The Kinship Matrix:
Inferring the Kinship Structure of a Population from its Demography

Christophe F. D. Coste[⋆,†,1], François Bienvenu[⋆,2,3,4,5], Victor Ronget[6], Juan-Pablo Ramirez-Loza[7], Sarah Cubaynes[7], and Samuel Pavard[6]

[1]Centre for Biodiversity Dynamics, Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway
[2]Department of Mathematics and Computer Science, Transilvania University of Brașov, 500091 Brașov, Romania.
[3]Institut des Sciences de l'Evolution de Montpellier (Université de Montpellier, CNRS, IRD, EPHE), F-34095 Montpellier, France.
[4]UMR AGAP (Université de Montpellier, CIRAD, INRAE, L'institut Agro), F-34398 Montpellier, France.
[5]Department of Statistics, University of Oxford, 24-29 St Giles', Oxford OX1 3LB, United Kingdom.
[6]Unité Eco-anthropologie (EA), Muséum National d'Histoire Naturelle, CNRS, Université Paris Diderot, F-75016, Paris, France.
[7]CEFE, CNRS, Univ. Montpellier, Univ. Paul Valéry Montpellier 3, EPHE, IRD, Montpellier, France.

June 25, 2021

Abstract

The familial structure of a population and the relatedness of its individuals are determined by its demography. There is, however, no general method to infer kinship directly from the life-cycle of a structured population. Yet this question is central to fields such as ecology, evolution and conservation, especially in contexts where there is a strong interdependence between familial structure and population dynamics. Here, we give a general formula to compute, from any matrix population model, the expected number of arbitrary kin (sisters, nieces, cousins, etc) of a focal individual *ego*, structured by the class of *ego* and of its kin. Central to our approach are classic but little-used tools known as genealogical matrices. Our method can be used to obtain both individual-based and population-wide metrics of kinship, as we illustrate. It also makes it possible to analyze the sensitivity of the kinship structure to the traits implemented in the model.

Keywords: kinship demography; relatedness; familial structure; structured populations; population projection matrices.

1 Introduction

The demography of a population determines its level of inbreeding and its kinship structure, i.e. the network of kin relationships between the individuals of the population. Population geneticists have long used non-overlapping generation models to study the effects of demography on kinship structure, for instance to estimate the frequency of consanguineous marriage (Barrai et al., 1962; Hajnal, 1963). Models with overlapping generations have also been considered – from the seminal work of Goodman et al. (1974) to the recent advances by Caswell (2019, 2020) – but they focus on specific kin relationships. Thus, a general method has yet to be developed to infer kinship structure from population dynamics in structured populations, and little is known about the influence of specific life history traits (or vital rates in general) on inbreeding levels, inclusive fitness and kinship/familial structures. Yet those questions are central to fields such as ecology, demography, evolution, genetics and conservation.

In this article, we pave the way towards such a general method by showing how classic tools from the study of matrix population models can be combined in a new way to infer kinship structure from steady-state population dynamics. More specifically, we explain how to compute the expected number of kin of a focal individual *ego*, as a function of the classes of *ego* and of its kin, directly from the projection matrix of a population. Examples of applications include studying how inbreeding depression leads to genetic Allee effects in populations of small effective sizes or modeling the eco-evolutionary demography of social species (where the vital rates of individuals may strongly depend on cooperative and competitive interactions with kin). Readily usable implementations of our methods are available at `10.5281/zenodo.4680715`, see SM.1.

⋆ Co-first authors.
† Corresponding author.

1.1 A brief history of kinship inference in structured populations

The first demographers to work on the estimation of kin distribution from the demographic rates of a structured population based their computations on age-structured, one-sex populations at the demographic steady state, and focused on a few key kin relationships (Le Bras, 1973; Goodman et al., 1974, discussed in Pavard and Coste, 2020). Their method, still widely used today, consists in computing the distribution of maternal age-at-childbirth from the Euler–Lotka equation (Lotka, 1939), and then use it to derive, for each age-class, the probability that the mother/grandmother of a focal individual is alive, as well as the expected number of daughters, sisters, aunts and nieces of that individual.

Shortly after, mathematicians attacked the problem from the field of branching processes. First, for unstructured populations without overlapping generations, using Bienaymé–Galton–Watson processes (Waugh, 1981; Pullum, 1982); then, by adding age as a category and considering multitype Crump-Mode-Jagers processes (e.g., Nerman and Jagers, 1984). However, the technical nature of these papers and their focus on sometimes abstract results has made them go largely unnoticed by demographers. For instance, despite being directly relevant to the matter, the work by Joffe and Waugh (1985) is only cited nine times in the mathematical literature – and not once in a biology or demography journal.

As of today, the inferred kin frequencies provided by Goodman et al. (1974), which have been in constant use since their publication, are still the state of the art (Pavard and Coste, 2020). These take the form of a limited number of *ad hoc* formulas, each specific to a kin relationship, that only apply to age-structured populations where at most one offspring can be produced at a time. Recently, Caswell (2019, 2020) has put this framework in matrix form and extended it to age×stage models, by giving a system of recursive equations expressing the number of kin at time $t+1$ as a function of the number of kin a time t. However, before the present work computing the expected number of any kin in any life cycle remained an open problem and a necessary step in the development of a general theory of the interplay between kinship and demographic processes (Pavard and Coste, 2020).

1.2 Specific challenges to overcome

Inferring the complete kinship structure of an arbitrary structured population poses several challenges. First, in general class-structured models the distribution of the class at birth is not as easily inferred as in age-structured populations, where it is obtained from the Euler–Lotka equation. We solve this problem by using the genealogical Markov chains first introduced by Demetrius to define population entropy (Demetrius, 1974, 1975). These somewhat little-known tools have recently proved useful to tackle various questions (Bienvenu and Legendre, 2015; Bienvenu et al., 2017).

The fact that several offspring can be produced at the same time constitutes another complication, because this prevents segregating genealogies as done in Goodman et al. (1974) and Caswell (2019, 2020). Accounting for the possibility of same-litter sisters requires special care and fine-grained information on reproduction.

Finally, the main difficulty is arguably the generalization of the calculations to arbitrary kin relationships. This requires combining two different timescales: the demographic timescale (where time is expressed in fixed projection intervals, such as one year) and the genealogical one (where time is expressed in number of generations). We do so by working out all the genealogies and associated trajectories of individuals in the life cycle that correspond to a given kin relationship.

1.3 The kinship matrices

In order to describe the kinship structure of a population, we need a generic notation for arbitrary kin relationships. In one-sex populations, we can use the following: we say that two individuals (i_1, i_2) are (g, q)-*kin* when their most recent common ancestor is separated from i_1 by g generations and from i_2 by q generations. For instance, the $(1, 0)$-kin of the focal individual *ego* is its mother, and its $(0, 1)$-kin are its daughters; similarly, its $(1, 1)$-kin are its sisters; its $(1, 2)$-kin its nieces; its $(2, 1)$-kin its aunts; etc (see Figure 1A). The same convention was used by Atkins (1974) and Pullum (1982), but with the order of g and q reversed.

We describe the kinship structure of a class-structured population through the *kinship matrices* $\mathbf{K}(g, q)$, where the (i, j)-th entry of $\mathbf{K}(g, q)$ is the expected number of (g, q)-kin alive in class i of *ego* in class j. From the point of view of *ego* – which is what matters for most applications – these matrices completely characterize the (expected) kinship structure of the population. An example of the genealogy of a focal individual, together with the corresponding population model and kinship matrix for daughters, is given in Figures 1B and 1C.

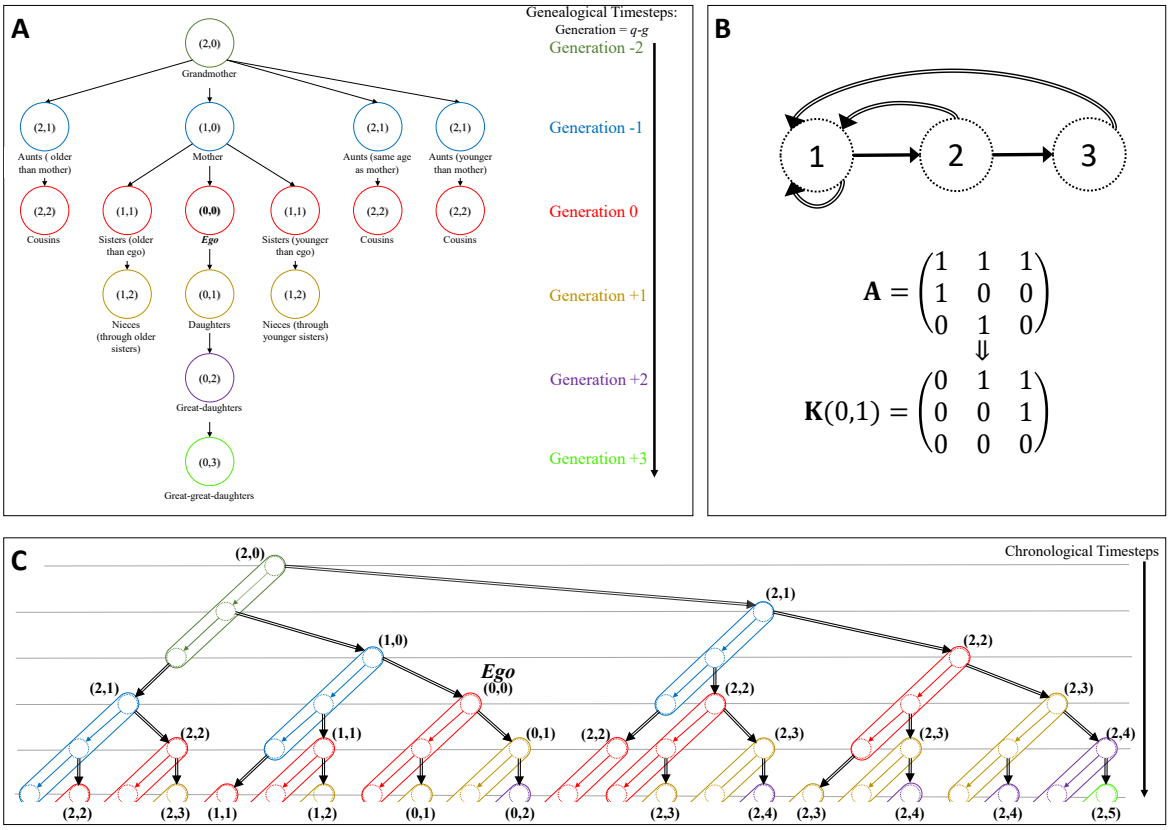


Figure 1: **(A)** Kin relationships, as described in terms of (g, q)-kin of the focal individual *ego*. Each line (and colour) corresponds to a generation. To get to the (g, q)-kin of *ego*, one has to go "up" g generations in order to reach the most recent common ancestor of *ego* and its (g, q)-kin; and then "down" q generations to reach the (g, q)-kin themselves. **(B)** Simple example of an age-structured life-cycle, with the corresponding projection matrix and kinship matrix $\mathbf{K}(0, 1)$ giving the expected number of daughters of *ego*. In this toy model, all individuals survive to age 3 and then die, and they produce exactly one offspring each year. Thus, individuals aged 2 have exactly one daughter (aged 1) and individuals aged 3 have exactly two daughters (one aged 1 and one aged 2). **(C)** The genealogy of *ego* for the population described in (B), with the corresponding kin structure. Each dotted circle corresponds to an individual at a given age, and each solid oval to an individual's life trajectory. The colors corresponds to the generation with respect to *ego*, as in (A). Here we have assumed that *ego* was born when its mother was aged 1, and that the mother of *ego* was born when its grandmother was aged 2.

The rest of this article is organized as follows: in Section 2, we explain how to compute the kinship matrices directly from the projection matrix of the population. We then illustrate the method in Box 2 using the life-cycle of the ground squirrel *Spermophilus dauricus* from Luo and Fox (1990); other detailed examples of applications can be found in SM.2. Finally, we conclude by discussing the limitations, implications and possible extensions of our method in Section 3.

2 Method

Our method is applicable to populations at the demographic steady state whose dynamics are governed by a matrix population model with a primitive (that is, irreducible and aperiodic) projection matrix $\mathbf{A} = (a_{ij})$. It can also be used with some reducible matrices (see SM.3), but here we assume primitivity, for simplicity. Recall that a_{ij} is the per-capita contribution of an individual in class j at time t to the abundance of class i at time $t+1$. The asymptotic growth rate λ is given by the dominant eigenvalue of $\mathbf{A}$ and the stable class distribution by the corresponding right eigenvector $\mathbf{w} = (w_i)$ such that $\sum_i w_i = 1$.

To compute the kinship matrix $\mathbf{K}(g, q)$, we decompose the projection matrix $\mathbf{A}$ into its survival and reproductive components – that is, into $\mathbf{A} = \mathbf{S} + \mathbf{F}$, where the *survival matrix* $\mathbf{S} = (s_{ij})$ is such that s_{ij} is the probability that an individual of class j survives into class i, and the *fertility matrix* $\mathbf{F} = (f_{ij})$ is such that f_{ij} is the expected number of offspring of class i produced by an individual of class j. Mathematically, we need reproduction and survival to be independent (it is otherwise not possible to retrace the life trajectory of an ancestor and then use this trajectory to compute its number of offspring). This assumption is never applicable in post-breeding models, where individuals have to survive a full projec-

tion interval before they can reproduce. We therefore recommend using pre-breeding models whenever possible and avoiding post-breeding models with our method. Birth-flow models can also be used as long as the covariance between survival and reproduction is kept in mind as a potentially significant source of discrepancy. Finally, note that one way to deal with covarying reproduction and survival (e.g, due to trade-offs) is to add (survival, fertility) classes to the model.

Finally, to take into account models with several offspring per time-step, we need the *same-litter newborn sisters matrix* $\mathbf{Z} = (z_{ij})$, where z_{ij} is the expected number of same-litter sisters of *ego* that have just been born in class i, knowing that *ego* is in class j. Unlike $\mathbf{S}$ and $\mathbf{F}$, the matrix $\mathbf{Z}$ is not always directly available, and to compute it one will typically need to make additional assumptions on the offspring distribution. We give expressions of $\mathbf{Z}$ under three of the most common scenarios: when at most one offspring is produced by projection interval; when there is only one class of offspring; and when the total number of offspring follows a Poisson distribution and each offspring is independently allocated to a newborn class.

The three matrices $\mathbf{S}$, $\mathbf{F}$ and $\mathbf{Z}$ are all we need to compute any kinship matrix $\mathbf{K}(g,q)$, as summarized in Box 1. Python and MATLAB implementations of the method are provided (see SM.1). In the rest of this section, we detail the reasoning behind the expression of $\mathbf{K}(g,q)$ and give examples of aggregated metrics of the kinship or relatedness structure of a population that can be derived from it. Concrete examples of application are given in Box 2 and SM.2.

Box 1: Expression of the kinship matrix $\mathbf{K}(g,q)$

- Input: let $\mathbf{A} = \mathbf{S} + \mathbf{F}$ be the projection matrix of the population, where $\mathbf{S}$ is the survival matrix and $\mathbf{F}$ is the fertility matrix. Let $\mathbf{Z}$ be the same-litter newborn sisters matrix (see Section 2.4).
- Compute the genealogical matrices $\mathbf{P_S} = \big(p_{\mathbf{S}}(i,j)\big)$ and $\mathbf{P_F} = \big(p_{\mathbf{F}}(i,j)\big)$ defined by

$$p_{\mathbf{S}}(i,j) \;=\; \frac{s_{ij}\, w_j}{\lambda\, w_i} \qquad \text{and} \qquad p_{\mathbf{F}}(i,j) \;=\; \frac{f_{ij}\, w_j}{\lambda\, w_i}\,,$$

  where λ is the asymptotic growth rate (dominant eigenvalue of $\mathbf{A}$) and $\mathbf{w}$ the stable class distribution vector (dominant right eigenvector of $\mathbf{A}$).
- For all integers $n \geq 0$ and $\ell \geq 1$, let $\mathcal{C}(n,\ell)$ denote the set of vectors $\mathbf{k} = (k_1, \ldots, k_\ell)$ of non-negative integers such that $k_1 + \cdots + k_\ell = n$ (when $\ell = 0$ or $n < 0$, let $\mathcal{C}(n,\ell) = \varnothing$).
- For any $\mathbf{k} = (k_1, \ldots, k_\ell)$, let $\mathbf{Q}(\mathbf{k}) = \mathbf{P_S}^{k_1}\, \mathbf{P_F} \cdots \mathbf{P_F}\, \mathbf{P_S}^{k_\ell}$ and

$$\mathbf{U}(t,g) \;=\; \sum_{\mathbf{k} \in \mathcal{C}(t-g,\, g)} \mathbf{Q}(\mathbf{k})\,,$$

  with the convention that $\mathbf{U}(0,0) = \mathbf{I}$, the identity matrix, and that $\mathbf{U}(t,g) = \mathbf{0}$ when the sum is empty.
- For any $\mathbf{k} = (k_1, \ldots, k_\ell)$, let $\mathbf{R}(\mathbf{k}) = \mathbf{S}^{k_1}\, \mathbf{F} \cdots \mathbf{F}\, \mathbf{S}^{k_\ell}$ and

$$\mathbf{D}(t,q) \;=\; \sum_{\mathbf{k} \in \mathcal{C}(t-1-q,\, q+1)} \mathbf{R}(\mathbf{k})\,,$$

  with the convention that $\mathbf{D}(t,q) = \mathbf{0}$ when the sum is empty.
- The kinship matrix $\mathbf{K}(g,q)$, whose (i,j)-th entry is the expected number of (g,q)-kin of class i of a focal individual of class j, is $\mathbf{K}(0,0) = \mathbf{I}$ and, for $(g,q) \neq (0,0)$,

$$\mathbf{K}(g,q) \;=\; \sum_{t \geq 0} \Big(\mathbf{D}(t,q)\, \mathbf{S}\, \mathbf{P_F}^{\top} \;+\; \mathbf{D}(t,q-1)\, \mathbf{Z} \;+\; \sum_{\alpha \geq 1} \mathbf{D}(t+\alpha, q-1)\, \mathbf{F}\, (\mathbf{P_S}^{\top})^{\alpha}\, \mathbf{Y}(g) \Big)\, \mathbf{U}(t,g)^{\top},$$

  where $\mathbf{Y}(g) = \mathbf{I}$ if $g = 0$ and $\mathbf{P_F}^{\top}$ otherwise.

2.1 The genealogical Markov chain

The central tool of our approach is the genealogical Markov chain introduced by Demetrius (1974, 1975) to define population entropy. Despite its early introduction in the field of matrix population models, this Markov chain has remained mostly associated with population entropy (see, e.g., Tuljapurkar, 1982, 1993) and has so far failed to become part of the standard toolbox of population ecologists and demographers. However, it has recently proved useful to study questions such as the generation time (Bienvenu and Legendre, 2015) and the optimal aggregation of classes (Bienvenu et al., 2017). Our work provides yet another example of the usefulness of this tool.

The transition matrix of the genealogical Markov chain associated with $\mathbf{A}$ is the matrix $\mathbf{P} = (p_{ij})$, where

$$p_{ij} \;=\; \frac{a_{ij}\, w_j}{\lambda\, w_i}\,. \tag{1}$$

This Markov chain describes the sequence of classes that we encounter as we go "up" the genealogy of the population by following the lineage of an individual, backwards in time. That is, p_{ij} is the probability that an individual in class i at time t was either alive and in class j at time $t-1$; or else had its mother in class j at that time. Thus, the probability distribution of the class of the ancestor (in the sense of a younger self or a genealogical ancestor), t time steps ago, of an individual currently in class i is given by $\mathbf{e}_i\mathbf{P}^t$, where $e_i(j) = 1$ if $j = i$ and 0 otherwise (note that the formalism of Markov chain uses right-multiplication to project probability distributions, as in $\mathbf{x}(t+1) = \mathbf{x}(t)\,\mathbf{P}$, whereas population projection matrices use left-multiplication). See Bienvenu et al. (2017) for more on genealogical Markov chains.

As we did with $\mathbf{A}$, we can split the transitions of the genealogical matrix into a survival and a reproductive component. We then have $\mathbf{P} = \mathbf{P_S} + \mathbf{P_F}$, where $\mathbf{P_S} = \big(p_{\mathbf{S}}(i,j)\big)$ and $\mathbf{P_F} = \big(p_{\mathbf{F}}(i,j)\big)$ are given by

$$p_{\mathbf{S}}(i,j) \;=\; \frac{s_{ij}\, w_j}{\lambda\, w_i} \qquad \text{and} \qquad p_{\mathbf{F}}(i,j) \;=\; \frac{f_{ij}\, w_j}{\lambda\, w_i}\,. \tag{2}$$

Accordingly, $p_{\mathbf{S}}(i,j)$ is the probability that an individual in class i at time t was alive and in class j at time $t-1$, while $p_{\mathbf{F}}(i,j)$ is the probability that it was born between times $t-1$ and t to a mother in class j. Thus, for instance, the (i,j)-th entry of $\mathbf{P_S}\mathbf{P_F}$ is the probability that the mother of an individual from class i was in class j two time-steps ago *and* that the first transition when going up the genealogy was a survival transition, while the second one was a reproductive transition.

We are now able to: (1) project the class of the ancestor of *ego* backwards in time (within generations using $\mathbf{P_S}$ and between generations using $\mathbf{P_F}$); and (2) project the descent of an ancestor forwards in time (within generations using $\mathbf{S}$ and between generations using $\mathbf{F}$). With this, our strategy to compute the number of (g,q)-kin of *ego* is to go "up" the genealogy to locate its g-th ancestor; and then "down" to compute the q-th descendants of that ancestor. However, to account for same-litter sisters, we also need to look at the $(g-1)$-th ancestor of *ego*.

2.2 Distribution of the $(g-1)$-th ancestor of *ego*

Let us start by assuming that $g \geq 1$ (the case $g = 0$ will be treated separately later). For $t \geq 1$, let $\mathbf{U}(t,g) =$ be the matrix whose (i,j)-th entry is the probability that $t-1$ time steps ago the $(g-1)$-th ancestor of *ego* was in class j, given that *ego* is currently in class i. To compute $\mathbf{U}(t,g)$, we need to consider all the ways that $g-1$ generations can pass in the span of $t-1$ time steps. For any integers $n \geq 0$ and $\ell \geq 1$, let

$$\mathcal{C}(n,\ell) \;=\; \Big\{(k_1,\ldots,k_\ell) \in \mathbb{N}^\ell : k_1 + \cdots + k_\ell = n\Big\} \tag{3}$$

denote the set of vectors of length ℓ whose components are non-negative integers that sum to n (with $\mathcal{C}(n,\ell) = \varnothing$ when $\ell = 0$ or $n < 0$). Such vectors are known as *weak compositions of the integer n into ℓ parts*, and there are efficient algorithms to list them (see SM.4). The weak compositions $\mathbf{k} \in \mathcal{C}(t-g,g)$ exactly encode the sequences of $t-1$ "survival" or "reproduction" events that contain $g-1$ "reproduction": indeed, $\mathbf{k} = (k_1,\ldots,k_g)$ corresponds to the sequence

$$\underbrace{\text{survival},\ldots,\text{survival}}_{k_1 \text{ times}},\ \text{reproduction},\ \cdots,\ \text{reproduction},\ \underbrace{\text{survival},\ldots,\text{survival}}_{k_g \text{ times}},$$

Thus, letting $\mathbf{Q}(\mathbf{k}) = \mathbf{P_S}^{k_1}\,\mathbf{P_F}\cdots\mathbf{P_F}\,\mathbf{P_S}^{k_g}$ – whose (i,j)-th entry is the probability that, starting from class i, after $k_1 + \ldots + k_g + g - 1$ steps one ends in class j after having encountered k_1 survivals, then one reproduction, then k_2 survivals, etc. – we have

$$\mathbf{U}(t,g) \;=\; \sum_{\mathbf{k}\in\mathcal{C}(t-g,g)} \mathbf{Q}(\mathbf{k})\,, \tag{4}$$

with the convention that $\mathbf{U}(t,g) = \mathbf{0}$ when the sum is empty (e.g., if $t < g$).

As examples of the formula, note that for $g = 1$ and $t \geq 1$, $\mathbf{U}(t,1) = \mathbf{P_S}^{t-1}$, and for $g = 2$ and $t = 2,3,4,\ldots$

$$\mathbf{U}(2,2) = \mathbf{P_F}, \qquad \mathbf{U}(3,2) = \mathbf{P_F}\mathbf{P_S} + \mathbf{P_S}\mathbf{P_F}, \qquad \mathbf{U}(4,2) = \mathbf{P_F}\mathbf{P_S}^2 + \mathbf{P_S}\mathbf{P_F}\mathbf{P_S} + \mathbf{P_S}^2\mathbf{P_F}, \qquad \ldots$$

Finally, letting $\mathbf{U}(t,g) = (u_{ij}(t,g))$, the probability, given that *ego* is currently in class i, that its g-th ancestor gave birth to its $(g-1)$-th ancestor in class ℓ, from class k and between times t and $t-1$ before

the present is $u_{i\ell}(t,g)\,p_{\mathbf{F}}(\ell,k)$. The reason why we need to look at both the $(g-1)$-th and the g-th ancestor of *ego*, rather than at its g-th ancestor only, is that we will need this information to account for the (g,q)-kin *ego* that are descendants of same-litter sisters of its $(g-1)$-th ancestor. Note that the probability, given that *ego* is currently in class i, that its g-th ancestor gave birth to its $(g-1)$-th ancestor from class k and between times t and $t-1$ before the present is the (i,k)-th entry of $\mathbf{U}(t,g)\mathbf{P}_{\mathbf{F}}$.

2.3 Expected number of q-th generation descendants

Let $\mathbf{D}(t,q)$ be the matrix whose (i,j)-th entry is the expected number of q-th generation descendants in class i, after $t-1$ time steps, of an individual in class j (not counting offspring produced before the first time-step). As before, to compute $\mathbf{D}(t,q)$ we need to consider all the ways that q generations can pass in the span of $t-1$ time steps. For $\mathbf{k}=(k_1,\ldots,k_\ell)$, let $\mathbf{R}(\mathbf{k}) = \mathbf{S}^{k_1}\mathbf{F}\cdots\mathbf{F}\,\mathbf{S}^{k_\ell}$. Recall that the expected number of descendants in class i of an individual in class j, after $t-1$ time steps (and counting the initial individual) is the (i,j)-th entry of the matrix

$$\mathbf{A}^{t-1} \;=\; \big(\mathbf{S}+\mathbf{F}\big)^{t-1} \;=\; \sum_{q=0}^{t-1}\sum_{\mathbf{k}\in\mathcal{C}(t-1-q,\,q+1)}\mathbf{R}(\mathbf{k})\,.$$

For instance, for $t-1=3$,

$$\mathbf{A}^3 \;=\; \underbrace{\mathbf{S}^3}_{q=0} \;+\; \underbrace{\mathbf{FS}^2+\mathbf{SFS}+\mathbf{S}^2\mathbf{F}}_{q=1} \;+\; \underbrace{\mathbf{F}^2\mathbf{S}+\mathbf{FSF}+\mathbf{SF}^2}_{q=2} \;+\; \underbrace{\mathbf{F}^3}_{q=3}\,.$$

In these expressions, the terms corresponding to a fixed value of q correspond to the q-th generation descendants of the initial individual. In other words,

$$\mathbf{D}(t,q) \;=\; \sum_{\mathbf{k}\in\mathcal{C}(t-1-q,\,q+1)}\mathbf{R}(\mathbf{k})\,, \tag{5}$$

again with the convention that $\mathbf{D}(t,q)=\mathbf{0}$ when the sum is empty (e.g, when $q<0$ or $t-1<q$).

2.4 Expected number of (g,q)-kin

Let *ego* be in class j, and recall that we assume $g\geq 1$. For t ranging over the positive integers and k,ℓ over the classes of the model, the events "the g-th ancestor of *ego* gave birth to its $(g-1)$-th ancestor in class ℓ, from class k, and $t-1$ time-steps before the present" form a complete system of events. Each of these events has probability $u_{j\ell}(t,g)\,p_{\mathbf{F}}(\ell,k)$. Thus, all we need in order to compute the expected number of (g,q)-kin of *ego* is the expected number of q-th generation descendants of an individual that was in class k at time t before the present and had an offspring born in class ℓ between times t and $t-1$. For this, we need to distinguish between descendants of younger, older and same-litter sisters of the $(g-1)$-th ancestor of *ego* (see Figure 2).

Descendants of younger sisters of the $(g-1)$-th ancestor

Between times t and $t-1$ before the present, the g-th ancestor of *ego* survives from class k to class m with probability s_{mk}. From there, the expected number of q-th generation descendants that it leaves in class i after the remaining $t-1$ time steps is $d_{im}(t,g)$, the (i,m)-th entry of the matrix $\mathbf{D}(t,g)$. The expected number of class-i (g,q)-kin of *ego* through younger sisters of its $(g-1)$-th ancestor is therefore $\sum_{t\geq 0}\sum_{k,\ell,m} d_{im}(t,q)\,s_{mk}\,u_{j\ell}(t,g)\,p_{\mathbf{F}}(\ell,k)$ – which, in matrix notation, is the (i,j)-th entry of

$$\mathbf{K}_{\mathrm{y}}(g,q) \;=\; \sum_{t\geq 1}\mathbf{D}(t,q)\,\mathbf{S}\,\mathbf{P}_{\mathbf{F}}^{\top}\mathbf{U}(t,g)^{\top}. \tag{6}$$

Descendants of older sisters of the $(g-1)$-th ancestor

To compute the expected number of descendants of older sisters of the $(g-1)$-th ancestor of *ego*, we have to look at the g-th ancestor at times $t+\alpha$ before the present, for each $\alpha\geq 1$, to see how many older sisters were born then and how many $(q-1)$-th generation descendants each of these left. Using the same reasoning as before, we get:

$$\text{For } g\geq 1,\quad \mathbf{K}_{\mathrm{o}}(g,q) \;=\; \sum_{t\geq 1}\sum_{\alpha\geq 1}\mathbf{D}(t+\alpha,q-1)\,\mathbf{F}\,(\mathbf{P}_{\mathbf{S}}^{\top})^{\alpha}\,\mathbf{P}_{\mathbf{F}}^{\top}\,\mathbf{U}(t,g)^{\top}. \tag{7}$$

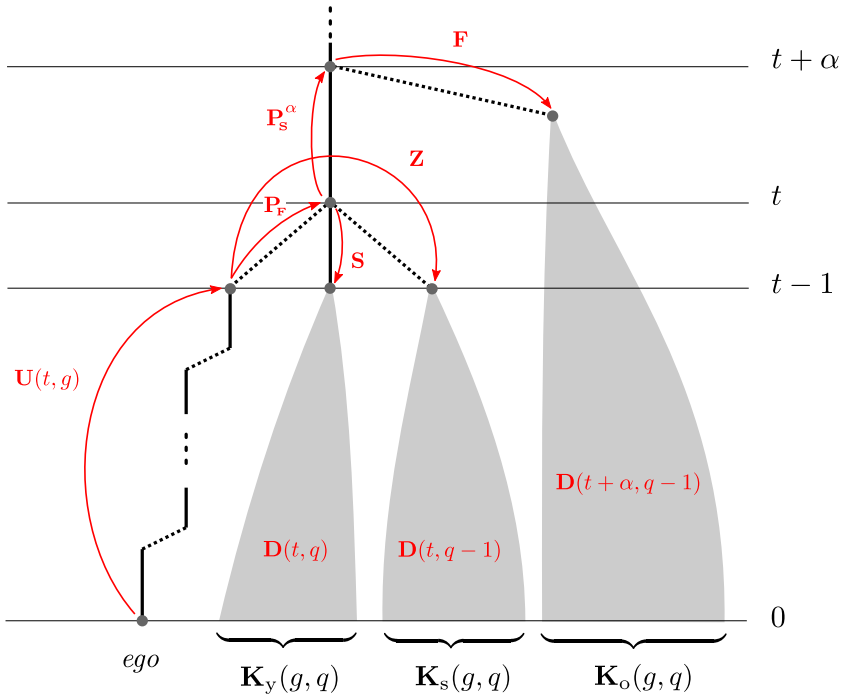


Figure 2: Genealogy of the population, showing the various contributions to the kinship matrix $\mathbf{K}(g,q)$. On the left is the lineage of *ego*, up to its g-th ancestor. The solid black lines correspond to survival and the dotted ones to reproduction. On the right are the q-th descendants of the g-th ancestor of *ego*, grouped according to whether they are descendants of younger, same-litter or older sisters of the $(g-1)-th$ ancestor of *ego*. The weights of the red arrows have to be multiplied along a path to get the corresponding contribution to the expected number of (g,q)-kin (matrix transposes have been omitted for simplicity).

Descendants of same-litter sisters of the $(g-1)$-th ancestor

Unlike for descendants of older and younger sisters, for which knowing about the g-th ancestor of *ego* was sufficient, here we also need to know about its $(g-1)$-th ancestor. This is because knowing that the g-th ancestor gave birth to the $(g-1)$-th one t time-steps before the present biases its number of offspring for this reproductive event. To understand this bias, we have to know the class in which the $(g-1)$-th ancestor was born. Indeed, if we sample an individual uniformly at random in class i and then look at its mother, then the mother is sampled proportionally to its number of offspring in class i. Let F_{ij} have the distribution of the number of offspring of class i of an individual of class j, and let $F_{ij}^{(\ell)}$ have the distribution of F_{ij} biased by $F_{\ell j}$, that is, $\mathbb{P}\big(F_{1j}^{(\ell)}=x_1,\dots,F_{Nj}^{(\ell)}=x_N\big)=\frac{x_\ell}{f_{\ell j}}\,\mathbb{P}\big(F_{1j}=x_1,\dots,F_{Nj}=x_N\big)$. Finally, let $\tilde{f}_{ij}^{(\ell)}=\mathbb{E}(F_{ij}^{(\ell)})-\mathbb{1}_{\{i=\ell\}}$ (where $\mathbb{1}_{\{i=\ell\}}=1$ if $i=\ell$ and 0 otherwise). With these quantities, the expected number of class-i (g,q)-kin of *ego* through same-litter sisters of its $(g-1)$-th ancestor is

$$\sum_{t\geq 1}\sum_{k,\ell,m} d_{im}(t,q-1)\,\tilde{f}_{mk}^{(\ell)}\,p_\mathbf{F}(\ell,k)\,u_{j\ell}(t,g) \;=\; \sum_{t\geq 0}\sum_{\ell,m} d_{im}(t,q-1)\,z_{m\ell}\,u_{j\ell}(t,g)\,,$$

where $z_{m\ell}=\sum_k \tilde{f}_{mk}^{(\ell)}\,p_\mathbf{F}(\ell,k)$. In matrix form, this is the (i,j)-th entry of

$$\mathbf{K}_\mathrm{s}(g,q) \;=\; \sum_{t\geq 1}\mathbf{D}(t,q-1)\,\mathbf{Z}\,\mathbf{U}(t,g)^\top. \tag{8}$$

What makes Equation (8) relevant is that the matrix $\mathbf{Z}$ has a concrete interpretation: as mentioned at the beginning of this section, z_{ij} is the expected number of same-litter sisters of a focal individual in class j that have just been born in class i. Thus, the matrix $\mathbf{Z}$ can be estimated from detailed census data. Alternatively, $\mathbf{Z}$ can be computed from the matrices $\mathbf{S}$ and $\mathbf{F}$ and the covariances $\mathrm{Cov}(F_{ij},F_{\ell j})$. Indeed, elementary calculations give

$$\tilde{f}_{ij}^{(\ell)} \;=\; \mathrm{Cov}(F_{ij},F_{\ell j})/f_{\ell j}+f_{\ell j}-\mathbb{1}_{\{i=\ell\}}, \tag{9}$$

from which $\mathbf{Z}$ can then be computed. A few special cases are worth pointing out. First, when at most one offspring is produced during a projection interval, we have $\mathbf{Z}=\mathbf{0}$.

Second, if there is a single newborn class, say class 1, then $z_{ij} = 0$ if $i \neq 1$ or $j \neq 1$, and

$$z_{11} \;=\; \sum_k \big(\mathrm{Var}(F_{1k})/f_{1k} + f_{1k} - 1\big) p_{\mathbf{F}}(1,k). \tag{10}$$

This covers the majority of the models used in practice. Third, when the numbers of offspring produced by an individual in the various classes are independent Poisson variables,

$$\mathbf{Z} \;=\; \mathbf{F}\mathbf{P}_{\mathbf{F}}^{\top} \tag{11}$$

This formula can be used by default when the covariances of the fertilities are not available.

Expression of the kinship matrix

Combining Equations (6), (7) and (8), we get the following expression for the kinship matrix $\mathbf{K} = \mathbf{K}_{\mathrm{y}} + \mathbf{K}_{\mathrm{o}} + \mathbf{K}_{\mathrm{s}}$, which is valid for all $q \geq 0$ but for $g \geq 1$ only:

$$\mathbf{K}(g,q) \;=\; \sum_{t\geq 1} \left(\mathbf{D}(t,q)\,\mathbf{S}\,\mathbf{P}_{\mathbf{F}}^{\top} \;+\; \mathbf{D}(t,q-1)\,\mathbf{Z} \;+\; \sum_{\alpha\geq 1} \mathbf{D}(t+\alpha,q-1)\,\mathbf{F}\,(\mathbf{P}_{\mathbf{S}}^{\top})^{\alpha}\,\mathbf{P}_{\mathbf{F}}^{\top} \right) \mathbf{U}(t,g)^{\top}.$$

For $g = 0$, the (g,q)-kin of *ego* are its q-th descendants (and *ego* itself when $q = 0$). Thus, $\mathbf{K}(0,0) = \mathbf{I}$ and, for $q > 0$,

$$\mathbf{K}(0,q) \;=\; \sum_{\alpha\geq 1} \mathbf{D}(\alpha,q-1)\,\mathbf{F}\,(\mathbf{P}_{\mathbf{S}}^{\top})^{\alpha}. \tag{12}$$

Note that, compared to Equation (7), there is no factor $\mathbf{P}_{\mathbf{F}}^{\top}$ to the right of $(\mathbf{P}_{\mathbf{S}}^{\top})^{\alpha}$.

In order to get a unique expression of $\mathbf{K}(g,q)$ that works for all $g \geq 0$, we let $\mathbf{U}(t,0) = \mathbf{0}$ for $t \geq 1$, in agreement with Equation (4) and the definition of weak compositions in Equation (3); but for $t = 0$ we use the special convention $\mathbf{U}(0,0) = \mathbf{I}$. Setting $\mathbf{Y}(g) = \mathbf{I}$ if $g = 0$ and $\mathbf{P}_{\mathbf{F}}^{\top}$ otherwise, this gives the following general expression for the kinship matrix, valid for any $(g,q) \neq (0,0)$:

$$\mathbf{K}(g,q) \;=\; \sum_{t\geq 0} \left(\mathbf{D}(t,q)\,\mathbf{S}\,\mathbf{P}_{\mathbf{F}}^{\top} \;+\; \mathbf{D}(t,q-1)\,\mathbf{Z} \;+\; \sum_{\alpha\geq 1} \mathbf{D}(t+\alpha,q-1)\,\mathbf{F}\,(\mathbf{P}_{\mathbf{S}}^{\top})^{\alpha}\,\mathbf{Y}(g) \right) \mathbf{U}(t,g)^{\top}. \tag{13}$$

Note that for $q = 0$, this formula reduces to

$$\mathbf{K}(g,0) \;=\; \sum_{t\geq 1} \mathbf{S}^{t}\,\mathbf{P}_{\mathbf{F}}^{\top}\mathbf{U}(t,g)^{\top} \tag{14}$$

and gives the probability that the g-th ancestor of *ego* is alive, structured by the class of *ego* and of the g-th ancestor.

2.5 Computing the kinship matrix in practice

Equation (13) gives the kinship matrix as an infinite sum. However, in numerical applications it will have to be approximated by a finite sum, for instance of the form

$$\begin{aligned} &\widehat{\mathbf{K}}(g,q\,;\,t_{\max},\alpha_{\max}) \\ &\quad= \sum_{t=0}^{t_{\max}} \left(\mathbf{D}(t,q)\,\mathbf{S}\,\mathbf{P}_{\mathbf{F}}^{\top} \;+\; \mathbf{D}(t,q-1)\,\mathbf{Z} \;+\; \sum_{\alpha=1}^{\alpha_{\max}} \mathbf{D}(t+\alpha,q-1)\,\mathbf{F}\,(\mathbf{P}_{\mathbf{S}}^{\top})^{\alpha}\,\mathbf{Y}(g) \right) \mathbf{U}(t,g)^{\top}. \end{aligned} \tag{15}$$

This raises the question of determining how many terms to sum, and of the precision of the resulting estimate. In SM.6, we give an explicit upper bound on $\|\mathbf{K}(g,q) - \widehat{\mathbf{K}}(g,q\,;\,t_{\max},\alpha_{\max})\|$ that can be used to chose $t_{\max}$ and $\alpha_{\max}$ so as to meet a target accuracy in the estimation. A procedure to do so automatically is implemented in our code (see SM.1). It should be pointed out, however, that the upper bound of SM.6 is rather crude and that the actual precision will typically be much better than it suggests.

Importantly, when there exists a maximal age ω in the model (that is, ω such that $\mathbf{S}^{\omega-1} \neq \mathbf{0}$ and $\mathbf{S}^{\omega} = \mathbf{0}$), the sums in Equation (13) are, in fact, finite: indeed, in that case, $\mathbf{U}(t,g) = \mathbf{0}$ for all $t > g\,\omega$; $\mathbf{D}(t,g) = \mathbf{0}$ for all $t > (q+1)\,\omega$; and $\mathbf{P}_{\mathbf{S}}^{\alpha} = \mathbf{0}$ for all $\alpha \geq \omega$. As a result,

$$\mathbf{K}(g,q) \;=\; \widehat{\mathbf{K}}(g,q\,;\,t_{\max},\alpha_{\max}), \qquad \text{with } \begin{cases} t_{\max} = \min\big(g\,\omega,\ (q+1)\,\omega\big) \\ \alpha_{\max} = \omega - 1. \end{cases} \tag{16}$$

This gives a useful heuristic to choose $t_{\max}$ and $\alpha_{\max}$ when the values obtained from the upper bound of SM.6 lead to prohibitively long calculations: first, find a reasonable estimate of the maximum age that one can expect to observe in the population, by picking $\hat{\omega}$ such that the probability that an individual reaches age $\hat{\omega}$ is less than some very small ε (e.g., by taking $\hat{\omega} = \inf\{t \geq 0 : \|\mathbf{S}^{t-1}\|_1 < \varepsilon\}$, where $\|\mathbf{X}\|_1 = \sup_j \sum_i |x_{ij}|$); then use the values of $t_{\max}$ and $\alpha_{\max}$ obtained by plugging $\hat{\omega}$ in Equation (16).

Although there is no mathematical guarantee – that is, other than that given in SM.6 – on the difference between $\mathbf{K}$ and $\widehat{\mathbf{K}}$ when using $t_{\max} = \min(g\,\hat{\omega}, (q+1)\,\hat{\omega})$ and $\alpha_{\max} = \hat{\omega} - 1$, this difference should be well within the modelling error of the projection model, as it will exclusively be due to individuals that live past $\hat{\omega}$ – which are allowed by the model but may not exist in practice. One way to test this is to compare the results obtained with the projection matrix $\mathbf{A}$ to those obtained with a model $\bar{\mathbf{A}}(\hat{\omega})$ where the maximal age has been capped (as discussed in SM.7). The biological relevance of $\bar{\mathbf{A}}(\hat{\omega})$ compared to that of $\mathbf{A}$ is something that can be assessed independently; and once this is done, the kinship matrix of $\bar{\mathbf{A}}(\hat{\omega})$ can be computed exactly using $t_{\max} = \min(g\,\hat{\omega}, (q+1)\,\hat{\omega})$ and $\alpha_{\max} = \hat{\omega} - 1$.

2.6 Aggregated measures of kinship and relatedness

Having a general expression for the number of any type of kin makes it possible to consider various aggregated measures of the kinship structure of a population. For instance, the matrix

$$\tilde{\mathbf{K}}(d) \;=\; \sum_{g+q=d} \mathbf{K}(g,q) \tag{17}$$

gives the number of kin that are at distance d from *ego* in their genealogical tree. If $\lambda > 1$, $\tilde{\mathbf{K}}(d)$ diverges as $d \to \infty$, whereas if $\lambda < 1$ it goes to zero. It is thus natural to wonder how $\|\tilde{\mathbf{K}}(d)\|$ grows/decreases as a function of d. In SM.9, we show that it grows like $\theta_{\mathrm{c}}^{\,d}$, where

$$\theta_{\mathrm{c}} \;=\; \inf\left\{\theta > 0 : \rho(\mathbf{S} + \theta^{-1}\mathbf{F})\,\rho(\mathbf{P_S} + \theta^{-1}\mathbf{P_F}) < 1\right\} \tag{18}$$

and where ρ denotes the spectral radius. This characterization makes it straightforward to compute θ_c numerically. In SM.9 we also explain why we expect to have

$$\theta_c \;\approx\; \frac{\lambda - s^\star}{\sqrt{\lambda} - s^\star} \quad \text{with } s^\star = \mathbf{vSw}\,, \tag{19}$$

where the vector of reproductive values $\mathbf{v}$ and the stable distribution $\mathbf{w}$ are such that $\mathbf{vw} = 1$.

One of the notable feature of θ_{c} is that it describes the global kinship structure, at a large-scale. This contrasts with the matrices $\tilde{\mathbf{K}}(d)$, which can only be computed for relatively small values of d (see SM.10) and thus only give us access to the local, small-scale structure of the kinship.

In applications, θ_{c} could be a useful descriptor for questions ranging from life history theory to inclusive fitness: for instance, Equation (19) shows that for a fixed value of $\lambda > 1$, the number of kin is maximized by increasing $s^\star$, i.e. by K-strategies; but that if $\lambda < 1$ it is optimized by decreasing $s^\star$, i.e. by r-strategies. The statistic θ_c could thus be used to compare life history patterns in terms of the r–K continuum, as an alternative to the generation time $T = 1/(\lambda - s^\star)$ (Bienvenu and Legendre, 2015). In the context of inclusive fitness, since the matrices $\mathbf{K}$ can be interpreted as measures of relatedness (see SM.11), θ_c is an indicator of how much relatedness is carried by various kin: for instance, the higher θ_{c}, the more relatedness is carried by close kin – but the smaller the relative contribution of close kin compared to that of distant ones.

Of course, $\tilde{\mathbf{K}}(d)$ and θ_{c} are only two examples of aggregated measures of kinship that can be defined from the matrices $\mathbf{K}$. In general, one may want to consider various statistics of the form $\sum_{g,q} f(\mathbf{K}(g,q), g, q)$, for functions f and values of (g,q) that depend on the specific application. If the application requires summing on large values of g and q, this may not be possible because of prohibitive computation times. In that case, one possible strategy could be to first turn the projection matrix into a 1×1 matrix projecting the unstructured population (see Bienvenu et al., 2017; Coste et al., 2017), and then use the efficiently computable formulas of $\mathbf{K}(g,q)$ for 1×1 models given in SM.8.

Box 2: Illustration

We illustrate our method on the most common model in animal ecology: the extended Leslie model, i.e. an age-structured model with a class that regroups all individuals older than a certain age. More complex examples are given in SM.2, to illustrate and discuss specific points; but once an appropriate projection matrix has been obtained the method is always the same. Recall that we need to consider a pre-breeding model – that is, in the case of a 3-class model, a projection matrix of the form

$$\mathbf{A} = \begin{pmatrix} s_0 m_1 & s_0 m_2 & s_0 m_3 \\ s_1 & 0 & 0 \\ 0 & s_2 & s_3 \end{pmatrix},$$

where classes 1 and 2 correspond to individuals that are about to turn 1 and 2 year(s) old, respectively, and class 3 corresponds to older individuals; s_0 is the probability that newborns survive their first year, s_i the probability that class-i individuals survive to the next year and m_i the average number of (female) newborns to which they give birth.

Using the life table for the ground squirrel *Spermophilus dauricus* from Luo and Fox (1990), we get the following survival and fertility matrices:

$$\mathbf{S} = \begin{pmatrix} 0 & 0 & 0 \\ 0.57 & 0 & 0 \\ 0 & 0.54 & 0.46 \end{pmatrix} \quad \text{and} \quad \mathbf{F} = \begin{pmatrix} 0.23 & 0.51 & 0.84 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The asymptotic growth rate is $\lambda = 1.00$ and the stable class distribution is $\mathbf{w} = (0.47, 0.27, 0.27)$. From this, Equation (2) gives

$$\mathbf{P_S} = \begin{pmatrix} 0 & 0 & 0 \\ 1.00 & 0 & 0 \\ 0 & 0.54 & 0.46 \end{pmatrix} \quad \text{and} \quad \mathbf{P_F} = \begin{pmatrix} 0.23 & 0.29 & 0.48 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

For instance, the (3,2)-th entry of $\mathbf{P_S}$ shows that, at the demographic steady state, 54% of class-3 individuals were in class 2 the previous year. Similarly, the (1, 1)-th entry of $\mathbf{P_F}$ indicates that a yearling has a 23% probability of having been born to a yearling the previous year. Since there is a single newborn class, we can use Equation (10) to compute the matrix $\mathbf{Z}$ under arbitrary assumptions on the fertilities. But here, in the absence of additional information, we will simply assume that they correspond to Poisson random variables, so that $\mathbf{Z} = \mathbf{F}\mathbf{P_F}^\top$. This yields

$$\mathbf{Z} = \begin{pmatrix} 0.60 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Let us now compute $\mathbf{K}(1,0)$, whose (i,j)-th entry is the probability that the mother of an individual of class j is currently alive and in class i. Using $t_{\max} = 9$ and $\alpha_{\max} = 0$, which as per SM.6 gives an approximation error $\|\widehat{\mathbf{K}} - \mathbf{K}\|_1 < 10^{-2}$, we get

$$\mathbf{K}(1,0) = \begin{pmatrix} 0 & 0 & 0 \\ 0.13 & 0 & 0 \\ 0.38 & 0.24 & 0.08 \end{pmatrix}.$$

This matrix indicates, for instance, that a yearling has a 13% probability that its mother is alive in class 2, and 38% that it is alive in class 3. Note that kin are structured by their *current* class. Thus, it is not possible to have yearling mothers here: even if the mother of *ego* was a yearling when *ego* was born, by now it will have moved to an adult class.

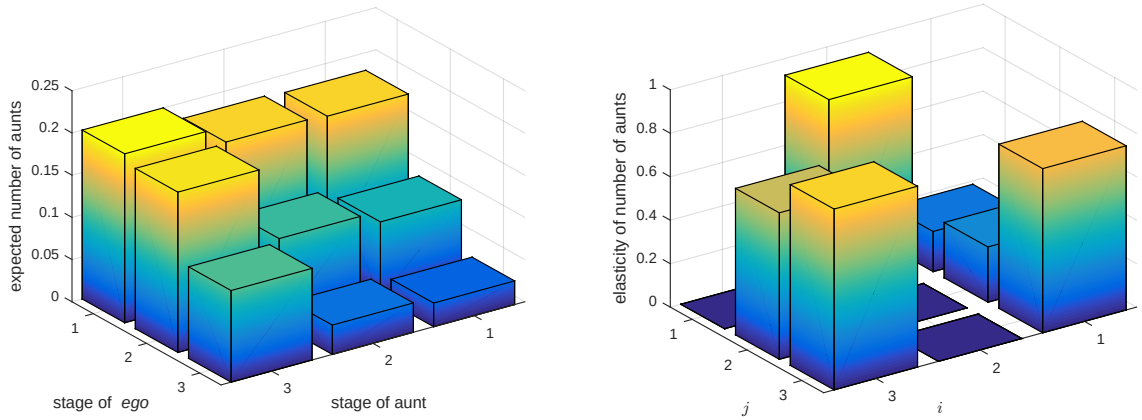
Figure 3a gives a graphical representation of $\mathbf{K}(2,1)$, the kinship matrix for aunts. The total number of aunts (irrespective of their class) of an individual sampled according to the stable class distribution is $\mathbf{1K}(2,1)\mathbf{w} = 0.36$, and numerically it is possible to compute its elasticities to the entries of the projection matrix, i.e. $\partial \ln \mathbf{1K}(2,1)\mathbf{w} / \partial \ln a_{ij}$. These are plotted in Figure 3b, which shows that the number of aunts is most affected by relative changes in the survival probabilities s_1 and s_3.

Finally, we get sensibly the same results when taking the values of $t_{\max}$ and $\alpha_{\max}$ suggested by Equation (16), with the maximal age $\hat{\omega} = 6$ reported in Luo and Fox (1990). This is also the case with other models that we studied (see SM.2), which tends to suggest our simple heuristic to chose $t_{\max}$ and $\alpha_{\max}$ based on a maximal age is adequate.

In SM.12, we compare the kinship matrices $\mathbf{K}(1,0)$ and $\mathbf{K}(2,1)$ to the output of individual-based simulations, confirming our formulas.

3 Discussion

In this article, we have developed a unified framework to compute the expected number of kin of a focal individual (as a function of the vital traits of the focal individual and of its kin), directly from the demographic rates of a population. The main feature of our method is that it applies to any kin relationship (mother-daughter, aunt-niece, cousin-cousin, etc) and any structured population described by a matrix population model – be it structured by age (as illustrated in Box 2), by stage (SM.2.1), by patch (SM.2.3), or by any conceivable type of discrete class. These theoretical results come with readily usable implementations that make them available to any ecologist. Here we discuss the main limitations and implications of this unified framework, and give some perspectives for future research.



(a) (b)

Figure 3: (a) graphical representation of the matrix $\mathbf{K}(2,1)$, whose entries give the expected number of aunts of *ego* as a function of its class and that of its aunts; (b) the elasticities of the expected total number of aunts of *ego* (sampled at random in the stable population), $\partial \ln \mathbf{1K}(2,1)\mathbf{w}/\partial \ln a_{ij}$.

Conceptual and practical limitations

Our work is set in the framework of matrix population models, and thus inherits its usual limitations. In particular, density-dependence and environmental stochasticity – which are not considered by our input model – are likely to be two of the main sources of discrepancies between our theoretical predictions and reality. However, because our mathematical reasoning is based directly on the stochastic model underlying the matrix population model (rather than on its deterministic description of the average population dynamics), demographic stochasticity is not a limitation: it is taken into account in our calculations, and although our final formulas correspond to expected values with respect to this stochasticity, our methods could also be used to compute, e.g, variances. This is discussed below.

Another important limitation of our approach is that it hinges on the assumption that the dynamics of the population has been at the steady state for an extended period of time. This limits the validity of our formulas in populations whose demography has been perturbed in the recent past – all the more so when studying distant kin relationships. Unfortunately, there does not seem to be an easy way to circumvent this using our techniques, and there may not be alternatives to recursive formulas for such models (Caswell and Song, 2021).

When applying our results, one should also pay attention to the fact that they are formulated for models where the fertilities are independent from adult survival. This assumption is required for the connection with the stochastic population model that implicitly underlies our calculations, but it is not compatible with post-breeding models. However, these can (usually easily) be reformulated as pre-breeding models, so this is not a major limitation of our work.

As implicit with matrix population models, we work in the context of monoparental genealogies. This setting could correspond either to a one-sex species, or to a two-sex species with a female-based demographic model. In the latter case, it should be noted that our formulas are based on a notion of kinship that only considers matrilines, and will thus differ from the traditional notion of kinship in a complete pedigree: for instance, the daughters of the brothers of the focal individual will not be counted as its nieces. Although this is not a limitation of our results, this is something important to have in mind when interpreting them. Incidentally, this specific notion of kinship can in some contexts be interpreted as a direct measure of relatedness. Indeed, as explained in SM.11, in an idealized setting (namely, a large panmictic population where males and females have similar life-histories and where no two individuals have the same two parents), discarding kin that are not based on matrilines has the same effect as the dilution of genetic material over successive generations. Thus, although our results were first and foremost developed to study kinship, they could have wider-ranging implications in evolutionary biology, by working as a measure of relatedness.

Finally, a practical limitation of our results is that the computing time of our formulas increases rapidly when we consider increasingly distant kin, and can become prohibitively long – especially for organisms with a "slow" strategy on the slow-fast continuum. Because close kin are what matter in most situations, this is, in general, hardly going to be a problem. However, should very distant kin be of interest, we provide two partial solutions to prohibitive computing times: first, in Section 2.6 we use our results to

derive an efficiently computable statistic to describe the large-scale kinship structure of a population; second, if one is ready to discard the structure of the population, then in SM.8 we give a simple formula for the expected number of kin that is trivial to compute in practice. Note that both of these results are made possible by the fact that we have a generic formula for the kinship, and would presumably not have been possible using ad hoc formulas for each type of kin.

Implications and perspectives

First and foremost, having a generic formula for arbitrary kin relationship makes it possible – and elementary – to define tailor-made measures of kinship to be used in a specific setting. For instance, anthropologists who would like to compare the number of lineal vs collateral kin can easily do so by summing the kinship matrices of their choice – or, more generally, combine them however they want to produce the best suited metric for their particular application.

Our methods can be used to compute, at least numerically, the sensitivity or the elasticity of the kinship structure to the traits implemented in the model. This is illustrated on a concrete example in Box 2. Because our formulas apply to any primitive projection matrix, they can be used to study models where classes are based on several traits, such as metapopulation models (Lebreton and Gonzáles-Dávila, 1993) or more general multitrait models (Roth and Caswell, 2016; Coste et al., 2017; Coste and Pavard, 2020). This makes it possible to study and compare the relative effects of different traits on the kinship. We illustrate this in SM.2.3, where we use a *stage* × *patch* model to study the effect of dispersal.

Finally, because our results make new use of classic tools to solve a long-standing problem, they could spur further theoretical developments in the field of matrix population models. An example of this is given by our introduction, in SM.5, of the previous-generation matrix (which is a dual version of the classic next-generation matrix of Cushing and Zhou, 1994) and of the Euler–Lotka matrices (which give two simple generalizations of the Euler–Lotka equation to arbitrarily structured populations). Despite being very natural from a mathematical and biological point of view, these quantities do not seem to have been identified as such before (although similar ideas can be found in Lebreton, 2005). Though not directly linked to kinship, these results are straightforward consequences of the framework we have developed to study it.

Our work offers several perspectives in terms of extensions, in particular to see if our methods can be extended beyond the framework of matrix population models. A natural candidate for this are integral projection models (Rees et al., 2014). Because the variations in the individual realizations of the vital rates have a relevant impact on kin structure (Coresh and Goldman, 1988; Tuljapurkar et al., 2020, 2021), another perspective is to go the beyond the "expected-value" description of the kinship structure given here – namely by deriving statistics that describe the demographic stochasticity underlying our calculations, instead of averaging it out. For instance, with our method, the classic results of Everett and Ulam (1948) on multitype Galton–Watson processes could be used to compute the variance of the number of kin. Finally, an important line of research that we plan to explore is the extension of our work to two-sex models. This is particularly challenging, as it requires leaving the framework of monoparental genealogies and will therefore have to be done using a combination of analytical and numerical methods.

Acknowledgments

This study is supported by a grant from the Agence Nationale de la Recherche (ANR-18-CE02-0011, MathKinD). CFDC was funded by Centre of Excellence grant from the Research Council of Norway (SFF-III, project no. 223257). FB was partly funded by Agence Nationale de la Recherche (ANR-16-CE27-0013), EPSRC Fellowship EP/N004833/1 and a Transilvania Fellowship from the Transilvania University of Brașov. We thank Arpat Ozgul for sharing data on the life cycle of the yellow-bellied marmot (Marmota flaviventris).

References

Atkins, J. R. (1974). On the fundamental consanguineal numbers and their structural basis. *American Ethnologist*, 1(1):1–31.

Barrai, I., Cavalli-Sforza, L., and Moroni, A. (1962). Frequencies of pedigrees of consanguineous marriages and mating structure of the population. *Annals of Human Genetics*, 25(4):347–377.

Bauduin, S., McIntire, E. J. B., and Chubaty, A. M. (2019). NetLogoR: a package to build and run spatially explicit agent-based models in R. *Ecography*, 42(11):1841–1849.

Bienvenu, F., Akçay, E., Legendre, S., and McCandlish, D. M. (2017). The genealogical decomposition of a matrix population model with applications to the aggregation of stages. *Theoretical Population Biology*, 115:69–80.

Bienvenu, F. and Legendre, S. (2015). A new approach to the generation time in matrix population models. *The American Naturalist*, 185(6):834–843.

Brault, S. and Caswell, H. (1993). Pod-specific demography of killer whales (Orcinus orca). *Ecology*, 74(5):1444–1454.

Caswell, H. (2001). *Matrix Population Models*. Sinauer Associates Inc., Sunderland, MA.

Caswell, H. (2019). The formal demography of kinship: A matrix formulation. *Demographic Research*, 41:679–712.

Caswell, H. (2020). The formal demography of kinship II: Multistate models, parity, and sibship. *Demographic Research*, 42:1097–1146.

Caswell, H. and Song, X. (2021). The formal demography of kinship iii: Kinship dynamics with time-varying demographic rates. *bioRxiv*.

Coresh, J. and Goldman, N. (1988). The effect of variability in the fertility schedule on numbers of kin. *Mathematical Population Studies*, 1(2):137–156.

Coste, C. F. D., Austerlitz, F., and Pavard, S. (2017). Trait level analysis of multitrait population projection matrices. *Theoretical Population Biology*, 116:47–58.

Coste, C. F. D. and 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.

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

Demetrius, L. (1974). Demographic parameters and natural selection. *Proceedings of the National Academy of Sciences*, 71(12):4645–4647.

Demetrius, L. (1975). Natural selection and age structured populations. *Genetics*, 79(3):535–544.

Everett, C. J. and Ulam, S. (1948). Multiplicative systems, I. *Proceedings of the National Academy of Sciences*, 34(8):403–405.

Goodman, L. A., Keyfitz, N., and Pullum, T. W. (1974). Family formation and the frequency of various kinship relationships. *Theoretical Population Biology*, 5(1):1–27.

Hajnal, J. (1963). Concepts of random mating and the frequency of consanguineous marriages. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 159(974):125–177.

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

Kemeny, J. G. and Snell, J. L. (1976). *Finite Markov chains*. Springer-Verlag, New York, 2nd edition.

Knuth, D. E. (2011). *Art of Computer Programming, Volume 4A*. Addison-Wesley Professional.

Le Bras, H. (1973). Parents, grands-parents bisaieux. *Population (French Edition)*, 1(1):9–38.

Lebreton, J.-D. (1990). Modelling density dependence, environmental variability, and demographic stochasticity from population counts: An example using Wytham Wood great tits. In *Population Biology of Passerine Birds*, pages 89–102. Springer, Berlin, Heidelberg.

Lebreton, J. D. (2005). Age, stages, and the role of generation time in matrix models. *Ecological Modelling*, 188(1):22–29.

Lebreton, J.-D. and Gonzáles-Dávila, G. (1993). An introduction to models of subdivided populations. *Journal of Biological Systems*, 01(04):389–423.

Lotka, A. J. (1939). *Théorie analytique des associations biologiques*. Hermann et Cie, Paris.

Luo, J. and Fox, B. J. (1990). Life-table comparisons between two ground squirrels. *Journal of Mammalogy*, 71(3):364–370.

Meyer, C. D. (2000). *Matrix analysis and applied linear algebra*, volume 71. SIAM.

Nerman, O. and Jagers, P. (1984). The stable doubly infinite pedigree process of supercritical branching populations. *Zeitschrift fuer Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 65(3):445–460.

Ozgul, A., Oli, M. K., Armitage, K. B., Blumstein, D. T., and Van Vuren, D. H. (2009). Influence of local demography on asymptotic and transient dynamics of a yellow-bellied marmot metapopulation. *American Naturalist*, 173(4):517–530.

Pavard, S. and Coste, C. F. D. (2020). Goodman, Keyfitz & Pullum (1974) and the population frequencies of kinship relationships. *Theoretical Population Biology*, 133:15–16.

Pullum, T. W. (1982). The eventual frequencies of kin in a stable population. *Demography*, 19(4):549–565.

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

Roth, G. and Caswell, H. (2016). Hyperstate matrix models: extending demographic state spaces to higher dimensions. *Methods in Ecology and Evolution*, 7(12):1438–1450.

Tuljapurkar, S. (1993). Entropy and convergence in dynamics and demography. *Journal of Mathematical Biology*, 31(3):253–271.

Tuljapurkar, S., Zuo, W., Coulson, T., Horvitz, C., and Gaillard, J. M. (2020). Skewed distributions of lifetime reproductive success: beyond mean and variance. *Ecology Letters*, 23(4):ele.13467.

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

Tuljapurkar, S. D. (1982). Population Dynamics in Variable Environments. IV. Weak ergodicity in the Lotka Equation. *Journal of Mathematical Biology*, 14:221–230.

Varga, R. S. (2000). *Matrix Iterative Analysis*, volume 27 of *Springer Series in Computational Mathematics*. Springer, Berlin, Heidelberg.

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

Young, K. E. and Keith, E. O. (2011). A comparative analysis of cetacean vital rates using matrix population modeling analysis of cetacean vital rates. *International Journal of Applied Science and Technology*, 1(6):261–277.

Supplementary Materials

SM.1 Python and MATLAB implementations of the formulas . . . 15
SM.2 Additional illustrations . . . 15
SM.3 Reducible projection matrices . . . 21
SM.4 Generating weak compositions . . . 22
SM.5 The previous-generation and Euler–Lotka matrices . . . 24
SM.6 Upper bound on $\|\mathbf{K}(g,q) - \widehat{\mathbf{K}}(g,q\,;t_{\max},\alpha_{\max})\|$ . . . 26
SM.7 Capping age in a model . . . 29
SM.8 Expression of $\mathbf{K}(g,q)$ for 1×1 models . . . 30
SM.9 Asymptotic behavior of $\tilde{\mathbf{K}}(d)$ as $d\to\infty$ . . . 30
SM.10 Computing time of $\widehat{\mathbf{K}}(g,q\,;\,t_{\max},\alpha_{\max})$ . . . 33
SM.11 Matrilineal kinship and relatedness . . . 35
SM.12 Comparison to individual-based simulations for the ground squirrel *Spermophilus dauricus* 36

SM.1 Python and MATLAB implementations of the formulas

We provide two implementations of the methods presented in this paper:

- A MATLAB implementation that can be used to reproduce all numerical calculations and figures presented in this manuscript. Using it requires knowing MATLAB.
- A Python implementation with a flexible, general-purpose interface. Python is a free, open-source programming language that is available on every platform. This implementation provides a well-documented library that can be used in Python scripts, but it also comes with a graphical user interface that requires no prior knowledge of any programming language to be used (see Figure 4).

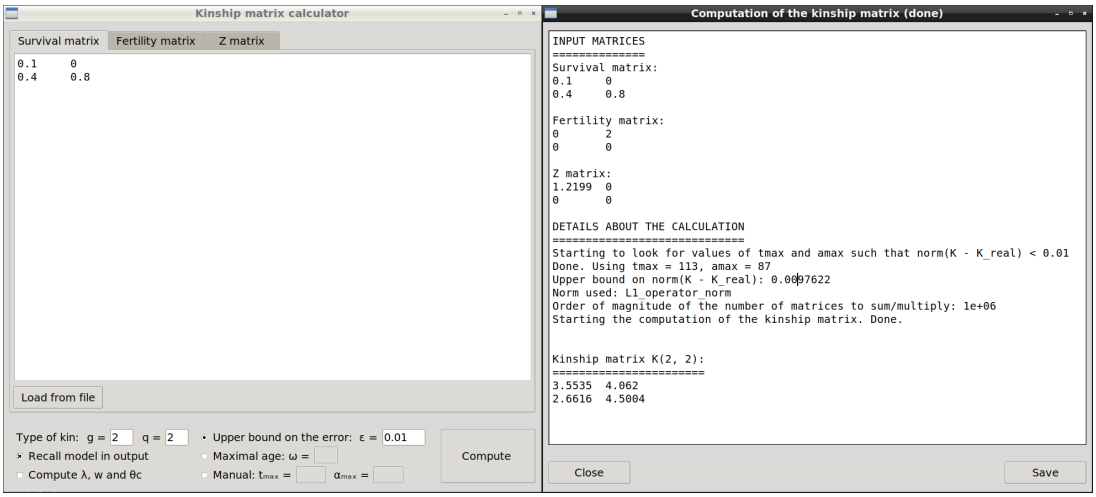


Figure 4: Graphical user interface provided with the Python implementation of the formulas.

Both implementations can be downloaded with detailed instructions from `10.5281/zenodo.4680715`.

SM.2 Additional illustrations

In Box 2, we have illustrated our method using an extended Leslie model. This choice is natural, because these are the most widely used models in animal ecology. However, this does not illustrate the fact that our method works for any population structure. This also did not give us the occasion to address relevant questions that can arise in some settings. Thus, in this section we give the following additional illustrations:

- The classic life cycle of the killer whale (*Orcinus orca*) from Brault and Caswell (1993). This is a fairly extreme example of how a "slow" strategy (high survival, low reproduction) can make the calculations computationally intensive. It is also used to discuss other questions (birth-flow and post-reproductive classes).

- The complex life cycle of the yellow-bellied marmot (*Marmota flaviventris*) from Ozgul et al. (2009). This is an example of a post-breeding model whose conversion to a pre-breeding model is non-trivial.
- A generic metapopulation model that illustrates how our methods can be used to study the effects of traits such as dispersal on kinship.

SM.2.1 A slow strategy: the killer whale *Orcinus orca*

Consider the following classic projection matrix for the killer whale *Orcinus orca*, taken from Brault and Caswell (1993):

$$\mathbf{A} = \begin{pmatrix} 0 & 0.0043 & 0.1132 & 0 \\ 0.9775 & 0.9111 & 0 & 0 \\ 0 & 0.0736 & 0.9534 & 0 \\ 0 & 0 & 0.0452 & 0.9804 \end{pmatrix}.$$

This is a so-called Lefkovitch matrix with four stages (newborn, juvenile, reproductive adult and post-reproductive adult).

Birth-flow

This model is what is known as a *birth-flow* model: since orcas do not have a fixed breeding season, individuals can be born at any time of the year. As a result of this, the fertilities (first row of the projection matrix) incorporate a part of the survival of the mother; and thus the number of offspring that an individual produces between time t and $t+1$ *cannot* rigorously be independent from its survival during that period.

However, one should keep in mind that, even in pre-breeding models, the independence of survival and reproduction is never perfectly met in practice because there is always a trade-off between survival and reproduction. Thus, just as any of the many working hypotheses of matrix population models, this independence is an "ideal" mathematical assumption that is required for the calculations to be exact and whose relevance will depend on the biological situation at hand. In a post-breeding model, it is typically not acceptable (especially given that it is almost always possible to convert such a model into a more suitable pre-breeding one); but in a birth-flow model, it can be – all the more so when survival is high and fertility is low, which tends to decorrelate them. The intuition for this is the following: when survival is high and reproduction is low, the death of pregnant mothers (which is the source of the non-independence that is problematic to us) is rare and can be neglected.

As a more formal argument, let S_i be the Bernoulli variable corresponding to the survival of an individual from class i from one year to the next (thus, letting $p_i = \mathbb{P}(S_i = 1)$, here $p_2 = 0.9847$ and $p_3 = 0.9986$), and let F_i be the number of offspring it leaves in the next year. Since orcas produce at most one offspring per year, F_i is a Bernoulli variable with parameter f_i (here with $f_2 = 0.0043$ and $f_3 = 0.1132$). Assuming that birth and death occur at constant rate (with at most one birth), standard calculations yield that $\mathbb{E}(F_i S_i) = 2fp/(1+p)$. Thus, the correlation coefficient between S_i and F_i is

$$\mathrm{corr}(S_i,\, F_i) \;=\; \frac{\mathbb{E}(S_i F_i) \,-\, \mathbb{E}(S_i)\mathbb{E}(F_i)}{\sqrt{\mathrm{Var}(S_i)\,\mathrm{Var}(F_i)}} \;=\; \sqrt{\frac{f_i\, p_i\, (1-p_i)}{(1+p_i)^2(1-f_i)}}\,.$$

Numerically, this gives $\mathrm{corr}(S_2, F_2) = 0.0041$ and $\mathrm{corr}(S_3, F_3) = 0.0067$, which tends to confirm our intuition about the low dependence between S_i and F_i. As a result, for our purposes the pre-breeding model gives a decent approximation of the birth-flow one and can be used to compute $\mathbf{K}$.

Reducibility

Here, the projection matrix $\mathbf{A}$ is *not* primitive, because of the post-reproductive class. However, it still has a dominant eigenvalue $\lambda = 1.0254$ with a unique associated right-eigenvector, namely $\mathbf{w} = (0.0370, 0.3161, 0.3229, 0.3240)$. Since $\mathbf{w} > 0$, as explained in SM.3 it is still possible to use Equation (2) to compute the genealogical matrices, yielding

$$\mathbf{P_S} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0.1115 & 0.8885 & 0 & 0 \\ 0 & 0.0703 & 0.9297 & 0 \\ 0 & 0 & 0.0439 & 0.9561 \end{pmatrix} \quad \text{and} \quad \mathbf{P_F} = \begin{pmatrix} 0 & 0.0358 & 0.9642 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Estimation of the kinship matrices

First, we need to compute the same-litter newborn sisters matrix $\mathbf{Z}$ (see Section 2.4). Since orcas give birth to at most one offspring a year, $\mathbf{Z} = \mathbf{0}$.

Next, we need to choose appropriate values of $t_{\max}$ and $\alpha_{\max}$. Here however, because of the high survival and low reproduction, the values of $(t_{\max}, \alpha_{\max})$ suggested by the upper bound of SM.6 are not practical: indeed, for the ℓ_1 operator norm $\|\mathbf{S}\|_1 \, \|\mathbf{P}_{\mathbf{s}}^{\top}\|_1 = 0.9986$, so that we would need to take, e.g., $t_{\max} = 9630$ and $\alpha_{\max} = 9523$ to compute $\widehat{\mathbf{K}}(1,0)$ with the guarantee that $\|\widehat{\mathbf{K}}(1,0) - \mathbf{K}(1,0)\|_1 < 10^{-3}$. This leads to prohibitively long calculations. For most other classic norms, $\|\mathbf{S}\| \, \|\mathbf{P}_{\mathbf{s}}^{\top}\| > 1$ so our upper bound on the error does not even converge.

Therefore, we use $t_{\max} = \min\big(g\,\hat{\omega},\, (q+1)\,\hat{\omega}\big)$ and $\alpha_{\max} = \hat{\omega} - 1$, as suggested in Section 2.5, with the maximum age $\hat{\omega} = 90$ reported by Brault and Caswell (1993). With this, we get the following matrices for mothers and grand-mothers, respectively

$$\widehat{\mathbf{K}}(1,0) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0.0327 & 0.0174 & 0.0073 & 0.0023 \\ 0.9219 & 0.6501 & 0.3885 & 0.1855 \\ 0.0436 & 0.2755 & 0.4215 & 0.4175 \end{pmatrix} \quad \text{and} \quad \widehat{\mathbf{K}}(2,0) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0.0070 & 0.0037 & 0.0016 & 0.0005 \\ 0.3799 & 0.2660 & 0.1577 & 0.0739 \\ 0.4261 & 0.4575 & 0.4325 & 0.3215 \end{pmatrix}.$$

For aunts and cousins, we get

$$\widehat{\mathbf{K}}(2,1) = \begin{pmatrix} 0.0451 & 0.0316 & 0.0188 & 0.0089 \\ 0.5988 & 0.4891 & 0.3304 & 0.1736 \\ 0.7629 & 0.7717 & 0.6764 & 0.4703 \\ 0.6817 & 0.8483 & 1.0094 & 1.0384 \end{pmatrix} \quad \text{and} \quad \widehat{\mathbf{K}}(2,2) = \begin{pmatrix} 0.0879 & 0.0894 & 0.0788 & 0.0550 \\ 0.7643 & 0.8655 & 0.8708 & 0.6955 \\ 0.6880 & 0.8895 & 1.0886 & 1.1000 \\ 0.4813 & 0.7121 & 1.1032 & 1.5474 \end{pmatrix}.$$

In the absence of useful upper bound on the error $\|\widehat{\mathbf{K}} - \mathbf{K}\|$, to get an idea of the reliability of these estimates we can proceed as suggested in Section 2.5 and compare them with the matrices obtained from the "capped" *age* $\times$ *stage* model $\bar{\mathbf{A}}(90)$ described in SM.7. Indeed, in the capped model the corresponding kinship matrices $\bar{\mathbf{K}}$ can be computed exactly using finite sums.

Here, the capped matrix $\bar{\mathbf{A}}(90)$ is a 265×265 matrix. Whether it gives an appropriate description of the population can be assessed by comparing its descriptors to those of $\mathbf{A}$ (for instance, $\bar{\lambda} = 1.0253$ vs $\lambda = 1.0254$ or $\bar{R}_0 = 1.9546$ vs $R_0 = 2.0131$, etc). One can also "fold" it over age to recover a model with the same population structure as $\mathbf{A}$ (see Bienvenu et al., 2017; Coste et al., 2017). This gives

$$\text{folded}(\bar{\mathbf{A}}(90)) \;=\; \begin{pmatrix} 0 & 0.0043 & 0.1132 & 0 \\ 0.9775 & 0.9111 & 0 & 0 \\ 0 & 0.0736 & 0.9531 & 0 \\ 0 & 0 & 0.0452 & 0.9770 \end{pmatrix}.$$

As expected, the survivals are slightly lower than in $\mathbf{A}$, in particular for post-reproductive adults (0.9770 vs 0.9804). But, overall, the comparison suggests that here the initial model $\mathbf{A}$ and the capped model $\bar{\mathbf{A}}(90)$ describe fairly similar population processes.

Comparing the kinship matrices $\widehat{\mathbf{K}}$ corresponding to $\mathbf{A}$ with the kinship matrices $\bar{\mathbf{K}}$ corresponding to $\bar{\mathbf{A}}(90)$ shows that neglecting individuals over 90 year old has a slight but acceptable impact on the estimation; compare for instance the total number of mothers in $\mathbf{A}$ vs in $\bar{\mathbf{A}}(90)$, i.e. $\mathbf{1}\widehat{\mathbf{K}}(1,0)\mathbf{w} = 0.7950$ vs $\mathbf{1}\bar{\mathbf{K}}(1,0)\bar{\mathbf{w}} = 0.7830$; of grand-mothers (0.5793 vs 0.5232); of aunts (1.9590 vs 1.9005); of cousins (2.9983 vs 2.9390), etc.

Comparison with another cetacean

The kinship matrices can be used to quantitatively compare the familial structure of orcas with that of other species. For instance, for the North Atlantic right whale (*Eubalaena glacialis*), using the projection matrix from Young and Keith (2011),

$$\mathbf{A}' \;=\; \begin{pmatrix} 0 & 0 & 0 & 0.3350 \\ 0.9200 & 0.8600 & 0 & 0 \\ 0 & 0.0800 & 0.8000 & 0.8300 \\ 0 & 0.0200 & 0.1900 & 0 \end{pmatrix},$$

we get the following kinship matrices for aunts and cousins:

$$\widehat{\mathbf{K}}'(2,1) = \begin{pmatrix} 0.0182 & 0.0144 & 0.0070 & 0.0071 \\ 0.1381 & 0.1165 & 0.0605 & 0.0609 \\ 0.3218 & 0.3114 & 0.2255 & 0.2262 \\ 0.0640 & 0.0618 & 0.0447 & 0.0448 \end{pmatrix} \quad \text{and} \quad \widehat{\mathbf{K}}'(2,2) = \begin{pmatrix} 0.0214 & 0.0208 & 0.0152 & 0.0152 \\ 0.1367 & 0.1371 & 0.1076 & 0.1079 \\ 0.2267 & 0.2450 & 0.2464 & 0.2465 \\ 0.0452 & 0.0488 & 0.0490 & 0.0490 \end{pmatrix}.$$

Despite similar growth rates for the two populations ($\lambda = 1.0254$ vs $\lambda' = 1.0000$), we see that killer whales have far more kin than North Atlantic right whales: compare for instance the total number of aunts ($\mathbf{1}\widehat{\mathbf{K}}(2,1)\mathbf{w} = 1.9590$ vs $\mathbf{1}\widehat{\mathbf{K}}'(2,1)\mathbf{w}' = 0.3886$) or of cousins (2.9983 vs 0.4273) of *ego* in both models. This general trend is confirmed by the fact that $\theta_{\mathrm{c}} = 1.3778$ and $\theta'_{\mathrm{c}} = 0.9996$ (cf Section 2.6).

This is unsurprising and can be related to the very different social structures of both species: indeed, killer whales live in pods with their kin and hunt cooperatively with them, whereas North Atlantic right whales are more solitary. Thus, there likely is a greater selective pressure for killer whales to have more kin than for North Atlantic right whales.

SM.2.2 A complex life cycle: the yellow-bellied marmot *Marmota flaviventris*

Consider the following life cycle for the marmot *Marmota flaviventris* from Ozgul et al. (2009):

$$\mathbf{A}_{\mathrm{post}} = \begin{array}{c} \begin{array}{cccc} \mathrm{j} & \mathrm{y} & \mathrm{p} & \mathrm{r} \end{array} \\ \begin{pmatrix} 0 & s_{\mathrm{y}}\psi_{\mathrm{y}}m_2 & s_{\mathrm{p}}\psi_{\mathrm{p}}m_3 & s_{\mathrm{r}}\psi_{\mathrm{r}}m_3 \\ s_{\mathrm{j}} & 0 & 0 & 0 \\ 0 & s_{\mathrm{y}}(1-\psi_{\mathrm{y}}) & s_{\mathrm{p}}(1-\psi_{\mathrm{p}}) & 0 \\ 0 & s_{\mathrm{y}}\psi_{\mathrm{y}} & s_{\mathrm{p}}\psi_{\mathrm{p}} & s_{\mathrm{r}} \end{pmatrix} \end{array},$$

where the four classes are juveniles, yearlings, pre-reproductive adults and reproductive adults; s_x is the probability that an individual of class x survives the year, ψ_x the probability that it reproduces at the end of it, and m_2 (resp. m_3) litter size of a 2 year-old (resp. ≥ 3 year-old). This models is a good example of the non-conventional (that is, neither Leslie nor strictly Lefkovitch) models that are increasingly used in animal ecology. As the subscript suggests, it is a post-breeding model. The non-independence between reproduction and survival is made apparent by the fact that factors such as s_x or ψ_x in the fertility of class x also appear in the third and/or fourth row of column x.

The reason why we present this model is that the conversion to a pre-breeding model is non-trivial and may at first glance seem to be an obstacle to the calculation the kinship matrices with our method. However, this is not the case: all we need to do is split the "reproductive adults" class into three subclasses so as to distinguish between reproductive adults based on whether they were yearlings, pre-reproductive adults or reproductive adults the previous year. Once this is done, the corresponding pre-breeding model is easily seen to be

$$\mathbf{A}_{\mathrm{pre}} = \begin{array}{c} \begin{array}{ccccc} \mathrm{y} & \mathrm{p} & \mathrm{r_j} & \mathrm{r_p} & \mathrm{r_r} \end{array} \\ \begin{pmatrix} 0 & 0 & s_{\mathrm{j}}m_2 & s_{\mathrm{j}}m_3 & s_{\mathrm{j}}\psi_{\mathrm{r}}m_3 \\ s_{\mathrm{y}}(1-\psi_{\mathrm{y}}) & s_{\mathrm{p}}(1-\psi_{\mathrm{p}}) & 0 & 0 & 0 \\ s_{\mathrm{y}}\psi_{\mathrm{y}} & 0 & 0 & 0 & 0 \\ 0 & s_{\mathrm{p}}\psi_{\mathrm{p}} & 0 & 0 & 0 \\ 0 & 0 & s_{\mathrm{r}} & s_{\mathrm{r}} & s_{\mathrm{r}} \end{pmatrix} \end{array}.$$

Here the classes are yearling, pre-reproductive adults, reproductive adults that were yearling the previous year, reproductive adults that were pre-reproductive adults the previous year, and reproductive adults that were already reproductive adults the previous year.

Once the projection matrix $\mathbf{A}_{\mathrm{pre}}$ has been obtained, the kinship matrices can be computed exactly as in Box 2 or in SM.2, without regard to the unusual form of the life cycle. Using the numerical values from Ozgul et al. (2009), we have the survival and fertility matrices

$$\mathbf{S} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0.5328 & 0.3434 & 0 & 0 & 0 \\ 0.1872 & 0 & 0 & 0 & 0 \\ 0 & 0.2436 & 0 & 0 & 0 \\ 0 & 0 & 0.7200 & 0.7200 & 0.7200 \end{pmatrix} \quad \text{and} \quad \mathbf{F} = \begin{pmatrix} 0 & 0 & 1.0412 & 1.1508 & 0.6997 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Assuming Poisson reproduction, we get – for instance – the following kinship matrix for sisters:

$$\mathbf{K}(1,1) = \begin{pmatrix} 0.8583 & 0.4413 & 0.5038 & 0.3177 & 0.1770 \\ 0.3463 & 0.5218 & 0.5762 & 0.4143 & 0.2390 \\ 0.0922 & 0.1345 & 0.1607 & 0.0826 & 0.0460 \\ 0.0594 & 0.0987 & 0.0844 & 0.1271 & 0.0766 \\ 0.1560 & 0.2684 & 0.2215 & 0.3611 & 0.3505 \end{pmatrix}.$$

If we want a coarser description of the population and would like to get the number of kin structured by the original classes of the models, if suffices to aggregate the classes $\mathrm{r_y}$, $\mathrm{r_p}$ and $\mathrm{r_r}$, i.e. to sum the corresponding entries after weighting each column by its asymptotic abundance. This yields:

$$\mathbf{K}_{\text{aggr.}}(1,1) \;=\; \begin{pmatrix} 0.8583 & 0.4413 & 0.2635 \\ 0.3463 & 0.5218 & 0.3403 \\ 0.3076 & 0.5016 & 0.5115 \end{pmatrix}.$$

Note that it is not possible to recover the number of kin of juveniles nor the number of juvenile kin, because these are not counted by the model.

SM.2.3 A metapopulation model: studying the impact of dispersal on kinship

Since our formulas apply to a very large class of projection models (see SM.3), they can be used to study models that are based on combinations of traits (referred to as multitrait models in Coste et al. (2017) and hyperstate models in Roth and Caswell (2016)), in particular metapopulation models (sensu Lebreton, 1990).

The model

Let us consider the following metapopulation model with two stages (juveniles and adults) and two patches (A and B), where juveniles can disperse:

$$\mathbf{M} \;=\; \begin{pmatrix} \mathbf{R}_{\mathrm{A}} & \mathbf{D}_{\mathrm{B}\to\mathrm{A}} \\ \mathbf{D}_{\mathrm{A}\to\mathrm{B}} & \mathbf{R}_{\mathrm{B}} \end{pmatrix} \;=\; \begin{pmatrix} 0 & f_{\mathrm{A}} & 0 & 0 \\ s_{\mathrm{j,A}}\,(1-d_{\mathrm{A}}) & s_{\mathrm{a,A}} & s_{\mathrm{j,B}}\,d_{\mathrm{B}} & 0 \\ 0 & 0 & 0 & f_{\mathrm{B}} \\ s_{\mathrm{j,A}}\,d_{\mathrm{A}} & 0 & s_{\mathrm{j,B}}\,(1-d_{\mathrm{B}}) & s_{\mathrm{a,B}} \end{pmatrix}$$

For such a model, the kinship matrices can be computed from $\mathbf{M}$ in order to understand the kinship structure within and across patches. To understand the effect of the metapopulation structure, we can compare it to the kinship structure associated to each patch in the absence of dispersal.

In order to work out a numerical example, let us fix some simple parameters for patches A and B. We choose them to correspond to a situation where patch A has slightly better probabilities of survival and of reproduction (say, $s_{\mathrm{j,A}} = 0.4$, $s_{\mathrm{a,A}} = 0.5$, $d_{\mathrm{A}} = 0.5$ and $f_{\mathrm{A}} = 0.8 \times 2.3$ for patch A, where f_{A} is the product of the probability of reproduction and of the average number of offspring that survive their first year, conditional on reproduction; and $s_{\mathrm{j,B}} = 0.3$, $s_{\mathrm{a,B}} = 0.4$, $d_{\mathrm{B}} = 0.4$ and $f_{\mathrm{B}} = 0.7 \times 2.3$ for patch B). This gives the projection matrix

$$\mathbf{M} \;=\; \begin{pmatrix} 0 & 1.84 & 0 & 0 \\ 0.20 & 0.50 & 0.12 & 0 \\ 0 & 0 & 0 & 1.61 \\ 0.20 & 0 & 0.18 & 0.40 \end{pmatrix}.$$

The corresponding genealogical matrices are

$$\mathbf{P_s} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0.34 & 0.48 & 0.18 & 0 \\ 0 & 0 & 0 & 0 \\ 0.35 & 0 & 0.27 & 0.39 \end{pmatrix} \quad \text{and} \quad \mathbf{P_F} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Finally, assuming Poisson reproduction, Equation (11) from 2.4 gives the following same-litter newborn sisters matrix:

$$\mathbf{Z} \;=\; \begin{pmatrix} 1.84 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1.61 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The previous generation matrix

Because there are several newborn classes here, the previous-generation matrix $\mathbf{G}$ (see SM.5) has an interesting structure:

$$\mathbf{G} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0.66 & 0 & 0.34 \\ 0 & 0 & 0 & 1 \\ 0 & 0.56 & 0 & 0.44 \end{pmatrix}.$$

This shows that adults found in patch A are likely have been born in that patch (66% probability), whereas in patch B they are more likely to have been born the other patch and dispersed from it (56% probability). If we look at g-th generation ancestors, then

$$\lim_{g\to\infty} \mathbf{G}^g = \begin{pmatrix} 0 & 0.62 & 0 & 0.38 \\ 0 & 0.62 & 0 & 0.38 \\ 0 & 0.62 & 0 & 0.38 \\ 0 & 0.62 & 0 & 0.38 \end{pmatrix},$$

meaning that a distant ancestor of *ego* has a 62% probability of having been born in patch A, irrespective of the stage or the patch of *ego*; in fact, the ancestral lineage of *ego* goes to and fro between patches A and B, spending on average 62% of its time in patch A.

Quantifying the effect of dispersal on kinship

Let $\mathbf{K}_{\text{ref}}$ be the kinship matrix in the absence of dispersal (i.e. $d_{\text{A}} = d_{\text{B}} = 0$) and let $\Delta\mathbf{K} = \mathbf{K} - \mathbf{K}_{\text{ref}}$. The matrix $\Delta\mathbf{K}$ enables us to quantify the effect of dispersal on the kinship structure. For instance, looking at cousins,

$$\Delta\mathbf{K}(2,2) = \begin{pmatrix} -0.9010 & -0.6205 & 0.6195 & 0.4101 \\ -0.4072 & -0.6160 & 0.1881 & 0.2785 \\ 0.5351 & 0.2880 & -0.1964 & -0.0661 \\ 0.2285 & 0.2750 & -0.0152 & 0.0120 \end{pmatrix}.$$

As expected, dispersal not only decreases the number of cousins within patch A, but it also decreases the overall number of cousins for individuals from patch A: if for $x \in \{\text{j}, \text{a}\}$ and $X \in \{\text{A}, \text{B}\}$ we let $\Delta_{\text{tot}}(x, X) = \sum_{(y,Y)} \Delta k_{(y,Y),(x,X)}$ denote the total number of kin (of any age and of any patch) that an individual in class x and patch X gains as a result of dispersion (where $\Delta k_{(y,Y),(x,X)}$ denotes the entries of $\Delta\mathbf{K}$); in other words, $\Delta_{\text{tot}}(x, X)$ is the sum of the elements in the (x, X) column of $\Delta\mathbf{K}$), then we see that $\Delta_{\text{tot}}(\text{j}, \text{A}) = -0.5447 < 0$ and $\Delta_{\text{tot}}(\text{a}, \text{A}) = -0.6735 < 0$. By contrast, although dispersal decreases the number of within-patch cousins for patch B, it actually increases the overall number of cousins for individuals from that patch: $\Delta_{\text{tot}}(\text{j}, \text{B}) = 0.5960 > 0$ and $\Delta_{\text{tot}}(\text{a}, \text{B}) = 0.6346 > 0$.

If instead of focusing on a particular kin relation (here cousins) we want to have a more global vision of this, we can compare $\theta_{\text{c}}(\mathbf{M}) = 1.0544$, $\theta_{\text{c}}(\mathbf{R}_{\text{A}}|_{d_{\text{A}}=0}) = 1.2087$ and $\theta_{\text{c}}(\mathbf{R}_{\text{B}}|_{d_{\text{B}}=0}) = 0.8963$. Again this shows that dispersal tends to decrease the number of kin for patch A and for the overall population, but that it increases it for patch B. Note that these values have to be interpreted in light of the growth rates $\lambda(\mathbf{M}) = 1.0382$, $\lambda(\mathbf{R}_{\text{A}}|_{d_{\text{A}}=0}) = 1.1436$ and $\lambda(\mathbf{R}_{\text{B}}|_{d_{\text{B}}=0}) = 0.9232$.

Finally, it is also possible to compute the sensitivities of the kinship structure to dispersal. For instance,

$$\frac{\partial}{\partial d_{\text{A}}}\mathbf{K}(2,2) = \begin{pmatrix} -1.4215 & -0.8322 & -0.9660 & -0.8658 \\ -0.4994 & -0.6303 & -0.3252 & -0.6232 \\ 0.5861 & 0.3071 & 0.6818 & 0.3408 \\ 0.0922 & 0.1146 & 0.1910 & 0.1437 \end{pmatrix}$$

and

$$\frac{\partial}{\partial d_{\text{B}}}\mathbf{K}(2,2) \begin{pmatrix} 0.2433 & 0.2715 & 0.7660 & 0.3194 \\ 0.1725 & 0.3063 & 0.2859 & 0.3373 \\ -0.2587 & -0.0838 & -0.1888 & -0.1099 \\ -0.1943 & -0.1399 & -0.0522 & -0.1369 \end{pmatrix}.$$

Similarly, $\partial\theta_{\text{c}}/\partial d_{\text{A}} = -0.2163$ and $\partial\theta_{\text{c}}/\partial d_{\text{B}} = 0.1384$.

In summary, this simple toy model shows that our methods give a precise and flexible way to quantify the effect of various life-history traits such as dispersal on kin structure.

Effect of categories on kinship

Another way to study the effect of various traits on kinship is to compare a model that implements a trait with the corresponding model that does not implement it (the "folded" model, see Coste et al., 2017; Coste and Pavard, 2020). Here, the model is based on two traits: stage and patch. Patch-related traits account for dispersal and the heterogeneity in vital rates between patches, whereas stage-related traits account for the heterogeneity in vital rates between stages. We assess the overall effect of each of these traits on kinship by comparing the kinship matrices of various models: the original metapopulation model $\mathbf{A}$ implementing boths traits; the folded-over-patch model; the folded-over-stage model; and the unstructured model (i.e. the corresponding 1×1 matrix). The corresponding values of the expected

number of aunts and cousins, as well as the probability that the mother is alive, are given in Table 1. They show that adding the patch structure to the unstructured model or to the stage-based model has little effect on kin frequencies. By contrast, adding the stage structure to the the unstructured model or to the patch-based model has a noticeable effect. This suggests that there is more heterogeneity between stages than between patches, and that stage is a more relevant trait for the kinship structure than patch.

kin	metapopulation	stage model	patch model	1x1 model
mother $\mathbf{K}(1,0)$	0.3406	0.3348	0.2871	0.2853
aunts $\mathbf{K}(2,1)$	0.6218	0.5690	0.3468	0.3402
cousins $\mathbf{K}(2,2)$	1.4600	1.3635	0.5090	0.4961

Table 1: Probability that the mother of *ego* is alive and expected number of aunts / cousins of *ego*, for four models: the stage×patch metapopulation, the corresponding stage-based model, the corresponding patch-based model, and the corresponding unstructured model.

SM.3 Reducible projection matrices

In order for the genealogical matrices $\mathbf{P}$, $\mathbf{P_s}$ and $\mathbf{P_F}$ given by Equations (1) and (2) to be well-defined, the projection matrix $\mathbf{A}$ must have:

(i) a positive eigenvalue λ that is greater in modulus than any other eigenvalue;

(ii) a unique right-eigenvector $\mathbf{w}$ associated with λ;

(iii) $\mathbf{w} > 0$.

If $\mathbf{A}$ is primitive – that is, irreductible and aperiodic – then the Perron-Frobenius theorem ensures that this is the case (see e.g. Meyer, 2000; Varga, 2000). Most population projection matrices encountered in practice are primitive, and thus we have relied on primivity to state our results in the main text, for simplificity. However, there are biologically relevant situations in which a non-primitive matrix may verify points (i–iii). Examples of such situations include, among others, the presence of post-reproductive classes (as is the case for the life-cycle of the killer whale *Orcinus orca* presented in SM.2.1); and meta-population models with "sink" populations.

In practice, the easiest way to determine whether our results can be applied is to compute the eigen-elements of the projection matrix and see whether points (i–iii) are verified. If points (i) and (ii) are verified but the dominant eigenvector $\mathbf{w}$ has some zero entries, then it is possible to remove the corresponding classes from the model and proceed with the resulting projection matrix.

In the rest of this section, we recall some elementary properties of reducible matrices that are helpful to easily identify classes of matrices for which the genealogical Markov chain is defined – and to better understand its ergodicity (or lack thereof). These results will also be useful in our study of the Euler–Lotka matrices in SM.5 and of the "capped-age" model of SM.7.

Assume that we have a reducible projection matrix $\mathbf{A}$. In that case, the life cycle can be decomposed into several irreducible components, with "flows" of individuals between these components. Formally, the projection matrix $\mathbf{A}$ can be written as a block triangular matrix by permuting its rows and columns:

$$\mathbf{A} = \begin{pmatrix} \tilde{\mathbf{A}}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{\Phi}_{21} & \tilde{\mathbf{A}}_2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{\Phi}_{m1} & \mathbf{\Phi}_{m2} & \cdots & \tilde{\mathbf{A}}_m \end{pmatrix}, \tag{20}$$

yielding what is sometimes known as a *normal form* of the matrix (see e.g. Varga, 2000, Section 2.3). In this normal form, the square matrices $\tilde{\mathbf{A}}_1, \ldots, \tilde{\mathbf{A}}_m$ correspond to the internal dynamics of the irreducible components of the life-cycle graph, and the matrices $\Phi_{k\ell}$ to the flows between those components. The eigenvalues of $\mathbf{A}$ are the eigenvalues of $\tilde{\mathbf{A}}_1, \ldots, \tilde{\mathbf{A}}_m$.

Assume that one of the matrices $\tilde{\mathbf{A}}_1, \ldots, \tilde{\mathbf{A}}_m$, say $\tilde{\mathbf{A}}_x$, is aperiodic – and therefore primitive – and that its dominant eigenvalue is greater in modulus than any of the eigenvalues of the other matrices $\tilde{\mathbf{A}}_k$, $k \neq x$. In that case, the internal dynamics of $\tilde{\mathbf{A}}_x$ will exponentially dominate that of the other matrices $\tilde{\mathbf{A}}_k$. Components that do not receive individuals from $\tilde{\mathbf{A}}_x$ will not contribute to the asymptotic constitution of the population; components that do will – but without affecting the exponential growth rate of the population. Note that the matrices $\tilde{\mathbf{A}}_k$, $k \neq x$, do not need to be aperiodic: the aperiodicity of $\tilde{\mathbf{A}}_x$

will "kill" any possible periodic behavior in the components that receive individuals from it (and, while the components that do not receive individuals from $\tilde{\mathbf{A}}_x$ *may* exhibit a periodic behavior, they do not contribute to the asymptotic composition of the population).

Formally, let λ be the dominant eigenvalue of the primitive matrix $\tilde{\mathbf{A}}_x$, with $\lambda > \rho(\tilde{\mathbf{A}}_k)$ for all $k \neq x$, and let $\tilde{\mathbf{w}}^{(x)}$ be the associated right-eigenvector of $\tilde{\mathbf{A}}_x$. Then, λ is an eigenvalue of $\mathbf{A}$ that is larger in modulus than any of its other eigenvalues, and the only associated (unscaled) right-eigenvector of $\mathbf{A}$ is $\mathbf{w} = (w_i)$, with

$$w_i \;=\; \begin{cases} 0 & \text{if } i \text{ cannot be reached from a class of } \tilde{\mathbf{A}}_x \\ \tilde{w}_i^{(x)} > 0 & \text{if } i \text{ is a class of } \tilde{\mathbf{A}}_x \\ y_i > 0 & \text{if } i \text{ can be reached from a class of } \tilde{\mathbf{A}}_x, \end{cases}$$

where by "i can be reached from j" we mean that there exists a a directed path from j to i in the life-cycle graph that corresponds to $\mathbf{A}$.

A prototypical example of this situation is given by the matrix

$$\mathbf{A} \;=\; \begin{pmatrix} \mu & 0 & 0 \\ \varphi & \lambda & 0 \\ \eta & \psi & \nu \end{pmatrix}, \quad \text{with } \lambda > \max(\mu, \nu).$$

Here, the irreducible components are the $1{\times}1$ matrices $\tilde{\mathbf{A}}_1 = (\mu)$, $\tilde{\mathbf{A}}_2 = (\lambda)$ and $\tilde{\mathbf{A}}_3 = (\nu)$. The component with the largest eigenvalue is $\tilde{\mathbf{A}}_2$, and it is (trivially) primitive. $\tilde{\mathbf{A}}_1$ sends individuals to $\tilde{\mathbf{A}}_2$ through φ and to $\tilde{\mathbf{A}}_3$ through ν; and $\tilde{\mathbf{A}}_2$ sends individuals to $\tilde{\mathbf{A}}_3$ through ψ. It is easy to check that the right-eigenvector of $\mathbf{A}$ associated with λ is

$$\mathbf{w} \;\propto\; \begin{pmatrix} 0 \\ \lambda - \nu \\ \psi \end{pmatrix}.$$

As expected, $w_1 = 0$ because $\tilde{\mathbf{A}}_1$ does not receive individuals from the dominant component $\tilde{\mathbf{A}}_2$, and $w_3 > 0$ if and only if $\psi > 0$, i.e. if and only if $\tilde{\mathbf{A}}_3$ receives individuals from $\tilde{\mathbf{A}}_2$. Note that neither the internal dynamics of $\tilde{\mathbf{A}}_1$ (that is, μ) nor the flows of individuals from it (that is, ϕ and η) appear in the expression of $\mathbf{w}$.

Let us close this section by recapitulating what to do in case the projection matrix is not primitive:

1. Identify the irreducible components and find the *dominant component*, i.e. the irreducible component with the largest eigenvalue. If there are several such irreducible components, our results do not apply because the stationary distribution (if it exists) will depend on the initial composition of the population.

2. If the dominant component is periodic, our results do not apply because there is no stationary distribution.

3. If the dominant component is aperiodic, then separate the projection matrix into two matrices:

   - The matrix $\mathbf{B}$ containing all of the classes that *can* be reached from the dominant component. This matrix has the same dominant eigenvalue as the original projection matrix $\mathbf{A}$ and a unique, positive dominant eigenvector. Thus, its genealogical matrices are given by Equations (1) and (2).
   - The matrix $\mathbf{C}$ containing all of the classes that *cannot* be reached from the dominant component. This matrix describes a population process taking place on a smaller scale as that of $\mathbf{A}$ and $\mathbf{B}$. If it is not primitive, go to step 1 to study it.

   The matrices $\mathbf{B}$ and $\mathbf{C}$ can then be studied independently.

SM.4 Generating weak compositions

Our formula for the kinship matrix requires summing on weak compositions, i.e. on the sets

$$\mathcal{C}(n, \ell) \;=\; \Big\{(k_1, \ldots, k_\ell) \in \mathbb{N}^\ell : k_1 + \cdots + k_\ell = n\Big\}.$$

In this section, we explain how to do so efficiently.

First, note that the weak compositions of the integer n into ℓ parts are essentially the combinations of $I_{n+\ell-1} = \{1, \ldots, n+\ell-1\}$ of size $\ell - 1$ (that is, subsets of $I_{n+\ell-1}$ with $\ell - 1$ elements). Indeed, the

combination $C = \{i_1, \dots, i_{\ell-1}\} \subset I_{n+\ell-1}$, with $i_1 < \cdots < i_{\ell-1}$, can be seen as the vector $\mathbf{u} = (u_p)$ of symbols "$\bullet$" and "$|$", of length $n + \ell - 1$, defined by: $u_p = |$ if $p \in C$ and $u_p = \bullet$ otherwise. This vector $\mathbf{u}$ can in turn be interpreted as a weak composition $\mathbf{k} = (k_1, \dots, k_\ell)$ as follows:

$$\mathbf{u} \;=\; \underbrace{\bullet\cdots\bullet}_{k_1}|\underbrace{\bullet\cdots\bullet}_{k_2}| \quad \cdots \quad |\underbrace{\bullet\cdots\bullet}_{k_\ell}$$

where k_p is the number of "$\bullet$" between the $(p-1)$-th and the p-th "$|$". This way, $k_1, \dots, k_\ell$ are indeed non-negative integers such that $k_1 + \cdots + k_\ell = n$. This shows that the number of weak compositions of n into ℓ parts is

$$\#\mathcal{C}(n, \ell) \;=\; \binom{n+\ell-1}{\ell-1},$$

a fact that we will be of constant use in various part of this document (SM.6, SM.8, SM.9 and SM.10). More to the point here, this shows that one can obtain a weak composition $\mathbf{k} = (k_1, \dots, k_\ell)$ from the combination $C = \{i_1, \dots, i_{\ell-1}\}$, $i_1 < \cdots < i_{\ell-1}$ by setting $i_0 = 0$, $i_\ell = n + \ell$, and then

$$k_p \;=\; i_p - i_{p-1} - 1\,.$$

Since many programming languages – such as Python, R or MATLAB – provide a way to generate combinations, this gives a simple way to generate weak compositions.

Alternatively, the efficient Algorithm 1 below (from Knuth, 2011, Section 7.2.1.3) can be used. This approach has the advantage that it is not necessary to store the weak compositions in memory (although we do it here with the variable *list*, for the sake of giving a language-independent presentation): they can be generated one after the other as we need them and discarded as we go (typically, using a *generator* in Python).

Finally, note that although the matrices $\mathbf{Q}(\mathbf{k}) = \mathbf{P}_{\mathbf{S}}^{k_1}\,\mathbf{P}_{\mathbf{F}} \cdots \mathbf{P}_{\mathbf{F}}\,\mathbf{P}_{\mathbf{S}}^{k_\ell}$ and $\mathbf{R}(\mathbf{k}) = \mathbf{S}^{k_1}\mathbf{F} \cdots \mathbf{F}\,\mathbf{S}^{k_\ell}$ could be implemented directly from combinations rather than from compositions (by setting $\mathbf{Q} = \mathbf{I}$ and then doing, for $p = 1$ to $n + \ell - 1$: if $p \in C$, then $\mathbf{Q} \leftarrow \mathbf{Q}\mathbf{P}_{\mathbf{F}}$; else $\mathbf{Q} \leftarrow \mathbf{Q}\mathbf{P}_{\mathbf{S}}$), this approach does not take advantage of possible optimizations in the computation of matrix powers. For instance, assuming that matrix multiplication has a time complexity $\Theta(1)$ and that matrix powers are computed via exponentiation by squaring, calculating $\mathbf{Q}$ from combinations has a time complexity $\Theta(n + \ell)$ whereas calculating it from weak compositions has a time complexity $\Theta(\ell \log(\max_p k_p))$. This can make a significant difference when n is large and ℓ is small – as is typically the case in our applications.

Algorithm 1: Generating weak compositions

```
Input  : n > 0 and k > 0.
Output: the weak compositions of n into k parts, as a list of vectors.
list ← {}
c ← (c_0, ..., c_{k-1}), with c_0 = n and c_i = 0 for all i > 0
if k = 1 then
    add c to list
else
    r ← 1
    continue ← true
    while continue do
        add a copy of c to list
        if c_0 = 0 then
            if r = k − 1 then
                continue ← false
            else
                c_0 ← c_r − 1
                c_r ← 0
                r  ← r + 1
        else
            c_0 ← c_0 − 1
            r  ← 1
        c_r ← c_r + 1
return list
```

SM.5 The previous-generation and Euler–Lotka matrices

The population projection matrix $\mathbf{A}$ projects the population one time step into the future, whereas the genealogical matrix $\mathbf{P}$ projects the lineage of an individual one time step into the past. Just as the so-called next-generation matrix $\mathbf{R} = \mathbf{F}(\mathbf{I}-\mathbf{S})^{-1}$ allows us to project the population one *generation* into the future (see e.g. Caswell (2001), Chapter 5), it is possible to define a matrix $\mathbf{G}$ that projects the lineage of an individual one generation into the past. We term this matrix the *previous-generation matrix*, and show that it has interesting connections with the Euler–Lotka equation.

Recalling the interpretation of the matrice $\mathbf{P_S}$ and $\mathbf{P_F}$ detailed in Section 2.1, we see that when going up the lineage of an individual *ego* from class i the probability that we reach its mother t time steps ago and in class j is the (i,j)-th entry of the matrix $\mathbf{P_S}^{t-1}\mathbf{P_F}$. Considering all possible t at which this mother can be reached shows that the probability that the mother of *ego* was in class j when she gave birth to *ego* is the (i,j)-th entry of the matrix

$$\mathbf{G} \;=\; \sum_{t\geq 1} \mathbf{P_S}^{t-1}\mathbf{P_F} \;=\; (\mathbf{I}-\mathbf{P_S})^{-1}\mathbf{P_F}\,. \tag{21}$$

Note that $(\mathbf{I}-\mathbf{P_S})^{-1}$ is always well-defined, and that it is the fundamental matrix of the absorbing Markov chain associated with $\mathbf{P_S}$ (see e.g. Kemeny and Snell (1976), Chapter V). The Markov chain associated with $\mathbf{G}$ is not ergodic (because non-reproductive classes are inaccessible), but its restriction to reproductive classes is. Thus, $\mathbf{G}$ has a stationary distribution $\boldsymbol{\nu}$ that can be expressed as a function of the stationary distribution $\boldsymbol{\pi}$ of $\mathbf{P}$ (which, as we may recall, is itself given by the class reproductive values $\pi_i = v_i w_i$):

$$\boldsymbol{\nu} \;=\; \boldsymbol{\pi}\mathbf{P_F}\,. \tag{22}$$

If this is not clear from the probabilistic interpretation, consider the following calculation:

$$\boldsymbol{\pi}\mathbf{P_F} \;=\; \boldsymbol{\pi}(\mathbf{P}-\mathbf{P_S}) \;=\; \boldsymbol{\pi}(\mathbf{I}-\mathbf{P_S})$$

and therefore

$$\boldsymbol{\pi}\mathbf{P_F}\mathbf{G} \;=\; \boldsymbol{\pi}(\mathbf{I}-\mathbf{P_S})(\mathbf{I}-\mathbf{P_S})^{-1}\mathbf{P_F} \;=\; \boldsymbol{\pi}\mathbf{P_F}\,.$$

A notable fact about the previous-generation matrix $\mathbf{G}$ is that it corresponds to the transition matrix of the genealogical Markov chain associated with the projection matrix

$$\mathbf{E} \;=\; (\lambda\mathbf{I}-\mathbf{S})^{-1}\mathbf{F}\,, \tag{23}$$

as proved in Proposition 3 below. Even though $\mathbf{E}$ is not primitive (because the non-reproductive classes of $\mathbf{A}$ correspond to "dead-ends" in the life-cycle graph associated to $\mathbf{E}$), it has a dominant eigenvalue and a unique dominant right-eigenvector: 1 is an eigenvalue of $\mathbf{E}$ and it is greater in modulus than any of its other eigenvalues; and $\mathbf{w}$, the stable class distribution of $\mathbf{A}$, is the unique right-eigenvector of $\mathbf{E}$ associated with the eigenvalue 1 (see Proposition 1). Thus, since $\mathbf{E}$ can also be written

$$\mathbf{E} \;=\; \sum_{t\geq 1} \lambda^t\,\mathbf{S}^{t-1}\,\mathbf{F}\,, \tag{24}$$

the equation $\mathbf{E}\mathbf{w} = \mathbf{w}$ is a generalization of the Euler–Lotka equation to arbitrarily structured populations. Indeed, in age-based models, the Euler–Lotka equation (Lotka, 1939) is

$$\sum_{t\geq 1} \lambda^t\,\ell(t)\,f(t) \;=\; 1\,, \tag{25}$$

where $\ell(t)$ is the probability of surviving to age t and $f(t)$ is the fertility at age t. But, in a general model: (1) the probability of surviving to age t depends both on the class in which the individual was born and on its class at age t, and it is given by the matrix $\mathbf{S}^t$; (2) the fertilities depend only on the class of the individual at age t (which may or may not include age as a variable) and are given by $\mathbf{F}$.

It should be noted, however, that although $\mathbf{E}$ provides a very natural generalization of the Euler–Lotka equation, another natural generalization is possible. Indeed, consider the matrix

$$\mathbf{E}^\star \;=\; \mathbf{F}(\lambda\mathbf{I}-\mathbf{S})^{-1} \;=\; \sum_{t\geq 1} \lambda^t\;\mathbf{F}\,\mathbf{S}^{t-1}\,. \tag{26}$$

Then, as shown in Proposition 1, $\mathbf{E}^\star$ also has dominant eigenvalue 1 and satisfies $\mathbf{v}\mathbf{E}^\star = \mathbf{v}$, where $\mathbf{v}$ is the vector of reproductive values of $\mathbf{A}$. As a result, the equation $\mathbf{v}\mathbf{E}^\star = \mathbf{v}$ can also be viewed as

a natural generalization of the Euler–Lotka equation. Finally, note that $\mathbf{E}^\star$ can be interpreted as the next-generation matrix associated with the projection matrix $\lambda^{-1}\mathbf{A}$.

Because of their interpretation as what are arguably the two most natural generalizations of the Euler–Lotka equation, we term the matrices $\mathbf{E}$ and $\mathbf{E}^\star$ the *Euler–Lotka matrices.*

We now close this section by proving our various claims about the matrices $\mathbf{G}$, $\mathbf{E}$ and $\mathbf{E}^\star$.

Proposition 1. *Let $\mathbf{A} = \mathbf{S} + \mathbf{F}$ be a primitive matrix with dominant eigenvalue λ and dominant left and right eigenvector $\mathbf{v}$ and $\mathbf{w}$, respectively. Assume that the matrix $\mathbf{S}$ is convergent and that $\mathbf{F} \neq \mathbf{0}$, and define the matrices*

$$\mathbf{E} \;=\; (\lambda\mathbf{I} - \mathbf{S})^{-1}\mathbf{F} \quad \textit{and} \quad \mathbf{E}^\star \;=\; \mathbf{F}(\lambda\mathbf{I} - \mathbf{S})^{-1}\,.$$

Then, $\mathbf{E}$ and $\mathbf{E}^\star$ have dominant eigenvalue 1 and

(i) $\mathbf{E}\,\mathbf{w} = \mathbf{w}$.

(ii) $\mathbf{v}\,\mathbf{E}^\star = \mathbf{v}$.

Moreover, those are the only dominant eigenvectors (right or left, respectively) of those matrices.

Proof. Let us start by proving (i) and (ii). First, note that

$$\mathbf{F}\mathbf{w} \;=\; (\mathbf{A} - \mathbf{S}\mathbf{w}) \;=\; (\lambda\mathbf{I} - \mathbf{S})\mathbf{w}\,,$$

and therefore

$$\mathbf{E}\mathbf{w} \;=\; (\lambda\mathbf{I} - \mathbf{S})^{-1}\mathbf{F}\mathbf{w} \;=\; (\lambda\mathbf{I} - \mathbf{S})^{-1}(\lambda\mathbf{I} - \mathbf{S})\mathbf{w} \;=\; \mathbf{w}\,,$$

proving (i). Similarly, we get (ii) by noting that $\mathbf{v}\mathbf{F} = \mathbf{v}(\lambda\mathbf{I} - \mathbf{S})$.

Let us now prove that $\mathbf{E}$ and $\mathbf{E}^\star$ have spectral radius 1. Let ρ denote the spectral radius. Note that since $\rho(\mathbf{AB}) = \rho(\mathbf{BA})$ for every matrices $\mathbf{A}$ and $\mathbf{B}$ (see e.g. Meyer, 2000, Exercice 7.1.19), we immediately have $\rho(\mathbf{E}) = \rho(\mathbf{E}^\star)$; moreover, from (i) and (ii), this spectral radius is at least 1.

Now, the "min-max" Collatz–Wielandt formula for non-negatives matrices (see Equation 8.3.3 in Meyer (2000) for the usual "max-min" version, and Exercice 8.2.9 for the "min-max" version) gives

$$\rho(\mathbf{E}) \;=\; \min_{\mathbf{x}>0}\,\max_i\,\frac{(\mathbf{E}\mathbf{x})_i}{x_i} \;\leq\; \max_i\,\frac{(\mathbf{E}\mathbf{w})_i}{w_i} \;=\; 1\,,$$

since $\mathbf{w} > 0$ and $\mathbf{E}\mathbf{w} = \mathbf{w}$.

To conclude the proof, we have to show that $\mathbf{E}$ (resp. $\mathbf{E}^\star$) has no other eigenvalue with modulus 1, and that $\mathbf{w}$ (resp. $\mathbf{v}$) is the only right-eigenvector (resp. left-eigenvector) associated with the eigenvalue 1. For this, consider the life-cycle graph associated with $\mathbf{E}$, and note that non-reproductive classes (that is, j such that $f_{ij} = 0$ for all i) are terminal vertices in that graph. By contrast, the set of reproductive classes is irreducible (to find a directed paths between two vertices in the graph of $\mathbf{E}$, consider the restriction of one such path in the graph of $\mathbf{A}$ – which is guaranteed to exists because $\mathbf{A}$ is irreducible – to the reproductive classes). The result then follows from SM.3. The matrix $\mathbf{E}^\star$ is treated similarly, but using newborn classes (that is, i such that there exists j with $f_{ij} > 0$) instead of reproductive ones. □

Remark 2. The fact that $\rho(\mathbf{E}^\star) = 1$ can also be seen from the interpretation of $\mathbf{E}^\star$ as the next-generation matrix of the projection matrix $\lambda^{-1}\mathbf{A}$: indeed, as such, $\rho(\mathbf{E}^\star)$ is the R_0 associated with the matrix $\lambda^{-1}\mathbf{A}$, whose asymptotic growth rate is 1. ◇

Proposition 3. *The previous-generation matrix $\mathbf{G} = (\mathbf{I} - \mathbf{P}_\mathbf{S})^{-1}\mathbf{P}_\mathbf{F}$ is the transition matrix of the genealogical Markov chain associated with the Euler–Lotka matrix $\mathbf{E} = (\lambda\mathbf{I} - \mathbf{S})^{-1}\mathbf{F}$.*

Proof. First, note that since, by Proposition 1, $\mathbf{E}$ has dominant eigenvalue 1 and a unique dominant right-eigenvector $\mathbf{w}$, its genealogical Markov chain is well-defined and given by Equation (1). Therefore, we only have to show that the entries of $\mathbf{G}$ and $\mathbf{E}$ satisfy:

$$(\mathbf{G})_{ij} \;=\; (\mathbf{E})_{ij}\frac{w_j}{w_i}.$$

For this, let us note that, for all $t \geq 1$,

$$(\mathbf{P}_\mathbf{S}^t)_{ij} \;=\; \frac{w_j}{w_i}\lambda^{-t}(\mathbf{S}^t)_{ij}. \tag{27}$$

Indeed, the case $t = 1$ is the definition of $\mathbf{P}_\mathbf{S} = (p_\mathbf{S}(i,k))$, see Equation (2). Thus, by induction,

$$(\mathbf{P}_\mathbf{S}^{t+1})_{ij} \;=\; \sum_k p_\mathbf{S}(i,k)(\mathbf{P}_\mathbf{S}^t)_{kj}\times \;=\; \sum_k \frac{s_{ik}w_k}{\lambda w_i}\times\frac{w_j}{w_k}\lambda^{-t}(\mathbf{S}^t)_{kj} \;=\; \frac{w_j}{w_i}\lambda^{-(t+1)}(\mathbf{S}^{t+1})_{ij}\,.$$

Now, from $\mathbf{G} = \sum_{t\geq 0} \mathbf{P}_\mathbf{S}^t \mathbf{P}_\mathbf{F}$ we get

$$\begin{aligned}
(\mathbf{G})_{ij} &= \sum_{t\geq 0}\sum_k (\mathbf{P}_\mathbf{S}^t)_{ik} \times (\mathbf{P}_\mathbf{F})_{kj} \\
&= \sum_{t\geq 0}\sum_k \frac{w_k}{w_i}\lambda^{-t}(\mathbf{S}^t)_{ik} \times f_{kj}\frac{w_j}{\lambda w_k} \\
&= \sum_{t\geq 0} \lambda^{-t+1}(\mathbf{S}^t\mathbf{F})_{ij}\frac{w_j}{w_i} \\
&= (\mathbf{E})_{ij}\frac{w_j}{w_i}\,,
\end{aligned}$$

concluding the proof. □

SM.6 Upper bound on $\|\mathbf{K}(g,q) - \widehat{\mathbf{K}}(g,q\,;t_{\max},\alpha_{\max})\|$

In this section, we provide an upper bound on the error that results from using a finite sum when computing the kinship matrix, as explained in Section 2.5. This upper bound can be used to determine values of $t_{\max}$ and $\alpha_{\max}$ that ensure we meet a given accuracy in our estimation of $\mathbf{K}$. It should be noted, however, that this upper bound is rather crude. As such, it will sometimes greatly overestimate the actual error – to the point of being uninformative. This is especially likely to happen in species with high survival and a low fertility.

Let $\|\cdot\|$ be any submultiplicative matrix norm, meaning that for any two matrices $\mathbf{A}$ and $\mathbf{B}$ we have $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \times \|\mathbf{B}\|$. Note that most of the "standard" matrix norms (the induced norms, the Frobenius norm, the nuclear norm, etc) are submultiplicative, one notable exception being the max norm. Set

$$s = \|\mathbf{S}\|, \quad f = \|\mathbf{F}\|, \quad \sigma = \|\mathbf{P}_\mathbf{S}^\top\|, \quad \phi = \|\mathbf{P}_\mathbf{F}^\top\|, \quad z = \|\mathbf{Z}\|.$$

For our upper bound to be useful, the norm $\|\cdot\|$ should be chosen so that $s\sigma < 1$, ideally with $s\sigma$ as small as possible. Note that the existence of such a norm is guaranteed by the fact that the matrices $\mathbf{S}$ and $\mathbf{P}_\mathbf{S}$ are convergent (that is, $\mathbf{S}^t \to \mathbf{0}$ and $\mathbf{P}_\mathbf{S}^t \to \mathbf{0}$ as $t \to \infty$). In particular, the ℓ_1 operator norm, given by

$$\|\mathbf{A}\|_1 \;=\; \max_j \sum_i |a_{ij}|$$

always satisfies $\|\mathbf{P}_\mathbf{S}^\top\|_1 \leq 1$ and $\|\mathbf{S}\|_1 < 1$, unless there exists a class from which individuals cannot die. However, in practice we do not necessarily recommend sticking to this norm, but instead comparing several of them and choosing the one that minimizes $s\sigma$.

Once a norm has been chosen and the quantities s, f, σ, ϕ and z have been computed accordingly, our upper bound on $\|\mathbf{K}(g,q) - \widehat{\mathbf{K}}(g,q\,;t_{\max},\alpha_{\max})\|$ is

$$\begin{aligned}
c(g,q)\Bigg[\Big(\mathbb{1}_{\{g=0\}} \;+\; g\big[T_{g+q-2}(0) - T_{g+q-2}(t_{\max})\big]\Big)\, q\, T_{q-1}(\alpha_{\max}) \\
+\; \Big(1 + \tfrac{z}{f\phi}\, q + q\, T_{q-1}(0)\Big)\, g\, T_{g+q-1}(t_{\max})\Bigg]\,,
\end{aligned} \tag{28}$$

where

$$c(g,q) \;=\; \frac{(\phi/\sigma)^g (f/s)^q}{g!\; q!} \quad \text{and} \quad T_k(n) = \sum_{t>n} t^k (s\sigma)^t\,.$$

Note that this upper bound goes to zero as $t_{\max}$ and $\alpha_{\max}$ go to infinity if and only if $s\sigma < 1$. To compute it numerically, T can be expressed using special functions; for instance,

$$T_k(n) \;=\; (s\sigma)^{(n+1)}\, \Phi(s\sigma, -k, n+1)\,,$$

where Φ denotes the Lerch transcendent, which is available in many programming languages (Python, R, Mathematica, etc). In languages that do not implement the Lerch transcendent, such as MATLAB, the polylogarithm can be used instead:

$$T_k(n) \;=\; \mathrm{Li}_{-k}(s\sigma) \;-\; \sum_{t=1}^{n} t^k (s\sigma)^t\,.$$

It is also relatively straightforward to design an ad hoc implementation of T.

The rest of this section is devoted to the calculation of the upper bound given in (28). Let us start by introducing some notation. As in the rest of this manuscript, we use the convention $\binom{n}{k} = 0$ when $n < 0$ or $k \notin \{0, \dots, n\}$, so that we always have

$$\#\mathcal{C}(n,k) \;=\; \binom{n+k-1}{k-1}.$$

We also use the habitual convention $1/k! = 0$ when $k = -1, -2, \dots$

Lemma 4.

(i) *For $g = 0$, $\|\mathbf{U}(0,0)^\top\| = 1$ and $\|\mathbf{U}(t,0)^\top\| = 0$ when $t > 0$.*

(ii) *For $g \geq 1$, and $t \geq 0$, $\|\mathbf{U}(t,g)^\top\| \leq \frac{\phi^{g-1}}{\sigma^g\,(g-1)!}\, t^{g-1}\sigma^t$.*

(iii) *For $t \geq 2$ and any q, $\|\mathbf{D}(t,q)\| \leq \frac{f^q}{s^{q+1}\,q!}\,(t-1)^q\, s^t$.*

(iv) *For $t \geq 0$ and any q, $\|\mathbf{D}(t,q)\| \leq \frac{f^q}{s^{q+1}\,q!}\, t^q\, s^t$.*

Proof. Point (i) is immediate, since $\mathbf{U}(t,0) = \mathbf{I}$ if $t = 0$ and $\mathbf{0}$ otherwise, and $\|\,\mathbf{I}\,\| = 1$ for any submultiplicative matrix norm.

For point (ii), using $\#\mathcal{C}(t-g,g) = \binom{t-1}{g-1}$ and the submultiplicativity of the norm gives

$$\|\mathbf{U}(t,g)^\top\| \;=\; \Big\|\sum_{\mathbf{k}\in\mathcal{C}(t-g,g)} (\mathbf{P}_\mathbf{S}^{k_1}\,\mathbf{P}_\mathbf{F}\cdots\mathbf{P}_\mathbf{F}\,\mathbf{P}_\mathbf{S}^{k_g})^\top\Big\| \;\leq\; \binom{t-1}{g-1}\sigma^{t-g}\phi^{g-1}\,,$$

and the result follows from the inequality $\binom{t-1}{g-1} \leq t^{g-1}/(g-1)!$. Similarly,

$$\|\mathbf{D}(t,q)\| \;\leq\; \binom{t-1}{q} s^{t-1-q} f^q\,,$$

proving (iii) and (iv) since $\binom{t-1}{q} \leq (t-1)^q/q!$ for $t \geq 2$ and $\binom{t-1}{q} \leq t^q/q!$ for $t \geq 0$. □

Proof of Inequality 28. With the notation of Section 2, let us write

$$\widehat{\mathbf{T}}(t\,;\alpha_{\max}) \;=\; \mathbf{D}(t,q)\,\mathbf{S}\,\mathbf{P}_\mathbf{F}^\top \;+\; \mathbf{D}(t,q-1)\,\mathbf{Z} \;+\; \sum_{\alpha=1}^{\alpha_{\max}} \mathbf{D}(t+\alpha,q-1)\,\mathbf{F}\,(\mathbf{P}_\mathbf{S}^\top)^\alpha\,\mathbf{Y}(g)$$

and $\mathbf{T}(t) = \lim_{\alpha_{\max}} \widehat{\mathbf{T}}(t\,;\alpha_{\max})$. Note that we have kept the dependency on g and q implicit to ease the notation, and will allow ourselves to do so with other matrices such as $\mathbf{K}$ and $\mathbf{U}$ when there is no risk of confusion. With this notation, we have

$$\mathbf{K} \;=\; \sum_{t\geq 0} \mathbf{T}(t)\,\mathbf{U}(t)^\top \quad\text{and}\quad \widehat{\mathbf{K}}(t_{\max},\alpha_{\max}) \;=\; \sum_{t=0}^{t_{\max}} \widehat{\mathbf{T}}(t\,;\alpha_{\max})\,\mathbf{U}(t)^\top,$$

and, therefore,

$$\mathbf{K} \;-\; \widehat{\mathbf{K}}(t_{\max},\alpha_{\max}) \;=\; \sum_{t=0}^{t_{\max}} \big(\mathbf{T}(t) - \widehat{\mathbf{T}}(t\,;\alpha_{\max})\big)\,\mathbf{U}(t)^\top \;+\; \sum_{t>t_{\max}} \mathbf{T}(t)\,\mathbf{U}(t)^\top.$$

Using, again, the submultiplicativity of the norm, this yields

$$\|\mathbf{K} \;-\; \widehat{\mathbf{K}}(t_{\max},\alpha_{\max})\| \;\leq\; \sum_{t=0}^{t_{\max}} \|\mathbf{T}(t) - \widehat{\mathbf{T}}(t\,;\alpha_{\max})\| \times \|\mathbf{U}(t)^\top\| \;+\; \sum_{t>t_{\max}} \|\mathbf{T}(t)\| \times \|\mathbf{U}(t)^\top\|\,. \tag{29}$$

Let us assume that $g \geq 1$ and treat the case $g = 0$ zero separately later. Then, $\mathbf{Y}(g) = \mathbf{P}_\mathbf{F}^\top$ and therefore

$$\mathbf{T}(t) - \widehat{\mathbf{T}}(t\,;\alpha_{\max}) \;=\; \sum_{\alpha>\alpha_{\max}} \mathbf{D}(t+\alpha,q-1)\,\mathbf{F}\,(\mathbf{P}_\mathbf{S}^\top)^\alpha\,\mathbf{P}_\mathbf{F}^\top\,.$$

Note that since $\mathbf{U}(0) = \mathbf{0}$, we can have $t \geq 1$ and $\alpha \geq 1$ in the sum. As a result, we can apply point (iii) of Lemma 4 to $\mathbf{D}(t+\alpha, q-1)$ and get

$$\|\mathbf{T}(t) - \widehat{\mathbf{T}}(t\,;\alpha_{\max})\| \;\leq\; \frac{f^q\,\phi}{s^q\,(q-1)!} \sum_{\alpha>\alpha_{\max}} (t+\alpha-1)^{q-1}\, s^{t+\alpha}\sigma^\alpha\,.$$

Using that $(t+\alpha-1)^{q-1} \leq t^{q-1}\alpha^{q-1}$ for all $t \geq 1$ and $\alpha \geq 1$ then gives

$$\|\mathbf{T}(t) - \widehat{\mathbf{T}}(t\,;\alpha_{\max})\| \;\leq\; \frac{f^q\,\phi}{s^q\,(q-1)!}\; t^{q-1}\, s^t \sum_{\alpha>\alpha_{\max}} \alpha^{q-1}\,(s\sigma)^\alpha\,.$$

Using point (ii) of Lemma 4 in this expression, we get

$$\begin{aligned} &\sum_{t=0}^{t_{\max}} \|\mathbf{T}(t) - \widehat{\mathbf{T}}(t\,;\alpha_{\max})\| \times \|\mathbf{U}(t)^\top\| \\ \leq\;& c(g,q)\,g\,q \sum_{t=1}^{t_{\max}} t^{g+q-2}\,(s\sigma)^t \sum_{\alpha>\alpha_{\max}} \alpha^{q-1}\,(s\sigma)^\alpha\,, \end{aligned} \tag{30}$$

where

$$c(g,q) \;=\; \frac{(\phi/\sigma)^g (f/s)^q}{g!\;q!}\,.$$

Similarly, applying Lemma 4 in

$$\mathbf{T}(t) \;=\; \mathbf{D}(t,q)\,\mathbf{S}\,\mathbf{P}_{\mathbf{F}}^\top \;+\; \mathbf{D}(t,q-1)\,\mathbf{Z} \;+\; \sum_{\alpha>0} \mathbf{D}(t+\alpha,q-1)\,\mathbf{F}\,(\mathbf{P}_{\mathbf{S}}^\top)^\alpha\,\mathbf{P}_{\mathbf{F}}^\top$$

yields, after some rearranging and using that $t^{q-1} \leq t^q$,

$$\|\mathbf{T}(t)\| \;\leq\; \frac{f^q\phi}{s^q\,q!}\Big(1 \;+\; \tfrac{z}{f\phi}\,q \;+\; q\sum_{\alpha>0}\alpha^{q-1}(s\sigma)^\alpha\Big)\,t^q\,s^t$$

Therefore,

$$\sum_{t>t_{\max}} \|\mathbf{T}(t)\| \times \|\mathbf{U}(t)^\top\| \;\leq\; c(g,q)\,g\,\Big(1 \;+\; \tfrac{z}{f\phi}\,q \;+\; q\sum_{\alpha>0}\alpha^{q-1}(s\sigma)^\alpha\Big)\sum_{t>t_{\max}} t^{q+g-1}(s\sigma)^t. \tag{31}$$

Plugging Inequalities (30) and (31) in (29) gives us the following upper bound for $g \geq 1$

$$\begin{aligned} &\|\mathbf{K} \;-\; \widehat{\mathbf{K}}(t_{\max},\alpha_{\max})\| \\ &\quad\leq\; c(g,q)\,g\,\Bigg[\sum_{t=1}^{t_{\max}} t^{g+q-2}\,(s\sigma)^t\; q\sum_{\alpha>\alpha_{\max}} \alpha^{q-1}\,(s\sigma)^\alpha \\ &\qquad\qquad +\; \Big(1 + \tfrac{z}{f\phi}\,q + q\sum_{\alpha>0}\alpha^{q-1}(s\sigma)^\alpha\Big)\sum_{t>t_{\max}} t^{q+g-1}(s\sigma)^t\Bigg]\,. \end{aligned} \tag{32}$$

Note that, because of the factor g, this upper bound is equal to 0 when $g = 0$.

Let us now turn to the case $g = 0$. In that case, the expression of the kinship matrix reduces to

$$\mathbf{K}(0,q) \;=\; \sum_{\alpha\geq 1} \mathbf{D}(\alpha,q-1)\,\mathbf{F}\,(\mathbf{P}_{\mathbf{S}}^\top)^\alpha\,,$$

see Equation (12) in the main text. Therefore,

$$\mathbf{K}(0,q) - \widehat{\mathbf{K}}(0,q\,;t_{\max},\alpha_{\max}) \;=\; \sum_{\alpha>\alpha_{\max}} \mathbf{D}(\alpha,q-1)\,\mathbf{F}\,(\mathbf{P}_{\mathbf{S}}^\top)^\alpha$$

and using point (iv) of Lemma 4 we get

$$\|\mathbf{K}(0,q) - \widehat{\mathbf{K}}(0,q\,;t_{\max},\alpha_{\max})\| \;\leq\; \frac{f^q}{s^q\,(q-1)!} \sum_{\alpha>\alpha_{\max}} \alpha^{q-1}\,(s\sigma)^\alpha\,,$$

which, since $c(0,q) = (f/s)^q/(q-1)!$, can also be written

$$\|\mathbf{K}(0,q) - \widehat{\mathbf{K}}(0,q\,;t_{\max},\alpha_{\max})\| \;\le\; c(g,q)\,q\sum_{\alpha>\alpha_{\max}}\alpha^{q-1}\,(s\sigma)^{\alpha}\,. \tag{33}$$

Combining (32) and (33), we get the following general upper bound

$$\begin{aligned}
&\|\mathbf{K} \,-\, \widehat{\mathbf{K}}(t_{\max},\alpha_{\max})\| \\
&\quad\le\; c(g,q)\Bigg[\Big(\mathbb{1}_{\{g=0\}} \,+\, g\sum_{t=1}^{t_{\max}} t^{g+q-2}\,(s\sigma)^t\Big)\,q\sum_{\alpha>\alpha_{\max}}\alpha^{q-1}\,(s\sigma)^{\alpha} \\
&\qquad\qquad +\;\Big(1+\tfrac{z}{f\phi}\,q + q\sum_{\alpha>0}\alpha^{q-1}(s\sigma)^{\alpha}\Big)\,g\sum_{t>t_{\max}} t^{g+q-1}(s\sigma)^t\Bigg]\,.
\end{aligned}$$

Letting

$$T_k(n) \;=\; \sum_{t>n} t^k(s\sigma)^t\,,$$

the right-hand side of this inequality can also be written

$$\begin{aligned}
c(g,q)\Bigg[&\Big(\mathbb{1}_{\{g=0\}} \,+\, g\big[T_{g+q-2}(0) - T_{g+q-2}(t_{\max})\big]\Big)\,q\,T_{q-1}(\alpha_{\max}) \\
&\qquad +\;\Big(1+\tfrac{z}{f\phi}\,q + q\,T_{q-1}(0)\Big)\,g\,T_{g+q-1}(t_{\max})\Bigg]\,,
\end{aligned}$$

yielding the upper bound (28). □

SM.7 Capping age in a model

In this section, we briefly describe a straightforward way to turn a model without a maximal age (that is, such that $\mathbf{S}^t \neq 0$ for all t) into a model with maximal age ω (that is, where $\mathbf{S}^{\omega-1} \neq \mathbf{0}$ and $\mathbf{S}^{\omega} = \mathbf{0}$). This could be useful, e.g, to estimate the precision of the calculation of the kinship matrix done by choosing $t_{\max}$ and $\alpha_{\max}$ as a function of ω, as discussed in Section 2.5. However, there are some caveats associated with this "capping of age", as the resulting model is non-trivially related to the original one.

First, given a projection matrix $\mathbf{A} = \mathbf{S} + \mathbf{F}$, choose a maximal age $\hat{\omega}$ (meaning that individuals will *not* reach age $\hat{\omega}$ in the capped model). This could be done:

- either using some a priori knowledge on the biology of the species / population under study;
- or from the model itself, by picking ε such that the probability of reaching age $\hat{\omega}$ is less than ε. This can be done as follows: letting $\mathbf{x} = (x_i)$ be defined by $x_i = 1$ if i is a newborn class and $x_i = 0$ otherwise, we have: $\hat{\omega} = \min\{\omega \ge 1 : \max(\mathbf{S}^{\omega}\mathbf{x}) < \varepsilon\}$.

Once $\hat{\omega}$ has been chosen, one way to obtain a projection matrix with maximal age $\hat{\omega}$ is the following. Start by considering the matrix $\mathbf{B}(\hat{\omega})$ defined, using block-notation, by

$$\mathbf{B}(\hat{\omega}) \;=\; \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ \vdots \\ \hat{\omega} \end{array}\!\!\begin{array}{c} \begin{array}{ccccc} 1 & 2 & \cdots & \hat{\omega}-1 & \hat{\omega} \end{array} \\ \begin{pmatrix} \mathbf{F} & \mathbf{F} & \cdots & \mathbf{F} & \mathbf{F} \\ \mathbf{S} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{S} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{S} & \mathbf{0} \end{pmatrix} \end{array}.$$

Such "block-Leslie" matrices are frequently used to convert *stage* classified models into *age* × *stage* classified models – see for instance the matrix $\mathbf{M}_{\mathrm{AS}}(p)$ of Lebreton (2005), which is almost identical to this one with the exception of the $(\hat{\omega},\hat{\omega})$-th block, which we take to be $\mathbf{0}$ whereas Lebreton takes it to be $\mathbf{S}$.

Note that the matrix $\mathbf{B}(\hat{\omega})$ is aperiodic, but that it is *not* irreducible. For this reason, it will be more convenient to work with the matrix $\bar{\mathbf{A}}(\hat{\omega})$ obtained by keeping only the classes of $\mathbf{B}(\hat{\omega})$ that have a non-zero asymptotic abundance, as explained in SM.3.

Because the projection matrix $\bar{\mathbf{A}}(\hat{\omega})$ is primitive and has a maximal age $\hat{\omega}$, its kinship matrix can be computed exactly by taking $t_{\max} = \min(g\,\hat{\omega}, (q+1)\,\hat{\omega})$ and $\alpha_{\max} = \hat{\omega} - 1$ in Equation (15). This yields

a kinship matrix with the same dimensions as $\bar{\mathbf{A}}(\hat{\omega})$. This kinship matrix $\bar{\mathbf{K}} = (\bar{k}_{ij})$ can be turned back into an $n \times n$ kinship matrix, where n is the dimension of $\mathbf{A}$, by summing the entries that correspond to the same class of $\mathbf{A}$ while weighting each entry $\bar{k}_{ij}$ by the asymptotic abundance $\bar{w}_j$.

However, some care is needed when interpreting the resulting kinship matrix: indeed, the capped matrix $\bar{\mathbf{A}}(\hat{\omega})$ describes different population dynamics than the matrix $\mathbf{A}$; for instance, it has a smaller asymptotic growth rate, a shorter generation time, etc. As a result, it may or may not better describe the population under study than the matrix $\mathbf{A}$, and it is the responsibility of the ecologist to determine whether this is the case.

Finally, note that if all descriptors of $\mathbf{A}$ and $\bar{\mathbf{A}}(\hat{\omega})$ are very similar, this tends to suggest that neglecting individuals that reach age $\hat{\omega}$ has little influence on the descriptors of the model, and could thus indicate that the kinship matrices obtained from $\mathbf{A}$ by taking $t_{\max} = \min(g\,\hat{\omega}, (q+1)\,\hat{\omega})$ and $\alpha_{\max} = \hat{\omega} - 1$ in Equation (15) are fairly accurate. This gives a way to get a rough idea of the precision of the estimate $\widehat{\mathbf{K}}$ when the upper bound of SM.6 is not useful.

SM.8 Expression of $\mathbf{K}(g, q)$ for 1×1 models

Assume that the projection matrix is a 1×1 matrix $\mathbf{A} = (s + f)$, with $f > 0$ and $0 \leq s < 1$, and let $\lambda = f + s$. In that case, the genealogical matrices $\mathbf{P_S}$ and $\mathbf{P_F}$, given by Equation (2) in the main text, are

$$\mathbf{P_S} = s/\lambda \quad \text{and} \quad \mathbf{P_F} = f/\lambda\,. \tag{34}$$

Since these matrices commute and that $\#\mathcal{C}(n, \ell) = \binom{n+\ell-1}{\ell-1}$, the expression of the matrix $\mathbf{U}(t, g)$ given by Equation (4) becomes: $\mathbf{U}(0, 0) = 1$ and, for any other values of t and g,

$$\mathbf{U}(t, g) \;=\; \sum_{\mathbf{k} \in \mathcal{C}(t-g, g)} \mathbf{P_S}^{k_1}\,\mathbf{P_F} \cdots \mathbf{P_F}\,\mathbf{P_S}^{k_g} \;=\; \binom{t-1}{g-1} \frac{s^{t-g} f^{g-1}}{\lambda^{t-1}}\,, \tag{35}$$

with the convention that $\binom{n}{k} = 0$ when $n < 0$ or $k \notin \{0, \ldots, n\}$. Similarly, Equation (5) gives

$$\mathbf{D}(t, q) \;=\; \sum_{\mathbf{k} \in \mathcal{C}(t-1-q,\, q+1)} \mathbf{S}^{k_1}\mathbf{F} \cdots \mathbf{F}\,\mathbf{S}^{k_{q+1}} \;=\; \binom{t-1}{q} s^{t-1-q} f^q\,. \tag{36}$$

Finally, the same-litter newborn sisters matrix $\mathbf{Z}$ (see 2.4) can be written $\mathbf{Z} = \tilde{f} f/\lambda$, where $\tilde{f} + 1$ is the expected number of offspring conditional on reproduction.

Plugging all of these quantities in Equation (13), we get the expression of $\mathbf{K}$ for 1×1 models. First, for $g = 0$,

$$\mathbf{K}(0, q) \;=\; \sum_{\alpha \geq 1} \binom{\alpha-1}{q-1} s^{\alpha-q} f^q \left(\frac{s}{\lambda}\right)^{\alpha}\,. \tag{37}$$

Second, for $g \geq 1$,

$$\begin{aligned}\mathbf{K}(g, q) \;=\; \sum_{t \geq 0} \Bigg[&\binom{t-1}{q} \frac{s^{t-q} f^{q+1}}{\lambda} \;+\; \binom{t-1}{q-1} \frac{s^{t-q} f^q \tilde{f}}{\lambda} \\ &+ \sum_{\alpha \geq 1} \binom{t+\alpha-1}{q-1} s^{t+\alpha-q} f^q \left(\frac{s}{\lambda}\right)^{\alpha} \frac{f}{\lambda} \Bigg] \binom{t-1}{g-1} \frac{s^{t-g} f^{g-1}}{\lambda^{t-1}},\end{aligned} \tag{38}$$

again with the convention that $\binom{n}{k} = 0$ when $n < 0$ or $k \notin \{0, \ldots, n\}$. Rearranging a bit, this second expression can also be written

$$\mathbf{K}(g, q) \;=\; \sum_{t \geq 0} \left[\binom{t-1}{q} \;+\; \binom{t-1}{q-1} \frac{\tilde{f}}{f} + \sum_{\alpha \geq 1} \left(\frac{s^2}{\lambda}\right)^{\alpha} \right] \left(\frac{s^2}{\lambda}\right)^{t}. \tag{39}$$

SM.9 Asymptotic behavior of $\tilde{\mathbf{K}}(d)$ as $d \to \infty$

In this section, we detail the mathematical reasoning between the introduction and characterization of the statistic θ_{c} discussed in Section 2.6. Loosely speaking, θ_{c} is such that, for any matrix norm,

$$\|\tilde{\mathbf{K}}(d)\| \;\propto\; \theta_{\mathrm{c}}^{d}\,.$$

More rigorously, $\theta_c = \inf\{\theta > 0 : \sum_d \theta^{-d}\tilde{\mathbf{K}}(d) < +\infty\}$. This justifies the appellation *exponential growth rate of* $\tilde{\mathbf{K}}(d)$. Our main result is that θ_c admits the following characterization, given as Equation (18) in the main text:

$$\theta_c \;=\; \inf\left\{\theta > 0 : \rho(\mathbf{S} + \theta^{-1}\mathbf{F})\,\rho(\mathbf{P_S} + \theta^{-1}\mathbf{P_F}) < 1\right\},$$

where ρ denotes the spectral radius (here, the dominant eigenvalue, since the matrices are primitive). This characterization makes it straightforward to compute θ_c numerically (e.g, using the bisection method, as implemented in our code, see SM.1).

Although the result for 1×1 models follows as a special case of the result for $n\times n$ models, we start by treating it on its own, in full detail. This is because the expression of $\mathbf{K}(g,q)$ for 1×1 models given in SM.8 and the fact that we are dealing with numerical (instead of matricial) series make the discussion less technical, while retaining its main idea; it is thus a useful stepping-stone for the $n\times n$ case. In addition to this, the simple expression of θ_c for 1×1 models deserves a special discussion.

SM.9.1 Expression of θ_c in 1×1 models

As in SM.8, assume that the projection matrix is a 1×1 matrix $\mathbf{A} = (s+f)$, with $f > 0$ and $0 \le s < 1$, and write $\lambda = f + s$. For any $\theta > 0$, let us consider the sum

$$\sum_{d\ge 0}\theta^{-d}\tilde{\mathbf{K}}(d) \;=\; \sum_{g\ge 0}\sum_{q\ge 0}\theta^{-(g+q)}\mathbf{K}(g,q)\,,$$

this equality being true because the summands are non-negative and thus the order of summation is irrelevant (Fubini-Tonelli theorem). Our aim is to get a necessary and sufficient condition on θ for this series to be convergent.

Plugging Equations (37) and (38) into $\sum_d \theta^{-d}\tilde{\mathbf{K}}(d)$, then using again Fubini-Tonelli's theorem to change the order of summation and rearranging a bit, we get

$$\begin{aligned}
\sum_{d\ge 0}\theta^{-d}\tilde{\mathbf{K}}(d) \;=\; & \sum_{\alpha\ge 1}\sum_{q\ge 0}\binom{\alpha-1}{q-1}\left(\frac{s^2}{\lambda}\right)^{\alpha}\left(\frac{f}{\theta s}\right)^{q} \\
& + \sum_{t\ge 0}\sum_{g\ge 1}\sum_{q\ge 0}\binom{t-1}{q}\binom{t-1}{g-1}\left(\frac{s^2}{\lambda}\right)^{t}\left(\frac{f}{\theta s}\right)^{g+q} \\
& + \sum_{t\ge 0}\sum_{g\ge 1}\sum_{q\ge 0}\binom{t-1}{q-1}\binom{t-1}{g-1}\left(\frac{s^2}{\lambda}\right)^{t}\left(\frac{f}{\theta s}\right)^{g+q}\tilde{f}/f \\
& + \sum_{t\ge 0}\sum_{\alpha\ge 1}\sum_{g\ge 1}\sum_{q\ge 0}\binom{t+\alpha-1}{q-1}\binom{t-1}{g-1}\left(\frac{s^2}{\lambda}\right)^{t+\alpha}\left(\frac{f}{\theta s}\right)^{g+q}.
\end{aligned}$$

Let us start by focusing on the first term of the right-hand side. Recalling that we use the convention $\binom{n}{k} = 0$ when $n < 0$ or $k \notin \{0,\dots,n\}$ and using the binomial theorem, we have, for every $\alpha \ge 1$,

$$\sum_{q\ge 0}\binom{\alpha-1}{q-1}\left(\frac{s^2}{\lambda}\right)^{\alpha}\left(\frac{f}{\theta s}\right)^{q} \;=\; \tfrac{f}{\theta s}\left(1+\tfrac{f}{\theta s}\right)^{-1}\left(\frac{(s+f/\theta)s}{\lambda}\right)^{\alpha}.$$

Therefore,

$$\sum_{q\ge 0}\theta^{-q}\mathbf{K}(0,q) < +\infty \iff (s+f/\theta)s < \lambda\,. \tag{40}$$

Finally, since $\lambda = s + f$,

$$(s+f/\theta)s = s+f \iff \theta = \frac{f}{1-s+f/s}\,.$$

This means that, as q increases, the number of q-th generation descendants of *ego* grows roughly like γ^q, where $\gamma = \frac{f}{1-s+f/s}$. Unsurprisingly, this exponential rate of increase γ is exactly the expected number of daughters of *ego* that are currently alive, that is,

$$\mathbf{K}(0,1) \;=\; \frac{f}{1-s+f/s}\,, \tag{41}$$

as can be checked from Equation (37).

Let us now turn to the second term of the right-hand side of the expression of $\sum_d \theta^{-d}\tilde{\mathbf{K}}(d)$. Again, for every $t \geq 1$,

$$\begin{aligned}
&\sum_{q\geq 0}\sum_{g\geq 1}\binom{t-1}{q}\binom{t-1}{g-1}\left(\frac{s^2}{\lambda}\right)^t\left(\frac{f}{\theta s}\right)^{g+q} \\
&\qquad = \frac{f}{\theta s}\times\left(\sum_{g\geq 1}\binom{t-1}{g-1}\left(\frac{f}{\theta s}\right)^{g-1}\right)\times\left(\sum_{q\geq 0}\binom{t-1}{q}\left(\frac{f}{\theta s}\right)^{q}\right)\times\left(\frac{s^2}{\lambda}\right)^t \\
&\qquad = \frac{f}{\theta s}\left(1+\frac{f}{\theta s}\right)^{2(t-1)}\left(\frac{s^2}{\lambda}\right)^t .
\end{aligned}$$

Therefore,

$$\sum_{t\geq 0}\sum_{q\geq 0}\sum_{g\geq 1}\binom{t-1}{q}\binom{t-1}{g-1}\left(\frac{s^2}{\lambda}\right)^t\left(\frac{f}{\theta s}\right)^{g+q} \;=\; \tfrac{f}{\theta s}\left(1+\tfrac{f}{\theta s}\right)^{-2}\sum_{t\geq 1}\left(\frac{(s+f/\theta)^2}{\lambda}\right)^t ,$$

which converges if and only if $(s+f/\theta)^2 < \lambda$. Similarly,

$$\sum_{t\geq 0}\sum_{q\geq 0}\sum_{g\geq 1}\binom{t-1}{q-1}\binom{t-1}{g-1}\left(\frac{s^2}{\lambda}\right)^t\left(\frac{f}{\theta s}\right)^{g+q}\frac{\tilde{f}}{f} \;=\; C_1\sum_{t\geq 1}\left(\frac{(s+f/\theta)^2}{\lambda}\right)^t$$

for some irrelevant constant C_1, and

$$\begin{aligned}
&\sum_{t\geq 0}\sum_{\alpha\geq 1}\sum_{q\geq 0}\sum_{g\geq 1}\binom{t+\alpha-1}{q-1}\binom{t-1}{g-1}\left(\frac{s^2}{\lambda}\right)^{t+\alpha}\left(\frac{f}{\theta s}\right)^{g+q} \\
&= C_2\sum_{t\geq 0}\sum_{\alpha\geq 1}\left(\frac{(s+f/\theta)^2}{\lambda}\right)^t\left(\frac{(s+f/\theta)s}{\lambda}\right)^{\alpha}
\end{aligned}$$

for some constant C_2. This last series converges if and only if $(s+f/\theta)^2 < \lambda$ and $(s+f/\theta)s < \lambda$; but, since $s < s + f\theta$, these two conditions are equivalent to $(s+f/\theta)^2 < \lambda$.

Putting the pieces together, we see that

$$\sum_{d\geq 0}\theta^{-d}\tilde{\mathbf{K}}(d) \;<\; +\infty \;\iff\; (s+f/\theta)^2 < \lambda = f+s\,. \tag{42}$$

Again, loosely speaking, this means that $\tilde{\mathbf{K}}(d)$ grows like θ_{c}^d, where

$$\theta_{\mathrm{c}} \;=\; \frac{f}{\sqrt{f+s}-s}\,. \tag{43}$$

(note that θ_{c} is always well-defined and positive because $f > 0$ and $0 \leq s < 1$). Interestingly, whenever $s \ll f$, be it because $s \ll 1$ or because $f \gg 1$, $\theta_{\mathrm{c}} \approx \sqrt{f}$ does not depend on s. Also note that θ_{c} can be written

$$\theta_{\mathrm{c}} \;=\; \frac{\lambda - s}{\sqrt{\lambda}-s}\,, \tag{44}$$

which is useful because it shows that:

(i) As expected, $\theta_{\mathrm{c}} > 1$ if and only if $\lambda > 1$.

(ii) For a fixed value of λ, θ_{c} is an increasing function of s when $\lambda > 1$ and a decreasing function of s when $\lambda < 1$.

In particular, the second point reflects the intuitive fact that, if there is an advantage to having closely related kin, then given two mutants with the same growth rate, the mutant with the higher survival is expected to be favored – in other words, that K strategies are expected to evolve when close kin help each other. What is perhaps surprising is that this tendency is reversed when the population is decreasing. It should be noted, however, that our setting assumes that the population has been at the demographic steady state for a long period of time, which undermines the relevance of our formulas when $\lambda < 1$.

SM.9.2 Characterization of θ_c in $n{\times}n$ models

Despite additional technicalities, the case of $n{\times}n$ models is very similar to that of $1{\times}1$ ones. Here the crucial observation is that

$$\sum_{g\geq 0}\theta^{-g}\mathbf{U}(t,g) \;=\; \theta^{-1}\sum_{g\geq 0}\sum_{\mathbf{k}\in\mathcal{C}(t-g,g)}\mathbf{P}_{\mathbf{S}}^{k_1}\,(\theta^{-1}\mathbf{P}_{\mathbf{F}})\cdots(\theta^{-1}\mathbf{P}_{\mathbf{F}})\,\mathbf{P}_{\mathbf{S}}^{k_g} \;=\; \theta^{-1}(\mathbf{P}_{\mathbf{S}}+\theta^{-1}\mathbf{P}_{\mathbf{F}})^t\,. \tag{45}$$

Similarly,

$$\sum_{q\geq 0}\theta^{-q}\mathbf{D}(t,q) \;=\; \begin{cases}\mathbf{0} & \text{if } t=0\\ (\mathbf{S}+\theta^{-1}\mathbf{F})^{t-1} & \text{for } t\geq 1.\end{cases} \tag{46}$$

Let us focus on the growth of the term $\mathbf{K}_{\mathrm{y}}$ in $\mathbf{K}=\mathbf{K}_{\mathrm{y}}+\mathbf{K}_{\mathrm{s}}+\mathbf{K}_{\mathrm{o}}$, that is,

$$\mathbf{K}_{\mathrm{y}}(g,q) \;=\; \sum_{t\geq 0}\mathbf{D}(t,q)\,\mathbf{S}\,\mathbf{P}_{\mathbf{F}}^{\top}\mathbf{U}(t,g)^{\top},$$

as given by Equation (6). Using the same reasoning as for $1{\times}1$ models, but with Equations (45) and (46) instead of the binomial theorem, we get

$$\begin{aligned}\sum_{d\geq 0}\theta^{-d}\tilde{\mathbf{K}}_{\mathrm{y}}(d) \;&=\; \sum_{g\geq 0}\sum_{q\geq 0}\theta^{-(g+q)}\mathbf{K}_{\mathrm{y}}(g,q)\\ &=\; \sum_{t\geq 0}\left(\sum_{q\geq 0}\theta^{-q}\mathbf{D}(t,q)\right)\mathbf{S}\,\mathbf{P}_{\mathbf{F}}^{\top}\left(\sum_{g\geq 0}\theta^{-g}\mathbf{U}(t,g)\right)^{\top}\\ &=\; \theta^{-1}\sum_{t\geq 1}\left(\mathbf{S}+\theta^{-1}\mathbf{F}\right)^{t-1}\mathbf{S}\,\mathbf{P}_{\mathbf{F}}^{\top}\left(\left(\mathbf{P}_{\mathbf{S}}+\theta^{-1}\mathbf{P}_{\mathbf{F}}\right)^{t}\right)^{\top}.\end{aligned} \tag{47}$$

Now, observe that the matrices $\mathbf{A}_\theta=\mathbf{S}+\theta^{-1}\mathbf{F}$ and $\mathbf{P}_\theta=\mathbf{P}_{\mathbf{S}}+\theta^{-1}\mathbf{P}_{\mathbf{F}}$ are both primitive (because $\mathbf{A}=\mathbf{S}+\mathbf{F}$, and thereby $\mathbf{P}=\mathbf{P}_{\mathbf{S}}+\mathbf{P}_{\mathbf{F}}$, are). As a result, letting $\rho(\,\cdot\,)$ denote their dominant eigenvalue, as $t\to\infty$,

$$\mathbf{A}_\theta^t \;\sim\; \rho(\mathbf{A}_\theta)^t\,\mathbf{C}_{\mathbf{A}_\theta} \quad\text{and}\quad \mathbf{P}_\theta^t \;\sim\; \rho(\mathbf{P}_\theta)^t\,\mathbf{C}_{\mathbf{P}_\theta} \tag{48}$$

for some irrelevant matrices $\mathbf{C}_{\mathbf{A}_\theta}$ and $\mathbf{C}_{\mathbf{P}_\theta}$. Combining this with (47), we get

$$\begin{aligned}\sum_{d\geq 0}\theta^{-d}\tilde{\mathbf{K}}_{\mathrm{y}}(d)<+\infty \;&\iff\; \sum_{t\geq 1}\big(\rho(\mathbf{A}_\theta)\,\rho(\mathbf{P}_\theta)\big)^t<+\infty\\ &\iff\; \rho(\mathbf{A}_\theta)\,\rho(\mathbf{P}_\theta)\;<\;1\,.\end{aligned}$$

The growth of the matrices $\mathbf{K}_{\mathrm{s}}$ and $\mathbf{K}_{\mathrm{o}}$ is treated similarly and yields the same criterion for convergence, completing our charaterization of θ_c.

Note that if $\mathbf{A}=(s+f)$ is a $1{\times}1$ matrix, then $\rho(\mathbf{A}_\theta)=s+f/\theta$ and $\rho(\mathbf{P}_\theta)=\frac{s+f/\theta}{s+f}$ and so we fallback on Equation (42).

Finally, because $f\colon\theta\mapsto\rho(\mathbf{A}_\theta)\rho(\mathbf{P}_\theta)$ is increasing in θ and easily evaluable numerically, this makes the numerical determination of θ_c straightforward and efficient. Also note that to get an initial estimate of θ_c it si possible to "collapse" $\mathbf{A}$ into a $1{\times}1$ model, as explained in Bienvenu et al. (2017), and then to use Equation (44). Using individualistic collapsing yields

$$\hat{\mathbf{A}}=\big(\hat{s}+\hat{f}\big),\quad \text{with } \hat{s}=\mathbf{1Sw} \text{ and } \hat{f}=\mathbf{1Fw}\,, \tag{49}$$

where $\mathbf{1}=(1,\ldots,1)$, whereas using genealogical collapsing – which is expected to be more accurate for the estimation of most biological descriptors – yields

$$\mathbf{A}^\star=(s^\star+f^\star),\quad \text{with } s^\star=\mathbf{vSw} \text{ and } f^\star=\mathbf{vFw}\,, \tag{50}$$

where the vector of reproductive values $\mathbf{v}$ is scaled in such a way that $\mathbf{vw}=1$.

SM.10 Computing time of $\widehat{\mathbf{K}}(g,q\,;\,t_{\max},\alpha_{\max})$

The computing time of the matrix $\widehat{\mathbf{K}}(g,q\,;\,t_{\max},\alpha_{\max})$, as given by Equation (15), increases rapidly as a function of g, q, $t_{\max}$ and $\alpha_{\max}$, and can be prohibitively long in practice. To help assess this, in this section

we provide rough estimates of the time complexity of a standard implementation of $\widehat{\mathbf{K}}(g, q\,;\, t_{\max}, \alpha_{\max})$.

SM.10.1 Setting and notation

We denote by $\mathcal{T}(\texttt{expr})$ the number of elementary operations required to compute the expression `expr`, where by "elementary operation" we refer to: matrix addition; matrix multiplication; matrix transposition; and matrix power. All of these operations are assumed to a have a $\Theta(1)$ time complexity, independently of the size n of the matrix, which we assume to be fixed. Should large matrices be studied, this $\Theta(1)$ could be replaced by the complexity of the matrix product, that is, $\Theta(n^x)$ for some $x \in [2, 3]$.

In reality, the time complexity of the matrix power $\mathbf{A}^k$ is $\Theta(\log k)$; but since taking this into account has the effect of adding a logarithmic factor $\Theta(\log(t_{\max} + \alpha_{\max}))$ and that, for all practical purposes, this factor can be assumed to be bounded, we use $\Theta(1)$ for simplicity.

Finally, to alleviate the notation, we write $f(t_{\max}, \alpha_{\max}) \asymp \texttt{expr}$ to indicate that $f(t_{\max}, \alpha_{\max}) = \Theta(\texttt{expr})$ as $t_{\max}$ and $\alpha_{\max}$ increase. Note that because we are interested in what happens not only when g and q are fixed but also when they increase with $t_{\max}$ and $\alpha_{\max}$, we will implicitly treat them as functions of $t_{\max}$ and $\alpha_{\max}$, rather than as constants.

SM.10.2 Calculation

First, observe that, since $\mathcal{T}(\mathbf{P}_{\mathbf{S}}^{k_1}\,\mathbf{P}_{\mathbf{F}} \cdots \mathbf{P}_{\mathbf{F}}\,\mathbf{P}_{\mathbf{S}}^{k_g}) \asymp g$ and $\#\mathcal{C}(t-g, g) = \binom{t-1}{g-1}$, for $g \geq 1$ we have

$$\mathcal{T}(\mathbf{U}(t, g)) \;\asymp\; g\binom{t-1}{g-1}. \tag{51}$$

Similarly, for $q \geq 1$,

$$\mathcal{T}(\mathbf{D}(t, g)) \;\asymp\; q\binom{t-1}{q}. \tag{52}$$

Now, recall Equation (15):

$$\begin{aligned}\widehat{\mathbf{K}}(g, q\,;\, t_{\max}, \alpha_{\max}) = \sum_{t=0}^{t_{\max}} \Bigg(&\mathbf{D}(t, q)\,\mathbf{S}\,\mathbf{P}_{\mathbf{F}}^{\top} \;+\; \mathbf{D}(t, q-1)\,\mathbf{Z} \\ &+ \sum_{\alpha=1}^{\alpha_{\max}} \mathbf{D}(t+\alpha, q-1)\,\mathbf{F}\,(\mathbf{P}_{\mathbf{S}}^{\top})^{\alpha}\,\mathbf{Y}(g)\Bigg)\,\mathbf{U}(t, g)^{\top}.\end{aligned}$$

Focusing on the case $g \geq 1, q \geq 1$, we get

$$\begin{aligned}&\mathcal{T}\big(\widehat{\mathbf{K}}(g, q\,;\, t_{\max}, \alpha_{\max})\big) \\ &\asymp \sum_{t=0}^{t_{\max}} \Bigg(\mathcal{T}\big(\mathbf{D}(t, q)\big) \;+\; \mathcal{T}\big(\mathbf{D}(t, q-1)\big) \;+\; \sum_{\alpha=1}^{\alpha_{\max}} \mathcal{T}\big(\mathbf{D}(t+\alpha, q-1)\big) \;+\; \mathcal{T}\big(\mathbf{U}(t, g)\big)\Bigg) \\ &\asymp \sum_{t=0}^{t_{\max}} \Bigg(q\binom{t-1}{q} \;+\; (q-1)\binom{t-1}{q-1} \;+\; \sum_{\alpha=1}^{\alpha_{\max}} (q-1)\binom{t+\alpha-1}{q-1} \;+\; g\binom{t-1}{g-1}\Bigg) \\ &\asymp q\sum_{t=0}^{t_{\max}}\sum_{\alpha=1}^{\alpha_{\max}} \binom{t+\alpha-1}{q-1} \;+\; g\sum_{t=0}^{t_{\max}} \binom{t-1}{g-1} \\ &\asymp (t_{\max} + \alpha_{\max} - q)\binom{t_{\max} + \alpha_{\max} + 1}{q} \;+\; (t_{\max} - g)\binom{t_{\max}}{g-1}.\end{aligned} \tag{53}$$

Let us focus on two specific cases: first, assume that g and q are constant. This is the most relevant scenario for practical applications, where g and q will typically be very small. In that case, as $t_{\max}$ and $\alpha_{\max}$ go to infinity,

$$\mathcal{T}\big(\widehat{\mathbf{K}}(g, q\,;\, t_{\max}, \alpha_{\max})\big) \;\asymp\; (t_{\max} + \alpha_{\max})^{q+1}\, t_{\max}^{g}. \tag{54}$$

In particular, if $t_{\max}$ and $\alpha_{\max}$ are chosen according to Equation (16), as a function of a maximal age ω (be it a "true" maximal age built in the model, or an inferred one used in the computation of $\widehat{\mathbf{K}}$), then

$$\mathcal{T}\big(\widehat{\mathbf{K}}(g, q\,;\, t_{\max}, \alpha_{\max})\big) \;\asymp\; \big(m\,\omega\big)^{g+q+1}\big(1 - 1/m\big)^{g}, \tag{55}$$

with $m = \min(g, q{+}1){+}1$. Assuming that the computing time start being prohibitive after 10^9 elementary

operations (which, of course, is going to be highly dependent on the specificities of the implementation of $\widehat{\mathbf{K}}$ and on the computing power available), Table 2 below gives the maximal value of ω for which the calculations can be performed as a function of g and q, based on Equation (55).

q \ g	0	1	2	3	4	5	6	7
0	—	31622	999	177	63	31	19	13
1	22360	629	105	36	17	10	7	5
2	793	72	24	12	7	4	3	2
3	149	26	9	5	3	2	2	1
4	54	13	5	2	2	1	1	1
5	28	8	3	2	1	1	1	0
6	17	6	3	1	1	0	0	0
7	12	4	2	1	1	0	0	0

Table 2: Largest value of ω for which $(m\,\omega)^{g+q+1}(1-1/m)^g < 10^9$, with $m = \min(g, q+1)+1$. This gives a rough idea of the values of g and q for which $\mathbf{K}(g,q)$ can be estimated, as function of the maximal age ω in the life cycle. More precise estimations can be obtained, e.g., using Equation (53) or even one of the more precise estimates that precede it.

Finally, let us focus on the worst-case scenario and assume that $q = \lfloor (t_{\max} + \alpha_{\max})/2 \rfloor$. In that case, Equation (53) gives

$$\mathcal{T}\big(\widehat{\mathbf{K}}(g,q\,;\,t_{\max},\alpha_{\max})\big) \asymp \sqrt{t_{\max}+\alpha_{\max}}\;2^{t_{\max}+\alpha_{\max}}\,. \tag{56}$$

Of course, the same reasoning shows that the time complexity is $\sqrt{t_{\max}}\,2^{t_{\max}}$ when $g = \lfloor t_{\max}/2 \rfloor$. This shows that there is no hope to be able to compute $\mathbf{K}(g,q)$ for large values of g and/or q. However, note that some information about the large-scale structure of the kinship can still be obtained through the statistics θ_{c} discussed in SM.9.

SM.11 Matrilineal kinship and relatedness

In this section, we briefly explain why – under some fairly stringent hypotheses – the matrilineal notion of kinship that we consider (that is, the expected of kin of *ego* in the genealogical tree of females) coincides with the usual notion of relatedness in population genetics (that is, the expected number of genes of *ego* that are identical by descent with those of its kin, including males) – and why it can, under more acceptable hypotheses, be seen as a lower bound on relatedness.

Consider the idealized setting of an infinite panmictic population where there is no difference in survival or reproduction between males and females, and where exactly one offspring is produced at each reproductive event (so that no two individuals share the same two parents). In that setting, *ego* has exactly 2^g distinct g-th generation ancestors in its complete pedigree – as opposed to one through matrilines only. Because the population is infinite and panmictic, these ancestors do not have shared descendants. Thus, the number of individuals that are q-th generation descendants of at least one of the g-th ancestors of *ego* is on average 2^g times the expected number of q-th descendants of any of those ancestors. Now, since there is no difference in survival and reproduction between males and females, the expected number of q-th descendants of an individual is 2^q times its number of q-th descendants through matrilines only. As a result, the expected number of (g,q)-kin in the complete pedigree of *ego* is 2^{g+q} times its number of matrilineal (g,q)-kin. Finally, since the relatedness of each of these (g,q)-kin with *ego* is $2^{-(g+q)}$, we see that the total amount of relatedness to *ego* that comes from its (g,q)-kin is, on average, exactly its number of matrilineal kin.

Of course, none of the assumptions on which this reasoning relies hold in the real world: populations are neither infinite nor panmictic, and a pair of individuals can give birth to more than one offspring – so that the pedigrees of the (g,q)-kin of *ego* will overlap. However, this overlap in itself is not a problem, as long as the expected number of descendants of each individual is the same. Indeed, assume for instance that two of the 2^g g-th ancestors of *ego* are, in fact, the same individual. Then, the contribution of those ancestors to the (g,q)-kin of *ego* is divided by two, but this is compensated by the fact that each of these individuals is twice as related to *ego* as if the two g-th ancestors had been different individuals. This reasoning can be generalized to the whole portion of the pedigree required to find the (g,q)-kin of *ego*.

Another consequence of the fact that the population is not infinite and panmictic is that the g-th ancestors of *ego*, even if they are all distinct, are bound to have some degree of relatedness that should be taken into account in an exact calculation of the relatedness, whereas we are ignoring it here. In other words, neglecting the relatedness of the least common ancestors of *ego* and its (g,q)-kin means

that we tend to underestimate the true relatedness, and this effect is more pronounced in consanguineous populations.

Finally, the effect of a difference in survival between males and females is more complex, and is bound to non-trivially affect the number of kin of *ego*. However, if the sex-ratio is 1:1, then for this to remain constant over time males and females should have, on average, the same number of offspring across the course of their life. Therefore, even if the distribution of the number of (g, q)-kin of *ego* is modified, its expected value should not change – and thus neither does their contribution to the relatedness.

Altogether, these considerations suggest that the number of matrilineal kin computed by our formulas can be seen as a direct measure of relatedness in population with low consanguinity, and as a lower bound on relatedness in more consanguineous populations.

SM.12 Comparison to individual-based simulations for the ground squirrel *Spermophilus dauricus*

To illustrate the validity of our formulas, we compare them to kinship matrices estimated from individual-based simulations. For this, we use the life-cycle of the ground squirrel *Spermophilus dauricus* presented in Box 2.

The simulations were done in R using the package NetLogoR (Bauduin et al., 2019), and correspond to a female-based population. At each time step, the individuals went through the following events, successively:

- **Reproduction:** the number of newborns produced by each female in stage i is drawn from a Poisson distribution with parameter m_i, where m_i is the expected number of (female) newborns produced by each (female) individual. The newborns are temporarily put in an auxiliary stage $i = 0$ that does not correspond to a class of the matrix population model.
- **Survival:** each individual in stage i survives with probability s_i, independently of the other individuals, and enters stage $i + 1$ (individuals in stage $i = 3$ remained in that stage). That is, conditional on the number n_i of individuals in stage i before the event, the number of individuals in stage $i + 1$ after the event is a binomial variable with parameters (n_i, s_i).

To run the simulations, we chose a maximum number of time-steps ensuring that, with high probability, for the values of (g, q) considered the simulated genealogy would contain the g-th and q-th ancestor of every individual alive at the end of the simulation. Here, using $T_{\max} = 50$ times steps was sufficient.

Next, we chose an initial population ensuring that, with high probability, the population would not go extinct before $T_{\max}$. Here we started with 1010 individuals at the stable stage distribution, that is, with an initial population vector $\mathbf{n} = (470, 270, 270)$.

Finally, we ran 300 independent replicates of the population process. For each of those replicates, at the last time step, we computed the empirical kinship matrices for mothers and aunts:

- $\mathbf{K}_{\text{emp}}(1, 0)$, whose (i, j)-th entry is fraction of individuals from class j whose mother is currently alive in stage i.
- $\mathbf{K}_{\text{emp}}(2, 1)$, whose (i, j)-th entry is per-capita number of aunts alive in stage i for individuals in class j.

We averaged the empirical kinship matrices over all 300 replicates and compared their average $\bar{\mathbf{K}}_{\text{emp}}$ to the expected value $\mathbf{K}$ given by the formulas of the main text. For mothers, we have

$$\bar{\mathbf{K}}_{\text{emp}}(1, 0) - \mathbf{K}(1, 0) = \begin{pmatrix} 0 & 0 & 0 \\ 9 \cdot 10^{-4} & 0 & 0 \\ -4 \cdot 10^{-4} & 0.0016 & 0.001 \end{pmatrix},$$

and for aunts we have

$$\bar{\mathbf{K}}_{\text{emp}}(2, 1) - \mathbf{K}(2, 1) = \begin{pmatrix} -0.0050 & 0.0010 & 0 \\ 0.0018 & -0.0010 & -7 \cdot 10^{-4} \\ 0.0018 & 0.0031 & -5 \cdot 10^{-4} \end{pmatrix}.$$

The empirical distribution of the sum of the columns of these matrices (that is the probability that the mother of *ego* is alive and its expected number of aunts, structured by the stage of *ego*) are depicted in Figures 5 and 6.

The code used in this Supplementary Material and the full output of the simulations can be found at `https://github.com/SCubaynes/Squirrel_IBM.git`.

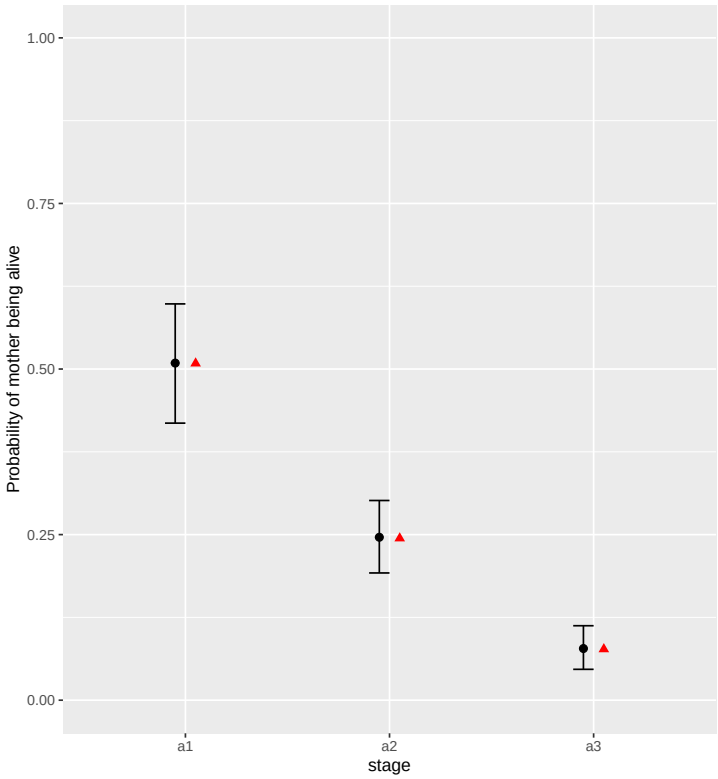


Figure 5: The empirical distribution of the probability that the mother of *ego* is alive, structured by the stage of *ego*. The black segments correspond to the 95% quantiles of the empirical distribution, the back dots to the empirical average, and the red triangles to the theoretical expected values given by our formulas.

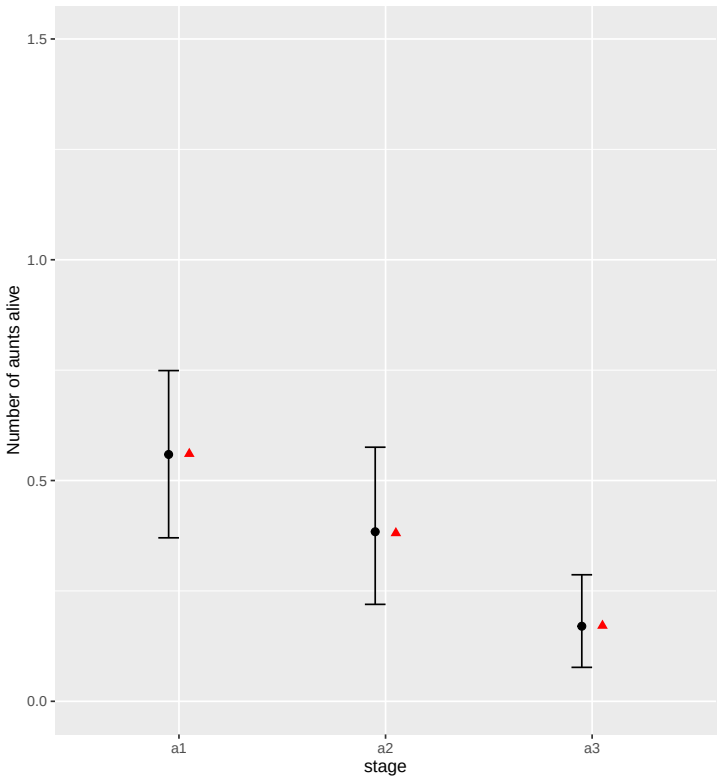


Figure 6: The empirical distribution of the average number of aunts of *ego*, structured by the stage of *ego*. The black segments correspond to the 95% quantiles of the empirical distribution, the back dots to the empirical average, and the red triangles to the theoretical expected values given by our formulas.