Received: 8 October 2024 | Accepted: 4 September 2025

DOI: 10.1111/2041-210X.70163

BRITISH ECOLOGICAL SOCIETY Methods in Ecology and Evolution

RESEARCH ARTICLE

A branching process approach to lifetime reproductive success of structured populations: Variance–covariance and distribution

Christophe F. D. Coste[1,2]

[1]Department of Terrestrial Ecology, Norwegian Institute for Nature Research, Trondheim, Norway

[2]Department of Biosciences, Swansea University, Singleton Park, UK

Correspondence
Christophe F. D. Coste
Email: christophe.coste@nina.no; c.f.d.coste@swansea.ac.uk; c.f.d.coste.ecology@gmail.com

Funding information
Natural Environment Research Council, Grant/Award Number: NE/W006731/1; Norges Forskningsråd, Grant/Award Number: 223257 and 343398

Handling Editor: Andres Lopez-Sepulcre

Abstract

1. Lifetime reproductive success (LRS) is a key metric in ecology and evolution. It measures the number of offspring produced by an individual during its lifetime. In epidemiology, it corresponds to the number of secondary cases generated by an infected individual.
2. For structured populations, it is crucial to understand how the distribution of LRS is shaped by survival, reproduction and other processes embedded in the projection model (such as dispersal or trait inheritance). Previous approaches have used diverse tools, such as Markov chains with rewards, to tackle this question, but generally consider only the total number of offspring produced, ignoring the distribution of their types (newborn states).
3. Here, we use the framework of branching processes to derive formulas for the variance–covariance matrices and probability-generating functions (allowing us to obtain the joint distribution) of LRS structured by the type of the parent and the type of the offspring. Furthermore, this framework leads to a simple algorithm providing a numerical approximation of the distribution of total LRS that does not require explicit expression of probability-generating functions.
4. We illustrate the power of the branching process approach, further, by studying the asymptotic behaviour of LRS, that is, the probability of producing many offspring. Finally, our general approach is applicable to any structured population model, and we provide R and Matlab code to facilitate implementation.

KEYWORDS
demographic stochasticity, lifetime reproductive output, lifetime reproductive success, multidimensional branching process, multi-type branching processes, net reproductive rate, probability-generating function, R0

1 | INTRODUCTION

In population ecology and demography, lifetime reproductive success (LRS), also known as lifetime reproductive output, quantifies the number of offspring produced by an individual over its lifetime. It is a key metric in population ecology and demography, as well as in epidemiology, where it represents the number of secondary cases generated by an infected individual in a fully susceptible population (Dublin & Lotka, 1925; Snyder et al., 2021; Tuljapurkar et al., 2020; Vegvari et al., 2022; Waples, 2023). Theoreticians,

---

over the last four decades, have investigated how to obtain the characteristics of the LRS of a population from its demography. In a structured population model, individuals are categorised by *states*, corresponding to the classes of one or several traits (e.g. age, stage, location, genotype and body mass). The demography of a structured population consists of the survival and fertility rates of individuals in each state and of the allocation of the survivor/newborn to a state. It therefore embeds ecological processes—such as dispersal/migration (if the structure is spatial), ageing (if the model is age-structured) and trade-offs—whose effects on the LRS are of crucial importance (see the introductions in Tuljapurkar et al., 2020, 2021). When the state-space and time are discrete, such projection models correspond, in their deterministic form, to matrix population models (Caswell, 2001). While the expectation of LRS for a structured population is rather simple to fathom—it corresponds to the 'next-generation matrix' (see Caswell, 2001; Cushing & Zhou, 1994; Diekmann et al., 1990)—its further moments (e.g. variance) and more generally its distribution are less easy to compute.

In structured populations, there is individual heterogeneity (individuals are not all considered identical); that is, there are several individual *states*. Following Ellner (2018), we call *types* the subset of this structure that corresponds to *newborn states*. Any individual in the population is therefore allocated a *type*, which corresponds to the *state* it was born in, and which remains, by definition, constant throughout its life trajectory. When several *types* (e.g. two sexes, several genotypes or a range of body mass at birth) are incorporated in a projection model, they may correspond to different life trajectory prospects and in particular to different LRS. In contrast, for a model with a single *type*, all individuals have the same prospect at birth: they are born in the same state (e.g. age 1, stage juvenile or instar egg) and share therefore the same distribution of LRS. Historically, population ecologists and demographers have primarily focused on models with a single *type* (Lefkovitch, 1965; Leslie, 1945; Lewis, 1942), where LRS corresponds to a univariate variable: the total number of offspring produced (total LRS). Theoreticians have developed formulas for its variance (and further moments) and a method to compute its distribution. For age-structured populations, the variance was formulated by Charlesworth and Williamson (1975) (see also Waples et al., 2011). This was extended for the more general framework of populations structured by stages by van Daalen and Caswell (2017). Recently, Tuljapurkar et al. (2020) published the first method to compute the distribution of LRS for such models.

In contrast, for models with several *types*, LRS can be decomposed by these *types*. Such models can incorporate inheritance and are critical for understanding the evolutionary demography of populations and bridging genetic and demographic projection models. If the *type* (e.g. body mass at birth) covaries with fitness (e.g. heavier newborns produce more offspring than lighter ones), reproduction may embed a quantity–quality trade-off: an individual producing a single large newborn may contribute more to the population's fitness than one producing two smaller newborns. By focusing solely on total LRS, rather than its distribution by offspring *types*, one overlooks these differences in fitness contributions. In models with multiple newborn states, the LRS is a multivariate variable: it corresponds to the number of offspring of each *type* produced by an individual (of a given *type*) during its entire lifetime. The variance of LRS for a model with multiple *types* is therefore a variance–covariance matrix and, as of today, a general formulation for it is lacking. More generally, a method has yet to be developed to obtain the joint distributions of the number of offspring of each *type*.

Two individuals in the same state, and therefore with the same vital rates, for example a probability to survive of 0.5, may have different fates: One may die while the other survives the time-step. This component of stochasticity is called demographic stochasticity (Engen & Sæther, 1998). It propagates to the population level, generating stochasticity in the number of offspring produced, that is, in the LRS. Branching processes (BPs) provide a framework that is well-suited to studying demographic stochasticity. Originally developed to model the extinction of family names (Bienaymé, 1845; Watson & Galton, 1875), BPs have remained an important tool, for mathematicians, to study the consequences of demographic stochasticity on population dynamics (Haccou et al., 2005). They have been used extensively, for example, to study kinship networks (Caswell, 2024; Jagers, 1982; Joffe & Waugh, 1985; Waugh, 1981). However, despite their relevance, BPs remain underutilised in ecology. The computation of the variance in LRS by Caswell (2011) and van Daalen and Caswell (2017) is based on Markov chains with rewards and the method of Tuljapurkar et al. (2020) to compute its distribution on convolutions and Fourier transforms. We believe that the markedly different formalisms of the mathematical frameworks underlying branching processes and projection models (compare Equation 2c with Equation 2a) have led many researchers studying population dynamics—whether mathematicians, economists, demographers or ecologists—to specialise in one approach or the other. However, BPs are multiplicative processes and, as such, they can also be modelled via matrices (of random variables, e.g. see Athreya & Ney, 1972; Pollard, 1966), which can be manipulated similarly to the more familiar (deterministic) matrix population models (Caswell, 2001).

In this paper, we use the BP framework, in its matrix form, to study the variance–covariance and distribution of LRS for a general population model that may embed several newborn states (*types*). We first briefly present the main results of BPs pertaining to the projection of abundances when accounting for demographic stochasticity and introduce a simplifying notation for such projections that bridges the divide between classic BP and matrix projection models notations. This allows us, from a simple stochastic equation relating the LRS to the time-step survival and reproduction, to produce formulas for its variance–covariance and its probability-generating function, a function from which the probability distribution of LRS can be computed. We go on to provide an algorithm to compute this distribution faster, without resorting explicitly to such functions, for total LRS. Finally, we demonstrate the power of the BP approach by studying, theoretically, the asymptotic behaviour of the probability distribution of LRS (i.e. the probability to produce a large number of offspring).

2 | DEMOGRAPHIC STOCHASTIC PROJECTIONS AS BRANCHING PROCESSES

In this section, we go through the most important aspects of the mathematical field of BPs for demographic projection and introduce a simplifying notation to facilitate their manipulations for ecologists. We summarise, in Box 1, the most relevant properties of BPs of structured populations, using the simplifying notation and focusing on the matrices of random variables projecting the population over time. These properties are relevant for both the projection of abundances (this section 2) and LRS (section 3). We also provide, in Box 2, an illustrated summary of the properties of probability-generating functions (*pgfs*) for unstructured and structured populations, and as Table 1, a summary of the notations used in this paper. Intermediate steps and proofs of the formulas provided in this manuscript (for this section, the following ones and the boxes) can be found in Supporting Information: Section I. We provide the code (R and Matlab) for the numeric illustrations, to produce the figures and to implement the algorithm of section 4 (Supporting Information: Section VII); the code is easily adaptable to any population projection model. The formulas of this section are well-known results of BPs (that can be found in books and articles, e.g. Athreya & Karlin, 1971; Haccou et al., 2005; Harris, 1963; Karlin & Taylor, 1975; Pollard, 1966).

2.1 | Unstructured population model with non-overlapping generations

Consider a population where all individuals are identical (unstructured model) and only live one time-step (i.e. the time-step is the generation, and generations do not overlap). In a constant environment, the population dynamics of this model is completely defined by the fertility (integer-valued) random variable f, corresponding to the number of offspring produced (in a given time-step, i.e. in a given generation) by an individual. We assume that the reproduction of every individual in every generation is a particular realisation of f, independent from that of the other individuals. The number of individuals in the population at time t, n_t, is then a stochastic process, which obeys the following stochastic projection equation:

$$n_{t+1} = \underbrace{f^{(1)} + f^{(2)} + \ldots + f^{(n_t)}}_{\text{sum of } n_t \text{ independent r.v.s}}, \tag{1a}$$

where $f^{(k)}$ is the random variable (r.v.) which takes the value of the number of individuals produced by the k^{th} individual of the population at time t. Each $f^{(k)}$ is ‘equal in distribution’ to f, denoted as

$$f^{(k)} \overset{d}{=} f \tag{1b}$$

The population abundance at time $t+1$ is, therefore, the sum of n_t independent and identically distributed (i.i.d.) random variables, each with the distribution of f. The random variable f can be considered as a ‘demographic stochastic projector’ (DSP), which projects the population abundance over time, by applying to each individual in the population independently. We simplify the above notation via the introduction of a new formalism for the demographic stochastic population projection of n_t by f:

$$n_{t+1} = f \Diamond n_t \tag{1c}$$

The n_t stochastic process is called a Galton–Watson BP. Note that the $\Diamond$ operator should not be confused with the product of two random variables.

2.1.1 | Expectation and variance

The following results stem from important properties of DSPs (summarised, in Box 1, for the general, structured case). From the expectation of f and n_t, one can, via the law of total expectation, compute that of n_{t+1}:

$$E\left(n_{t+1}\right) = E(f)E\left(n_t\right), \tag{1d}$$

which leads to $E\left(n_t\right) = E(f)^t n_0$. If $E(f) > 1$ (respectively, $E(f) < 1$), the BP is said to be supercritical (resp., subcritical) and the expected population size increases (resp., decreases) exponentially over time. Via the law of total variance, we have, further:

$$var\left(n_{t+1}\right) = var(f)E\left(n_t\right) + E(f)^2 var\left(n_t\right). \tag{1e}$$

2.1.2 | Probability-generating function

The probability-generating function (*pgf*, see Box 2) of a (univariate, integer-valued) r.v., such as f, is a function of a dummy variable (that we denote x):

$$\varphi_f(x) = \sum_{k=0}^{+\infty} Pr(f = k)x^k \tag{1f}$$

From the *pgf* of a r.v., one can compute its probability distribution (Box 2, Equation 6b). An important property of these functions is that the *pgf* of the demographic stochastic projection of a r.v. by another one corresponds to the composition of their *pgfs* (Equation 3e, in Box 1, for the general, structured case).

$$\varphi_{n_{t+1}}(x) = \varphi_{f \Diamond n_t}(x) = \varphi_{n_t}\left(\varphi_f(x)\right), \tag{1g}$$

which allows, in principle, to compute the probability distribution of n_t for any t.

2.2 | Structured population model with non-overlapping generations

Let us now consider the case where the population is structured (i.e. heterogeneous: individuals are not all identical) and generations, as per the Galton–Watson BP, do not overlap (all individuals

BOX 1 Notations for, and properties of, demographic stochastic models for structured populations.

All proofs to be found in Supporting Information: Section I and notations in Table 1. We denote with a $\diamond$ symbol, the projection over time of a (stochastic) population abundance vector $\boldsymbol{n}$ of length s (a multivariate integer-valued r.v.) by a 'demographic stochastic projector' (DSP) $\mathbf{A}$ of dimension $a \times s$ (a matrix of integer-valued random variables) where $A(i, j)$ represents the number of i individuals ($1 \le i \le a$) produced by a j individual ($1 \le j \le s$) via process $\mathbf{A}$:

$$\mathbf{A} \Diamond \boldsymbol{n} \quad = \quad \underbrace{\mathbf{A}_1^{(1)} + \mathbf{A}_1^{(2)} + \ \dots \ + \mathbf{A}_1^{(n(1))}}_{\substack{\text{sum of } n(1) \text{ i.i.d. sto. vectors} \\ \text{each with the distrib. of } \mathbf{A}_1}} \ + \ \underbrace{\mathbf{A}_2^{(1)} + \mathbf{A}_2^{(2)} + \ \dots \ + \mathbf{A}_2^{(n(2))}}_{\substack{\text{sum of } n(2) \text{ i.i.d. sto. vectors} \\ \text{each with the distrib. of } \mathbf{A}_2}} \ + \ \dots \ + \ \underbrace{\mathbf{A}_s^{(1)} + \mathbf{A}_s^{(2)} + \ \dots \ + \mathbf{A}_s^{(n(s))}}_{\substack{\text{sum of } n(s) \text{ i.i.d. sto. vectors} \\ \text{each with the distrib. of } \mathbf{A}_s}} \ , \tag{3a}$$

where $\mathbf{A}_j$ is the stochastic vector (of length a) corresponding to the j^{th} column of $\mathbf{A}$. Matrix $\mathbf{A}$ is the (horizontal) vector of these **independent** stochastic (vertical) vectors: $\mathbf{A} = \left[\mathbf{A}_1, \mathbf{A}_2, \ \dots \ , \mathbf{A}_s\right]$. $\mathbf{A} \diamond \boldsymbol{n}$ is an integer-valued stochastic vector of length a. Projecting over several time-steps via the same DSP $\mathbf{A}$ requires $a = s$ ($\mathbf{A}$ is then a square matrix of r.v.s). If $a = s = 1$, then $\mathbf{A}$ and $\boldsymbol{n}$ are integer-valued, univariate (i.e. scalar) r.v.s modelling an unstructured population. If $s = 1$ but $a > 1$, $\mathbf{A}$ is a stochastic vector. If $a = 1$ but $s > 1$, it is a (horizontal) vector of univariate r.v.s.

Expectation: From Equation (3a), via the law of total expectation, we have

$$\mathbf{E}(\mathbf{A} \diamond \boldsymbol{n}) = \mathbf{E}(\mathbf{A})\mathbf{E}(\boldsymbol{n}) \tag{3b}$$

This is the well-known result that the projection over time of the vector of expected population abundances is obtained by the (classic) matrix multiplication, on the left, of that vector by $\mathbf{E}(\mathbf{A})$, the expectation of $\mathbf{A}$, called a matrix projection model (Caswell, 2001). This equality justifies the use of the simplifying $\Diamond$ notation, as it behaves similarly, for demographic stochastic processes, to the matrix multiplication for deterministic ones.

Variance–covariance: The variance of a stochastic vector, such as $\boldsymbol{n}$ (of length s), is a $s \times s$ variance–covariance matrix that we denote $\boldsymbol{var}(\boldsymbol{n})$, and that can be *vectorised* (i.e. its columns stacked upon one another, via the vec operator, see Henderson & Searle, 1981): $\boldsymbol{V}(\boldsymbol{n}) = \boldsymbol{vec}(\boldsymbol{var}(\boldsymbol{n}))$, producing a (vertical) vector of length s^2. Since the columns of DSP $\mathbf{A}$ are independent stochastic vectors, we can extend the $\boldsymbol{V}$ notation and organise the variance of the various $\mathbf{A}_j$ into one matrix of size $a^2 \times s$: $\boldsymbol{V}(\mathbf{A}) = \left[\boldsymbol{V}(\mathbf{A}_1), \boldsymbol{V}(\mathbf{A}_2), \ \dots \ , \boldsymbol{V}(\mathbf{A}_s)\right]$. Via the law of total variance:

$$\boldsymbol{V}\left(\mathbf{A} \Diamond \boldsymbol{n}\right) = (\boldsymbol{E}(\mathbf{A}) \otimes \boldsymbol{E}(\mathbf{A}))\boldsymbol{V}(\boldsymbol{n}) + \boldsymbol{V}(\mathbf{A})\boldsymbol{E}(\boldsymbol{n}) \tag{3c}$$

Probability-generating function: We define the *pgf* of DSP $\mathbf{A}$, as the (horizontal) vector of the *pgf*s of the various stochastic vectors $\mathbf{A}_j$ (see Box 2):

$$\boldsymbol{\varphi}_{\mathbf{A}}(\boldsymbol{x}) = \left[\varphi_{\mathbf{A}_1}(\boldsymbol{x}), \varphi_{\mathbf{A}_2}(\boldsymbol{x}), \ \dots \ , \varphi_{\mathbf{A}_s}(\boldsymbol{x})\right] \tag{3d}$$

Note that the horizontal vector of *pgfs*, $\boldsymbol{\varphi}_{\mathbf{A}}$ is not denoted with the transpose sign in order to alleviate notations. From the law of total expectation, we have

$$\varphi_{\mathbf{A} \Diamond \boldsymbol{n}}(\boldsymbol{x}) = \varphi_{\boldsymbol{n}}\left(\boldsymbol{\varphi}_{\mathbf{A}}(\boldsymbol{x})\right) \tag{3e}$$

We extend the *'equivalent in distribution'* notation to DSPs:

$$\mathbf{A} \overset{d}{=} \mathbf{B} \text{ is equivalent to } \quad \forall j, \mathbf{A}_j \overset{d}{=} \mathbf{B}_j \ , \tag{3f}$$

where $\mathbf{A}$ and $\mathbf{B}$ must have same dimensions. Note that

$$\left[\text{for any stochastic vector } \boldsymbol{n}, \ \ \mathbf{A} \Diamond \mathbf{n} \overset{d}{=} \mathbf{B} \Diamond \mathbf{n}\right] \quad \text{is equivalent to} \quad \left[\mathbf{A} \overset{d}{=} \mathbf{B}\right]. \tag{3g}$$

BOX 1 (Continued)

Total DSP: From DSP $\mathbf{A}$ (or from a stochastic vector), one can consider 'Total $\mathbf{A}$', that is the vector of univariate r.v.s $\boldsymbol{a}^T = \mathbf{1}^T\mathbf{A}$:

$$\begin{cases} \mathbf{E}(a^T) = \mathbf{1}^T\mathbf{E}(\mathbf{A}) \\ \mathbf{V}(a^T) = \mathbf{1}^T\mathbf{V}(\mathbf{A}) \\ \varphi_{a^T}(y) = \varphi_{\mathbf{A}}(y, y, \dots, y) \end{cases} \tag{3h}$$

From stochastic vector $\boldsymbol{n}$, one can consider $n(i)$, the number of individuals in state i, with $var(n(i)) = [\mathbf{var}(\boldsymbol{n})](i,i)$ and $\varphi_{n(i)}(y) = \varphi_{\mathbf{n}}(1, 1, \dots, y, \dots, 1)$ (the y is in position i).

Sum of two DSPs: If $\mathbf{A}$ and $\mathbf{B}$ are two DSPs of same dimensions, one can consider their sum, defined as $\mathbf{A} + \mathbf{B} = [\mathbf{A}_1 + \mathbf{B}_1, \mathbf{A}_2 + \mathbf{B}_2, \cdots]$. Then, $(\mathbf{A} + \mathbf{B}) \Diamond \boldsymbol{n} = \mathbf{A} \Diamond \boldsymbol{n} + \mathbf{B} \Diamond \boldsymbol{n}$ and $\mathbf{E}(\mathbf{A} + \mathbf{B}) = \mathbf{E}(\mathbf{A}) + \mathbf{E}(\mathbf{B})$. If further they are such that, for every j, the fates of a given individual in j via $\mathbf{A}_j$ and $\mathbf{B}_j$ are **independent** (i.e. $\boldsymbol{cov}(\mathbf{A}_j, \mathbf{B}_j) = \mathbf{0}$), then $\mathbf{A}$ and $\mathbf{B}$ are said to be **independent**, which we denote $\boldsymbol{cov}(\mathbf{A}, \mathbf{B}) = \mathbf{0}$. In that case,

$$\begin{cases} \mathbf{V}(\mathbf{A}+\mathbf{B}) = \mathbf{V}(\mathbf{A}) + \mathbf{V}(\mathbf{B}) \\ \varphi_{(\mathbf{A}+\mathbf{B})}(\mathbf{x}) = \varphi_{\mathbf{A}}(\mathbf{x}) \bullet \varphi_{\mathbf{B}}(\mathbf{x}) = \varphi_{\mathbf{A}}(\mathbf{x})\boldsymbol{diag}(\varphi_{\mathbf{B}}(\mathbf{x})) \end{cases} \tag{3i}$$

with $\bullet$ denoting the termwise product and ***diag*** turning a vector into a diagonal matrix.

Successive DSPs: One can apply DSP $\mathbf{B}$ (of size $b \times a$) to the stochastic vector $\mathbf{A} \Diamond \mathbf{n}$ (Equation 3a). By extending the $\Diamond$ notation so that $\mathbf{B} \Diamond \mathbf{A} = [\mathbf{B} \Diamond \mathbf{A}_1, \mathbf{B} \Diamond \mathbf{A}_2, \cdots, \mathbf{B} \Diamond \mathbf{A}_s]$, we have that the $\Diamond$ operator is **associative**, that is $\mathbf{B} \Diamond (\mathbf{A} \Diamond \boldsymbol{n}) = (\mathbf{B} \Diamond \mathbf{A}) \Diamond \boldsymbol{n} = \mathbf{B} \Diamond \mathbf{A} \Diamond \boldsymbol{n}$. One can then consider directly the properties of DSP $\mathbf{B} \Diamond \mathbf{A}$, e.g.:

$$\begin{cases} E(\mathbf{B} \Diamond \mathbf{A}) = E(\mathbf{B})E(\mathbf{A}) \\ V(\mathbf{B} \Diamond \mathbf{A}) = (\mathbf{E}(\mathbf{B}) \otimes \mathbf{E}(\mathbf{B}))\boldsymbol{V}(\mathbf{A}) + \boldsymbol{V}(\mathbf{B})\boldsymbol{E}(\mathbf{A}) \\ \varphi_{(\mathbf{B} \Diamond \mathbf{A})}(\mathbf{x}) = \varphi_{\mathbf{A}}(\varphi_{\mathbf{B}}(\mathbf{x})) \\ \boldsymbol{cov}(\mathbf{A}, \mathbf{C}) = \mathbf{0} \text{ implies } \boldsymbol{cov}(\mathbf{B} \Diamond \mathbf{A}, \mathbf{C}) = \mathbf{0} \text{ for any DSP } \mathbf{C} \text{ of same dimension than } \mathbf{A} \end{cases} \tag{3j}$$

die before the end of the time-step). In such a model, an individual remains in its (newborn) state its entire life: all states are newborn states, that is *types*. If there are s states in the population, its dynamics in a constant environment is fully described by s stochastic vectors (i.e. by s multivariate r.v.s) of length s, that we denote $\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_s$. Vector $\mathbf{F}_j$ corresponds to the number of offspring produced by a type-j parent, structured by the types of the offspring. In particular the i^{th} entry of $\mathbf{F}_j$, denoted $F(i,j)$, is the random variable of the number of type-i offspring that a type-j parent produces. In any given time-step, the reproduction of an individual of type j is governed solely by $\mathbf{F}_j$, whose realisation is independent from that of other individuals (of the same type or other types). The vector of abundances at time t, $\boldsymbol{n}_t$, a stochastic vector of non-negative integers of length s, obeys the following stochastic projection equation

$$\boldsymbol{n}_{t+1} = \underbrace{\mathbf{F}_1^{(1)} + \mathbf{F}_1^{(2)} + \dots + \mathbf{F}_1^{(n_t(1))}}_{\text{sum of } n_t(1) \text{ independent r.v.}} + \dots + \underbrace{\mathbf{F}_s^{(1)} + \mathbf{F}_s^{(2)} + \dots + \mathbf{F}_s^{(n_t(s))}}_{\text{sum of } n_t(s) \text{ independent r.v.}}, \tag{2a}$$

where $\mathbf{F}_j^{(l)}$ corresponds to the reproduction of the l^{th} individual of type j in the population at time t $(1 \le l \le n_t(j))$ and has the distribution of $\mathbf{F}_j$. Extending the $\Diamond$ demographic stochastic projection notation to structured populations (Box 1, Equation 3a), we can write Equation (2a) as

$$\boldsymbol{n}_{t+1} = \mathbf{F} \Diamond \boldsymbol{n}_t \tag{2b}$$

where $\mathbf{F}$, the (horizontal) vector of the various (vertical, independent) stochastic vectors $\mathbf{F}_j$, that is $\mathbf{F} = [\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_s]$, is a matrix, of dimension $s \times s$, of integer-valued random variables. $\mathbf{F}$ is a DSP that projects the (stochastic) vector of abundances over time. The $\boldsymbol{n}_t$ stochastic process is called a Multi-Type BP.

2.2.1 | Expectation and variance

From the expectation of $\mathbf{F}$, that we denote $\overline{\mathbf{F}} = \mathbf{E}(\mathbf{F})$, we can project, from Equation (3b), the vector of expected abundances $\overline{\boldsymbol{n}}_t$ over time:

$$\overline{\boldsymbol{n}}_{t+1} = \overline{\mathbf{F}}\overline{\boldsymbol{n}}_t \tag{2c}$$

BOX 2 Probability-generating functions

Probability-generating function of a random variable: The *pgf* of a (univariate, integer-valued) random variable a is defined as:

$$\varphi_a(x) = E\left(x^a\right) = \sum_k Pr(a = k)x^k. \tag{6a}$$

It is a function of a dummy unknown x and represents the probability distribution of r.v. a. Parameter k goes through all non-negative integer values. From the *pgf* of a random variable, one can obtain, in turn, its probability distribution (hence the name), via successive derivatives taken at $x = 0$: We have $Pr(a = 0) = \varphi_a(0)$ and, in general,

$$Pr(a = k) = \frac{1}{k!}\varphi_a^{(k)}(0) \tag{6b}$$

where $\varphi_a^{(i)}$ denotes the i^{th} derivative of φ_a with respect to x.

Example: If a is a Poisson of parameter 1, we have $\varphi_a(x) = \sum_k \frac{e^{-1}}{k!}x^k = e^{x-1}$. In this particular case, all successive derivatives of the *pgf* are equal, ie. $\forall i, \varphi_a^{(i)}(x) = e^{x-1} = \varphi_a(x)$, which leads back to $Pr(a = k) = \frac{1}{k!}\varphi_a^{(k)}(0) = \frac{1}{k!}\varphi_a(0) = \frac{e^{-1}}{k!}$.

Probability-generating function of a random vector: The *pgf* of a stochastic (vertical) vector $\boldsymbol{a}$ of length b is defined as $\varphi_{\boldsymbol{a}}(\mathbf{x}) = E\left(x_1^{a(1)}x_2^{a(2)} \dots x_b^{a(b)}\right)$, that is:

$$\varphi_{\boldsymbol{a}}(\mathbf{x}) = \sum_{\boldsymbol{k}} Pr(\boldsymbol{a} = \boldsymbol{k})x_1^{k(1)}x_2^{k(2)} \dots x_b^{k(b)} = \sum_{k(1),k(2),\dots,k(b)} Pr\left(\begin{bmatrix} a(1) \\ a(2) \\ \dots \\ a(b) \end{bmatrix} = \begin{bmatrix} k(1) \\ k(2) \\ \dots \\ k(b) \end{bmatrix}\right)x_1^{k(1)}x_2^{k(2)} \dots x_b^{k(b)} \tag{6c}$$

It is a function of a dummy vector of length b, that we denote $\mathbf{x} = \left[x_1, x_2, \dots, x_b\right]$ and represents the probability distribution of $\boldsymbol{a}$ (i.e. the joint probability distributions of the $a(1), \dots, a(b)$). Each parameter $k(j)$ goes through all non-negative integer values. From the *pgf* of a random vector, one can obtain, in turn, its probability distribution, via successive derivatives taken at $\mathbf{x} = \mathbf{0}$: We have $Pr(\boldsymbol{a} = \mathbf{0}) = \varphi_{\boldsymbol{a}}(\mathbf{0})$ and, in general:

$$Pr(\boldsymbol{a}=\boldsymbol{k})=\frac{1}{k(1)!k(2)!\dots k(b)!}\varphi_{\boldsymbol{a}}^{(k(1),k(2),\dots,k(b))}(\mathbf{0}) \tag{6d}$$

where $\varphi_{\boldsymbol{a}}^{(i_1,i_2,\dots,i_b)}$ denotes the partial (mixed) derivative: $\varphi_{\boldsymbol{a}}^{(i_1,i_2,\dots,i_b)} = \frac{\partial^{i_1+i_2+\dots+i_b}\varphi_{\boldsymbol{a}}}{\partial x_1^{i_1}\partial x_2^{i_2}\dots\partial x_b^{i_b}}$.

Example: If $b = 2$ and $\boldsymbol{a}$ can take only three realisations – say $Pr\left(\boldsymbol{a} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \frac{1}{2}$, $Pr\left(\boldsymbol{a} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \frac{1}{3}$ and $Pr\left(\boldsymbol{a} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \frac{1}{6}$ – then

$$\varphi_{\boldsymbol{a}}(\mathbf{x}) = Pr\left(\boldsymbol{a} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right)x_1^0x_2^0 + Pr\left(\boldsymbol{a} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)x_1^0x_2^1 + Pr\left(\boldsymbol{a} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)x_1^1x_2^1 = \frac{1}{2} + \frac{1}{3}x_2 + \frac{1}{6}x_1x_2$$

.

We have $\varphi_{\boldsymbol{a}}^{(0,0)}(\mathbf{x}) = \varphi_{\boldsymbol{a}}(\mathbf{x})$, $\varphi_{\boldsymbol{a}}^{(0,1)}(\mathbf{x}) = \frac{d\left(\frac{1}{2} + \frac{1}{3}x_2 + \frac{1}{6}x_1x_2\right)}{dx_2} = \frac{1}{3} + \frac{1}{6}x_1$, $\varphi_{\boldsymbol{a}}^{(1,1)}(\mathbf{x}) = \frac{\partial^2\varphi_{\boldsymbol{a}}}{\partial x_1\partial x_2} = \frac{d\left(\frac{1}{6}x_2\right)}{dx_2} = \frac{1}{6}$ and for all other values of i_1 and i_2, $\varphi_{\boldsymbol{a}}^{(i_1,i_2)}(\mathbf{x}) = 0$. This leads back to $Pr\left(\boldsymbol{a} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \varphi_{\boldsymbol{a}}(\mathbf{0}) = \frac{1}{2}$, $Pr\left(\boldsymbol{a} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \varphi_{\boldsymbol{a}}^{(0,1)}(\mathbf{0}) = \frac{1}{3}$ and $Pr\left(\boldsymbol{a} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \varphi_{\boldsymbol{a}}^{(1,1)}(\mathbf{0}) = \frac{1}{6}$.

One can, further, project the variance–covariance of the population abundance vector (Equation 3c):

$$\boldsymbol{V}\left(\boldsymbol{n}_{t+1}\right) = \left(\overline{\mathbf{F}} \otimes \overline{\mathbf{F}}\right)\boldsymbol{V}\left(\boldsymbol{n}_t\right) + \boldsymbol{V}(\mathbf{F})\overline{\boldsymbol{n}}_t, \tag{2d}$$

where $\boldsymbol{V}(\mathbf{F}) = \left[\boldsymbol{V}\left(\mathbf{F}_1\right), \boldsymbol{V}\left(\mathbf{F}_2\right), \dots, \boldsymbol{V}\left(\mathbf{F}_s\right)\right]$ with $\boldsymbol{V}\left(\mathbf{F}_j\right)$ the vectorised version of the variance–covariance matrix of $\mathbf{F}_j$: $\boldsymbol{V}\left(\mathbf{F}_j\right) = \boldsymbol{vec}\left(\boldsymbol{var}\left(\mathbf{F}_j\right)\right)$ (Henderson & Searle, 1981) and where $\otimes$ is the Kronecker (or direct) product. $\boldsymbol{V}(\mathbf{F})$ is the (horizontal) vector of the variance–covariance of

TABLE 1 Table of notations.

Variable	Dimension	Definition
$\mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_b]$	$a \times b$	A demographic stochastic projector (DSP) representing a demographic process. It is a matrix of integer-valued random variables that consists of an horizontal vector of b independent random (vertical) vectors
$\mathbf{A}_j$ (j^{th} column of $\mathbf{A}$)	$a \times 1$	Random vector representing process $\mathbf{A}$ for individuals in state j
$\overline{\mathbf{A}} = E(\mathbf{A})$	$a \times b$	Expectation of matrix of random variables $\mathbf{A}$
Total $\mathbf{A}$: $\boldsymbol{a}^T = \mathbf{1}^T\mathbf{A}$	$1 \times b$	DSP where each column is univariate; $a(j) = \sum_i A(i,j)$ is the total number of individuals (irrespective of their state) produced by a j individual, via process $\mathbf{A}$.
$\boldsymbol{V}(\mathbf{A}_j) = \boldsymbol{vec}(\boldsymbol{var}(\mathbf{A}_j))$	$a^2 \times 1$	Variance–covariance matrix, in its vectorised form, of random vector $\mathbf{A}_j$
$\boldsymbol{V}(\mathbf{A}) = [\boldsymbol{V}(\mathbf{A}_1), \ldots, \boldsymbol{V}(\mathbf{A}_b)]$	$a^2 \times b$	Variance of $\mathbf{A}$ as the (horizontal) vector of the variance of its columns
$\varphi_{\mathbf{A}_j}(x_1, \ldots, x_a)$	1	Probability-generating function (*pgf*) of random vector $\mathbf{A}_j$
$\boldsymbol{\varphi}_{\mathbf{A}} = [\varphi_{\mathbf{A}_1}, \ldots, \varphi_{\mathbf{A}_b}]$	$1 \times b$	*pgf* of $\mathbf{A}$ as the (horizontal) vector of the *pgf*s of its columns
States and types		
s	1	Number of states structuring the population
d (with $1 \le d \le s$)	1	d is the number of newborn states (*types*) in the model, that is, the number of rows of $\overline{\mathbf{F}}$ that are non-zero
Particular DSPs		
$\mathbf{F}$	$s \times s$	Reproduction: $F(i,j)$ is the number of offspring in state i produced by an individual in state j during a time-step
$\mathbf{S}$	$s \times s$	Survival: $S(i,j) = 1$ implies that an individual in state j survives the time-step and transitions to state i
$\mathbf{R}$	$s \times s$	Eventual number of offspring: $R(i,j)$ is the number of offspring in i that an individual currently in j will produce eventually. If j is a *type*, it is the LRS of a *type*-j individual in i offspring
$\boldsymbol{r}^T = \mathbf{1}^T\mathbf{R}$	$1 \times s$	Total eventual number of offspring: $r(j)$ is the eventual total number of offspring that an individual currently in state j will produce. If j is a *type*, it is the total LRS of a *type*-j individual
$\mathbf{N}$	$s \times s$	$N(i,j)$ is the eventual number of time-steps spent, before death, in state i for an individual currently in state j
Symbol		**Definition**
^{T}		Matrix transposition
$\boldsymbol{e}_i$		Vector of 0s (of the appropriate length), except for a 1 in position i
$\mathbf{1}$		Vector of 1s (of the appropriate length)
$\mathbf{0}$		Vector of 0s (of the appropriate length)
$\otimes$		Kronecker (or direct) product
$\bullet$		Termwise (or Hadamard) product
$\Diamond$		Demographic Stochastic Projection operator
diag ()		Operator turning a vector into a diagonal matrix, with the vector as its diagonal
vec ()		Operator turning a matrix into a vertical vector by stacking the columns of the former on top of one another
$\overset{d}{=}$		Equality in distribution

the various $\mathbf{F}_j$ considered in their vectorised form. This implies that the (i,l) element of $\boldsymbol{var}(\mathbf{F}_j)$—which corresponds to $Cov(F(i,j), F(l,j))$, the covariance of the number of i and l newborn produced by a j individual—is to be found in position $i + s(l-1)$ of $\boldsymbol{V}(\mathbf{F}_j) = \boldsymbol{vec}(\boldsymbol{var}(\mathbf{F}_j))$ (e.g. see Equation e). Note that this entry can also be reached directly via the Kronecker product:

$$\boldsymbol{var}(\mathbf{F}_j)(i,l) = \boldsymbol{e}_i^T \boldsymbol{var}(\mathbf{F}_j)\boldsymbol{e}_l = (\boldsymbol{e}_l \otimes \boldsymbol{e}_i)^T \boldsymbol{vec}(\boldsymbol{var}(\mathbf{F}_j)) = (\boldsymbol{e}_l \otimes \boldsymbol{e}_i)^T \boldsymbol{V}(\mathbf{F}_j), \qquad (2e)$$

where $\boldsymbol{e}_i$ is the (deterministic) vector with 0s everywhere except for a 1 in position i. Equation (2d) is an important result that can be found specifically as equation (26) in (Pollard, 1966) and under a non-vectorised form as equation (4.2), on page 37, of (Harris, 1963).

2.2.2 | Probability-generating function

Applying Equation (3e) to the projection equation (Equation 2b) yields the *pgf* of $\boldsymbol{n}_{t+1}$ as a function of that of $\boldsymbol{n}_t$:

$$\varphi_{\boldsymbol{n}_{t+1}}(\mathbf{x}) = \varphi_{\mathbf{F}\Diamond\boldsymbol{n}_t}(\mathbf{x}) = \varphi_{\boldsymbol{n}_t}\left(\boldsymbol{\varphi}_{\mathbf{F}}(\mathbf{x})\right), \tag{2f}$$

from which one can compute the joint probability distributions of all elements of $\boldsymbol{n}_t$ for all t (Box 2, Equation 6d).

2.3 | Structured population model with overlapping generations

We now wish to extend the Multi-Type BP (where the population is structured but generations do not overlap) to a process where generations overlap. This can be done several ways, for example, by considering an alternative DSP to **F** that turns the Multi-Type BP into a Crump–Mode–Jagers BP, by allowing individuals to reproduce several times during a generation (Crump & Mode, 1968, 1969; Jagers, 1975). Here, we choose a different approach and extend to demographic stochastic models the framework used by population ecologists to study expected abundances. We consider a time-step in chronological time (e.g. 1 year) by contrast to the generation time used in the Multi-Type BP. DSP **F** now corresponds therefore to reproduction during that time-step, and individuals can reproduce several times during their lifetime following the same DSP (which can embed several stages).

2.3.1 | Survival DSP

This implies also that individuals can survive that time-step, which we model via an additional DSP: **S**. Stochastic vector $\mathbf{S}_j$ (the j^{th} column of the matrix of r.v.s **S**) corresponds to a multivariate Bernoulli random variable: $\sum_i S_{i,j}$ is either equal to 0 (the individual in state j dies) or 1 (the individual survives and transitions to the unique state l such that $S_{l,j} = 1$). The survival process is therefore completely characterised by its expectation $\bar{\mathbf{S}}$. Indeed, $Pr\left(\mathbf{S}_j = \mathbf{e}_i\right) = \bar{S}(i,j)$ and $Pr\left(\mathbf{S}_j = \mathbf{0}\right) = 1 - \sum_i \bar{S}(i.j)$. As a consequence (proof in Supporting Information: Section: I.3),

$$\left\{\begin{array}{l} \boldsymbol{V}(\mathbf{S}) = \mathbf{J}\bar{\mathbf{S}} - \left(\bar{\mathbf{S}} \otimes \bar{\mathbf{S}}\right)\mathbf{J} \\ \boldsymbol{\varphi}_{\mathbf{S}}(\boldsymbol{x}) = \mathbf{1}^T + \left(\boldsymbol{x}^T - \mathbf{1}^T\right)\bar{\mathbf{S}} \end{array}\right. , \tag{4}$$

with $\mathbf{J} = \sum_i \left(\boldsymbol{e}_i \otimes \boldsymbol{e}_i\right)\boldsymbol{e}_i^T$, a dummy matrix of dimension $s^2 \times s$, and where T denotes transposition. This equation, together with other intermediate results, useful formulas used in this method and the main results obtained in this manuscript, is reproduced in Table 2.

2.3.2 | The state-dependent multi-type BP

Using directly the introduced $\Diamond$ notation (Box 1, Equation 3a), the demographic stochastic population projection equation for such a model is

$$\boldsymbol{n}_{t+1} = (\mathbf{F} + \mathbf{S}) \Diamond \boldsymbol{n}_t. \tag{5a}$$

For simplicity, we consider, in the main text of this manuscript, that DSPs **F** and **S** are independent (and relax this assumption in Supporting Information: Section II). In other words, we assume that, for any j, there is no covariance between the reproduction and the survival of a given individual in state j: $\boldsymbol{cov}\left(\mathbf{F}_j, \mathbf{S}_j\right) = \mathbf{0}$. In such a model, and contrary to the non-overlapping generation Multi-Type BP (Equation 2b), the number of states an individual can be born into, that is, the number of *types* d, can be inferior to the number of states, s. There can exist states that are not *types*: i is such a state if $F(i,j) = 0$ for all j. We refer to the stochastic process described by Equation (5a) as a State-Dependent Multi-Type BP as it can be, in general, both Multi-Type (when $d > 1$) and State-Dependent (the state of an individual at the next time-step depends on its current state, implying that the state of an individual can vary over its life trajectory). It can also be purely State-Dependent, if $d = 1$: **F** has only one row that is non-zero. It can be purely Multi-Type, if $d = s$ and $\bar{\mathbf{S}}$ is a diagonal matrix. We consider here the general case where there are several *types* and states that are not *types*, as per the following simple illustration, that we will use for the remainder of this method.

2.3.3 | Illustration

We illustrate the BP approach with a population structured by trait body size, with $s = 2$ states: state 1 corresponds to *small* individuals and state 2 to *large* individuals. Each *small* individual has a probability 0.5 to survive; specifically, it survives the time-step and remains *small* with probability 0.3, and it survives and transitions towards the *large* state with probability 0.2. Each *large* individual has a probability 0.7 to survive (and remain *large*). This information is summarised via the expectation of survival:

$$\bar{\mathbf{S}} = \begin{bmatrix} 0.3 & 0 \\ 0.2 & 0.7 \end{bmatrix} \tag{a}$$

or, equivalently, its *pgf* (Equation 4):

$$\boldsymbol{\varphi}_{\mathbf{S}}\left(x_1, x_2\right) = \left[\varphi_{\mathbf{S}_1}\left(x_1, x_2\right),\ \varphi_{\mathbf{S}_2}\left(x_1, x_2\right)\right] = \left[0.3x_1 + 0.2x_2 + 0.5,\ 0.7x_2 + 0.3\right] \tag{b}$$

One can also directly compute the variance of the survival process (Equation 4):

TABLE 2 Table of mathematical formulas.

Formula	Interpretation	Reference
Main result		
$\mathbf{R} \overset{d}{=} \mathbf{F} + (\mathbf{R} \Diamond \mathbf{S})$	Stochastic LRS as function of survival and reproduction	Equation 7
Results pertaining to independent survival and reproduction		
$\mathbf{V}(\mathbf{R}) = \left(\mathbf{V}(\mathbf{F}) + (\bar{\mathbf{R}} \otimes \bar{\mathbf{R}})\mathbf{V}(\mathbf{S})\right)(\mathbf{I} - \bar{\mathbf{S}})^{-1}$	var.-cov. in LRS as function of survival and reproduction	Equation 10
$\mathbf{V}(\mathbf{R}) = \mathbf{V}(\mathbf{F})\bar{\mathbf{N}} + (\bar{\mathbf{F}} \otimes \bar{\mathbf{F}})\boldsymbol{V}(\mathbf{N})$	var.-cov. in LRS as function of time-spent in states and reproduction	Equation 15
$\varphi_{\mathbf{R}}(\boldsymbol{x}) = \varphi_{\mathbf{F}}(\boldsymbol{x}) \bullet \left(\varphi_{\mathbf{S}}(\varphi_{\mathbf{R}}(\boldsymbol{x}))\right)$	*pgf* of LRS as function of survival and reproduction	Equation 16
$\varphi_{\mathbf{R}}(\boldsymbol{x}) = \left(\varphi_{\mathbf{F}}(\boldsymbol{x}) + \mathbf{1}^T\mathbf{D}_x - \mathbf{1}^T\right)\mathbf{D}_x^{-1}$	Direct expression of the *pgf* of LRS; $\mathbf{D}_x = \mathbf{I} - \bar{\mathbf{S}}\boldsymbol{diag}(\varphi_{\mathbf{F}}(\boldsymbol{x}))$	Equation 17
Extension to general case (covarying survival and reproduction, Supporting Information: Section II)		
$\boldsymbol{V}(\mathbf{R}) = [\boldsymbol{V}(\mathbf{F}) + (\bar{\mathbf{R}} \otimes \bar{\mathbf{R}})\boldsymbol{V}(\mathbf{S}) - [(\mathbf{I} \otimes \bar{\mathbf{R}}) +$ $(\bar{\mathbf{R}} \otimes \mathbf{I})\mathbf{K}^{(s,s)}]\mathbf{C}](\mathbf{I} - \bar{\mathbf{S}})^{-1}$	var.-cov. in LRS as function of survival, reproduction and their covariance ($\mathbf{C}_j = \boldsymbol{vec}(\boldsymbol{cov}(\mathbf{S}_j, \mathbf{F}_j))$); extends Equation (15) to general case; $\mathbf{K}^{(s,s)} = \sum_{i,j}(\boldsymbol{e}_i \otimes \boldsymbol{e}_j)(\boldsymbol{e}_j \otimes \boldsymbol{e}_i)^T$ is a commutation matrix	equation 43, Supporting Information: Section II
$\varphi_{\mathbf{R}_j}(\boldsymbol{x}) = \varphi_{\mathbf{F}_j,\mathbf{S}_j}(\boldsymbol{x}, \varphi_{\mathbf{R}}(\boldsymbol{x}))$	Extends Equation (16) to the general case; $\varphi_{\mathbf{F}_j,\mathbf{S}_j}\left(\boldsymbol{a}^T, \boldsymbol{b}^T\right) = \sum_{k,l} Pr((\mathbf{F}_j, \mathbf{S}_j) = (\boldsymbol{k}, \boldsymbol{l}))a_1^{k_1}a_2^{k_2} \ldots a_s^{k_s}b_1^{l_1} \ldots b_s^{l_s}$	equation 44, Supporting Information: Section II
Intermediate results and useful formulas		
If $\bar{\mathbf{S}}$ converges, $\sum_{k=0}^{\infty} \bar{\mathbf{S}}^k = (\mathbf{I} - \bar{\mathbf{S}})^{-1} = \bar{\mathbf{N}}$	Convergence of survival matrix	equation 38, Supporting Information: Section I
$\bar{\mathbf{R}} = \bar{\mathbf{F}}(\mathbf{I} - \bar{\mathbf{S}})^{-1}$	Expectation of LRS	Equation 9
$\bar{\mathbf{N}} = (\mathbf{I} - \bar{\mathbf{S}})^{-1}$	Expectation of time-spent in the various states	equation 49, Supporting Information: Section IV
$\boldsymbol{V}(\mathbf{N}) = (\bar{\mathbf{N}} \otimes \bar{\mathbf{N}})\boldsymbol{V}(\mathbf{S})(\mathbf{I} - \bar{\mathbf{S}})^{-1} = (\bar{\mathbf{N}} \otimes \bar{\mathbf{N}})\boldsymbol{V}(\mathbf{S})\bar{\mathbf{N}}$	Variance–covariance of time-spent in the various states	equation 50, Supporting Information: Section IV
$\boldsymbol{V}(\mathbf{S}_j) = \boldsymbol{vec}(\boldsymbol{diag}(\bar{\mathbf{S}}_j)) - \bar{\mathbf{S}}_j \otimes \bar{\mathbf{S}}_j$	Variance of Bernoulli stochastic vector (such as $\mathbf{S}_j$)	equation 34, Supporting Information: Section I
$\mathbf{J} = \sum_i(\boldsymbol{e}_i \otimes \boldsymbol{e}_i)\boldsymbol{e}_i^T$	Dummy matrix	equation 25, Supporting Information: Section I
$\boldsymbol{V}(\mathbf{S}) = \mathbf{J}\bar{\mathbf{S}} - (\bar{\mathbf{S}} \otimes \bar{\mathbf{S}})\mathbf{J}$	Variance–covariance of survival as a function of its expectation	Equation 4
$\varphi_{\mathbf{S}}(\boldsymbol{x}) = \mathbf{1}^T + \left(\boldsymbol{x}^T - \mathbf{1}^T\right)\bar{\mathbf{S}}$	*pgf* of survival as a function of its expectation	Equation 4
$\mathbf{A} \Diamond \boldsymbol{e}_j \overset{d}{=} \mathbf{A}_j$	$\mathbf{A}_j$ as projection of $\boldsymbol{e}_j$ via DSP $\mathbf{A}$	equation 27, Supporting Information: Section I
$\mathbf{A} \Diamond \boldsymbol{n} \overset{d}{=} \sum_j \mathbf{A}_j \Diamond n(j)$	decomp. of DSP projection of stochastic vector $\boldsymbol{n}$ by $\mathbf{A}$	equation 33, Supporting Information: Section I
$\boldsymbol{c}^T \bullet \boldsymbol{d}^T = \boldsymbol{c}^T\boldsymbol{diag}(\boldsymbol{d}) = \left(\boldsymbol{c}^T \otimes \boldsymbol{d}^T\right)\mathbf{J}$	Relationship between Kronecker and termwise products and the diag operator	equation 26, Supporting Information: Section I
$(\mathbf{AB}) \otimes (\mathbf{CD}) = (\mathbf{A} \otimes \mathbf{C})(\mathbf{B} \otimes \mathbf{D})$	Relationship between classic and Kronecker products	equation 32, Supporting Information: Section I
$\boldsymbol{vec}(\mathbf{ABC}) = \left(\mathbf{C}^T \otimes \mathbf{A}\right)\boldsymbol{vec}(\mathbf{B})$	Relationship between vec operator and Kronecker product	equation 36, Supporting Information: Section I

$$\boldsymbol{V}(\mathbf{S}) = [\boldsymbol{V}(\mathbf{S}_1), \boldsymbol{V}(\mathbf{S}_2)] = \begin{bmatrix} 0.21 & 0 \\ -0.06 & 0 \\ -0.06 & 0 \\ 0.16 & 0.21 \end{bmatrix} \tag{c}$$

Each *small* individual produces at most one newborn per time-step (Bernoulli reproduction): it gives birth to one *small* offspring with probability 0.5. *Large* individuals produce either one *small* or one *large* offspring, at every time-step, with equal probabilities. This corresponds to the following *pgf* (Box 2, Equation 6c):

$$\varphi_{\mathbf{F}}(x_1, x_2) = [\varphi_{\mathbf{F}_1}(x_1, x_2),\ \varphi_{\mathbf{F}_2}(x_1, x_2)] = [0.5x_1 + 0.5,\ 0.5x_1 + 0.5x_2] \tag{d}$$

One can also characterise the reproduction process via its expectation and variance. As reproduction is Bernoulli, in this particular case, we can directly use the same formula as for the survival process (Equation 4) for the reproductive variance, which gives

$$\overline{\mathbf{F}} = \begin{bmatrix} 0.5 & 0.5 \\ 0 & 0.5 \end{bmatrix} \quad \text{and} \quad \boldsymbol{V}(\mathbf{F}) = \begin{bmatrix} 0.25 & 0.25 \\ 0 & -0.25 \\ 0 & -0.25 \\ 0 & 0.25 \end{bmatrix} \tag{e}$$

2.3.4 | Expectation

One can project, from Equation (3b), the vector of expected abundances $\overline{\boldsymbol{n}}_t$ over time:

$$\overline{\boldsymbol{n}}_{t+1} = (\overline{\mathbf{F}} + \overline{\mathbf{S}})\overline{\boldsymbol{n}}_t, \tag{5b}$$

which is the deterministic population projection equation at the core of the matrix population projection framework (Caswell, 2001). The deterministic projector is often considered as a single matrix, $\overline{\mathbf{M}} = \overline{\mathbf{F}} + \overline{\mathbf{S}}$, which dominant eigenvalue λ determines whether the BP is overcritical ($\lambda > 1$) or subcritical ($\lambda < 1$).

2.3.5 | Variance

From Equation (3c), and as reproduction and survival are independent processes (Equation 3i), we have:

$$\boldsymbol{V}(\boldsymbol{n}_{t+1}) = (\overline{\mathbf{M}} \otimes \overline{\mathbf{M}})\boldsymbol{V}(\boldsymbol{n}_t) + (\boldsymbol{V}(\mathbf{F}) + \boldsymbol{V}(\mathbf{S}))\overline{\boldsymbol{n}}_t \tag{5c}$$

Illustration: Let us assume a deterministic initial population vector $\boldsymbol{n}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ (therefore $\boldsymbol{V}(\boldsymbol{n}_0) = \mathbf{0}$). From Equation (5c), we have

$$\boldsymbol{V}(\boldsymbol{n}_1) = (\boldsymbol{V}(\mathbf{F}) + \boldsymbol{V}(\mathbf{S}))\boldsymbol{n}_0 = \begin{bmatrix} 0.46 & 0.25 \\ -0.06 & -0.25 \\ -0.06 & -0.25 \\ 0.16 & 0.46 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.71 \\ -0.31 \\ -0.31 \\ 0.62 \end{bmatrix}, \text{ that is,}$$

$$\boldsymbol{var}(\boldsymbol{n}_1) = \begin{bmatrix} 0.71 & -0.31 \\ -0.31 & 0.62 \end{bmatrix}$$

At $t = 1$, the number of individuals in state *small* covaries negatively with the number of individuals in state *large*: this is caused by the reproduction of large individuals (that produce *either* one large or one small individual) and the survival of *small* individuals (which *either* remain small or become large).

2.3.6 | Probability-generating function

Applying Equation (3e) to the projection equation (Equation 5a) yields

$$\varphi_{\boldsymbol{n}_{t+1}}(\mathbf{x}) = \varphi_{\boldsymbol{n}_t}\left(\boldsymbol{\varphi}_{\mathbf{F}+\mathbf{S}}(\mathbf{x})\right).$$

As **S** and **F** are independent, we have (see Box 1, Equation 3i)

$$\boldsymbol{\varphi}_{\mathbf{F}+\mathbf{S}}(\mathbf{x}) = \boldsymbol{\varphi}_{\mathbf{F}}(\mathbf{x}) \bullet \boldsymbol{\varphi}_{\mathbf{S}}(\mathbf{x}),$$

where $\bullet$ denotes the termwise (or Hadamard) product of two vectors and therefore

$$\varphi_{\boldsymbol{n}_{t+1}}(\mathbf{x}) = \varphi_{\boldsymbol{n}_t}\left(\boldsymbol{\varphi}_{\mathbf{F}}(\mathbf{x}) \bullet \boldsymbol{\varphi}_{\mathbf{S}}(\mathbf{x})\right) \tag{5d}$$

Illustration: As $\boldsymbol{n}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, we have (see Box 2, Equation 6c)

$$\varphi_{\boldsymbol{n}_0}(x_1, x_2) = x_1 x_2 \tag{f}$$

Further (by definition, as $\bullet$ denotes the termwise product):

$$\boldsymbol{\varphi}_{\mathbf{F}}(x_1,x_2) \bullet \boldsymbol{\varphi}_{\mathbf{S}}(x_1,x_2) = \left[\boldsymbol{\varphi}_{\mathbf{F}_1}(x_1,x_2)\boldsymbol{\varphi}_{\mathbf{S}_1}(x_1,x_2),\ \boldsymbol{\varphi}_{\mathbf{F}_2}(x_1,x_2)\boldsymbol{\varphi}_{\mathbf{S}_2}(x_1,x_2)\right] \tag{g}$$

From Equation (5d) applied to Equations (b), (d), (f) and (g), this implies that

$$\begin{aligned} \varphi_{\boldsymbol{n}_1}(x_1,x_2) &= \varphi_{\boldsymbol{n}_0}\left(\boldsymbol{\varphi}_{\mathbf{F}_1}(x_1,x_2)\boldsymbol{\varphi}_{\mathbf{S}_1}(x_1,x_2),\ \boldsymbol{\varphi}_{\mathbf{F}_2}(x_1,x_2)\boldsymbol{\varphi}_{\mathbf{S}_2}(x_1,x_2)\right) \\ &= \boldsymbol{\varphi}_{\mathbf{F}_1}(x_1,x_2)\boldsymbol{\varphi}_{\mathbf{S}_1}(x_1,x_2)\boldsymbol{\varphi}_{\mathbf{F}_2}(x_1,x_2)\boldsymbol{\varphi}_{\mathbf{S}_2}(x_1,x_2) \\ &= \frac{1}{400}\left(21x_1^3x_2 + 9x_1^3 + 35x_1^2x_2^2 + 71x_1^2x_2 + 24x_1^2\right. \\ &\quad \left. + 14x_1x_2^3 + 76x_1x_2^2 + 65x_1x_2 + 15x_1 + 14x_2^3 + 41x_2^2 + 15x_2\right) \end{aligned} \tag{h}$$

As the reproduction, in this illustration, is finite (i.e. there is a maximum number of offspring, α, that can be produced by an individual per time-step, here $\alpha = 1$), $\boldsymbol{\varphi}_{\mathbf{F}}(\mathbf{x})$ is a vector of polynomials of finite degree, and therefore, $\varphi_{\boldsymbol{n}_1}$ is produced directly as a polynomial (of finite degree) in x_1 and x_2, which allows to 'read' directly, in the polynomials' coefficients, the probability distribution of the number of individuals at $t = 1$ (Box 2, Equation 6d). For example, $Pr\left(\boldsymbol{n}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}\right) = \frac{21}{400}$, $Pr\left(\boldsymbol{n}_1 = \begin{bmatrix} 3 \\ 0 \end{bmatrix}\right) = \frac{9}{400}$ and $Pr\left(\boldsymbol{n}_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}\right) = \frac{35}{400}$.

3 | LIFETIME REPRODUCTIVE SUCCESS

We remain in the Multi-Type State-Dependent BP framework (Equation 5a), and wish to infer, from the knowledge of the DSPs **F** and **S** driving the stochastic population dynamics, the moments and, in general, the distribution of the lifetime reproductive success (LRS) of the various *types* in the population.

3.1 | LRS and eventual number of offspring

We denote **R**, the $s \times s$ matrix of integer-valued random variables, where $R(i,j)$ is the number of offspring in state i, an individual currently in state j will eventually produce, that is, will produce at the current time-step, and provided it survives, at the next time-step, etc., until it eventually dies. **R** is evidently like **F** or **S**, a DSP: one can for instance consider the stochastic vector $\mathbf{R} \Diamond \boldsymbol{n}$ corresponding to the eventual number of offspring (structured by *type*) produced by the individuals making up abundance vector $\boldsymbol{n}$.

Consider, without any loss of generality, that the d *types* make up the first d positions in the state structure (i.e. rows $d+1$ to s of $\overline{\mathbf{F}}$ are made up only of 0s, contrary to the rows above). While $R(i,j)$ is defined for all states i and j ($1 \leq i,j \leq s$), it is only non-zero for states i that are *types* (that is $1 \leq i \leq d$): one can only produce offspring in newborn states. Further, if j is a *type* (that is $1 \leq j \leq d$), then $\mathbf{R}_j$, the random vector of the eventual number of offspring produced by an individual currently in state j (and structured by the *type* of the offspring produced), is also its LRS. In that case, $R(i,j)$, the eventual number of offspring in state i produced by an individual in newborn state j, *is* the LRS of a *type*-j individual (in i-newborns). If the model is purely State-Dependent (i.e. there is only one *type*: $d=1$), then the LRS random variable is a scalar corresponding to $R(1,1)$.

3.2 | A simple stochastic formula

The eventual number of offspring (structured by their *types*) produced by an individual in state j at time t has the distribution of the stochastic vector $\mathbf{R}_j$. It is also the number of offspring produced by this individual at the current time-step, which has the distribution of $\mathbf{F}_j$, plus the eventual number of offspring it will produce from the next time-step onwards, providing it survives the current one, which has the distribution of $\mathbf{R} \Diamond \mathbf{S}_j$ (the environment is constant over time, so is the distribution of **R**), therefore:

$$\mathbf{R}_j \overset{d}{=} \mathbf{F}_j + \left(\mathbf{R} \Diamond \mathbf{S}_j\right)$$

As this is true for any state j, (see Box 1, Equation 3f), we have, simply:

$$\boxed{\mathbf{R} \overset{d}{=} \mathbf{F} + \left(\mathbf{R} \Diamond \mathbf{S}\right)}. \tag{7}$$

3.3 | A direct formula for the variance–covariance of the LRS

Taking the expectation of both sides of Equation (7), we have (Box 1, Equation 3j)

$$\overline{\mathbf{R}} = \overline{\mathbf{F}} + \overline{\mathbf{R}\mathbf{S}} \tag{8}$$

that is, $\overline{\mathbf{R}}\left(\mathbf{I} - \overline{\mathbf{S}}\right) = \overline{\mathbf{F}}$, where **I** is the identity matrix. Providing that $\mathbf{I} - \overline{\mathbf{S}}$ is invertible, that is, providing that $\overline{\mathbf{S}}$ converges—a condition equivalent, from an ecological standpoint, to the fact that every individual dies eventually—this yields the well-known formula for $\overline{\mathbf{R}}$, the 'next-generation' matrix (Caswell, 2001):

$$\overline{\mathbf{R}} = \overline{\mathbf{F}}\left(\mathbf{I} - \overline{\mathbf{S}}\right)^{-1} \tag{9}$$

Illustration: For the population of the illustration, we have

$$\overline{\mathbf{R}} = \begin{bmatrix} 0.5 & 0.5 \\ 0 & 0.5 \end{bmatrix} \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.3 & 0 \\ 0.2 & 0.7 \end{bmatrix} \right)^{-1} = \begin{bmatrix} 1.19 & 1.67 \\ 0.48 & 1.67 \end{bmatrix}, \tag{i}$$

which implies, in particular, that the expected LRS of an individual of *type small* consists in 1.19 offspring of *type small* and 0.48 offspring of *type large*.

Similarly, one can take the variance of both sides of Equation (7) and via Equations (3j) and (3i) of Box 1, accounting in particular for the independence of **S** and **F**, we obtain

$$\mathbf{V}(\mathbf{R}) = \mathbf{V}(\mathbf{F}) + \left(\overline{\mathbf{R}} \otimes \overline{\mathbf{R}}\right)\mathbf{V}(\mathbf{S}) + \mathbf{V}(\mathbf{R})\overline{\mathbf{S}}, \qquad \text{which leads to}$$

$$\boxed{\mathbf{V}(\mathbf{R}) = \left(\mathbf{V}(\mathbf{F}) + \left(\overline{\mathbf{R}} \otimes \overline{\mathbf{R}}\right)\mathbf{V}(\mathbf{S})\right)\left(\mathbf{I} - \overline{\mathbf{S}}\right)^{-1}} \tag{10}$$

3.3.1 | Total LRS

This formula extends the results of van Daalen and Caswell (2017) where the authors focus on the total LRS, $\boldsymbol{r}^T$, a (horizontal) vector of univariate random variables providing the eventual total number of *all* offspring produced (by individuals in the various states). We have, simply $r(j) = \sum_i R(i,j)$, that is $\boldsymbol{r}^T = \mathbf{1}^T\mathbf{R}$. Who can do more can do less: with the same reasoning than for the (structured) LRS, one can obtain a stochastic equation for the total LRS. Denoting $\boldsymbol{f}^T = \mathbf{1}^T\mathbf{F}$, the vector of total time-step reproduction, this is:

$$\boldsymbol{r}^T \overset{d}{=} \boldsymbol{f}^T + \boldsymbol{r}^T \Diamond \mathbf{S}\ , \tag{11}$$

which leads to (Box 1, Equation 3j):

$$\boldsymbol{V}\left(\boldsymbol{r}^T\right) = \left(\mathbf{V}\left(\boldsymbol{f}^T\right) + \left(\overline{\boldsymbol{r}}^T \otimes \overline{\boldsymbol{r}}^T\right)\mathbf{V}(\mathbf{S})\right)\left(\mathbf{I} - \overline{\mathbf{S}}\right)^{-1} = \mathbf{1}^T\boldsymbol{V}(\mathbf{R}) \tag{12}$$

Equation (12) can be shown to correspond to equation (22) in (van Daalen & Caswell, 2017) (see Supporting Information: Section III). Equation (12) allows to compute directly the variance in (total) LRS, which only requires knowledge of the variance of total reproduction ($\boldsymbol{f}^T$, in contrast to the structured reproduction **F** required for $\boldsymbol{V}(\mathbf{R})$, Equation 10). It shows also that $\boldsymbol{V}\left(\boldsymbol{r}^T\right)$ can be obtained directly by summing the columns of $\boldsymbol{V}(\mathbf{R})$ which corresponds to the classic result:

$$\begin{aligned} var(r(j)) = var\left(\sum_i R(i,j)\right) &= \sum_{i,l} cov(R(i,j), R(l,j)) \\ &= \sum_i var(R(i,j)) + 2\sum_{i<l} cov(R(i,j), R(l,j)) \end{aligned} \tag{13}$$

3.3.2 | Illustration

With regards to the illustration, Equation (10) yields:

$$\mathbf{V}(\mathbf{R}) = \left(\begin{bmatrix} 0.25 & 0.25 \\ 0 & -0.25 \\ 0 & -0.25 \\ 0 & 0.25 \end{bmatrix} + \left(\begin{bmatrix} 1.19 & 1.67 \\ 0.48 & 1.67 \end{bmatrix} \otimes \begin{bmatrix} 1.19 & 1.67 \\ 0.48 & 1.67 \end{bmatrix} \right) \begin{bmatrix} 0.21 & 0 \\ -0.06 & 0 \\ -0.06 & 0 \\ 0.16 & 0.21 \end{bmatrix} \right)$$

$$\left(\mathbf{I} - \bar{\mathbf{S}}\right)^{-1} = \begin{bmatrix} 1.87 & 2.78 \\ 0.88 & 1.11 \\ 0.88 & 1.11 \\ 1.36 & 2.78 \end{bmatrix}, \tag{j}$$

which implies for instance that the variance–covariance matrix of the LRS of a *small type* is

$$\boldsymbol{var}\left(\mathbf{R}_1\right) = \begin{bmatrix} 1.87 & 0.88 \\ 0.88 & 1.36 \end{bmatrix}$$

For the LRS, the numbers of offspring produced of each *type* covary positively, despite the opposite being true for the time-reproduction (covariance was negative for large individuals and 0 for small ones). This is because the effect of time-step reproduction on LRS, in this case, is outweighed by that of survival: the production of each *type* covaries positively with the age-at-death, and therefore, they covary positively with one another. Following Equation (12), we can obtain the variance of the total LRS of a *small type* by summing all the elements in $\boldsymbol{var}\left(\mathbf{R}_1\right)$, which gives $var(r(1)) = 5$.

3.3.3 | Decomposition according to stochasticity in reproduction and survival (or age-at-death)

Equation (10) provides a way to compute the variance–covariance of LRS, but also gives its decomposition according to demographic variance in reproduction ($\mathbf{V}_\mathbf{F}(\mathbf{R})$) and demographic variance in survival ($\mathbf{V}_\mathbf{S}(\mathbf{R})$):

$$\begin{cases} \mathbf{V}(\mathbf{R}) = \mathbf{V}_\mathbf{F}(\mathbf{R}) + \mathbf{V}_\mathbf{S}(\mathbf{R}) \quad \text{with} \\ \mathbf{V}_\mathbf{F}(\mathbf{R}) = \mathbf{V}(\mathbf{F})\left(\mathbf{I} - \bar{\mathbf{S}}\right)^{-1} \quad \text{and} \quad \mathbf{V}_\mathbf{S}(\mathbf{R}) = \left(\bar{\mathbf{R}} \otimes \bar{\mathbf{R}}\right)\mathbf{V}(\mathbf{S})\left(\mathbf{I} - \bar{\mathbf{S}}\right)^{-1} \end{cases} \tag{14}$$

In Supporting Information: Section IV, we apply the method developed here for $\mathbf{R}$, to a different DSP, $\mathbf{N}$, where $N(i,j)$ is the number of time-steps spent, before death, in state i for an individual currently in state j. For a *type*, the 'total $\mathbf{N}$', $\boldsymbol{d}^T = \mathbf{1}^T\mathbf{N}$, is then the age-at-death. The expectation and variance we obtain for $\boldsymbol{d}^T$ correspond to the results in (Steiner & Tuljapurkar, 2012). The expectation and variance we obtain for $\mathbf{N}$ allow us to interpret Equation (10), alternatively, as the decomposition of the variance in LRS according to the demographic variance in reproduction and in the time spent in the various states:

$$\mathbf{V}(\mathbf{R}) = \mathbf{V}(\mathbf{F})\bar{\mathbf{N}} + \left(\bar{\mathbf{F}} \otimes \bar{\mathbf{F}}\right)\boldsymbol{V}(\mathbf{N}), \tag{15}$$

which extends the results in (Steiner & Tuljapurkar, 2012, see Supporting Information: Section III). We provide the code to compute the variance–covariance of LRS, for R and Matlab (see Supporting Information: Section VII).

3.4 | Joint probability distribution of LRS

From the stochastic LRS formula (Equation 7), we can move beyond summary statistics such as the mean and variance, and examine the full distribution of LRS directly. This is important because, as shown by Tuljapurkar et al. (2020), the first moments of LRS often fail to capture the shape of its distribution—for instance, many projection models predict a mode at zero that is not inferable from the expectation or variance. In this study, we analyse LRS structured by both offspring and parental *types*, allowing us to go further and investigate the *joint* distribution of LRS across offspring *types*. This joint distribution offers crucial insights for various questions in ecology and evolution, including inheritance patterns—that is, how parental *type* influences offspring *type*—and the presence of trade-offs, as illustrated below. Equation (7) is equivalent to (see Box 1, Equation 3j):

$$\varphi_\mathbf{R}(\mathbf{x}) = \varphi_\mathbf{F}(\mathbf{x}) \bullet \left(\varphi_\mathbf{S}\left(\varphi_\mathbf{R}(\mathbf{x})\right)\right) \tag{16}$$

which leads to (Box 1 and Equation 4):

$$\varphi_\mathbf{R}(\mathbf{x})\left(\mathbf{I} - \bar{\mathbf{S}}\boldsymbol{diag}\left(\varphi_\mathbf{F}(\mathbf{x})\right)\right) = \varphi_\mathbf{F}(\mathbf{x}) - \mathbf{1}^T\bar{\mathbf{S}}\boldsymbol{diag}\left(\varphi_\mathbf{F}(\mathbf{x})\right),$$

where $\boldsymbol{diag}(\varphi_\mathbf{F}(\mathbf{x}))$ is a diagonal matrix of functions, of size $s \times s$, worth 0 away from the diagonal and which (j,j) entry corresponds to the function $\varphi_{\mathbf{F}_j}(\mathbf{x})$. On the domain of interest of the *pgfs*, the matrix of functions $\mathbf{D}_\mathbf{x} = \mathbf{I} - \bar{\mathbf{S}}\boldsymbol{diag}(\varphi_\mathbf{F}(\mathbf{x}))$ is invertible provided $\bar{\mathbf{S}}$ converges (see Supporting Information: Section I). Denoting this inverse $\mathbf{D}_\mathbf{x}^{-1}$ (i.e. $\mathbf{D}_\mathbf{x}^{-1}$ is the $s \times s$ matrix of functions of $\mathbf{x}$, such that $\mathbf{D}_\mathbf{x}^{-1}\mathbf{D}_\mathbf{x} = \mathbf{I}$) the above equation is equivalent to:

$$\boxed{\varphi_\mathbf{R}(\mathbf{x}) = \left(\varphi_\mathbf{F}(\mathbf{x}) + \mathbf{1}^T\mathbf{D}_\mathbf{x} - \mathbf{1}^T\right)\mathbf{D}_\mathbf{x}^{-1}} \tag{17}$$

From the explicit formula for $\varphi_\mathbf{R}(\mathbf{x})$ (Equation 17), one can compute all the moments and co-moments of LRS, and most importantly, via successive partial derivations, obtain its (joint) distribution (see Box 2). This requires to be done (by hand or) via a computer program with symbolic/functional capabilities. We provide the code in Matlab (equipped with its *symbolic toolbox*) to perform this operation (see Supporting Information: Section VII). In the next section, we develop an algorithm that allows to approximate the distribution of the total LRS via numeric computations (which can be implemented in any programming environment). In Supporting Information: Section VI, we go further than the study of the decomposition of LRS according to parental type performed in the main text of this manuscript and investigate the

LRS of a typical individual, that is, of a newborn taken at random in a cohort.

3.4.1 | Illustration

From the *pgf* for reproduction (Equation d) and the expectation of survival (Equation a), we have

$$\mathbf{D}_x = \mathbf{I} - \bar{\mathbf{S}}\boldsymbol{diag}\left(\varphi_{\mathbf{F}}(\mathbf{x})\right) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.3 & 0 \\ 0.2 & 0.7 \end{bmatrix}\begin{bmatrix} 0.5x_1+0.5 & 0 \\ 0 & 0.5x_1+0.5x_2 \end{bmatrix}$$

$$= \frac{1}{20}\begin{bmatrix} 17-3x_1 & 0 \\ -2\left(x_1+1\right) & 20-7x_1-7x_2 \end{bmatrix}$$

Solving the system of equations corresponding to $\mathbf{D}_x^{-1}\mathbf{D}_x = \mathbf{I}$, we obtain

$$\mathbf{D}_x^{-1} = 20\begin{bmatrix} \dfrac{1}{17-3x_1} & 0 \\ 2\dfrac{\left(x_1+1\right)}{\left(17-3x_1\right)\left(20-7x_1-7x_2\right)} & \dfrac{1}{20-7x_1-7x_2} \end{bmatrix}$$

Equation (17) leads to

$$\varphi_{\mathbf{R}}\left(x_1,x_2\right) = \left(\left[0.5x_1+0.5, 0.5x_1+0.5x_2\right] + \frac{1}{20}\left[15-5x_1, 20-7x_1-7x_2\right] - [1,1]\right)\mathbf{D}_x^{-1},$$

that is,

$$\varphi_{\mathbf{R}}\left(x_1,x_2\right) = \left[\frac{\left(x_1+1\right)\left(100-29x_1-29x_2\right)}{\left(17-3x_1\right)\left(20-7x_1-7x_2\right)}, \frac{3x_1+3x_2}{20-7x_1-7x_2}\right] \tag{k}$$

From this, the probability, for a newborn to produce 0 offspring (during its entire lifetime, Box 1), is, for an individual of *type small*, $Pr\left(\mathbf{R}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \varphi_{\mathbf{R}_1}(0,0) = \frac{100}{17^{*}20} = \frac{5}{17} \approx 0.2941$, and for a large individual, $Pr\left(\mathbf{R}_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \varphi_{\mathbf{R}_2}(0,0) = 0$ (see Box 2). One can also compute (Box 2):

$$Pr\left(\mathbf{R}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$$

$$= \left.\frac{\partial \varphi_{\mathbf{R}_2}\left(x_1,x_2\right)}{\partial x_1}\right|_{(x_1,x_2)=(0,0)} = \left.\frac{3\left(20-7x_1-7x_2\right)+7\left(3x_1+3x_2\right)}{\left(20-7x_1-7x_2\right)^2}\right|_{(x_1,x_2)=(0,0)}$$

$$= \frac{3}{20}$$

And since, in this particular case, $\varphi_{\mathbf{R}_2}\left(x_1,x_2\right) = \varphi_{\mathbf{R}_2}\left(x_2,x_1\right)$, we have directly, $Pr\left(\mathbf{R}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \frac{3}{20}$. We can go one step further and compute

$$Pr\left(\mathbf{R}_2 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}\right)$$

$$= \frac{1}{2!}\left.\frac{\partial^2 \varphi_{\mathbf{R}_2}\left(x_1,x_2\right)}{\partial x_1^2}\right|_{(x_1,x_2)=(0,0)} = \frac{1}{2}\left.\frac{42\left[\left(20-7x_1-7x_2\right)+7\left(3x_1+3x_2\right)\right]}{\left(20-7x_1-7x_2\right)^3}\right|_{(x_1,x_2)=(0,0)}$$

$$= \frac{21}{400}$$

For the same reason than above,

$Pr\left(\mathbf{R}_2 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}\right) = Pr\left(\mathbf{R}_2 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}\right) = \frac{21}{400} = 0.0525$. We have, further,

$$Pr\left(\mathbf{R}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$$

$$= \frac{1}{1!}\frac{1}{1!}\left.\frac{\partial^2 \varphi_{\mathbf{R}_2}\left(x_1,x_2\right)}{\partial x_1 \partial x_2}\right|_{(x_1,x_2)=(0,0)} = \left.\frac{42\left[\left(20-7x_1-7x_2\right)+7\left(3x_1+3x_2\right)\right]}{\left(20-7x_1-7x_2\right)^3}\right|_{(x_1,x_2)=(0,0)}$$

$$= \frac{21}{200}$$

In general, since $\boldsymbol{\varphi}_{\mathbf{R}_2}\left(x_1,x_2\right) = \boldsymbol{\varphi}_{\mathbf{R}_2}\left(x_2,x_1\right)$, we have that $Pr\left(\mathbf{R}_2 = \begin{bmatrix} a \\ b \end{bmatrix}\right) = Pr\left(\mathbf{R}_2 = \begin{bmatrix} b \\ a \end{bmatrix}\right)$. This symmetry is absent from $\mathbf{R}_1$. We let Matlab compute further partial derivatives and evaluate them at $(0,0)$; we plot the obtained joint probability distributions in Figure 1a (for $\mathbf{R}_1$) and Figure 1b (for $\mathbf{R}_2$). We also consider the total LRS produced by these two *types* of individuals and we get from Equation (3h) in Box 1 and the *pgf* of **R** (Equation k):

$$\boldsymbol{\varphi}_{\mathbf{r}^T}(x) = \left[\varphi_{r(1)}(x), \varphi_{r(2)}(x)\right] = \boldsymbol{\varphi}_{\mathbf{R}}(x,x)$$
$$= \left[\frac{(x+1)(100-58x)}{(17-3x)(20-14x)}, \frac{6x}{20-14x}\right] \tag{l}$$

From this, we have that $Pr(r(2)=0) = \varphi_{r(2)}(0) = 0$; obviously $Pr(r(2)=0) = Pr\left(\mathbf{R}_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right)$. We also get $Pr(r(2)=1) = \left.\frac{d\varphi_{r(2)}(x)}{dx}\right|_{x=0} = \frac{6}{20} = 0.3$ which we could have gotten directly from $Pr(r(2)=1) = Pr\left(\mathbf{R}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) + Pr\left(\mathbf{R}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)$. Further computations lead to the probability distributions displayed in Figure 1c (for $r(1)$) and Figure 1d (for $r(2)$).

Figures 1c,d, which represent the total LRS for each parental *type*, could in principle be obtained using the method of Tuljapurkar et al. (2020), by evaluating the distribution of total LRS from a given newborn state. In contrast, Figures 1a,b, which show the joint distribution (over offspring *types*), reflects the additional capabilities of the method presented here. Figures 1c,d reveal that *large-type* parents produce more offspring, and with higher variance, than *small-type* parents—something we could already infer by summing the columns of $\bar{\mathbf{R}}$ (Equation i) and $\mathbf{V}(\mathbf{R})$ (Equation j). However, these moments alone do not capture the fact that all

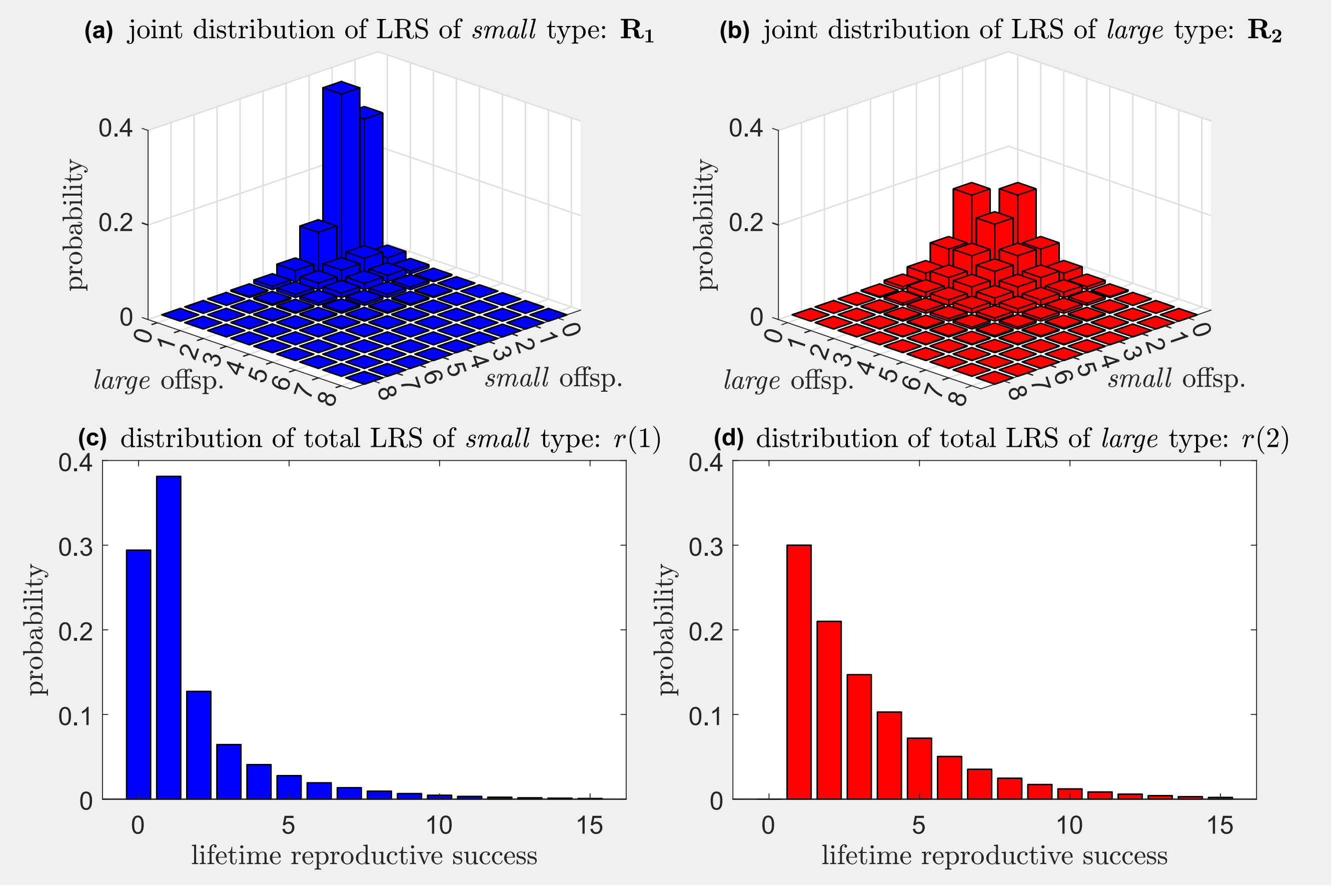

FIGURE 1 Probability distribution of lifetime reproductive success (LRS) for the illustration. Top row: Joint distribution of the number of offspring of *type small* and *large* produced by (a) a parent of *type small* and (b) a parent of *type large*. Bottom row: Distribution of the total LRS produced by (c) a parent of *type small* and (d) a parent of *type large*.

large-*type* individuals produce offspring ($Pr\left(\sum_i R(i,2)=0\right)=0$) while approximately 29.41% of small *types* do not produce any. The mode—that is, the most probable LRS—is 1 for both *types*, but the probability of producing more offspring declines much more steeply for small-*type* parents indicating greater skew. Figure 1b goes further by revealing that, for large-*type* parents, the distribution of LRS is symmetric with respect to offspring *types*: the probability of producing, during its lifetime, n_s *small* and n_l *large* offspring is the same as producing n_l *small* and n_s *large* offspring. This symmetry is absent for small-*type* parents: from $\bar{\mathbf{R}}$ (Equation i) we already know that a *small-type* parent is expected to produce more *small* newborn than *large* ones. Furthermore, the fact that Figures 1a,b differ from each other confirms that the number of offspring of each *type* produced depends on parental *type*: there is inheritance of the *type* in that population model. Finally, in this illustration, small *types* are clearly of lower quality than the large ones. Therefore, the asymmetry across offspring *types* in the LRS of *small-type* parents corresponds to a quantity–quality trade-off: a *small-type* individual will, on average, produce more lower quality (*small*) offspring than higher quality (*large*) ones. In this context, Figure 1a displays the properties of this trade-off.

4 | EXTENSION I: ALGORITHM FOR THE DISTRIBUTION OF TOTAL LRS

4.1 | Eventual number of offspring as a function of age

Stochastic Equations (7) and (11) allow, when turned into vectors of *pgfs* (e.g. see Equation 17), to compute the joint probability distribution of LRS and the distribution of total LRS. However, this computation (a) requires using a symbolic/functional programming software and (b) can be quite long, all the more so for the computation of the probability to produce a large number of offspring for a model with many states. Thanks to the Branching Process approach, we can develop a numeric algorithm (fast and implementable in any computer program) that provides an approximation of this distribution. We describe it here in the case of the total LRS.

While Equation (11) consists of a stochastic formula for the eventual total number of offspring structured by state, $\mathbf{r}^T$, it is also possible to construct such a formula for the total eventual number of offspring structured by state *and* age. Let us denote $\mathbf{r_a}^T$ the horizontal vector of (integer-valued, univariate) random variables, where

$r_a(j)$ is the total number of offspring an individual, currently aged a and in state j, will produce eventually before dying. The eventual number of offspring of an individual aged a is the number produced at that age plus, provided it survives, its eventual number of offspring at age $a+1$. Therefore,

$$\mathbf{r}_a^T \overset{d}{=} \mathbf{f}^T + \mathbf{r}_{a+1}^T \Diamond \mathbf{S} \tag{18a}$$

The LRS of an individual of *type* j *is* its total eventual number of offspring when in that state and in the first age class. Therefore, for each *type* j:

$$r(j) \overset{d}{=} r_1(j). \tag{18b}$$

4.2 | Age-structured populations with finite reproduction

In the general case, Equation (18a) does not allow to compute $\mathbf{r}_1^T$, from which to obtain the total LRS (Equation 18b). However, when the population is (explicitly) age-structured (age is *a* trait of the model, not necessarily the only one), it implements a *maximum age* ω, beyond which no individual survives. For individuals aged ω, the eventual number of offspring corresponds, simply, to the offspring produced during that final time-step:

$$\mathbf{r}_\omega^T \overset{d}{=} \mathbf{f}^T. \tag{18c}$$

With the properties of this 'initial value' (Equation 18c), one can, by recursion (Equation 18a) obtain those of the total LRS (Equation 18b). Turning the above system of stochastic equations (Equation 18) into its equivalent system of *pgf*s, leads to an iterative computation of the various *pgf*s, starting at $a=\omega$ and ending at $a=1$ (Equation 4 and Box 1):

$$\left(\begin{array}{l} \boldsymbol{\varphi}_{\mathbf{r}_\omega^T}(x) = \boldsymbol{\varphi}_{\mathbf{f}^T}(x) \\[1ex] \boldsymbol{\varphi}_{\mathbf{r}_a^T}(x) = \boldsymbol{\varphi}_{\mathbf{f}^T}(x) \bullet \boldsymbol{\varphi}_{\mathbf{q}_{a+1}^T}(x), \text{ for all } 1 \le a < \omega, \text{ with} \\[1ex] \qquad \boldsymbol{\varphi}_{\mathbf{q}_a{}^T}(x) = \boldsymbol{\varphi}_{\mathbf{S}}\left(\boldsymbol{\varphi}_{\mathbf{r}_a^T}(x)\right) = \mathbf{1}^T - \mathbf{1}^T\overline{\mathbf{S}} + \boldsymbol{\varphi}_{\mathbf{r}_a^T}(x)\overline{\mathbf{S}} \\[1ex] \qquad\qquad \varphi_{r(j)}(x) = \varphi_{r_1(j)}(x) \end{array}\right. \tag{19}$$

If, moreover, reproduction is finite—that is there exists α, so that no individual in that model can produce more than α offspring in a given time-step—each element of $\boldsymbol{\varphi}_{\mathbf{f}^T}(x)$ can be written as a polynomial in x of degree α or less. It implies also, by induction, that the s *pgfs* in $\boldsymbol{\varphi}_{\mathbf{r}_a^T}$ can be written as polynomials of (maximum) degree $\alpha(\omega - a + 1)$; this is equivalent to saying that the maximum total eventual number of offspring produced by an individual aged a is $\alpha(\omega - a + 1)$. In particular, the *pgf* of total LRS, $\boldsymbol{\varphi}_{\mathbf{r}_1^T}(x)$, reached via the iteration of Equation (19), is a vector of polynomials of maximum degree $\alpha\omega$. This implies that we can read directly the distribution of total LRS in the coefficients of the various $\varphi_{r_1(j)}(x)$, for $1 \le j \le d$. This is not possible in general with the formulation of the *pgf* of total LRS derived from Equation (11), which leads to ratios of functions (as per Equation I).

As a consequence, one can compute the iteration of Equation (19) and obtain the distribution of total LRS, without resorting to symbolic programming, by using classic operations on numeric matrices. Let us call $\mathscr{F}$ the numeric matrix such that $\mathscr{F}(i,j) = Pr(f(j) = i-1)$ is the probability that an individual in state j produces $i-1$ offspring in a given time-step. It corresponds to the coefficients of the polynomials in $\boldsymbol{\varphi}_{\mathbf{f}^T}(x)$ as $\varphi_{f(j)}(x) = \sum_{i=0}^{\alpha} \mathscr{F}(i+1,j)x^i$. Similarly, one can model $\boldsymbol{\varphi}_{\mathbf{r}_a^T}(x)$ via $\mathscr{R}_a$ where $\mathscr{R}_a(i,j) = Pr\left(r_a(j) = i-1\right)$. The iteration in *pgfs* of Equation (19) then becomes an iteration in numeric matrices:

$$\left(\begin{array}{l} \qquad\qquad \mathscr{R}_\omega = \mathscr{F} \\[1ex] \mathscr{R}_a = \mathscr{F} * \mathscr{Q}_{a+1}, \text{for all } 1 \le \mathrm{a} < \omega, \text{ with} \\[1ex] \qquad\qquad \mathscr{Q}_a = \mathscr{S} + \mathscr{R}_a\overline{\mathbf{S}} \\[1ex] \qquad\qquad \mathscr{R} = \mathscr{R}_1 \end{array}\right. , \tag{20}$$

where $*$ denotes convolution (i.e. $\mathscr{R}_a(i,j) = \sum_{k+l=i+1} \mathscr{F}(l,j)\mathscr{Q}_{a+1}(k,j)$) which yields the coefficients of the product of two polynomials and $\mathscr{S} = \boldsymbol{e}_1\left(\mathbf{1}^T - \mathbf{1}^T\overline{\mathbf{S}}\right)$, the numeric matrix which is zero apart from its first row corresponding to $\left(\mathbf{1}^T - \mathbf{1}^T\overline{\mathbf{S}}\right)$. This iteration ($\omega - 1$ steps) leads to $\mathscr{R}$, the probability distribution of total LRS for that model. Specifically, for any *type* j, $\mathscr{R}(i,j)$ is the probability that its total LRS is $i-1$.

4.3 | General case

The algorithm of Equation (20) provides a way to approximate the distribution of total LRS in the general case where the population projection model does not have finite reproduction (e.g. because it is a Poisson), and/or does not embed a maximum age. If age is not a trait of the model, one can, for a given ϵ_1, choose ω, such that, for any *type* j, $\sum_i \overline{S}^{\omega}(i,j) < \epsilon_1$, that is, such that the proportion of individuals surviving past age ω is inferior to ϵ_1 whatever their *type*. Similarly, if reproduction is not finite, one can, for a given ϵ_2, choose α, such that, for any state j, $Pr(f(j) > \alpha) < \epsilon_2$, that is, such that the probability to produce more than α offspring at a given time-step is at most ϵ_2. Of course ω and α can/should also be chosen according to the known ecology of the population. With these values, Equation (20) will yield an approximation for the probability distribution for total LRS, which deviation from the true LRS is controlled by ϵ_1 and ϵ_2. We provide the code for this algorithm (R and Matlab, see Supporting Information: Section VII).

Illustration: In the population model of the illustration, reproduction is finite with $\alpha = 1$. From the *pgfs* of reproduction, one can build $\mathscr{F} = \begin{bmatrix} 0.5 & 0 \\ 0.5 & 1 \end{bmatrix}$. There is no maximum age in that model, as age is not a trait (for any age a, the probability for an individual to live past age a is not zero). However, the probability for a *type small* or *type large* individual to live past age $\omega = 65$ is inferior to $\epsilon_1 = 1 \times 10^{-10}$. With this value of ω, we run the algorithm (Equation 20), starting with $\mathscr{R}_{65} = \mathscr{F}$. We have $\mathcal{Q}_{65} = \mathcal{S} + \mathscr{R}_{65}\bar{\mathbf{S}} = \begin{bmatrix} 0.5 & 0.3 \\ 0 & 0 \end{bmatrix} + \mathscr{R}_{65}\begin{bmatrix} 0.3 & 0 \\ 0.2 & 0.7 \end{bmatrix} = \begin{bmatrix} 0.65 & 0.3 \\ 0.35 & 0.7 \end{bmatrix}$. We perform the $\mathscr{R}_{64} = \mathscr{F} * \mathcal{Q}_{65}$ convolution, column by column. For the first column, this is

$$\begin{cases} \mathscr{R}_{64}(1,1) = \mathscr{F}(1,1)(t)\mathcal{Q}_{65}(1,1) = 0.325 \\ \mathscr{R}_{64}(2,1) = \mathscr{F}(2,1)(t)\mathcal{Q}_{65}(1,1) + \mathscr{F}(1,1)(t)\mathcal{Q}_{65}(2,1) = 0.5 \\ \mathscr{R}_{64}(3,1) = \mathscr{F}(2,1)(t)\mathcal{Q}_{65}(2,1) = 0.175 \end{cases}$$

Performing the convolution similarly for the second column leads to $\mathscr{R}_{64} = \begin{bmatrix} 0.325 & 0 \\ 0.5 & 0.3 \\ 0.175 & 0.7 \end{bmatrix}$. Repeating these operations 63 times, we reach $\mathscr{R}$, which plot corresponds to Figures 1c,d. Computationally, this method is very similar to that of Tuljapurkar et al. (2020) except that it proceeds through individual ages in the opposite direction. The algorithm described here considers that the distribution of the eventual number of offspring of an individual aged a, can be computed from that of an individual aged $a+1$ and initiates the process at the maximum age ω. In Tuljapurkar et al. (2020), the authors consider that the distribution of LRS is the mixing of the distribution of LRS conditional to age-at-death. The latter is first computed for individual dying in the first time-step, which has the distribution of reproduction. From this, they can compute the distribution of LRS conditional to age-at-death being 2, by convoluting the distribution of LRS conditional to age-at-death being 1 with that of reproduction, etc.

5 | EXTENSION II: ASYMPTOTIC BEHAVIOUR OF LRS PROBABILITY DISTRIBUTION

In this section, we illustrate the power of the BP approach by studying the probability to produce many offspring. First, we study this theoretically for unstructured populations, showing that the tail of the LRS distribution is geometric. Then, we show how this result can be useful to study the asymptotic behaviour of LRS probability distributions for a classic stage-structured model that (a) contains one type only, (b) does not embed a maximum age and (c) incorporates a stage for which the probability to be in, for any individual, tends towards 1 as age increases.

5.1 | Unstructured model

Consider a model where the population is unstructured and generations overlap: it corresponds to the BP process of Equation (5a) with $s = 1$. Each individual survives the time-step with probability $\bar{\mathfrak{s}}$ and reproduces, at each time-step, according to r.v. f. The LRS of this population model is a univariate r.v. r, such that (Equation 7) $r \overset{d}{=} f + \left(r \diamondsuit \mathfrak{s}\right)$, which leads to

$$\varphi_r(x) = \varphi_f(x)\varphi_{\mathfrak{s}}\left(\varphi_r(x)\right) = \varphi_f(x)(\bar{\mathfrak{s}}\left(\varphi_r(x) - 1\right) + 1) \tag{21}$$

Denoting $\mathcal{P}_k = Pr(r = k)$ and $\mathscr{F}_k = Pr(f = k)$, we can replace $\varphi_r(x)$ (respectively $\varphi_f(x)$) by $\sum_k \mathcal{P}_k x^k$ (respectively $\sum_k \mathscr{F}_k x^k$) in Equation (21). This yields (proof in Supporting Information: Section I), for all i:

$$\mathcal{P}_i\left(1 - \bar{\mathfrak{s}}\mathscr{F}_0\right) = \left(1 - \bar{\mathfrak{s}}\right)\mathscr{F}_i + \bar{s}\sum_{k=0}^{i-1} \mathcal{P}_k \mathscr{F}_{i-k} \tag{22}$$

which allows to construct recursively the probability distribution of LRS. Indeed, for $i = 0$, we get $\mathcal{P}_0 = \frac{(1-\bar{\mathfrak{s}})\mathscr{F}_0}{1-\bar{s}\mathscr{F}_0}$, for $i = 1$, $\mathcal{P}_1 = \frac{(1-\bar{\mathfrak{s}})\mathscr{F}_1 + \bar{s}\mathcal{P}_0\mathscr{F}_1}{1-\bar{\mathfrak{s}}\mathscr{F}_0}$, etc.

Asymptotic behaviour: The *pgf* framework provides a simple way to analyse the behaviour of $\mathcal{P}_k$ for large k. We show in Supporting Information: Section I, that we have

$$\lim_{i\to\infty} \frac{\mathcal{P}_{i+1}}{\mathcal{P}_i} = \rho, \text{where } \rho \text{ is solution of } \varphi_f\left(\rho^{-1}\right) = \frac{1}{\bar{\mathfrak{s}}} \tag{23}$$

This solution exists, and is unique, in all cases (as $\varphi_f(x)$ is a strictly growing function of x, with $\varphi_f(1) = 1$, $\frac{1}{\bar{s}} \geq 1$ and $\varphi_f(x) \to \infty$ when $x \to \infty$) but for $f = 0$. For f Bernoulli, we have $\varphi_f\left(\rho^{-1}\right) = \bar{f}\rho^{-1} + \left(1 - \bar{f}\right)$, which leads to $\rho = \frac{\bar{f}\bar{\mathfrak{s}}}{1-\bar{\mathfrak{s}}(1-\bar{f})}$. For f a Poisson (of parameter $\bar{f}$), we have $\varphi_f\left(\rho^{-1}\right) = e^{\bar{f}(\rho^{-1}-1)}$, which leads to $\rho = \frac{\bar{f}}{\bar{f} - ln(\bar{\mathfrak{s}})}$ and therefore

$$\text{for } f \text{ a Poisson, } \lim_{i\to\infty} \frac{\mathcal{P}_{i+1}}{\mathcal{P}_i} = \frac{\bar{f}}{\bar{f} - \ln\left(\bar{\mathfrak{s}}\right)} \tag{24}$$

5.2 | Asymptotic behaviour of the LRS of evergreen tree *Tsuga canadensis*

This constitutes a method that can be generalised to study the asymptotic behaviour of LRS for structured populations, but is sufficient to understand that of many stage-structured models as we illustrate now. We consider the annual demography of the evergreen tree *Tsuga canadensis*, which is used as an illustration by Tuljapurkar et al. (2020) and which fertility and survival matrices are given in their A.12 Supporting Information. This population is structured by $s = 6$ states (which are stages) the last 4 reproducing as Poisson. All offspring produced are in stage 1, therefore $d = 1$ and the total LRS corresponds to $r(1)$. From any stage but stage 6, individuals can remain in the same stage or transition to the next one; individuals in stage 6, if they survive, remain in that stage. We use the algorithm of Equation (20) to produce the distribution of the LRS (i.e. of $r(1)$) as well as the distributions of the eventual number of offspring produced by an individual in later stages ($r(j)$ for $2 \leq j \leq 6$) and display them in Figure 2 on a log scale. In Figure 2a, we consider the probability to produce up to 20 offspring; in Figure 2b, we extend it to 10,000 offspring.

Figure 2a shows the large 'zero-mode' of the LRS — $Pr(r(1)=0) \gg Pr(r(1)=1)$ — which is studied in Tuljapurkar et al. (2020) and caused, here, by the very low survival of the non-reproductive years (it can be observed for $r(1)$ and $r(2)$, but not for the reproductive stages $j \geq 3$). When considering the probability distribution of LRS, we see (Figure 2b) that asymptotically (for large k), $ln(Pr(r(1)=k))$ decreases linearly with k and that the same is true for all $r(j)$, with the same slope. This does not come as a surprise: stage 6 individuals remain in that stage until their die and the reproduction is Poisson therefore, from Equation (24), and by denoting $\overline{f_6}$ and $\overline{s_6}$ the expected fertility and survival of stage 6 individuals:

$$\lim_{k\to\infty}\frac{\Pr(r(6)=k+1)}{\Pr(r(6)=k)}=\frac{\overline{f_6}}{\overline{f_6}-\ln(\overline{s_6})}=\rho_6=0.99838$$

Therefore, for large k,

$$ln(Pr(r(6))=k+1) \approx ln(Pr(r(6)=k))+ln(\rho_6).$$

We display, as a thick magenta line, in Figure 2b, the $y(k)=\rho^k$ asymptote. Because, in this model, the reproduction of stage 6 individuals is much larger than that of other stages, and because individuals in stage 6 remain in that stage until they die, the probabilities $Pr(r(j)=k)$ for large k and for any j are governed by stage 6, and in particular the asymptotic probability distribution of the LRS:

$$\lim_{k\to\infty}\frac{\Pr(r(1)=k+1)}{\Pr(r(1)=k)}=\rho_6$$

This example illustrates the power of the BP approach: In practice, it is often unnecessary to compute the exact probability of extremely large LRS values, such as $k=10{,}000$, in order to understand the shape and tail behaviour of the LRS distribution. Instead, it is sufficient to (a) determine the asymptotic rate at which probabilities decrease, and (b) calculate the LRS probabilities for a modest range of offspring numbers, until the progression ratio converges sufficiently closely to its asymptote. For the Tsuga case, for example, we observe that $\left|\frac{Pr(r(1)=624)}{Pr(r(1)=623)}-\rho_6\right| \approx 1\times10^{-5}$ is small enough that we can approximate, $\forall k>624$,

$$Pr(r(1)=k) \approx Pr(r(1)=624)\rho_6^{k-624}$$

Applied to $k=10000$, this is $Pr(r(1)=10000) \approx 2.7915\times10^{-13}$ which is indeed close enough to the exact value (see Figure 2b) $Pr(r(1)=10000)=2.7899\times10^{-13}$, with a relative error inferior to 6×10^{-4}. We consider further properties of the asymptotic behaviour of LRS for Tsuga in Supporting Information: Section I.

6 | DISCUSSION

In this paper, via the framework of Branching Processes, we study the effects of demographic stochasticity on the LRS of a structured population. In the general case, where there are multiple *types* (newborn states) in a population, the LRS of each *type* is structured according to the *type* of its offspring. We provide a simple stochastic equation (Equation 7) relating LRS, survival and reproduction. To make this equation easier to use and emulate, we introduce, beforehand, a novel notation for the (demographic stochastic) projection of abundances over

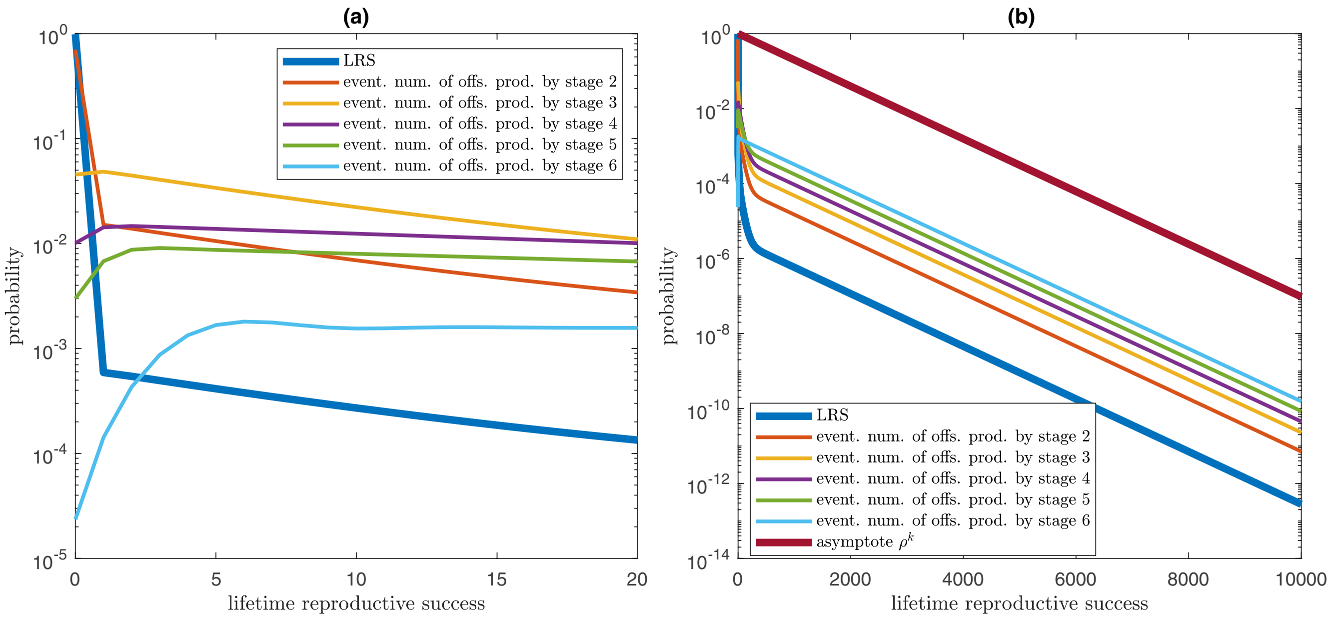

FIGURE 2 Probability distribution of lifetime reproductive success (LRS) for *Tsuga Canadiensis*. Probabilities are displayed on a log scale; panel (a) considers $Pr(r(j)=k)$ for $0 \leq k \leq 20$, panel (b) for $0 \leq k \leq 10{,}000$. The distribution of LRS ($r(1)$) is displayed in thick blue curve. The distributions of the eventual number of offspring $r(j)$ produced by an individual in stage j, for $2 \leq j$, are displayed via thinner curves. On panel (b), we also display the asymptotic slope ρ^k (thick magenta curve).

time, that uses the symbol ◊, where a classic, deterministic, projection would use matrix multiplication (Box 1). The relationship obtained (Equation 7) can lead to all moments of LRS as functions of the moments of reproduction and survival. In particular, it provides a direct formula for the variance–covariance of the various *types* produced during lifetime as a function of the variance in time-step reproduction and survival (Equation 10). The stochastic equation also directly provides an expression for the (multivariate) probability-generating functions for the different *types* (Equation 17), from which one can compute the exact joint distribution of LRS for the different *types*. As this computation requires a computer program dealing with functions and can be quite long, we provide, in Extension I, a numeric algorithm to generate the distribution of the total LRS (the total number of offspring produced, irrespective of their *types*), which we implement in R and Matlab.

6.1 | Multi-type projection models

In order to bridge the gap between the models of demography and genetics, it is important to be able to incorporate, within the classic framework of population projection matrices (Caswell, 2001), models incorporating various *types*. In a model with a single *type*, such as a population structured by age (only), there is no inheritance: all newborns are 'identical'. This is not the case of a Multi-Type model where one can study the mother-offspring correlation in *type*. In a model structured by genotypes or breeding values, the total number of offspring produced is a key component of fitness, but the way these newborns are distributed among *types* is crucial from an evolutionary *and* dynamical perspective. Some models of population ecology also embed multiple *types*. This is particularly true for Integral Projection Models (Metcalf et al., 2013; Rees et al., 2014; Vindenes et al., 2011) that embed quantitative traits, such as body mass: newborns can start their life with different body masses, and therefore be of different *types*, with consequences propagating towards subsequent survival and/or reproduction and therefore LRS. An important question, for instance, that requires understanding the properties of the LRS structured by *types*, and not just the total LRS, is that of its inheritance and in particular its heritability (its genetic inheritance) (Austerlitz & Heyer, 1998; Gustafsson, 1986; Merilä & Sheldon, 2000; Pettay et al., 2005). In essence, the present work contributes to the ongoing effort among ecologists to integrate demographic projection models with those of quantitative (e.g. see Coulson et al., 2017; Simmonds et al., 2020) and population (e.g. see de Vries & Caswell, 2019a, 2019b) genetics.

6.2 | Opportunity for selection

Despite claims (e.g. in Waples, 2023; Caswell, 2011) that being able to compute the variance in total LRS from the time-step demography of a structured population model allows one to compute the 'opportunity for selection' defined by Crow (1958), this is incorrect, as Crow's model assumes non-overlapping generations. As mentioned in the introduction, for a population projected over a chronological time-step, in a constant environment, it is the long-term population growth rate, and not the LRS, that evolution maximises. As a consequence, the opportunity for selection can only be appropriately computed via the stochastic population growth rate and its variance (caused by demographic stochasticity and often called 'demographic variance'; Engen et al., 1998), as illustrated by previous works on selection gradients (Lande, 1982) and effective size (Engen et al., 2005). Still, the distribution of LRS and of the population growth rate are linked. We show (in Supporting Information: Section V) that computing the probability of extinction of a population from the *pgf* of the LRS, as per (Tuljapurkar & Zuo, 2022), is equivalent to the classic method focusing on the *pgf* of the time-step projection process.

6.3 | Theoretical questions

Thanks to their simple forms, the stochastic equations derived from the BP approach are useful beyond numerical computation. They enable clear interpretation of how different ecological processes contribute to the moments and distribution of LRS. In Equation (15), for example, the variance–covariance in LRS is decomposed into variance in reproduction and variance in the time spent in the various states. For total LRS, this corresponds to a decomposition into variance in reproduction and in age-at-death. In Equation (10), the decomposition focuses on variance in survival and reproduction. Each of these components can be further decomposed. For instance, survival includes both survival per se and transitions between states; its variance can thus be broken down into the variance of these sub-processes. In spatially structured models, this yields a decomposition of LRS variance into contributions from reproduction, survival, dispersal rates and dispersal routes. Crucially, the BP approach facilitates theoretical analysis—as demonstrated, for example, in the study of the asymptotic behaviour of LRS (i.e. the probability of producing k offspring for very large k) in Extension II. As the relationship between models with multiple *types* and the fixed component of individual heterogeneity becomes clearer, the BP framework could also enrich recent studies that decompose LRS variance into contributions from fixed heterogeneity and demographic stochasticity (Snyder et al., 2021; Snyder & Ellner, 2018; Steiner & Tuljapurkar, 2020, 2012; Tuljapurkar & Steiner, 2010; van Daalen & Caswell, 2017, 2020). By accounting for multiple offspring *types*, our approach extends these analyses to incorporate inheritance of *types* and quantity–quality trade-offs.

6.4 | Consequences of the inputs

While theoretical advancements towards simpler formulas or faster computations of moments and distributions of LRS for more general models are important for advancing population ecology, demography and epidemiology, the quality of their outputs is, as always, dependent on the quality of the input, that is, of the projection model used. Information not implemented in the latter, will not be reflected in the former. Conversely, general assumptions of matrix population projection models, generally not made explicit in studies, such that the independence of the various states, will have consequences on the distribution of LRS. A model structured by age only, for example, assumes that survival and reproduction are independent from past reproduction; in other words, that there is no cost of reproduction in that population. In reality, trade-offs exist in many populations and they have strong effects on the variance and distribution of LRS (Coste & Pavard, 2020); still, they are extremely rarely incorporated in population projection models. Considering, for example, that 'every 26-year-old woman, without exception, produces one baby with a probability of 0.25 and zero babies with a probability of 0.75'(van Daalen & Caswell, 2017) assumes implicitly that the probability for a woman of a given age to produce its k^{th} offspring is independent from k and will therefore produce a biased variance in LRS compared with a model accounting for, for example, the parity progression ratios of that population (Retherford et al., 2010). Similarly, when we—along with Tuljapurkar et al. (2020) and van Daalen and Caswell (2017)—compute the distribution or variance of LRS of *Tsuga canadiensis* from a model structured by stage only, we assume implicitly that all individuals currently in a stage have the same probability to exit it during the current time-step (irrespective of the time spent in the stage), which can have vast implications for the distribution LRS.

6.5 | BPs and demographic stochasticity

Finally, this paper aims at facilitating the manipulation of BPs for ecologists—especially via the $\Diamond$ simplifying notation—so that they can derive, compute and study the moments, distribution and other characteristics of variables made random because of demographic stochasticity. The branching process framework is widely used in genetics (e.g. see Campbell, 2003; Lambert, 2006). Apart from a few notable exceptions—such as the study of density-dependent populations (Lambert, 2005), of extinction probabilities (Jeppsson & Forslund, 2012), of unstructured two-sex models (Daley et al., 1986; Hull, 2001)—and despite having been promoted for more than 50 years (Jagers, 1975, 1995; Pollard, 1966), the use of BPs in ecology is rare and researchers tend to resort to other, generally more complex, approaches as exemplified by the case of LRS. We encourage its broader use by theoretical ecologists to study demographic stochastic metrics. Of particular relevance would be the extension of this manuscript towards the study of LRS for a general, structured, two-sex population; it would allow, among other outputs, to understand the effects of the mating function on the joint distribution of the number of offspring (structured by *types*, including by sex).

ACKNOWLEDGEMENTS

This research was funded by the Natural Environment Research Council (grant no.: NE/W006731/1) and by the Norwegian Research Council (Norges Forskningsråd, grant/award numbers: 223257 and 343398). It also benefited from the support of the Chaire 'Modélisation Mathématique et Biodiversité of Veolia–Ecole Polytechnique–MNHN–Fondation X'. We thank Steinar Engen and three anonymous reviewers for very valuable comments.

CONFLICT OF INTEREST STATEMENT

The author declares no conflict of interest.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210X.70163.

DATA AVAILABILITY STATEMENT

No new data are used. All data used to illustrate the method are public. The code (R and Matlab) used for the illustration and to produce the figures is available via https://doi.org/10.5281/zenodo.17123366 (Coste, 2025); it is easily adaptable to any population projection model. A description of the code can be found in the online Supporting Information (Section VII).

ORCID

Christophe F. D. Coste https://orcid.org/0000-0003-3680-5049

REFERENCES

Athreya, K. B., & Karlin, S. (1971). Branching processes with random environments, II—Limit theorems. *Annals of Mathematical Statistics*, *42*(6), 1843–1858.

Athreya, K. B., & Ney, P. E. (1972). *Branching processes*, *volume 11*. Springer Berlin Heidelberg.

Austerlitz, F., & Heyer, E. (1998). Social transmission of reproductive behavior increases frequency of inherited disorders in a young-expanding population. *PNAS*, *95*(25), 15140–15144.

Bienaymé, I. J. (1845). De la loi de multiplication et de la durée des familles. *Société Philomatique de Paris*, *5*, 37–39.

Campbell, R. B. (2003). A logistic branching process for population genetics. *Journal of Theoretical Biology*, *225*(2), 195–203.

Caswell, H. (2001). *Matrix population models*. Sinauer Associates Inc.

Caswell, H. (2011). Beyond R0: Demographic models for variability of lifetime reproductive output. *PLoS One*, *6*(6), e20809.

Caswell, H. (2024). The formal demography of kinship VI: Demographic stochasticity and variance in the kinship network. *Demographic Research*, *51*, 1201–1256.

Charlesworth, B., & Williamson, J. A. (1975). The probability of survival of a mutant gene in an age-structured population and implications for the evolution of life-histories. *Genetical Research*, *26*(1), 1–10.

Coste, C. F., & Pavard, S. (2020). Analysis of a multitrait population projection matrix reveals the evolutionary and demographic effects of a life history trade-off. *Ecological Modelling*, *418*, 108915.

Coste, C. F. D. (2025). VarCovarDistributionLRS: code for article "A Branching Process Approach to Lifetime Reproductive Success of Structured Populations: Variance-Covariance and Distribution".

Coulson, T., Kendall, B. E., Barthold, J., Plard, F., Schindler, S., Ozgul, A., & Gaillard, J.-M. M. (2017). Modeling adaptive and nonadaptive responses of populations to environmental change. *The American Naturalist*, *190*(3), 313–336.

Crow, J. F. (1958). Some possibilities for measuring selection intensities in man. *Human Biology; an International Record of Research*, *30*(1), 1–13.

Crump, K., & Mode, C. J. (1969). A general age-dependent branching process. II. *Journal of Mathematical Analysis and Applications*, *25*(1), 8–17.

Crump, K. S., & Mode, C. J. (1968). A general age-dependent branching process. I. *Journal of Mathematical Analysis and Applications*, *24*(3), 494–508.

Cushing, J. M., & Zhou, Y. (1994). The net reproductive value and stability in matrix population models. *Natural Resource Modeling*, *8*(4), 297–333.

Daley, D. J., Hull, D. M., & Taylor, J. M. (1986). Bisexual Galton–Watson branching processes with superadditive mating functions. *Journal of Applied Probability*, *23*(3), 585–600.

de Vries, C., & Caswell, H. (2019a). Selection in two-sex stage-structured populations: Genetics, demography, and polymorphism. *Theoretical Population Biology*, *130*, 160–169.

de Vries, C., & Caswell, H. (2019b). Stage-structured evolutionary demography: Linking life histories, population genetics, and ecological dynamics. *The American Naturalist*, *193*(4), 545–559.

Diekmann, O., Heesterbeek, J. A., & Metz, J. A. (1990). On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology*, *28*(4), 365–382.

Dublin, L. I., & Lotka, A. J. (1925). On the true rate of natural increase. *Journal of the American Statistical Association*, *20*(151), 305–339.

Ellner, S. P. (2018). Generation time in structured populations. *The American Naturalist*, *192*(1), 105–110.

Engen, S., Bakke, O., & Islam, A. (1998). Demographic and environmental stochasticity-concepts and definitions. *Biometrics*, *54*(3), 840–846.

Engen, S., Lande, R., & Saether, B.-E. E. (2005). Effective size of a fluctuating age-structured population. *Genetics*, *170*(2), 941–954.

Engen, S., & Sæther, B.-E. (1998). Stochastic population models: Some concepts, definitions and results. *Oikos*, *83*(2), 345.

Gustafsson, L. (1986). Lifetime reproductive success and heritability: Empirical support for Fisher's fundamental theorem. *The American Naturalist*, *128*(5), 761–764.

Haccou, P., Jagers, P., & Vatutin, V. A. (2005). *Branching processes*. Cambridge University Press.

Harris, T. (1963). *The theory of branching processes*. Springer Berlin.

Henderson, H. V., & Searle, S. R. (1981). The vec-permutation matrix, the vec operator and Kronecker products: A review. *Linear and Multilinear Algebra*, *9*(4), 271–288.

Hull, D. M. (2001). A reconsideration of Lotka's extinction probability using a bisexual branching process. *Journal of Applied Probability*, *38*(3), 776–780.

Jagers, P. (1975). *Branching processes with biological applications*. John Wiley and Sons.

Jagers, P. (1982). How probable is it to be first born? And other branching-process applications to kinship problems. *Mathematical Biosciences*, *59*(1), 1–15.

Jagers, P. (1995). Branching processes as population dynamics. *Bernoulli*, *1*(1/2), 191.

Jeppsson, T., & Forslund, P. (2012). Can life history predict the effect of demographic stochasticity on extinction risk? *American Naturalist*, *179*(6), 706–720.

Joffe, A., & Waugh, W. A. O. (1985). The kin number problem in a multitype Galton-Watson population. *Journal of Applied Probability*, *22*(1), 37–47.

Karlin, S., & Taylor, H. E. (1975). *A first course in stochastic processes* (2nd ed.). Academic Press.

Lambert, A. (2005). The branching process with logistic growth. *The Annals of Applied Probability*, *15*(2), 1506–1535.

Lambert, A. (2006). Probability of fixation under weak selection: A branching process unifying approach. *Theoretical Population Biology*, *69*(4), 419–441.

Lande, R. (1982). A quantitative genetic theory of life history evolution. *Ecology*, *63*(3), 607–615.

Lefkovitch, L. P. (1965). The study of population growth in organisms grouped by stages. *Biometrics*, *21*(1), 1–18.

Leslie, P. H. (1945). On the use of matrices in certain population mathematics. *Biometrika*, *33*, 183–212.

Lewis, E. G. (1942). On the generation and growth of a population. *Sankhya*, *6*(part 1), 93–96.

Merilä, J., & Sheldon, B. C. (2000). Lifetime reproductive success and heritability in nature. *The American Naturalist*, *155*(3), 301–310.

Metcalf, C. J. E., McMahon, S. M., Salguero-Gómez, R., & Jongejans, E. (2013). *IPMpack*: An R package for integral projection models. *Methods in Ecology and Evolution*, *4*(2), 195–200.

Pettay, J. E., Kruuk, L. E. B., Jokela, J., & Lummaa, V. (2005). Heritability and genetic constraints of life-history trait evolution in preindustrial humans. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(8), 2838–2843.

Pollard, J. H. J. (1966). On the use of the direct matrix product in analysing certain stochastic population models. *Biometrika*, *53*(3/4), 397–415.

Rees, M., Childs, D. Z., & Ellner, S. P. (2014). Building integral projection models: A user's guide. *Journal of Animal Ecology*, *83*(3), 528–545.

Retherford, R., Ogawa, N., Matsukura, R., & Eini-Zinab, H. (2010). Multivariate analysis of parity progression-based measures of the total fertility rate and its components. *Demography*, *47*(1), 97–124.

Simmonds, E. G., Cole, E. F., Sheldon, B. C., & Coulson, T. (2020). Testing the effect of quantitative genetic inheritance in structured models on projections of population dynamics. *Oikos*, *129*(4), 559–571.

Snyder, R. E., & Ellner, S. P. (2018). Pluck or luck: Does trait variation or chance drive variation in lifetime reproductive success? *The American Naturalist*, *191*(4), E90–E107.

Snyder, R. E., Ellner, S. P., & Hooker, G. (2021). Time and chance: Using age partitioning to understand how luck drives variation in reproductive success. *The American Naturalist*, *197*(4), E110–E128.

Steiner, U. K., & Tuljapurkar, S. (2012). Neutral theory for life histories and individual variability in fitness components. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(12), 4684–4689.

Steiner, U. K., & Tuljapurkar, S. (2020). Drivers of diversity in individual life courses: Sensitivity of the population entropy of a Markov chain. *Theoretical Population Biology*, *133*, 159–167.

Tuljapurkar, S., & Steiner, U. K. (2010). Dynamic heterogeneity and life histories. *Annals of the New York Academy of Sciences*, *1204*(1), 65–72.

Tuljapurkar, S., & Zuo, W. (2022). Mutations and the distribution of lifetime reproductive success. *Journal of the Indian Institute of Science*, *102*(4), 1269–1275.

Tuljapurkar, S., Zuo, W., Coulson, T., Horvitz, C., & Gaillard, J. M. (2020). Skewed distributions of lifetime reproductive success: Beyond mean and variance. *Ecology Letters*, *23*(4), 748–756.

Tuljapurkar, S., Zuo, W., Coulson, T., Horvitz, C., & Gaillard, J. M. (2021). Distributions of LRS in varying environments. *Ecology Letters*, *24*(September 2020), 1328–1340.

van Daalen, S. F., & Caswell, H. (2017). Lifetime reproductive output: Individual stochasticity, variance, and sensitivity analysis. *Theoretical Ecology*, *10*(3), 355–374.

van Daalen, S. F., & Caswell, H. (2020). Variance as a life history outcome: Sensitivity analysis of the contributions of stochasticity and heterogeneity. *Ecological Modelling*, *417*(September 2019), 108856.

Vegvari, C., Abbott, S., Ball, F., Brooks-Pollock, E., Challen, R., Collyer, B. S., Dangerfield, C., Gog, J. R., Gostic, K. M., Heffernan, J. M., Hollingsworth, T. D., Isham, V., Kenah, E., Mollison, D., Panovska-Griffiths, J., Pellis, L., Roberts, M. G., Scalia Tomba, G., Thompson, R. N., & Trapman, P. (2022). Commentary on the use of the reproduction number R during the COVID-19 pandemic. *Statistical Methods in Medical Research*, *31*(9), 1675–1685.

Vindenes, Y., Engen, S., & Saether, B.-E. (2011). Integral projection models for finite populations in a stochastic environment. *Ecology*, *92*(5), 1146–1156.

Waples, R. S. (2023). Partitioning variance in reproductive success, within years and across lifetimes. *Ecology and Evolution*, *13*(11), 1–25.

Waples, R. S., Do, C., & Chopelet, J. (2011). Calculating N e and N e/N in age-structured populations: A hybrid Felsenstein-Hill approach. *Ecology*, *92*(7), 1513–1522.

Watson, H. W., & Galton, F. (1875). On the probability of the extinction of families. *The Journal of the Anthropological Institute of Great Britain and Ireland*, *4*(1875), 138–144.

Waugh, W. A. O. (1981). Application of the Galton-Watson process to the kin number problem. *Advances in Applied Probability*, *13*(4), 631–649.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Supporting Information: Section I. Proofs and intermediary steps.

Supporting Information: Section II. Covarying survival and reproduction.

Supporting Information: Section III. Alternative formulations for the variance in LRS.

Supporting Information: Section IV. Lifetime spent in various states and age at death.

Supporting Information: Section V. Extinction probability.

Supporting Information: Section VI. LRS of newborn at random in a cohort.

Supporting Information: Section VII. Code (description).

Supporting Information: Section VIII. References (cited in Supporting Information).

How to cite this article: Coste, C. F. D. (2025). A branching process approach to lifetime reproductive success of structured populations: Variance–covariance and distribution. *Methods in Ecology and Evolution*, *00*, 1–21. https://doi.org/10.1111/2041-210X.70163