The Annals of Applied Probability
1995, Vol. 5, No. 4, 1061–1086

THE ASYMPTOTIC EVOLUTION OF THE GENERAL STOCHASTIC EPIDEMIC[1]

BY GESINE REINERT

University of Southern California, Los Angeles

Generalizing Sellke's construction, a general stochastic epidemic with non-Markovian transition behavior is considered. At time $t = 0$, the population of total size K consists of aK individuals that are infected by a certain disease (and infectious); the remaining bK individuals are susceptible with respect to that disease. An initially susceptible individual i, when infected (call A_i^K its time of infection), stays infectious for a period of length r_i, until it is removed. An initially infected individual i stays infected for a period of length $\hat{r}_i$ until it is removed. Removed individuals can no longer be affected by the disease. A deterministic approximation as (as $K \to \infty$) to the empirical measure

$$\xi_K = \frac{1}{K}\sum_{i=1}^{aK} \delta_{(0,\hat{r}_i)} + \frac{1}{K}\sum_{i=1}^{bK} \delta_{(A_i^K, A_i^K + r_i)},$$

describing the average path behavior, is established using Stein's method.

Introduction. The general stochastic epidemic (GSE) is a complicated birth–death process where the temporal evolution of one individual depends "uniformly" on those of the others. The following construction, based on Sellke's [11] approach, yields an epidemic model that generalizes the GSE. However, as it is of the same type, we apply the term GSE to this more general model.

A population with total size K is considered. At time $t = 0$, aK of these individuals are infected by a certain disease (and infectious; the infectious period and the period of being infected are assumed to coincide); the remaining $bK = (1 - a)K$ individuals are susceptible to that disease. Infectious individuals will get removed after some time, for example, by lifelong immunity or death, and are then no longer affected by that disease. (Thus, we have an SIR model.)

Let $(l_i, r_i)_{i \in \mathbf{N}}$ be a family of positive i.i.d. random vectors and let $(\hat{r}_i)_{i \in \mathbf{N}}$ be a family of positive, independent random variables. Assume that the families $(l_i, r_i)_{i \in \mathbf{N}}$ and $(\hat{r}_i)_{i \in \mathbf{N}}$ are mutually independent.

An initially infected individual i stays infectious for a period of length $\hat{r}_i$; then it is removed. (That the $\hat{r}_i$ need not be identically distributed reflects the possibility that an infected individual has already been infectious for a certain period before, at time $t = 0$, it is observed.) An initially susceptible

Received January 1993; revised March 1995.
[1]Work supported in part by the DAAD and by the Schweizer Nationalfonds Grant 20-31262.91.
AMS 1991 *subject classifications*. 60K30, 60G57, 92D30.
Key words and phrases. General stochastic epidemic, empirical measures, Stein's method.

individual i, once infected, stays infectious for a period of length r_i, until it is removed. Furthermore, an initially susceptible individual i accumulates exposure to infection with a rate that depends on the evolution of the epidemic; if the total exposure reaches l_i, the individual i becomes infected. The possible dependence between l_i and r_i for each fixed i reflects the fact that both the resistance to infection and the duration of the infection may, for a fixed individual, depend on its physical constitution.

An initially susceptible individual i gets infected as soon as a certain functional, depending on the course of the epidemic, exceeds the individual's level l_i of resistance. Denote its infection time by A_i^K. To be more precise, if $Z_K(t)$ denotes the proportion of infected individuals present in the population at time $t \in \mathbf{R}_+$, then A_i^K is given by

$$A_i^K = \inf\left\{ t \in \mathbf{R}_+ : \int_{(0,t]} \lambda(s, Z_K)\, ds = l_i \right\},$$

for a certain function λ.

Since, for epidemics, the length of the infectious period of an individual is usually very small compared to its life length, we neglect births and removals that are not caused by the disease, as well as any age dependence of the infectivity or the susceptibility. Furthermore, the population is idealized to be homogeneously mixing. Despite these restrictions, there are many useful applications of the model (cf. [4] and [2]), and the process is still quite simple. This last fact helps one to understand the underlying method, which is an application of Stein's method for proving convergence of stochastic processes (cf. [3] and [9]).

In special cases, there are already some asymptotic results for the proportion of susceptible and infectious individuals. However, the previous results were obtained for cases where l_i and r_i are independent and where the transition behavior is "Markovian," that is, $\mathscr{L}(l_1) = \exp(1)$. This case, in the special form $\lambda(t, x) = \lambda(x(t))$, was analyzed by Wang [14, 15]. For general λ, Solomon [13] has discussed a related, age-dependent population model that deals only with one class of individuals. The very special case $\lambda(t, x) = x(t)$ and $(\hat{r}_i), (r_i)$ being i.i.d. $\exp(\rho)$ yields the classical GSE, as constructed by Sellke [11].

In this paper, we not only discuss a more general model, but also describe the asymptotic evolution in a more detailed form. We investigate, for $K \to \infty$, the empirical measure

$$\xi_K = \frac{1}{K} \sum_{i=1}^{aK} \delta_{(0,\hat{r}_i)} + \frac{1}{K} \sum_{i=1}^{bK} \delta_{(A_i^K, A_i^K + r_i)},$$

considered as a substochastic measure on $[0,\infty)^2$, where the half-open interval $[a, b) \subset [0,\infty)$ is represented by the point $(a, b) \in [0,\infty)^2$. (In general, δ_a shall denote the Dirac measure at the point a.) In this way we obtain the asymptotic average path behavior. Usually, in epidemic models, the proportion of infected individuals and the proportion of susceptible individuals are investigated. These quantities can easily be reconstructed via ξ_K. For instance, if $t \geq 0$,

$$\xi_K([0,t] \times (t,\infty)) = \frac{1}{K}\sum_{i=1}^{aK} 1_{[0,\hat{r}_i)}(t) + \frac{1}{K}\sum_{i=1}^{bK} 1_{[A_i^K, A_i^K + r_i)}(t) =: I_K(t)$$

describes the proportion infected at time t. Moreover, we can also investigate quantities like

$$\xi_K([0,s] \times (t,\infty]), \qquad t > s,$$

giving the proportion of individuals that were infected before time s and are not removed before time t, that is, the infectivity at time t in the population resulting from individuals that were infected before time s. Thus we gain new insights concerning the behavior of the epidemic.

In Section 1, the results on the asymptotic behavior of the GSE are presented, a heuristic argument is given to explain how the results are obtained and some applications are indicated. Section 2 concerns the connection with "classical" results for the GSE, as given by Wang [14]. Finally, Section 3 contains the proofs.

1. Results and heuristics.

1.1. *Assumptions.* As described in the Introduction, let $(l_i, r_i)_{i \in \mathbf{N}}$ be a family of positive i.i.d. random vectors, let Ψ be the common distribution function of the $(l_i)_{i\in\mathbf{N}}$, let Φ be the common distribution function of the $(r_i)_{i\in\mathbf{N}}$, let $(\hat{r}_i)_{i\in\mathbf{N}}$ be a family of positive, independent random variables with distribution functions $(\hat{\Phi}_i)_{i\in\mathbf{N}}$ and assume the $(l_i, r_i)_{i\in\mathbf{N}}, (\hat{r}_i)_{i\in\mathbf{N}}$ to be mutually independent (whereas, for each fixed i, l_i and r_i may be dependent). Let $D_+ = \{x\colon [0,\infty) \to [-1,1]$ right continuous with left-hand limits$\}$ and let $\lambda\colon \mathbf{R}_+ \times D_+ \to \mathbf{R}_+$ be the "accumulation" function. Then, for an initially susceptible individual i, its infection time A_i^K is given by

$$A_i^K = \inf\left\{t \in \mathbf{R}_+ \colon \int_{(0,t]} \lambda(s, Z_K)\, ds = l_i\right\}, \tag{1}$$

with

$$Z_K(t) = \frac{1}{K}\sum_{j=1}^{aK} 1_{[0,\hat{r}_j)}(t) + \frac{1}{K}\sum_{j=1}^{bK} 1_{[A_j^K, A_j^K + r_j)}(t)$$

being the proportion of infected individuals present in the population at time $t \in \mathbf{R}_+$. [We use the notation $1_C(t)$ to denote the indicator function on the set C. The notation $I[t \in C]$ refers to the indicator of a set, not considered as a function.] This gives a recursive definition of the A_i^K's: If $l_{(j)}$ is the jth order statistic of $l_1, \dots, l_{bK}$, corresponding to the individual i_j, say, then

$$A_{i_j}^K = \inf\left\{ t \in \mathbf{R}_+ : \int_{(0,t]} \lambda\left(s, \frac{1}{K} \sum_{m=1}^{aK} 1_{[0, \hat{r}_m)} + \frac{1}{K} \sum_{k=1}^{j-1} 1_{[A_{i_k}^K, A_{i_k}^K + r_{i_k})} \right) ds = l_{(j)} \right\}.$$

This completes the description of the model. Furthermore, we make some technical assumptions.

1. There is a probability measure $\hat{\mu}$ on $\mathbf{R}_+$ such that, for all $T \in \mathbf{R}_+$,

$$\sup_{0 \le t \le T} \left| \frac{1}{aK} \sum_{i=1}^{aK} \mathbf{P}[\hat{r}_i \le t] - \hat{\mu}([0,t]) \right| \to 0 \qquad (K \to \infty).$$

   Denote its distribution function by $\hat{\Phi}$.
2. The function λ: $\mathbf{R}_+ \times D_+ \to \mathbf{R}_+$ satisfies, for all $t \in \mathbf{R}_+$, $x, y \in D_+$:
   (a) $\lambda(t, x) = \lambda(t, x_t)$, where, for $t, u \in \mathbf{R}_+$, $x \in D_+$, $x_t(u) = x(t \wedge u)$.
   (b) There is a positive constant α such that

$$|\lambda(t,x) - \lambda(t,y)| \le \alpha \sup_{0 \le s \le t} |x(s) - y(s)|.$$

   (c) There is a positive constant γ such that $\sup_{0 \le s \le t} \lambda(s, x) \le \gamma$.
3. There is a positive constant β such that, for each $x \in \mathbf{R}_+$, $\Psi_x(t) := \mathbf{P}[l_1 \le t | r_1 = x]$ satisfies, for all $s, t \in \mathbf{R}_+$,

$$|\Psi_x(t) - \Psi_x(s)| \le \beta |t - s|.$$

The basic tool to obtain the desired convergence result is Stein's method in the form of Theorem 4.5, combined with Proposition 4.9, in Reinert [9]. This is a fairly obvious generalization of Corollary 2.6 in Reinert [10]. Let $E \subset \mathbf{R}^2$ be a locally compact Hausdorff space with a countable basis. Denote by $M^b(E)$ the space of bounded Radon measures on E and by $M_1(E)$ the subspace of all positive Radon measures on E with total mass less than or equal to 1, and for $\phi \in C_c(E)$, $\nu \in M^b(E)$, $\langle \nu, \phi \rangle = \int \phi \, d\nu$.

PROPOSITION 1.1. *Let $(\eta_K)_{K \in \mathbf{N}}$ be a family of random elements with values in $M_1(E)$ and let $\mu \in M_1(E)$. If for all $m \in \mathbf{N}$, $f \in C_b^\infty(\mathbf{R}^m)$, $\phi_1, \dots, \phi_m \in C_b^\infty(E)$, we have for all $\phi \in C_b^\infty(E)$,*

$$\text{(2)} \qquad \mathbf{E}[f(\langle \eta_K, \phi_1 \rangle, \dots, \langle \eta_K, \phi_m \rangle) \langle \mu - \eta_K, \phi \rangle] \to 0 \qquad (K \to \infty),$$

then

$$\mathscr{L}(\eta_K) \Rightarrow_W \delta_\mu \qquad (K \to \infty).$$

1.2. *Results.* We now want to apply this proposition. First, in Theorems 1.2 and 1.3, we find a candidate for the measure μ. In general, the natural choice would be $\mu = \lim_{K\to\infty} \mathbf{E}[\eta_K]$, where the limit is taken in the vague topology. Theorem 1.4 then verifies the convergence in (2).

For the required measure μ, observe that during the course of the epidemic not necessarily (hopefully) every susceptible will get infected; $A_i^K = \infty$ for some i is possible. Therefore, if such a μ exists, it will in general not be a probability measure but a positive measure with total mass less than or equal to 1. Furthermore, as the existence of $\mathbf{E}r_i$ or $\mathbf{E}\hat{r}_i$, $i \in \mathbf{N}$, is not assumed, we restrict the observations to finite intervals $[0,T] \times [0,T]$ for a $T \in \mathbf{R}_+$ arbitrary and fixed. This leads to some notation. For $T \in \mathbf{R}_+$, put $[0,T]^2 = [0,T] \times [0,T]$ and $\mathscr{B}_T = \mathscr{B}([0,T]^2)$. Let $\nu \in M_1(\mathbb{R}_+^2)$. Then

$$\nu^T = \nu|_{\mathscr{B}_T}$$

is the restriction of ν on $\mathscr{B}_T$ [hence, $\nu^T \in M_1([0,T]^2)$]. For $A \in \mathscr{B}(\mathbf{R}^2)$, put

$$\nu^T(A) = \nu\left(A \cap [0,T]^2\right);$$

this defines ν^T also on $\mathscr{B}(\mathbf{R}^2)$. If in addition X is a random element with $\mathscr{L}(X) = \nu$, then, for all $T \in \mathbf{R}_+$, $f \in L_1(\nu)$, $A \in \mathscr{B}(\mathbf{R}^2)$,

$$\mathbf{E}^T f(X) = \int f(x)\,\nu^T(dx),$$

$$\mathbf{P}^T[X \in A] = \int 1_A(x)\,\nu^T(dx),$$

$$\mathscr{L}^T f(X) = \mathscr{L}(f(X))|_{\mathscr{B}_T}$$

are the corresponding restrictions. Our aim is to show a weak law of large numbers (w.l.l.n.) type of result for ξ_K. For that purpose, we define, for $f \in C(\mathbf{R}_+, \mathbf{R})$, $t \in \mathbf{R}_+$, an operator $\mathscr{Z}$ and an operator L:

$$\begin{aligned} &\mathscr{Z}f(t) = a\left(1 - \hat{\Phi}(t)\right) + b\Psi(f(t)) - b\int_{(0,t]} \Psi_x(f(t-x))\mathbf{P}[r_1 \in dx], \\ (3)\quad & \\ &Lf(t) = \int_{(0,t]} \lambda(s, \mathscr{Z}f)\,ds \end{aligned}$$

(as $\mathscr{Z}f \in D_+$, the latter expression is defined). Let $\|f\|_T = \sup_{s\le T}|f(s)|$ denote the supremum norm on $C([0,T])$. Then we can prove the following results.

THEOREM 1.2. *For $T \in \mathbf{R}_+$, the equation*

$$f(t) = \int_{(0,t]} \lambda(s, \mathscr{Z}f)\,ds, \qquad 0 \le t \le T, \tag{4}$$

has a unique solution G_T. This solution can be obtained by an iteration procedure: Choose an arbitrary $f_0 \in C([0,T])$ and put $f_1 = Lf_0$, $f_n = Lf_{n-1}$ for $n \in \mathbf{N}$. Then

$$\|f_n - G_T\|_T \le \frac{(b/2)^n}{1 - b/2}\,\frac{\left(1 + 4\alpha\beta T(n+1)\right)^{T/\eta+2} - 1}{4\alpha\beta T(n+1)}\|f_0 - Lf_0\|_T,$$

where

$$\eta = \sup\left\{t \le T: \int_{(0,T]} (1 + \Phi(s))\, ds \le \frac{1}{2\alpha\beta}\right\}.$$

THEOREM 1.3. *For $T \in \mathbf{R}_+$, let G_T be the solution of* (4) *and $\tilde{\mu}^T \in M_1(\mathbb{R}_+^2)$ be given for $r, s \in (0, T]$ by*

$$\tilde{\mu}^T([0,r] \times [0,s]) = \int_{(0,(s-r)\vee 0]} \Psi_x(G_T(r))\mathbf{P}^T[r_1 \in dx] + \int_{((s-r)\vee 0, s]} \Psi_x(G_T(s-x))\mathbf{P}^T[r_1 \in dx].$$

Put

$$\mu^T = a(\delta_0 \times \hat{\mu})^T + b\tilde{\mu}^T.$$

Then

$$\frac{1}{K}\sum_{i=1}^{aK} \mathscr{L}^T((0, \hat{r}_i)) + \frac{1}{K}\sum_{i=1}^{bK} \mathscr{L}^T\big((A_i^K, A_i^K + r_i)\big) \Rightarrow_v \mu^T \qquad (K \to \infty).$$

Note that it might be more intuitive to think of $\tilde{\mu}^T([0,r] \times [0,s])$ in the form

$$\tilde{\mu}^T([0,r] \times [0,s]) = \mathbf{P}^T[l_1 \le G_T(r), l_1 \le G_T(s - r_1)].$$

THEOREM 1.4. *Let μ^T be as in Theorem* 1.3. *Then, for all $T \in \mathbf{R}_+$,*

$$\mathscr{L}(\xi_K^T) \Rightarrow_w \delta_{\mu^T} \qquad (K \to \infty).$$

1.3. *Heuristics.* The problem consists essentially of approximating the average distribution of the infective periods $(1/K)\sum_{i=1}^{bK}\mathscr{L}((A_i^K, A_i^K + r_i))$, as it reflects the dependence structure of the process. This distribution is determined by the averages $(1/bK)\sum_{i=1}^{bK}\mathbf{P}[A_i^K \le s,\ A_i^K + r_i \le t]$, $s, t \in \mathbf{R}_+$. For $s, t \in \mathbf{R}_+$, let

$$H_K(s,t) = \frac{1}{bK}\#\{i \le bK: A_i^K \le s,\ A_i^K + r_i \le t\},$$

$$H(s,t) = \lim_{K\to\infty} H_K(s,t),$$

$$h(s,t) = \frac{\partial}{\partial s}\frac{\partial}{\partial t}H(s,t),$$

and suppose that H and h exist. Then we have, for all $s > 0$,

$$\frac{1}{bK}\sum_{i=1}^{bK} 1_{[A_i^K, A_i^K + r_i)}(s) \approx H(s,\infty) - H(s,s).$$

Furthermore, by our first assumption, we have, for all $s > 0$,

$$\frac{1}{aK}\sum_{i=1}^{aK} 1_{[0,\hat{r}_i)}(s) \approx 1 - \hat{\Phi}(s).$$

Thus,

$$Z_K(s) \approx a\big(1 - \hat{\Phi}(s)\big) + b\big(H(s,\infty) - H(s,s)\big) =: Z(s). \tag{5}$$

Let

$$\tilde{G}(t) = \int_{(0,t]} \lambda(s, Z)\,ds.$$

Then, from (1) and (5), we have $A_i^K \approx \tilde{G}^{-1}(l_i)$, where $f^{-1}(y) = \inf\{x\colon f(x) = y\}$, and $\inf \varnothing = \infty$. Applying the s.l.l.n.,

$$\begin{aligned} H(s,t) &\approx \frac{1}{bK}\#\big\{i \le bK\colon \tilde{G}^{-1}(l_i) \le s, \tilde{G}^{-1}(l_i) + r_i \le t\big\} \\ &\approx \mathbf{P}\big[l_i \le \tilde{G}(s), l_i \le \tilde{G}(t - r_i)\big] \\ &= \int_{(0,(t-s)\vee 0]} \Psi_x\big(\tilde{G}(s)\big)\mathbf{P}[r_1 \in dx] \\ &\quad + \int_{((t-s)\vee 0, t]} \Psi_x\big(\tilde{G}(t-x)\big)\mathbf{P}[r_1 \in dx]. \end{aligned}$$

Hence, recalling (3), $Z \approx \mathscr{Z}\tilde{G}$. This motivates the integral equation (4) as a way of solving for $\tilde{G}$ and thus for Z also. Intuitively speaking, $a(1 - \hat{\Phi}(s))$ describes the proportion of initially infected that are at time s not yet removed, $b\Psi(G_T(s))$ describes the proportion of new infected individuals up to time s and $b\int_{(0,s]}\Psi_x(G_T(s-x))\mathbf{P}[r_1 \in dx]$ describes the proportion of new infected individuals that are already removed at time s.

The proof of Theorem 1.2 is technical, the basic procedure being much the same as for proving the Picard–Lindelöf theorem (see, e.g., [1], pages 104–106). Observe, however, that the Picard–Lindelöf theorem cannot be applied directly, because we do not yet know whether $\lambda(s, \mathscr{Z}f)$, as a function of f, satisfies a Lipschitz condition; this was only assumed to hold for $\lambda(s, f)$. Thus, we still have to prove a contraction property for $\lambda(s, \mathscr{Z}f)$. Then, as for the Picard–Lindelöf theorem, the contraction theorem gives both the uniqueness and the claimed iteration procedure and thus finishes the proof.

Once Theorem 1.2 is established, for Theorem 1.3 we can reason as follows. For $T \in \mathbf{R}_+$ fixed and $B_1 = [u_1, u_2] \times [0, v] \in \mathscr{B}(\mathbf{R}_+^2)$,

$$\begin{aligned} \frac{1}{aK}\sum_{i=1}^{aK}\mathbf{P}^T\big[(0,\hat{r}_i) \in B_1\big] &= \frac{1}{aK}\sum_{i=1}^{aK}\mathbf{P}^T\big[0 \in [u_1,u_2], \hat{r}_i \in [0,v]\big] \\ &\approx (\delta_0 \times \hat{\mu})^T(B_1). \end{aligned}$$

For $B_2 = [0,u] \times [0,v] \in \mathscr{B}(\mathbf{R}_+^2)$, we have, as calculated above,

$$\begin{aligned}\frac{1}{bK}\sum_{i=1}^{bK}\mathbf{P}^T\big[\big(A_i^K, A_i^K + r_i\big) \in B_2\big] &\approx \int_{(0,(v-u)\vee 0]}\Psi_x\big(\tilde{G}(u)\big)\mathbf{P}^T[r_1 \in dx]\\ &\quad + \int_{((v-u)\vee 0, v]}\Psi_x\big(\tilde{G}(v-x)\big)\mathbf{P}^T[r_1 \in dx]\\ &= \tilde{\mu}^T(B_2),\end{aligned}$$

and Theorem 1.3 follows by making the approximations rigorous.

Now the third theorem can be obtained with help of Proposition 1.1. Let $T \in \mathbf{R}_+$ and $f \in C_b^\infty(\mathbf{R})$, $\phi, \psi \in C_b^\infty([0,T]^2)$. Then, for the above μ^T,

$$\begin{aligned}&\mathbf{E}\Bigg[f\Bigg(\Bigg\langle\frac{1}{K}\sum_{i=1}^{aK}\delta_{(0,\hat{r}_i)}^T + \frac{1}{K}\sum_{i=1}^{bK}\delta_{(A_i^K, A_i^K + r_i)}^T, \phi\Bigg\rangle\Bigg)\\ &\quad\times\Bigg\langle\mu^T - \Bigg(\frac{1}{K}\sum_{i=1}^{aK}\delta_{(0,\hat{r}_i)}^T + \frac{1}{K}\sum_{i=1}^{bK}\delta_{(A_i^K, A_i^K + r_i)}^T\Bigg), \psi\Bigg\rangle\Bigg]\\ &\approx \mathbf{E}\Bigg[f\Bigg(\Bigg\langle a(\delta_0 \times \hat{\mu})^T + b\int_{(0,\infty)}\int_{(0,\infty)}h(u, u+v)\,\delta_{(u,u+v)}^T\,dv\,du, \phi\Bigg\rangle\Bigg)\\ &\qquad\times\Bigg\langle\mu^T - \Bigg(\frac{1}{K}\sum_{i=1}^{aK}\delta_{(0,\hat{r}_i)}^T + \frac{1}{K}\sum_{i=1}^{bK}\delta_{(A_i^K, A_i^K + r_i)}^T\Bigg), \psi\Bigg\rangle\Bigg]\\ &\approx f\Bigg(\Bigg\langle a(\delta_0 \times \hat{\mu})^T + b\int_{(0,\infty)}\int_{(0,\infty)}h(u,u+v)\,\delta_{(u,u+v)}^T\,dv\,du, \phi\Bigg\rangle\Bigg)\\ &\quad\times\Bigg\langle\tilde{\mu}^T - b\int_{(0,\infty)}\int_{(0,\infty)}h(u,u+v)\,\delta_{(u,u+v)}^T\,dv\,du, \psi\Bigg\rangle,\end{aligned}$$

and the last term is equal to 0 due to Theorem 1.3. With Proposition 1.1, this would prove the assertion. In the proof, however, as the existence of h is not ensured, we will approximate $\delta^2 h(u,v)$ by $H(u+\delta, v+\delta) - H(u+\delta, v) - H(u, v+\delta) + H(u,v)$, and we will have to restrict ourselves on $\mathscr{B}_T$.

1.4. *Some applications.* Typically, in epidemic models, the proportion of infected individuals and the proportion of susceptible individuals are investigated. As mentioned in the Introduction, these quantities can be reconstructed via ξ_K. For $t > 0$,

$$\xi_K([0,t] \times (t,\infty)) =: I_K(t)$$

gives the proportion of infected at time t,

$$\xi_K((t,\infty] \times [0,\infty]) = \frac{1}{K}\sum_{i=1}^{bK} I[A_i^K > t] =: S_K(t)$$

gives the proportion of susceptibles at time t and

$$\xi_K([0,t] \times [0,t]) =: R_K(t)$$

is the proportion of removed at time t. We can determine the limiting behavior of these quantities, since from Theorem 4.2 of Kellenberg ([7], page 32) we obtain the following lemma.

LEMMA 1.5. *Let $T \in \mathbf{R}_+$ and $\Gamma \in \mathscr{B}(\mathbf{R}_+^2)$ be such that $\mu^T(\partial\Gamma) = 0$. Then*

$$\xi_K^T(\Gamma) \to_{\mathbf{P}} \mu^T(\Gamma) \qquad (K \to \infty).$$

Furthermore, $\xi_K([0,\infty] \times [0,\infty]) = 1$. By Theorem 1.3, (16) below and this lemma we hence have, for all $t \in \mathbf{R}_+$ with $\hat{\mu}(\{t\}) = 0$ and for all $T > t$ with $\mathbf{P}[r_1 = T - t] = 0$,

$$\begin{aligned} R(t) &= \mathbf{P}\text{-}\lim_{K\to\infty} R_K(t) \\ &= a\hat{\Phi}(t) + \frac{b}{c}\int_{(0,t]} \Psi_x(G_T(t-x))\mathbf{P}^T[r_1 \in dx], \end{aligned} \tag{6}$$

$$\begin{aligned} S(t) &= \mathbf{P}\text{-}\lim_{K\to\infty} S_K(t) \\ &= b(1 - \Psi(G_T(t))), \end{aligned}$$

$$\begin{aligned} I(t) &= \mathbf{P}\text{-}\lim_{K\to\infty} I_K(t) \\ &= a(1 - \hat{\Phi}(t)) \\ &\quad + b\left(\Psi(G_T(t)) - \int_{(0,t]} \Psi_x(G_T(t-x))\mathbf{P}^T[r_1 \in dx]\right). \end{aligned} \tag{7}$$

In this way classical quantities like the total size and the maximum size of the epidemic can easily be determined. Moreover, our results provide additional information about the epidemic. Suppose, for example, that an epidemic is known to be taking place in a region, and that after some time t_0 every remaining susceptible in that region is immunized. Thus there are no new cases, although infectives may still be present. To decide at what time the region, which was probably put under quarantine, can be opened to the public again, we are interested in estimating the remaining infectivity in the population at times $s > t_0$. This is given by $\xi_K([0,t_0] \times (s,\infty))$, which is the proportion of individuals that were infected before the time t_0 and are still present in the population at time s. For large K and T,

$$\begin{aligned} &\xi_K([0,t_0] \times (s,\infty)) \\ &\quad \approx a(1 - \hat{\Phi}(s)) + b\left(\Psi(G_T(t_0)) - \int_{(0,s]} \Psi_x(G_T(s-x))\mathbf{P}^T[r_1 \in dx]\right). \end{aligned}$$

Thus, as soon as this expression is smaller than a certain critical level, we can abandon the isolation.

2. Comparison with known results. Our asymptotic results can be compared with those obtained by Wang [14], which seem to be the most general ones for the GSE known so far. As is made more explicit in [15], Wang [14] considers a population of total size $N = K$ and, in our notation, made the following assumptions:

1. $\mathbf{P}$[a particular susceptible individual becomes infected during the time interval $[t, t + \delta t]] = \lambda(I_N(t))\,\Delta t + o(\Delta t)$, for a function λ that is positive, bounded and Lipschitz on $[0,1]$, and $\lambda(0) = 0$.
2. $\mathbf{P}$[a particular infected individual stays infected for at least a period of length $t] = F(t)$, for a function F with $F(0) = 1$ and $F(t) \searrow 0$ as $t \to \infty$.
3. At time 0 there are $NI_N(0) = x(N)$ infected individuals $s_1, \dots, s_{x(N)}$ present, where s_i represents also the total time that the ith individual has been infected up to time 0. There is assumed to exist a positive density $q \in L_1(\mathbf{R}_+)$ such that, for all $s \in \mathbf{R}_+$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{x(N)} 1_{[0,s]}(s_i) = \int_0^s q(u)\,du.$$

With

$$g(s,t) = \frac{F(t+s)}{F(s)} I[F(s) > 0],$$

$$\gamma(t) = \int_0^\infty g(s,t) q(s)\,ds,$$

Wang proves that $(I_N(t), I_N(t) + R_N(t))$ converges to the unique positive solution $(P(t), B(t))$ of the system

$$(8) \qquad \begin{aligned} P(t) &= \gamma(t) + \int_{(0,t]} \lambda(P(u))(1 - B(u))F(t-u)\,du, \\ B(t) &= P(0) + \int_{(0,t]} \lambda(P(u))(1 - B(u))\,du, \end{aligned}$$

in the sense that, for every $\varepsilon > 0$,

$$\lim_{N\to\infty} \mathbf{P}\left[\sup_{u\in[0,T]} |I_N(u) - P(u)| + |I_N(u) + R_N(u) - B(u)| > \varepsilon \right] = 0.$$

Now consider in our model the special case that the (l_i) have an exp(1) distribution, that l_i and r_i are independent for each i, that there are $s_i \in \mathbf{R}_+$, $i \in \mathbf{N}$, such that, with $\Phi(t) = 1 - F(t)$,

$$\mathbf{P}[\hat{r}_i > t] = \frac{1 - \Phi(t + s_i)}{1 - \Phi(s_i)}, \qquad i \in \mathbf{N},$$

that $\lambda(t, x) = \lambda(x(t))$ and that there is a positive density $q \in L_1(\mathbf{R}_+)$ with

$$\lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{aK} 1_{[0,s]}(s_i) = \int_0^s q(u)\, du.$$

Then the second assumption of Wang's model is obviously satisfied. Furthermore, by the lack of memory of exponentials,

$$\begin{aligned} &\mathbf{P}[\text{the initially susceptible } i \text{ gets infected in } [t, t + \Delta t]] \\ &\quad = \mathbf{P}\big[A_i^K \in [t, t + \Delta t] \,\big|\, A_i^K \geq t\big] \\ &\quad = \lambda(I_K(t))\, \Delta t + o(\Delta t). \end{aligned}$$

Thus, Wang's first assumption is also fulfilled. To see that the restrictions on $\hat{r}_i$ are a special case of our restrictions, observe that

$$\begin{aligned} \frac{1}{K} \sum_{i=1}^{aK} \mathbf{P}[\hat{r}_i > t] &= \frac{1}{K} \sum_{i=1}^{aK} \frac{1 - \Phi(t + s_i)}{1 - \Phi(s_i)} \\ &\to \int \frac{1 - \Phi(t + x)}{1 - \Phi(x)} q(x)\, dx \\ &=: a\hat{\mu}([t, \infty)). \end{aligned}$$

Thus, $\hat{\mu}$ is a probability measure on $\mathbf{R}_+$ and $1 - \hat{\mu}([t, \infty)) =: \hat{\Phi}(t)$ is a continuous distribution function, as the set of discontinuity points of Φ has Lebesgue measure 0. Therefore, the convergence of $\hat{\Phi}_K(t) := (1/aK)\sum_{i=1}^{aK} \mathbf{P}[\hat{r}_i \leq t]$ to $\hat{\Phi}(t)$ is uniform (cf. [5], page 265, Lemma 3). Thus, Wang's model is a special case of the one we consider.

To see that in this special case the resulting expressions coincide, it suffices to prove that $(I(t), I(t) + R(t))$ satisfy (8). (In what follows we suppress the subscript T and the superscript T.) In view of (7) and (6), we have to show that, for $\gamma(t) = a(1 - \hat{\Phi}(t))$ and $\Psi(x) = 1 - e^{-x}$,

$$\begin{aligned} a + b(1 - e^{-G(t)}) &= a + b\int_{(0,t]} \lambda(I(s)) e^{-G(s)}\, ds, \\ I(t) &= \gamma(t) + b\int_{(0,t]} \lambda(I(s)) e^{-G(s)} (1 - \Phi(t - s))\, ds. \end{aligned} \tag{9}$$

Observe that $G'(t) = \lambda(I(t))$ by construction. Thus (9) is obviously satisfied (for the second equation apply integration by parts). This proves the coincidence of our deterministic approximation with that Wang obtained for this special case. However, our formulation covers a much wider class of models and gives much more detailed information about the process.

3. Proofs.

PROOF OF THEOREM 1.2. First observe that every solution f: $[0, T] \to \mathbf{R}$ of (4) is continuous, as for all $s, t \in [0, T]$,

$$|f(t) - f(s)| = \left|\int_{(s,t]} \lambda(u, \mathscr{Z}f)\, du\right| \le \gamma|t - s|.$$

By the same argument, $Lf \in C([0, T])$. Furthermore, $(C([0, T]), \|\cdot\|_T)$ is a complete metric space. In what follows, we basically proceed as in the proof of the Picard–Lindelöf theorem. That is, we show that Lf is a contraction on small intervals and then employ the contraction theorem. This gives us unique solutions of (4) and the iteration procedure on small intervals; their composition proves the assertion.

Let $f, g \in C([0, T])$. Then, for all $t \in \mathbf{R}_+$,

$$\begin{aligned}
&|Lf(t) - Lg(t)|\\
&\quad \le \alpha \int_{(0,t]} \sup_{u \le s} |\mathscr{Z}f(u) - \mathscr{Z}g(u)|\, ds\\
&\quad = \alpha b \int_{(0,t]} \sup_{u \le s} \Bigg| \Psi(f(u)) - \Psi(g(u))\\
&\qquad\qquad + \int_{(0,u]} \{\Psi_x(g(u-x)) - \Psi_x(f(u-x))\}\\
&\qquad\qquad\qquad \times P[r_1 \in dx] \Bigg|\, ds \qquad (10)\\
&\quad \le \alpha\beta b \|f - g\|_t \int_{(0,t]} (1 + \Phi(s))\, ds.
\end{aligned}$$

Now put

$$\eta = \sup\left\{t \le T: \int_{(0,t]} (1 + \Phi(s))\, ds \le \frac{1}{2\alpha\beta}\right\}.$$

Then, for $t \le \eta$,

$$\|Lf - Lg\|_t \le \frac{b}{2}\|f - g\|_t.$$

Thus L is a contraction on $C([0, \eta])$. By the contraction theorem, (4) has a unique solution G_η on $C([0, \eta])$.

For $k \in \mathbf{N}$, define the space of continuous extensions on $C([0, (k+1)\eta])$ of a unique solution of (4) on $C([0, k\eta])$ as

$$C_\eta([0, (k+1)\eta]) = \{f \in C([0, (k+1)\eta]): Lf(x) = f(x) \text{ for } x \le k\eta\}.$$

Now we proceed by induction. Suppose L is a contraction on $C_\eta([0, k\eta])$. Then, for all $f, g \in C_\eta([0,(k+1)\eta])$,

$$Lf(t) - Lg(t) = 0, \qquad t \le k\eta,$$

and for $k\eta < t \le (k+1)\eta$, by (10),

$$\begin{aligned}
&|Lf(t) - Lg(t)|\\
&\quad \le \alpha b \left\{ \int_{(k\eta, t]} \sup_{k\eta < u \le s} \left| \Psi(f(u)) - \Psi(g(u)) \right.\right.\\
&\qquad \left.\left. + \int_{(0, u-k\eta]} \{\Psi_x(g(u-x)) - \Psi_x(f(u-x))\} \mathbf{P}[r_1 \in dx] \right| ds \right\} \qquad (11)\\
&\quad \le \alpha\beta b \|f - g\|_t \int_{(k\eta, t]} (1 + \Phi(s - k\eta))\, ds\\
&\quad \le \frac{b}{2} \|f - g\|_t,
\end{aligned}$$

which proves the contraction property on every $C_\eta([0, k\eta])$. Let $k \in \mathbf{N}$ be such that $k\eta < T \le (k+1)\eta$. Then it follows that (4) has a unique solution G_T on $C_\eta([0, T])$. Because every solution of (4) must be in $C_\eta([0, T])$, we have also uniqueness on $C([0, T])$.

The contraction theorem also gives us the following iteration procedure, for $k \in \mathbf{N}$. Choose an arbitrary $f_0 \in C_\eta([0,(k+1)\eta])$ and put $f_1 = Lf_0$, $f_n = Lf_{n-1}$, for $n \in \mathbf{N}$. Then, if $G_{(k+1)\eta}$ is the unique solution on $C_\eta([0,(k+1)\eta])$,

$$\|f_n - G_T\|_{(k+1)\eta} \le \frac{(b/2)^n}{1 - b/2} \|f_0 - Lf_0\|_{(k+1)\eta}.$$

For the claimed iteration procedure, choose an $f_0 \in C([0, T])$. Let $f_1 = Lf_0$ and $f_n = Lf_{n-1}$, for $n \in \mathbf{N}$. Due to the contraction theorem, we have

$$\|f_n - G_T\|_\eta \le \frac{(b/2)^n}{1 - b/2} \|f_0 - Lf_0\|_\eta.$$

Furthermore, for $k \in \mathbf{N}$, to get an estimate on $\|f_n - G_T\|_{(k+1)\eta}$, put

$$g_0(s) = G_T(s) I[s \le k\eta] + f_0(s) I[s > k\eta].$$

Let $g_1 = Lg_0$ and $g_n = Lg_{n-1}$, for $n \in \mathbf{N}$. Then,

$$\begin{aligned}
\|f_n - G_T\|_{(k+1)\eta} &\le \|Lf_{n-1} - Lg_{n-1}\|_{(k+1)\eta} + \|g_n - G_T\|_{(k+1)\eta}\\
&\le \|Lf_{n-1} - Lg_{n-1}\|_{(k+1)\eta} + \frac{(b/2)^n}{1 - b/2} \|g_0 - Lg_0\|_{(k+1)\eta},
\end{aligned}$$

as $g_0 \in C_\eta([0,(k+1)\eta])$. For $s \le k\eta$, as $Lg_0(s) = LG_T(s) = G_T(s) = g_0(s)$,

$$\begin{aligned}\|g_0 - Lg_0\|_{(k+1)\eta} &= \sup_{s \in (k\eta,(k+1)\eta]} |f_0(s) - Lg_0(s)| \\ &\le \|f_0 - Lf_0\|_{(k+1)\eta} + \sup_{s \in (k\eta,(k+1)\eta]} |Lf_0(s) - Lg_0(s)|.\end{aligned}$$

Moreover,

$$\begin{aligned}\|Lf_{n-1} - Lg_{n-1}\|_{(k+1)\eta} \\ \le \|f_n - G_T\|_{k\eta} + \sup_{s \in (k\eta,(k+1)\eta]} |Lf_{n-1}(s) - Lg_{n-1}(s)|.\end{aligned}$$

Now, for $k\eta < t \le (k+1)\eta$, $t \le T$, we have by (10),

$$\begin{aligned}|Lf_n(t) - Lg_n(t)| \le \alpha\beta b\bigg\{&\int_{(0,t]} \|f_n - g_n\|_s \, ds \\ &+ \int_{(0,t]} \int_{(0,s]} |f_n(s-x) - g_n(s-x)| \mathbf{P}[r_1 \in dx] \, ds\bigg\}.\end{aligned}$$

Splitting up the integrals yields

$$\begin{aligned}&|Lf_n(t) - Lg_n(t)| \\ &\quad\le \alpha\beta b\bigg\{t\|f_n - g_n\|_{k\eta} + (t - k\eta) \sup_{k\eta < s \le t} |f_n(s) - g_n(s)| \\ &\qquad + \|f_n - g_n\|_{k\eta} \int_{(0,k\eta]} \Phi(s) \, ds \\ &\qquad + \|f_n - g_n\|_{k\eta} \int_{(k\eta,t]} (\Phi(s) - \Phi(s - k\eta)) \, ds \qquad (12)\\ &\qquad + \sup_{k\eta < s \le t} |f_n(s) - g_n(s)| \int_{(k\eta,t]} \Phi(s - k\eta) \, ds\bigg\} \\ &\quad\le \alpha\beta b 2T \|f_n - G_T\|_{k\eta} + \frac{b}{2} \sup_{k\eta < s \le t} |Lf_{n-1}(s) - Lg_{n-1}(s)|,\end{aligned}$$

where we employed the definition of η. For $n = 0$, we get by construction that for $k\eta < t \le (k+1)\eta$, $t \le T$,

$$|Lf_0(t) - Lg_0(t)| \le 2\alpha\beta b T \|f_0 - G_T\|_{k\eta}. \qquad (13)$$

Solving the recursion given by (12) and (13) yields

$$\sup_{k\eta < t \le (k+1)\eta} |Lf_n(t) - Lg_n(t)| \le 2\alpha\beta b T \sum_{i=0}^{n} \|f_{n-i} - G_T\|_{k\eta} \left(\frac{b}{2}\right)^i.$$

Combining these estimates, we obtain the recursion

$$\|f_n - G_T\|_{(k+1)\eta} \le \|f_n - G_T\|_{k\eta} + 2\alpha\beta bT \sum_{i=0}^{n-1} \|f_{n-1-i} - G_T\|_{k\eta}\left(\frac{b}{2}\right)^i$$
$$+ \frac{(b/2)^n}{1-b/2}\left(\|f_0 - Lf_0\|_{(k+1)\eta} + 2\alpha\beta bT\|f_0 - G_T\|_{k\eta}\right),$$
$$\|f_n - G_T\|_\eta \le \frac{(b/2)^n}{1-b/2}\|f_0 - Lf_0\|_\eta.$$

For an estimate on $\|f_n - G_T\|_{(k+1)\eta}$, suppose that, for a $c_k \in \mathbf{R}_+$,

$$\|f_n - G_T\|_{k\eta} \le \frac{(b/2)^n}{1-b/2}\|f_0 - Lf_0\|_{k\eta} c_k.$$

Then

$$\|f_n - G_T\|_{(k+1)\eta} \le \frac{(b/2)^n}{1-b/2}\|f_0 - Lf_0\|_{(k+1)\eta} c_{k+1},$$

where

$$c_{k+1} = c_k(1 + 4\alpha\beta T(n+1)) + 1.$$

Finally, solving this recursion on c_k with $c_0 = 1$, we get

$$\|f_n - G_T\|_T \le \frac{(b/2)^n}{1-b/2}\|f_0 - Lf_0\|_T \frac{(1 + 4\alpha\beta T(n+1))^{T/\eta+2} - 1}{4\alpha\beta T(n+1)}.$$

This proves the iteration procedure and thus the assertion. □

PROOF OF THEOREM 1.3. This proof consists, to a large extent, of justifying the heuristics. Let $T \in \mathbf{R}_+$ be fixed and let G_T be the unique solution of (4) in $C([0,T])$. Let $B_1 = [u_1, u_2] \times [0, v] \in \mathscr{B}(\mathbf{R}_+^2)$. Then, due to the assumptions,

$$\frac{1}{aK}\sum_{i=1}^{aK} \mathbf{P}^T[(0,\hat{r}_i) \in B_1] = \frac{1}{aK}\sum_{i=1}^{aK} \delta_0([u_1,u_2])\mathbf{P}^T[\hat{r}_i \le v]$$
$$\to (\delta_0 \times \hat{\mu})^T(B_1).$$

Thus,

$$\frac{1}{K}\sum_{i=1}^{aK} \mathscr{L}^T((0,\hat{r}_i)) \Rightarrow_w a(\delta_0 \times \hat{\mu})^T.$$

For the second sum $(1/K)\sum_{i=1}^{bK}\mathscr{L}^T((A_i^K, A_i^K + r_i))$, consider $B_2 = [0, u] \times [0, v] \in \mathscr{B}(\mathbf{R}_+^2)$. Then, as by symmetry all A_i^K and r_i have the same distribution,

$$\begin{aligned}\frac{1}{bK}\sum_{i=1}^{bK}\mathbf{P}^T\big[(A_i^K, A_i^K + r_i) \in B_2\big] &= \int_{\mathbf{R}_+}\mathbf{P}^T\big[A_1^K \le u,\, A_1^K + r_1 \le v \mid r_1 = x\big]\mathbf{P}[r_1 \in dx] \\ &= \int_{(0,(v-u)\vee 0]}\Psi_x(G_T(u))\mathbf{P}^T[r_1 \in dx] \\ &\quad + \int_{((v-u)\vee 0, v]}\Psi_x(G_T(v-x))\mathbf{P}^T[r_1 \in dx] + R_1 + R_2,\end{aligned}$$

where

$$R_1 = \int_{(0,(v-u)\vee 0]}\big\{\mathbf{P}^T\big[A_1^K \le u \mid r_1 = x\big] - \Psi_x(G_T(u))\big\}\mathbf{P}^T[r_1 \in dx],$$

$$R_2 = \int_{((v-u)\vee 0, v]}\big\{\mathbf{P}^T\big[A_1^K \le v - x \mid r_1 = x\big] - \Psi_x(G_T(v-x))\big\}\mathbf{P}^T[r_1 \in dx].$$

Thus, if we can show that, for all $T \in \mathbf{R}_+$,

$$\sup_{0\le t\le T}\int_{(0,T]}\big|\mathbf{P}^T\big[A_1^K \le t \mid r_1 = x\big] - \Psi_x(G_T(t))\big|\mathbf{P}^T[r_1 \in dx] \to 0 \qquad (K \to \infty), \tag{14}$$

it would follow that $R_1 \to 0$, $R_2 \to 0$ $(K \to \infty)$ and thus

$$\frac{1}{bK}\sum_{i=1}^{bK}\mathbf{P}^T\big[(A_i^K, A_i^K + r_i) \in B_2\big] \to \tilde{\mu}^T(B_2) \qquad (K \to \infty).$$

This would establish the assertion. Thus it is sufficient to prove (14).

Proof of (14). For $t \le T \in \mathbf{R}_+$ fixed, let

$$F_K(t) = \int_{(0,t]}(\lambda(s, Z_K))\,ds.$$

Then $F_K \in C([0, T])$ and is nondecreasing and

$$A_i^K = F_K^{-1}(l_i).$$

Thus, for all $x \in \mathbf{R}_+$,

$$\begin{aligned}&\mathbf{P}^T\big[A_1^K \le t \mid r_1 = x\big] - \Psi_x(G_T(t)) \\ &\quad = \mathbf{P}^T\big[l_1 \le F_K(t) \mid r_1 = x\big] - \mathbf{P}^T\big[l_1 \le G_T(t) \mid r_1 = x\big],\end{aligned}$$

and therefore

$$\begin{aligned}
&\int_{(0,T]} \left|\mathbf{P}^T[A_1^K \le t | r_1 = x] - \Psi_x(G_T(t))\right| \mathbf{P}^T[r_1 \in dx] \\
&\quad \le \int_{(0,T]} \left\{\mathbf{P}^T[G_T(t) < l_1 \le F_K(t) | r_1 = x]\right. \\
&\qquad\qquad \left. + \mathbf{P}^T[F_K(t) < l_1 \le G_T(t) | r_1 = x]\right\} \mathbf{P}^T[r_1 \in dx] \\
&\quad \le \mathbf{P}^T[G_T(t) < l_1 \le F_K(t)] + \mathbf{P}^T[F_K(t) < l_1 \le G_T(t)].
\end{aligned}$$

As for all $\varepsilon > 0$,

$$\begin{aligned}
&\mathbf{P}[G_T(t) < l_1 \le F_K(t)] + \mathbf{P}[F_K(t) < l_1 \le G_T(t)] \\
&\quad \le \mathbf{P}[G_T(t) < l_1 \le G_T(t) + \varepsilon] + 2\mathbf{P}[|G_T(t) - F_K(t)| \ge \varepsilon] \\
&\qquad + \mathbf{P}[G_T(t) - \varepsilon < l_1 \le G_T(t)] \\
&\quad \le 2\beta\varepsilon + 2\mathbf{P}[\|G_T - F_K\|_T \ge \varepsilon],
\end{aligned} \tag{15}$$

it suffices for (14) to show that

$$\|G_T - F_K\|_T \to_{\mathbf{P}} 0 \qquad (K \to \infty). \tag{16}$$

□

PROOF OF (16). As G_T is characterized by being the fixed point of the operator L in the proof of Theorem 1.2, we look for a similar characterization for F_K. Define

$$\mathscr{H}_T = \left\{h: h(t) = \int_{(0,t]} \lambda(s, g)\, ds \text{ for a step function } g: [0, \infty) \to [0,1], t \le T\right\}.$$

Then, a.s. every realization of F_K is in $\mathscr{H}_T$. For $h \in \mathscr{H}_T$, put

$$\mathscr{Z}_K h(t) = \frac{1}{K}\sum_{i=1}^{aK} 1_{[0, \hat{r}_i)}(t) + \frac{1}{K}\sum_{i=1}^{bK} 1_{[h^{-1}(l_i), h^{-1}(l_i) + r_i)}(t),$$

$$L_K h(t) = \int_{(0,t]} \lambda(s, \mathscr{Z}_K h)\, ds.$$

Then, $L_K F_K = F_K$ a.s. by construction. Therefore, for all $t \le T$ and for a.s. every realization of F_K, we have

$$\begin{aligned}
|F_K(t) - G_T(t)| &= |L_K F_K(t) - L G_T(t)| \\
&\le \sup_{h \in \mathscr{H}_T} |L_K h(t) - L h(t)| + |L F_K(t) - L G_T(t)|.
\end{aligned}$$

Suppose for the moment that we have shown

$$\sup_{0 \le t \le T} \sup_{h \in \mathscr{H}_T} |L_K h(t) - L h(t)| \to_{\mathbf{P}} 0 \qquad (K \to \infty). \tag{17}$$

Then, for $\eta = \sup\{t \le T: \int_{(0,t]}(1 + \Phi(s))\,ds \le 1/(2\alpha\beta)\}$, we have by the contraction property of L on $C([0,\eta])$ (see the proof of Theorem 1.2)

$$\|G_T - F_K\|_\eta \le \left\| \sup_{h \in \mathscr{H}_T} |L_K h - Lh| \right\|_T + \frac{b}{2}\|G_T - F_K\|_\eta,$$

that is,

$$\|G_T - F_K\|_\eta \le \frac{1}{1 - b/2}\left\| \sup_{h \in \mathscr{H}_T} |L_K h - Lh| \right\|_T,$$

and therefore, with (17),

$$\|G_T - F_K\|_\eta \to_{\mathbf{P}} 0 \qquad (K \to \infty).$$

For $T > \eta$, we proceed as in the proof of Theorem 1.2. Suppose that for $k \in \mathbf{N}$ we have shown

$$\|G_T - F_K\|_{k\eta} \to_{\mathbf{P}} 0 \qquad (K \to \infty). \tag{18}$$

Then for $k\eta < T \le (k+1)\eta$, we have a.s.

$$\begin{aligned}
&|L_K F_K(t) - LG_T(t)| \\
&\quad \le \alpha b \beta \Big\{ k\eta \|G_T - F_K\|_{k\eta} + (t - k\eta) \sup_{k\eta < u \le t} |F_K(u) - G_T(u)| \\
&\qquad + \int_{(k\eta, t]} \sup_{k\eta < u \le t} |F_K(u) - G_T(u)| \Phi(s - k\eta)\,ds \\
&\qquad + \int_{(k\eta, t]} \|G_T - F_K\|_{k\eta} \sup_{u \le s} (\Phi(u) - \Phi(u - k\eta))\,ds \Big\} \\
&\quad \le \alpha\beta b \Big\{ \|G_T - F_K\|_{k\eta}(k+1)\eta \\
&\qquad + \sup_{k\eta < u \le t} |F_K(u) - G_T(u)| \int_{(k\eta, t]} (1 + \Phi(s))\,ds \Big\}.
\end{aligned}$$

where we employed (11) for the second inequality. Thus, as

$$\begin{aligned}
\|G_T - F_K\|_{(k+1)\eta} \le{}& \left\| \sup_{h \in \mathscr{H}_T} |L_K h - Lh| \right\|_T + \|G_T - F_K\|_{k\eta} \\
&+ \sup_{k\eta < u \le (k+1)\eta} |LF_K(u) - LG_T(u)|,
\end{aligned}$$

we would have

$$\begin{aligned}
\|G_T - F_K\|_{(k+1)\eta} &\le \frac{1}{1 - b/2}\left\{ \left\| \sup_{h \in \mathscr{H}_T} |L_K h - Lh| \right\|_T \right. \\
&\qquad \left. + \|G_T - F_K\|_{k\eta}\left(1 + \alpha\beta\frac{b}{2}(k+1)\eta\right)\right\} \\
&\to_{\mathbf{P}} 0 \qquad (K \to \infty)
\end{aligned}$$

due to (16) and (17). Thus, if (17) holds, it follows by induction that (18) is true for all $k \in \mathbf{N}$, and this proves (16). Therefore it is sufficient to prove (17). □

PROOF OF (17). The proof of (17) is based on a Glivenko–Cantelli argument. First observe that, for every $h \in \mathscr{H}_T$, $t \le T$ and a.s. every realization of L_K,

$$\begin{aligned} |L_K h(t) - Lh(t)| &\le \alpha T \sup_{s \le T} |\mathscr{Z}_K h(s) - \mathscr{Z} h(s)| \\ &\le \alpha T \sup_{s \le T} (a R_1(s) + b R_2(s)), \end{aligned}$$

where

$$\begin{aligned} R_1(t) &= \left| \frac{1}{aK} \sum_{i=1}^{aK} 1_{[0,\hat{r}_i)}(t) - (1 - \hat{\Phi}(t)) \right|, \\ R_2(t) &= \left| \frac{1}{bK} \sum_{i=1}^{bK} 1_{[h^{-1}(l_i), h^{-1}(l_i) + r_i)}(t) - \Psi(h(t)) \right. \\ &\qquad \left. + \int_{(0,t]} \Psi_x(h(t-x)) \mathbf{P}[r_1 \in dx] \right|. \end{aligned}$$

For $R_1(t)$ observe that

$$\begin{aligned} R_1(t) &\le \left| \frac{1}{aK} \sum_{i=1}^{aK} (I[\hat{r}_i > t] - \mathbf{P}[\hat{r}_i > t]) \right| \\ &\quad + \left| \frac{1}{aK} \sum_{i=1}^{aK} \mathbf{P}[\hat{r}_i > t] - \lim_{K \to \infty} \frac{1}{aK} \sum_{i=1}^{aK} \mathbf{P}[\hat{r}_i > t] \right|. \end{aligned}$$

The first summand tends to 0 ($K \to \infty$) uniformly in t due to the Glivenko–Cantelli theorem for the nonidentically distributed case (cf. [12]), and the second summand tends to 0 ($K \to \infty$) uniformly in t due to the assumption. Thus, independently of h,

$$\sup_{t \le T} R_1(t) \to 0 \quad \text{a.s.}\ (K \to \infty).$$

For $R_2(t)$ first observe that

$$R_2(t) = \left| \frac{1}{bK} \sum_{i=1}^{bK} 1_{[h^{-1}(l_i), h^{-1}(l_i) + r_i)}(t) - \mathbf{E} 1_{[h^{-1}(l_1), h^{-1}(l_1) + r_1)}(t) \right|.$$

We now use a Glivenko–Cantelli theorem that is closely related to Theorem II.2.2 of Pollard ([8], page 8).

PROPOSITION 3.1. *Suppose* $(\zeta_i)_{i \in \mathbf{N}} = (\zeta_i^1, \dots, \zeta_i^d)_{i \in \mathbf{N}}$ *are i.i.d.* $\mathbf{R}^d$*-valued random elements with common distribution function* $F(x) = \mathbf{P}[\zeta_1^j \le x_j,\ j =$

$1, \dots, d]$, $x \in \mathbf{R}^d$ *(for fixed* i, $\zeta_i^1, \dots, \zeta_i^d$ *need not be independent). Let* $\mathscr{F}$ *be a class of integrable functions* f: $\mathbf{R}^d \to \mathbf{R}$ *with the following approximation property. For all* $\varepsilon > 0$, *there is a finite set* $\mathscr{F}_\varepsilon$ *such that for all* $f \in \mathscr{F}$ *there are functions* $f_{\varepsilon, l}, f_{\varepsilon, u} \in \mathscr{F}_\varepsilon$ *with*

$$f_{\varepsilon,l}(x) \le f(x) \le f_{\varepsilon,u}(x), \qquad x \in \mathbf{R}^d, \quad \textit{and} \quad \mathbf{E}[f_{\varepsilon,u}(\zeta_1) - f_{\varepsilon,l}(\zeta_1)] \le \phi(\varepsilon),$$

for a function $\phi > 0$ *with* $\phi(\varepsilon) \to 0$ $(\varepsilon \to 0)$. *Then*

$$\sup_{f \in \mathscr{F}} \left| \frac{1}{n} \sum_{i=1}^n f(\zeta_i) - \mathbf{E} f(\zeta_1) \right| \to 0 \quad a.s.\ (n \to \infty).$$

PROOF. It is sufficient to show that

$$\liminf_{n \to \infty} \inf_{f \in \mathscr{F}} \left(\frac{1}{n} \sum_{i=1}^n f(\zeta_i) - \mathbf{E} f(\zeta_1) \right) \ge 0$$

and

$$\limsup_{n \to \infty} \sup_{f \in \mathscr{F}} \left(\mathbf{E} f(\zeta_1) - \frac{1}{n} \sum_{i=1}^n f(\zeta_i) \right) \ge 0.$$

Let $\varepsilon > 0$ be fixed. For $f \in \mathscr{F}$ choose $f_\varepsilon \in \mathscr{F}_\varepsilon$ such that $f_\varepsilon \le f$ and $\mathbf{E} f_\varepsilon(\zeta_1) \ge \mathbf{E} f(\zeta_1) - \phi(\varepsilon)$. Then

$$\begin{aligned}
&\liminf_{n \to \infty} \inf_{f \in \mathscr{F}} \left(\frac{1}{n} \sum_{i=1}^n f(\zeta_i) - \mathbf{E} f(\zeta_1) \right) \\
&\quad \ge \liminf_{n \to \infty} \inf_{f \in \mathscr{F}} \left(\frac{1}{n} \sum_{i=1}^n f_\varepsilon(\zeta_i) - \mathbf{E} f(\zeta_1) \right) \\
&\quad \ge \liminf_{n \to \infty} \inf_{f_\varepsilon \in \mathscr{F}_\varepsilon} \left(\frac{1}{n} \sum_{i=1}^n f_\varepsilon(\zeta_i) - \mathbf{E} f_\varepsilon(\zeta_1) \right) + \inf_{f \in \mathscr{F}} (\mathbf{E} f_\varepsilon(\zeta_1) - \mathbf{E} f(\zeta_1)) \\
&\quad \ge \liminf_{n \to \infty} \min_{f_\varepsilon \in \mathscr{F}_\varepsilon} \left(\frac{1}{n} \sum_{i=1}^n f_\varepsilon(\zeta_i) - \mathbf{E} f_\varepsilon(\zeta_1) \right) + \phi(\varepsilon) \\
&\quad \ge 0,
\end{aligned}$$

as $\mathscr{F}_\varepsilon$ is a finite set, and

$$\liminf_{n \to \infty} \left(\frac{1}{n} \sum_{i=1}^n f_\varepsilon(\xi_i) - \mathbf{E} f_\varepsilon(\zeta_1) \right) = 0$$

by the s.l.l.n. Thus the first inequality is proven. The second inequality follows by the same argument (choose $f_\varepsilon \ge f$). □

To apply this proposition, note that

$$1_{[h^{-1}(l_i), h^{-1}(l_i) + r_i)}(t) = I\left[\zeta_i^1 \le h(s)\right] I\left[\zeta_i^1 > h\left(s - \zeta_i^2\right)\right]$$

with $\zeta_i = (l_i, r_i)$. [Thus, $(\zeta_i)_{i\in\mathbf{N}}$ are i.i.d.] Choose

$$\mathscr{F} = \{f_{h,s}: \mathbf{R}^2 \to \mathbf{R},\ f_{h,s}(x_1, x_2) = I[x_1 \le h(s)]I[x_1 > h(s - x_2)] \quad \text{for some } h \in \mathscr{H}_T,\ s \le T\}.$$

The proof of the approximation property for $\mathscr{F}$ follows by standard integral approximation for $h \in \mathscr{H}_T$ and by discrete approximation of $0 \le s \le T$. Thus,

$$\begin{aligned}\sup_{s\le T}\sup_{h\in\mathscr{H}_T} R_2(t) &= \sup_{s\le T}\sup_{h\in\mathscr{H}_T}\left|\frac{1}{bK}\sum_{i=1}^{bK} f_{h,s}(\zeta_i) - \mathbf{E}f_{h,s}(\zeta_1)\right|\\ &= \sup_{f\in\mathscr{F}}\left|\frac{1}{bK}\sum_{i=1}^{bK} f(\zeta_i) - \mathbf{E}f(\zeta_1)\right|\\ &\to 0 \quad \text{a.s. } (K \to \infty).\end{aligned}$$

Together with the result for $R_1(t)$, this proves (17) and thus completes the proof of Theorem 1.3. □

PROOF OF THEOREM 1.4. In view of Proposition 1.1, it suffices to show that for all $T \in \mathbf{R}_+$ and for all $m \in \mathbf{N}$, $f \in C_b^\infty(\mathbf{R}^m)$, $\phi_1, \dots, \phi_m, \psi \in C_b^\infty([0, T]^2)$,

$$\mathbf{E}\big[f(\langle \xi_K^T, \phi_1\rangle, \dots, \langle \xi_K^T, \phi_m\rangle)\langle \mu^T - \xi_K^T, \psi\rangle\big] \to 0 \qquad (K \to \infty).$$

For this, we proceed as sketched in the heuristics. Let $f, \phi_1, \dots, \phi_m, \psi$ be as above and μ, $\hat{\mu}$, $\tilde{\mu}$ as in Theorem 1.3. Then

$$\begin{aligned}&\mathbf{E}\big[f(\langle \xi_K^T, \phi_1\rangle, \dots, \langle \xi_K^T, \phi_m\rangle)\langle \mu^T - \xi_K^T, \psi\rangle\big]\\ &\quad = \mathbf{E}\left[f\left(\left\langle a(\delta_0 \times \hat{\mu})^T + \frac{1}{K}\sum_{i=1}^{bK} \delta^T_{(A_i^K, A_i^K + r_i)}, \phi_l\right\rangle, l = 1, \dots, m\right)\right.\\ &\qquad\qquad \left.\times\langle \mu^T - \xi_K^T, \psi\rangle\right] + R_1\end{aligned}$$

with

$$\begin{aligned}R_1 = \mathbf{E}\Bigg[\Bigg(&f(\langle \xi_K^T, \phi_l\rangle, l = 1, \dots, m)\\ &-f\left(\left\langle a(\delta_0 \times \hat{\mu})^T + \frac{1}{K}\sum_{i=1}^{bK} \delta^T_{(A_i^K, A_i^K + r_i)}, \phi_l\right\rangle, l = 1, \dots, m\right)\Bigg)\\ &\times\langle \mu^T - \xi_K^T, \psi\rangle\Bigg].\end{aligned}$$

Now let $\delta > 0$ be arbitrary and fixed; put, as proposed in the heuristics,

$$H_T(s,t) = \tilde{\mu}^T([0,s] \times [0,t])$$

and

$$\begin{aligned} \Delta_{k,m} H_T &= H_T((k+1)\delta, (k+m+1)\delta) - H_T((k+1)\delta, (k+m)\delta) \\ &\quad - H_T(k\delta, (k+m+1)\delta) + H_T(k\delta, (k+m)\delta) \qquad (19) \\ &= \tilde{\mu}^T([k\delta, (k+1)\delta) \times [(k+m)\delta, (k+m+1)\delta)). \end{aligned}$$

Then

$$\begin{aligned} &\mathbf{E}[f(\langle \xi_K^T, \phi_l \rangle, l = 1, \dots, m) \langle \mu^T - \xi_K^T, \psi \rangle] \\ &\quad = \mathbf{E}\Bigg[f\Bigg(\Bigg\langle a(\delta_0 \times \tilde{\mu})^T + b \sum_{k,m=0}^{\infty} \Delta_{k,m} H_T \delta^T_{(k\delta,(k+m)\delta)}, \phi_l \Bigg\rangle, l = 1, \dots, m \Bigg) \\ &\qquad \times \langle \mu^T - \xi_K^T, \psi \rangle \Bigg] + R_1 + R_2, \end{aligned}$$

where

$$\begin{aligned} R_2 = \mathbf{E}\Bigg[\Bigg\{ &f\Bigg(\Bigg\langle a(\delta_0 \times \hat{\mu})^T + \frac{1}{K} \sum_{i=1}^{bK} \delta^T_{(A_i^K, A_i^K + r_i)}, \phi_l \Bigg\rangle, l = 1, \dots, m \Bigg) \\ &- f\Bigg(\Bigg\langle a(\delta_0 \times \hat{\mu})^T + b \sum_{k,m=0}^{\infty} \Delta_{k,m} H_T \delta^T_{(k\delta,(k+m)\delta)}, \phi_l \Bigg\rangle, l = 1, \dots, m \Bigg) \Bigg\} \\ &\times \langle \mu^T - \xi_K^T, \psi \rangle \Bigg]. \end{aligned}$$

Using the same approximations for ξ_K^T in $\langle \mu^T - \xi_K^T, \psi \rangle$, we obtain

$$\begin{aligned} &\mathbf{E}[f(\langle \xi_K^T, \phi_l \rangle, l = 1, \dots, m) \langle \mu^T - \xi_K^T, \psi \rangle] \\ &\quad = \mathbf{E}\Bigg[f\Bigg(\Bigg\langle a(\delta_0 \times \hat{\mu})^T + b \sum_{k,m=0}^{\infty} \Delta_{k,m} H_T \delta^T_{(k\delta,(k+m)\delta)}, \phi_l \Bigg\rangle, l = 1, \dots, m \Bigg) \\ &\qquad \times \Bigg\langle b\tilde{\mu}^T - b \sum_{k,m=0}^{\infty} \Delta_{k,m} H_T \delta^T_{(k\delta,(k+m)\delta)}, \psi \Bigg\rangle \Bigg] + R_1 + R_2 + R_3 + R_4, \end{aligned}$$

where

$$\begin{aligned} R_3 = \mathbf{E}\Bigg[&f\Bigg(\Bigg\langle a(\delta_0 \times \tilde{\mu})^T + b \sum_{k,m=0}^{\infty} \Delta_{k,m} H_T \delta^T_{(k\delta,(k+m)\delta)}, \phi_l \Bigg\rangle, l = 1, \dots, m \Bigg) \\ &\times \Bigg\langle a(\delta_0 \times \hat{\mu})^T - \frac{1}{K} \sum_{i=1}^{aK} \delta^T_{(0,\hat{r}_i)}, \psi \Bigg\rangle \Bigg], \end{aligned}$$

$$R_4 = \mathbf{E}\Bigg[f\Bigg(\Big\langle a(\delta_0 \times \hat{\mu})^T + b \sum_{k,m=0}^{\infty} \Delta_{k,m} H_T \delta^T_{(k\delta,(k+m)\delta)}, \phi_l \Big\rangle, l = 1, \dots, m \Bigg)$$
$$\times \Big\langle b\tilde{\mu}^T - \frac{1}{K} \sum_{i=1}^{bK} \Big(\sum_{k,m=0}^{\infty} \Delta_{k,m} H_T \delta^T_{(k\delta,(k+m)\delta)} - \delta^T_{(A_i^K, A_i^K + r_i)} \Big), \psi \Big\rangle \Bigg].$$

The last term is deterministic: Put

$$R_5 = f\Bigg(\Big\langle a(\delta_0 \times \hat{\mu})^T + b \sum_{k,m=0}^{\infty} \Delta_{k,m} H_T \delta^T_{(k\delta,(k+m)\delta)}, \phi_l \Big\rangle, l = 1, \dots, m \Bigg)$$
$$\times \Big\langle b\tilde{\mu}^T - b \sum_{k,m=0}^{\infty} \Delta_{k,m} H_T \delta^T_{(k\delta,(k+m)\delta)}, \psi \Big\rangle.$$

It suffices to show that the remainders tend to zero, as $\delta \to 0$ and $K \to \infty$.

Estimation of the remainders. *Estimation of R_1 and R_3.* In R_1 and R_3, we approximated $(1/K)\sum_{i=1}^{aK} \delta^T_{(0,\hat{r}_i)}$ by $a(\delta_0 \times \hat{\mu})^T$. Hence, R_1 and R_3 can be dominated in the same way. With Taylor's expansion we have

$$|R_1| \le 2\|\psi\| \, \|Df\| \sum_{l=1}^{m} \mathbf{E} \left| \Big\langle \frac{1}{K} \sum_{i=1}^{aK} \delta^T_{(0,\hat{r}_i)} - a(\delta_0 \times \hat{\mu})^T, \phi_l \Big\rangle \right|,$$
$$|R_3| \le \|f\| \mathbf{E} \left| \Big\langle \frac{1}{K} \sum_{i=1}^{aK} \delta^T_{(0,\hat{r}_i)} - a(\delta_0 \times \hat{\mu})^T, \psi \Big\rangle \right|,$$

where we suppressed the subscript ∞ in the norms. For all $\psi \in C_b^\infty([0, T]^2)$, we have by the Cauchy–Schwarz inequality and the independence of the $\hat{r}_i$'s,

$$\begin{aligned}
&\mathbf{E} \left| \Big\langle \frac{1}{aK} \sum_{i=1}^{aK} \delta^T_{(0,\hat{r}_i)} - (\delta_0 \times \hat{\mu})^T, \psi \Big\rangle \right| \\
&\le \left\{ \operatorname{Var}^T \Big(\frac{1}{aK} \sum_{i=1}^{aK} \psi(0, \hat{r}_i) \Big) \right\}^{1/2} + \left| \frac{1}{aK} \sum_{i=1}^{aK} \mathbf{E} \psi(0, \hat{r}_i) - \big\langle (\delta_0 \times \hat{\mu})^T, \psi \big\rangle \right| \\
&\le \Big(\frac{1}{aK} \Big)^{1/2} \|\psi\| + \left| \Big\langle \frac{1}{aK} \sum_{i=1}^{aK} \mathscr{L}((0, \hat{r}_i)) - \lim_{K\to\infty} \frac{1}{aK} \sum_{i=1}^{aK} \mathscr{L}((0, \hat{r}_i)), \psi \Big\rangle \right| \\
&\to 0 \qquad (K \to \infty)
\end{aligned}$$

due to the assumptions. Hence, $R_1 \to 0$ and $R_3 \to 0$ $(K \to \infty)$.

Estimation of R_2 and R_4. These remainders deal with the approximation of $(1/K)\sum_{i=1}^{bK}\delta^T_{(A_i^K, A_i^K+r_i)}$ by $b\sum_{k,m=0}^{\infty}\Delta_{k,m}H_T\delta^T_{(k\delta,(k+m)\delta)}$. We have by Taylor's expansion

$$|R_2| \le 2\|\psi\|\,\|Df\| \times \sum_{i=1}^{m}\mathbf{E}\left|\left\langle\frac{1}{K}\sum_{i=1}^{bK}\left(\delta^T_{(A_i^K, A_i^K+r_i)} - \sum_{k,m=0}^{\infty}\Delta_{k,m}H_T\delta^T_{(k\delta,(k+m)\delta)}\right), \phi_l\right\rangle\right|,$$

$$|R_4| \le \|f\|\mathbf{E}\left|\left\langle\frac{1}{K}\sum_{i=1}^{bK}\left(\delta^T_{(A_i^K, A_i^K+r_i)} - \sum_{k,m=0}^{\infty}\Delta_{k,m}H_T\delta^T_{(k\delta,(k+m)\delta)}\right), \psi\right\rangle\right|.$$

Abbreviate the indicator function

$$1_{(k,m,\delta)} = 1_{[k\delta,(k+1)\delta)} \times [(k+m)\delta, (k+m+1)\delta).$$

With the same notation as in the proof of Theorem 1.3, observe that for every $\psi \in C_b^\infty([0,T]^2)$, by Taylor's expansion and the Cauchy–Schwarz inequality,

$$\begin{aligned}
&\mathbf{E}\left|\left\langle\frac{1}{K}\sum_{i=1}^{bK}\left(\delta^T_{(A_i^K, A_i^K+r_i)} - \sum_{k,m=0}^{\infty}\Delta_{k,m}H_T\delta^T_{(k\delta,(k+m)\delta)}\right), \psi\right\rangle\right| \\
&\quad \le \mathbf{E}^T\left|\frac{1}{K}\sum_{i=1}^{bK}\sum_{k,m=0}^{\infty}\big(\psi(A_i^K, A_i^K+r_i) - \psi(k\delta,(k+m)\delta)\big) \times 1_{(k,m,\delta)}(A_i^K, A_i^K+r_i)\right| \\
&\qquad + \mathbf{E}^T\left|\frac{1}{K}\sum_{i=1}^{bK}\sum_{k,m=0}^{\infty}\Big\{\psi(k\delta,(k+m)\delta) \times 1_{(k,m,\delta)}\big(F_K^{-1}(l_i), F_K^{-1}(l_i)+r_i\big) - \mathbf{E}\big[1_{(k,m,\delta)}\big(G_T^{-1}(l_i), G_T^{-1}(l_i)+r_i\big)\big]\Big\}\right| \\
&\quad \le 2b\|D\psi\|\delta + \mathbf{E}^T\left|\frac{1}{K}\sum_{i=1}^{bK}\sum_{k,m=0}^{\infty}\Big\{\psi(k\delta,(k+m)\delta) \times \Big(1_{(k,m,\delta)}\big(F_K^{-1}(l_i), F_K^{-1}(l_i)+r_i\big) - 1_{(k,m,\delta)}\big(G_T^{-1}(l_i), G_T^{-1}(l_i)+r_i\big)\Big)\Big\}\right| \\
&\qquad + \mathbf{E}^T\left|\frac{1}{K}\sum_{i=1}^{bK}\sum_{k,m=0}^{\infty}\psi(k\delta,(k+m)\delta) \times \Big\{1_{(k,m,\delta)}\big(G_T^{-1}(l_i), G_T^{-1}(l_i)+r_i\big)\right.
\end{aligned}$$

$$-\mathbf{E}\left[1_{(k,m,\delta)}\left(G_T^{-1}(l_i), G_T^{-1}(l_i) + r_i\right)\right]\Big\}\Bigg|$$

$$\leq 2b\|D\psi\|\delta + 2b\|\psi\|\frac{T}{\delta}\{\beta\varepsilon + \mathbf{P}[\|F_K - G_T\|_T \geq \varepsilon]\}$$

$$+\mathbf{E}^T\left[\left\{\operatorname{Var}^T\left(\frac{1}{K}\sum_{i=1}^{bK}\sum_{k,m=0}^{\infty}\theta_{k,i,m,\delta}\,\middle|\, r_j, j \in \mathbf{N}\right)\right\}^{1/2}\right]$$

$$\leq 2\left\{b\|D\psi\|\delta + b\|\psi\|\frac{T}{\delta}\{\beta\varepsilon + \mathbf{P}[\|F_K - G_T\|_T \geq \varepsilon]\}\right.$$

$$\left. + \frac{\sqrt{b}}{c\sqrt{K}}\|\psi\|\right\},$$

for all $\varepsilon > 0$, writing $\theta_{k,i,m,\delta} = \psi(k\delta, (k+m)\delta)1_{(k,m,\delta)}(G_T^{-1}(l_i), G_T^{-1}(l_i) + r_i)$ and using (15) and (16). Hence $\lim_{\delta\to 0}\lim_{K\to\infty}|R_2| = 0$ and $\lim_{\delta\to 0}\lim_{K\to\infty}|R_4| = 0$.

Estimation of R_5. Finally, R_5 reflects the approximation of $\tilde{\mu}^T$ by the sum $\sum_{k,m=0}^{\infty}\Delta_{k,m}H_T\,\delta^T_{(k\delta,(k+m)\delta)}$. We have

$$|R_5| \leq b\|f\|\left|\langle\tilde{\mu}^T, \psi\rangle - \sum_{k,m=0}^{\infty}\psi(k\delta, (k+m)\delta)\Delta_{k,m}H_T\right|$$

$$\leq 2b\|f\|\left(\|D\psi\|\delta + \left|\sum_{k,m=0}^{T/\delta}\psi(k\delta, (k+m)\delta)\right.\right.$$

$$\times\{\tilde{\mu}^T([k\delta, (k+1)\delta)$$

$$\times[(k+m)\delta, (k+m+1)\delta))$$

$$\left.\left. - \Delta_{k,m}H_T\}\right|\right),$$

where we employed Taylor's expansion. By (19), the last summand is 0. Hence, $\lim_{\delta\to 0} R_5 = 0$. This completes the proof. □

PROOF OF LEMMA 1.5. Let $C_c = \{f: [0,T]^2 \to \mathbf{R}_+ \text{ continuous}\}$ and let ξ be a random element on $\mathbf{R}^2$ with $\mathscr{L}(\xi) = \delta_{\mu^T}$. As we have from $\mathscr{L}(\xi_K^T) \Rightarrow_w \delta_{\mu^T}$ $(K \to \infty)$ that

$$\langle\mathscr{L}(\xi_K^T), f\rangle \to_d \langle\delta_{\mu^T}, f\rangle,$$

for all $f \in C_c$, it follows by Theorem 4.2 and Lemma 4.3 of Kallenberg ([7], page 32) that

$$\xi_K^T(B) \to_d \xi(B),$$

for all $B \in \mathscr{B}_\xi = \{B \in \mathscr{B}_T,\ \xi(\partial B) = 0 \text{ a.s.}\}$. Now, for all $A \in \mathscr{B}_T$, $B \in \mathscr{B}(\mathbf{R}^2)$,

$$\mathbf{P}[\,\xi_K(B) \in A] = \int_{M_1(\mathbb{R}_+^2)} I[\eta(B) \in A]\mathbf{P}[\,\xi_K(B) \in d\eta]$$
$$= I[\,\mu^T(B) \in A],$$

that is, $\xi_K(B) = \delta_{\mu^T(B)}$ a.s.; hence, $\mathscr{L}(\xi_K^T(B)) \Rightarrow_w \delta_{\mu^T(B)}$. The assertion now follows from Proposition 11.1.3 of Dudley ([6], page 305). □

Acknowledgments. I would like to thank my thesis advisor A. D. Barbour for many fruitful discussions and for his continuing encouragement and support. Furthermore, I would like to thank the referee for many helpful remarks.

REFERENCES

[1] AMANN, H. (1990). *Ordinary Differential Equations: An Introduction to Nonlinear Analysis*. de Gruyter, Berlin.

[2] BAILEY, N. T. J. (1975). *The Mathematical Theory of Infectious Diseases and Its Applications*, 2nd ed. Griffin, London.

[3] BARBOUR, A. D. (1990). Stein's method for diffusion approximations. *Probab. Theory Related Fields* **84** 297–322.

[4] BERARD, D. R., BURGESS, J. W. and COSS, R. G. (1981). Plasticity of dendritic spine formation: a state-dependent stochastic process. *Int. J. Neurosci.* **13** 93–98.

[5] CHOW, Y. S. and TEICHER, H. (1988). *Probability Theory*, 2nd ed. Springer, New York.

[6] DUDLEY, R. M. (1989). *Real Analysis and Probability*. Wadsworth & Brooks/Cole, Belmont, CA.

[7] KALLENBERG, O. (1986). *Random Measures*, 4th ed. Academic Press, London.

[8] POLLARD, D. (1984). *Convergence of Stochastic Processes*. Springer, New York.

[9] REINERT, G. (1994). A weak law of large numbers for empirical measures via Stein's method, and applications. Ph.D. dissertation, Univ. Zürich.

[10] REINERT, G. (1995). A weak law of large numbers for empirical measures via Stein's method. *Ann. Probab.* **23** 334–354.

[11] SELLKE, T. (1983). On the asymptotic distribution of the size of a stochastic epidemic. *J. Appl. Probab.* **20** 390–394.

[12] SHORACK, G. R. (1979). The weighted empirical process of row independent random variables with arbitrary distribution functions. *Statist. Neerlandica* **33** 169–189.

[13] SOLOMON, W. (1987). Representation and approximation of large population age distributions using Poisson random measures. *Stochastic Process. Appl.* **26** 237–255.

[14] WANG, F. S. J. (1975). Limit theorems for age and density dependent stochastic population models. *J. Math. Biol.* **2** 373–400.

[15] WANG, F. S. J. (1977). A central limit theorem for age- and-density dependent population processes. *Stochastic Process. Appl.* **5** 173–193.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SOUTHERN CALIFORNIA
DRB 155
LOS ANGELES, CALIFORNIA 90089-1113