. Wahrscheinlichkeitsth.* **51**, 269–286.
[45] SPRUYT, V. (2014). How to draw a covariance error ellipse? Available at https://www.visiondummy.com/2014/04/draw-error-ellipse-representing-covariance-matrix.
[46] VAN DER HOFSTAD, R. (2010). Percolation and random graphs. In *New Perspectives in Stochastic Geometry*, Oxford University Press, pp. 173–247.
[47] VAN DER HOFSTAD, R. (2017). *Random Graphs and Complex Networks*, Vol. 1. Cambridge University Press.
[48] VOLZ, E. (2008). SIR dynamics in random networks with heterogeneous connectivity. *J. Math. Biol.* **56**, 293–310.
[49] WIERMAN, J. C. AND MARCHETTE, D. J. (2004). Modeling computer virus prevalence with a susceptible–infected–susceptible model with reintroduction. *Comput. Statist. Data Anal.* **45**, 3–23.

https://doi.org/10.1017/apr.2022.52 Published online by Cambridge University Press