Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER Stochastic Processes and their Applications 113 (2004) 173 – 197

stochastic processes and their applications

www.elsevier.com/locate/spa

Approximating the Reed–Frost epidemic process[☆]

A.D. Barbour[a,*], Sergey Utev[b]

[a] Abteilung für Angewandte Mathematik, Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

[b] School of Mathematical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, UK

Received 15 October 2002; received in revised form 7 April 2003; accepted 30 March 2004

Available online 3 May 2004

Abstract

The paper is concerned with refining two well-known approximations to the Reed–Frost epidemic process. The first is the branching process approximation in the early stages of the epidemic; we extend its range of validity, and sharpen the estimates of the error incurred. The second is the normal approximation to the distribution of the final size of a large epidemic, which we complement with a detailed local limit approximation. The latter, in particular, is relevant if the approximations are to be used for statistical inference.

© 2004 Elsevier B.V. All rights reserved.

MSC: 92D30; 60K99; 60F05

Keywords: Reed–Frost epidemic process; Local limit approximation; Asymptotic relative closeness; Total variation; Final size distribution; Branching process approximation

1. Introduction

The Reed–Frost epidemic process $\{(S^{(n)}(r), I^{(n)}(r)), r \geqslant 0\}$ is a discrete time S–I–R model, in which the index n denotes the initial number of susceptibles $S^{(n)}(0)$ and the initial number $I^{(n)}(0)$ of infectives is denoted by i_n. The process evolves according to a Markovian recursion: given that $S^{(n)}(r) = s$ and $I^{(n)}(r) = i$, then

$$S^{(n)}(r+1) \sim \mathrm{Bi}(s, (1-q^{(n)})^i); \quad I^{(n)}(r+1) := s - S^{(n)}(r+1) \tag{1.1}$$

[☆] Work supported in part by Schweizer Nationalfonds Projekte Nr. 20-61753.00 and 20-67909.02

[*] Corresponding author. Tel.: +41-1-635-5846; fax: +41-1-635-5705.
E-mail address: adb@amath.unizh.ch (A.D. Barbour).

0304-4149/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.spa.2004.03.013

with $q^{(n)} = n^{-1}\lambda$, for some fixed mean reproduction number $\lambda > 0$. The interpretation of (1.1) is as follows. At any given time r, any pair of the $n + i_n$ individuals may come into contact, with the $\binom{n+i_n}{2}$ possible contact events being realized independently with probability $q^{(n)}$. A susceptible who has been in contact with at least one infective becomes an infective at time $r+1$; infectives at time r become 'removed' (immune and no longer infectious) at time $r+1$. Thus the $S^{(n)}(r+1)$ susceptibles at time $r+1$ are those of the s susceptibles at time r who escape contact with the i infectives present at time r; those that do not escape infection become the $I^{(n)}(r+1)$ infectives at time $r+1$.

The Reed–Frost model is one of the simplest epidemic models, and is used as a template for constructing many more sophisticated variants. Despite this, its structure is sufficiently complicated that more tractable approximations are still needed, if its behaviour is to be understood. In this paper, we are interested in the detail of two simpler approximations to particular aspects of the process: the branching approximation in the early stages, and the central limit theorem for the final size (see [7]). Both are used to provide approximate likelihoods for use in statistical analyses, though this is only justifiable if the true and approximate likelihoods are known to be (sufficiently) close. This seems actually not to have so far been established (see [4] for supporting arguments, and [2] for further applications of the approximate likelihoods or pseudo likelihood methodology); our aim is to do so. In Section 2, we introduce a measure of closeness which is tailored to the likelihood, and show that the epidemic process and a branching process approximation to it are close in this sense until the number of susceptibles has fallen by an amount of order $O(n^\alpha)$, for any $\alpha < \frac{2}{3}$; this is actually rather better than previous approximations in total variation, which were only proved for $\alpha < 1/2$ [1,6]. We also establish that the likelihoods then agree to within a relative error of order $O(n^{-(1-3\alpha/2)}\log^2 n)$, except possibly on a set (identifiable from the data) of very small asymptotic probability. In Section 3, we turn to the final size, proving that the relative error in approximating its point probabilities by a discretized normal distribution is small enough to justify the use of the normal density to approximate the true probabilities in likelihood calculations.

2. Relative closeness in the branching approximation

We begin by defining a concept of closeness designed for statistical applications. Let P and Q be non-negative measures defined on a measurable space $(\mathscr{X}, \mathscr{F})$, and set

$$0 \leqslant \delta(x) := \frac{\mathrm{d}P}{\mathrm{d}Q}(x) \leqslant \infty.$$

We say that P and Q are ε-relatively close with tolerance η, $\mathrm{RC}(\varepsilon, \eta)$ for short, if there exists a set $R \in \mathscr{F}$ such that

$$P(R^{\mathrm{c}}) \leqslant \eta, \quad Q(R^{\mathrm{c}}) \leqslant \eta \quad \text{and} \quad \sup_{x \in R} |\log \delta(x)| \leqslant \varepsilon. \tag{2.1}$$

We think of R as being the set of 'typical' outcomes, R^{c} as being the exceptional set. Similarly, we say that sequences of measures $(P_n, n \geqslant 1)$ and $(Q_n, n \geqslant 1)$ are asymptotically ε_n-relatively close with tolerance η_n, ARC(ε_n, η_n) for short, if P_n and Q_n are RC(ε'_n, η'_n), and $\varepsilon'_n = \mathrm{O}(\varepsilon_n)$ and $\eta'_n = \mathrm{O}(\eta_n)$ as $n \to \infty$, and that sequences $(X^{(n)},\ n \geqslant 1)$ and $(Y^{(n)},\ n \geqslant 1)$ of random elements with values in $\mathscr{X}$ are ARC(ε_n, η_n) if their probability distributions $(\mathscr{L}(X^{(n)}),\ n \geqslant 1)$ and $(\mathscr{L}(Y^{(n)}),\ n \geqslant 1)$ are ARC(ε_n, η_n). Note that if probability measures P and Q are RC(ε, η), then $d_{\mathrm{TV}}(P, Q) \leqslant (\mathrm{e}^{\varepsilon} - 1) + \eta$, whereas, if $d_{\mathrm{TV}}(P, Q) = \varepsilon$, then P and Q are RC($-\log(1 - \varepsilon/\eta), 2\eta$) for any $\eta > \varepsilon$.

In statistical applications, one would typically have families of probability distributions $\{(P_n^{\theta}, Q_n^{\theta});\ \theta \in \Theta\}$ whose elements $X^{(n)}$ and $Y^{(n)}$ were ARC(ε_n, η_n) uniformly in θ. Typically, P_n^{θ} would be the distribution of the actual model generating the data at parameter value θ, and Q_n^{θ} a simpler approximation to it; the closeness of the distributions would then be used to justify a likelihood derived from the approximation Q_n^{θ} being used for inference. To protect against large errors being introduced in this way, one should keep η_n extremely small; in this paper, we shall always arrange to have $\eta_n = \varpi(n)$, where $\varpi(n)$ denotes a generic quantity of order $\mathrm{O}(n^{-r})$ for all $r > 0$.

In this respect, the notion of ARC is rather more flexible than that of total variation distance. By proving that $d_{\mathrm{TV}}(\mathscr{L}(X^{(n)}), \mathscr{L}(Y^{(n)})) = \mathrm{O}(\varepsilon_n)$, it follows that $X^{(n)}$ and $Y^{(n)}$ are ARC($-\log(1 - \varepsilon_n/\eta_n), \eta_n$) for any $\eta_n > \varepsilon_n$, but it is very much more useful to know, as may often also be the case, that this is because, in fact, $X^{(n)}$ and $Y^{(n)}$ are ARC(ε_n, η_n) for very small η_n.

If the random elements $X^{(n)}$ and $Y^{(n)}$ are processes, the notion of ARC(ε_n, η_n) can be extended to include the time interval over which closeness is to be measured. We then say that $X^{(n)}$ and $Y^{(n)}$ are ARC(ε_n, η_n) up to time $T^{(n)}$ if the elements $\pi_{T^{(n)}}(X^{(n)})$ and $\pi_{T^{(n)}}(Y^{(n)})$ are ARC(ε_n, η_n), where, for $x = (x_0, x_1, \ldots)$, we define $\pi_t(x) := (x_0, x_1, \ldots, x_t)$. Stopping times can also be included in analogous fashion. Some further useful properties related to ARC are discussed in Appendix.

Our aim in this section is to show that the population of infectives in the Reed–Frost epidemic process defined in (1.1) and a branching process with Poisson offspring distribution Po(λ) are ARC($\varepsilon_n, \varpi(n)$) up to the time at which the n^{α}th infection (birth) occurs, for any $\alpha < \frac{2}{3}$, where $\varepsilon_n = \mathrm{O}(n^{-(1-3\alpha/2)} \log^2 n)$. These two processes can be expected to be initially close in distribution, since each infective contacts a binomially $\mathrm{Bi}(n + i_n - 1, n^{-1}\lambda) \approx \mathrm{Po}(\lambda)$ distributed number of other individuals before being removed, and these are all infectives in the next generation if they were previously susceptible, a very likely event in the early stages of an outbreak, when almost all individuals are still susceptible. An elementary coupling argument, based on this idea and using birthday problem asymptotics, suggests that the branching process approximation should remain good as long as the number of infectives is of order $\mathrm{o}(n^{1/2})$. Here, we wish to derive an approximation which goes substantially further. To do so, we start by considering some simpler sequences of pairs of processes.

Lemma 2.1. *Let $(Z_{j1}, j \geqslant 0)$ and $(Z_{j2}, j \geqslant 0)$ be sequences of independent random variables, with $Z_{j1} \sim \mathrm{Bi}(m_n, p_n)$ and $Z_{j2} \sim \mathrm{Po}(m_n p_n)$, where $m_n p_n \leqslant \lambda$ and $c_1 n \leqslant m_n \leqslant C_1 n$ for some $c_1 > 0$ and $\lambda, C_1 < \infty$. Define processes $X^{(n)}$ and $Y^{(n)}$ taking values in $\mathbf{Z}$ by*

$$X_{j+1}^{(n)} = X_j^{(n)} - Z_{j1} \quad \text{and} \quad Y_{j+1}^{(n)} = Y_j^{(n)} - Z_{j2}, \quad j \geqslant 0,$$

where $X_0^{(n)} = Y_0^{(n)}$. *Suppose that*

$$T^{(n)} \leqslant C_2 n/\log^4 n \quad \textit{for some } C_2 > 0. \tag{2.2}$$

Then $X^{(n)}$ *and* $Y^{(n)}$ *are* $\mathrm{ARC}(\varepsilon_n, \varpi(n))$ *up to time* $T^{(n)}$, *where* $\varepsilon_n = T^{(n)} n^{-1} \log^4 n$.

Proof. For $k = (k_0, k_1, \ldots, k_{T^{(n)}}) \in \mathbf{Z}^{T^{(n)}+1}$ satisfying $k_0 \geqslant k_1 \geqslant \cdots k_{T^{(n)}} \geqslant 0$, write $r_j = k_{j-1} - k_j$, $1 \leqslant j \leqslant T^{(n)}$, and set $X_0^{(n)} = Y_0^{(n)} = k_0$ a.s. Then

$$\frac{\mathbb{P}[\pi_{T^{(n)}}(X^{(n)}) = k]}{\mathbb{P}[\pi_{T^{(n)}}(Y^{(n)}) = k]} = \exp\left(\sum_{j=1}^{T^{(n)}} \log \psi_j\right),$$

where, with $q_n = 1 - p_n$,

$$\psi_j = \frac{\mathrm{Bi}(m_n, p_n)\{r_j\}}{\mathrm{Po}(m_n p_n)\{r_j\}} = \binom{m_n}{r_j} p_n^{r_j} q_n^{m_n - r_j} \bigg/ \{\mathrm{e}^{-m_n p_n}(p_n m_n)^{r_j}/r_j!\},$$

so that

$$\sum_{j=1}^{T^{(n)}} \log \psi_j = T^{(n)} m_n (p_n + \log q_n) + \sum_{j=1}^{T^{(n)}} \left(\left\{ \sum_{i=1}^{r_j - 1} \log(1 - i/m_n) \right\} - r_j \log q_n \right). \tag{2.3}$$

The first term in (2.3) is of order $\mathrm{O}(T^{(n)} n p_n^2) = \mathrm{O}(T^{(n)} n^{-1})$ as $n \to \infty$. Now define the set

$$R_n = \{k \in \mathbf{Z}_+^{T^{(n)}+1}\colon\ 0 \leqslant k_{i-1} - k_i \leqslant \log^2 n,\ 1 \leqslant i \leqslant T^{(n)}\}.$$

Clearly, because $m_n p_n$ is uniformly bounded, and because of the Chernoff bounds for binomial and Poisson random variables,

$$\mathbb{P}[\pi_{T^{(n)}}(Y^{(n)}) \notin R_n] = \varpi(n); \quad \mathbb{P}[\pi_{T^{(n)}}(X^{(n)}) \notin R_n] = \varpi(n). \tag{2.4}$$

Then, for $k \in R_n$ and $r_j = k_{j-1} - k_j$ as before, the second term in (2.3) is uniformly of order

$$\sum_{j=1}^{T^{(n)}} \{m_n^{-1} r_j^2 + r_j p_n\} = \mathrm{O}(T^{(n)} n^{-1} \log^4 n),$$

completing the proof. □

Lemma 2.2. *For each* $n \geqslant 1$, *define processes* $X^{(n)}$ *and* $Y^{(n)}$ *taking values in* $\mathbf{Z}$ *by*

$$X_{j+1}^{(n)} = X_j^{(n)} - Z_{j1}^{(n)} \quad \textit{and} \quad Y_{j+1}^{(n)} = Y_j^{(n)} - Z_{j2}^{(n)}, \quad j \geqslant 0,$$

where $X_0^{(n)} = Y_0^{(n)}$ *a.s., and* $(Z_{j1}^{(n)}, j \geqslant 0)$ *and* $(Z_{j2}^{(n)}, j \geqslant 0)$ *are sequences of random variables defined as follows. The* $Z_{j2}^{(n)} \sim \mathrm{Bi}(m_n, p_n)$ *are independent random variables, with the sequences* m_n *and* p_n *as for Lemma* 2.1. *Then, for* $j \geqslant 0$, *conditionally on*

$\mathscr{F}_j^{(n)} := \sigma(X_0^{(n)}, Z_{01}^{(n)}, \ldots, Z_{j-1,1}^{(n)})$, *$Z_{j1}^{(n)}$ is sampled independently from the binomial* $\mathrm{Bi}(M_{nj}, p_n)$ *distribution, where M_{nj} is $\mathscr{F}_j^{(n)}$-measurable and satisfies* $0 \leqslant m_n - M_{nj} \leqslant \Delta^{(n)}$ *a.s. for all $j \geqslant 0$. If $T^{(n)}$ and $\Delta^{(n)}$ are such that*

$$T^{(n)} \geqslant (2\lambda \mathrm{e})^{-2} \log^4 n, \quad \varepsilon_n := \sqrt{T^{(n)}}\Delta^{(n)} n^{-1} \log^2 n \to 0 \quad \textit{as } n \to \infty, \tag{2.5}$$

then the processes $X^{(n)}$ and $Y^{(n)}$ are $\mathrm{ARC}(\varepsilon_n, \varpi(n))$ *up to time $T^{(n)}$.*

Proof. If $\Delta^{(n)} = 0$, there is nothing to prove, so we assume henceforth that $\Delta^{(n)} \geqslant 1$. The proof then runs much as in the previous lemma, starting with

$$\frac{\mathbb{P}[\pi_{T^{(n)}}(X^{(n)}) = k]}{\mathbb{P}[\pi_{T^{(n)}}(Y^{(n)}) = k]}$$

$$= \prod_{j=1}^{T^{(n)}} \frac{\mathbb{P}[X_j^{(n)} = k_j | \pi_{j-1}(X^{(n)}) = (k_0, \ldots, k_{j-1})]}{\mathbb{P}[Y_j^{(n)} = k_j | \pi_{j-1}(Y^{(n)}) = (k_0, \ldots, k_{j-1})]} = \exp\left(\sum_{j=1}^{T^{(n)}} \log \psi_j\right),$$

where now

$$\psi_j = \frac{\mathrm{Bi}(m_{nj}, p_n)\{r_j\}}{\mathrm{Bi}(m_n, p_n)\{r_j\}} = \binom{m_{nj}}{r_j} p_n^{r_j} q_n^{m_{nj}-r_j} \Big/ \binom{m_n}{r_j} p_n^{r_j} q_n^{m_n - r_j}$$

$$= q^{m_{nj}-m_n} \binom{m_{nj}}{r_j} \Big/ \binom{m_n}{r_j},$$

and $m_{nj} := M_{n,j-1}(k_0, \ldots, k_{j-1})$. Hence

$$\sum_{j=1}^{T^{(n)}} \log \psi_j = \sum_{j=1}^{T^{(n)}} \Bigg\{ (m_{nj} - m_n)(p_n + \log q_n)$$
$$+ \left(p_n(m_n - m_{nj}) + \sum_{i=0}^{r_j - 1} \log\{1 - (m_n - m_{nj})/(m_n - i)\} \right) \Bigg\}. \tag{2.6}$$

The first term in (2.6) is immediately of order

$$\mathrm{O}(T^{(n)} \Delta^{(n)} p_n^2) = \mathrm{O}(\varepsilon_n^2 / \{\Delta^{(n)} \log^4 n\}) = \mathrm{o}(\varepsilon_n).$$

For the second term, define the set

$$R_n = \Bigg\{ (k_0, \ldots, k_{T^{(n)}}) \colon 0 \leqslant k_{i-1} - k_i \leqslant \log^2 n, \ 1 \leqslant i \leqslant T^{(n)};$$
$$\left| \sum_{j=1}^{T^{(n)}} (k_{j-1} - k_j - m_n p_n)(m_n - m_{nj}) \right| \leqslant \Delta^{(n)} \{\log n\}^2 \sqrt{T^{(n)}} \Bigg\}.$$

The random variable

$$S_n := \sum_{j=1}^{T^{(n)}} (Z^{(n)}_{j-1,1} - M_{n,j-1}p_n)(m_n - M_{n,j-1})$$

is a martingale; the weights $m_n - M_{n,j-1}$ satisfy $0 \leqslant m_n - M_{n,j-1} \leqslant \Delta^{(n)}$ a.s., and the random factors $Z^{(n)}_{j-1,1} - M_{n,j-1}p_n$ have centred binomial distributions, conditional on $X^{(n)}_0, \dots, X^{(n)}_{j-1}$. Hence, using $m_n p_n \leqslant \lambda$, the moment generating function $\mathbb{E}(e^{\theta S_n})$ of S_n is simply bounded in $\theta \leqslant 1/\Delta^{(n)}$ by

$$\mathbb{E}(e^{\theta S_n}) \leqslant \exp\{\tfrac{1}{2}\lambda e \theta^2 (\Delta^{(n)})^2 T^{(n)}\}.$$

Provided that $\log^2 n \leqslant 2\lambda e \sqrt{T^{(n)}}$, we can thus choose $\theta := \pm\{\log n\}^2/\{2\lambda e \Delta^{(n)}\sqrt{T^{(n)}}\}$ to show that

$$\mathbb{P}[|S_n| \geqslant \tfrac{1}{2}\Delta^{(n)}\{\log n\}^2\sqrt{T^{(n)}}] \leqslant 2\exp\{-\{\log n\}^4/8\lambda e\} = \varpi(n).$$

Then we also have

$$\sum_{j=1}^{T^{(n)}} (m_n - M_{n,j-1})^2 p_n = O(T^{(n)}(\Delta^{(n)})^2 n^{-1}) = O(\Delta^{(n)}\{\log n\}^2\sqrt{T^{(n)}}\{\varepsilon_n/\log^4 n\})$$

$$\leqslant \tfrac{1}{2}\Delta^{(n)}\{\log n\}^2\sqrt{T^{(n)}}$$

for all n large enough. These considerations, together with the Chernoff bounds on the tails of the binomial distribution, show that $\mathbb{P}[\pi_{T^{(n)}}(X^{(n)}) \notin R_n] = \varpi(n)$; the argument for $Y^{(n)}$ is a little easier, because M_{nj} is replaced by m_n.

Now, returning to the second term in (2.6), for $k \in R_n$ and $1 \leqslant j \leqslant T^{(n)}$, we have

$$p_n(m_n - m_{nj}) + \sum_{i=0}^{r_j-1} \log\{1 - (m_n - m_{nj})/(m_n - i)\}$$

$$= m_n^{-1}(m_n - m_{nj})(m_n p_n - r_j) + \eta_{nj},$$

where

$$\eta_{nj} := (m_n - m_{nj})\sum_{i=0}^{r_j-1}\left(\frac{1}{m_n} - \frac{1}{m_n - i}\right)$$

$$+ \sum_{i=0}^{r_j-1}\left(\log\left\{1 - \frac{m_n - m_{nj}}{m_n - i}\right\} + \frac{m_n - m_{nj}}{m_n - i}\right)$$

$$= O(n^{-2}r_j(m_n - m_{nj})\{(m_n - m_{nj}) + r_j\})$$

$$= O(n^{-2}(\log n)^2\Delta^{(n)}(\Delta^{(n)} + \log^2 n)),$$

uniformly for $k \in R_n$, so that

$$\sum_{j=1}^{T^{(n)}} |\eta_{nj}| = O(T^{(n)}n^{-2}(\Delta^{(n)})^2\log^4 n) = O(\varepsilon_n^2) = o(\varepsilon_n),$$

again uniformly for $k \in R_n$. But then, for $k \in R_n$,

$$\left| \sum_{j=1}^{T^{(n)}} m_n^{-1}(m_n - m_{nj})(m_n p_n - r_j) \right| \leqslant m_n^{-1} \Delta^{(n)} \{\log n\}^2 \sqrt{T^{(n)}} = \mathrm{O}(\varepsilon_n);$$

hence it follows that

$$\sum_{j=1}^{T^{(n)}} \log \psi_j = \mathrm{O}(\varepsilon_n),$$

uniformly for $k \in R_n$. This completes the proof of the lemma. □

Lemma 2.3. *Define processes $X^{(n)}$ and $Y^{(n)}$ taking values in $\mathbf{Z}$ by*

$$X_{j+1}^{(n)} = X_j^{(n)} - Z_{j1}^{(n)} \quad \text{and} \quad Y_{j+1}^{(n)} = Y_j^{(n)} - Z_{j2}^{(n)}, \quad j \geqslant 0,$$

where, with m_n, p_n as before, $X_0^{(n)} = Y_0^{(n)} = m_n$ and the $Z_{j2}^{(n)} \sim \mathrm{Bi}(m_n, p_n)$ are independent; and now, conditionally on $\mathscr{F}_j^{(n)}$, $Z_{j1}^{(n)}$ is sampled independently from $\mathrm{Bi}(X_j^{(n)}, p_n)$. If

$$T^{(n)} \geqslant (2\lambda \mathrm{e})^{-2} \log^4 n; \quad \varepsilon_n := \{T^{(n)}\}^{3/2} \{\log n\}^2 n^{-1} \to 0 \quad \text{as } n \to \infty, \tag{2.7}$$

then it follows that $X^{(n)}$ and $Y^{(n)}$ are $\mathrm{ARC}(\varepsilon_n, \varpi(n))$ up to time $T^{(n)}$.

Proof. The lemma follows from Lemma 2.2. Define a further process $X'^{(n)}$ by setting $X_0'^{(n)} = X_0^{(n)}$ and $X_{j+1}'^{(n)} = X_j'^{(n)} - Z_{j1}'^{(n)}$ for $j \geqslant 0$, where, given $(X_i'^{(n)}, 0 \leqslant i \leqslant j)$, $Z_{j1}'^{(n)}$ is sampled from the binomial distribution $\mathrm{Bi}(\max\{X_j'^{(n)}, m_n - \Delta^{(n)}\}, p_n)$, with $\Delta^{(n)} = 2\lambda T^{(n)}$. Since $Y^{(n)}$ is stochastically smaller than $X^{(n)}$ and $\sum_{j=1}^{T^{(n)}} Z_{j-1,2}^{(n)} \sim \mathrm{Bi}(m_n T^{(n)}, p_n)$, it follows that

$$\mathbb{P}[m_n - X_{T^{(n)}}^{(n)} > \Delta^{(n)}] = \varpi(n),$$

and hence, defining

$$R_n = \{(k_0, \ldots, k_{T^{(n)}}):\ 0 \leqslant k_{i-1} - k_i \leqslant \log^2 n,\ 1 \leqslant i \leqslant T^{(n)};\ k_0 - k_{T^{(n)}} \leqslant \Delta^{(n)}\},$$

that $X^{(n)}$ and $X'^{(n)}$ are $\mathrm{ARC}(0, \varpi(n))$ up to time $T^{(n)}$. Now apply Lemma 2.2 to the processes $X'^{(n)}$ and $Y^{(n)}$, recalling that $\Delta^{(n)} = 2\lambda T^{(n)}$. □

Lemma 2.4. *For each $n \geqslant 1$, let $(Z_{j1}^{(n)}, j \geqslant 0)$ and $(Z_{j2}^{(n)}, j \geqslant 0)$ be sequences of independent random variables, with $Z_{j1}^{(n)} \sim \mathrm{Po}(\lambda_n')$ and $Z_{j2}^{(n)} \sim \mathrm{Po}(\lambda_n)$, where $0 \leqslant \lambda_n - \lambda_n' \leqslant n^{-1}\Delta^{(n)}$ and where $\lambda_n \asymp 1$ as $n \to \infty$. Define processes $X^{(n)}$ and $Y^{(n)}$ taking values in $\mathbf{Z}$ by*

$$X_{j+1}^{(n)} = X_j^{(n)} - Z_{j1}^{(n)} \quad \text{and} \quad Y_{j+1}^{(n)} = Y_j^{(n)} - Z_{j2}^{(n)}, \quad j \geqslant 0,$$

where $X_0^{(n)} = Y_0^{(n)}$. Then if

$$\varepsilon_n := n^{-1} \Delta^{(n)} \sqrt{T^{(n)}} \log^2 n \to 0 \quad \text{as } n \to \infty,$$

the processes $X^{(n)}$ and $Y^{(n)}$ are $\mathrm{ARC}(\varepsilon_n, \varpi(n))$ up to time $T^{(n)}$.

Proof. We start with

$$\psi_j = \frac{\mathrm{Po}(\lambda_n')\{r_j\}}{\mathrm{Po}(\lambda_n)\{r_j\}} = \{\mathrm{e}^{-\lambda_n'}(\lambda_n')^{r_j}/r_j!\}/\{\mathrm{e}^{-\lambda_n}(\lambda_n)^{r_j}/r_j!\},$$

yielding

$$\begin{aligned}\sum_{j=1}^{T^{(n)}} \log\psi_j &= T^{(n)}(\lambda_n - \lambda_n') + \sum_{j=1}^{T^{(n)}} r_j[\log(\lambda_n') - \log(\lambda_n)] \\ &= T^{(n)}\{(\lambda_n - \lambda_n') + \lambda_n[\log(\lambda_n') - \log(\lambda_n)]\} \\ &\quad + \sum_{j=1}^{T^{(n)}} (r_j - \lambda_n)[\log(\lambda_n') - \log(\lambda_n)]. \end{aligned} \tag{2.8}$$

As before, the first term in (2.8) is of order $\mathrm{O}(T^{(n)}\{n^{-1}\varDelta^{(n)}\}^2) = \mathrm{o}(\varepsilon_n^2)$ as $n \to \infty$. Now define the set

$$R_n = \left\{ (k_0, \ldots, k_{T^{(n)}}):\ \left| \sum_{i=1}^{T^{(n)}} (k_{i-1} - k_i - \lambda_n) \right| \leqslant \log^2 n\sqrt{T^{(n)}} \right\}.$$

Note that

$$T^{(n)}|\lambda_n - \lambda_n'| = \mathrm{O}(T^{(n)}n^{-1}\varDelta^{(n)}) = \mathrm{O}(\log^2 n\sqrt{T^{(n)}}\{\varepsilon_n/\log^4 n\}) = \mathrm{o}(\log^2 n\sqrt{T^{(n)}}),$$

and that

$$\mathbb{P}\left[\left| \sum_{j=1}^{T^{(n)}} (Z_{j-1,1}^{(n)} - \lambda_n') \right| > \frac{1}{2}\log^2 n\sqrt{T^{(n)}} \right] = \varpi(n),$$

in view of the Chernoff bounds for Poisson random variables. Hence it follows that $\mathbb{P}[\pi_{T^{(n)}}(X^{(n)}) \notin R_n] = \varpi(n)$; the corresponding argument for $Y^{(n)}$ is easier. Then, for $k \in R_n$, the second term in (2.8) is of order

$$\begin{aligned}\mathrm{O}\left(|\log(\lambda_n') - \log(\lambda_n)| \left| \sum_{j=1}^{T^{(n)}} (k_{i-1} - k_i - \lambda_n) \right| \right) &= \mathrm{O}(n^{-1}\varDelta^{(n)}\log^2 n\sqrt{T^{(n)}}) \\ &= \mathrm{O}(\varepsilon_n),\end{aligned}$$

uniformly for $k \in R_n$, completing the proof. □

The Reed–Frost epidemic process can be formulated as a sequence of processes $\widehat{X}^{(n)}$, defined as is $X^{(n)}$ in Lemma 2.3, having $\widehat{X}_0^{(n)} = n$ and with $\mathrm{Bi}(\widehat{X}_j^{(n)}, n^{-1}\lambda)$ innovations, $j \geqslant 0$; time $j \geqslant 0$ is to be interpreted as the number of removals. The numbers of susceptibles and infectives $(S^{(n)}(r), I^{(n)}(r))$ in the epidemic process (1.1) at genuine time steps $r = 0, 1, \ldots$ are then derived from the initial number $i_n = I^{(n)}(0)$ of infectives and from values of the $\widehat{X}^{(n)}$-process at (random) times $J_1, J_2, \ldots$, where

$J_{r+1} := n + i_n - \widehat{X}^{(n)}_{J_r}$ and $J_0 := 0$: the recursion is given by

$$S^{(n)}(r) = \widehat{X}^{(n)}_{J_r} \quad \text{and} \quad I^{(n)}(r) = J_{r+1} - J_r, \quad r \geqslant 0, \tag{2.9}$$

see [7, relation (29)]. The branching process approximation $(\tilde{S}^{(n)}, \tilde{I}^{(n)})$ with which it is to be compared can be constructed in analogous fashion, starting from a process $\tilde{Y}^{(n)}$ having $\tilde{Y}^{(n)}_0 = n$ and $\tilde{Y}^{(n)}_{j+1} = \tilde{Y}^{(n)}_j - \tilde{Z}_j$, where the innovations $(\tilde{Z}_j,\ j \geqslant 0)$ are independent Poisson Po(λ)-distributed random variables, and using the recursion

$$\tilde{S}^{(n)}(r) = \tilde{Y}^{(n)}_{\tilde{J}_r} \quad \text{and} \quad \tilde{I}^{(n)}(r) = \tilde{J}_{r+1} - \tilde{J}_r, \quad r \geqslant 0,$$

where $\tilde{J}_{r+1} := n + i_n - \tilde{Y}^{(n)}_{\tilde{J}_r}$ and $\tilde{J}_0 = 0$. The branching process is the second component $\tilde{I}^{(n)}$, and, if i_n is the same for all n, so is the distribution of $\tilde{I}^{(n)}$.

By comparing the innovations, the approximation can only be expected to be reasonable as long as $\widehat{X}^{(n)}_j \approx n$. It has previously been justified until the time r when $n - S^{(n)}(r)$ first exceeds n^{α} in [1,6], for $\alpha < 1/2$. Here, we extend the range of approximation to allow any $\alpha < \frac{2}{3}$, and indeed up to times of order $o(\{n/\log^2 n\}^{2/3})$, and give bounds on the accuracy of the approximation. To do so, fixing any positive sequence $t^{(n)}$, define

$$U^{(n)} := \min\{r \geqslant 0\colon\ n + i_n - S^{(n)}(r) \geqslant t^{(n)} \text{ or } I^{(n)}(r) = 0\},$$

and set $S^{(n)}_*(r) := S^{(n)}(r \wedge U^{(n)})$; define $\tilde{U}^{(n)}$ and $\tilde{S}^{(n)}_*$ analogously. Then the following theorem makes the approximation precise.

Theorem 2.5. *Assume that $t^{(n)}$ satisfies (2.7) and define $\varepsilon_n := \{t^{(n)}\}^{3/2}\{\log n\}^2 n^{-1} \to 0$. Then $S^{(n)}_*$ and $\tilde{S}^{(n)}_*$ are* ARC$(\varepsilon_n, \varpi(n))$.

Remark. It follows from Lemma 2.4 that Theorem 2.5 is also true if $\lambda = \lambda_n \to \lambda_0$ depends on n in such a way that $|\lambda_n - \lambda_0| \leqslant n^{-1}\Delta^{(n)}$, but now with

$$\varepsilon_n = \{t^{(n)}\}^{1/2}(\Delta^{(n)} + t^{(n)})\{\log n\}^2 n^{-1},$$

provided that $\varepsilon_n \to 0$.

Proof. Note that, if $i_n \geqslant t^{(n)}$, there is nothing to prove. Otherwise, we start by observing that

$$\mathbb{P}[J_{U^{(n)}} \geqslant 2\lambda t^{(n)}] = \varpi(n); \quad \mathbb{P}[\tilde{J}_{\tilde{U}^{(n)}} \geqslant 2\lambda t^{(n)}] = \varpi(n). \tag{2.10}$$

To prove the first of these, note that, for each r, $J_{r+1} - J_r$ is stochastically smaller than a sum of $J_r - J_{r-1}$ independent random variables with distributions Bi$(n, n^{-1}\lambda)$. Hence, from the Chernoff bounds,

$$\mathbb{P}[J_{r+1} \geqslant 2\lambda t^{(n)} \,|\, J_{r-1} < J_r \leqslant t^{(n)}] = \varpi(n),$$

and the probability is zero if $J_{r-1} = J_r$. Since the event $\{J_{r-1} < J_r \leqslant t^{(n)}\}$ can occur at most $t^{(n)}$ times, the first claim in (2.10) follows. The second is proved similarly.

Since $T^{(n)} = 2\lambda t^{(n)}$ satisfies (2.2) and (2.7), Lemmas 2.1 and 2.3 can be combined to show that the processes $\widehat{X}^{(n)}$ and $\tilde{Y}^{(n)}$ defined above are ARC$(\varepsilon_n, \varpi(n))$ up to time

$t^{(n)}$: Lemma 2.1 shows that $\tilde{Y}^{(n)}$ and a process $\bar{X}^{(n)}$ with $\mathrm{Bi}(n, n^{-1}\lambda)$ innovations are $\mathrm{ARC}(\varepsilon_n, \varpi(n))$ up to time $t^{(n)}$, and Lemma 2.3 that $\bar{X}^{(n)}$ and $\widehat{X}^{(n)}$ are $\mathrm{ARC}(\varepsilon_n, \varpi(n))$ up to time $t^{(n)}$. The proof is completed easily using (2.10) by expanding the set R_n if necessary, and applying Corollary A.2. □

Corollary 2.6. *Suppose that* $t^{(n)} = Kn^{\alpha}\log^b n$, *where* $0 \leqslant \alpha \leqslant \frac{2}{3}$ *and we assume that* $b \geqslant 4$ *when* $\alpha = 0$ *and* $b < -4/3$ *when* $\alpha = \frac{2}{3}$. *Then* $S_*^{(n)}$ *and* $\tilde{S}_*^{(n)}$ *are* $\mathrm{ARC}(\varepsilon_n, \varpi(n))$ *with*

$$\varepsilon_n = n^{-(1-3\alpha/2)}\log^{2+3b/2} n.$$

Remark. The stopping time $U^{(n)}$ is actually the time when the total of infectives and removals first exceeds $t^{(n)}$, if $I^{(n)}$ does not previously reach 0, being the time at which $n - S^{(n)}(r)$ first exceeds $t^{(n)} - i_n$. Hence Corollary 2.6 implies a similar result for the processes until the number of removals $n - S^{(n)}(r)$ first exceeds $Kn^{\alpha}\log^b n$, as long as $i_n = \mathrm{O}(n^{\alpha}\log^b n)$.

Corollary 2.6 shows that the Reed–Frost epidemic remains close in distribution to a branching process with Poisson offspring distribution until either the epidemic terminates or the number of removals exceeds a level of magnitude $\mathrm{O}(n^{\alpha}\log^b n)$. A subcritical epidemic process, with $\lambda < 1$, typically dies out very soon, so that the corollary is enough to show that the two processes remain close in distribution for all time, and hence, in particular, that their final size distributions are close. We give a brief sketch of the reasoning; the detailed argument parallels the corresponding part of the proof of Theorem 3.3.

The argument is based on the observation that, with the above construction of the Reed–Frost epidemic, $I_n(r) = 0$ when $n + i_n - \widehat{X}_{J_r}^{(n)} = J_r$, and that the total size of the epidemic is then just $n + i_n - S_n(r) = J_r$. In terms of the underlying process $\widehat{X}^{(n)}$, this translates into the statement that the total size of the epidemic is equal to the value of the stopping time

$$\tau^{(n)} := \min\{j\colon\ n + i_n - \widehat{X}_j^{(n)} \leqslant j\}. \tag{2.11}$$

Now $\widehat{X}_j^{(n)} \sim \mathrm{Bi}(n, \{1 - \lambda/n\}^j)$, so that $n - \widehat{X}_j^{(n)}$ has mean close to λj for $j \ll n$, and the binomial Chernoff bounds then show that $\mathbb{P}[\tau^{(n)} > T^{(n)}] = \varpi(n)$, where

$$T^{(n)} := \max(2i_n(1-\lambda)^{-1}, \log^4 n).$$

This fact, in combination with Corollaries 2.6 and A.2, leads to the following corollary, improving related results in [6].

Corollary 2.7. *Let* $Y_\infty^{(n)}$ *denote the total progeny in the Galton–Watson process with offspring distribution* $\mathrm{Po}(\lambda)$, *starting with* i_n *individuals, and let* $R_\infty^{(n)}$ *denote the total size of the Reed–Frost epidemic, starting with n susceptibles and* i_n *infectives, and having mean reproduction number* λ. *Then if* $\lambda < 1$ *and* $i_n = \mathrm{O}(n^{\alpha}\log^b n)$, *and if* $\{i_n\}^{3/2}\{\log n\}^2 n^{-1} \to 0$, *then* $\mathscr{L}(R_\infty^{(n)})$ *and* $\mathscr{L}(Y_\infty^{(n)})$ *are* $\mathrm{ARC}(\varepsilon_n, \varpi(n))$, *with* $\varepsilon_n := \{i_n \vee \log^4 n\}^{3/2}\{\log n\}^2 n^{-1}$.

3. Relative closeness in the normal approximation

Over longer time intervals than those considered in the previous section, it is unrealistic to hope to approximate the Reed–Frost epidemic in the sense of ARC by any simpler process. However, for summaries of the trajectory, this may still be possible, and in particular for the final size $n-S^{(n)}(\infty)$ of a large epidemic, a statistic frequently used in practice. Here, we show that the distribution of the final size and a discretized normal distribution are ARC, a result which can then be applied to justify the use of the normal approximation in likelihood inference based on the final size.

The basis for proving this is Theorem 3.2, which shows that the distribution of a random sum of independent and identically distributed integer valued random variables and an appropriately matched discretized normal distribution are ARC. The theorem requires the distributions of the random variables involved to have exponential moments, a condition which presents no problems in the Reed–Frost context. However, for more general epidemic models, such a condition could cause difficulties. We therefore begin with a somewhat simpler result, establishing a local limit theorem for the random sum under much weaker conditions.

Theorem 3.1. *For each $n \geqslant 1$, let $(\xi_j^{(n)}, j \geqslant 1)$ be independent, identically distributed integer valued random variables with $\mathbb{E}\xi_1^{(n)} = m_1^{(n)}$ and $\mathbb{E}\{\xi_1^{(n)}\}^2 = m_2^{(n)} < \infty$, and such that $\xi_1^{(n)} \to_{L_2} \xi_1$, where $\mathbb{E}\xi_1 = m_1$, $\mathbb{E}\xi_1^2 = m_2$ and ξ_1 has lattice span 1. Let $(S_n, n \geqslant 1)$ be non-negative random variables, independent of the sequence $(\xi_j^{(n)}, j \geqslant 1)$, with Laplace-transform $\varphi_n(\psi) := \mathbb{E}\{e^{-\psi S_n}\}$. Suppose that, for sequences a_n and b_n satisfying*

$$s_n^2 := a_n\{m_2^{(n)} - (m_1^{(n)})^2\} + b_n(m_1^{(n)})^2 \sim ns^2 \quad \text{as } n \to \infty \tag{3.1}$$

for some $s > 0$, and for a sequence of functions ε_n satisfying

$$\lim_{\psi \to 0; \Re\psi \geqslant 0} \sup_n |\varepsilon_n(\psi)| = 0,$$

the quantities

$$\delta_{n1}(c) := \sup_{\psi: 0 \leqslant \Re\psi \leqslant c\{n^{-1}\log n\}^{1/2}} \left| \varphi_n(\psi) - \exp\left\{ -\psi a_n + \frac{1}{2}\psi^2 b_n + n\psi^2 \varepsilon_n(\psi) \right\} \right|,$$

$$\delta_{n2}(c) := \sup_{\psi: \Re\psi \geqslant c\{n^{-1}\log n\}^{1/2}} |\varphi_n(\psi)| \tag{3.2}$$

are such that $\delta_{ni}(c) = o(n^{-1/2})$ for any $c > 0$ and for $i = 1, 2$. Then $Z_n := \sum_{j=1}^{S_n} \xi_j^{(n)}$ satisfies a local limit theorem:

$$\lim_{n \to \infty} \sup_k \left| \sqrt{n}\mathbb{P}[Z_n = k] - \frac{1}{s}\phi\left(\frac{k - a_n m_1^{(n)}}{s\sqrt{n}} \right) \right| = 0;$$

where $\phi(x) := \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ denotes the standard normal density.

Remark. The constants a_n and b_n are to be expected to be close to $\mathbb{E}S_n$ and $\operatorname{Var} S_n$, respectively. No direct assumption has been made as to their relative magnitudes, though it follows from (3.1) that $a_n = \mathrm{O}(n)$, and that $b_n = \mathrm{O}(n)$ also if $m_1 \neq 0$. However, a local limit theorem cannot hold if $S_n = 0$ with too large a probability; hence one would expect to see a condition of the form $a_n \gg \sqrt{b_n}$ in the theorem as well, suggesting that the mean of S_n is many standard deviations away from zero. This is actually implied here by the condition $\delta_{n2}(c) = \mathrm{o}(n^{-1/2})$; in Theorem 3.2, a more explicit assumption is made.

Proof. Let $w^{(n)}(t) := \mathbb{E}\{\mathrm{e}^{\mathrm{i}t\xi_1^{(n)}}\}$, $t \in \mathbb{R}$, be the characteristic function of the random variables $\xi_j^{(n)}$. Since $\xi_1^{(n)} \to_{L_2} \xi_1$, it follows that

$$w^{(n)}(t) = 1 + \mathrm{i}tm_1^{(n)} - \tfrac{1}{2}\, t^2 m_2^{(n)} + t^2 \eta^{(n)}(t),$$

where $\lim_{t\to 0} \sup_n |\eta^{(n)}(t)| = 0$, and also, because ξ_1 has lattice span 1, that

$$\sup_n |w^{(n)}(t)| \leqslant \mathrm{e}^{-ct^2}, \qquad |t| \leqslant \pi, \tag{3.3}$$

for some $c > 0$ and all n large enough. Then

$$\mathbb{E}\{\mathrm{e}^{\mathrm{i}tZ_n}\} = \mathbb{E}\{w^{(n)}(t)^{S_n}\} = \mathbb{E}\{\mathrm{e}^{S_n \log w^{(n)}(t)}\},$$

and $\Re \log w^{(n)}(t) \leqslant 0$ for $|t| \leqslant \pi$ because $|w^{(n)}(t)| \leqslant 1$. Thus, taking $\psi = -\log w^{(n)}(t)$ in (3.2), we have

$$\begin{aligned} &|\mathbb{E}\{\mathrm{e}^{\mathrm{i}tZ_n}\} - \exp\{a_n \log w^{(n)}(t) + \tfrac{1}{2}\, b_n \{\log w^{(n)}(t)\}^2 \\ &\quad + n\{\log w^{(n)}(t)\}^2 \varepsilon_n(-\log w^{(n)}(t))\}| \\ &\qquad \leqslant \delta_{n1}(c_1) \end{aligned}$$

uniformly in $0 \leqslant -\Re \log w^{(n)}(t) \leqslant c\{n^{-1}\log n\}^{1/2}$, where also

$$\log w^{(n)}(t) = \mathrm{i}tm_1^{(n)} - \tfrac{1}{2}\, t^2 m_2^{(n)} - \tfrac{1}{2}(\mathrm{i}tm_1^{(n)})^2 + t^2 \hat{\eta}^{(n)}(t)$$

with $\hat{\eta}^{(n)}$ satisfying $\lim_{t\to 0} \sup_n |\hat{\eta}^{(n)}(t)| = 0$. Hence it follows that

$$|\mathbb{E}\{\mathrm{e}^{\mathrm{i}tZ_n}\} - \exp\{\mathrm{i}ta_n m_1^{(n)} - \tfrac{1}{2}\, s_n^2 t^2 + (a_n + n)t^2 \hat{\varepsilon}_n(t)\}| \leqslant \delta_{n1}(c_1) \tag{3.4}$$

uniformly in $0 \leqslant |t| \leqslant \sqrt{c_2}\{n^{-1}\log n\}^{1/4}$ and for all sufficiently large n, where $\hat{\varepsilon}_n(t)$ defined by

$$\begin{aligned} (a_n + n)\hat{\varepsilon}_n(t) &:= n\{t^{-1} \log w^{(n)}(t)\}^2 \varepsilon_n(-\log w^{(n)}(t)) + a_n \hat{\eta}^{(n)}(t) \\ &\quad + \tfrac{1}{2}\, b_n(\{t^{-1}\log w^{(n)}(t)\}^2 + \{m_1^{(n)}\}^2) \end{aligned} \tag{3.5}$$

satisfies $\lim_{t\to 0} \sup_n |\hat{\varepsilon}_n(t)| = 0$, and $c_2 := 4c_1/\{m_2 - m_1^2\}$. Thus the characteristic function of Z_n is close to that of the normal; the remaining argument consists of showing that the approximation is good enough to prove a local limit theorem.

First, from the definition of $\hat{\varepsilon}_n$, we can find $k_0, t_0 > 0$ such that

$$\inf_{|t| \leqslant t_0} \left\{ \frac{1}{2}\, s_n^2 - (a_n + n)|\hat{\varepsilon}_n(t)| \right\} \geqslant k_0 n, \tag{3.6}$$

hence, from (3.4), it follows that

$$|\mathbb{E}(e^{itZ_n})| \leqslant \delta_{n1}(c_1) + \varpi(n) \quad \text{uniformly in } t_1(n) \leqslant |t| \leqslant t_2(n) \tag{3.7}$$

with $t_1(n) := n^{-1/2}\log n$ and $t_2(n) := \min\{t_0, \sqrt{c_2}\{n^{-1}\log n\}^{1/4}\}$. For $t_2(n) < |t| \leqslant \pi$, from (3.3),

$$|\mathbb{E}(e^{itZ_n})| \leqslant \mathbb{E}\{|w^{(n)}(t)|^{S_n}\} \leqslant \mathbb{E}(e^{-ct^2 S_n}) = \varphi_n(ct^2) \leqslant \delta_{n2}(cc_2), \tag{3.8}$$

for all n sufficiently large. Combining (3.7) and (3.8), and writing $\delta_n := \delta_{n1}(c_1) + \delta_{n2}(cc_2)$, it thus follows that

$$\begin{aligned}
&\sqrt{n}\left|\mathbb{P}[Z_n = k] - s_n^{-1}\phi\left(\frac{k - a_n m_1^{(n)}}{s_n}\right)\right| \\
&\quad\leqslant \left|\frac{\sqrt{n}}{2\pi}\int_{-\pi}^{\pi} e^{-itk}\left(\mathbb{E}\{e^{itZ_n}\} - \exp\left\{ita_n m_1^{(n)} - \frac{1}{2}s_n^2 t^2\right\}\right) dt\right| \\
&\qquad + \frac{\sqrt{n}}{2\pi}\left|\int_{|t|>\pi} e^{-s_n^2 t^2/2}\, dt\right| \\
&\quad\leqslant \frac{\sqrt{n}}{2\pi}\int_{-\pi}^{\pi}\left|\left(\mathbb{E}\{e^{itZ_n}\} - \exp\left\{ita_n m_1^{(n)} - \frac{1}{2}s_n^2 t^2\right\}\right)\right| dt + \varpi(n) \\
&\quad\leqslant \frac{\sqrt{n}}{2\pi}\int_{-t_1(n)}^{t_1(n)} \exp\left\{-\frac{1}{2}s_n^2 t^2\right\}|e^{(a_n+n)t^2\hat{\varepsilon}_n(t)} - 1|\, dt + n^{1/2}\delta_n + \varpi(n) \\
&\quad= O\left(n^{1/2}\delta_n + n^{3/2}\sup_{|t|\leqslant t_1(n)}|\hat{\varepsilon}_n(t)|\int_0^\infty t^2 e^{-k_0 t^2 n}\, dt\right) + \varpi(n) \\
&\quad= O\left(n^{1/2}\delta_n + \sup_{|t|\leqslant t_1(n)}|\hat{\varepsilon}_n(t)|\right) + \varpi(n) = o(1),
\end{aligned} \tag{3.9}$$

uniformly in k, from (3.6) and because $\lim_{t\to 0}\sup_n|\hat{\varepsilon}_n(t)|=0$. Finally, since $n^{-1}s_n^2 \to s^2$, it follows that

$$\sup_x |s_n^{-1}\sqrt{n}\phi(xs_n^{-1}) - s^{-1}\phi(xs^{-1}n^{-1/2})| = O(|n^{-1}s_n^2 - s^2|) = o(1), \tag{3.10}$$

uniformly in $x \in \mathbb{R}$, completing the proof. □

The local limit theorem is a purely asymptotic measure of the closeness of the density of Z_n to a discretized normal density, implying only that the two measures are ARC(ε_n, η_n) for some unspecified $\varepsilon_n, \eta_n \to 0$. We now strengthen the hypotheses of Theorem 3.1, and improve the conclusion to ARC closeness with explicit convergence rate.

Theorem 3.2. *The setting is as for Theorem* 3.1, *but with the definition of* $\delta_{n1}(c)$ *modified to*

$$\delta_{n1}(c) := \sup_{\psi: |\Re\psi| \leqslant c\{n^{-1}\log n\}^{1/2}} \left| \varphi_n(\psi) - \exp\left\{ -\psi a_n + \frac{1}{2}\psi^2 b_n + n\psi^2 \varepsilon_n(\psi) \right\} \right|,$$

with the function ε_n *analytic in this strip. Assume in addition that*

$$a_n \gg \{n \log n\}^{1/2} \quad \text{and} \quad b_n = \mathrm{O}(n),$$

that, for any $c > 0$, $\delta_{n1}(c) = \mathrm{O}(\mathrm{e}^{-\alpha a_n^2/n}) = \varpi(n)$ *for some* $\alpha = \alpha(c) > 0$, *and that* $\delta_{n2}(c) = \varpi(n)$. *Assume also that* $v_n(t) := \mathbb{E}(\mathrm{e}^{t\xi_1^{(n)}})$ *exists for all* $|t| \leqslant t_3$ *for some* $t_3 > 0$, *and that* $\limsup_{n\to\infty}\{v_n(t_3) + v_n(-t_3)\} < \infty$. *Then the sequences of measures* P_n *and* Q_n *defined by*

$$P_n(k) = \mathbb{P}[Z_n = k] \quad \text{and} \quad Q_n(k) = s_n^{-1}\phi(u_{n,k}), \quad k \in \mathbf{Z},$$

where $u_{n,k} := s_n^{-1}(k - a_n m_1^{(n)})$, *are* $\mathrm{ARC}(v_n, \varpi(n))$, *where*

$$v_n := \log^2 n \sup_{t \in \mathbf{C}: |t| \leqslant 2t_1(n)} |\hat{\varepsilon}_n(t)|, \quad t_1(n) = n^{-1/2}\log n$$

and $\hat{\varepsilon}_n(t)$ *is as defined in* (3.5), *provided that* $v_n \geqslant n^{-\alpha}$ *for some* $\alpha > 0$.

Proof. Arguing much as for (3.9), we derive

$$s_n \mathbb{P}(Z_n = k) = \frac{1}{2\pi}\int_{-s_n t_1(n)}^{s_n t_1(n)} \exp(-\mathrm{i}yu_{n,k} - y^2/2 + \hat{\varepsilon}_n(y/s_n)(a_n + n)y^2 s_n^{-2})\,\mathrm{d}y + \varpi(n),$$

uniformly in

$$k \in R_n' := \{k: s_n^{-1}|k - a_n m_1^{(n)}| \leqslant \tfrac{1}{2}s_n t_1(n)\}.$$

Putting

$$u = u_{n,k}, \quad d = s_n t_1(n) \quad \text{and} \quad L(y) = \exp\{y^2 \hat{\varepsilon}_n(y/s_n)(a_n + n)s_n^{-2}\},$$

the integral can be written as

$$\frac{1}{\sqrt{2\pi}}\int_{-d}^{d} \phi(y)\mathrm{e}^{-\mathrm{i}yu} L(y)\,\mathrm{d}y.$$

By Cauchy's formula, this is equal to the integral along the contour consisting of the three chords

γ_1 : joining $-d$ to $-d - \mathrm{i}u$; γ_2 : joining $-d - \mathrm{i}u$ to $d - \mathrm{i}u$;

γ_3 : joining $d - \mathrm{i}u$ to d,

and $L(y) = 1 + \mathrm{O}(v_n)$ uniformly along the contour. The integrals along γ_1 and γ_3 are immediately seen to be of order

$$\mathrm{O}(|u|\exp\{-3d^2/8\}) = \varpi(n),$$

since $|u| \leqslant d/2$, and the remaining integral, with $y = z - \mathrm{i}u$, is just

$$\begin{aligned}
&\phi(u) \int_{-d}^{d} \phi(z) L(z - \mathrm{i}u)\,\mathrm{d}z \\
&\quad = \phi(u) \int_{-d}^{d} \phi(z)[1 + \mathrm{O}(v_n)]\,\mathrm{d}z = \phi(u)\{(1 + \mathrm{O}(v_n)) + \varpi(n)\} \\
&\quad = \phi(u)\{1 + \mathrm{O}(v_n)\},
\end{aligned}$$

since $v_n \geqslant n^{-\alpha}$. Combining these estimates, we establish the approximation

$$s_n \mathbb{P}[Z_n = k] = \phi(u_{n,k})(1 + \eta_{n,k}^{(1)}) + \eta_{n,k}^{(2)}, \tag{3.11}$$

with $\eta_{n,k}^{(1)} = \mathrm{O}(v_n)$ and $\eta_{n,k}^{(2)} = \varpi(n)$ uniformly for $k \in R_n'$.

Now let

$$A_1 := \{k \in R_n' :\ \log \mathbb{P}[Z_n = k] > \log(s_n^{-1}\phi(u_{n,k})) + 2c^* v_n\},$$

where c^* is such that $|\eta_{n,k}^{(1)}| \leqslant c^* v_n$ uniformly in n and k. Then, using (3.11), we have

$$\begin{aligned}
\log Q_n(A_1) + 2c^* v_n &< \log \mathbb{P}[Z_n \in A_1] \\
&\leqslant \log\{Q_n(A_1)(1 + c^* v_n) + s_n^{-1}|R_n'|\eta_n^{(2)}\},
\end{aligned} \tag{3.12}$$

where $\eta_n^{(2)} = \max_{k \in R_n'} |\eta_{n,k}^{(2)}|$, and so

$$c^* v_n \leqslant \log\left(1 + \frac{|R_n'|\eta_n^{(2)}}{s_n Q_n(A_1)}\right) \leqslant \frac{|R_n'|\eta_n^{(2)}}{s_n Q_n(A_1)}.$$

Hence, and from (3.12), both $Q_n(A_1)$ and $\mathbb{P}[Z_n \in A_1]$ are of order $\varpi(n)$. A similar argument covers the case of

$$A_2 := \{k \in R_n' :\ \log(s_n^{-1}\phi(u_{n,k})) > \log \mathbb{P}[Z_n = k] + 2c^* v_n\},$$

so that we shall have proved that $\mathscr{L}(Z_n)$ and Q_n are ARC$(v_n, \varpi(n))$ by taking R_n to be the set $R_n' \setminus \{A_1 \cup A_2\}$, provided only that $\mathbb{P}[Z_n \notin R_n'] = \varpi(n)$, since the corresponding bound for Q_n is immediate.

For this final step, we use Laplace-transform bounds. Observe that for real t with $|t| \leqslant \{n^{-1}\log n\}^{1/2}$, much as for (3.4),

$$\begin{aligned}
\mathbb{E}\{\mathrm{e}^{t(Z_n - a_n m_1^{(n)})}\} &= \mathrm{e}^{-t a_n m_1^{(n)}} \varphi_n(-\log v_n(t)) \\
&\leqslant \exp\{knt^2\} + \mathrm{e}^{-t a_n m_1^{(n)}} \delta_{n1}(c_3),
\end{aligned}$$

for some $k < \infty$ and $c_3 > 0$. Taking $t = t_n = \{n^{-1}\log n\}^{1/2}$, it follows easily that

$$\begin{aligned}
&\mathbb{P}\left[Z_n - a_n m_1^{(n)} > \frac{1}{2} s_n^2 t_1(n)\right] \\
&\quad \leqslant \exp\left\{-\frac{1}{2} t_n s_n^2 t_1(n)\right\} \{\exp\{knt_n^2\} + \mathrm{e}^{-t_n a_n m_1^{(n)}} \delta_{n1}(c_3)\} \\
&\quad = \varpi(n),
\end{aligned}$$

since we have

$$t_n s_n^2 t_1(n) \asymp \log^{3/2} n \gg knt_n^2 = O(\log n),$$

$$t_n a_n |m_1^{(n)}| \asymp a_n \{n^{-1} \log n\}^{1/2} \ll n^{-1} a_n^2,$$

and because of the bound on $\delta_{n1}(c)$. A similar bound for $\mathbb{P}[Z_n - a_n m_1^{(n)} < -\frac{1}{2} s_n^2 t_1(n)]$ is obtained by taking $t = -t_n$, and the theorem is proved. □

We now wish to apply the above results to the Reed–Frost epidemic, using the construction in (2.9). The idea is to express the final size as being close enough to the sum of a random number of independent and identically distributed random variables. This is actually based on a second branching process approximation. For times r near the end of the epidemic, $I^{(n)}(r)$ reduces to values of much smaller order than n, and $S^{(n)}(r)$ becomes close to a value $n\theta$, where θ is the proportion of susceptibles typically remaining at the end of a major outbreak. At such times, the number of contacts with *susceptibles* made by a given infective is close to having a $\mathrm{Po}(\lambda\theta)$ distribution, and $\lambda\theta < 1$. Hence each infective at time r gives rise to a rather small total number of cases over the remaining duration of the epidemic, having mean $1/(1 - \lambda\theta)$, and, if $I^{(n)}(r)$ is small enough, these numbers for the different infectives are close to being independent and identically distributed. This idea is actually exploited from within the process $\widehat{X}^{(n)}$ of the construction in (2.9), which makes for a simpler argument; note that the final size is given, as in (2.11), by

$$\tau^{(n)} := \min\{j \geqslant 0 : \widehat{X}_j^{(n)} + j \geqslant n + i_n\}, \tag{3.13}$$

where $\widehat{X}_0^{(n)} = S^{(n)}(0) = n$ denotes the initial number of susceptibles, and $I^{(n)}(0) = i_n$ the initial number of infectives.

Our normal approximation theorem concerns the distribution of $\tau^{(n)}$ conditional on there having been a large outbreak; for small epidemics, the branching process approximation in Theorem 2.5 can be used. A large outbreak has negligible probability of occurring unless $\lambda > 1$; if this condition is satisfied, we define the epidemic to have been large if $\tau^{(n)} > n(1 - 1/\lambda)$, and denote the corresponding conditional probability by $\mathbb{P}_L$.

Theorem 3.3. *Assume that $\lambda > 1$ and $n^{-1} i_n \to \imath$ with $0 \leqslant \imath < \infty$. Define the sequences of measures $(P_n, Q_n,\ n \geqslant 1)$ on $\mathbf{Z}_+$ by*

$$P_n(k) = \mathbb{P}_L[\tau^{(n)} = k] \quad \text{and} \quad Q_n(k) = \frac{1}{\sigma\sqrt{n}}\, \phi\left(\frac{k - n(1-\theta_n)}{\sigma\sqrt{n}}\right), \qquad k \geqslant 0.$$

Then $(P_n,\ n \geqslant 1)$ and $(Q_n,\ n \geqslant 1)$ are $\mathrm{ARC}(\varepsilon_n, \varpi(n))$ with

$$\varepsilon_n = |n^{-1} i_n - \imath| \log^2 n + n^{-1/4} \log^5 n,$$

where $\sigma^2 = \theta(1-\theta)/(1-\lambda\theta)^2$, $\theta = \theta(\imath)$ and $\theta_n = \theta(n^{-1} i_n)$, with $\theta(w)$ the solution in $(0,1)$ of the equation $\theta - \lambda^{-1} \log \theta = 1 + w$.

The theorem immediately implies a local limit approximation.

Corollary 3.4. *Under the conditions of Theorem* 3.3,

$$\sup_k \left| \sqrt{n}\,\mathbb{P}_L[\tau^{(n)} = k] - \frac{1}{\sigma}\,\phi\left(\frac{k - n(1-\theta_n)}{\sigma\sqrt{n}}\right)\right|$$
$$= \mathrm{O}(|n^{-1}i_n - \imath|\log^2 n + n^{-1/4}\log^5 n).$$

Remark. A similar rate of convergence in the (distributional) central limit theorem was established in [3] for sparse undirected random graphs, by using the martingale approximation approach.

Proof. The first step is to show that $n^{-1}\tau^{(n)}$ is either very small or else near to the value $n^{-1}(1-\theta_n)$. Rather than using coupling arguments as in [5], we apply simple large deviation bounds applied to the process $\widehat{X}^{(n)}$: in what follows, we always omit the hat, and write just $X^{(n)}$. Since $X_j^{(n)} \sim \mathrm{Bi}(n, (1-n^{-1}\lambda)^j)$, with mean $\mathbb{E}X_j^{(n)} \leqslant n\mathrm{e}^{-j\lambda/n}$ and variance $\mathrm{Var}(X_j^{(n)}) \leqslant n(1-\mathrm{e}^{-j\lambda/n}) \leqslant j\lambda$, it follows from the Bernstein inequality (see for example [8, p. 193]) that

$$\mathbb{P}[\tau^{(n)} = j] \leqslant \mathbb{P}[X_j^{(n)} = n + i_n - j] \leqslant \exp\left\{-\frac{nf_n^2(t)}{2\lambda t + 2f_n(t)/3}\right\},$$

where $t = j/n$ and

$$1 + \sup_n(n^{-1}i_n) = C \geqslant f_n(t) = (1 + n^{-1}i_n - t - \mathrm{e}^{-\lambda t}) > 0,$$

provided that $0 < t < 1 - \beta_n$, where $\beta_n = \theta_n - i_n$ and θ_n is as defined in the statement of the theorem. Note also that

$$f_n(t) \geqslant \begin{cases} \frac{1}{2}(\lambda-1)t & \text{if } 0 \leqslant t \leqslant 1 - 1/\lambda, \\ (1-\beta_n - t)(1-\lambda\mathrm{e}^{1-\lambda}) & \text{if } 1 - 1/\lambda \leqslant t \leqslant 1 - \beta_n, \end{cases}$$

so that

$$\mathbb{P}[\tau^{(n)} = j] \leqslant \begin{cases} K\mathrm{e}^{-jc_1(\lambda)} & \text{if } 0 \leqslant t \leqslant 1 - 1/\lambda, \\ K\exp\{-c_2(\lambda)n^{-1}(n(1-\beta_n) - j)^2\} & \text{if } 1 - 1/\lambda \leqslant t \leqslant 1 - \beta_n \end{cases}$$

for

$$c_1(\lambda) = \frac{(\lambda-1)^2/2}{4\lambda + 2(\lambda-1)/3} \quad \text{and} \quad c_2(\lambda) = \frac{(1-\lambda\mathrm{e}^{1-\lambda})^2}{2\lambda(1-1/\lambda) + 2C/3}.$$

Hence

$$\mathbb{P}[\log^2 n \leqslant \tau^{(n)} \leqslant u_n] = \varpi(n), \tag{3.14}$$

if we define

$$u_n = n(1-\beta_n) - \lfloor\sqrt{n}\log^2 n\rfloor = n(1-\theta_n) + i_n - \lfloor\sqrt{n}\log^2 n\rfloor.$$

We next examine the distribution of the modified stopping time

$$\tilde{\tau}^{(n)} := \min\{j \geqslant u_n : X_j^{(n)} + j \geqslant n + i_n\},$$

where the process is not stopped at the end of a 'small' epidemic. Then it follows from the above that $X_{u_n}^{(n)} \sim \mathrm{Bi}(n, (1-n^{-1}\lambda)^{u_n})$ and that

$$\mathbb{P}[X_{u_n}^{(n)} \geqslant n + i_n - u_n] = \varpi(n); \tag{3.15}$$

also, since

$$\mathbb{E}X_{u_n}^{(n)} = n(1-n^{-1}\lambda)^{u_n} = n\theta_n \exp\{\lambda n^{-1/2}\log^2 n\} + \mathrm{O}(1), \tag{3.16}$$

it follows that

$$n + i_n - u_n - X_{u_n}^{(n)} = \sqrt{n}\log^2 n(1-\lambda\theta_n) + \{\mathbb{E}X_{u_n}^{(n)} - X_{u_n}^{(n)}\} + \mathrm{O}(\log^4 n),$$

and hence, from the Chernoff bounds, that

$$\mathbb{P}[n + i_n - u_n - X_{u_n}^{(n)} > 2\sqrt{n}\log^2 n] = \varpi(n). \tag{3.17}$$

In particular, from (3.15), it follows that

$$d_{\mathrm{TV}}(\mathscr{L}(\tilde{\tau}^{(n)}), \mathscr{L}(\tau^{*(n)})) = \varpi(n), \tag{3.18}$$

where

$$\tau^{*(n)} := u_n + \tau(X_{u_n}^{(n)}, [n + i_n - u_n - X_{u_n}^{(n)}]_+) \tag{3.19}$$

and $\tau(m,i)$ denotes the final size of an epidemic starting with m susceptibles and i infectives.

Now $\tau(X_{u_n}^{(n)}, [n + i_n - u_n - X_{u_n}^{(n)}]_+)$ is zero if $[n + i_n - u_n - X_{u_n}^{(n)}]_+ = 0$, and is otherwise the first time j at which $n + i_n - u_n \leqslant Y_j^{(n)} + j$, where $Y^{(n)}$ is a process like $X^{(n)}$ in Lemma 2.3, satisfying the recursion $Y_{j+1}^{(n)} = Y_j^{(n)} - Z_j^{(n)}$, $j \geqslant 0$, with initial value $Y_0^{(n)} = X_{u_n}^{(n)}$, where, conditional on $\mathscr{F}_j^{(n)}$, $Z_j^{(n)} \sim \mathrm{Bi}(Y_j^{(n)}, n^{-1}\lambda)$. What is more, we have $Y_j^{(n)} \sim \mathrm{Bi}(Y_0^{(n)}, (1-n^{-1}\lambda)^j)$, implying from the Chernoff bounds that, writing $t^{(n)} = 3\sqrt{n}\log^2 n$,

$$\mathbb{P}[Y_{t^{(n)}}^{(n)} + t^{(n)} \geqslant Y_0^{(n)} + 2\sqrt{n}\log^2 n] = 1 - \varpi(n), \tag{3.20}$$

which, with (3.17), implies that $\tau(X_{u_n}^{(n)}, [n + i_n - u_n - X_{u_n}^{(n)}]_+) \leqslant t^{(n)}$ with probability $1 - \varpi(n)$. Hence, for our purposes, we can approximate the distribution of the stopping time $\tau(X_{u_n}^{(n)}, [n + i_n - u_n - X_{u_n}^{(n)}]_+)$ by working instead with any process which is $\mathrm{ARC}(\varepsilon_n, \varpi(n))$ close to $Y^{(n)}$ up to time $t^{(n)}$, since ARC-closeness of the processes carries over to the distributions of the stopping times of interest to us, by Corollary A.2.

We now apply Lemmas 2.1, 2.3 and 2.4 with $T^{(n)} = t^{(n)} = 3\sqrt{n}\log^2 n$ to the process $Y^{(n)}$ defined above and to $Y'^{(n)}$ defined by

$$Y_{j+1}'^{(n)} = Y_j'^{(n)} - Z_j'^{(n)}, \quad j \geqslant 0; \quad Y_0'^{(n)} = Y_0^{(n)} = X_{u_n}^{(n)},$$

where $Z_j'^{(n)} \sim \mathrm{Po}(\lambda\theta_n)$. Note that, in applying the lemmas, we have $m_n = X_{u_n}^{(n)}$, so that the mean of the Poisson innovations in Lemma 2.1 is $n^{-1}\lambda X_{u_n}^{(n)}$, and Lemma 2.4 is used to replace this mean with $\lambda\theta_n$. This can be done since, from (3.16) and the

Chernoff bounds for the Binomial distribution,

$$\mathbb{P}[|n^{-1}X^{(n)}_{u_n} - \lambda\theta_n| > c_4 n^{-1/2}\log^2 n] = \varpi(n),$$

for some $c_4 < \infty$. Thus $Y^{(n)}$ and $Y'^{(n)}$ are ARC$(\varepsilon'_n, \varpi(n))$ up to time $t^{(n)}$, where, with $\Delta^{(n)} = c_4 n^{1/2}\log^2 n$,

$$\varepsilon'_n = \{t^{(n)}\}^{1/2}(t^{(n)} + \Delta^{(n)})\{\log n\}^2 n^{-1} = \mathrm{O}(n^{-1/4}\log^5 n).$$

In consequence, with $Y_\infty(J)$ defined by

$$Y_\infty(J) := \min\left\{ t \geqslant 0 : \sum_{j=1}^{t}(1 - Z'^{(n)}_t) \geqslant J \right\},$$

the sequences of measures P'_n and Q'_n defined by

$$P'_n(k) = \mathbb{P}[\tau(X^{(n)}_{u_n}, [n + i_n - u_n - X^{(n)}_{u_n}]_+) = k],$$

$$Q'_n(k) = \mathbb{P}[Y_\infty([n + i_n - u_n - X^{(n)}_{u_n}]_+) = k],$$

are ARC$(\varepsilon'_n, \varpi(n))$; thus the problem is reduced to proving the ARC-closeness of

$$\mathscr{L}(u_n + Y_\infty([n + i_n - u_n - X^{(n)}_{u_n}]_+)) \text{ and } Q_n.$$

Now, as for Corollary 2.7, $Y_\infty(J)$ is just the total progeny in the Galton–Watson process with offspring distribution Po$(\lambda\theta_n)$, starting with J individuals, and so

$$Y_\infty(J) = \sum_{i=1}^{J} \xi_i,$$

where $\{\xi_i\}$ is an i.i.d. sequence distributed as $Y_\infty(1)$; in particular, we have

$$\xi_1 \geqslant 1 \text{ a.s.}, \quad \mathbb{P}[\xi_1 = 1] = \mathrm{e}^{-\lambda\theta_n}, \quad \mathbb{E}\xi_1 = (1 - \lambda\theta_n)^{-1},$$

$$\mathbb{E}(\mathrm{e}^{t\xi_1}) < \infty \quad \text{for all } t \leqslant \lambda\theta_n - 1 - \log(\lambda\theta_n). \tag{3.21}$$

Thus the ξ_i satisfy the relevant conditions in Theorem 3.2, with $m_1^{(n)} = (1 - \lambda\theta_n)^{-1}$, and we now wish to show that $[n + i_n - u_n - X^{(n)}_{u_n}]_+$ satisfies the conditions for S_n in that theorem, so that it can be applied to approximate $\mathscr{L}(u_n + Y_\infty([n + i_n - u_n - X^{(n)}_{u_n}]_+))$.

To show that $S_n := [n + i_n - u_n - X^{(n)}_{u_n}]_+$ indeed satisfies the conditions, we use the fact that $X^{(n)}_{u_n} \sim \mathrm{Bi}(n, p_n)$ with $p_n := (1 - n^{-1}\lambda)^{u_n}$; this suggests using the Laplace transform of $\widehat{S}_n := n + i_n - u_n - X^{(n)}_{u_n}$ as the approximation to $\varphi_n(\psi) = \mathbb{E}\{\mathrm{e}^{-\psi S_n}\}$, which from the definitions of u_n, θ_n and p_n, gives

$$a_n := (n + i_n - u_n) - n p_n = \sqrt{n}\log^2 n(1 - \lambda\theta_n) + \mathrm{O}(\log^4 n) \gg \{n\log n\}^{1/2},$$

$$b_n := n p_n(1 - p_n) \sim n\theta(1 - \theta) = \mathrm{O}(n),$$

$$\varepsilon_n(\psi) := \psi^{-2}\{\log(1 + p_n(\mathrm{e}^{\psi} - 1)) - \psi p_n - \tfrac{1}{2}\psi^2 p_n(1 - p_n)\},$$

satisfying all the conditions relating to a_n, b_n and $\varepsilon_n(\cdot)$, with $s^2 := \theta(1-\theta)(1-\lambda\theta)^{-2}$ and $v_n := \sup_{|t| \leqslant 2t_1(n)} |\hat{\varepsilon}_n(t)| \asymp n^{-1/2}\log n$. It therefore remains to show that the quantities $\delta_{n1}(c)$ and $\delta_{n2}(c)$ are suitably small.

For $\delta_{n2}(c)$, it is enough to note that, in $\psi_0 := \Re\psi \geqslant 0$,

$$|\mathbb{E}(e^{-\psi S_n})| \leqslant \exp\{-\tfrac{1}{2}a_n\psi_0\} + \mathbb{P}[\widehat{S}_n < \tfrac{1}{2}a_n] = \varpi(n),$$

uniformly in $\psi_0 \geqslant c\{n^{-1}\log n\}^{1/2}$, because of the definition of a_n and the Chernoff bounds on the binomial distribution. For $\delta_{n1}(c)$, note that, again in $\psi_0 := \Re\psi \geqslant 0$,

$$|\mathbb{E}(e^{-\psi S_n}) - \mathbb{E}(e^{-\psi\widehat{S}_n})| \leqslant \mathbb{E}\{e^{-\psi_0\widehat{S}_n} I[\widehat{S}_n \leqslant 0]\},$$

and that this, by considering ratios of successive probabilities in the binomial distribution, is at most of order

$$(n/a_n)\exp\{-a_n^2/(2np_n(1-p_n))\} = O(e^{-\alpha a_n^2/n}) = \varpi(n)$$

for any $0 < \alpha < 1/\{2\theta(1-\theta)\}$, uniformly in $0 \leqslant \psi_0 \leqslant c\{n^{-1}\log n\}^{1/2}$ for any $c > 0$; if $\Re\psi < 0$, the Chernoff bound alone suffices. Applying Theorem 3.2, we see that $\mathscr{L}(u_n + Y_\infty([n + i_n - u_n - X_{u_n}^{(n)}]_+))$ and Q_n'' are $\mathrm{ARC}(v_n, \varpi(n))$, where

$$Q_n''(k) := \frac{1}{s_n}\phi\left(\frac{k - u_n - a_n m_1^{(n)}}{s_n}\right);$$

then, applying Lemma A.3 to replace s_n by $\sigma\sqrt{n}$ and $u_n + a_n m_1^{(n)}$ by $n(1-\theta_n)$, we deduce that Q_n'' and Q_n are $\mathrm{ARC}(\varepsilon_n'', \varpi(n))$ with

$$\varepsilon_n'' = O(n^{-1/2}\log^4 n + |n^{-1}i_n - \imath|), \tag{3.22}$$

since

$$n^{-1}s_n^2 = p_n(1-p_n)\{m_1^{(n)}\}^2 + O(n^{-1/2}\log^2 n)$$
$$= \theta_n(1-\theta_n)(1-\lambda\theta_n)^{-2} + O(n^{-1/2}\log^2 n),$$

$$|\theta_n - \theta| = O(|n^{-1}i_n - \imath|) \quad \text{and} \quad |n(1-\theta_n) - u_n - a_n m_1^{(n)}| = O(\log^4 n).$$

Combining the previous approximations, we have thus shown that $\mathscr{L}(\tilde{\tau}^{(n)})$ and Q_n are $\mathrm{ARC}(\varepsilon_n, \varpi(n))$ close.

We now wish to replace the distribution of $\tilde{\tau}^{(n)}$ by the distribution of $\tau^{(n)}$ conditional on $\tau^{(n)} \geqslant n(1 - 1/\lambda)$. From (3.14), with probability $1 - \varpi(n)$, $\tau^{(n)}$ differs from $\tilde{\tau}^{(n)}$ only if $X_j^{(n)} + j = n + i_n$ for some $1 \leqslant j \leqslant \lfloor\log^2 n\rfloor$, so that there is essentially no correction to be made if $i_n > \lfloor\log^2 n\rfloor$. Otherwise, we have

$$\mathbb{P}_L[\tau^{(n)} = k] = \frac{\mathbb{P}[\tilde{\tau}^{(n)} = k] - \sum_{l=i_n}^{\lfloor\log^2 n\rfloor}\mathbb{P}[\tilde{\tau}^{(n)} = k \mid \tau^{(n)} = l]\mathbb{P}[\tau^{(n)} = l]}{\mathbb{P}[\tau^{(n)} > \lfloor\log^2 n\rfloor]} + \varpi(n)$$

$$= \frac{\mathbb{P}[\tilde{\tau}^{(n)} = k] - \sum_{l=i_n}^{\lfloor\log^2 n\rfloor}P_{n,l}(k)\mathbb{P}[\tau^{(n)} = l]}{\mathbb{P}[\tau^{(n)} > \lfloor\log^2 n\rfloor]} + \varpi(n), \tag{3.23}$$

where $P_{n,l}$ denotes $\mathbb{P}[\tilde{\tau}^{(n)} \in \cdot \mid \tau^{(n)} = l]$. The sequence of measures $\mathscr{L}(\tilde{\tau}^{(n)})$ has already been approximated by Q_n. To approximate the measures $P_{n,l}(k)$ for each $l \geqslant i_n$, the conditional distribution of $\tilde{\tau}^{(n)}$ given $\tau^{(n)} = l$ can be analyzed exactly as before, except that the new process $X^{(n)}$ starts at time l with value $n + i_n - l$ instead of at time zero with value n, and it follows by the argument above that, for each fixed l, the measures $P_{n,l}$ and $Q_{n,l}$ are ARC$(\varepsilon_n, \varpi(n))$ close, uniformly in l, where

$$Q_{n,l}(k) = \frac{1}{\sigma}\,\phi\left(\frac{k - a_n^l}{\sigma\sqrt{n}}\right),$$

and $|a_n^l - a_n| \leqslant Kl$ for some $K > 0$. Hence, by Lemma A.3 applied to the measures $Q_{n,l}$ and Q_n for each l, the measures $P_{n,l}$ and Q_n are ARC$(\varepsilon_{n,l}, \varpi(n))$, where

$$\varepsilon_{n,l} = \mathrm{O}(\varepsilon_n + \ln^{-1/2} \log n).$$

It remains only to apply Lemma A.4 to the measures Q_n and $(P_{n,l} : n \geqslant 1, l = i_n, \ldots, \lfloor \log^2 n \rfloor)$, with $r_{n,l} := \mathbb{P}[\tau^{(n)} = l]/\mathbb{P}[\tau^{(n)} \leqslant \lfloor \log^2 n \rfloor]$, and then Lemma 4.5 to complete the proof. □

Appendix A. Properties of ARC

The definition of relative closeness depends on the typical set R being $\mathscr{F}$-measurable. While this may seem a trivial requirement, it needs to be considered when taking functionals. If P and Q are RC(ε, η) and $f : (\mathscr{X}, \mathscr{F}) \to (\mathscr{Y}, \mathscr{G})$ is measurable, it may well be that $R \notin f^{-1}(\mathscr{G})$, so that Pf^{-1} and Qf^{-1} are not automatically RC(ε, η). The following lemma examines the relationship in more detail.

Lemma A.1. *Suppose that measures P and Q are RC(ε, η) close in $(\mathscr{X}, \mathscr{F})$. Let $A \in \mathscr{F}$ be such that $|\log Q(A) - \log P(A)| > \varepsilon + t$, for some $t > 0$. Then*

$$\max\{P(A), Q(A)\} \leqslant \eta(1 + t^{-1}).$$

As a result, if $f : (\mathscr{X}, \mathscr{F}) \to (\mathscr{Y}, \mathscr{G})$ is measurable, then the measures Pf^{-1} and Qf^{-1} are RC$(\varepsilon + t, 2\eta(1 + t^{-1}))$ *for any $t > 0$.*

Proof. Suppose that $\log Q(A) + \varepsilon + t < \log P(A)$. Then $Q(A) < P(A)$ and

$$\begin{aligned}\log Q(A) + \varepsilon + t &< \log(P(A \cap R) + P(A \setminus R)) \\ &\leqslant \log Q(A) + \varepsilon + \log\left(1 + \frac{P(A \setminus R)}{P(A) - P(A \setminus R)}\right),\end{aligned}$$

where R denotes the set on which $|\log(\mathrm{d}P/\mathrm{d}Q)| \leqslant \varepsilon$. Hence

$$t < \frac{P(A \setminus R)}{P(A) - P(A \setminus R)},$$

and thus $Q(A) < P(A) < P(A \cap R^c)(1 + t^{-1})$. Applying the first part of the lemma to the sets

$$A_1 = \left\{ y : \frac{d(Qf^{-1})}{d(Pf^{-1})}(y) \leqslant e^{-\varepsilon - t} \right\} \quad \text{and} \quad A_2 = \left\{ y : \frac{d(Pf^{-1})}{d(Qf^{-1})}(y) \leqslant e^{-\varepsilon - t} \right\}$$

proves the second part of the lemma. □

We shall only make use of the following simple consequence.

Corollary A.2. *Suppose that the measures P_n and Q_n are* ARC$(\varepsilon_n, \varpi(n))$ *in* $(\mathscr{X}, \mathscr{F})$, *where $\varepsilon_n > n^{-\alpha}$ for some $\alpha > 0$. Then, for any measurable $f : (\mathscr{X}, \mathscr{F}) \to (\mathscr{Y}, \mathscr{G})$, the measures $P_n f^{-1}$ and $Q_n f^{-1}$ are* ARC$(\varepsilon_n, \varpi(n))$ *also.*

The next lemma states simple ARC-closeness properties of discrete normal densities. The proof involves only elementary calculations.

Lemma A.3. (i) *Let sequences of measures $(P_n : n \geqslant 1)$ and $(Q_n : n \geqslant 1)$ be defined by*

$$P_n(k) = \phi\left(\frac{k - a_n}{s_n}\right) \quad \textit{and} \quad Q_n(k) = \phi\left(\frac{k - b_n}{s_n}\right).$$

Then, if $\varepsilon_n := |a_n - b_n| s_n^{-1} \log n \to 0$ as $n \to \infty$, the measures P_n and Q_n are ARC$(\varepsilon_n, \varpi(n))$.

(ii) *Let sequences of measures $(P_n : n \geqslant 1)$ and $(Q_n : n \geqslant 1)$ be defined by*

$$P_n(k) = \phi\left(\frac{k - a_n}{s_n}\right) \quad \textit{and} \quad Q_n(k) = \phi\left(\frac{k - a_n}{\sigma\sqrt{n}}\right).$$

Then, if $\varepsilon_n = |n^{-1}s_n^2 - \sigma^2| \log^2 n \to 0$ as $n \to \infty$, the measures P_n and Q_n are ARC$(\varepsilon_n, \varpi(n))$.

In the next lemma, we give conditions under which ARC-closeness is preserved under mixtures.

Lemma A.4. *Let Q and $(P_j : j \geqslant 1)$ be measures such that, for each j, P_j and Q are* RC(ε_j, η_j), *and let $(r_j : j \geqslant 1)$ be non-negative real numbers summing to* 1. *Then $P^* := \sum_{j \geqslant 1} r_j P_j$ and Q are* RC$(\varepsilon + t, 2\eta(1 + t^{-1}))$ *for any $t > 0$, where $\varepsilon := \max_{j \geqslant 1} \varepsilon_j$ and $\eta := \sum_{j \geqslant 1} r_j \eta_j$.*

Proof. The proof is much as for Lemma A.1. Let

$$A := \{x \in \mathscr{X} \colon (dP^*/dQ)(x) < -(\varepsilon + t)\}.$$

Then $\log P^*(A) + \varepsilon + t < \log Q(A)$, and so

$$\log P^*(A) + \varepsilon + t < \log\left(\sum_{j \geqslant 1} r_j \{Q(A \cap R_j) + Q(A \setminus R_j)\}\right),$$

where the R_j are the typical sets in the RC-comparisons of P_j and Q. Hence it follows that either $Q(A) \leqslant \eta$ or

$$\begin{aligned}\log P^*(A) + \varepsilon + t &< \log\left(\sum_{j\geqslant 1} r_j Q(A\cap R_j)\right) + \log(1+\eta/(Q(A)-\eta))\\ &\leqslant \log\left(\sum_{j\geqslant 1} r_j P(A)\mathrm{e}^{\varepsilon_j}\right) + \eta/(Q(A)-\eta)\\ &\leqslant \log P^*(A) + \varepsilon + \eta/(Q(A)-\eta),\end{aligned}$$

and thus $P^*(A) < Q(A) < \eta(1+t^{-1})$. The argument for

$$A' := \{x\in\mathscr{X}\colon\ (\mathrm{d}P^*/\mathrm{d}Q)(x) > \varepsilon + t\}$$

is similar, and the lemma follows easily. □

The final lemma allows linear combinations with some negative coefficients.

Lemma A.5. *Let P_1, P_2 and Q be measures on $(\mathscr{X},\mathscr{F})$ such that P_j and Q are* RC(ε_j,η_j) *for $j = 1,2$. Let $\theta > 0$, and define $P^* := (1+\theta)P_1 - \theta P_2$. Suppose also that $\varepsilon_1, \varepsilon_2$ and θ are such that*

$$\max\{\varepsilon_1,\varepsilon_2\} \leqslant 1/2, \quad (1+\theta)\mathrm{e}^{-\varepsilon_1} - \theta\mathrm{e}^{\varepsilon_2} > 1/2,$$
$$\varepsilon := (1+\theta)(\mathrm{e}^{\varepsilon_1}-1) + \theta(1-\mathrm{e}^{-\varepsilon_2}) \leqslant 7. \tag{A.1}$$

Then, for any $t > 0$, the measures P^* and Q are RC$(8\varepsilon + t, 2\eta(8t^{-1}+2))$, *with $\eta := (1+\theta)\eta_1 + \theta\eta_2$.*

Proof. The proof makes frequent use of the following inequalities: if measures Q_1 and Q_2 are RC(ε',η') with typical set R', then

$$(Q_2(A)-\eta')\mathrm{e}^{-\varepsilon'} \leqslant Q_1(A) = Q_1(A\cap R') + Q_1(A\setminus R') \leqslant Q_2(A)\mathrm{e}^{\varepsilon'} + \eta'. \tag{A.2}$$

We begin by taking

$$A := \{x\in\mathscr{X}\colon\ (\mathrm{d}P^*/\mathrm{d}Q)(x) < -(8\varepsilon + t)\}.$$

Then, from (A.2),

$$\begin{aligned}&\log P^*(A) + 8\varepsilon + t < \log Q(A)\\ &\quad = \log\{(1+\theta)Q(A) - \theta Q(A)\}\\ &\quad \leqslant \log\{(1+\theta)(P_1(A)\mathrm{e}^{\varepsilon_1} + \eta_1) - \theta(P_2(A)-\eta_2)\mathrm{e}^{-\varepsilon_2}\}\\ &\quad \leqslant \log\{P^*(A) + (1+\theta)P_1(A)(\mathrm{e}^{\varepsilon_1}-1) + \theta P_2(A)(1-\mathrm{e}^{-\varepsilon_2}) + \eta\}.\end{aligned} \tag{A.3}$$

Now note that, from (A.1) and (A.2), if $Q(A) > 4\eta$, then

$$\begin{aligned} P^*(A) &\geqslant (1+\theta)\mathrm{e}^{-\varepsilon_1}(Q(A)-\eta_1) - \theta(\mathrm{e}^{\varepsilon_2}Q(A)-\eta_2) \\ &\geqslant Q(A)\{(1+\theta)\mathrm{e}^{-\varepsilon_1} - \theta\mathrm{e}^{\varepsilon_2}\} - \eta \\ &\geqslant \tfrac{1}{4}Q(A), \end{aligned}$$

hence, from (A.3), we have

$$\begin{aligned} &\log P^*(A) + 8\varepsilon + t \\ &\quad < \log P^*(A) + (4/Q(A))\{(1+\theta)P_1(A)(\mathrm{e}^{\varepsilon_1}-1) + \theta P_2(A)(1-\mathrm{e}^{-\varepsilon_2}) + \eta\} \\ &\quad \leqslant \log P^*(A) + (4\eta/Q(A)) + 4\{(1+\theta)\mathrm{e}^{\varepsilon_1}(\mathrm{e}^{\varepsilon_1}-1) + \theta\varepsilon^{\varepsilon_2}(1-\mathrm{e}^{-\varepsilon_2})\} \\ &\qquad + (4\eta/Q(A))\max\{\mathrm{e}^{\varepsilon_1}-1, 1-\mathrm{e}^{-\varepsilon_2}\} \\ &\quad \leqslant \log P^*(A) + (4\eta/Q(A))\{1+\max(\mathrm{e}^{\varepsilon_1}-1, 1-\mathrm{e}^{-\varepsilon_2})\} + 8\varepsilon, \end{aligned}$$

giving

$$P^*(A) < Q(A) < (4\eta/t)\{1+\max(\mathrm{e}^{\varepsilon_1}-1, 1-\mathrm{e}^{-\varepsilon_2})\}.$$

For the set

$$A' := \{x \in \mathscr{X}\colon (\mathrm{d}P^*/\mathrm{d}Q)(x) > 8\varepsilon + t\},$$

we have, from (A.2),

$$\begin{aligned} \log Q(A') + 8\varepsilon + t &< \log P^*(A') \\ &\leqslant \log\{(1+\theta)(Q(A')\mathrm{e}^{\varepsilon_1}+\eta_1) - \theta(Q(A')\mathrm{e}^{-\varepsilon_2}-\eta_2) \\ &\leqslant \log Q(A') + \varepsilon + \eta/Q(A'), \end{aligned}$$

giving $Q(A') \leqslant t^{-1}\eta$. Hence also, from (A.2),

$$P^*(A) \leqslant Q(A)(1+\varepsilon) + \eta \leqslant \eta\{t^{-1}(1+\varepsilon)+1\},$$

completing the proof, since $\varepsilon \leqslant 7$. □

References

[1] F.G. Ball, P. Donnelly, Strong approximations for epidemic models, Stochastic. Process. Appl. 55 (1995) 1–21.

[2] F. Ball, O. Lyne, Parameter estimation for SIR epidemics in households, Bulletin of the International Statistical Institution, 52nd Session Contributed Papers, Vol. LVIII, Book 2, 1999, pp. 251–252.

[3] D. Barraez, S. Boucheron, W. Fernandez de la Vega, On the fluctuations of the giant component, Combin. Probab. Comput. 9 (2000) 287–304.

[4] T. Britton, N.G. Becker, Design issues for studies of infectious diseases, J. Statist. Plan. Inference 96 (2001) 41–66.

[5] Cl. Lefèvre, S. Utev, Poisson approximation for the final state of a generalized epidemic process, Ann. Probab. 23 (1995) 1139–1162.

[6] Cl. Lefèvre, S. Utev, Branching approximation for the collective epidemic model, Method. Comput. Appl. Probab. 1 (1999) 211–228.

[7] A. Martin-Löf, Symmetric sampling procedures, general epidemic processes and their threshold limit theorems, J. Appl. Probab. 23 (1986) 265–282.

[8] D. Pollard, Convergence of Stochastic Processes, Springer, New York, 1984.