

1 **Identifiability of Age-Dependent Branching**
2 **Processes from Extinction Probabilities and**
3 **Number Distributions**

4 **Pak-Wing Fok¹ and Tom Chou²**

5 ¹Department of Mathematics, University of Delaware, Newark, DE 19716

6 ²Departments of Biomathematics and Mathematics, UCLA, Los Angeles, CA
7 90095-1766

8 E-mail: pakwing@udel.edu, tomchou@ucla.edu

9 **Abstract.** Consider the single-species, independent-particle, Bellman-Harris
10 branching process, defined by a progeny number distribution, and a particle
11 lifetime distribution. In this paper, we explore the existence and uniqueness of
12 the inverse problem, where one wishes to solve for the progeny number or lifetime
13 distribution given information about the total number distribution. Results
14 are derived for the uniqueness and reconstructibility of these two distributions
15 from two types of information: the extinction time probability of the entire
16 process (extinction time distribution), and the distribution of the total number
17 of particles at a fixed time. Assuming perfect information of either type, we
18 seek to determine if the progeny number distribution or the lifetime probability
19 distribution functions can be uniquely determined. We demonstrate that the
20 distribution of extinction times allows us to formally determine either the progeny
21 number distribution *or* the lifetime distribution. Furthermore, we show that these
22 constructions are unique. For a known total number distribution, given at a
23 specific time, we show that the lifetime distribution is locally unique and that the
24 progeny distribution is globally unique. Our results are presented through four
25 theorems, each describing the constructions in the four distinct cases. on times
26 allows us to formally determine either the progeny number distribution *or*
27 the lifetime distribution. Furthermore, we show that these constructions are unique.
28 For a known total number distribution, given at a specific time, we show that
29 the lifetime distribution is locally unique and that the progeny distribution is
30 globally unique. Our results are presented through four theorems, each describing
31 the constructions in the four distinct cases.
32
33
34
35
36
37
38
39
40
41
42

43 PACS numbers: 02.50.Ey,05.10.Gg,02.30.Rz

44 AMS classification scheme numbers: 45Q05,92B05,60G99

45 **1. Introduction**

46 Branching processes have proved useful in many applications, including modeling cell
 47 proliferation and differentiation, propagation of surnames, and analyzing radioactive
 48 chain reactions [1, 2, 3, 4]. A mathematical model for the branching of a single
 49 species of identical particles is the Bellman-Harris process, defined by a particle lifetime
 50 distribution and a progeny number distribution. The lifetime distribution determines
 51 the statistics of the time a particle waits before branching, while the progeny number
 52 distribution determines the statistics of how many offspring particles each parent
 53 particles generates at each branching event. Schematics of the single-species branching
 54 process are shown in Fig. 1. Here, the process is initiated with a single parent particle
 55 that branches at a time between τ and $\tau + d\tau$ with probability $G(\tau)d\tau$. Each branching
 56 event results in the parent particle giving birth to $k \geq 2$ new particles with probability
 57 b_k , or dying with probability b_0 . All particle lifetimes are independently and identically
 58 drawn from the probability distribution function $G(\tau)$.

59 Mathematically, the branching process is conveniently described via the
 60 probability generating function (pgf) $F(z, t)$ for the probability $f_n(t)$ of observing
 61 n total particles at time t :

62
$$F(z, t) = \sum_{n=0}^{\infty} f_n(t) z^n.$$

63 Note that normalization $\sum_{n=0}^{\infty} f_n(t) = 1$ yields $F(z = 1, t) = 1$.

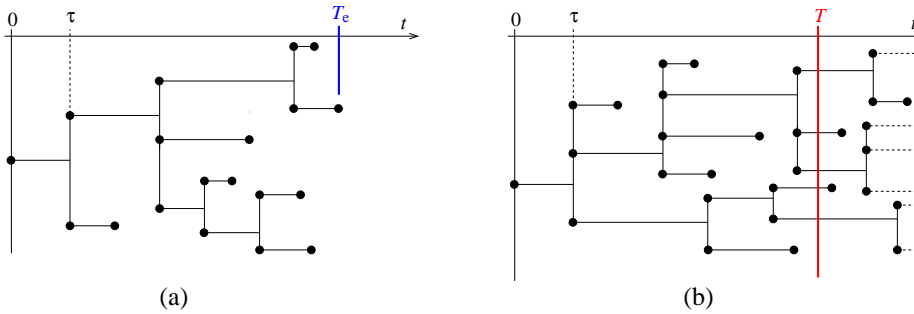


Figure 1. (a) A realization of a Bellman-Harris branching process that went extinct at time T_e . If the process dies in finite time, can the distribution of extinction times be used to determine the progeny or lifetime distributions? (b) A realization of a Bellman-Harris branching process with a known number distribution of particles at time $t = T$. In the depicted realization there are five surviving particles at time T . Can the total number distribution at a fixed time be used to determine the progeny or lifetime distributions? In principle, both types of “data” can only be obtained from an infinite number of measurements.

64 For completeness, we follow standard derivations of the age-dependent,
 65 continuous-time Bellman-Harris branching process by considering the process to be
 66 initially seeded by a single parent particle [1, 2, 3, 4]. Upon defining $F(z, t|\tau)$ as the
 67 generating function of the process conditioned on the original parent particle having
 68 first branched between time τ and $\tau + d\tau$, we find

69
$$F(z, t|\tau) = \begin{cases} z, & t < \tau \\ B[F(z, t - \tau)], & t \geq \tau, \end{cases} \quad (1)$$

70 where

$$71 \quad B[z] = \sum_{k=0}^{\infty} b_k z^k, \quad (2)$$

72 is the generating function for the progeny number distribution $b_k \geq 0$. The coefficients
 73 of z^k in $B[z]$ correspond to the probability that k particles are born at each branching
 74 event. For processes with finite maximum offspring and branching probabilities
 75 $\{b_0, b_1, \dots, b_N\}$, not all zero, our convention in eq. (2) is to take $b_k = 0$ for $k > N$.
 76 Generating $k = 1$ particle at the time of branching is simply equivalent to renewing
 77 the particle's lifespan. This generally does not occur in cell proliferation or nuclear
 78 chain reactions (the main applications we have in mind) and so throughout this paper,
 79 we always assume that the particles either die, or generate $k \neq 1$ particles at each
 80 branching event. *Therefore we always take $b_1 = 0$.* Averaging eq. (1) over the lifetime
 81 distribution $G(\tau)$, we find

$$82 \quad F(z, t) \equiv \int_0^{\infty} F(z, t|\tau)G(\tau)d\tau = z \int_t^{\infty} G(\tau)d\tau + \int_0^t B[F(z, t-\tau)]G(\tau)d\tau. \quad (3)$$

83 This branching process is defined by two parameter functions, \mathbf{b} , the vector of progeny
 84 number probabilities, and $G(\tau)$, the probability density function (pdf) of lifetimes for
 85 each particle before it branches. Given a set \mathbf{b} and a lifetime distribution $G(\tau)$, eq. (3),
 86 along with the initial condition $F(z, 0) = z$, can be solved to find a $F(z, t)$, from which
 87 $f_n(t)$ can be generated.

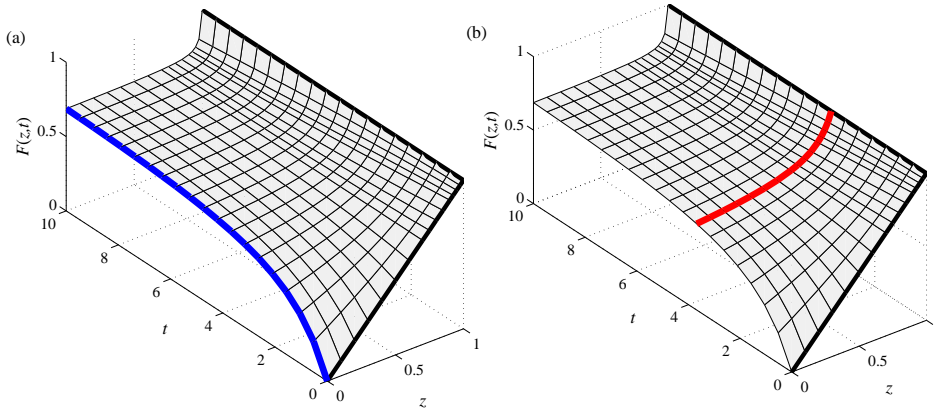


Figure 2. Reconstruction of a branching process from two types of functional information. (a) The extinction time distribution, $F(0, t) \equiv f_0(t)$ (shown in blue). (b) The number distribution at a given time, $F(z, T)$ (shown in red). Thick black lines indicate common, known, information about the surface $F(z, t)$: $F(z, 0) = z$ and $F(1, t) = 1$.

88 Equation (3), and its multispecies generalization have been well-studied by
 89 researchers across many disciplines [5, 6, 7, 8], particularly in cell biology. In the
 90 forward problem one knows the set of progeny probabilities $\{\mathbf{b}\}$ and the lifetime
 91 distribution function $G(t)$, and wishes to calculate the total particle probability

distribution $f_n(t)$ [3], the moments of the particle numbers (for example the mean $\langle n(t) \rangle = \sum_{n=0}^{\infty} n f_n(t)$), or the extinction probabilities of the process [9]. Often, exponentially-distributed lifetimes are assumed, turning eq. (3) into a nonlinear ODE. This nonlinear equation can be further reduced to a series of linear ODEs describing the evolution of the moments of the particle number.

Often however, the underlying details of the branching process are not precisely known *a priori*, but statistical properties of the branching process can be measured. Thus, inference and the inverse branching problem are also important topics. Maximum likelihood approaches have been used to estimate parameters defining simple branching processes [10, 11, 12, 13]. Estimation of parameters in a more sophisticated branching model was done in [14] where the distribution of times to death and to reproduction may differ [15]. These studies consider branching processes under specific parameters and attempt to estimate them from measurements of quantities such as the mean particle number [16]. Non-parametric approaches have also been explored. For example, in [17] the authors use a Bayesian approach to estimate offspring and lifetime distributions given the whole history of the process up to some fixed time.

In this paper, we consider the formal inverse problem of finding the set \mathbf{b} or the lifetime distribution $G(\tau)$ from *perfect* measurements of the branching process. Rather than estimating parameters, we simply explore the mathematical uniqueness of the inverse problem defined by the nonlinear integral eq. (3) and two types of perfect information. One could also consider the inverse problem of simultaneously finding $G(\tau)$ and \mathbf{b} . However, it can be shown that a branching process with $b_1 > 0$ and exponentially distributed lifetimes generates identical statistics to a branching process with $b_1 = 0$ and a rescaled exponential lifetime distribution. Thus the problem of reconstructing both $G(\tau)$ and \mathbf{b} is not unique.

The first known function to use is the cumulative distribution of extinction times $F(0, t)$. Note that $\lim_{t \rightarrow \infty} F(0, t) < 1$ in general because the process may never extinguish. The extinction probability is positive only if $b_0 > 0$. Therefore the function $F(z, t)$ is known at $t = 0$, where $F(z, 0) = z$; at $z = 1$, where $F(z = 1, t) = 1$ from normalization; and at $z = 0$ where $F(z = 0, t)$ is known. These known parts of the function $F(z, t)$ are indicated by the thick black and blue curves in Fig. 2(a). In particular, $F(0, 0) = 0$.

Another possible type of information is the total particle number distribution of the branching processes at *one* fixed time T . This may in principle be derived from an infinite number of samples of the process. The resulting distribution of particle numbers at time T , $f_n(T)$, yields exactly the pgf $F(z, T) = \sum_{n=0}^{\infty} f_n(T) z^n$. In this case, $F(z, t)$ is known at $t = 0$, where $F(z, 0) = z$; at $z = 1$, where $F(z = 1, t) = 1$ from normalization; and for $0 \leq z \leq 1$ at time $0 < T < \infty$, as indicated by the thick black and red curves in Fig. 2(b).

The remainder of this paper essentially explores the conditions under which the rest of the function $F(z, t)$ can be reconstructed from these two types of information. Specifically, we wish to determine if the probabilities \mathbf{b} and/or lifetime distribution $G(\tau)$ can be extracted from the thick curves in Figs. 2, and if so, whether or not they are unique.

Throughout this paper, we will refer to eq. (3) and assume that $G(t)$ is a continuous probability distribution function ($G(t) \geq 0$ if $t \geq 0$ and $\int_0^{\infty} G(t) dt = 1$) and $B[z]$ is a bounded, infinitely differentiable function on $z \in [0, 1]$ through eq. (2). Furthermore, we will freely use the fact that a probability distribution is uniquely

141 defined by its Laplace Transform [21] (i.e. $\mathcal{L}\{F_1(t)\} = \mathcal{L}\{F_2(t)\}$ if and only if
 142 $F_1(t) = F_2(t)$ for distributions $F_1(t)$ and $F_2(t)$ and \mathcal{L} denotes Laplace Transform).

143 1.1. Technical Lemmas

144 We first prove some technical lemmas that are used often in our subsequent proofs.

145 **Lemma 1.** *Let $F(z, t)$ be a solution to eq. (3) and let $G(t)$ be a continuous pdf. Let*
 146 *$P(t)$ be a function that is not identically zero and is continuous for all $t \geq 0$. Then*

147 (i) *The pgf $F(z, t)$ satisfies $0 \leq F(z, t) \leq 1$ and is continuous on $(z, t) \in (0, 1) \times$
 148 $(0, \infty)$. Furthermore, it has infinitely many z -derivatives on $(0, 1)$.*

149 (ii) *For any fixed $0 \leq t < \infty$, the functions in the set $\{F^j(z, t)\}$, $j = 0, 1, \dots, \infty$ are*
 150 *linearly independent in $z \in (0, 1)$.*

151 (iii) *For fixed $t > 0$, the functions $1, z$ and $F^j(z, t)$, $j = 2, 3, \dots, \infty$ which we compactly*
 152 *write as $\{1, z, F^j(z, t)\}$, form an infinite, linearly independent set in $z \in (0, 1)$.*

153 (iv) *The infinite set of functions*

$$154 \left\{ \int_0^t P(\tau) d\tau, \int_0^t z P(\tau) d\tau, \int_0^t F^2(z, t - \tau) P(\tau) d\tau, \dots, \int_0^t F^j(z, t - \tau) P(\tau) d\tau, \dots \right\} \quad (4)$$

155 *form a linearly independent set in $z \in (0, 1)$ for all $t > 0$.*

156 *Proof.* (i) A well-known property of pgfs like $F(z, t)$ is that they are continuous,
 157 infinitely differentiable and strictly monotonically increasing in z (e.g. see
 158 Theorem 1 in [1]). Also, the number of particles at finite time t in a Bellman-
 159 Harris process is always finite and this is equivalent to the statement that
 160 $F(1, t) = 1$ [1]. Since $F(0, t) \geq 0$, we have $0 \leq F(z, t) \leq 1$ and F is a bounded
 161 function. Furthermore, $F(z, t) \equiv \sum_{n=0}^{\infty} z^n f_n(t)$ must also be continuous for $t > 0$.
 162 If F is not continuous in t , then $\int_0^t B[F(t - \tau)]G(\tau) d\tau$ in (3) must be discontinuous
 163 as $\int_t^{\infty} G(\tau) d\tau$ is clearly continuous. Therefore $B[F(t - \tau)]G(\tau)$ cannot be a
 164 bounded function since all bounded functions have continuous indefinite integrals.
 165 Therefore F cannot be bounded since B is a bounded function and we have a
 166 contradiction. Hence $F(z, t)$ must be continuous for $t > 0$ and $0 < z < 1$.

167 (ii) Consider the equation

$$168 \sum_{j=0}^{\infty} c_j F^j(z, t) = 0. \quad (5)$$

169 For a fixed t , there exist constants $F_{min} < F_{max}$ such that $F(z, t) : [0, 1] \rightarrow$
 170 $[F_{min}, F_{max}]$ bijectively. The monomials $\{1, F, F^2, \dots\}$ form a basis for infinitely
 171 differentiable functions on $[F_{min}, F_{max}]$. Therefore, regarding (5) as a function
 172 in F , we have $c_j = 0$ for $j \geq 0$ and so the $F^j(z, t)$ are linearly independent on
 173 $z \in [0, 1]$.

174 (iii) From (ii), $\{1, F^2(z, t), F^3(z, t), \dots\}$ are linearly independent in z since any subset
 175 of a linearly independent set is itself linearly independent. The functions 1
 176 and z are clearly linearly independent. We claim that the Wronskian of z
 177 and $F^j(z, t)$, $j \geq 2$ is not identically zero in $(0, 1)$, so that supplementing
 178 $\{1, F^2(z, t), F^3(z, t), \dots\}$ with $\{z\}$ keeps the set linearly independent.

179 We assume that the Wronskian of z and $F^j(z, t)$ ($j \geq 2$) is identically zero and
 180 aim for a contradiction. Then

$$181 \quad z \frac{\partial}{\partial z} F^j(z, t) - F^j(z, t) = 0,$$

182 for all $z \in (0, 1)$ which implies $F^j(z, t) = C(t)z$ for some function $C(t)$, or
 183 $F(z, t) = C^{1/j}(t)z^{1/j}$ with $j \geq 2$. However, expanding about $z = 0$, $F(z, t) =$
 184 $\sum_{n=0}^{\infty} z^n f_n(t)$ has no fractional powers of z . This gives us a contradiction.

185 (iv) Let

$$186 \quad c_1 z \int_0^t P(\tau) d\tau + \sum_{m \neq 1}^{\infty} c_m \int_0^t F^m(z, t - \tau) P(\tau) d\tau = 0.$$

187 To show linear independence, we need to show that $c_m = 0$ for all m . Upon
 188 taking Laplace Transforms, we find

$$189 \quad c_1 \frac{z \tilde{P}(s)}{s} + \sum_{m \neq 1}^{\infty} c_m \widetilde{F^m}(z, s) \tilde{P}(s) = 0, \quad \text{Re}(s) > 0.$$

190 Since $P(t)$ is not identically zero, $\tilde{P}(s)$ is not identically zero. After dividing by
 191 $\tilde{P}(s)$ and taking the inverse Laplace Transform of both sides, we find

$$192 \quad c_1 z + \sum_{m \neq 1}^{\infty} c_m F^m(z, t) = 0.$$

193 From (iii), the functions $\{1, z, F^2(z, t), \dots, F^m(z, t), \dots\}$ are linearly independent
 194 in z . Therefore $c_m = 0$ for all m , and the set of functions in (4) are also linearly
 195 independent.
 196 □

197 2. Reconstruction from extinction time distributions

198 In this section, we first consider eq. (3) when the cumulative extinction probability
 199 $F(z = 0, t)$ is perfectly measured (see Fig. 2(a)). We will use the following lemma in
 200 the two theorems of this section:

201 **Lemma 2.** Assume that $h(x)$ and all its derivatives exist on $(0, \infty]$ and vanish at
 202 $x = \infty$. Also, let $h(x) \sim x^k$ as $x \rightarrow 0^+$ where $k > -1$. Then as $\lambda \rightarrow \infty$,

$$203 \quad \int_0^{\infty} h(x) e^{\pm i \lambda x} dx = O(\lambda^{-(k+1)}),$$

204 *Proof.* A proof is provided in [18]. □

205 2.1. Reconstruction of lifetime distribution $G(t)$ given progeny number distribution \mathbf{b}

206 If the progeny number distribution \mathbf{b} , and the corresponding generating function $B[z]$
 207 is known, is there more than one lifetime distribution function $G(t)$ that yields the
 208 same extinction time statistics? To answer this question, we prove the following:

209 **Theorem 1** (Reconstruction and uniqueness of lifetime probabilities from extinction
 210 probabilities). For eq. (3), assume that the coefficients of $B[z]$ (see eq. (2)) b_k ,
 211 $k = 0, 1, \dots, \infty$ are known, $b_0 > 0$, and there is an underlying probability distribution
 212 of lifetimes $G(t)$ that gives rise to $F(0, t)$, the cumulative extinction time distribution.
 213 Furthermore assume $F(0, t) = O(t)$ as $t \rightarrow 0^+$. Then $G(t)$ is uniquely given in terms
 214 of $F(0, t)$ by

$$215 \quad G(t) = \int_{\Gamma} \frac{\tilde{F}(0, s)e^{st}}{\tilde{K}(s)} \frac{ds}{2\pi i}, \quad \tilde{K}(s) = \mathcal{L}\{B[F(0, t)]\}, \quad (6)$$

216 where we use both tilde and \mathcal{L} to denote the Laplace Transform and the contour Γ lies
 217 to the right of all the poles of the integrand. If $F_1(0, t)$ and $F_2(0, t)$ are extinction
 218 time distributions and $G_1(t)$ and $G_2(t)$ are the associated lifetime distributions, then
 219 $F_1(0, t) = F_2(0, t) \Rightarrow G_1(t) = G_2(t)$.

220 *Proof.* We first show that eq. (6) is true for any cumulative extinction time distribution
 221 $F(0, t)$. Upon taking the Laplace transform of the Bellman-Harris equation (3) at
 222 $z = 0$, we find

$$223 \quad \tilde{F}(0, s) = \tilde{K}(s)\tilde{G}(s), \quad (7)$$

224 and $\tilde{K}(s)$ is the Laplace transform of the function $K(t) \equiv B[F(0, t)]$. All three
 225 Laplace-transformed functions in (7) exist providing $\text{Re}(s) > 0$. Upon inverse Laplace
 226 transforming $\tilde{G}(s)$, we find

$$227 \quad G(t) = \int_{\gamma-i\infty}^{\gamma+i\infty} \frac{\tilde{F}(0, s)e^{st}}{\tilde{K}(s)} \frac{ds}{2\pi i}, \quad (8)$$

228 where $\gamma > 0$. To show that the integral (8) converges for all $t \geq 0$, let $s = \gamma + i\mu$,
 229 $\mu \in \mathbb{R}$. Lemma 2 implies

$$\tilde{F}(0, \gamma + i\mu) = \int_0^{\infty} e^{-i\mu t} g_0(t) dt = O(\mu^{-2}), \quad (9)$$

$$\tilde{K}(\gamma + i\mu) = \int_0^{\infty} e^{-i\mu t} g_1(t) dt = O(\mu^{-1}), \quad (10)$$

230 as $\mu \rightarrow \pm\infty$, where $g_0(t) = e^{-\gamma t} F(0, t) = O(t)$ and $g_1(t) = e^{-\gamma t} \sum_{m \neq 1} b_m F^m(0, t) =$
 231 $O(1)$ as $t \rightarrow 0^+$. Since

$$232 \quad G(t) = \frac{e^{\gamma t}}{2\pi} \int_{-\infty}^{\infty} e^{i\mu t} \frac{\tilde{F}(0, \gamma + i\mu)}{\tilde{K}(\gamma + i\mu)} d\mu,$$

233 we see that $G(t)$ is proportional to the inverse Fourier Transform of $\tilde{F}(0, \gamma + i\mu)/\tilde{K}(\gamma +$
 234 $i\mu)$. This ratio is square integrable at $\mu = \pm\infty$ by (9) and (10). Furthermore
 235 $\tilde{F}(0, \gamma + i\mu)/\tilde{K}(\gamma + i\mu)$ is regular for finite μ since the contour in (8) lies to the
 236 right of all singularities in the integrand. Therefore the inverse Fourier transform
 237 always exists and $G(t)$ converges for all t – in particular $G(t) \equiv 0$ when $t < 0$.

238 What remains is for us to show that this $G(t)$ is unique. Assume there are two
 239 lifetime distributions $G_1(t)$ and $G_2(t)$ that give rise to extinction probabilities $F_1(0, t)$

240 and $F_2(0, t)$. If $F_1 = F_2$, we can subtract the corresponding integral equations (eq. 3)
241 evaluated at $z = 0$ to find

$$242 \quad 0 = \int_0^t B[F(0, t - \tau)](G_1(\tau) - G_2(\tau))d\tau.$$

243 This equation can be written in terms of the Laplace transforms of $K(t) \equiv B[F(0, t)]$
244 and $G_1(t) - G_2(t)$:

$$245 \quad \tilde{K}(s)(\tilde{G}_1(s) - \tilde{G}_2(s)) = 0.$$

246 If $b_0 > 0$, then $F(0, t) > 0$ and $B[F(0, t)] > 0$. Therefore, $\tilde{K}(s) > 0$, implying
247 $\tilde{G}_1(s) = \tilde{G}_2(s)$ and $G_1(t) = G_2(t)$ (two distributions are identical if and only if
248 their Laplace Transforms are identical [21]). As long as an extinction occurs with
249 nonzero probability, and one lifetime distribution $G(t)$ leads to a specific extinction
250 time distribution $F(0, t)$, then no other different lifetime distribution can lead to the
251 same extinction time distribution. \square

252 In Theorem 1, we assumed that there exists an underlying Bellman-Harris process
253 with associated branching probabilities $\{b_j\}$ and lifetime distribution function $G(t)$
254 that generated the cumulative extinction time distribution $F(0, t)$; the main result of
255 the theorem was that $G(t)$ can be found from $F(0, t)$ and $\{b_j\}$ through (6). However,
256 if we are simply given $F(0, t)$ without any conditions on the existence of G , we also
257 need to show that (6) is indeed a pdf, i.e. $G(t) > 0$ for all $t > 0$ and $\int_0^\infty G(t)dt = 1$.
258 Conditions for $G(t)$ to be a pdf in terms of its Laplace transform can be found in [21],
259 namely:

- 260 • $\tilde{F}(0, s)/\tilde{K}(s)$ has infinitely many derivatives in s and $(-1)^n \frac{d^n}{ds^n} \frac{\tilde{F}(0, s)}{\tilde{K}(s)} \geq 0$ when
261 $s > 0$ for every integer n .
- 262 • $\lim_{s \rightarrow 0} \frac{\tilde{F}(0, s)}{\tilde{K}(s)} = 1$.

263 If the conditions of Theorem 1 *and* these two conditions above are satisfied, then the
264 extinction time data $F(0, t)$ corresponds to a Bellman-Harris branching process and
265 the corresponding pdf of lifetimes $G(t)$ can be found through (6).

266 2.2. Reconstruction of progeny number distribution \mathbf{b} given $G(t)$

267 Now, consider the complementary problem where the lifetime distribution $G(t)$ is
268 known. Can one find a unique progeny number distribution \mathbf{b} for each cumulative
269 extinction probability distribution? We prove

270 **Theorem 2** (Reconstruction and uniqueness of progeny number probabilities from
271 extinction probabilities). *Consider eq. (3) with an unknown underlying progeny*
272 *number distribution b_j , $j = 0, 1, \dots, \infty$. With a known continuous lifetime distribution*
273 *$G(t)$ such that $\lim_{t \rightarrow 0^+} G(t) = 0$, assume the process generates a cumulative extinction*
274 *time distribution $F(0, t)$ which is strictly increasing in t . Then the b_j are (i) uniquely*
275 *determined by $F(0, t)$ and (ii) formally given in terms of the cumulative extinction*
276 *time distribution by*

$$b_j = K(0), \quad (11)$$

$$b_j = \lim_{t \rightarrow 0} \frac{K(t) - \sum_{k=0}^{j-1} b_k F^k(0, t)}{F^j(0, t)}, \quad j > 0, \quad (12)$$

277 where

$$278 \quad K(t) = \int_{\Gamma} \frac{\tilde{F}(0, s)e^{st}}{\tilde{G}(s)} \frac{ds}{2\pi i},$$

279 and Γ lies to the right of all singularities of the integrand and tilde denotes Laplace
280 Transform.

281 *Proof.* We first prove uniqueness. Consider two branching processes with the same
282 lifetime distribution $G(t)$, but with different sets of progeny number probabilities
283 \mathbf{b} and \mathbf{b}^* corresponding to the coefficients of $B[z]$ and $B^*[z]$ respectively. If these
284 two branching processes generate the same extinction time distribution $F(0, t)$, the
285 difference between the corresponding eqs. (3) becomes

$$286 \quad \int_0^t (B[F(0, t - \tau)] - B^*[F(0, t - \tau)]) G(\tau) d\tau = 0.$$

287 Introducing $K(t) = B[F(0, t)]$, $K^*(t) = B^*[F(0, t)]$, and $\tilde{K}(s)$, $\tilde{K}^*(s)$ as the
288 corresponding Laplace transforms – valid for $\text{Re}(s) > 0$ – we find

$$289 \quad (\tilde{K}(s) - \tilde{K}^*(s))\tilde{G}(s) = 0, \quad (13)$$

290 where $\tilde{G}(s)$ is also defined for $\text{Re}(s) > 0$. Since $\tilde{G}(s) \neq 0$ when $\text{Re}(s) > 0$, condition
291 (13) is satisfied only if $\tilde{K}(s) = \tilde{K}^*(s)$, which means $B[F(0, t)] = B^*[F(0, t)]$. After
292 expanding $B[F] = b_0 + \sum_{j=2}^{\infty} b_j F^j$ and $B^*[F] = b_0^* + \sum_{j=2}^{\infty} b_j^* F^j$ we find

$$293 \quad (b_0 - b_0^*) + \sum_{j=2}^{\infty} (b_j - b_j^*) F^j(0, t) = 0,$$

294 which must hold for $0 \leq F(0, t) \leq F_{max}$ where $F_{max} = \lim_{t \rightarrow \infty} F(0, t)$. From lemma
295 1(ii), since $\{F^j\}$, $j = 0, 1, 2, \dots, \infty$ are linearly independent on $[0, F_{max}]$, $b_j = b_j^*$ for
296 $j = 0, 1, 2, \dots, \infty$, i.e. $\mathbf{b} = \mathbf{b}^*$.

297 We now show how to determine the unique progeny number distribution \mathbf{b} from
298 a known extinction time distribution $F(0, t)$. From eq. (3),

$$299 \quad F(0, t) = \int_0^t K(t - \tau) G(\tau) d\tau, \quad (14)$$

300 where $K(t) \equiv B[F(0, t)]$. By taking Laplace transforms of (14), we find $\tilde{F}(0, s) =$
301 $\tilde{K}(s)\tilde{G}(s)$ when $\text{Re}(s) > 0$, leading to

$$302 \quad K(t) = \int_{\gamma - i\infty}^{\gamma + i\infty} \frac{\tilde{F}(0, s)e^{st}}{\tilde{G}(s)} \frac{ds}{2\pi i}, \quad \gamma > 0. \quad (15)$$

303 Using Lemma 2, we find

$$\tilde{F}(0, \gamma + i\mu) = \int_0^{\infty} e^{-i\mu t} \{e^{-\gamma t} F(0, t)\} dt = O(\mu^{-2}), \quad (16)$$

$$\tilde{G}(\gamma + i\mu) = \int_0^{\infty} e^{-i\mu t} \{e^{-\gamma t} G(t)\} dt = O(\mu^{-1}), \quad (17)$$

304 as $\mu \rightarrow \pm\infty$ since $e^{-\gamma t}F(0, t) = O(t)$ and $e^{-\gamma t}G(t) = O(1)$ as $t \rightarrow 0^+$. Furthermore,
 305 since

$$306 \quad K(t) = e^{\gamma t} \int_{-\infty}^{\infty} \frac{\tilde{F}(0, \gamma + i\mu)}{\tilde{G}(\gamma + i\mu)} e^{i\mu t} \frac{d\mu}{2\pi},$$

307 $K(t)$ is proportional to an inverse Fourier Transform. At $\mu = \pm\infty$, $\tilde{F}(0, \gamma + i\mu)/\tilde{G}(\gamma +$
 308 $i\mu)$ is square integrable by (16) and (17). It is also regular for finite μ since the
 309 contour in (15) lies to the right of all singularities in the integrand. Therefore
 310 $\tilde{F}(0, \gamma + i\mu)/\tilde{G}(\gamma + i\mu)$ is square integrable on $\mu \in [-\infty, +\infty]$ and the integral (15)
 311 converges for all t (and in particular, $K(t) \equiv 0$ for $t < 0$).

312 The probabilities b_j can be extracted from $F(0, t)$ through

$$313 \quad K(t) = b_0 + b_2 F^2(0, t) + b_3 F^3(0, t) + \dots \quad (18)$$

314 Since $F(0, 0) = 0$, the single-particle decay probability $b_0 = K(0)$. Likewise, b_j for
 315 $j > 0$ are reconstructed recursively from (18) by the relations

$$316 \quad b_j = \lim_{t \rightarrow 0} \frac{K(t) - \sum_{k=0}^{j-1} b_k F^k(0, t)}{F^j(0, t)}.$$

317

□

318 We have shown in this theorem that a completely known extinction time
 319 distribution function $F(0, t)$ allows one to uniquely find constants $\{b_i\}$ so that eq.
 320 (3) is satisfied. If it is known *a priori* that there are probabilities $\{b_i\}$ that generated
 321 $F(0, t)$ in a Bellman-Harris process, they can be reconstructed through formulas (11)
 322 and (12). However, if $F(0, t)$ is an arbitrary function, application of (11) and (12)
 323 will not guarantee that the resulting $\{b_i\}$ are probabilities (e.g. some of the b_i
 324 could be negative, or they may not sum to 1). Therefore $b_i \geq 0, \sum_i b_i = 1$ with
 325 b_i defined through (11) and (12) are necessary conditions for $F(0, t)$ to correspond
 326 to the cumulative extinction probability for a Bellman-Harris branching process. In
 327 principle, these conditions define the class of cumulative distribution functions $F(0, t)$
 328 that can arise from Bellman-Harris processes.

329 3. Reconstruction from a single number distribution

330 In this section, we consider the reconstruction of either the lifetime distribution $G(t)$ or
 331 the progeny number distribution \mathbf{b} from a known total number probability distribution
 332 at a single fixed time $0 < T < \infty$. Since $f_n(T)$ are known for all n , the generating
 333 function $F(z, T)$ is also precisely known for all $0 \leq z \leq 1$, as shown in Fig. 2(b). Our
 334 results focus on uniqueness of the reconstruction.

335 3.1. Uniqueness of lifetime distribution $G(t)$ given \mathbf{b}

336 **Lemma 3.** Let $B[z] = \sum_{m=0}^{\infty} b_m z^m$ and $G(t)$ give rise to $F(z, t)$ in eq. (3). Let
 337 $T > 0$ and $P(t)$ be a continuous function. Then

$$338 \quad \int_0^T (B[F(z, T - \tau)] - z) P(\tau) d\tau \equiv 0, \quad (19)$$

339 for all $0 \leq z \leq 1$ if and only if $P(t) \equiv 0$ for $0 \leq t \leq T$.

340 *Proof.* Eq. (19) is obviously true if $P(t) \equiv 0$ on $[0, T]$. Assume that eq. (19) holds.
 341 Since $B[z]$ is convergent for all $0 \leq z \leq 1$, we can interchange integration and
 342 summation and rewrite eq. (19) as

$$343 \quad \sum_{k=0}^{\infty} c_k \phi_k(z) = 0,$$

344 where

$$\begin{aligned} \phi_0(z) &= \int_0^T P(\tau) d\tau, \\ \phi_1(z) &= z \int_0^T P(\tau) d\tau, \\ \phi_k(z) &= \int_0^T F^k(z, T - \tau) P(\tau) d\tau, \quad k > 1, \end{aligned}$$

345 and $c_0 = b_0$, $c_1 = -1$, and $c_k = b_k$ for $k > 1$. Clearly, not all the coefficients c_k are
 346 zero and therefore $\{\phi_k(z)\}$ is not a linearly independent set. This contradicts Lemma
 347 1(iv), unless $P(t) \equiv 0$ for $0 \leq t \leq T$. \square

348 **Lemma 4.** *Consider the integral equation*

$$349 \quad y(t) = \int_0^t y(t - \tau) H(\tau) d\tau + cP(t), \quad (20)$$

350 where $H(t)$ and $P(t)$ are functions that are continuous and positive for all $t > 0$. Then
 351 $y(t) \geq 0$ for $t > 0$ if $c \geq 0$.

352 *Proof.* Because of the continuity of H and P , it can be shown [23] that $y(t)$ is
 353 continuous for $t > 0$. Let $c > 0$. Then $y(0) > 0$. Suppose that $y(t)$ is not positive for
 354 all t . Since y is continuous, there is a point $t^* > 0$ such that $y(t^*) = 0$ and $y(t) > 0$
 355 for $t < t^*$. Then eq. (20) implies

$$356 \quad \int_0^{t^*} y(t^* - \tau) H(\tau) d\tau = -cP(t^*) < 0.$$

357 But since $H(t) > 0$, the integral is positive and we have a contradiction. Therefore
 358 $y(t) > 0$ for all t . The proof for $c < 0$ is similar. \square

359 The next lemma is a result describing the stability of particular “trajectories”.
 360 A trajectory refers to a solution of (3) for a particular value of z . Equivalently, it is
 361 a “slice” of $F(z, t)$ at a constant value of z . Since every trajectory obeys the same
 362 equation (3), the fate of a trajectory depends only on its initial condition $F(z, 0) = z$.

363 The next lemma suggests that solutions to (3) can be understood in terms of the
 364 stability of individual trajectories, analogous to the theory of autonomous differential
 365 equations.

366 **Lemma 5.** *Consider the integral equation (3) and let $G(t)$ be continuous for $t > 0$.
 367 Then:*

368 (i) *When $B'[1] > 1$, $B[z] - z = 0$ has two distinct roots at $z = z^* < 1$ and $z = 1$
 369 and $F(z, t) = z^*, 1$ are two constant-in-time trajectories of eq. (3). Furthermore
 370 the first trajectory is (almost) globally attracting: $F(z, t) \rightarrow z^*$ monotonically if
 371 $0 \leq z < 1$. This case is commonly known as the supercritical case.*

372 (ii) When $B'[1] = 1$ or $B'[1] < 1$, $z = 1$ is the only root of $B[z] - z = 0$ and $F(z, t) \rightarrow 1$
 373 for all $0 \leq z \leq 1$. These are also known as the critical ($B'[1] = 1$) and subcritical
 374 ($B'[1] < 1$) cases.

375 *Proof.* Since $B''[z] > 0$ for all $z \in [0, 1]$, $B[z] - z$ can have either 0, 1 or 2 distinct roots.
 376 But $z = 1$ is clearly a root because $\sum_{m=0}^{\infty} b_m = 1$. Therefore $B[z] - z$ either has 1 or
 377 2 roots. If $b_0 = 0$ then the second root is $z^* = 0$. Now consider $b_0 > 0$. If $B'[1] > 1$, it
 378 is clear that $B[z] - z < 0$ for $z = 1^-$. Since $B[0] > 0$, there must be a second root in
 379 $(0, 1)$ by the Intermediate Value Theorem. We call this root z^* . If $B'[1] = 1$ or < 1 ,
 380 then $z = 1$ is the only root.

381 We now treat the two subcases $B'[1] > 1$ and $B'[1] = 1, < 1$ separately.

382 (i) Upon taking time derivatives of (3), we have

$$383 \quad \frac{\partial F(z, t)}{\partial t} = (B[z] - z)G(t) + \int_0^t \frac{\partial F}{\partial t}(z, t - \tau)B'[F(z, t - \tau)]G(\tau)d\tau, \quad (21)$$

384 where $\frac{\partial F}{\partial t}$ is treated as the unknown and F is the solution to (3). We look
 385 for a trajectory $F(z, t)$ that is constant in t for a fixed z . Then (21) implies
 386 $B[z] - z = 0$ so the only constant solutions are $F(z^*, t) = z^*$ and $F(1, t) = 1$. The
 387 kernel of (21) $B'[F(z, t - \tau)]G(\tau)$ is continuous in both arguments t and τ since
 388 G is continuous by assumption and F is continuous by Lemma 1(i). Therefore,
 389 by Lemma 4, we have $\frac{\partial F}{\partial t} \geq 0$ if $B[z] - z \geq 0$. Hence $F(z, t)$ is monotonically
 390 increasing in t if $z < z^*$ and monotonically decreasing if $1 > z > z^*$. Uniqueness of
 391 F in (3) for $t > 0$ [1] means that solutions cannot cross: increasing (decreasing)
 392 solutions are bounded from above (below) by z^* and so these solutions must
 393 asymptote to a constant, F_0 say. In (3), as $t \rightarrow \infty$, $z \int_t^{\infty} G(\tau)d\tau \rightarrow 0$ and
 394 $\int_0^t B[F(t - \tau)]G(\tau)d\tau \rightarrow B[F_0] \int_0^{\infty} G(\tau)d\tau$. Therefore F_0 satisfies $B[F_0] - F_0 = 0$
 395 or $F_0 = z^*, 1$. Therefore $F(z, t) \rightarrow z^*$ if $0 \leq z < 1$ and $F(z, t) = 1$ if $z = 1$.

396 (ii) If $z = 1$, then $B[z] - z = 0$ and $F(1, t) = 1$ is the only constant solution to (21).

397 If $z < 1$, $B[z] - z > 0$. Using Lemma 4 on (21), we find that $\frac{\partial F(z, t)}{\partial t} > 0$. Using
 398 a similar argument to (i), we conclude that $F(z, t) = 1$ is globally attracting.

399 \square

400 **Theorem 3** (Local uniqueness of lifetime probabilities from number distribution).

401 For equation (3), let $T > 0$ be the time at which $F(z, t)$ is measured for $0 \leq z \leq 1$.
 402 Consider an infinitesimal perturbation $\delta G(t)$ of the lifetime distribution $G(t)$ leading to
 403 a corresponding perturbation $\delta F(z, T)$. Let $\int_0^{\infty} \int_0^{\infty} G^2(t - \tau)dt d\tau < \infty$, with $G(t) = 0$
 404 if $t < 0$. Then $\delta F(z, T) \neq 0 \iff \delta G(t) \neq 0$.

405 *Proof.* Linearizing equation (3), $\delta G(t)$ and $\delta F(z, t)$ satisfy

$$406 \quad \delta F(z, t) - \int_0^t K(t, \tau, z)\delta F(z, \tau)d\tau = \delta S(z, t), \quad (22)$$

407 where

$$408 \quad \delta S(z, t) = \int_0^t d\tau (B[F(z, t - \tau)] - z)\delta G(\tau), \quad (23)$$

409 and

$$410 \quad K(t, \tau, z) = \frac{dB[F(z, \tau)]}{dF}G(t - \tau).$$

411 Equation (22) is a linear Volterra equation of the second kind. From Lemma 5, for a
 412 fixed z , $F(z, t)$ is bounded by 1 on $0 \leq t \leq \infty$ and $B'[z]$ is increasing in z . Therefore,
 413 since

$$\begin{aligned} \int_0^\infty \int_0^\infty |K^2(t, \tau, z)| dt d\tau &= \int_0^\infty \int_0^\infty \left| \frac{dB[F(z, \tau)]}{dF} G(t - \tau) \right|^2 dt d\tau, \\ &< B'[1]^2 \int_0^\infty \int_0^\infty |G(t - \tau)|^2 dt d\tau, \end{aligned}$$

414 the kernel $K(t, \tau, z)$ is square integrable on $0 \leq t, \tau \leq \infty$ and (22) can be solved in
 415 terms of a resolvent kernel Γ :

$$416 \quad \delta F(z, t) = \delta S(z, t) + \int_0^t \Gamma(t, \tau, z) \delta S(z, \tau) d\tau, \quad (24)$$

417 where Γ is defined through the Liouville-Neumann series:

$$\begin{aligned} \Gamma(t, \tau, z) &= K(t, \tau, z) + \int_\tau^t dt' K(t, t', z) K(t', \tau, z) \\ &\quad + \int_\tau^t dt' \int_\tau^{t'} dt'' K(t, t', z) K(t', t'', z) K(t'', \tau, z) + \dots \end{aligned}$$

418 Evaluating equation (24) at $t = T$, we have

$$419 \quad \delta F(z, T) = \delta S(z, T) + \int_0^T \Gamma(T, \tau, z) \delta S(z, \tau) d\tau. \quad (25)$$

420 To show $\delta G(\tau) \neq 0 \Rightarrow \delta F(z, T) \neq 0$, assume $\delta F(z, T) \equiv 0$ and aim for a contradiction.
 421 Then $\delta F(0, T) = 0$ and

$$422 \quad \delta S(0, T) + \int_0^T \Gamma(T, \tau, 0) \delta S(0, \tau) d\tau = 0. \quad (26)$$

423 From (23), the Laplace Transforms of δS and δG are related through $\delta \tilde{S}(0, s) =$
 424 $\mathcal{L}\{B[F(0, t)]\} \delta \tilde{G}(s)$, so that any given $\delta S(0, \tau)$ is obtained by taking $\delta G(\tau) =$
 425 $\mathcal{L}^{-1}\{\delta \tilde{S}(0, s) / \mathcal{L}\{B[F(0, t)]\}\}$: since $\delta G(\tau)$ can be freely chosen, $\delta S(0, \tau)$ can also be
 426 freely chosen. Let $\delta S(0, \tau)$ have the property that $\delta S(0, T) = 0$, but $\delta S(0, \tau) > 0$ for
 427 $0 \leq \tau < T$. Because $\Gamma(T, \tau, z) \geq 0$ for $\tau \in [0, \infty]$, $z \in [0, 1]$, (26) gives a contradiction.
 428 Essentially, (26) requires $\Gamma(T, \tau, 0)$ to be proportional to a delta distribution located
 429 at $\tau = T$, which contradicts the continuity of $\Gamma(T, \tau, 0)$.

430 Finally, from eq. (23) and lemma 3, $\delta G(t) \equiv 0$ implies that $\delta S(z, t) \equiv 0$ for all
 431 $(z, t) \in [0, 1] \times [0, T]$ and so eq. (25) implies that $\delta F(z, T) \equiv 0$ for $0 \leq z \leq 1$. \square

432 3.2. Reconstruction of progeny distribution \mathbf{b} given lifetime distribution $G(t)$

433 **Lemma 6.** *Let $G(t)$ and $P(t)$ be continuous for $t > 0$. For the two integral equations*

$$\begin{aligned} y_1(t) &= c_1 \int_0^t y_1(t - \tau) G(\tau) d\tau + P(t), \\ y_2(t) &= c_2 \int_0^t y_2(t - \tau) G(\tau) d\tau + P(t), \end{aligned}$$

434 *assume that $c_1, c_2 > 0$. Then $y_1(t) \geq y_2(t)$ for all $t > 0$ if $c_1 \geq c_2$.*

435 *Proof.* Let $c_2 > c_1 > 0$. Then for $t > 0$ we have

$$\begin{aligned} y_1(t) - y_2(t) &= \int_0^t [c_1 y_1(t - \tau) - c_2 y_2(t - \tau)] G(\tau) d\tau, \\ &< c_2 \int_0^t [y_1(t - \tau) - y_2(t - \tau)] G(\tau) d\tau, \end{aligned}$$

436 so that

$$437 \quad y_1(t) - y_2(t) = c_2 \int_0^t [y_1(t - \tau) - y_2(t - \tau)] G(\tau) d\tau + C(t),$$

438 where $C(t) < 0$ for all $t > 0$. Since $y_1(t)$ and $y_2(t)$ are continuous for $t > 0$ [23],
439 $C(t) \equiv y_1(t) - y_2(t) - c_2 \int_0^t [y_1(t - \tau) - y_2(t - \tau)] G(\tau) d\tau$ is also continuous for $t > 0$.
440 By Lemma 4, $y_1(t) - y_2(t) < 0$. The proof for $c_2 < c_1$ is similar. \square

441 **Lemma 7.**

442 *Consider the two integral equations*

$$\begin{aligned} y_1(t) &= \int_0^t y_1(t - \tau) G(\tau) d\tau + c_1 P(t) + Q(t), \\ y_2(t) &= \int_0^t y_2(t - \tau) G(\tau) d\tau + c_2 P(t) + Q(t), \end{aligned}$$

443 *and assume $Q(t)$ is continuous and $G(t)$ and $P(t)$ are positive and continuous for*
444 *$t > 0$. Then $y_1(t) \geq y_2(t)$ for all $t > 0$ if $c_1 \geq c_2$.*

445 *Proof.* The proof follows immediately if we consider the integral equation for $y_1 - y_2$
446 and apply Lemma 4. \square

447 **Theorem 4** (Global uniqueness of progeny number probabilities from a number
448 distribution). *Consider two branching processes with the same continuous lifetime*
449 *distribution $G(t)$. Suppose their progeny number distributions, defined by their*
450 *generating functions $B_1[z]$ and $B_2[z]$, give rise to the pgfs of two number distributions*
451 *$F_1(z, T)$ and $F_2(z, T)$ at time $t = T$. Then $F_1(z, T) = F_2(z, T)$ for all z if and only if*
452 *$B_1[z] \equiv B_2[z]$.*

453 *Proof.* Our proof relies on a local analysis of the trajectories of (3) near the equilibrium
454 points $z = z^*$. Recall that such points satisfy $B[z] - z = 0$. Unlike in lemma 5, here
455 we allow $0 \leq z^* \leq 1$: in the critical and subcritical cases, $z^* = 1$ is attracting and in
456 the supercritical case $0 \leq z^* < 1$ is attracting. Although our method is local in z , it is
457 *global* in t . The idea is to linearize eq. (3) about the equilibrium points and show that
458 when $B_1[z] \not\equiv B_2[z]$, if z is close to z^* then $F_1(z, t) \neq F_2(z, t)$ for $t > 0$. Moreover, we
459 show that if $B_1[z] \equiv B_2[z]$, $F_1(z, t) = F_2(z, t)$ for $t \geq 0$ and $z \in [0, 1]$. In other words:

- 460 (i) When $B_1[z] \equiv B_2[z]$, $F_1(z, T) = F_2(z, T)$ for all $T \geq 0$ and $0 \leq z \leq 1$.
461 (ii) When $B_1[z] \not\equiv B_2[z]$, there cannot exist a $T > 0$ such that $F_1(z, T) = F_2(z, T)$
462 for all $0 \leq z \leq 1$.

463 Item (i) is quickly proved as follows: first suppose that $B_1[z] \equiv B_2[z]$. Since
 464 the solution to eq. (3) is unique for a given branching function [19], we must have
 465 $F_1(z, t) = F_2(z, t)$ on $[0, 1] \times [0, \infty]$ and clearly $F_1(z, T) = F_2(z, T)$.

466 The remainder of the proof is to show (ii). Let z_1^* be any equilibrium point of
 467 $B_1[z]$ and z_2^* be any equilibrium point of $B_2[z]$. If $z_1^* \neq z_2^*$, then without loss of
 468 generality, let $z_1^* < z_2^*$ and fix a point z satisfying $z_1^* < z < z_2^*$. Then by Lemma 5,
 469 $F_1(z, t) \rightarrow z_1^*$ is a strictly decreasing trajectory in t while $F_2(z, t) \rightarrow z_2^*$ is a strictly
 470 increasing trajectory. Therefore $F_1(z_1^*, t) \neq F_2(z_2^*, t)$ for $t > 0$ so there cannot exist a
 471 $T > 0$ such that $F_1(z, T) = F_2(z, T)$ for all $0 \leq z \leq 1$.

472 Now suppose that $z_1^* = z_2^* = z^*$ (this would be the case, for example, when neither
 473 branching process is supercritical and $z_1^* = z_2^* = 1$). For $|\delta z| \ll 1$, the Bellman-Harris
 474 integral equation is still satisfied at $z = z^* + \delta z$:

$$475 \quad F_j(z^* + \delta z, t) = (z^* + \delta z) \int_t^\infty G(\tau) d\tau + \int_0^t B_j[F_j(z^* + \delta z, t - \tau)] G(\tau) d\tau, \quad (27)$$

476 for $j = 1, 2$. Expanding F_1, B_1 and F_2, B_2 in a Taylor series about z^* , we find

$$F_1(z^* + \delta z, t) = \sum_{n=0}^{\infty} \frac{F_1^{(n)}(z^*, t)}{n!} \delta z^n, \quad (28)$$

$$B_1[z^* + \delta z] = \sum_{n=0}^{\infty} \frac{B_1^{(n)}[z^*]}{n!} \delta z^n, \quad (29)$$

$$F_2(z^* + \delta z, t) = \sum_{n=0}^{\infty} \frac{F_2^{(n)}(z^*, t)}{n!} \delta z^n, \quad (30)$$

$$B_2[z^* + \delta z] = \sum_{n=0}^{\infty} \frac{B_2^{(n)}[z^*]}{n!} \delta z^n, \quad (31)$$

477 where $F'(z, t) = \partial_z F(z, t)$ and $F^{(m)}(z, t) = \partial_z^m F(z, t)$. Our strategy is to substitute
 478 (28)-(31) into (27), and equate powers of δz^n to find *linear* integral equations for
 479 $F_j^{(n)}(z^*, t)$.

480 If $B_1[z] \neq B_2[z]$, some terms of their Taylor series must differ. Suppose
 481 $B_1^{(n)}[z^*] = B_2^{(n)}[z^*]$ for $n = 0, 1, \dots, m-1$ but $B_1^{(m)}[z^*] \neq B_2^{(m)}[z^*]$. We complete the
 482 proof by showing that $F_1^{(m)}(z^*, t) \neq F_2^{(m)}(z^*, t)$ for $t > 0$.

483 • Case $m = 1$

484 Suppose $B_1[z^*] = B_2[z^*]$ and $B_1'[z^*] \neq B_2'[z^*]$. Then at $O(\delta z^0)$, $F_1^{(0)}(z^*, t) =$
 485 $F_2^{(0)}(z^*, t) = z^*$ are solutions to eq. (3). However, at $O(\delta z)$, F_1' and F_2' satisfy

$$F_1'(z^*, t) = \int_t^\infty G(\tau) d\tau + B_1'[z^*] \int_0^t F_1'(z^*, t - \tau) G(\tau) d\tau, \quad (32)$$

$$F_2'(z^*, t) = \int_t^\infty G(\tau) d\tau + B_2'[z^*] \int_0^t F_2'(z^*, t - \tau) G(\tau) d\tau. \quad (33)$$

486 By Lemma 6, $F_1'(z^*, t) \neq F_2'(z^*, t)$ for $t > 0$ so $F_1(z, t) \neq F_2(z, t)$ for $t > 0$ when
 487 $z = z^* + \delta z$.

488 • Case $m = 2$

489 For this case, $B_1[z^*] = B_2[z^*]$, $B_1'[z^*] = B_2'[z^*]$, $F_1^{(0)}(z^*, t) = F_2^{(0)}(z^*, t)$ and (32)

490 and (33) imply $F_1'(z^*, t) = F_2'(z^*, t) \equiv F'(z^*, t)$: we have to go to higher order in
 491 δz to show $F_1(z, t) \neq F_2(z, t)$.

492 Suppose that $B_1'[z^*] = B_2'[z^*] = B'[z^*]$ but that $B_1''[z^*] \neq B_2''[z^*]$. Then at
 493 $O(\delta z^2)$, F_1'' and F_2'' satisfy

$$F_1''(z^*, t) = B'[z^*] \int_0^t F_1''(z^*, t - \tau) G(\tau) d\tau + B_1''[z^*] P(t),$$

$$F_2''(z^*, t) = B'[z^*] \int_0^t F_2''(z^*, t - \tau) G(\tau) d\tau + B_2''[z^*] P(t),$$

494 where $P(t) \equiv \int_0^t (F'(z^*, t - \tau))^2 G(\tau) d\tau$ is continuous. It is clear that the second
 495 derivative of B is now responsible for distinguishing the two solutions $F_1(z^* + \delta z, t)$
 496 and $F_2(z^* + \delta z, t)$. Upon using Lemma 7, $F_1''(z^*, t) \neq F_2''(z^*, t)$ for $t > 0$ and
 497 therefore $F_1(z, t) \neq F_2(z, t)$ for $t > 0$.

498 • Case $m > 2$

499 Suppose that $B_1^{(n)}[z^*] = B_2^{(n)}[z^*] = B^{(n)}[z^*]$ for $n = 0, 1, \dots, m - 1$, $B_1^{(m)}[z^*] \neq$
 500 $B_2^{(m)}[z^*]$ and $F_1^{(n)}(z^*, t) = F_2^{(n)}(z^*, t) = F^{(i)}(z^*, t)$ for $n = 0, 1, \dots, m - 1$. To
 501 order $O(\delta z^m)$ we have

$$F_1^{(m)}(z^*, t) = B'[z^*] \int_0^t F_1^{(m)}(z^*, t - \tau) G(\tau) d\tau + B_1^{(m)}[z^*] P_m(t) + Q_m(t), \quad (34)$$

$$F_2^{(m)}(z^*, t) = B'[z^*] \int_0^t F_2^{(m)}(z^*, t - \tau) G(\tau) d\tau + B_2^{(m)}[z^*] P_m(t) + Q_m(t), \quad (35)$$

502 where

$$P_m(t) = \int_0^t \left[\frac{\partial F}{\partial z}(z^*, t - \tau) \right]^m G(\tau) d\tau,$$

$$Q_m(t) = \int_0^t G(\tau) R_m[F', \dots, F^{(m-1)}; B'', \dots, B^{(m-1)}] d\tau,$$

503 and $R_m[F'(z^*, t - \tau), \dots, F^{(m-1)}(z^*, t - \tau); B''[z^*], \dots, B^{(m-1)}[z^*]]$ is an algebraic
 504 function whose exact form is not important. Applying Lemma 7 on eqs. (34)
 505 and (35), we conclude that $F_1^{(m)}(z^*, t) \neq F_2^{(m)}(z^*, t)$ for $t > 0$.

506 □

507 **4. Conclusions**

508 We have investigated the reconstructibility of lifetime distributions and progeny
 509 number distributions of a single-species Bellman-Harris process. We assumed perfectly
 510 known data of two types: an extinction time distribution (equivalent to the survival
 511 probability of the processes), and the total number distribution at a single fixed time
 512 T . We find that for a given extinction time distribution, both the lifetime distribution
 513 function $G(\tau)$ and progeny number distribution \mathbf{b} can be found provided the other
 514 is given. Moreover, these distributions are unique and injective with respect to the
 515 cumulative extinction probability $F(0, t)$.

516 For perfect number distributions, known at a single fixed time T , we show local
 517 identifiability of the lifetime distribution $G(\tau)$ only for $\tau < T$. In fact, by writing eq.
 518 (3) with $t = T$ as

$$519 F(z, T) = z + \int_0^T (B[F(z, T - \tau)] - z) G(\tau) d\tau,$$

520 it is easy to see that two functions $G(\tau)$ that are identical for $\tau \leq T$, but not for $\tau > T$,
 521 would yield the same number distribution $f_n(T)$. However, we find that if the lifetime
 522 distribution is known, the progeny number distribution is uniquely determined by the
 523 number distribution $F(z, T)$.

524 Again, we emphasize that Theorems 1 and 2 give reconstructions for sensible
 525 $G(t)$ and $\{b_j\}$ only if the exact data were derived from an underlying Bellman-Harris
 526 process. Formulas (6) and (11)-(12) assume that the underlying process that generated
 527 $F(0, t)$ is a Bellman-Harris process. For an arbitrary set of exact data $F(0, t)$, however,
 528 there is no guarantee that $G(t)$ is a pdf or $\{b_j\}$ are a set of probabilities. For example,
 529 if $F(0, t)$ is generated by a process where particles branch after living for a random
 530 amount of time *and* are removed at a time-dependent rate, applying formulas (6) and
 531 (11)-(12) could generate $G(t)$ that do not integrate to unity or $\{b_j\}$ that do not sum
 532 to unity. In fact, whether or not sensible $G(t)$ or $\{b_j\}$ can be obtained would provide
 533 arguments for model selection. We leave as future work the task of characterizing in
 534 more detail those functions $F(0, t)$ that are the cumulative extinction probabilities of
 535 some underlying Bellman-Harris branching process.

536 Our results define which parameter functions of the Bellman-Harris process can be
 537 uniquely determined. These findings may guide numerical and statistical approaches
 538 to inference of branching processes, even though perfect data is not available. Possible
 539 extensions of our analyses include the utilization of perfectly measured pgfs at
 540 multiple discrete times $F(z, T_1), F(z, T_2), \dots$, to better reconstruct the parameterizing
 541 functions in the branching model.

542 This work was supported by the National Science Foundation through grants
 543 DMS-1021818 and DMS-1032131, and the Army Research Office through grant
 544 58386MA.

545 References

- 546 [1] Kimmel M and Axelrod D E 2002 *Branching Processes in Biology* (New York: Springer-Verlag)
 547 [2] Athreya K B and Ney P E 1972 *Branching Processes* (New York: Springer)
 548 [3] Allen L J S 2003 *An Introduction to Stochastic Processes with Applications to Biology* (Upper
 549 Saddle NJ: Pearson Education Inc)
 550 [4] Harris T E 1948 Branching Processes *Ann. Math. Statist.* **19** 474-494
 551 [5] Sorarrain O M, Boggio R and Cordo C 1979 An actual Bellman-Harris epidemic process. The
 552 infection of septoria in barley or wheat *J. theor. Biol.* **80** 425-433
 553 [6] Zilman A, Ganusov V V and Perelson A S 2010 Stochastic Models of Lymphocyte Proliferation
 554 and Death *PLoS One* **5** 9
 555 [7] Antal T and Krapivsky P L 2010 Exact solution of a two-type branching process: clone size
 556 distribution in cell division kinetics *J. Stat. Phys.-Theory and Experiment* P07028
 557 [8] Klein A M, Doupe D P, Jones P H and Simons B D 2007 Kinetics of cell division in epidermal
 558 maintenance *Physical Review E* **76** 021910
 559 [9] Jacobson M E 1985 Computation of extinction probabilities for the Bellman-Harris branching
 560 process *Mathematical Biosciences* **77** 173-177
 561 [10] Hyrien O, Mayer-Pröschel M, Noble M and Yakovlev A 2005 A Stochastic Model to Analyze
 562 Clonal Data on Multi-Type Cell Populations *Biometrics* **61** 199-207
 563 [11] Hyrien O 2007 Pseudo-likelihood estimation for discretely observed multitype Bellman-Harris
 564 branching processes *J. Statistical Planning and Inference* **137** 1375-1388
 565 [12] Bhat B R and Adke S R 1981 Maximum Likelihood Estimation for Branching Processes with
 566 Immigration *Advances in Applied Probability* **13** 498-509
 567 [13] Kale M M and Deshmukh S R 1993 Maximum likelihood estimation in branching processes with
 568 continuous state space *Statistics and Prob. Letters* **18** 245-251
 569 [14] Waugh W A O'N 1955 Age-dependent birth and death process *Biometrika* **42** 291-306

- 570 [15] Nedelman J, Downs H and Pharr P 1987 Inference for an age-dependent, multitype branching-
571 process model of mast cells *Journal of Mathematical Biology* **25** 203–226
- 572 [16] Dion J P 1975 Estimation of the Variance of a Branching Process *The Annals of Statistics* **3**
573 1183–1187
- 574 [17] Johnson R A, Susarla V and van Ryzin J 1979 Bayesian non-parametric estimation for age-
575 dependent branching processes *Stochastic Processes and their Applications* **9** 307–318
- 576 [18] Erdelyi A, *Asymptotic Expansions* 1956 (Dover)
- 577 [19] Kiffe T and Stecher M 1978 Existence and uniqueness of solutions to abstract volterra integral
578 equations *Proceedings of the American Mathematical Society* **68** 169–175
- 579 [20] Ablowitz M J and Fokas A. S. 1997 *Complex Variables* (Cambridge University Press)
- 580 [21] Feller W *An Introduction to Probability Theory and its Applications, Vol. 2* 1957 (Wiley)
- 581 [22] Tricomi FG 1957 *Integral Equations* (New York: Interscience Publishers)
- 582 [23] Kress R 1989 *Linear Integral Equations* (Springer-Verlag Berlin)