J. Appl. Probab. **60,** 1079–1095 (2023)
*doi:*10.1017/jpr.2022.112

A MODEL FOR AN EPIDEMIC WITH CONTACT TRACING AND CLUSTER ISOLATION, AND A DETECTION PARADOX

JEAN BERTOIN ,* *University of Zurich*

Abstract

We determine the distributions of some random variables related to a simple model of an epidemic with contact tracing and cluster isolation. This enables us to apply general limit theorems for super-critical Crump–Mode–Jagers branching processes. Notably, we compute explicitly the asymptotic proportion of isolated clusters with a given size amongst all isolated clusters, conditionally on survival of the epidemic. Somewhat surprisingly, the latter differs from the distribution of the size of a typical cluster at the time of its detection, and we explain the reasons behind this seeming paradox.

Keywords: Crump–Mode–Jagers branching process; structured population model

2020 Mathematics Subject Classification: Primary 60J80
Secondary 92D25

1. Introduction

Predicting and controlling the evolution of epidemics has motivated mathematical contributions for a long time and generated a huge literature; let us merely point to the lecture notes [4] and references therein. Models involving contact tracing and isolation, which aim at reducing the transmissibility of infections, have raised significant interest; see in particular [1], [2], [3], [7], [13], [14], [15], and [17], among others. Below we present a toy model in this framework, which is clearly oversimplified (many important aspects such as the possibility of recovery, the age-dependency of the contamination rate, the spatial locations and displacements of infected individuals are not taken into account) and likely unrealistic for practical applications, but which is solvable in the sense that many quantities of interest can be computed explicitly. This model is close to the one introduced recently by Bansaye, Gu, and Yuan [2], as will be discussed in the final section of this article.

We take into account only infected individuals, implicitly assuming that there is an infinite reservoir of healthy individuals susceptible to becoming infected at some point. There is no death or recovery, but we distinguish between contagious individuals and those who have been isolated and have therefore ceased to spread the epidemic. The infected population grows with time as new individuals are contaminated; we suppose that a newly infected individual is always contaminated by a single contagious individual. Imagine further that when a contamination occurs, it can be either *traceable*, for instance in the case of a contamination between two relatives, or *untraceable*, for instance in the case when it occurs during a public event involving two unrelated individuals. At any time, there is thus a natural partition

Received 8 January 2022; revision received 3 October 2022.
* Postal address: Institute of Mathematics, University of Zurich, Winterthurerstrasse 190, 8057 Zürich, Switzerland. Email address: jean.bertoin@math.uzh.ch

of the infected population into *clusters*, where two individuals are parts of the same cluster if and only if the contamination path between those individuals can be fully traced. Finally, we suppose that individuals are randomly tested, and when a contagious individual is detected, then one isolates its entire cluster instantaneously. A newly infected individual is always contagious until it has been isolated, and then it ceases to contaminate further individuals forever. We stress the distinction between detection, which acts on individuals, and isolation, which follows from detection of a contagious individual and applies to a whole cluster. Clusters consisting of contagious individuals are called active, and then isolated after detection of an infected individual.

We now turn this model into a simple stochastic evolution depending on three parameters, namely:

- $\gamma > 0$, the contamination rate of a contagious individual,
- $p \in (0, 1)$, the probability of traceability for a contamination event,
- $\delta > 0$, the rate of detection for a contagious individual.

In words, the probability that a contagious individual at time t contaminates some healthy individual during the time interval $[t, t + \mathrm{d}t]$ is $\gamma\, \mathrm{d}t$, and when this occurs, the probability that the contamination is traceable is p. Simultaneously, the probability that a contagious individual is detected during the time interval $[t, t + \mathrm{d}t]$ is $\delta\, \mathrm{d}t$. We suppose that these events are mutually independent, simultaneously for all times and all contagious individuals. In particular, the probability that an active cluster of size s at time t is put into isolation during the time interval $[t, t + \mathrm{d}t]$ is $s\delta\, \mathrm{d}t$. Finally, we suppose for simplicity that at the initial time $t = 0$, there is a single infected individual in the population, which we call the ancestor.

The epidemic eventually stops once all contagious individuals have been isolated, and we shall see that this occurs almost surely if and only if the rate of detection is greater than or equal to the rate of untraceable contaminations, i.e. $\delta \geq (1 - p)\gamma$. Note that this is independent of the rate $p\gamma$ of traceable contaminations. We are mostly interested in the super-critical case $\delta < (1 - p)\gamma$ when the epidemic survives forever with strictly positive probability.

Our main results in the super-critical regime specify to our setting general limit theorems for Crump–Mode–Jagers branching processes. They show that the number of active, respectively isolated, clusters counted with some characteristic grows exponentially fast in time with exponent $\alpha = \alpha(\gamma, p, \delta)$ given by the Malthusian parameter. The limits after rescaling involve as a universal factor (i.e. independent of the chosen characteristics) the terminal value of the so-called intrinsic martingale. As a consequence, conditionally on survival of the epidemic, the empirical distribution of the sizes of active clusters (respectively of isolated clusters) converges as time goes to infinity. More precisely, we will show that the proportion of clusters of given size $k \geq 1$ amongst all active clusters at time t tends to

$$c_a(1 - \delta/\rho)^{k-1}\mathrm{B}(1 + \alpha/\rho, k) \tag{1.1}$$

as $t \to \infty$, whereas this proportion amongst isolated clusters at time t tends to

$$c_i(1 - \delta/\rho)^{k-1}\mathrm{B}(\alpha/\rho, k + 1), \tag{1.2}$$

where $\rho = \delta + p\gamma$, B denotes the beta function, and c_a and c_i are the normalization factors.

Concretely, the only observable variables at a given time in this model are the isolated clusters, since, by definition, the active ones have not yet been detected. Our results point

https://doi.org/10.1017/jpr.2022.112 Published online by Cambridge University Press

towards the following rather surprising feature (at least for non-specialists of general branching processes or structured population models). Since, loosely speaking, isolated clusters are independent with the same distribution, one might expect that when the epidemic has spread for a long time, the empirical distribution of the isolated clusters should be close to the law of a typical isolated cluster, that is, the cluster generated by a typical contagious individual at the time when it is isolated. However, it is easy to see that the size of a typical isolated cluster has the geometric distribution with success parameter δ/ρ, so the probability that a typical isolated cluster has size k equals

$$\frac{\delta}{\rho}(1-\delta/\rho)^{k-1},$$

which differs from (1.2). This is the detection paradox alluded to in the title of this work, and which will be explained in the last section. Note that the bias factor $\mathrm{B}(\alpha/\rho, k+1)$ in (1.2) decays as k increases, which entails that in the large time limit, the empirical isolated cluster is in fact stochastically smaller than the typical cluster.

The plan of this article is as follows. In Section 2 we explain how the model can be recast in terms of a Crump–Mode–Jagers branching process by focusing on the clusters. In Section 3 we describe the evolution of a typical cluster as a Yule process stopped at the time when it exceeds an independent geometric variable. This enables us to derive a number of related statistics explicitly, in particular regarding the point process of untraceable contaminations which are induced. Our main results on the large time behavior of the epidemic in the super-critical regime are presented in Section 4; they are merely deduced from a well-known general limit theorem for Crump–Mode–Jagers branching processes using the explicit formulas of Section 3. In Section 5 we first compare the model by Bansaye, Gu, and Yuan, and their approach, with ours. In particular, we observe that the Malthusian parameter α and the limiting distribution (1.1) solve a natural eigenproblem when the evolution of active cluster sizes is viewed as an age-structured population model. We then explain the detection paradox, and finally, we briefly discuss the relation between (1.1) and the classical Yule–Simon distribution.

2. The Crump–Mode–Jagers branching process of clusters

Although we introduced the epidemic model from the perspective of individuals, it will be convenient for its analysis to look at clusters and their evolutions as time passes. Specifically, imagine that we create an (unoriented) edge between the infector and the infected at the time when a contamination occurs; each edge is further labeled traceable or untraceable, depending on the type of contamination. If we ignore the labels of edges, this endows the infected population at any given time with a genealogical tree structure which is rooted at the ancestor. Plainly this tree structure grows as new individuals are contaminated and new traceable or untraceable edges are added. The reader may find Figure 1 useful to visualize the notions that will be introduced.

Any pair of infected individuals is connected by a unique segment in the tree, which we call the contamination path. Two individuals belong to the same cluster if and only if their contamination path contains only traceable edges, and more generally, the number of untraceable edges along the contamination path between two individuals only depends on the two clusters to which these individuals belong. We then define the generation of a given cluster as the number of untraceable edges along the contamination path between any individual in that cluster and the ancestor.

We next observe that backtracking contaminations endows clusters with a natural genealogy, which in turn enables us to view the epidemic model as a so-called Crump–Mode–Jagers

https://doi.org/10.1017/jpr.2022.112 Published online by Cambridge University Press

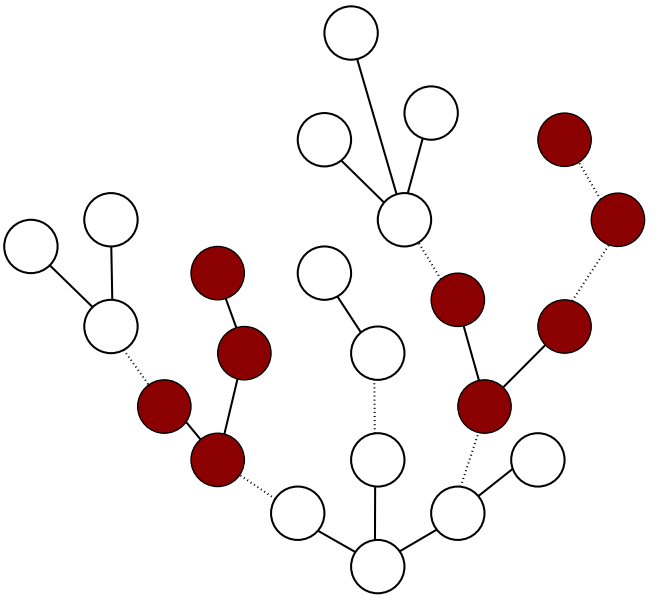

FIGURE 1. Graphical representation of the epidemic at a given time. The ancestor is the vertex at the bottom of the figure. Vertices in red represent contagious individuals, vertices in white individuals who have been isolated. Full edges indicate traceable contaminations, and dotted ones untraceable contaminations. Clusters consist of subsets of vertices connected by full edges. In turn, clusters are connected by dotted edges. There are four active clusters: two at the first generation with sizes 4 and 3, one at the third generation with size 1, and one at the fourth generation with size 1. There are four isolated clusters: the ancestor cluster with size 5, one cluster with size 2 at the first generation, and two clusters with sizes 4 and 3 at the second generation.

branching process; see [9, Chapter 6] as well as [10], [11], and [16] for classical background, and also [6, Section 5] for a more recent survey with further references. For this purpose, we shall index each cluster by a finite sequence of positive integers, that is, by a vertex u of the Ulam–Harris tree $\mathcal{U} = \bigcup_{n=0}^{\infty} \mathbb{N}^n$, such that the length $|u|$ of u corresponds to the generation of the cluster. By convention, the empty sequence $\varnothing$ with length 0 is used to label the cluster containing the ancestor. Clusters at the first generation are those such that there is a single untraceable edge along the contamination path from an infected individual in this cluster to the ancestor. They are indexed by $\mathbb{N}^1 = \mathbb{N} = \{1, 2, \ldots\}$ according to increasing order of their birth times, that is, times at which an individual in the ancestral cluster causes an untraceable contamination and generates a new cluster. For the sake of definitiveness, we agree that when the ancestral cluster generates only k untraceable contaminations until it is isolated, then the clusters indexed by $k+1, k+2, \ldots$ are fictitious clusters born at time ∞. This is only a formality, and of course we shall only be concerned with non-fictitious clusters. By an obvious iteration, we label clusters at the nth generation by $u = (u_1, \ldots, u_n) \in \mathbb{N}^n$ for any $n \geq 0$. It should be plain that the genealogy of clusters does not change with time, in the sense that once a cluster is born, its label will remain the same in the future.

For any $u \in \mathcal{U}$, if the cluster labeled by u is not fictitious, then we write ζ_u for the age (that is, the time elapsed from the birth time) at which this cluster is isolated. We also write ξ_u for the simple point process on $[0, \infty)$ of the ages at which this cluster is involved with untraceable

https://doi.org/10.1017/jpr.2022.112 Published online by Cambridge University Press

contaminations, that is, it generates new clusters. So $\xi_u([0, t])$ is the number of children clusters generated when the cluster reaches age t, and in particular, $\xi_u([\zeta_u, \infty)) = 0$. Finally, we write $C_u = (C_u(t), t \geq 0)$, where $C_u(t)$ is the size of this cluster at time $t < \zeta_u$, and we agree that $C_u(t) = 0$ whenever $t \geq \zeta_u$. In words, C_u is the process of the number of contagious individuals in that cluster as a function of its age (recall that infected individuals are no longer contagious once they have been isolated); in particular, $C(\zeta -)$ is the size reached by the cluster at the time when it is detected. The most relevant information about the evolution of clusters and hence about the epidemic is encoded by the family of pairs

$$\mathbf{C}_u = (C_u, \xi_u), \quad u \in \mathcal{U},$$

where we agree for definitiveness that $C_u \equiv 0$ and $\xi_u \equiv 0$ when the cluster indexed by u is fictitious. Of course, $\mathbf{C}_u$ does not enable us to recover the subtree structure of the cluster indexed by u, but this is irrelevant for the questions we are interested in.

It should be intuitively clear that the distribution of the ancestral cluster $\mathbf{C}_\varnothing$ determines that of the whole process $(\mathbf{C}_u)_{u\in\mathcal{U}}$. More precisely, let us write $\mathbf{C} = (C, \xi)$ for a pair distributed as $\mathbf{C}_\varnothing$, which we think of as describing the evolution of a *typical cluster*. Then it is readily checked that conditionally on $\xi_\varnothing(\mathbb{R}_+) = k$, $\mathbf{C}_1, \ldots, \mathbf{C}_k$ are k independent and identically distributed (i.i.d.) copies of $\mathbf{C}$ that are further independent of $\mathbf{C}_\varnothing$. More generally, it follows by iteration that for every $n \geq 1$ and any $u^1, \ldots, u^k \in \mathbb{N}^n$, conditionally on the event that none of the clusters $\mathbf{C}_{u^1}, \ldots, \mathbf{C}_{u^k}$ are fictitious (which is measurable with respect to the family $(\mathbf{C}_v\colon |v| < n)$), $\mathbf{C}_{u^1}, \ldots, \mathbf{C}_{u^k}$ are k i.i.d. copies of $\mathbf{C}$ that are further independent of $(\mathbf{C}_v\colon |v| < n)$. In other words, $(\mathbf{C}_u)_{u\in\mathcal{U}}$ generates a Crump–Mode–Jagers branching process where the evolution of typical elements is distributed as $\mathbf{C}$.

Remark 2.1. If we interpret the isolation time ζ of an active cluster as the death time, and if we further view the size C(t) at time t as measuring some 'age' of the cluster, in the loose sense that this quantity grows with time until death occurs, then we are essentially in the framework of age-structured population models; see e.g. [12, Section II.E]. This aspect will be useful in the forthcoming Section 5.1. In this area, we further refer to [7] for a different model for contact tracing in an epidemic in terms of a disease age-structured population.

Remark 2.2. The arguments in this section are rather robust, in the sense that they remain valid for more sophisticated versions of the model. For instance, one could incorporate recovery, let the contamination rates depend on the age of the infection, etc. However, the quantitative results in the next section are much more fragile; notably, the calculations for the key Lemma 3.1 there cannot be adapted even to deal with recovery or death.

3. Statistics of a typical cluster

We discuss here some basic statistics of the typical cluster $\mathbf{C} = (C, \xi)$ in terms of the parameters (γ, p, δ) of the model. Recall that the integer-valued process C is absorbed at 0 at the time

$$\zeta = \inf\{t \geq 0\colon C(t) = 0\}$$

when this cluster is detected and isolated, and that ξ is the point process of times at which untraceable contaminations occur.

It is now convenient to set

$$\rho := \delta + p\gamma, \tag{3.1}$$

https://doi.org/10.1017/jpr.2022.112 Published online by Cambridge University Press

and recall that a Yule process with rate $\rho > 0$ refers to a pure birth process with birth rate ρk from any state $k \geq 1$ and started from 1.

Lemma 3.1. *The process* $(C(t), t \geq 0)$ *has the same law as*

$$(\mathbf{1}_{\{Y(t)\leq G\}} Y(t), t \geq 0),$$

where $Y = (Y(t), t \geq 0)$ *is a Yule process with rate* ρ*, and* G *is an independent geometric variable with success probability* δ/ρ*, i.e. with tail distribution function*

$$\mathbb{P}(G > k) = (1 - \delta/\rho)^k, \quad k \geq 0.$$

Proof. The process C is a continuous-time Markov chain on $\mathbf{Z}_+ = \{0, 1, \ldots\}$, which starts from 1 at time 0 and is absorbed at the cemetery state 0. Recall that only traceable contaminations contribute to the growth of the cluster, that they occur at rate $p\gamma$ per contagious individual, and that each individual in the cluster is detected at rate δ.

We see that when the chain is at some state $k \geq 1$, its next jump occurs after a waiting time with the exponential distribution with parameter $k(p\gamma + \delta) = k\rho$, and independently of this waiting time, the state after that jump is $k + 1$ with probability $p\gamma/\rho$, and 0 with complementary probability δ/ρ. In particular, the size reached by the cluster when it is isolated is a geometric variable with success probability δ/ρ. Our claim follows from the classical properties of independent exponential variables. □

Lemma 3.1 shows in particular that the size $C(\zeta-)$ of the typical isolated cluster has the geometric distribution with success probability δ/ρ. The one-dimensional marginal laws of the typical cluster size process as well as the joint distribution of the time of isolation ζ and $C(\zeta-)$ follow readily.

Corollary 3.1. *For every* $t \geq 0$*, we have*

$$\mathbb{P}(C(t) = k) = (1 - \delta/\rho)^{k-1}(1 - \mathrm{e}^{-\rho t})^{k-1}\, \mathrm{e}^{-\rho t} \quad \textit{for } k \geq 1$$

and

$$\mathbb{P}(C(t) = 0) = \mathbb{P}(\zeta \leq t) = 1 - \frac{\rho}{\rho + \delta(\mathrm{e}^{\rho t} - 1)}.$$

Furthermore, we also have

$$\mathbb{P}(C(\zeta-) = k, t \geq \zeta) = \frac{\delta}{\rho}(1 - \delta/\rho)^{k-1}(1 - \mathrm{e}^{-\rho t})^k \quad \textit{for } k \geq 1.$$

Proof. It suffices to write for $k \geq 1$ that

$$\begin{aligned}\mathbb{P}(C(t) = k) &= \mathbb{P}(Y(t) = k, G \geq k)\\ &= \mathbb{P}(Y(t) = k)\mathbb{P}(G \geq k),\end{aligned}$$

and recall that $Y(t)$ has the geometric distribution with success probability $\mathrm{e}^{-\rho t}$. Then summation for $k \geq 1$ yields the second formula of the statement. We get the third formula similarly, writing for $k \geq 1$ that

$$\mathbb{P}(C(\zeta-) = k, t \geq \zeta) = \mathbb{P}(G = k, Y(t) > k) = \mathbb{P}(G = k)\mathbb{P}(Y(t) > k).$$ □

https://doi.org/10.1017/jpr.2022.112 Published online by Cambridge University Press

We next turn our attention to the point process ξ at which new clusters are generated, and write

$$Z_1 := \xi(\mathbb{R}_+)$$

for the total number of (non-fictitious) clusters that the typical cluster begets. Its distribution is obtained by a slight variation of the argument for Lemma 3.1, and this entails the criterion for extinction of the epidemic that was stated in the Introduction.

Lemma 3.2. *The variable $1 + Z_1$ follows the geometric distribution with success probability $\delta/((1-p)\gamma + \delta)$. In particular, $Z_1 \in L^r(\mathbb{P})$ for all $r \geq 1$,*

$$\mathbb{E}(Z_1) = (1-p)\gamma/\delta,$$

and as a consequence, the total number of infected individuals is finite (in other words, the epidemic eventually ceases) almost surely if and only if

$$(1-p)\gamma \leq \delta.$$

Proof. Fix some arbitrary time $t \geq 0$, and work conditionally on the event that at time t, the typical cluster has size $k \geq 1$ and is still active. Consider the first event after time t at which either there is a new traceable or untraceable contamination, or the cluster is detected. The probability that this event is due to an untraceable contamination is $(1-p)\gamma/(\gamma+\delta)$, whereas the probability that this event is due to detection is $\delta/(\gamma+\delta)$. In the remaining case, the size of the cluster increases by one unit.

The probabilities above depend on neither t nor k, and it follows by iteration that if we now introduce the first instant τ after t at which either an untraceable contamination occurs or the cluster is detected, then independently of $C(\tau)$, the probability that τ is the time of an untraceable contamination equals $((1-p)\gamma)/((1-p)\gamma+\delta)$ (this is the failure probability). Another iteration yields our first claim, and the formula for the first moment of Z_1 follows.

Finally, if we write Z_n for the number of (non-fictitious) clusters at the nth generation, then $(Z_n, n \geq 0)$ is a Galton–Watson process with reproduction law distributed according to Z_1. So if $\delta < (1-p)\gamma$, there is a strictly positive probability that this Galton–Watson process survives for ever, in which case the total number of infected individuals is obviously infinite. Otherwise, the Galton–Watson process eventually becomes extinct almost surely: there are only finitely many (non-fictitious) clusters, each consisting of finitely many infected individuals. □

Last, we introduce the intensity measure μ of the point process ξ:

$$\mu(t) := \mathbb{E}(\xi([0,t])), \quad t \geq 0.$$

Corollary 3.2. *For every $t \geq 0$, there is the identity*

$$\mu(t) = (1-p)\frac{\gamma}{\delta}\left(1 - \frac{1}{1+\delta(\mathrm{e}^{\rho t}-1)/\rho}\right).$$

https://doi.org/10.1017/jpr.2022.112 Published online by Cambridge University Press

Proof. Indeed, the conditional probability given the process C, that an untraceable contamination event occurs during the time interval $[t, t+\mathrm{d}t]$, equals $(1-p)\gamma C(t)\,\mathrm{d}t$, and as a consequence,

$$\mathrm{d}\mu(t) = (1-p)\gamma\mathbb{E}(C(t))\,\mathrm{d}t.$$

We deduce from Corollary 3.1 that

$$\mathbb{E}(C(t)) = \frac{\mathrm{e}^{\rho t}}{(1+\delta(\mathrm{e}^{\rho t}-1)/\rho)^2}.$$

The formula in the statement follows. □

We note that letting $t\to\infty$ in Corollary 3.2 yields

$$\mathbb{E}(\xi(\mathbb{R}_+)) = (1-p)\gamma/\delta,$$

in agreement with Lemma 3.2.

4. The Malthusian behavior

We shall assume throughout this section that

$$\delta < (1-p)\gamma, \tag{4.1}$$

so that the epidemic survives with strictly positive probability. More precisely, one immediately deduces from Lemma 3.2 that the probability of extinction equals $\delta/((1-p)\gamma)$, which is the smallest solution to the equation $\mathbb{E}(x^{Z_1}) = x$. We shall derive here the main results of this work, simply by specifying in our setting some fundamental results of Nerman [16] on the asymptotic behavior of Crump–Mode–Jagers branching processes with random characteristics. We start by introducing some of the key actors in this framework.

Consider the Laplace transform of the intensity measure of untraceable contaminations for a typical cluster:

$$\mathcal{L}(x) = \int_0^\infty \mathrm{e}^{-xt}\mathrm{d}\mu(t), \quad x\geq 0.$$

Since $\mathcal{L}(0) = (1-p)\gamma/\delta > 1$, the equation $\mathcal{L}(x) = 1$ possesses a unique solution $\alpha = \alpha(\gamma, p, \delta)\in(0,\infty)$, called the *Malthusian parameter*. That is, thanks to Corollary 3.2,

$$(1-p)\gamma\int_0^\infty \frac{\mathrm{e}^{(\rho-\alpha)t}}{(1+\delta(\mathrm{e}^{\rho t}-1)/\rho)^2}\,\mathrm{d}t = 1,$$

or equivalently, in a slightly simpler form, using the change of variables $x = \mathrm{e}^{-\rho t}$,

$$(1-p)\gamma\rho\int_0^1 \frac{x^{\alpha/\rho}}{((\rho-\delta)x+\delta)^2}\,\mathrm{d}x = 1. \tag{4.2}$$

We further set

$$\beta = -\mathcal{L}'(\alpha) = (1-p)\gamma\int_0^\infty t\frac{\mathrm{e}^{(\rho-\alpha)t}}{(1+\delta(\mathrm{e}^{\rho t}-1)/\rho)^2}\,\mathrm{d}t; \tag{4.3}$$

plainly $\beta\in(0,\infty)$.

https://doi.org/10.1017/jpr.2022.112 Published online by Cambridge University Press

Next, it is convenient to use the notation

$$\langle m, f\rangle := \sum_{n=1}^{\infty} f(n)m(n),$$

where $m = (m(n), n \in \mathbb{N})$ is a finite measure on $\mathbb{N}$ and $f\colon \mathbb{N} \to \mathbb{R}_+$ is a generic non-negative function. We introduce two important measures m^a and m^i, related to typical active and isolated clusters respectively, by

$$\langle m^a, f\rangle = \int_0^{\infty} \mathrm{e}^{-\alpha t}\mathbb{E}(f(C(t)), t < \zeta)\,\mathrm{d}t$$

and

$$\langle m^i, f\rangle = \int_0^{\infty} \mathrm{e}^{-\alpha t}\mathbb{E}(f(C(\zeta-)), \zeta \le t)\,\mathrm{d}t.$$

These two measures can be determined explicitly from Corollary 3.1, using the notation

$$\mathrm{B}(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} = \int_0^1 s^{x-1}(1-s)^{y-1}\mathrm{d}s, \quad x, y > 0,$$

for the beta function. Indeed, from the change of variables $\mathrm{e}^{-\rho t} = s$, we then obtain that, for every $k \ge 1$,

$$\begin{aligned} m^a(k) &= (1-\delta/\rho)^{k-1}\int_0^{\infty} \mathrm{e}^{-\alpha t}(1-\mathrm{e}^{-\rho t})^{k-1}\,\mathrm{e}^{-\rho t}\mathrm{d}t \\ &= \frac{1}{\rho}(1-\delta/\rho)^{k-1}\mathrm{B}(1+\alpha/\rho, k) \end{aligned} \tag{4.4}$$

and

$$\begin{aligned} m^i(k) &= \frac{\delta}{\rho}(1-\delta/\rho)^{k-1}\int_0^{\infty} \mathrm{e}^{-\alpha t}(1-\mathrm{e}^{-\rho t})^{k}\mathrm{d}t \\ &= \frac{\delta}{\rho^2}(1-\delta/\rho)^{k-1}\mathrm{B}(\alpha/\rho, k+1). \end{aligned} \tag{4.5}$$

Finally, we introduce

$$W_n = \sum_{u\in\mathbb{N}^n} \mathrm{e}^{-\alpha\sigma_u}, \quad n \ge 0,$$

where σ_u stands for the birth time of the cluster labeled by u (so that $\sigma_u = \infty$ and $\mathrm{e}^{-\alpha\sigma_u} = 0$ if this cluster is fictitious). The process $(W_n, n \ge 0)$ is a martingale, often referred to as the intrinsic martingale; see Jagers [9, Chapter 6]). Using the inequality $W_1 \le \xi(\mathbb{R}_+)$ and Lemma 3.2, we see that

$$\mathbb{E}(W_1^2) < \infty,$$

and the uniform integrability of the intrinsic martingale follows; see e.g. [10, Theorem 6.1]. We furthermore recall that its terminal value W_∞ is strictly positive on the event that the epidemic survives, and of course $W_\infty = 0$ on the event that the epidemic eventually ceases.

https://doi.org/10.1017/jpr.2022.112 Published online by Cambridge University Press

For the sake of simplicity, we focus on a few natural statistics of the epidemic at time $t \geq 0$. Given a generic non-negative function $f\colon \mathbb{N} \to \mathbb{R}_+$, we agree implicitly that $f(0) = 0$ and write

$$A^f(t) = \sum f(C_u(t - \sigma_u)),$$

where the sum is taken over all vertices u in the Ulam–Harris tree $\mathcal{U}$, such that the cluster labeled by u is born at time $\sigma_u \leq t$ (note that only active clusters at time t contribute to the sum). Turning our attention to isolated clusters, we similarly write

$$I^f(t) = \sum f(C_u(\zeta_u -))\mathbf{1}_{\{\sigma_u + \zeta_u \leq t\}},$$

where the sum is taken over all clusters that are isolated at time t. In the Crump–Mode–Jagers terminology, A^f and I^f are known as the processes counted with the random characteristics

$$\phi^a\colon t \mapsto f(C(t))\mathbf{1}_{\{t<\zeta\}} \quad \text{and} \quad \phi^i\colon t \mapsto f(C(\zeta -))\mathbf{1}_{\{t\geq\zeta\}}, \tag{4.6}$$

respectively.

Recall the notation above, and notably (4.2), (4.3), (4.4), and (4.5). We can now state the main result of this work.

Theorem 4.1. *Assume* (4.1) *and let* $f\colon \mathbb{N} \to \mathbb{R}_+$ *with* $f(n) = O(\mathrm{e}^{bn})$ *for some* $b < -\log(1 - \delta/\rho)$. *The following limits then hold almost surely and in* $L^1(\mathbb{P})$*:*

$$\lim_{t\to\infty} \mathrm{e}^{-\alpha t} A^f(t) = \beta^{-1} \langle m^a, f\rangle W_\infty$$

and

$$\lim_{t\to\infty} \mathrm{e}^{-\alpha t} I^f(t) = \beta^{-1} \langle m^i, f\rangle W_\infty.$$

In particular, taking $f(n) = n$ yields the first-order asymptotic behavior of the total number of contagious (respectively isolated) individuals as time goes to infinity.

Proof. The claim of almost sure convergence is seen from Theorem 5.4 in Nerman [16]; we just need to verify Conditions 5.1 and 5.2 there. For the first, we simply write

$$\int_t^\infty \mathrm{e}^{-\alpha s}\xi(\mathrm{d}s) \leq \mathrm{e}^{-\alpha t}\xi(\mathbb{R}_+),$$

and recall from Lemma 3.2 that $Z_1 = \xi(\mathbb{R}_+)$ is integrable. This ensures that

$$\mathbb{E}\left(\sup_{t\geq 0} \mathrm{e}^{\alpha t} \int_t^\infty \mathrm{e}^{-\alpha s}\xi(\mathrm{d}s)\right) < \infty.$$

For the second, we assume for simplicity that $|f(n)| \leq \mathrm{e}^{bn}$ for all $n \geq 1$ without loss of generality. The random characteristics in (4.6) can be bounded for all $t \geq 0$ by

$$|\phi^a(t)| \leq \exp(bC(\zeta -)) \quad \text{and} \quad |\phi^i(t)| \leq \exp(bC(\zeta -)).$$

Observe that

$$\mathbb{E}(\exp(bC(\zeta -))) < \infty,$$

https://doi.org/10.1017/jpr.2022.112 Published online by Cambridge University Press

since we know from Lemma 3.1 that $C(\zeta-)$ has the geometric distribution with success probability δ/ρ, and $(1-\delta/\rho)\,\mathrm{e}^b < 1$. It follows immediately that Condition 5.2 of [16] holds for both ϕ^a and ϕ^i.

We next turn our attention to convergence in $L^1(\mathbb{P})$. This in turn relies on [16, Corollary 3.3], and we have to check equations 3.1 and 3.2 therein. The latter are both straightforward from the bounds established in the first part of this proof. □

Remark 4.1. Iksanov, Kolesko, and Meiners [8] have recently obtained a remarkable central limit theorem for general Crump–Mode–Jagers branching processes counted with random characteristics, which specifies the fluctuations of Nerman's law of large numbers. Of course, it would be interesting to apply their results to our setting; however, in order to do so, one needs information about the possible roots to the equation $\mathcal{L}(z) = 1$ in the complex strip $\alpha/2 \le z \le \alpha$, which does not seem easy to obtain even though the intensity μ is explicitly known.

We next turn our attention to the empirical distribution of the sizes of active, respectively isolated, clusters. We denote the function identical to 1 on $\mathbb{N}$ by $\mathbf{1}$, so that in the above notation, $A^{\mathbf{1}}(t)$ and $I^{\mathbf{1}}(t)$, respectively, are the number of active and of isolated clusters at time t. We then define the empirical distributions $\Pi^a(t)$ and $\Pi^i(t)$ for a generic function $f\colon \mathbb{N} \to \mathbb{R}_+$ by

$$\langle \Pi^a(t), f\rangle = A^f(t)/A^{\mathbf{1}}(t)$$

and

$$\langle \Pi^i(t), f\rangle = I^f(t)/I^{\mathbf{1}}(t).$$

We also introduce the normalized probability measures on $\mathbb{N}$:

$$\pi^a = m^a/\langle m^a, \mathbf{1}\rangle \quad \text{and} \quad \pi^i = m^i/\langle m^i, \mathbf{1}\rangle.$$

Thanks to (4.4) and (4.5), these are given explicitly by

$$\pi^a(k) = c_a(1-\delta/\rho)^{k-1}\mathrm{B}(1+\alpha/\rho, k) \quad \text{for all } k \ge 1, \tag{4.7}$$

with

$$1/c_a = \sum_{j=1}^{\infty} (1-\delta/\rho)^{j-1}\mathrm{B}(1+\alpha/\rho, j),$$

and

$$\pi^i(k) = c_i(1-\delta/\rho)^{k-1}\mathrm{B}(\alpha/\rho, k+1) \quad \text{for all } k \ge 1, \tag{4.8}$$

with

$$1/c_i = \sum_{j=1}^{\infty} (1-\delta/\rho)^{j-1}\mathrm{B}(\alpha/\rho, j+1).$$

We can now state the convergence of the empirical distributions.

Corollary 4.1. *Assume* (4.1). *Then, conditionally on the event that the epidemic survives forever, we have almost surely*

$$\lim_{t\to\infty} \Pi^a(t) = \pi^a \quad \text{and} \quad \lim_{t\to\infty} \Pi^i(t) = \pi^i.$$

https://doi.org/10.1017/jpr.2022.112 Published online by Cambridge University Press

Proof. Indeed, recall that $W_\infty > 0$ a.s. conditionally on survival of the epidemic. We derive from Theorem 4.1 that, on this event,

$$\lim_{t\to\infty} A^f(t)/A^{\mathbf{1}}(t) = \langle \pi^a, f\rangle \quad \text{and} \quad \lim_{t\to\infty} I^f(t)/I^{\mathbf{1}}(t) = \langle \pi^i, f\rangle$$

for every bounded function $f\colon \mathbb{N}\to\mathbb{R}$. □

In words, Corollary 4.1 states that conditionally on survival of the epidemic, the empirical distributions of active cluster sizes and of isolated clusters sizes converge to π^a and π^i, respectively, as time goes to infinity. We shall therefore think of the latter as describing asymptotically the average distributions of the sizes of active clusters and isolated clusters, respectively. It is interesting to point out similarities between Corollary 4.1 and earlier results by Deijfen [5, Theorem 1.1 and Example 2] on the asymptotic degree distribution for certain random evolving networks. More specifically, in this model new vertices arrive in continuous time, are connected to an existing vertex with probability proportional to the so-called fitness of that vertex, and vertices then die at rates depending on their accumulated in-degrees. Although the model considered by Deijfen is different from ours, it also bears a clear resemblance, and the similarities between the results (and the methods as well) should not come as a surprise.

It is also interesting to observe from formulas (4.7) and (4.8) and the elementary identity

$$\frac{\alpha}{\rho}\mathrm{B}(\alpha/\rho, k+1) = k\mathrm{B}(1+\alpha/\rho, k)$$

that

$$\pi^i(k) = \frac{k\pi^a(k)}{\sum_{j=1}^\infty j\pi^a(j)} \quad \text{for all } k\geq 1.$$

In words, the average distribution of the sizes of isolated clusters is the size-biased version of that of active clusters. This relation stems from the fact that the rate at which an active cluster becomes isolated is proportional to its size. Since the empirical distribution of active cluster sizes converges to π^a, the empirical distribution of isolated cluster sizes must converge to the size-biased version of π^a. We refer to Corollary 1 of [2] and its proof for details of a rigorous argument.

5. Concluding comments

5.1. Comparison with a model of Bansaye, Gu, and Yuan, and an eigenproblem

This work was inspired by a recent manuscript of Bansaye *et al.* [2], in which they introduced a similar model for epidemics with contact tracing and cluster isolation. The main difference from the present article is that in [2], contaminations are always traceable initially, but traceability gets lost at some fixed rate. In other words, edges in the contamination tree are traceable when they first appear, and become untraceable after exponentially distributed time-laps, independently of the other edges. As a consequence, clusters do not only grow when new contamination events occur, but also split when a traceable edge becomes untraceable.

Bansaye *et al.* investigate the large time asymptotic behavior of the epidemic using different tools, namely they first analyze a deterministic eigenproblem for a growth–fragmentation–isolation equation that is naturally related to their setting; furthermore they also rely on known properties of random recursive trees. They establish results similar to our Theorem 4.1 and

https://doi.org/10.1017/jpr.2022.112 Published online by Cambridge University Press

Corollary 4.1 in terms of these eigenelements; the statements in [2] are, however, less precise than ours, as no explicit formulas for the eigenelements are given (only their existence is established).

In our setting, using the notation of Section 4, the expectation of linear functionals of clusters at a given time yields a family $(\nu_t, t \geq 0)$ of measures on $\mathbb{N}$ given by

$$\langle \nu_t, f \rangle = \mathbb{E}(A^f(t)),$$

where $f\colon \mathbb{N} \to \mathbb{R}_+$ is a generic bounded function. From the dynamics of the epidemic, we get the evolution equation

$$\mathrm{d}\langle \nu_t, f \rangle = \langle \nu_t, \mathcal{A}f \rangle \, \mathrm{d}t, \tag{5.1}$$

with

$$\mathcal{A}f(k) = k(p\gamma(f(k+1) - f(k)) + (1-p)\gamma f(1) - \delta f(k)); \tag{5.2}$$

the initial condition ν_0 is the Dirac mass at 1 since we assume that the epidemic starts from a single contagious individual. Specifically, in (5.2), the term $kp\gamma(f(k+1) - f(k))$ accounts for the growth of a cluster from size k to $k+1$, which occurs with rate $kp\gamma$. The term $k(1-p)\gamma f(1)$ stems from the birth of new clusters of size 1 (i.e. an untraceable contamination) induced by a cluster of size k, which occurs with rate $k(1-p)\gamma$, and finally $-k\delta f(k)$ for the isolation of a cluster of size k, which occurs with rate $k\delta$. This formula for the infinitesimal generator $\mathcal{A}$ should be compared with Lemma 1 in [2], and notably equation (4.15) therein.

Predominantly, growth–fragmentation equations (and more generally, evolution equations) cannot be solved explicitly, and most works in this area are concerned with the large time asymptotic behavior of its solutions; see [2] for some references. Roughly speaking, the paradigm, which stems from the Perron–Frobenius theorem for matrices with positive entries, is to resolve the eigenproblem for the infinitesimal generator, that is, to determine the principal eigenvalue (i.e. the eigenvalue with the largest real part) and its left eigenfunctions. The principal eigenvalue is identified as the Malthusian parameter, and the left eigenfunction (viewed as a measure) yields the so-called asymptotic profile, that is, in our setting, the measure m^a in Theorem 4.1. So the analysis carried out in the present Section 4 solves this eigenproblem indirectly for (5.2), the solution being given by (4.2) and (4.7). Specifically, it holds for all bounded $f\colon \mathbb{N} \to \mathbb{R}_+$ that

$$\langle \mathcal{A}^\top \pi^a, f \rangle := \langle \pi^a, \mathcal{A}f \rangle = \alpha \langle \pi^a, f \rangle.$$

However, it does not seem straightforward to check this identity directly, and we shall provide more details below.

Let $\nu_t(k)$ denote the expected number of active clusters of size $k \geq 1$ at time t. From the dynamics of the epidemic (see the discussion following (5.2)), we have that for $k \geq 2$

$$\frac{\partial \nu_t(k)}{\partial t} + p\gamma(k\nu_t(k) - (k-1)\nu_t(k-1)) = -\delta k \nu_t(k), \tag{5.3}$$

whereas for $k = 1$

$$\frac{\partial \nu_t(1)}{\partial t} + p\gamma \nu_t(1) = (1-p)\gamma \sum_{j=1}^{\infty} j\nu_t(j) - \delta \nu_t(1). \tag{5.4}$$

Of course, (5.3) and (5.4) are equivalent to the evolution equation (5.1). From the point of view of age-structured population models (recall Remark 2.1), these should be viewed as a version of the McKendrick–von Foerster PDE; see [12, equations (23.4) and (23.5)].

https://doi.org/10.1017/jpr.2022.112 Published online by Cambridge University Press

Following [12, Chapter 23], it is then natural to search for a special solution to (5.3) and (5.4) in the form $\nu_t(k) = \mathrm{e}^{rt}\nu(k)$ for some $r > 0$ and some measure ν on $\mathbb{N}$, which of course amounts to solving the eigenproblem $\mathcal{A}^\top \nu = r\nu$. Recall the notation (3.1); from (5.3) and (5.4), we first get the linear recurrence equation

$$\nu(k) = \frac{p\gamma(k-1)}{r+\rho k}\nu(k-1), \quad k \geq 2, \tag{5.5}$$

and then, for $k = 1$, the identity

$$(r+\rho)\nu(1) = (1-p)\gamma \sum_{j=1}^{\infty} j\nu(j). \tag{5.6}$$

We readily deduce from (5.5) and well-known properties of the beta function B that

$$\nu(k) = c(1-\delta/\rho)^{k-1}\mathrm{B}(1+r/\rho, k), \quad k \geq 1, \tag{5.7}$$

where $c > 0$ is some arbitrary constant. We note that $(r+\rho)\nu(1)/c = \rho$, and can now determine r by rewriting (5.6) in the form

$$\begin{aligned}
\rho &= (1-p)\gamma \sum_{j=1}^{\infty} j(1-\delta/\rho)^{j-1}\mathrm{B}(1+r/\rho, j) \\
&= (1-p)\gamma \sum_{j=1}^{\infty} \int_0^1 j(1-\delta/\rho)^{j-1}(1-x)^{j-1}x^{r/\rho}\,\mathrm{d}x \\
&= (1-p)\gamma \int_0^1 \frac{x^{r/\rho}}{(1-(1-\delta/\rho)(1-x))^2}\,\mathrm{d}x \\
&= (1-p)\gamma\rho^2 \int_0^1 \frac{x^{r/\rho}}{((\rho-\delta)x+\delta)^2}\,\mathrm{d}x.
\end{aligned}$$

We have recovered (4.2), which determines the Malthusian parameter. So $r = \alpha$, and we conclude from (5.7) and (4.7) that ν and π^a are indeed proportional.

The calculations above are reminiscent of those for the Leslie model [12, Chapter 22], and in particular (4.2) can be thought of as an Euler–Lotka equation [12, equation (20.6)].

5.2. A detection paradox

We next discuss in more detail the detection paradox mentioned in the Introduction. Imagine that we rank the clusters in increasing order of their birth times rather than indexing them by the Ulam–Harris genealogical tree as we did previously. This sequence is infinite if and only if the epidemic survives, and conditionally on that event, its elements are independent, each being distributed as the typical cluster. One may then expect from the law of large numbers that as time goes to infinity, the limit π^i of the empirical distribution of the sizes of isolated clusters should coincide with the law of the size of a typical cluster at the time when it is detected, that is, by Lemma 3.1, the geometric distribution with success probability δ/ρ. However, (4.8) shows that this is not the case, and more precisely, π^i is a biased version of the geometric law, where the bias is given by a beta function.

https://doi.org/10.1017/jpr.2022.112 Published online by Cambridge University Press

The naive argument above of course has a flaw, which stems from the fact that the empirical distribution of the isolated clusters at a given time corresponds to a partial sum of clusters which are listed in increasing order of their detection times rather than their birth times. This reordering tends to list first clusters which are quickly detected and hence had little time to grow, which hints at the feature that the average isolated cluster size is dominated stochastically by the size of a typical isolated cluster. Nonetheless, reordering alone is not sufficient to explain the detection paradox: the second crucial ingredient is the exponential growth, and more precisely the fact that the number of clusters, say for simplicity born during the time interval $[t, t+1]$, is of the same order as the number of all the clusters born before time t, no matter how large t is. A significant proportion of clusters born during $[t, t+1]$ are detected before time $t+1$; due to the time constraint, these clusters have on average a smaller size than the typical cluster when it is isolated, and this explains the seeming paradox.

For a better understanding of the mechanisms at work in the explanation above, it may be useful to consider the following elementary example. Consider a Poisson point process on $\mathbb{R}_+ \times (0, \infty)$ whose atoms are denoted generically by (b, ℓ), and which has intensity $\mathrm{e}^b \mathrm{d}b\lambda(\mathrm{d}\ell)$, where λ is some probability measure on $(0, \infty)$. We think of (b, ℓ) as an individual born at time b and with lifespan ℓ. Imagine that we want to estimate the lifespan distribution λ, that is, more specifically, the quantity $\langle \lambda, f \rangle$ for an arbitrary bounded continuous function $f\colon (0, \infty) \to \mathbb{R}$, from the observation of the population up to some large time t. If we could observe the lifespan of an individual at the time when it is born, then this would be an easy matter. Indeed, it then suffices to compute the empirical mean of $f(\ell)$ for individuals (b, ℓ) born at time $b \le t$, and it is readily checked by Poissonian computation that this quantity converges almost surely to $\langle \lambda, f \rangle$ as $t \to \infty$. But of course it is unrealistic to assume that the lifespan can be observed at the birth of an individual, and let us instead assume that lifespan can be observed at death only.

The total number of dead individuals at time t has the Poisson distribution with parameter

$$\int_0^t \mathrm{e}^b \lambda((0, t-b])\, \mathrm{d}b \sim \mathrm{e}^t \int_{(0,\infty)} \mathrm{e}^{-\ell} \lambda(\mathrm{d}\ell) \quad \text{as } t \to \infty.$$

More generally, it is easily checked that if we write $\langle M(t), f \rangle$ for the empirical mean of $f(\ell)$ computed for all individuals who are dead at time t, that is, such that $b + \ell \le t$, then

$$\lim_{t \to \infty} \langle M(t), f \rangle = \langle \lambda_1, f \rangle,$$

where

$$\lambda_1(\mathrm{d}\ell) = \frac{\mathrm{e}^{-\ell} \lambda(\mathrm{d}\ell)}{\int_{(0,\infty)} \mathrm{e}^{-s} \lambda(\mathrm{d}s)}.$$

In other words, the empirical mean $\langle M(t), f \rangle$ is a consistent estimator of $\langle \lambda_1, f \rangle$ rather than of $\langle \lambda, f \rangle$.

We stress that this detection paradox disappears for a version of this model where the intensity of the Poisson point process only grows sub-exponentially in time – say for simplicity it is given by $b^r \mathrm{d}b\lambda(\mathrm{d}\ell)$ for some $r > 0$. The same calculation as above easily shows that the empirical mean of $f(\ell)$ computed for all individuals who are dead by time t then does converge to $\langle \lambda, f \rangle$ as $t \to \infty$.

https://doi.org/10.1017/jpr.2022.112 Published online by Cambridge University Press

5.3. Relation to the Yule–Simon distribution

In 1955, following G. Udny Yule [19], Herbert A. Simon [18] introduced an elementary model depending on a parameter $q \in (0, 1)$ (that accounts for the memory of the model), which today would be referred to as an algorithm with preferential attachment. Simon's algorithm produces a random text, that is, a long string of words $w_1 \dots w_n$, as follows. Once the first word w_1 has been written, for each $j = 1, \dots, n-1$, w_{j+1} is copied from a uniform sample from $w_1, \dots, w_j$ with probability q, and with complementary probability $1-q$, w_{j+1} is a new word different from all the preceding words. Simon proved that for every fixed $k \geq 1$, the expected proportion of different words that have been written exactly k times in the text converges as $n \to \infty$ towards

$$\sigma_q(k) = \frac{1}{q} \mathrm{B}(1 + 1/q, k). \tag{5.8}$$

The probability measure on $\mathbb{N}$, $\sigma_q = (\sigma_q(k), k \geq 1)$, is known as the Yule–Simon distribution with parameter $1/q$. Comparing (4.7) with (5.8), we can now view the average distributions of the sizes of active clusters m^a as an exponentially tilted version of the Yule–Simon distribution with parameter α/ρ.

In this direction, we observe that the limiting case of our model with $\delta = 0$, which corresponds to a degenerate case where detection is absent, merely rephrases Simon's algorithm with memory parameter $q = p$. When there is no detection, the evolution of a typical cluster is just that of a Yule process with rate $p\gamma$ (without killing). Then the intensity measure of birth times of new clusters given by $\mu(\mathrm{d}t) = (1-p)\gamma\, \mathrm{e}^{p\gamma t}\mathrm{d}t$ and the Malthusian parameter can be identified by solving

$$(1-p)\gamma \int_0^\infty \mathrm{e}^{(p\gamma - \alpha)t}\mathrm{d}t = 1.$$

Plainly we have $\alpha = \gamma$ and the parameter of the Yule–Simon distribution is simply $\alpha/\rho = 1/p$. In this setting, the degenerate case of Corollary 4.1 for $\delta = 0$ can be viewed as a strong version of Simon's result, where the convergence is almost sure and not just for the expectation. See [6, Example B.11] for a closely related discussion in the setting of Yule's original model of evolution of species, which is a bit different but nonetheless also yields the Yule–Simon distribution (5.8).

Acknowledgements

I would like to thank Vincent Bansaye for pointing out some similarities to age-structured models and, in particular, the existence of explicit solutions to eigenproblems for the latter. I am also grateful to two anonymous referees for their careful reading of the first version of this work and their constructive comments.

Funding information

The author acknowledges partial support from Swiss National Science Foundation grants 188693.

Competing interests

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

References

[1] BALL, F. G., KNOCK, E. S. AND O'NEILL, P. D. (2011). Threshold behaviour of emerging epidemics featuring contact tracing. *Adv. Appl. Prob.* **43**, 1048–1065.

[2] BANSAYE, V., GU, C. AND YUAN, L. (2022). A growth–fragmentation–isolation process on random recursive trees. Available at arXiv:2109.05760. To appear in *Ann. Appl. Prob*.

[3] BARLOW, M. (2020). A branching process with contact tracings. Available at arXiv:2007.16182.

[4] BRITTON, T. AND PARDOUX, E. (2019). *Stochastic Epidemic Models with Inference* (Lecture Notes Math. **2255**). Springer, Cham.

[5] DEIJFEN, M. (2010). Random networks with preferential growth and vertex death. *J. Appl. Prob.* **47**, 1150–1163.

[6] HOLMGREN, C. AND JANSON, S. (2017). Fringe trees, Crump–Mode–Jagers branching processes and *m*-ary search trees. *Prob. Surv.* **14**, 53–154.

[7] HUO, X. (2015). Modeling of contact tracing in epidemic populations structured by disease age. *Discrete Contin. Dyn. Syst. Ser. B* **20**, 1685–1713.

[8] IKSANOV, A., KOLESKO, K. AND MEINERS, M. (2021). Asymptotic fluctuations in supercritical Crump–Mode–Jagers processes. Available at arXiv:2109.00867.

[9] JAGERS, P. (1975). *Branching Processes with Biological Applications* (Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics). John Wiley, London, New York, Sydney.

[10] JAGERS, P. (1989). General branching processes as Markov fields. *Stoch. Process. Appl.* **32**, 183–212.

[11] JAGERS, P. AND NERMAN, O. (1984). The growth and composition of branching populations. *Adv. Appl. Prob.* **16**, 221–259.

[12] KOT, M. (2003). *Elements of Mathematical Ecology*. Cambridge University Press.

[13] LAMBERT, A. (2021). A mathematical assessment of the efficiency of quarantining and contact tracing in curbing the COVID-19 epidemic. *Math. Model. Nat. Phenom.* **16**, 53.

[14] MÜLLER, J. AND KRETZSCHMAR, M. (2021). Contact tracing: old models and new challenges. *Infect. Disease Model.* **6**, 222–231.

[15] MÜLLER, J., KRETZSCHMAR, M. AND DIETZ, K. (2000). Contact tracing in stochastic and deterministic epidemic models. *Math. Biosci.* **164**, 39–64.

[16] NERMAN, O. (1981). On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrscheinlichkeitsth.* **57**, 365–395.

[17] OKOLIE, A. AND MÜLLER, J. (2020). Exact and approximate formulas for contact tracing on random trees. *Math. Biosci.* **321**, 108320.

[18] SIMON, H. A. (1955). On a class of skew distribution functions. *Biometrika* **42**, 425–440.

[19] YULE, G. U. (1925). A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. *Phil. Trans. R. Soc. London B* **213**, 21–87.

https://doi.org/10.1017/jpr.2022.112 Published online by Cambridge University Press