The costs of reproduction in evolutionary demography : an application of Multitrait Population Projection Matrix models

Christophe Coste

Thèse de doctorat de l'Université Sorbonne Paris Cité
Préparée à l'Université Paris Diderot

Ecole Doctorale - Frontières du Vivant (ED 474)

Chaire - Modélisation Mathématique et Biodiversité

Laboratoire d'Éco-anthropologie et Ethnobiologie (UMR 7206)
Muséum National d'Histoire Naturelle - Musée de l'Homme

Thèse de doctorat

discipline : Mathématiques appliquées et sciences sociales

Spécialité : Eco-Evolution Mathématique

présentée par

Christophe Coste

The costs of reproduction in evolutionary demography:

An application of Multitrait Population Projection Matrix models

---

dirigée par Frédéric Austerlitz et Samuel Pavard

Présentée et soutenue publiquement (sous réserve d'acceptation des rapporteurs), au Musée de l'Homme, le 20 novembre 2017, devant le jury composé de :

M. Jean-Michel Gaillard	Directeur de recherche(CNRS)	rapporteur
M. Jacob Moorad	Lecturer (University of Edinburgh)	rapporteur
Mme Emmanuelle Porcher	Professeure (MNHN)	examinatrice
M. Laurent Lehmann	Professeur (Université de Lausanne)	examinateur
M. Bernard Cazelles	Professeur (UPMC)	examinateur
M. Frédéric Austerlitz	Directeur de recherche (CNRS)	directeur de thèse
M. Samuel Pavard	Maitre de coférence (MNHN)	co-directeur de thèse

Laboratoire d'Éco-anthropologie et Ethnobiologie
Muséum National d'Histoire Naturelle
UMR 7206
Musée de l'Homme
17 place du Trocadéro
75016 Paris

École doctorale Frontières du vivant ED 474
Faculté de Médecine Paris Descartes
CRI, 24 rue du faubourg Saint Jacques, 75014 PARIS

Chaire Modélisation Mathématique et Biodiversité
École polytechnique, 91128 Palaiseau

Acknowledgments et remerciements

I first wish to thank Jean-Michel Gaillard and Jacob Moorad for having accepted to report on my thesis, and Laurent Lehmann, Emmanuelle Porcher and Bernard Cazelles to examine it. I really appreciate you taking the time to consider my work.

And second, I thank my two (wonderful) supervisors Frédéric Austerlitz and Samuel Pavard, for their help, direction and advice during this PhD.

I have been very lucky to be able to work on this thesis. First, I have been very fortunate to be funded by the "chaire MMB". Je remercie particulièrement Sylvie Méléard pour sa confiance, ainsi que Pierre-Henri Gouyon et Vincent Bansaye, qui faisaient partie de mon comité de thèse, pour leurs conseils.

I have also been lucky to be part of the "Frontières du Vivant" doctoral school. I have learned a lot in its courses, clubs and workshops. Special thanks to David Tareste for his help and Stéphane Douady for the mathematical fun !

Ecologically speaking, I am also indebted to Minus van Baalen for very valuable advice.

Life historically speaking, il me faut aussi remercier l'équipe de biométrie de Lyon (je ne sais pas si je dois vous mentionner ici ou parmi les amis ?) qui n'a pas hésité à prendre du temps pour partager son savoir. En particulier, merci Jean-Michel, merci Jeff.

Evolutionary demographically speaking, I wish to deeply thank Tulja(purkar), Uli (Steiner) and Tim Coulson for precious pieces of advice early in my PhD.

J'ai, enfin, eu la chance de pouvoir faire cette thèse au sein du Musée de l'Homme. Je remercie tout particulièrement les membres de l'UMR7206, dont j'ai été fier de faire partie, et, en particulier, son directeur Serge Bahuchet. Merci à Evelyne et Fred de m'avoir accueilli dans leur équipe, et tous les autres chercheuses, chercheurs de l'équipe pour de nombreux excellents moments, et quelques vols perdus. J'ai énormément apprécié de passer ces années en votre compagnie et j'espère pouvoir, dans le futur, collaborer avec nombre d'entre vous. J'ai aussi une petite pensée pour les membres émérites de mon ancien bureau au jardin des plantes - Bernard Dupaigne et Pierre Darlu - et pour leurs stimulantes discussions.

Surtout, je remercie mes compagnons de galère, les doctorantes et doctorants. Celles qui ont fini, Elise (la reine du swing), Julie, Florence, etc. Celles et ceux qui sont en train de terminer, Nina, Valentin, Goki , Vanessa (si si!) et ceux pour qui le port n'est pas encore en vue mais presque: Bérénike, Lou, Jérémy et Emma.

La thèse c'est super, mais aussi c'est long, et surtout, c'est parfois difficile. Je n'y serais pas arrivé sans mes amis, et parmi eux, tout particulièrement, Hélène, Lulla, Laure, Julia, Rita, Jessica, le Nikus (aka le curd), Tibo Lakish, Derf Duvaltin, Fredo Dupch, Coco, Zrah, Flora, Fadi Latrumpet, Seb et Delph. Merci du fond du cœur (petite larme, snif !)

Ni sans ma famille, les frangins et les frangines, les beaux-frères et belles-soeurs, les neveux et les nièces. Je cite tout particulièrement Manon, Thibaut, Clément, Loulou et Hugo.

Je remercie surtout mes parents, Christian et Michèle Coste, qui m'ont hébergé, nourri et soutenu pendant les moments difficiles, et dieu sait qu'il y en a eu. Merci pour votre patience et votre aide infiniment précieuse.

Last but not least, merci à toi petit Théo pour ton amour indéfectible. Vu ton adoration pour Lewis Carroll et les aventures de sa petite Alice, tu pourrais bien un jour, je l'espère, devenir un grand biologiste de l'évolution !

Merci aussi aux grands anciens qui ont essayé de comprendre la complexité de la nature - les Darwin, Williams, Hamilton - sans qui non plus cette thèse n'existerait pas mais que je ne pourrai jamais remercier de vive voix. Merci aussi à ceux dont je me suis inspiré mais que je n'ai pas (encore) rencontré: Ronald Lee, Brian Charlesworth, et Stephen C Stearns dont la théorie m'a permis de mieux comprendre ce qui sous-tend la diversité du monde naturel qui nous entoure. C'est donc aussi dame Nature qu'il faut que je remercie. Grâce à une multiplicité de positionnements d'espèces sur une multitude de trade-offs, elle a généré la biodiversité inégalée des montagnes du Jura et en a fait le havre de paix et de beauté qui m'a permis d'écrire une bonne partie de ces lignes.

Enfin, mon remerciement le plus important est pour toi Sam. Sans ta patience, ton soutien, ton enthousiasme, cette thèse n'aurait jamais vu le jour.

Paris, le 25 Septembre 2017

Contents

The costs of reproduction in evolutionary demography: An application of Multitrait Population Projection Matrix models

Costs of reproduction are pervasive in life history theory. Through this constraint, the reproductive effort of an organism at a given time negatively affects its later survival and fertility. For life historians, they correspond mostly to a physiological trade-off that stems from an allocative process, occurring at each time-step, at the level of the individual. For evolutionary demographers, they are essentially about genetic trade-offs, arising from a genetic variance in a pleiotropic gene acting antagonistically on early-age and late-age fitness components. The study, from an evolutionary demographic standpoint, of these mechanisms and of the relative, cross and joint effects of physiological and genetic costs, is the aim of this thesis.

The close examination of Williams (1966)'s original definition of the physiological costs of reproduction led us to produce a theoretical design of their apparatus that accounts for both their mechanistic and evolutionary mechanisms. This design allowed us to make predictions with regards to the strength of costs of reproduction for various positions of organisms on three life-history spectra: slow-fast, income-capital breeders and quality-quantity.

From Stearns (1989b)'s tryptic architecture of life history trade-offs –that divides their structure into the genotypic level, the intermediate structure and the phenotypic level – we devised a general framework, which models the possible cohabitation of both physiological and genetic costs. From this, we inferred differing detectability patterns of both types of costs according to the environmental conditions, their variance and individual stochasticity. We could also establish that both costs buffer environmental variations, but with varying time windows of effect. Their dissimilarity emerges also from the differences between mathematical projection models specific to each cost. A new family of evolutionary models is therefore required to implement both physiological and genetic trade-offs.

We then describe the vector-based construction method for such a model which we call Multitrait Population Projection Matrix (MPPM) and which allows incorporating both types of costs by embedding them as traits into the matrix. We extend the classical sensitivity analysis techniques of evolutionary demography to MPPMs. Most importantly, we present a new analysis tool for both life history and evolutionary demography: the *Trait Level Analysis*. It consists in comparing pairs of models that share the same asymptotic properties. Such ergodic equivalent matrices are produced by *folding*, an operation that consists in reducing the number of traits of a multi-trait model, by averaging transitions for the traits *folded* upon, whilst still preserving the asymptotic flows. The *Trait Level Analysis* therefore allows, for example, to measure the evolutionary importance of costs of reproduction by comparing models incorporating them with *folded* versions of these models from which the costs are absent.

Using classical and new methods to compute fitness moments – selection gradient, variance in reproductive success, environmental variance - in models with and without the costs, we can show their effects on various demographic and evolutionary measures. We reveal, in this way, the combined effects of genetic and physiological costs on the vital rates of an age-structured population. We also demonstrate how physiological costs affect both components of effective selection, as they flatten the slope of selection gradients and increase the effective size of a population. Finally, we show how their buffering of environmental and demographic variance confer greater resilience to populations experiencing physiological costs of reproduction.

Lastly, we hint at the extension of such a multitrait model towards a new evolutionary demographic field, studying the coevolution of kinship distributions and biodemography, we call *Kinship Demography*. *Kinship demography* considers both the effects of kinship distribution on the demography of the population, which occur via the intra- and intergenerational transfers between kin, and the reciprocal influence of vital rates on the distribution of kin.

Keywords: Evolutionary demography, Life-history theory, Multitrait Population Projection Matrix, MPPM, Trait-level analysis, Kinship models, Physiological trade-off, Genetic trade-off, Costs of reproduction.

Les coûts de la reproduction en démographie évolutive:
Une application des modèles de Matrices de Projection de Population Multitrait

Les coûts de la reproduction sont un compromis biologique (*trade-off*) fondamental en théorie des histoires de vie. Par ce compromis, le succès, pour un organisme, d'un évènement de reproduction réduit sa survie et sa fertilité futures. Pour les écologues, ce *trade-off* correspond principalement à un compromis physiologique résultant d'un processus d'allocation ayant lieu à chaque instant et au niveau de chaque individu. Au contraire, en démographie évolutive, il est envisagé comme un *trade-off* génétique découlant du polymorphisme génotypique d'un gène pléiotropique agissant de manière antagoniste sur la reproduction aux jeunes âges et la fitness aux âges élevés. L'étude des mécanismes des coûts de la reproduction, physiologiques et génétiques, de leur possible cohabitation et de leur effets relatifs, croisés et conjoints est le sujet de cette thèse.

Un examen attentif de la définition originelle des coûts de la reproduction par Williams (1966), nous permet de construire un modèle théorique des coûts physiologiques intégrant leurs aspects mécaniques *et* évolutifs. Cette construction nous permet d'induire l'intensité des coûts de la reproduction selon la position d'un organisme sur trois continuums d'histoire de vie: "slow-fast", "income-capital breeders" et "quantity-quality". A partir de la décomposition, par Stearns (1989b), de l'architecture des contraintes d'histoire de vie en trois parties – le niveau génotypique, la structure intermédiaire et le niveau phénotypique - nous étendons notre modèle conceptuel pour y intégrer à la fois des *trade-offs* physiologiques et génétiques. Cela nous permet d'inférer les effets de l'environnement, de sa variance et de la stochasticité individuelle sur la détectabilité de chaque famille de coûts. La différence entre coûts physiologiques et génétiques se retrouve également dans leur modélisation mathématique. Il est donc nécessaire de développer de nouveaux modèles permettant d'incorporer coûts physiologiques *et* génétiques.

Nous proposons ensuite une méthode vectorielle de construction d'un tel type de modèle, que nous appelons Matrice de Projection de Population Multitrait (MPPM). Ce dernier peut implémenter chaque type de coût en l'intégrant dans la matrice en tant que trait. Nous étendons ensuite aux MPPMs les techniques d'analyse de sensibilité, standards en démographie évolutive, des modèles à un trait aux MPPMs. Surtout, nous décrivons un nouvel outil d'analyse, pertinent en théorie des histoires de vie et en démographie évolutive: la Trait Level Analysis. Elle consiste à comparer des modèles qui partagent les mêmes propriétés asymptotiques. Ceci est rendu possible par le *repliement* d'une MPPM selon certains traits, une opération qui réduit le nombre de traits du modèle en moyennant ses transitions selon les abondances ergodiques relatives. Ainsi, la Trait Level Analysis permet de mesurer l'importance évolutive des coûts de la reproduction en comparant des modèles implémentant ces coûts, avec des versions ergodiquement équivalentes de ces modèles mais *repliées* selon les traits supportant les compromis.

Nous utilisons des méthodes, classiques et nouvelles, de calculs des moments de la fitness – gradient de sélection, variance du succès reproducteur, variance environnementale – que nous appliquons aux modèles avec coûts et sans coûts afin de mesurer leurs effets démographiques et évolutifs. Nous présentons les effets conjoints des coûts physiologiques et génétiques sur la distribution par âge des taux vitaux d'une population. Nous montrons également comment les coûts physiologiques influencent les deux composants de la *sélection efficace*, en aplatissant le gradient de sélection d'un côté et en accroissant la taille efficace de la population de l'autre. Enfin, nous démontrons comment l'effet tampon des coûts sur les variances environnementales et démographiques améliore la résilience d'une population soumise aux coûts physiologiques de la reproduction.

Finalement, nous montrons en quoi ce modèle évolutif, implémentant des compromis, est pertinent pour étudier les relations entre structures d'apparentement et démographie. Ce champ disciplinaire, appelé *Kinship Demography*, s'intéresse à la fois aux conséquences démographiques des structures d'apparentement d'une population (par exemple du fait des transferts de ressources intra- et inter-générationnels entre apparentés) et à l'influence réciproque des taux vitaux d'une population sur ces structures de parenté.

Mots-clefs : Démographie évolutive, Théorie des histoires de vie, Modèles matriciels de populations, Modèles multitrait, Trait-level Analysis, Kinship models, Compromis physiologique, Compromis génétique, Coûts de la reproduction, Ecologie théorique.

Preambule

Published in 1651, the first edition of William Harvey's *De generatione animalium* (On animal generation) shows, on its front page (fig 1), Jupiter on its throne, opening an egg-shaped box from which escape living beings of all shapes and forms : a crocodile, a bird, a cricket, a deer, a man and even what looks like a plant (Harvey, 1651). On the box itself, three words, *ex ovo omnia* (*all life comes from the egg*). This was echoed, two hundred years later by Pasteur's *Omne vivum ex vivo* (Pasteur, 1862). Denying spontaneous generation for all species - as *all life is from life* - these two authors provided a link between offspring and adults. By doing so, they added a transition between generations to the within-generation study of organisms' ontogenesis and therefore established the universality of the life cycle.

(a) (b)

Figure 1: Front cover (fig 1a) and detail (fig 1b) of the 1651 version of Harvey's *De generatione animalium* (On animal generation).

Introduction

Half a century after their initial formulations, the theories of evolution (Darwin, 1859) and of the laws of genetics (Mendel, 1865) were reconciled in the *modern synthesis* (as later denoted by Huxley (1942)). This synthesis prompted the emergence of evolutionary biology and within it, population genetics – studying genetic differences within and between populations (Fisher, 1930; Haldane, 1941) - and evolutionary ecology – studying population biology from an evolutionary perspective. Within the latter, a new theory was born - Life History Theory (LHT) - that examines the life cycle of organisms in the light of evolution. It would develop steadily in the 1950s (Cole, 1954; Haldane, 1957; Lack, 1954; Maynard Smith, 1958; Medawar, 1952; Williams, 1957). Life cycles (or life histories as they are also called) are categorized by *traits* (in general one *trait*; stage for instance), where each value of the *trait* correspond to a specific life history *state* (stage rosette for example). The specificity of a life history lies therefore in the transition probabilities between the various *states* of the life cycle, which we call *vital rates*.

As population genetics considers genetic variation from an evolutionary perspective, so does Life History Theory with variations between life histories. The difference between the life cycles of two populations correspond to differing life history strategies, i.e. to different combinations of *vital rates*. As a matter of fact, the differences in strategies are often characterized by a few key components of the life history that some call life history "traits". Because these “traits” are generally of different demographic natures (some are *traits*, some *transitions*, see Stearns (1992)’s list below), we shall follow in the steps of Lamont Cole and refer to them as (pertinent) life history *features* (Cole, 1954). We illustrate this nomenclature in fig. 2.

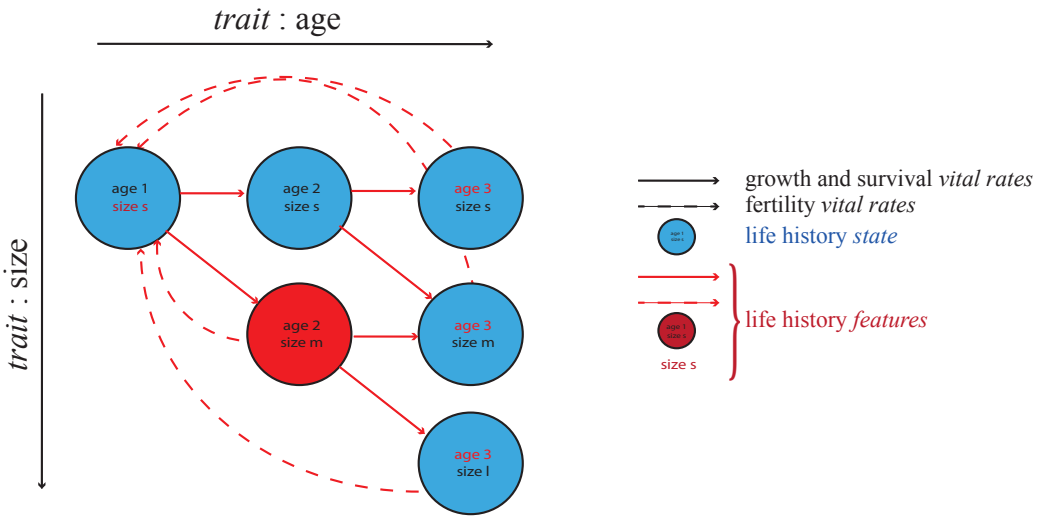


Figure 2: Life cycle. In this life cycle, structured by *traits* age and size, we represent the *states* as circles representing the combinations of *trait* values and the *vital rates* (the transitions between states) as arrows, plain for survival and growth transitions and dotted for fertility rates. Some of these *states* and transitions correspond to the key life history *features* of Stearns (1992) (he calls life history "traits"). When the case, they have been colored in red. They are: size at birth, growth pattern (the plain arrows), age and size at maturity (the *state* in red corresponds to one of these), age (and size-)-specific reproductive investments (the dotted arrows), length of life (*states* of maximum age).

In order to be ecologically meaningful, the choice of the *trait* categorizing a life history should fall on that which best segregates the organism's lifetime into *states* of great demographic and evolutionary importance. For that reason, in most cases, the *traits* used are age, stage and size. In humans for instance, age is a clear determinant of a woman's position in the key reproductive segments of her life history: immature, reproductive and post-reproductive periods. To the contrary, in trees, age is not such a great determinant of life history, but *trait* stage allows to divide the life history of the organism into low survival *states* (e.g., seed) and high survival ones (e.g., mature trees). Stage is also the *trait* of choice for organisms undergoing metamorphosis. Other *traits* can be considered as key determinants of life history, like size (for which the *vital rates* are growth rates), location (*vital rates* are dispersal rates), sex etc. As a matter of fact, according to Stearns (1992), among these *traits*, two are of paramount importance across all forms of life: age and size. This can be seen from his identification of seven crucial life history *features* that most significantly affect an organism's level of fitness: size at birth, growth pattern, age and size at maturity, number of offspring, age (and size-)-specific reproductive investments, age (and size-)-specific mortality schedules and length of life. We have illustrated this by highlighting some of these key *features* in the simple age and size life cycle of figure 2.

This slight difference in concept, between *traits* characterizing life histories and pertinent life history *features* - which correspond to certain *states* and *vital rates* as defined on these traits - would lead to a branching in LHT between two main axes of research: life history optimality theory, focusing on trade-offs between life history *features*, and evolutionary demography, based on *traits*.

Life history optimality theory

Focusing on life history *features*, life historians want to analyze their variations between individuals within a population or within a species, as well as between species. First, at the interface with population and quantitative genetics, by trying to understand in what measure these life history variations are phenotypic expressions of genetic variations. Second, by questioning how the change in a specific *feature* may be affected by the variation in another and therefore by identifying the relationships between the various life history *features*. Since for any given organism, all life history characteristics cannot be maximized simultaneously – this would lead to the so-called 'Darwinian demon' (Law, 1979) – they are constrained. These constrains, potentially caused by a limited resource common to different functions of the organism, result in what are called 'trade-offs'. Trade-offs are not the prerogative of LHT. They are at the core of quantitative genetics, where they are a major component of the **G** matrix of additive variance covariance (Lande, 1982). Life History Theory, for its part, focuses specifically on the relationships between pertinent life history *features* and in particular between the *vital rates* (the transitions in the life cycle) associated with them and that are called fitness components (fertility rates, survival rates, growth rates, etc.). These 'life history trade-offs' at the core of life history theory - relating pairs of *features* both with demographic significance and intraspecific variance - occur within all organisms. For example, in a population, some individuals will allocate more energy than others to reproduction at a certain time but at the cost of future fertility or survival; this trade-off is denoted as 'costs of reproduction' (Williams, 1966) Some will produce more offspring than others, but theirs will be frailer (smaller or less cared for, for instance); this is the quantity-quality trade-off (Lack, 1947). Other major life history trade-offs include age vs size at maturity (Stearns and Koella, 1986). Whilst these are arguably ubiquitous in nature, some are specific to certain species, taxa or taxonomic kingdoms like the trade-off between water-use-efficiency and relative-growth-rate in plants (Angert et al., 2014).

However, because variations in life history features are deemed larger between species that within populations, these trade-offs are often solely considered at the between-species level, for which trade-offs are more about evolved strategies than allocation of current resources. As a matter of fact, the evolved positions of species on key trade-offs are often used as central statistics of their life history strategies. For example, some mammal species are slow (they promote longevity at the cost of fertility), some are fast (opposing characteristics) (Gaillard et al., 1989; Stearns, 1983). Some primate species have more offspring than others, but these offspring are usually smaller (Walker et al., 2008). These slow-fast and quantity-quality spectra, form, with others (e.g., the iteroparous-semelparous spectrum), a family of life-history continua. This multidimensional family has gradually taken over the role of life history classifier formerly held by the unidimensional "r/K" characterization. This family of life history continua is in constant evolution. First, as addition to the family of trade-offs stem from the theory (see the dimensionless quantities of Charnov (2002)). Second, as theoretical and empirical results provide new ways to quantify the position of an organism on a known trade-off (Gaillard et al., 2005) sometimes with contrasted results (Oli and Dobson, 2003).

At the species level, the focus of life historians on trade-offs - deemed to stem from genetic correlations caused by antagonistic pleiotropy or genetic linkage - gave rise to the life history optimality theory. If a population is allowed time to adapt to a given environment, it will settle at a specific point on a life-history

strategy continuum - considered evolutionary stable - that maximizes its fitness (Parker and Maynard Smith, 1990). In reality, of course, an organism is located on an infinity of life history continua and only considering one pair of life history features as being optimized by natural selection causes interpretation issues. The position of an organism on a continuum may be evolutionary stable without being optimal, if for instance it is constrained by the positions on other continua, as was illustrated thanks to a famous architectural analogy by Gould and Lewontin (1979).

Evolutionary demography

Focusing on the *trait*(*s*) pacing the life cycle, life historians have turned to the tools of mathematical demography. Amid the proliferation of demographical studies on the growth rates of populations of the second half of the 19th century and the first half of the 20th century (Verhulst, 1845; Pearl and Reed, 1920), Alfred Lotka, building on earlier work by Euler (Euler, 1760), produced the famous Euler-Lotka equation (Alfred J. Lotka, 1925). This equation relates the life history transitions (the *vital rates*) of an age-structured population – the fertility and mortality rates – to the asymptotic growth rate of the population, the Malthusian parameter (Malthus, 1798). This discovery was concomitant with Fisher's equation of the latter with Darwinian fitness (Fisher, 1930) which meant that the fitness of a population could be directly extracted from its life-cycle. However skepticism grew amongst biologists about the perceived inadequacy or oversimplification of such theoretical population models (Allee, 1934; Salt, 1936). This situation led LaMont Cole, twenty years later, to express his concern: "the [...] analysis of the ways in which differences between the life histories of species may result in different characteristics of their populations has remained relatively unexplored" (Cole, 1954). New major advances in this new field of Life History Theory – in which Hutchinson placed his hopes in and christened biodemography (Hutchinson, 1948) - would finally occur at the turn of the 1960s. Initial efforts by Hamilton (1966) and Lewontin (1965) were generalized in a strikingly simultaneous effort by Demetrius (1969), Emlen (1970) and Goodman (1971) to provide "the sensitivity of the intrinsic growth rate to changes in the age-specific birth and death rates": evolutionary demography was born.

This breakthrough could certainly not have occurred without the advent of the population projection matrix by Lewis (1942) and Leslie (1945) and the demonstration by Keyfitz and Murphy (1967) that life histories were equivalently described by the "old" continuous-time equations and the "new" discrete time matrices. Written in matrix form, the life histories of age-structured populations were not only easier to visualize (the vital rates appearing clearly as different entries of the matrix) and to project over time (a "simple" multiplication of a population vector by the Leslie matrix providing the population vector at the next time-step) but also to analyze.

In 1978, using the powerful tools of linear algebra, Caswell (1978) related the sensitivities of λ (the asymptotic growth rate or dominant eigenvalue of the projection matrix) to changes in vital rates, to its associated right-eigenvector (the vector of asymptotic abundances, denoted $\boldsymbol{w}$) and left-eigenvector (the vector of reproductive values, denoted $\boldsymbol{v}$).At the same time, a fuller formalization of λ as Darwinian fitness, under simplifying assumptions, was provided (Charlesworth, 1980; Demetrius, 1981). This led to heated disputes among ecologists, in particular with respect to the alternative use of the net reproductive rate $\mathbf{R_0}$ (see, for instance, Nur, 1984; Stenseth, 1984), but these were finally settled (see Murray, 1992). Therefore, sensitivity analysis tools allowed to start providing answers to Cole's question by enabling to measure the relative contribution of certain *vital rates*, *states* and other life history *features* and sub-cycles of the life-history, to the fitness of the population (de Kroon et al., 1986; van Groenendael et al., 1994).

In parallel, the shift, for life history modeling, from continuous-time equations to matrix form allowed to expand the scope of implementable *traits*. Lefkovitch (1965) devised the first stage-based population projection matrix that now bears his name, and Usher (1966) a matrix for population characterized by size. Linear algebra, the mathematics underlying projection matrices, allowed to extend the results of age-structured models to population characterized by any *trait*. The Euler-Lotka equation was generalized into the characteristic equation of the projection matrix, from which ergodic growth rate, abundances and reproductive values could be obtained similarly. Other key demographic measures were then extended. For instance, Charlesworth (1980) showed that generation time was the first derivative of the characteristic polynomial whatever the *trait* used. Other efforts were more laborious but still proved useful as the extraction of age parameters for stage-structured populations by Cochran and Ellner (1992).

First attempt at reconciling trade-off optimality theory and evolutionary demography

A first bridge between these two slightly differing focuses of LHT - on one side, the optimality (of fitness under trade-offs constrains) theory and, on the other side, the evolutionary demographic approach – was initiated by the first sensitivity analyses of optimal life histories (Caswell, 1982a,b; Caswell and Real, 1987; Law, 1979; Schaffer, 1974). This mixed approach applies the perturbation analysis tools of evolutionary demography to the Evolutionary Stable Strategy position of optimality theory. The reasoning is that, if the organism is in ESS, its fitness λ is optimal, and its total derivative is therefore zero. This application of optimality theory to evolutionary demography provides, therefore, a multivariate relationship between all the $\frac{\partial \lambda}{\partial vr_i}$ (the sensitivities of fitness to vital rates) in the life cycle, a relationship assumed to represent the projection of trade-offs on *vital rates* and *states*. However this reconciliation attempt can be considered to be only partial, for two reasons.

First, because the relationship it provides is between all transitions (between all *states*) in the life history, and not merely between those related to the pertinent life history *features* that connect the trade-offs. Considered globally, it shows every *vital rate* as trading-off with all others and is therefore hard to make sense of. Focusing, to the contrary, on pairs of transitions - assuming all other transitions constant - it yields pairwise relationships between vital rates sensitivities that should not, ecologically speaking, be considered independently from the others (Caswell, 1982a).

Second, because at the time of this incorporation of population-level constrains in evolutionary demography, only a subset of trade-offs were actually, but largely unknowingly, considered. Indeed, the negative correlations between life history features observed at the population and species level - now called genetic (sometimes evolutionary) trade-offs - were then considered to be manifestations, at the level of the population, of the trade-offs occurring within each of its individuals - the physiological trade-offs.

Genetic and physiological trade-offs

This misinterpretation is epitomized by the use of the name of the cause (trade-off) for the consequence (a negative correlation). As Roff and Fairbairn (2007) point out: "the term 'trade-off' may be used to describe the functional relationship between two traits or the statistical correlation between the traits". Therefore the same consequences, the negative correlations between life history features, were deemed to stem from the same mechanism, although at different levels of study. This confusion prompted ecology theoreticians to step forward and start to disentangle the variety of underlying mechanisms producing trade-offs (Partridge, 1992; Roff, 1992; Stearns, 1992, 1989a,b). Since their work, we can differentiate the mechanisms of physiological and genetic trade-offs. The former act at the individual level through a physiological allocation mechanism which generates competition between different functions of the organisms, themselves affecting different fitness components. The latter stem from genetic variance in genes having antagonistically pleiotropic effects on different life history *features*, and therefore acts at the level of the population .

Despite all their efforts and clarifications, many questions remain open with regards to the roles of these two families of trade-offs. It is now suspected that they can act simultaneously (Flatt and Heyland, 2011; Kirkwood and Rose, 1991; Partridge et al., 1991), but some authors wonder, still, whether physiological trade-offs are not automatic by-products of genetic trade-offs (Gavrilov and Gavrilova, 2002; Rodríguez et al., 2017). Whilst progress has been made on the effects of each of these two families, the upcoming challenge for life historians is to understand the relative roles of physiological and genetic trade-offs and whether/how they influence one other. As Braendle et al. (2011) put it "while most traditional life history research is based on mathematical, statistical, and phylogenetic approaches without explicit reference to underlying mechanisms, today's principal research challenge is to fill this gap".

These questions on the respective roles of genetic and physiological trade-offs, have now been transposed to the field of senescence theories, where they have aroused a very large interest; the antagonistic pleiotropy theory of Williams (1957) playing the role of genetic trade-off (between early-life and late-life fitness) and the disposable soma theory of Thomas Kirkwood (Kirkwood and Holliday, 1979; Kirkwood and Rose, 1991) that of physiological trade-off (between maintenance and reproduction) (Hammers et al., 2013; Lemaître et al., 2015; Robins and Conneely, 2014; Shefferson et al., 2017).

New attempt at reconciling trade-off optimality theory and evolutionary demography

In order to take up the gauntlet thrown by Braendle et al. (2011), it is therefore now needed, on the theoretical side, to conceive of evolutionary demographic models incorporating both physiological and genetic trade-offs. In other words, a new life history optimization theory is currently required, where the trade-offs implemented are physiological. This necessitates, however, to be able to implement several traits in an evolutionary demography model. As pointed out above, incorporating (respectively inferring) a genetic trade-off into (resp. from) a life history projecting matrix as was done via the classical life history optimization theory only requires one *trait*. Both the trade-off and the matrix are about the genotypic level. For a given genotype, the optimized position on the genetic trade-off is illustrated by specific vital rates - that would differ from those of a different genotypes with a different strategy – that only need one *trait* to characterize them. To the contrary, a physiological trade-off, working at the level of the individual, occurs within the genotypic level, not between genotype-related life history strategies (each represented by its own projection matrix). Put simply, a physiological trade-off at that level is not a position on a continuum, but a real bivariate constraint, that therefore requires two *traits* to be implemented. This necessity poses the double - technical and conceptual – challenge of incorporating and interpreting multiple traits in the evolutionary demography model of choice, the projection matrix.

The study of the incorporation of a second *trait* in a projection matrix model began in earnest at the turn of the 1960s (Goodman, 1969; Le Bras, 1970; Rogers, 1969). New techniques were later introduced to facilitate this multitrait implementation (Caswell, 2009; Hunter and Caswell, 2005). In these models, however, the addition of traits was merely a way to improve the scrutiny of the model in order to get a finer understanding of the population dynamics. The ergodic abundances extracted from these were now categorized by combinations of *traits*, by age and size, or by stage and location for instance. Similarly, the sensitivities of fitness could now be computed multidimensionally (Caswell, 2012). However both a theory of multitrait matrices and a concept of "sensitivity of fitness to traits" – akin to the sensitivity of fitness to vital rates provided by Demetrius (1969); Emlen (1970); Goodman (1971) – are still lacking. In that context, the ability of adding traits to a matrix is, from an evolutionary demography point of view, pointless, as is the quest of an evolutionary demographic grasp of physiological trade-offs.

Aim of this thesis

In order to advance towards the disentanglement of physiological and genetic trade-offs and thus to better understand the evolutionary consequences of physiological trade-offs, I have had to try and tackle the aforementioned (life history) conceptual, (evolutionary demography modeling) theoretical and computational challenges. I did this in the context of “the most prominent of all life history trade-offs” (Stearns, 1989a), the costs of reproduction, whereby reproduction of an organism negatively affects its later survival and fertility.

Costs of reproduction

This trade-off is pervasive: it relates life history *features* - survival and fertility – that are *vital rates* and that are part of any life history model, whatever the *trait* used. This trade-off is also general: most life history trade-offs that connect two specific life history *features* of an organism can be considered as special cases of the costs of reproduction. Indeed, it is hard to conceive a way to impact fitness of an organism without, eventually, affecting either its mortality or its reproduction. As a matter of fact, it has been argued, that the costs of reproduction, together with the quality-quantity trade-off, constitute the key life history trade-offs, and that "all other trade-offs can be considered examples of these two major trade-offs" (Koivula et al., 2003; Lessells, 1991).

The choice of the costs of reproduction was also promoted by their direct relation with the aforementioned theories of senescence (Bell, 2011; Hendry and Berg, 1999; Jasienska, 2009; Orell and Belda, 2002) and by the fact that the general problematic around the nature and roles of genetic and physiological trade-offs has been made explicit for these costs. As Edward and Chapman (2011) formulate it " ... there is relatively little [...] work in this area so far. This is an important oversight because it is not yet clear whether physiological and evolutionary trade-offs occur via the same underlying mechanisms. It would be interesting to know, for example, whether individuals selected for early- or late-age reproduction retain equal capacity to express physiological trade-offs; that is, whether the effects underlying these different kinds of trade-offs are additive."

Plan of the manuscript

My thesis in eco-evolutionary mathematics aims to progress towards answers to Edward and Chapman (2011)'s questions. It is therefore broken down into three parts, corresponding to the three different questions identified earlier – clarifying the trade-offs concepts within life history theory, extending evolutionary demography model theory and finally computing evolutionary consequences of physiological trade-offs. I also extend, as a concise overture, these reflections towards the study of the coevolution of demographic and kinship parameters, in the context of anthropology.

Chapter 1: Costs of reproduction, concepts and methods

In the first introductive chapter, we start from Williams (1966)'s initial definition of the costs of reproduction to generate a theoretical model for physiological costs of reproduction. These are, by definition, physiological, but also, inevitably, evolutionary and both aspects need to be reflected in a theoretical approach of the costs of reproduction. We then extend this framework to incorporate genetic costs of reproduction as well. Both costs can cohabit in a population, and we draw, from the conceptual model, patterns for their detectability at different levels: the individual, the population and between populations. We then relate the different mechanisms underlying genetic and physiological costs to the different classes of models which they are generally implemented in. We finally show that, in order to incorporate both physiological and genetic costs in a single evolutionary demographic template, a new class of model is required, the multitrait population projection matrix (MPPM) and we draw the first outlines of the implementation of physiological and genetic costs in this new model.

Chapter 2: Trait level analysis of multitrait population projection matrices

In this chapter, presented in its article version as published in Theoretical Population Biology, we develop a construction method for MPPMs. This vector-based method allows to computationally efficiently model populations characterized by numerous traits with large distributions. We extend sensitivity analyses towards these models. Then, we present a new analysis tool for evolutionary demography: the *Trait Level Analysis*. It enables to compare demographic properties of a model characterizing a population by certain *traits* with its, ergodic-equivalent, *folded* model only implementing a subset of the traits. By doing so, *Trait Level Analysis* allows to measure the relative evolutionary importance of the different traits in an MPPM. The scope of this new tool is very large, but amidst its uses, it allows to measure the demographic and evolutionary consequences of a trade-off by folding upon the traits implementing it.

Chapter 3: The demographic and evolutionary consequences of physiological costs of reproduction

The life history concepts of chapter I and the mathematical tools of chapter II, enable us to construct an evolutionary model implementing both physiological costs and genetic costs of reproduction. After setting up the model, we provide the necessary calculation tools – some of them new – to extract, from multitrait models, several fitness measures such as the selection gradients (the sensitivity of fitness to vital rates), the net reproductive rate ($\mathbf{R_0}$), its variance (the variance in lifetime reproductive output $\sigma^2_{\mathcal{LRO}}$), and the demographic (σ^2_d) and environmental (σ^2_e) variances. We can then use these tools to compare these various fitness measures between the full model implementing the costs of reproduction and the folded models, implementing only type of costs, physiological or genetic, or implementing none. The combination of the trait level analysis and the fitness measures computations allow to gauge the evolutionary and demographic effects of the costs of reproduction.

Chapter 4: Kinship demography

In chapter 4, we discuss a new field, to which the concepts, methods and tools of the preceding chapters will benefit, that of *kinship demography*. In social species, *kinship demography* deals with the effects of the kinship distribution in a population on its demography (transfers of resources between kin affect the vital rates of both the giver and the receiver) and with the reciprocal influence of demography on kinship distribution. As such, *kinship demography* is in contact with a wide range of fields. Focusing on humans, we discuss these connected research domains and review the specific benefits on a child of kin aliveness and the specific costs on parents of caring for their offspring. Finally we hint at two extensions. First, of the theoretical model implementing costs of reproduction of chapter 1, to implement transfers between related individuals. Second, of the fitness measures in multitrait models (chapter 3) to provide inferred distributions of kin.

Chapter 1

The costs of reproduction, from a general theory to evolutionary models

Contents

1.1 Introduction

Without constraints on life histories, nature would be invaded by non-competing Darwinian demons (Law, 1979); an organism which can simultaneously maximize all fitness components. Life history as a field would then have no reason to exist. In life history theory such constraints are called trade-offs. They "represent the costs paid in the currency of fitness when a beneficial change in one trait is linked to detrimental change in another" (Stearns, 1989a). Among these life-history trade-offs, the cost(s) of reproduction (singular and plural are equivalently used) is the most prominent (Stearns, 1989a) as the traits it connects are directly the two highest level components of fitness : survival and fertility. Another trade-off relates quality and quantity of offspring and it has been argued that all other trade-offs are particular cases of either one of the two (Lessells, 1991).

Costs of reproduction as an allocation mechanism

When coining the term, Williams (1966) based his definition of the costs of reproduction on Fisher (1930)'s *reproductive value* of an individual. In an age-structured population, if, for a given genotype of fitness λ (the asymptotic growth rate of sub-population with this genotype) and maximum longevity ω, the expected fertility and survival rates at age j are f_j and s_j, then the reproductive value of an individual aged i (i.e., the present value of all its expected future offspring) is $v_i = \sum_{j=i}^{\omega} f_j \lambda^{i-j} (\prod_{k=i}^{j-1} s_k)$. Williams (1966) formalized the costs of reproduction as the trade-off, at the level of the individual and at each age i, between the *reproductive effort* - the portion of its reproductive value "immediately at stake", simply taken as the current fertility rate f_i - and the *residual reproductive value* which is then $\sum_{j=i+1}^{\omega} f_j \lambda^{i-j} (\prod_{k=i}^{j-1} s_k)$.

Such a cost is an application of the principle of allocation (Cody, 1966; Lack, 1954; Orton, 1929) to reproductive value and was fully formalized by Gadgil and Bossert (1970) under the term "cost of reproductive effort". Therefore Williams' characterization of the costs of reproduction is clearly based on a repartition mechanism whereby the allocation of available resources infers costs on the individual's survival and fertility. Following Partridge et al. (1991), we further categorize such allocative trade-offs - they call *functional constraints* - in either *physiological* or *ecological* costs of reproduction.

The former represent "functional constraints internal to the organism", the latter originate from the organism's environment when trying to increase its current reproductive effort. When deciding to "forage one more time instead of retiring for the night" Williams (1966)'s robin invests more resource towards its current brood at the *physiological cost* of having less metabolic resource for its own survival and at the *ecological cost* of being killed by a predator. Naturally the difference in such costs is progressive : *physiological costs* are affected by the environment and *ecological costs* have physiological causes and effects. But at the far ends of the spectrum, we can differentiate these costs by whether the actual allocation process of reproductive value consists in a split (*physiological* costs) or in a bet (*ecological* costs).

As a matter of fact it can be argued that *ecological costs* of reproduction are not really about an allocation towards higher or lower vital rates at the next(s) time-step(s), but about an all-or-nothing gamble taken by the individual to freely increase residual reproductive value at the cost of losing it all. As a consequence, *ecological costs* to pay are larger (disproportionate to the potential increase in reproductive value), shorter-term (mostly immediate) and far more dependent on the environment at the time the risk is taken. *Ecological costs* are, therefore, key components of the study of ecosystems. *Physiological costs* for their part, as they gradually connect the different fitness components of a life cycle, are the clay from which life history is mold. They are, therefore, the subject of this chapter.

The two sides of costs of reproduction

An individual life trajectory is the stochastic realization, in a specific environment, of the individual's genotype. Each genotype will contain guidelines leading its bearers along a specific life-history strategy evolved by its ancestors. The different life-history-strategy-genotypes in the population will only marginally differ from the broader strategy evolved at the species level (along with variants of other life history traits). Therefore the allocation towards reproductive effort, for an individual at a given time-step, will both depend on the state of the individual, encompassing its past reproductive and environmental histories, and on the life history its lineage has evolved.

That physiological costs are two-sided was certainly already obvious to Williams as he asserted his fundamental definition of costs of reproduction. The power of his phrase comes from the admixture of two semantic fields that are generally formalized separately. Indeed only resources can be allocated, and only towards functional mechanisms. However reproductive value is not a resource, and the residual reproductive value not a biological function of the organism. Rather, these are evolutionary concepts that

embed the theoretical expectation of all future vital rates as evolved by the organism's genotype. In such a phrase Williams manages therefore to evoke at the same time and thus in a seemingly circular manner, a *forward* process whereby allocation, at a given time-step, depends on the *past* realized life trajectory of the individual and a *backward* process whereby this allocation also depends on the individual's expected *future* life trajectory.

A closer look at the phrase highlights the contradictions such a mixture is doomed to generate. If the individual constantly "decides" to allocate towards its residual reproductive value, it is deemed to live forever. However its reproductive value, transferred untouched from one period to the next, is finite, implying the organism has a maximum age. From a physiological point of view, this implies the organism allocates resources that it has yet to (ever) acquire. As a matter of fact, using concepts loosely, Williams manages to highlight, in a single sentence, the two main drivers of an allocation process at the individual level. At each time-step, the reproductive effort will be determined by both its past realized environmental and reproductive histories and by its expected future life trajectory as embedded in its genotype. This highlights the fact that life-history strategies are certainly major drivers of the mechanisms of physiological costs of reproduction, while physiological costs are major drivers of individual life-histories.

Further this hints at several important features of the costs of reproduction, rarely found in the literature. First, the existence of different types of resource capitals, either built *forward*, or managed *backward*; themselves certainly related to different type of resources. A novel theory of the costs of reproduction could stem from such a difference, to better link empirical knowledge at the species' level to theoretical predictions of the physiological and genetic structure of the costs of reproduction. Second, the double sides of physiological costs should also be accounted for when assessing the emergence of these costs at the phenotypic level and their detectability along different life-trajectory segments.

From physiological costs of reproduction to negative correlations and ... not back

The variance in genotypic life-history strategies inside a population may have confounding effects on the identification of costs of reproduction. Indeed, whilst Williams' definition clearly depicts an allocative mechanism at the level of the individual, the term "costs of reproductions" is often used as soon as a negative correlation appears between vital rates, at different ages, aggregated at the level of the individual, population, species or across taxa. This is because the accumulation of time-step physiological costs over the life of an individual may translate into negative correlations between various fitness components both at that level and when aggregated at higher levels.

By contrast, such negative correlations need no allocative physiological trade-off to occur. Observing that mortality rates correlate with fertility rates, at the level of the population, does not necessarily imply the action of underlying physiological costs of reproduction. First, because, in a changing environment, such correlations may also be due to pleiotropic genes (each acting on vital rates at certain ages) that have crossing reaction norms (dynamic linkages). Second, since, even in constant environment, such negative relationships can also be generated by variance in genotypic life-history strategies. Whenever two alleles of a gene driving life history allocation strategy (or more generally of a gene with pleiotropic effects on different fitness components) cohabit in the population, negative correlations between vital rates at different ages will emerge. Because they have similar phenotypic consequences than the physiological costs, the genetic variance in such a gene is also considered to be a cost of reproduction called genetic costs. This a general concept that applies to all trade-offs : genetic (or evolutionary) trade-offs occur when, in a population in a constant environment, there is genetic variance in a gene that has antagonistic pleiotropic effects on the traits connected by the trade-off.

Both physiological and genetic costs have been shown to occur in nature. However, as Edward and Chapman (2011) are asking, "it is not yet clear whether physiological and evolutionary trade-offs occur via the same underlying mechanisms. It would be interesting to know, for example, whether individuals selected for early- or late-age reproduction retain equal capacity to express physiological trade-offs; that is, whether the effects underlying these different kinds of trade-offs are additive". We would go even further and claim that it is as yet unknown whether they are different manifestations of the same underlying evolutionary process, or altogether different, sometimes potentially opposing, mechanisms.

In real life, we only encounter costs of reproduction through their phenotypic expressions.van Noordwijk and de Jong (1986) have shown how, with regards to the physiological process of acquiring/allocating energy, detectability was marred when the variance in allocation between individuals is swamped by the variance in acquisition.Houle (1991) studied this phenomenon at the genetic level. Such analyses are important since detectability of trade-offs tell us about the underlying processes at play, and because they

provide a general level of expectation as to whether particular costs can be expected to be observable (Metcalf, 2016). However physiological costs are not all about energy, and since physiological and genetic costs stem from different mechanisms, we do not expect the same factors to allow each underlying cost to become phenotypical and further to become detectable. It has thus become necessary to disentangle the cross-effects of factors on the detectability of the various costs of reproduction, for - as Jessica Metcalf puts it - understanding how such processes "push distributions of traits around" and "thinking clearly about the drivers of this variation can be rather counterintuitive" (Metcalf, 2016). It is however necessary to understand the actual mechanisms driving phenotypic costs and detection at the level of the individual and the population, and their evolutionary consequences.

Models for trade-offs

Difference in core mechanisms also imply these different costs -physiological and genetic - will be accommodated by different models. Most life history models for trade-offs solely focus on genetic trade-offs. This is the case of early theoretical models by Schaffer (1974) and Taylor et al. (1974), analyzed and put in context by Pianka and Parker (1975), extended by Charlesworth and Leon (1976) and many others. For instance, when Bell (1980) invokes costs of reproduction to analyze the emergence of semelparity in iteroparous organisms, he really studies the invasibility of alternative alleles of a gene antagonistically acting on early reproduction and late fitness. Equivalent approaches have used evolutionary demography's projection matrix to implement or study genetic costs. In particular, optimality theory allows to infer the strength of these costs from the matrix of vital rates of the organism supposedly at ESS (Caswell, 1982b, 1984; Van Tienderen, 1995). This method is akin to the multivariate quantitative genetics approach which enables to anticipate the change, over the near evolutionary future, in the mean value of a pair of traits from the (genetic) trade-off between these traits as embedded in the $\mathbf{G}$ genetic covariance matrix (Charnov, 1989). And indeed Charlesworth (1990) has shown that these two approaches are equivalent under certain conditions.

Physiological costs, for their part, were incorporated early in the theory, but modeled much later. This is because of both the complexity of implementing such complex mechanisms and because, for most ecologists, as Dhobzhansky puts it "nothing [...] makes sense except in the light of evolution". Because physiological costs of reproduction occur at the level of the individual, they have been modeled via Individual-Based Models (IBM, also known as agent-based models or, in demography, microsimulations), which offer valuable information (see the incorporation of key components of survival physiological costs of reproduction in order to estimate their demographic consequences in Proaktor et al., 2008). Because they involve complex (for instance metabolic) pathways, they require complex modeling to be accurate (illustrated by the complexity of fish bioenergetics modeling in (Jørgensen et al., 2016)). In an IBM, every particle is tracked at all times. If the processes determining its fate and the fates of its offspring are complex, it will be difficult to infer an accurate distribution of the stochastic growth rate for such a population; let alone its sensitivity to vital rates. Without such selection gradients, any understanding of the evolutionary (recent) past and (near) future of the organism is made harder, limiting the interest of such models (but see Lee, 2008).

An evolutionary framework for physiological costs is therefore increasingly needed. Some have tried to bridge the gap between these different classes of models. This is, for instance, the case of McNamara, Houston, Mangel and Clark (see for instance McNamara and Houston, 1986; Mangel and Clark, 1986; Houston et al., 1988; McNamara and Houston, 1996; Clark and Mangel, 2000), who have developed a state-dependent model framework, embedded in dynamic programming theory. Such models incorporate a level of scrutiny closer to the individual (they are "state-based") that allow implementing physiological trade-offs, and backward induction, to infer optimal strategies maximizing fitness. Such a tool is important as it allows to implement trade-offs and stochasticity, whilst still being able to compute all types of evolutionary measures. However they are not evolutionary models in the sense that they do not allow, for each genotype and environment, to agglomerate all state-specific rates and constraints in a single equation that relates all vital rates and trade-offs with fitness (growth rate). Matrix models do provide such a powerful tool - it is the characteristic equation, named Euler-Lotka for age-structured models - that allows to provide an evolutionary-neutral framework in which one can ponder the relative evolutionary importance of vital rates, and most importantly a trade-offs themselves. The much needed evolutionary framework for physiological costs should therefore be based on population projection matrix theory. It should additionally allow the implementation of any type of physiological and genetic costs. It would then go a long way towards answering such questions regarding the nature, cohabitation, detectability and evolution of both sides of trade-offs.

Plan

In this chapter, drawing on the seminal theoretical works of Bell (1980); Gadgil and Bossert (1970); Partridge et al. (1991); Roff (1992) and most importantly Stearns (1989b) we first establish a consistent and unifying theory of *physiological costs of reproduction*, that - whilst kept as simple and parsimonious as possible - can incorporate all the major inputs of these costs, such as the environment, the life-history strategy of the organism and individual stochasticity. To do so, we base the allocation formula - of resources towards reproduction - at the core of physiological costs, on (i) two capitals (related to the two sides of costs aforementioned) and on (ii) the position of organisms on major life-history strategy spectra (Slow-Fast,Income-Capital-Breeding and Quantity-Quality). From this formula, we then derive expected secondary determinants of the costs, and make predictions with regards to the different manifestations of such costs depending on life-history strategy.

We then extend this physiological mechanism that lies in Stearns (1989b)'s intermediate structure by adding genetic variance at the genotypic level. We principally focus on the variance of two genes. First the allocation gene, which variance generates a gradient of heritable iso-fitness life-history strategies in the population, mainly characterized by their respective positions on an intraspecific Slow-Fast Continuum. Second the acquisition gene, which variance generates a gradient in overall fitness between genotypes. We show that the *genetic costs of reproduction* arising from the variance in allocation strategy and the physiological allocative process itself, can combine to form a general mechanism we call *physiological costs of reproduction with genetic basis*. We then indicate that genetic costs can, however, emerge without any need for an underlying physiological mechanism.

This mapping of costs of reproduction, physiological and genetic, is then used to make predictions with respect to their detectability at different levels (individual, intrapopulation and interpopulation) and to the influence of two major drivers of their emergence : the environment (its absolute level and its variance) and individual stochasticity.

Finally, we discuss the different types of mathematical models that are adapted to these costs and show that the dichotomy in core mechanism (physiological vs genetic) is reflected by a dichotomy in model families (individual-based vs projection matrix models). We then go on and provide initial steps towards the construction of a model that can bridge the gap between these two families of model in order to model *physiological costs of reproduction with genetic basis*.

In the discussion, we show that this fundamental dichotomy has further repercussions in the field of evolutionary senescence theories and this helps us to discuss the relatedness of these two types of costs, and to interpret their fundamental hermeneutical differences.

1.2 A general theory of costs of reproduction

In order to determinate and model the architecture of the physiological costs of reproduction, we shall refer to Stearns (1989a)'s classification of trade-offs. According to him, trade-offs can be studied at three levels, "the phenotypic level, the genotypic level and the intermediate structure": the phenotypic level is where selection acts, the genotypic level drives heredity, and the intermediate structure, filling up the entire space in between, "modulates the expression of genetic trade-offs [... depending on ...] environmental conditions".

When focusing on allocative costs of reproduction it is quite clear than no specific genetic polymorphism is required to account for the generic mechanism, the principle of allocation operating at its core. In other words, genetic variance is not required for the cost to operate. The cost is a physiological feature encountered by all individuals, at all times in all environments, whatever their genotype. Indeed for Partridge et al. (1991) cheetahs' life histories, although "lacking significant genetic variance", would be absurdly deemed not to involve trade-offs. Thus, the core, allocative mechanism of physiological costs lay in the intermediate structure.

1.2.1 Intermediate structure

The intermediate structure of *physiological costs of reproduction*would consist, for a given individual life trajectory, in the compounding of successive allocations that combine with the encountered environmental series to produce the phenotypic level. At each time-step, for a given individual, in a given environment, the physiological mechanism consists in a machinery involving *resources* - requiring or not to be acquired - and their *allocation* towards *reproductive effort*. The various patterns of these three important components will shape the diversity of physiological costs. This variety is the reason, we believe, for the use of plural "costs" by Williams (1966). They will also allow us to categorize these costs with regards to the timing

and strength of their phenotypic expression.

Capitals, resources and mechanisms

As mentioned in the introduction page 8, the allocation process of the costs of reproduction is two-sided. At each time, the allocation towards reproductive effort is a function of both a forward process projecting the effects of realized past life-trajectory towards the present, and a backward process materializing at the current time, all the future reproductive efforts as expected by the evolved life-history of its lineage.

We argue that these two aspects are not merely two alternative ways to conceive the costs of reproduction. Rather they are related to two different costs acting on two different capitals, blended in Williams' definition under the name of reproductive value, both continuously and simultaneously affecting life-trajectories. We further argue that this categorization is fundamental to better understand the schedule of the costs of reproduction throughout life-history, as well as to structure life-history within empirically grounded continua : the slow-fast continuum, the income-capital breeding spectrum and the quantity-quality continuum.

Two different capitals drive *physiological costs of reproduction* We introduce two concepts of capitals, each corresponding to one side of the physiological costs. The fluctuating capital (FC) is built forward. It starts empty and thus resources need to be acquired before any allocation can occur. The ratchet-capital (RC) is managed backwards. It consists of a lifelong budget that reduces each time as the organism divests the resource towards reproductive effort. We shall consider the physiological costs to be about the constant uses of both capitals, each of which required to remain positive under penalty of death. Since the FC fluctuates over time, it is associated with resources (we shall call them FC resources) than can (and need to) be acquired like energy. By contrast the RC is associated to any resource (RC resource) that cannot, like time (see Lorenzini et al., 2011). Resources combining properties of both capitals, like metabolism, would have effects on both.

Defining evolved lifetime reproductive effort Let us first us define quantities describing the central trajectory of reproductive efforts, evolved at the level of the population. Let $f(a)$ and $s(a)$ be the expected fertility and survival rate at age (or any state parameter) a as evolved by the individual's ancestors. Because in many species effort to produce independent offspring is spread over time, before and after birth for instance (see section 1.2.1), let us define res the *reproductive effort schedule* which represents the time distribution of reproductive effort required to produce one independent offspring. For example, res could be a distribution centered on birth which also encompasses efforts before birth (mating, gestation, incubation, etc.) and after birth (lactation, parental care, etc.). From s (survival rate) let us first construct $e(a) = \sum_a s(t)$, the life expectancy at age a. Then we can write that the *reproductive effort schedule* convolves with fertility rates by age/state $f(a)$ to produce the lifetime distribution of all reproductive efforts re :

$$re(a) = (f * res)(a) = \int_{t=0}^{e(0)} f(t).res(t-a).dt \quad , \tag{1.1}$$

where $*$ represents the mathematical convolution of two distributions. In this equation $f(t).res(t-a)$ represents the effort produced $a-t$ time-steps from the fertility event at time t represented by $f(t)$. Because we focus here on central evolved life-history strategy, we integrate the convoluted reproductive efforts over the central evolved lifetime, i.e. from birth to life expectancy $e(0)$. Summed over life, it extends the fertility lifetime schedule f into the lifetime reproductive schedule re. Therefore re represents the way reproduction is structured over the expected lifetime, expressed by e, of an average individual. It tells us, for a given life expectancy $e(0)$, whether reproductive efforts are concentrated early or later in life, cease before or continue until old ages, are clustered over a short time span, or fanned out throughout life, are stacked over a short $e(0)$ or thinly spread over a longer life expectancy. Thus, the (e, re) joint distribution completely determines the life history strategy of an individual's lineage. Specific life history characteristics, like the positions on the slow-fast continuum (SFC) and the semelparity/iteroparity spectrum and to a lesser extent the income/capital breeding (ICB) continuum, are all moments of (e, re).

We denote rc the backwards cumulative distribution of re. It is summed from the population life expectancy $e(0)$ to the specific age of the organism a, such that $rc(a) = \sum_{j=a}^{e(0)} re(j) + \alpha$; with α the non-reproductive baseline periodic costs simplified as a constant over time. The addition of α in the definition of rc to account for other non-reproductive functions that may also have to receive efforts from the organism renders the model easier to fathom, but, in theory, is not necessary if one considers that every effort an organism has evolved to deliver has to promote reproduction one way or another. Natural selection

would indeed prevent, at the genotypic level, any effort that does not eventually lead to the multiplication of an individual's gene. At the level of the individuals' life trajectories this is not the same obviously, as sterile individuals are still able to produce the efforts necessary for their survival without resulting in any reproduction. Actually, such individuals, as they don't incur costs of reproduction, will actually live longer than the fecund elements of the population, as exemplified by the increased lifespan of Korean eunuchs (Min et al., 2012). The same reasoning prompts us to sum the expected reproductive efforts over the life expectancy evolved by the organism in order to generate, backwards, the expected total effort produced $rc(0)$ and not longevity. On the contrary, the *actual* RC of an individual - continuously reduced as the *realized* reproductive efforts unfold over life - will need to be tracked until the maximum longevity of the organism.

Defining individual realized reproductive effort We have so far defined three quantities, res, re and rc, common for all individuals in a lineage, characterizing the central life history strategy towards reproduction. Let us now see how these relate to the individual life trajectories in relation with the two capitals defined previously RC and FC.

At birth (this very simple model does not take into account any cost of onthogenesis) the ratchet capital is maximum and worth $RC_0 = rc(0)$. All individuals of the same (e, re) will thus share the same initial level of RC. However their life trajectories will soon diverge as at each time t, RC would be diminished by the portion of RC resource (say time) allocated to the reproductive effort the organism is able to produce at that period : RE_t. Such a capital is analogous to the maintenance capital of Kirkwood and Rose (1991), with a stronger focus on reproductive effort. Then, simply,

$$RC_{t+1} = RC_t - RE_t - \alpha \tag{1.2}$$

Therefore, the difference between RC_t and $rc(a)$ for an individual aged a at time t is $\sum_{a=1}^{t} RE_a - re(a)$ which corresponds to the accumulated divergence between expected and realized reproductive efforts.

Conversely, the FC is zero at birth $FC_0 = 0$ (as, again, onthogenesis is not modeled here). Then, at each time-step, FC resources (say energy) - necessary for reproduction - would be acquired and added to the FC capital, a portion of which then spent on reproductive effort. We denote the FC resources acquired from the environment at time t, Env_t. They can possibly be stored, with efficiency $stor$, if the organism has evolved the capacity to build FC resources reserves.

Making the assumption that reproductive effort has proportional effects on both capitals, we scale them so that the REs in each system need not be (but bearing in mind that RC and FC are made or different resources and thus counted in different units of measurement). Then

$$FC_{t+1} = stor.(FC_t + Env_t - RE_t) \tag{1.3}$$

In order to predict the relative order of magnitude of FC and RC, let us consider an organism with a storage capability that is low or even non-existent (i.e. $stor \approx 0$). Then FC will be almost reset after each reproductive effort, $FC_t \approx 0$. During each time-step the level of FC will fluctuate between 0, before acquiring the resource, and $\overline{Env}$ (the level of such resource to be acquired in an average environment) before producing the reproductive effort. Since there is no point for such an organism in not spending its acquired resources before its capital is reset at 0, we would expect its mean reproductive effort to be approximately $\overline{RE} \approx \overline{Env}$. Simplifying also the RC process by setting α at 0, this implies that $RC_0 = \sum_t RC_t - RC_{t+1} = \sum_t RE_t = e_0.\overline{RE}$. And thus, whilst FC is fluctuating around the level of $\overline{RE}$ (the mean periodic acquisition of resource), RC, at birth, is of the order of magnitude of the total reproductive effort an organism is expected to produce in its lifetime in the mean environment.

The intermediate structure component of figure 1.1 illustrates the differences in magnitude between FC and RC capitals and how they are impacted by reproductive efforts (green arrows).

Ratchet Capital is affected by position in the slow-fast continuum The way the total potential effort, RC_0 is spread over lifetime is measurable by the ratio:

$$sfc = \frac{\sum i.re(i)}{\sum re(i)} = \frac{\sum i.re(i)}{RC_0} = \frac{\sum i.re(i)}{e(0).\overline{RE}} \tag{1.4}$$

It depends on the life history (e, re) this organism has evolved and in particular on its position on the continuum called slow-fast (SFC) (Gaillard et al., 1989; Stearns, 1983; Promislow and Harvey, 1990). Indeed sfc is (inversely) related, as a ratio of life timing to reproductive effort, to the F/a ratio where a is the age at first reproduction and F the fertility rate, used to categorize mammals (Oli and Dobson, 2003). It is even more related to generation time, another (and arguably stronger) indicator of the position on

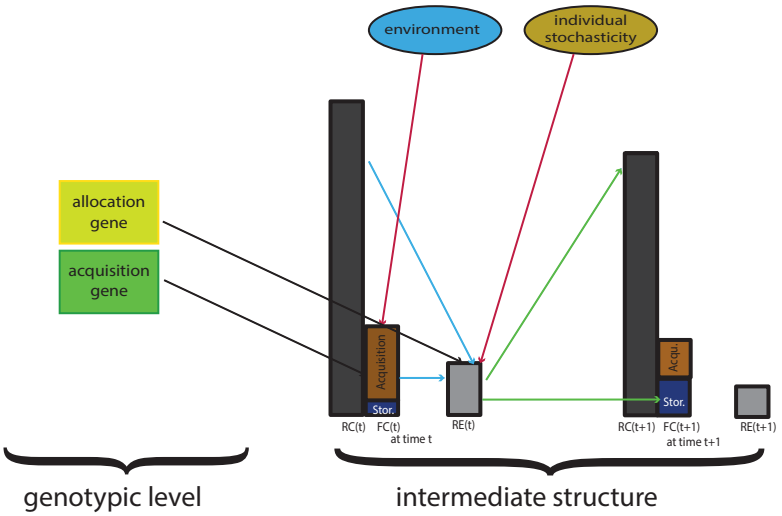


Figure 1.1: Representation of the genotypic level and intermediate structure of *physiological costs of reproduction with genetic basis*. In the intermediate structure, the organism is represented by two capitals the Ratchet Capital (RC) and the Fluctuating Capital (FC). The latter being replenished, every time-step, by acquired FC resource and both are diminished by the reproductive effort produced at time t, $RE(t)$ (green arrows symbolize the reduction of capitals by reproductive efforts). The reproductive effort itself is derived, via the allocation process, from the levels of both capitals (blue arrows depict the effects of both capitals as main drivers of the allocation process). The two main processes generating $RE(t)$, the allocation process and the acquisition of FC resources are affected by the environment and by chance (red arrows) and, at the genotypic level, by the variance in (allocation $\times$ acquisition)(black arrows).

the SFC (see Gaillard et al., 2005) (sfc is actually equal to generation time in the simplified case where the entire reproductive effort schedule is concentrated at time of birth, i.e. $res = \delta_0$ (Dirac distribution), implying both reproductive effort and fertility lifetime distribution are equal, $re = f$).

We extend this denomination here for any organism (mammal or not) to describe the pace at which it distributes its reproductive effort re over its expected lifetime trajectory e. This characterization is affiliated to genetic costs of reproduction (see in the introduction page 8) : for a given environment, two genotypes inferring the same fitness, can cohabit in an organism, ; the "fast" allele would promote fertility (at the cost of survival), the "slow" one longevity at the cost of fecundity.

The connection between RC and SFC has important consequences for the way this capital is managed over time and the costs of reproduction it incurs. RC costs, because of the compounding of the inherent ratchet effect of reproductive effort on capital, will result in long-term and mostly late-life effects. After a lifetime of erosion, the accumulation of reproductive effort brings the RC close to levels at which neither fertility nor survival can be sustained. The delay of such effects being proportional to sfc. Conversely, at young ages, the high ratio of $\frac{RC(t)}{\overline{RE}} \approx e(0)$ implies that RC costs will be little, buffered by the high level of the capital, all the more so for slow (high sfc) individuals. In a nutshell, we expect slow organisms to experiment important long-term and very little short term physiological costs related to RC, compared to fast organisms.

Fluctuating Capital is affected by position in the income-capital breeding continuum
Since it starts life at 0, the FC will have no such long-term buffer effect. This is especially true for organisms that cannot store the FC resource (say energy). Indeed, if $stor = 0$, then at the beginning of each period, i.e. at the beginning of a new cycle of acquisition followed by allocation of the resource, $FC(t) = 0$. Such FC costs will have no effects beyond the end of the period, and will thus mainly consist in the reproductive effort negatively affecting survival until the following acquisition period. If $stor > 0$ though, then the portion of the unused FC that can be stored and therefore carried over to the next period, will buffer and delay the FC costs. Since, as discussed above in section 1.2.1, the size of the FC is of the order of magnitude of $\overline{RE}$, such effects will then, contrary to RC costs, be short- to mid-term. Because of the acquisition component of FC, they will depend heavily on the environment. With sufficient storage capacity, the FC will be able to buffer part of the effect of environmental variance on reproductive effort. Conversely this means that the effect of the costs, though delayed by storage, will still be strongly

dependent on the environment. Thus the timing, phenotype and detection of FC costs will strongly depend on storage facility *stor*, i.e. on the organism's position on the income-capital breeder (ICB) spectrum.

This was studied by Stearns (1989b) in an article where, focusing on acquirable resource (he calls energy, we call FC resource), he distinguishes FC costs between "behavioral costs" and "physiological costs". Behavioral costs are associated with the economics and accounting-related concepts of *income breeders* (IB) and *direct costing* first used in an ecological framework by (Drent and Daan, 1980; Sibly and Calow, 1983, 1984) : for *income breeders*, which "high metabolic rate leave little room for storage", the cost of reproduction is *direct*, "drawn out of current revenue" and relate to behavioral mechanisms (e.g. foraging)(Stearns, 1989b). The *Tokophrya* studied Kent (1981) which "produces one offspring for each Paramecium eaten" is an example of "income breeder". "Physiological costs" are associated with the concepts of *capital breeders* (CB) and *absorption costing*. *Capital breeders*, like the red deer studied by Clutton-Brock et al. (1983), physiologically allocate a portion of the FC resource pool to be *absorbed* by the current reproductive effort and the unused portion can be stored for future reproduction, in the FC capital. As they can capitalize energy over time, successive breeding attempts will likely share the same resource pool (the FC capital), thus both delaying and buffering the costs and the environmental effects.

In this way, the physiological costs associated with both RC and the FC of Income Breeders (IB-FC), are costs of cumulative reproductive efforts, with short, middle and long-term effects. They are very different in nature but operate similar mechanisms. One main difference being that we expect FC costs to be phenotypically more important in the short and middle-term (the higher the storage capacity *stor*, the longer the effects), as they are about reserves that are readily available should the environment become detrimental. Rather RC costs - as RC capital is much larger than FC in early/mid-life - will only have strong effects in late-life, with a delay related to sfc. The other difference is the strong dependence of the FC capital on the environment, through its acquisition process. However, because any allocation towards reproductive effort will deplete both capitals, current and past environments - via respectively the allocation and storage processes of FC- will be strong determinant of reproductive effort and therefore the RC costs are also, albeit much less, environmentally-dependent.

As a conclusion, we expect the strength and delay of the physiological costs of reproduction to be vastly determined by the organism's position on the Income-Capital Breeding and Slow-Fast continua.

Ratchet and Fluctuating Capitals in nature In the wild, a vast array of timing and strength of costs of reproduction have been observed. Pre-industrial humans, have been shown to exhibit both short term and long-term costs. The former are mostly to be found in the literature on "maternal depletion syndrome", showing the negative correlations displayed between pregnancies and maternal health (Butte and King, 2005) but with limited effect on overall reproductive success (Gurven et al., 2016). But mostly, pre-industrial humans have been investigated for longer term effects. Hayward et al. (2015) have showed, for instance, that early-life fecundity - measured as the number of children produced before age 25 - is positively correlated with mortality rate throughout the remaining life of the mother. Westendorp and Kirkwood (1998) and Thomas et al. (2000) have furthermore established the negative relationship between overall number of children and longevity. As have Gagnon et al. (2009) in a studying of the population of ancient Québec.

The Soay sheep - a capital breeder - where shown by Tavecchia et al. (2005) to display effects of breeding success at maturity on survival with effects throughout life, and stronger at young and old ages. This is in accordance with our model where the FC costs of early reproduction would hinder survival at that time and shortly thereafter with decreasing effect over time, whilst RC costs would take effect much later.

To the contrary, the small passerine Parus Montanus - an income breeder, thus lacking long-term energy capital - was demonstrated by (Orell and Belda, 2002) to suffer only long-term costs (early breeding impacts females' survival rates aged 5 years or older but not in the years following breeding). This agrees with our model where IC do not suffer from mid-term FC costs (only immediate), but from late-life RC costs.

Indeed, another trademark of FC costs for organisms able to store its acquirable resource, is their sometimes sporadic effects throughout the life of individuals. This is remarkable, for instance in such a capital breeder as the caribou, Rangifer Tarandus, which every 4 years on average, enters "reproductive pause", in order to "compensate for the [...] costs of gestation and lactation" (Cameron, 1994). A somewhat similar behavior in perennial plants, the masting strategy - whereby seed production is periodically massively reduced, has been related to such an energy cost by Venner et al. (2016). Masting differ however with reproductive pause by the total synchronicity of its occurrence at the population level. This may be explained by the larger still environmental-dependency of such plants (especially comparatively with other factors, like age) and/or by the synchronizing effect on reproductive effort of cross-pollination (Venner et al., 2016).

The special case of income-breeders FC capital IB-FC costs, through the preponderant importance of the environment and of the very narrow time window of its effects (only until the next feeding season) are extremely similar in patterns, if not in nature, to the ecological costs described in the introduction page 8. This is, we think, the reason why Stearns (1989b) calls them "behavioral costs", whilst calling CB-FC simply "physiological costs". Following in his steps, we shall in this article mainly focus on costs with a potential lasting effect, namely the RC and CB-FC costs, calling them simply physiological costs.

It should be noted however, that via their effect on immediate survival ecological/IB-FC costs, have also long-term effects from a live-history trajectory perspective. At the individual level, once the bet of the current reproductive effort is won, there is obviously no physiological cost impacting future fitness. However, at the organism level, the expected number of offspring in late life, via its dependency on survival, is affected by this potential ecological cost. For such costs, we could therefore consider the "survival gauge" - the accumulated mortality risks an individual takes in order to reproduce - as a capital, akin to FC and RC (with the main difference that, instead of being gradually emptied, it is entirely put at stake for each reproductive effort). Thus theoretically, the model for physiological costs of this review are readily extendable to ecological/IB-FC costs.

Allocation process

In most analysis and models of costs of reproduction, the allocation process towards reproduction is predicted and implemented as a function of many factors such as age, stage, size and environmental conditions. In our two-capital model, such drivers can be regarded as second level dependencies of the main parameters that determine, we think, the allocation process : the capitals and the life history strategy.

Capitals and life history strategy drive the allocation process towards allowed reproductive effort The way the two capitals combine is certainly a complex process that will prove very difficult to put in simple equations. However, with parsimony in mind, we can write the maximum reproductive effort, allowed by the capitals, at time t, $aRE(t)$, as the following combinations of both capitals, cumulative lifetime reproductive schedule (rc) and position on the ICB ($stor$). Both capitals are required for a reproductive effort, therefore the maximum allowed reproductive effort will be limited by the capital with the *min*imum level:

$$aRE(t) \approx \min(FC(t) - stor.\overline{RE}, RC(t) - K.rc(t)) \quad , \tag{1.5}$$

where $0 \leq K \leq 1$ is a parameter accounting for the (inverse of the) latitude allowed to an organism with regards to possible deviations of its Ratchet Capital RC away from the evolved expected capital level rc. In this equation, the allocation is, *first*, dependent on both capitals $FC(t)$ and $RC(t)$ being provided for. *Second*, the dependency on $RC(t)$ will be altered by $K.rc(t)$ which represent the lowest limit the RC is allowed to reach at that time as defined by both the mean reproductive schedule evolved by its ancestors rc and the latitude parameter K. If the organism's flexibility is large for an individual to drift from the central trajectory (K small), then RC will only play a preponderant role in late-life (senescence) as discussed in section 1.2.1. By contrast, if no leeway is given to the individual trajectories ($K \approx 1$), the effects of RC may be felt much sooner than that. *Third*, the dependency on $FC(t)$ is modulated by $stor$, the storage capacity of the organism. As with RC and sfc, $stor$ corresponds to an evolved adaptation of the FC. For this role to be implemented mechanistically, a high limit on the reproductive effort needs to be added in order for the buffer effects of the storage capacity to be activated. We chose $stor.\overline{RE}$ for this high limit, as it has to be positively related to both $stor$ itself and to the mean effort allowed by the mean environment $\overline{RE} = \overline{Env}$ (see section 1.2.1).

Equation 1.5 can be simplified when considering the particular case where the allowed reproductive efforts do not vary with age (i.e., $re(t) \approx \bar{re}$) then, from eq. 1.4 page 13, $rc(t) = \sum_{i=t}^{e(0)} re(t) \approx (e(0) - t).\bar{re}$ and $sfc \approx \frac{e(0)}{2}$. And then we can write the maximum allowed reproductive effort at time t as :

$$aRE(t) \approx \min(FC(t) - stor.\overline{RE}, RC(t) - \bar{re}.K(2.sfc - t)) \tag{1.6}$$

Equation 1.6 shows the allocation towards reproductive effort to be, as intended, only dependent on the capitals (FC,RC) with evolved life history strategy indicators (sfc, $stor$ and K) delimiting the degree of latitude the capitals have on the process. As can be read from the equation, for organisms that are fast (low sfc) and limited in storage (low $stor$), the allocation becomes a direct function of the levels of the capitals : $aRE(t) \approx min(FC(t), RC(t))$. For a long-lived and/or capital-breeding organism, the high values of sfc and $stor$ control the reproductive effort expenditure and thus safeguard the future buffering capabilities these parameters allow.

In other words, RC will have effects whenever the environment encountered by an individual (on average $\overline{RE}$) can be vastly different from the one its species evolved in (on average $\bar{re}$). And, in particular, if $\bar{re} \ll \overline{RE}$, implying $RC(t) \ll rc(t)$, and if K is large. In a such a case RC would be the limiting capital (i.e., the capital generating the costs), even early in life, in order to preserve flexibility in the organism evolved slow schedule. This mechanism would be beneficial for the organism were the patterns of environmental variations, it has evolved in, to recur.

Equation 1.6 thus shows that high *stor* and sfc move the allocation away from a direct function of the minimum of both capitals. However it also reveals, via the product of K by $(2.sfc - t)$ that for slow organisms (high sfc) a larger latitude with regards to the evolved strategy (ie., a low K) may be allowed. And all the more so when t is low, that is in early life. The sheer generation time of slow organisms enables them a larger drift from the central evolved trajectory than fast individuals.

In equation 1.6, there is a clear symmetry between $FC/stor/\overline{RE}$ on one side and $RC/K.sfc/\bar{re}$ on the other. There is also, however, a major difference with respect to dependency on time t that brings to light the two sides of physiological costs discussed in the introduction section 1.1 . A capital that moves *forward* $FC(t)$ and another capital that is, surely, spent as time passes *forward* $RC(t)$ but managed, *backwards*, i.e., controlled by the evolved life history strategy of the organism $\bar{re}.K(2.sfc - t)$.

Overall, the higher *stor* and sfc, the more subtle and complex we expect the allocation process to be and the less directly dependent on current environmental conditions $Env(t)$. To the contrary, for fast organisms, unable to store resources (i.e. with small sfc and *stor* ≈ 0), the allocation would be much simpler - $aRE(t) = min(FC(t), RC(t)) = min(Env(t), RC_0 - \sum_{a=1}^{t} Env(a))$ - and almost directly depend on the current environment.

Secondary divers of the allocation process towards allowed reproductive effort Such an allocation function, as the ones put in equations 1.5 or 1.6, hopefully encompasses the diversity of reproductive efforts, between species, populations, individuals, and at the level of the individual throughout its life that have been observed and studied. It can also allow us to determine some of the secondary drivers of allocation to reproductive efforts, and how we may expect the reproductive effort to change over an individual's life.

Other components of life history strategy. This allocation model is life-history-strategy-dependent by construction and in particular incorporation of rc in equation 1.5. Simplifying this allocation process (from eq.1.5 to eq.1.6) evidences the effect of the position of an organism on the SFC and the ICB (via sfc and *stor*) on allocation : slower organisms and capital breeders can spare capital (respectively in the RC and as storage in the FC). However for an organism, the evolved sequence of $rc(t)$ actually encompasses all aspects of the life history reproductive schedule of an organism. Therefore all other (than the SFC and the ICB) conceivable strategy spectra, like obviously the degree of iteroparity/semelparity - seemingly difficult to reconcile with the the SFC (Dobson and Oli, 2007) - also drive the allowed reproductive effort $aRE(t)$ (see for instance (Calow, 1979) for a detailed analysis of the effects of semelparity/iteroparity on the schedule of reproductive efforts).

The environment. The environment, via the effect on acquisition, is evidently also a strong driver of reproductive effort. That the allocation of energy towards reproduction is a function, in relative terms also, of the acquisition levels and thus of the environment has been established in nature (see for instance Erikstad et al., 1998; Christians, 2000; King et al., 2011) and studied theoretically (Fischer et al., 2009; Descamps et al., 2016). This is accounted for in our model, as $FC(t)$ here corresponds to the fluctuating capital after feeding; it is made of the resources just acquired from the current environment added to the stored resources from previous seasons.

Age, state and terminal investment. On the income side of *stor* (i.e. low *stor*) provided that RC allows it, there is no reason for the organism to save more than what is required for its immediate survival, and thus the allocation is expected to be dependent on the absolute level of FC. This is also expected to be the case for the CB-FC and the RC as the latter reaches zero (and thus expectedly at old ages, or in general, in "poor" states). Since the organism is expected to die soon, it seems optimal to allocate as much to reproductive effort as possible, as a last push on fitness; a phenomenon denoted as the terminal investment strategy. The relationship between allocation towards reproduction (reproductive effort) and age or state, has been studied theoretically (e.g. by Fisher, 1930; Pianka and Parker, 1975)). In particular Charlesworth and Leon (1976) uses the reproductive effort model devised by Schaffer (1974) to provide conditions on life histories that favor the generally expected increase in reproductive effort with age. (Clutton-Brock, 1984) provide an empirical review of the weak demonstrations of such "terminal investments", which is however well established in fish (Constantz, 1974). Specific life histories, albeit

long-lived, may however maximize reproductive effort, even in relative terms, before reaching late age or poor health, as seems to be the case for male red deer which peaks at prime age (Yoccoz et al., 2002).

From allowed to realized reproductive effort: position on the quality-quantity spectrum So far, in this section, we have explored the drivers of the allocation of capitals/resources towards *allowed* (i.e., *maximum* possible) reproductive effort $aRE(t)$, the effort the organism can sustain given its current state. The actual effort, the *realized* reproductive effort, $RE(t)$ will maximized by $aRE(t)$, but may also, by chance, be much lower. This individual stochasticity in the realization of reproductive effort is likely to be driven by basic reproductive effort bre, the indivisible reproductive effort it takes to produce one independent offspring. More specifically, at each time-step t, even if both capitals allow for a reproductive effort $aRE(t)$ to be produced, such an effort would be pointless, if $aRE(t) < bre$. Instead, the resource will be reinvested in the capital and, if possible, used for later reproductive efforts.

The actual realized reproductive effort $RE(t)$ is thus a random variable, distributed between 0 and $aRE(t)$ in steps of size bre. Individuals may, by chance, only produce a portion of the reproductive effort their resources allow; and all the more so if bre is large. We can therefore introduce a new quantity - $gr = \frac{bre}{\overline{re}}$ the ratio of basic reproductive effort to mean reproductive effort - that we call the *granularity of reproductive effort* of an organism. This indicator gr, that drives the strength of the effect of individual stochasticity on reproductive effort, clearly positions the organism on the quantity-quality spectrum (Lack, 1947; Smith and Fretwell, 1974). Organisms with very low gr will produce many seeds per period and individual stochasticity will have little effect on reproductive effort : $RE(t) \approx aRE(t)$. Conversely, an organism with very high gr ($gr \approx 1$) will only be able to produce 1 offspring per period. Even with available resources, if it misses that opportunity, $RE(t)$ for that period will be 0 and the unused resources will be reinvested in the capitals.

Indirectly, this model thus predicts that capital-breeding, which we know to be a response to constantly varying environments as storage buffers the environmental variations, may also have co-evolved with a shift of organisms towards the quality side of the quantity-quality spectrum. Indeed a high $stor$ provides cushion from the overall variance of $FC(t)$, that is (as can be seen from its periodic balance-sheet summarized in equation 1.3) from the variance in its proceeds - $Env(t)$ - and its expenditure $Re(t)$, an increasing function of both the allowed effort $aRE(t)$ and the granularity gr.

In summary, we thus expect the allocation towards reproductive efforts to be determined by both capitals themselves and the organism's position in the SFC, the ICB and the quantity/quality spectra. Because RC capital decreases with age, and that both capitals can be said to constitute (a part of) the "state" of the organism, we thus expect reproductive efforts to be functions of (among many others) the age and the state of the individual in the framework of its species evolved life history. This relates perfectly to Williams definition where the costs are about both the life history of its species, driving the general allocative strategy, and the environment-dependent state of the individual tailoring a specific life-trajectory around this baseline strategy.

Furthermore, The combined consideration of allocation equation 1.6, capitals mechanisms equations (eq. 1.3 and 1.2)and the stochastic process turning allowed effort aRE into realized reproductive effort RE, hints at two buffering characteristics of the costs of reproduction that we will underline more specifically later when considering detectability (section 1.2.3). From eq. 1.3, we know that $FC(t)$ would benefit from an above average environment $Env(t) \gg \overline{RE}$, from eq 1.6 we know that this large $FC(t)$ will not be entirely spent (if $stor > 0$), and a subsequent poor environment will therefore be compensated : *physiological costs buffer environmental variance*. Similarly, if, by chance, $RE(t) \ll aRE(t)$, then from eq. 1.3 and eq. 1.2, we know the capitals are unchanged, but new acquisition from $Env(t+1)$ (for FC) and lower $rc(t+1)$ as time passes (for RC) imply that both sides of eq 1.6 will be large than before : $aRE(t+1) > aRE(t)$: *physiological costs buffer individual variance.*

Components affected : survival and fertility costs of reproduction

The separation of the costs of reproduction in "survival cost of reproduction" and "fecundity costs of reproduction" was made early by Bell (1980) in a study of the emergence of semelparity, as evolved from iteroparity in organisms encountering various costs of reproduction. Most empirical studies however focus mainly on one these costs like (Tavecchia et al., 2005) on survival cost and (Bell et al., 1977) on reproductive cost.

As previously discussed, IB will incur large and immediate FC costs. Therefore the next breeding season will be preceded by a feeding season resetting the FC. Such costs will be immediate survival costs. The reserves of CB will allow them to reduce such an immediate survival risk, but will generate delayed reproductive costs. Indeed, the fat stored by a CB will help him buffer its mortality risk due to the current reproductive effort, before it is able to feed again, by allowing to draw on reserves. By doing so, however,

the CB will reduce the resources available to it for reproduction in the following periods, with the effect decreasing with time (as the accumulation of newly acquired resources buffers the costs) and which duration is related to *stor* itself. Conversely, we expect the RC costs to be felt, mainly at old ages, especially for organisms on the slow side of the SFC, on both fertilities and survivals .

Our reasoning with regards to the determining effect of life history strategies on costs of reproduction, seems consistent with a review from (Hamel et al., 2010) whereby fast-living small rodents (like the bank vole) demonstrate higher immediate survival costs and lower reproductive costs of reproduction than long-lived ungulates. The model their propose to analyze such data predicts that the level of the costs to pay on survival or fertility depends on the variance of the fitness component for the specific life history of the organism : fast organisms have larger variance on survival, and thus that is where it the costs should lay. In their study however, since most fast-lived rodents are actually income-breeders, and most slow-paced ungulates are capital-breeders, their data confirms the expectation of stronger short-term survival costs for the former and stronger mid-term reproductive costs for the latter. On closer inspection, both approaches can also be found to be connected: the variance inferred on the allocation towards current survival for an income-breeder by reproductive effort is much higher than that of a capital breeder since the latter can buffer this variance and pass it on towards reserves to be used for future reproductive efforts.

Importance of the reproductive effort schedule

Reproductive efforts consists of all expenditures required to turn food into independent mature offspring. In most species they obviously consist in efforts made around production of offspring. In this case, costs of reproduction are often equated with costs of reproductive success. And indeed most studies focus on breeding success, since it is both a central component of reproductive effort and relatively easy to observe. However studying survival costs of reproduction in the black-legged *kittiwake*, Aubry et al. (2011) have shown that breeding attempts was a better predictor of future survival than clutch size, brood size or breeding success.

In many species, reproductive efforts have to be made long before birth, via the onthogenesis of the reproductive system, and also other secondary sexual characters are such cases, such that physiological costs of reproduction are sometimes paid long before reproduction ever occurs. In red deer, for instance, a major component of reproductive success is the size of the young adult which directly influences its chances of breeding. Because of the delay between this reproductive effort and breeding, these costs of reproduction have the peculiarity of applying to survival at early ages (even before maturity): large males have a higher mortality rates before reaching adulthood (Clutton-Brock et al., 1985).

In most species also, reproductive efforts do not stop at birth when offspring cannot feed nor protect themselves yet. In mammals, lactation is a substantial component of the reproductive effort and (Clutton-Brock et al., 1989) has even shown that in wild red dear (cervus elaphus) costs of gestation are slight compared to those of lactation. All such behaviors delaying the bulk of reproductive effort after birth are regrouped under the term of parental care. Parental care is an integral part of reproductive efforts, and can be very costly ((van den Berghe, 1992; Santos and Nakagawa, 2012)), especially in altricial species at the slow end of the slow-fast continuum, like humans; see (Gross, 2005) that studies the evolution of parental care in the general framework of Williams' principle. The importance of postnatal reproductive efforts can be inferred from behaviors such as juvenile wastage, common to both animals (Tait, 1980) and plants (Stephenson, 1980; Gosling, 1986)

More generally, the length of the entire reproductive process, spanning from birth of parent until long after birth of offspring, dilates the time window of the costs, whether immediate or delayed. Introduced in section 1.2.1, the reproductive effort schedule $res(t)$ represents, for an organism, the time distribution of reproductive effort required to produce one independent offspring. Such a distribution can be centered at time of birth of offspring, in which case $res(t), \quad t < 0$ represent efforts produced before birth and $res(t), \quad t > 0$ post-natal efforts (characterizing res by time difference from birth of offspring is a simplification as many other parameters certainly play a role, as for instance age of parent).

As we saw, the *reproductive effort schedule* convolves with fertility rates by age/state $f(a)$ to produce the lifetime distribution of all reproductive efforts (equation 1.1). Apart from the theoretical case where res is only non-zero at time of birth - that is when $res = \delta_0$, where δ is the Dirac delta function - the time distribution of reproductive efforts re will stretch wider than that of the so-called reproductive period embedded in f. Indeed, if the reproductive effort schedule is spread far and wide before and after birth - i.e., if $res > 0$ for $t \ll 0$ and $t \gg 0$ - the lifetime distribution of all reproductive efforts will span far wider than the fertility schedule.

This means that if, for the production of an offspring, the reproductive effort components exerting the

largest costs are not easily identified, it will difficult to distinguish immediate, short-, mid- and long-term costs. As a matter of fact, if $res(t)$ is mainly related to one specific effort - like gestation - or two major consecutive efforts - like gestation immediately followed by lactation - it will be easier to relate reproductive efforts and their costs and to assess the level and delay of the latter, than if res is evenly spread over time, made of multiple small efforts over a long period of time. For a female ungulate, the reproductive effort schedule is mainly concentrated at the level of the season, in the time window of which it, in turns, acquires resources, then breeds, gestates, gives birth and finally feeds and cares for its offspring(s) of the year. For such an individual, it will be possible to equate death during that season or reduced fertility in the next season(s) with costs of that particular reproductive period. For a 35-year old human female, it is nearly impossible to the designate, as the origin of a drop in fitness (reproductive pause, health deterioration, ...), the complicated gestation she currently has, or the feeding of her newborn child, or even the care she takes for her first three children that still depend on her.

1.2.2 Genotypic level of the costs of reproduction

In the previous section, we have investigated the mechanism of physiological costs of reproduction, which occur in Stearns' intermediate structure. If such a mechanism does not need genetic variance to occur, genotypic polymorphism located at the genotypic level - and constituting, inter alia, the *genetic costs of reproduction*(genetic variance in allocation or in an antagonistically pleiotropic gene acting on vital rates) - may obviously still be involved. Such a genetic variance would affect, together with the *physiological costs of reproduction*of the intermediate structure, the phenotypic level where costs of reproduction are observed. A realistic population model for costs of reproduction therefore needs to be able to implement genetic variance.

A genetic basis for *physiological costs of reproduction*

In the context of *physiological costs of reproduction*explored in section 1.2.1 where life trajectories are determined by the combinations of the effects of reproductive efforts on the capitals (equations 1.2 and 1.3) and the effects of the capitals on reproductive effort (the allocation process, described in equation 1.6), only two functions are dependent on the individual. First the acquisition process turning current environmental conditions into $Env(t)$ in equation 1.3. Second, the allocation function itself (eq. 1.5 and 1.6) which also depends on the life history strategy rc evolved by the population (with manifestations sfc and $stor$ among others). As a matter of fact, in section 1.2.1 discussing the intermediate structure of the costs and in particular in section 1.2.1 describing the allocation process, we have already hinted at that influence by calling rc the reproductive schedule evolved by the lineage of the individual, which may differ between different lineages in the population. Therefore genetic variance in either allocation and acquisition would have effects on the costs of reproduction (black arrows in figure 1.1).

We call *allocation* gene, a gene acting on the process of allocation itself, where two different alleles would, everything else being equal, allocate towards different level reproductive efforts. The genetic variance in such an allocation gene would generate a gradient of heritable life history strategies in the population.

To simplify our analysis of the topology of the genotypic level of costs of reproduction, we project the distribution of all possible reproductive schedules rc onto one of its moment, sfc , which positions the lineage on the SFC. In other terms, we summarize the diversity of lifetime reproductive schedule into its sfc expression. Then, the genetic variance in the *allocation* gene corresponds to the variance, within the population, in the slow-fast continuum (horizontal axis of figure 1.2)

We can restrain the gradients of life-pace strategies induced by the variance in *allocation* to be iso-fitness by adding another, "orthogonal" gene, the *acquisition* gene, acting on the acquisition of the resource itself (vertical axis of figure 1.2). Polymorphism in this gene would generate a variance in overall fitness in the population we call robustness (more robust individuals can acquire more resource and thus survive *and* reproduce better). For the different alleles of the acquisition gene to cohabit, the effect of such a gene need to be strongly environment-dependent with crossing reaction norms : the robust genotype in a given environment needs to be the frail one in another, since otherwise it would quickly invade the population (as do the "super-flies" of Reznick et al., 2000).

The addition in our model of variance in *acquisition* - whereby two individuals with different *acquisition* but the same *allocation* would differ in overall fitness, i.e. in investment towards both fertility *and* survival - allows thus to refine our definition of the effects of the variance in *allocation* : two individuals with different *allocation* but the same *acquisition* would differ in the way they allocate the *same* amount of resources towards either current fertility *or* survival (i.e. prospective future reproduction). This hints at the possibility to use a different set of axis to position the different *allocation* $\times$ *acquisition* genotypes of a population : *fertility* and *survival* (diagonal axes on figure 1.2). Robust genotypes (like G1 on figure

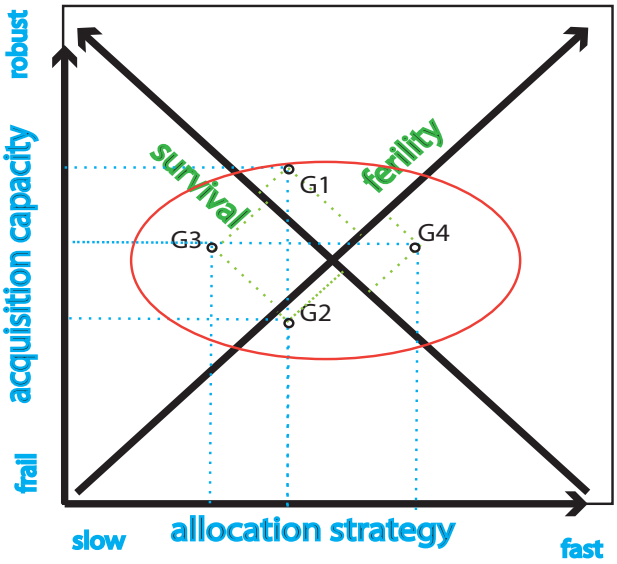


Figure 1.2: Genotypic map : 4 genotypes of a population, G1, G2, G3 and G4, are represented on a genotypic map according to two sets of coordinates (equivalent and related by a 45° change of basis), the *allocation* × *acquisition* corresponding to variances on the slow-fast continuum and in robustness and the *fertility* × *survival* corresponding to relative investment towards survival or fertility. The two sets are Genotype G1 is the fittest (highest acquisition capacity) with a central position on the slow-fast continuum. Genotype G2 has the same position on the slow-fast continuum (the relative investments towards fertility and survival are the same), but is less fit (both survival and fertility are lower in absolute terms). Genotypes G3 and G4 have have the same, intermediary, position on the robustness axis (fitter than G2, less fit than G1) and are therefore iso-fitness. They however differ by their position on the slow-fast continuum. G3 is a slow organism favoring survival at the cost of fertility. G4 is fast with opposite investments. Variance on the *acquisition* axis will be allowed by environmental variance but kept in check, in the long evolutionary run, by selection. On the contrary, variance along the *allocation* axis will only be limited by the extent of life history variations a population is able to sustain before it looses the capability of interbreeding (speciation). And thus it is ultimately the latter variance which is measured between species. Even, within species however we then expect, in general,t he variance along the *allocation* axis to be larger than the variance in *acquisition* (ellipse shape). If it is not the case (because of a large environmental variance for instance) the *genetic costs of reproduction*at play in the population will not be detectable (van Noordwijk and de Jong, 1986; Houle, 1991)

.

1.2) would invest more towards survival *and* towards fertility than frail ones (G2 on the figure). Faster genotypes (G4) would promote fertility, but at the cost of survival, whilst slow ones (G3) would favor longevity at the cost of current reproduction.

Such a mechanism with both an allocative intermediate structure and variance at the genotypic level, we call *physiological costs of reproduction with genetic basis* and their overall architecture is depicted in figure 1.1. As such trade-offs have a genetic basis, they now can evolve. Particular environments will favor particular allocation strategies and the genotypes that have higher fitness in that environment. In that regard, the purely physiological mechanism with no genetic basis described in section 1.2.1 can be considered to be the result of an environment stable in a particular state for long evolutionary times.

Adding variance at the genotypic level, does not change the capital allocation mechanism at the core of the intermediate structure. Indeed such a mechanism operates at the level of the individual not at the level of the population. However the allocation strategies and the acquisition capabilities now vary according to the individual's situation on the *allocation* × *acquisition* genotypic landscape (see 1.2). Individuals of the same lineage will have same acquisition and allocation processes, acquiring the same amount of resources in the same environment and allocating these same capitals towards equal reproductive efforts. The only differences in life history trajectory for such clones sharing the same environment would be due to chance.

As a matter of fact, the presence of an (acquisition/) allocation physiological process in the intermediate structure is not necessary to generate the variance in *robustness* × *slow-fast* (as they could equivalently be called when not referring to any physiological process) of figure 1.2.

Genetic non-allocative costs of reproduction

In the introduction, section 1.1, we called *genetic costs of reproduction*, the variance in genes that are antagonistically pleiotropic with regards to investments towards fertility and survival. As such genes express their variance in the various genotypes of the population, a negative correlation arises between early and late fitness, similar to that produced by *physiological costs of reproduction*. This similarity in phenomena prompts ecologists to use a common term for the mechanisms, costs of reproduction, physiological for the latter, genetic for the former (see figure 1.3).

The genetic basis of the *physiological costs of reproduction with genetic basis* just encountered in section 1.2.2 is thus also, in itself, a genetic cost. However we can also, at least theoretically, construct genetic costs that do not require any physiological allocative mechanism to occur. We call them *genetic non-allocative costs of reproduction*. In such costs the pleiotropic gene would have a direct antagonistic effect on both fertility and survival instead of directly promote one fitness component at the indirect cost of another. In particular, *genetic non-allocative costs of reproduction* would not be associated with a resource or a capital that needs to be shared amongst various functions. As such *genetic non-allocative costs of reproduction* may be related to *ecological costs* (see introduction section 1.1). A simplistic gene which would express itself via the coloration of the skin in a specific color attracting both mates and predators would belong to that category.
In figure 1.3, we represent the topology of costs of reproduction we happened upon in this chapter. As we can see from that figure, *physiological costs of reproduction with genetic basis*are included in the *genetic costs of reproduction*. However, this is not a double inclusion inducing equality, because of the *genetic non-allocative costs of reproduction* we have just discussed.

1.2.3 Phenotypic level and detectability

In this section, we discuss and analyze the effect of several parameters - chiefly the environment, its variance, demographic variance and genetic variances - on the emergence of the costs of reproduction at the phenotypic level, and on their detectability. A summary is presented in box 1.1.

From physiological to fitness costs : the effects of absolute level of environmental abundance on physiological costs

Physiological costs of reproduction are deemed to occur at all times in all organisms. By definition of the Fluctuating and Ratchet Capitals devices described in section 1.2.1, the fate of an individual will depend on such gauges. However, parameters might dampen the propagation of the effects of capital spending at time t to the capital itself at time $t+1$. This is chiefly the case for the environment with regards to FC (see equation 1.3). If the environment is good, the FC will be easily replenished thus canceling any effect of past reproductive effort, rendering the physiological costs not only undetectable to us, but mostly undetectable to natural selection: such physiological costs are not turned into fitness costs of reproduction

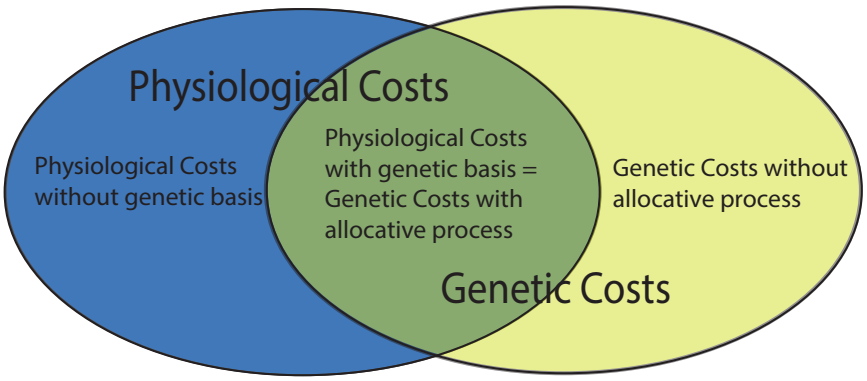


Figure 1.3: Topology of costs of reproduction. In order for negative correlations between early fertility and late fitness to emerge from a population, one of two mechanisms is required, either a physiological allocative mechanism or variance in an antagonistically pleiotropic gene (acting on such vital rates). The first mechanism, we call *physiological costs of reproduction*, are represented by the blue ellipse. The second, we call *genetic costs of reproduction*, by the yellow one. When the physiological costs, laying in Stearns' intermediate structure, also have a genetic basis, the phenotypic level will be the output of both physiological and genetic costs. Such costs are called *physiological costs of reproduction with genetic basis*and are therefore to be found at the (green) intersection of both ellipses. Away from that intersection, we find the physiological costs for which there is no genetic variance discussed in section 1.2.2 and the genetic costs with no allocation process of section 1.2.2.

in the sense of Hamel et al. (2010).

For IB, lacking the storage allowing transmission of balance of FC from one period to the next, the environment is the main driver of reproductive life history, with the role of FC costs confined to the short period between breeding and feeding. For CB, FC propagates the costs over to the next time-steps. If the environment remains "good" for the length of time reserves can be stored by such an organism, these physiological trade-offs will also not generate fitness costs. As we saw, the environment, via the effect of FC on reproductive effort 1.5, also impacts the RC trajectories (from 1.2). However, contrary to FC, the RC will always diminish as the organism makes reproductive efforts. Actually since good environments are those permitting the allocation towards reproductive efforts, and since they also allow the organism to survive until the late ages at which the RC costs are mostly felt, we would expect the latter to be as strong if not stronger fitness costs is such conditions.

In a simple simulation, in appendix 1.5.2, we show how the absolute level of environmental abundances drives the correlation between successive fertility realization. We also see that it is also strongly affected by the changes in environmental conditions.

From fitness costs to detectable ones : the effects of environmental and demographic variance at the level of the individual

Fitness costs, occurring at the level of the individual, should be observable there. However detectability, evidently proportional to the strength of the fitness costs themselves, is expected to result from the confrontation of two opposing forces : environmental and individual variance.

environmental variance Because of the strong dependency of the strength of the FC costs on the environment (see equation 1.3), environmental variance will likely blur, at the level of the individual and over time, the detectability of phenotypic costs. This is obvious for an Income Breeder, which FC costs occurs within the time-step of the season, making IB-FC costs akin to ecological costs. But even for a Capital Breeder, which fat reserves propagate costs from one breeding season to the next, it will be difficult to extract from reproductive trajectories, in a volatile environment, a strong signal for costs. This

is because the capitalization of resources in RC and CB-FC buffers the costs but also the environmental variance; thus making costs impossible to detect if the exact environmental variance, and the effect of the environment on the allocation process are unknown. The individual would draw in such reserves when needed, and the phenotype would reflect both the environmental conditions and the physiological costs. Such costs are fitness costs but the changes in successive environmental conditions have to be known in order for these fitness costs to be detectable. Otherwise the costs detected will likely be much smaller than the underlying physiological costs really are, or even go undetected. This may seem counter-intuitive as, without changes in the environment, all individuals would, time-step after time-step, allocate very similar amounts of resources to successive reproductive efforts thus hindering the physiological costs; whereas changes in environmental conditions will generate different patterns of allocation, making them seemingly more detectable. The ambiguity is removed by saying that if the exact environmental time series is known along with its relation to the life trajectory of the individual, then environmental variance may be considered to help reveal the fitness costs, but otherwise, and thus in general, conceal them. Moreover, as we shall see in the next section, the assumption of constant and equal allocations in constant environment omits the diversifying effect of demographic variance on life history trajectories.

In appendix section 1.5.2, we illustrate the effects of environmental variance and absolute level on detectability of CB-FC costs of reproduction in a simple model that shows how the correlations between successive reproductive efforts are raised by both parameters.

Overall therefore fitness costs will only reflect the physiological trade-offs in poor environments where scarcity of acquired resources means some functions may have to be drastically reduced or even shut down. And they will only be detected if this "poor" environment remains relatively constant over time. Indeed, if not controlled for, environmental variance conceals the costs. To the contrary, in the particular cases where both the environmental time series and the environment's exact effects of the costs are both known, fluctuations can increase detectability of the costs. However, we expect individual stochasticity to perform much better as a revealer of costs of reproduction.

Individual stochasticity To the contrary, at the level of the individual (or rather the genotype) detectability of the costs will be enhanced by demographic variance (also called chance or individual stochasticity). Indeed, the process of giving birth to one offspring (and the process of surviving to the next time-step) is subject to individual stochasticity : even if the capitals are large enough to be able to produce a reproductive effort aRE, random events (external, internal to the organism) may reduce the allocation ($RE(t) < aRE(t)$ or even prevent it ($RE(t) = 0$), transferring back the unused budget to the capitals.

As we just saw, for individuals of the same *acquisition* × *allocation* genotype (clones), sharing the same environment, and therefore the same capital levels, the costs of reproduction could only be detected, in a cohort analysis (tracking over time individuals experimenting equal environment) if both the exact environmental series and its effects on the individuals are known. Without individual stochasticity, a period analysis - comparing different individuals at the level of the time-steps - would not detect any cost: all individuals of the same genotype would have the exact same trajectories. In the general case, where the effects of the environment on the acquisition/allocation are not precisely known, it is individual stochasticity that will generate the initial differences between clones, that will then be further propagated over lifetime, with a snowball effect, by the costs themselves. At the level of the single time-step however, the individual stochasticity, may have a local blurring effect akin to that of environmental variance : if two individuals of the same genotype have by chance reached the same state with the same capitals, one may reproduce and not the other, by chance again, thus seemingly clouding any inference of physiological costs. However, as soon as the horizon is extended over several time-steps, individual variance is the fuel of detectability of physiological costs at the level of the genotype, as soon as the required environmental conditions (not too good, not too variable) are met. In summary, we predict that individual stochasticity will reveal the physiological costs at the genotype/individual level when trajectories are observed longitudinally. At the time-step level however the variance it generates will hinder the costs.

The difference in effects of both variances, environmental and demographic, can be better understood by observing their level of actions. Both impact all aspects of physiological costs., However, we think, the influence of the environment is stronger on the acquisition than on the allocation process, whilst, in general, allocation process will be more prone to individual stochasticity than acquisition. This is because, in general, granularity of reproductive effort (related to the mean number of offspring per season) is likely to be higher than the granularity of FC resources (related to the number of basic FC resources acquired per season). Moreover, the acquisition process, related to a parameter that is difficult to track precisely (the environment), occurs before the allocation process. It therefore acts on a lever that does not directly determine the costs (eq. 1.3 vs eq. 1.5), and thus environmental variance modulates and conceals the effects of the costs. Conversely, the allocation process is related to a parameter that is easier to observe and measure (the realization of reproductive effort, akin to reproductive success). Moreover it acts exactly

where the costs are produced (eq. 1.5). Consequently demographic variance does not modulate the costs, it only delays them. This prediction of opposing effects of environmental and demographic variance on detectability with regards to FC costs is akin to the result of (van Noordwijk and de Jong, 1986) : variance in acquisition - mainly caused in our model by environmental variance - conceals trade-offs; whilst variance in allocation - mainly driven by demographic variance - reveals them.

As mentioned above, we expect demographic variance to play a lesser role on acquisition than allocation. This will be especially the case for organisms which granularity of reproductive effort is large, with storage capacities and diverse and "small" sources of energy. The strength of the effect of such individual stochasticity on the costs of reproduction is commensurate with the granularity of reproductive efforts of the given organism, measured as $gr = \frac{bre}{\bar{re}}$ introduced in section 1.2.1 page 18. For organisms at the quantity end of the quantity-quality continuum, i.e. with small granularity $gr << 1$, the effects of demographic variance will be little. On the contrary, for organisms, producing offspring of very high quality, in very low quantity - i.e., where $gr \approx 1$ - we expect the effect of demographic variance on detectability to be much stronger. This is similar to the effect on variance in expected financial capital of two friends who decide to split 1 euro by either flipping one 1 euro coin head or tails, or by flipping a hundred 1-cent coins. As a consequence, we predict the costs of reproduction to be easier found in quality organism. The physiological costs are not weaker for quantity organisms but there mainly driven by environmental conditions; most individuals in the same environment will incur the same costs, therefore making them harder to emerge. Compare, for instance, the reproductive pause (called masting) of oak trees occurring at the same time for all trees in a patch and the reproductive pause of ungulates. In a the latter case (a quality organism) the pause is clearly related to costs of reproduction, whilst the simultaneity of the former (a quality organism) invites other interpretations than the costs for the reproductive behavior (see discussion and references in section 1.2.1 p.15).

As we know, the position of an organism on the quantity/quality line is not related to the other life-history strategy indicators - SFC, ICB and semelparity/iteroparity spectrum for instance - discussed before. Using the wording of this article, contrary to sfc and other indicators of life history pace, the quality-quantity indicator is not a moment of rc. Slow organisms can occupy both ends the quantity-quality spectrum, like humans and trees, adding another dimension to the diversity of expression and detection of costs of reproduction.

Detectability at higher levels

At the level of the population At the level of the population, individual stochasticity will here again act as a revealer and genetic variance as a concealer of the *physiological costs of reproduction*. However at that level, *genetic costs of reproduction* will also come into play as the differences in trajectories in the population will also result from genetic variance. Negative correlations between fitness components would be further enhanced by the genotypic variance in allocation strategies. Therefore, it would not be possible when focusing on inter-trajectories data points alone (comparing data between individuals, and forsaking intra-trajectories analysis) to distinguish between the expression of the physiological mechanism (the physiological cost) or the *allocation* genotypic polymorphism (the genetic cost).

Contrary to physiological costs, detectability of genetic costs will suffer from demographic variance, as it would induce noise around the gradient of life history strategies that genetic costs establish in the population. But most importantly, detectability of genetic costs will also depend on the environment in a manner that is dictated by the reaction norms -as defined by Woltereck (1909) - of the *allocation* gene with regards to traits *survival* and *fertility* (see firegure 1.2). If these reaction norms cross, that is, if a relatively slow genotype in one environment is relatively fast in another, then environmental variance will reduce the measurable negative correlations between early fertility rates and late vital rates. If they do not cross, i.e. if the different genotypes have a consistent relative strategy (towards fertility or towards survival) across environments, the negative correlations will not be blurred by environmental variance.

In both cases, contrary to physiological costs, there is no reason to believe that their detectability depends on the environment being poor. Some organisms may have evolved reaction norms for the allocation gene whereby, even in very favorable conditions, some alleles invest less in reproductive effort than others. We therefore predict that short or mid-term costs that are easily detectable in relatively good environments are more likely to be genetic costs. To the contrary, we expect such costs detected *only* in bad environments to stem from the physiological intermediate structure.

Reaction norms for the acquisition gene, on the other side, are expected to cross : selection would not allow genotypes with strong differences in fitness, consistently in all environments, to cohabit (figure 1.2). This was discussed by Reznick et al. (2000) hinting at the fact that the "super-fleas" emerging from Spitze (1991)'s study, dominating others in all components of fitness, were only "super" in a specific environment,

but inferior in others. Such a genetic variance in acquisition capabilities, i.e. in robustness, will limit and even sometimes conceal the negative correlations stemming from physiological and genetic costs of reproduction. This is in conformation with the results from van Noordwijk and de Jong (1986) adapted to the genotypic level by Houle (1991). More recent studies have fine-tuned such predictions. Descamps et al. (2016) have, for instance, shown that the dependency of the relative allocation of resource on the acquisition level - which is the case in our model where allocation depends on capitals levels themselves depending on the acquisition - has further implications on detectability of the underlying physiological costs.

Antagonistic pleiotropic trade-offs as "dynamic linkages". As we have just discussed, as genotypes interact with the environment to produce phenotypes, the negative correlations (the costs) in a given environment might become positive in another. Turning things around, instead of considering these as functions of 2 variables (the response in the trait of one genotype in one environment) as varying trait response to genotype in varying environments, we have also regarded these same functions as varying trait response to the environment for varying genotypes, that is as the continuous reaction norms - expressing phenotypic plasticity - for varying genotypes as described in (Stearns, 1989b). And we have shown that the effects of the environmental variance on detectability of *genetic costs of reproduction* would mainly depend on whether the reaction norms of the various alleles of the pleiotropic gene do cross or not. In that approach, however, the trait tacitly referred to, is actually two traits (investment towards *survival* and *fertility*). We have considered them to be one trait only, as the pleiotropic mechanism cause the change in one to be compensated by an opposite change in the other.
However, we can also consider non-pleiotropic genes acting on vital rates. In that framework, the (trait, environment, genotype) relationships discussed above become more complex as both traits have to be considered; it becomes a function of three variables. For ease of understanding, let us consider two genes only, one acting solely on *survival*, the other solely on *fertility*. They thus correspond to the alternatives axis of figure 1.2 (diagonal axes). Importantly, variance in either or both genotypes will not infer negative correlations in the population. All combinations of high and low *survival* and/or high and low *fertility* can be found. The variance in such genes does not, therefore, constitute a *genetic costs of reproduction*. However, if reaction norms of both genes cross, negative correlations may appear in changing environments. Indeed, individuals with low *fertility* and high *survival* in an environment would then have opposite features in another. If environmental conditions shift over time, the cohort study of these individuals will exhibit apparent costs of reproduction. In a theoretical study of phenotypic plasticity, Stearns (1989a) has investigated how the shapes of such reaction norms, in particular when they cross, can seemingly lead to trade-offs. Because they lack any repartition mechanism (either allocative in the physiological intermediate structure, or pleiotropic at the genotypic level), Stearns refuse them the status of trade-offs, and suggests to call them "dynamic linkages". We predict this to have strong implications on the detectability of costs at the population level. If detectability is increased when environmental variance seemingly increases, the costs detected are likely to be "dynamic linkages". In the opposite case, they are likely to physiological costs which detectability is impaired by high environmental variance. Finally, if the strength of the costs detected varies very little between different environments, the negative correlations are likely to be due genetic costs with relatively constant reaction norms.

As a conclusion, forsaking "dynamic linkages", the level of the population is a battleground, whereby both physiological costs and genetic costs are deemed to coexist. The population level is therefore the ideal level at which to study the interactions between physiological and genetic costs. Because of the effects of environmental variance on the former at the level of each individual/genotype, and of acquisition heterogeneity on the latter between individuals/genotypes, the combination of both costs in the population may not result in negative relations between early fertility and late fitness. Even if it does, it will be hard to disentangle the effects due to each cost unless individuals are tracked through life and costs appear in particular environmental conditions; for instance in conditions known to be good for the studied species (genetic costs), or only in conditions known to be poor (physiologial costs).

Between populations Whereas variance in *acquisition* is kept in check by selection, the iso-fitness variance in *allocation* can extend (figure 1.2). Over evolutionary time, as the differences in evolved life history strategies in the population have grown large, and since, for instance, the population is large enough for different locations to sustain different environments favoring different life history strategies, speciation will occur. The differences, larger still, in environments encountered by the split populations will further increase the variance in allocation strategy at this inter-population level (see figure 1.4). Within each population, there will be variance in acquisition, as allowed by the local environmental variance, and variance in allocation around the central strategy evolved by the species. Between populations however, the differences between the mean strategies evolved by each species given its mean local environment, will predominate, making the costs of reproduction easier to detect at the inter-population level than at the intra-population one.

We can simply formalize this by applying to the environmental variance the law of total variance to such a system : $Var(Env) = E(Var(Env|Pop)) + Var(E(Env|Pop))$, where Pop designates the population an individual belongs to, Env is a multidimensional vector characterizing the environmental conditions. The first component of the sum therefore computes the mean (over all populations) of their internal environmental variance. It is the intra-population variance $\sigma^2_{intra} = E(Var(Env|Pop))$. The second component is the variance of the populations' mean environments, the inter-population variance : $\sigma^2_{inter} = Var(E(Env|Pop)$. And thus $Var(Env) = \sigma^2_{intra} + \sigma^2_{inter}$. σ^2_{intra} corresponds to short term variations of the environment, at the level of the population, that will generate variance in acquisition at the level of the physiological costs. σ^2_{inter} corresponds to settled differences in environments between population to which life histories have adapted. Such a variance will translate into a gradient in allocation strategies at the genetic costs level. And thus we can display both variances along the *allocation* and *acquisition* axes of the genotypic map (figure 1.2). The detectability of inter-population costs of reproduction, will depend on the ratio of intra-population variance to total environmental variance as $1 = \frac{\sigma^2_{intra}}{Var(Env)} + \frac{\sigma^2_{inter}}{Var(Env)}$, i.e. on the ratio of dimensions, on the genotypic map, of the grouping of all individuals (fig 1.2).

We would expect the various species of a "comparative method" study to have adapted to environments that are more diversified (between the species) than they are fluctuating (within each species), all the more so if the species are phylogenetically very distant. Statistically, this means we expect the variance in mean environments σ^2_{inter} to be larger that the intra-population variances σ^2_{intra}, thus generally avoiding the type II statistical error of rejecting the existence of the genetic trade-offs (σ^2_{inter}) because of a large σ^2_{intra}. As a matter of fact, in the literature, most demonstrated trade-offs stem from interspecific "comparative data" analyses, whilst negative correlations often fail to emerge from intraspecific studies. This was, in the case of the quantity-quality trade-off, demonstrated by Bernardo (1996) and Christians (2000).

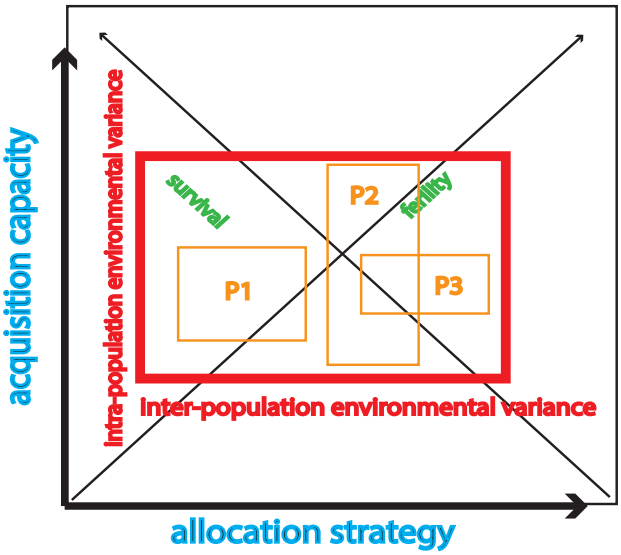


Figure 1.4: Inter-population genotypic map : we depict the position of several populations, P1, P2 and P3, on the *acquisition* × *allocation* genotypic map. Each population will have an acquisition variance related to its environmental variance. Its mean position on the allocation axis (or slow-fast continuum), on the other hand, is an adaptation related to the mean environment this population has evolved in (P1 is a slower organism than P2, itself slightly slower than P3). At the level of the inter-population study, the variance in allocation is therefore akin to the variance in mean environments between population, whereas the the variance in acquisition is associated with the mean of the intra-population environmental variances. Thus we expect such a study, if the populations studied are distant enough in the tree of life or adapted to different enough environments, to display a larger variance in allocation than in acquisition and consequently to make the *genetic costs* detectable, whereas it may not be the case for the populations of the study themselves (here for instance P2 has a larger variance in acquisition than allocation)

Table 1.1: Summary of detectability patterns for physiological genetic costs of reproduction

Physiological Costs are	- revealed in bad and concealed in good environments
	- revealed by demographic variance in longitudinal analysis
	- concealed by environmental variance if fluctuations are not controlled for
⇒ important at the individual/genotype level (only cost there), together with Genetic costs at the population level, but loses importance in interspecific studies	
Genetic Costs are	- revealed by variance in *allocation* gene (related to mean environment encountered by population over evolutionary time).
	- concealed by variance in *acquisition* or *robustness* gene (allowed by environmental variance)
⇒ irrelevant at the level of the individual/genotype, important at the population level together with Physiological costs, primordial in interspecific studies	

Detectability and phenotypic plasticity. The difference in detectability in populations and between species may be related to the seemingly irreconcilable positions taken by Waddington (1953) and Wright (1931) and Waddington on genetic/evolutionary effects of phenotypic plasticity, as disentangled by (Stearns, 1989b). Wright saw phenotypic plasticity as "reducing the amount of genetic change in evolution", Waddington saw it as "creating more opportunity for genetic change". As elegantly proved by Stearns, those views "only vary because they apply to different time scales and evolutionary situations". Stearn's synthesis is consistent with our approach where at the population level, the *allocation* gene has limited variance around the strategy evolved by the species, but the flexibility of the physiological allocative process buffering environmental variance at the horizon of the organism's longevity generates phenotypic plasticity. At the taxa level, wide changes in environments will have fixed very different *allocation* genotypes across populations thus effectively transferring the mechanism generating the phenotypic plasticity from the intermediate structure to the genetic level.

Detectability in nature

When working at the level of the individual trajectories, the *physiological costs of reproduction*costs are indeed, as expected, detectable. This is demonstrated by many studies on birds using brood manipulation (see, for instance, Boonekamp et al., 2014; Dijkstra et al., 1990). When manipulation of the reproductive effort - artificially generating individual stochasticity - is not feasible, it is still possible to inflect the allocation by worsening the environment, modifying environmental abundance, as Maynard-Smith did when studying costs of reproduction in *Drosophilia* (Maynard Smith, 1958).

Most studies however focus on the level of the population, where theory expects the *physiological costs of reproduction*and *genetic costs of reproduction*to be hindered, in their producing of negative correlations between early fertility and late fitness components, by environmental variance and genotypic acquisition variance (i.e. robustness heterogeneity). On aggregate at the level of the population, this may even induce a positive correlation whereby individuals - as robust ones gradually take over frail individuals in successive age classes - seem to actually benefit from reproduction , as in (Hamel et al., 2009), a phenomenon known as selective disappearance or frailty effect.

As expected as well, the *physiological costs of reproduction*are less easy to detect in "good" environments; for instance in captivity (Tarín et al., 2014; Ricklefs and Cadena, 2007; Kengeri et al., 2013) or in environments that are clear of evepidemics (Garnier et al., 2015) or not cold or dense enough (Tavecchia et al., 2005; Hamel et al., 2009). On the contrary, studies comparing different populations or species do display negative correlations, as the different populations have evolved different mean allocation strategies, adapted to different environments and life histories; as for lizards in (Tinkle, 1969). Charnov (2002)'s classification of life history using dimensionless indicators illustrated by "life-history cube" can be interpreted as a theoretical proof of the interest of between-clades comparative studies in order to detect physiological or genetic trade-offs.

1.3 Models : Towards an evolutionary model for physiological trade-offs with genetic basis

In this section, we first relate the two main trade-off mechanisms, physiological trade-offs (occurring continuously at the level of individuals) and genetic trade-offs (which are negative correlations between pair of traits at the level of the population) to two families of models, Agent-based and Matrix projection models. We then show that these model properties can be combined, thanks to the development of what we call Multitrait Population Projection Matrices (MPPMs), in order to model physiological trade-offs in an evolutionary context. Thereafter we hint at a way to incorporate, in an MPPM, the various components of physiological and genetic costs of reproduction as defined and described in 1.2.

1.3.1 Existing models implementing different aspects of costs of reproduction

Two main type of models are adapted to model the two main types of cost of reproduction we have brought to light: *physiological costs of reproduction*and *genetic costs of reproduction*.

Individual Based Models for physiological costs

As already mentioned in the introduction, the cumulative acquisition/allocation process of *physiological costs of reproduction*, working at the level of the individual, of *physiological costs of reproduction*, can be modeled via Individual-Based Models (IBM also called agent-based models) that track each specific individual during every step of its life-history. See for instance an agent-based model for the costs of reproduction in ungulates by Proaktor et al. (2008) and another one investigating the implications of acquisition-dependency of resource allocation by (Descamps et al., 2016).

Individual-based models or microsimulations as they are known in demography, can indeed account for such allocative costs by implementing, at the level of each individual, specific allocation and acquisition processes that are functions of the environment and the acquired and stored resources. They also make it possible to incorporate heterogeneity classes in the population. The levels - for an individual - of its Ratchet and Fluctuating Capitals (defined in section 1.2.1) would, along with its age and other life-history traits, define its individual state. The output of such a model consists in the stochastic response, that is the new state of the individual and its offspring, to the different random processes affecting the organism's life history. Among such processes, in the case of an IBM modeling costs of reproduction, would one find, at least, an acquisition process (the process turning a genotype in a given environment into $Env(t)$) and most importantly an allocation process, as for instance the one defined in equation 1.5.

Generating many runs, over long running times for given or stochastic environments, an individual-based model will provide expectancy and variance of many demographic parameters. Thanks to their level of details, such models are more precise and more flexible population projectors than matrices (Van Imhoff and Post, 1998). But, contrary to matrices, projecting the population as a whole, they find it very difficult to demonstrate the generalization of simulation results and to qualitatively ponder the weights of the various parameters that influence the population fitness (Caswell and John, 1992). Sensitivities and elasticities, measuring the effects of any vital rate on any individual demographic measures (net reproductive rate, reproductive value) and any population asymptotic measure (growth rate, abundances) that are at the core of evolutionary demography are population projection matrices' bread and butter (de Kroon et al., 1986; Caswell, 2001; van Tienderen, 2000).

Population projection matrices for genetic costs

Matrices are the ideal tool to model *genetic costs of reproduction*, as their elementary elements are the vital rates for a given genotype and environment. These matrices, whether modeling age-structured (Leslie, 1945), stage-structured (Lefkovitch, 1965) or size-structured (Usher, 1966) populations allow to project the population over any amount of time-steps. Most importantly, in evolutionary demography, they allow to calculate the asymptotic growth rate, abundances and reproductive values of each state (i.e. class or category) of the population and the sensitivity of these ergodic measures (Caswell, 2001, 1978; Demetrius, 1969). Such models are used to investigate how the life-history parameters (chiefly fertility and survival rates) can optimize fitness (measured via the net reproductive rate or the population growth rate) when constrained by genetic trade-offs like the genetic costs of reproduction. This has been used, for instance, in order to understand in what conditions semelparity can evolutionary emerge and fix (Bell, 1980; Cole, 1954).

Optimality theory Conversely, turning the argument around and considering that the category-specific vital rates observed for a population are the manifestations of an Evolutionary Stable Strategy (Parker and Maynard Smith, 1990), some authors then use matrix models to deduce the constrains between various traits influencing fitness. Population projection matrices are useful model for structured populations as they enable easy sensitivity and elasticity analysis of ergodic growth rate λ (the maximal eigenvalue of the matrix) to vital rates (the entries of the matrix) as shown by Caswell (1978). Considering this ergodic growth rate - taken as fitness - to be (locally) optimal implies that vital rates changes are constrained by their sensitivity values, and that a positive change in, say, fertility at age α, $f(\alpha)$ would infer a negative change in survival at age β, $s(\beta)$, with the ratio of changes (i.e. the constrain) equal to the ratio of sensitivities : $\frac{\partial\lambda}{\partial f(\alpha)}/\frac{\partial\lambda}{\partial s(\beta)}$; (see Caswell, 1984, 1982c; Van Tienderen, 1995, for detailed analysis).

Quantitative genetics approach Such optimization models, revealing genetic trade-offs inherent to particular life histories, appear very similar to the quantitative genetics approach. Indeed, the variance-covariance genetic matrix $\mathbf{G}$ on vector of traits $\boldsymbol{tr} = \begin{bmatrix} f(\alpha) \\ s(\beta) \end{bmatrix}$ is such that the generation change in the mean value of these traits is given by $\frac{d\boldsymbol{tr}}{dt} = \mathbf{G}.\nabla\lambda$ where $\nabla\lambda = \begin{bmatrix} \partial\lambda/\partial f(\alpha) \\ \partial\lambda/\partial s(\beta) \end{bmatrix}$ is the vector of selection gradients for the traits or, in evolutionary demographic terms, the vector of sensitivities (Lande, 1982). At ESS, $\frac{d\boldsymbol{tr}}{dt} = 0$ and thus $\mathbf{G}$ is a function of the ratio of sensitivities: $\frac{\partial\lambda}{\partial f(\alpha)}/\frac{\partial\lambda}{\partial s(\beta)}$ used in optimality theory. A complete comparison of the evolutionary optimality theory and quantitative genetics approaches was performed by Charlesworth (1990). He demonstrated that "under suitable conditions (including weak selection), useful approximate formulas for the relations between the functional constraints and the additive genetic variance-covariance matrix can be derived [which]... can be used to show that the conditions for equilibrium under selection according to the two different approaches are approximately equivalent".

Extension of matrix models to stochastic matrix models Even though population-based and using mean population vital rates as inputs, matrices are still a model of choice when asking the consequences, at the level of the population, of environmental stochasticity (Tuljapurkar, 1990a, 1986b; Tuljapurkar et al., 2003; Tuljapurkar, 1989) and individual stochasticity (Caswell and Sánchez Gassen, 2015; Engen et al., 2005a; Lande et al., 2003; Shpak, 2007; Shpak et al., 2013; Vindenes et al., 2008). Indeed, the field of evolutionary demography does not concern itself with the fate of particular individuals in a population, or with the effect of a specific segment of an environmental series. As its name indicates, it focuses on evolution, and therefore on evolutionary time windows and on the level on which evolution is at work : the population. However, it still needs to account for the long-term and population-wise effects of individual and environmental stochasticity. Specifically, their contracting effect on the population stochastic growth rate (taken as fitness), as demonstrated by Tuljapurkar (1990a) and Engen et al. (2005a), is of primordial importance to evolutionary demography. We shall exhibit, in chapter 3, how matrix models can yield such quantities.

Two irreconcilable models for two irreconcilable costs ?

It is clear from the inspection of these models, that the differences between *physiological* and *genetic costs of reproduction* in core mechanisms, evolvability, detectability, action time horizon are reflections of a deeper, ontological, difference in concepts and principle that seem hard to reduce and which is further echoed by the very different modeling approaches (Peck, 2004).

Whether two sides of the same coin, or orthogonal processes, *physiological* and *genetic costs of reproduction*are nonetheless, albeit theoretically, able to co-exist as demonstrated by the conjectural construction of *physiological costs of reproduction with genetic basis*. In that case, they certainly also interact with one another. Is one cost the cause, the consequence of the other one ? Do they have concurrent or opposite effects on phenotypical correlations, on they own mechanisms ? In order to advance towards the answers to such fundamental questions for costs of reproduction and senescence in particular and life history theory and trade-offs in general, we need to be able to build a model fit for evolutionary demography and thus genetic trade-offs, but with a narrower scrutiny level than a basic projection matrix, that would allow to get closer to the individual level and be able to implement physiological trade-offs between traits.

Simply put, we need to develop matrix models that are almost individual-based. This can be done via the addition of (potentially numerous) additional traits to basic age or stage-structured matrices in a framework we call multitrait matrices.

1.3.2 Towards an implementation of *physiological costs of reproduction with genetic basis* in a multitrait framework

Most matrix models indeed project populations where organisms are characterized by one (Lefkovitch, 1965; Leslie, 1945; Usher, 1969) sometimes two (Goodman, 1969; Rogers, 1966) but very rarely more

traits. The very recent development of methodologies to develop models with arbitrarily high number of states – called 'hyperstate matrices' by Roth and Caswell (2016) allows to create models with a great level of scrutiny, whilst still retaining all the evolutionary demography features of simpler matrices. In chapter 2, where we call such matrices Multitrait Population Projection Matrices (MPPMs), we develop an alternative construction method with computational complexity in mind. Most importantly we develop tools that enable to make sense of dynamic and evolutionary role of the traits (hence the name Trait Level Analysis) and therefore of the trade-offs connecting them (chapter 2).

In this section, we investigate how the general *physiological costs* depicted in section 1.2.1, possibly with a *genetic basis* described in section 1.2.2, are to be incorporated in an MPPM $\mathbf{M}$; or more precisely in a suite of MPPMs $\mathbf{M^e}$, where $\mathbf{e} \in \mathcal{E}$ the set of all possible environments for the studied population. In some cases, where environmental variance is not deemed central to a particular study, one may focus on the sole $\mathbf{M^{\bar{e}}}$, the model for the mean environment, simply noted $\mathbf{M}$. In order to do this, three families of traits will be incorporated into the model. First $\mathcal{B}$ the basic trait(s) that best determine the life history of the organism. Second $\mathcal{G}$ the genotypic traits that will allow to implement *hidden heterogeneity* and in particular the *genetic costs of reproduction*. And third, $\mathcal{D}$ the family of traits enabling to incorporate *dynamic heterogeneity*, and in particular the *physiological costs of reproduction*.

Hidden vs dynamic heterogeneity

By segregating components of categorization of an organism as being corresponding to either *hidden heterogeneity* or *dynamic heterogeneity* family of traits, we follow an important dividing line in life history theory. Individual heterogeneity, pervasive in most organisms and corresponding to the "variation observed in a trait among individuals within a given population" (Plard et al., 2012) is major determinant of population dynamics (Bjørnstad and Hansen, 1994). It is decomposed in two components.

First, *Hidden heterogeneity*, which accounts for "fixed at birth" *heterogeneity* - also called, when focusing on survival, "frailty" or "robustness" - is the expression of differences in individuals that are unobservable directly, and only inferred via the alleged effects on vital rates. *Hidden heterogeneity* thus corresponds, among other things, to differences in genotype(s) (epigenetics and early environmental effects are other determinants of *hidden heterogeneity*). In the context of costs of reproduction, *hidden heterogeneity* relates to *genetic costs of reproduction* (the variance in allocation genotype) and to the variance in acquisition genotype we call, for that reason, variance in "robustness".

Second, *dynamic heterogeneity* characterizes the differences arising between individuals (of the same genotype) as their life-trajectory unfolds. It is a product of individual stochasticity, another name for chance, the 'invisible hand' behind the differences in life history trajectory that can occur between two clones in the exact same given environments. In the context of costs of reproduction, it corresponds to the stochastic component of *physiological costs of reproduction*related to the granularity of reproductive effort.

Historically, this split has rarely been taken into account in empiricists' matrix models since, in the wild, it is particularly difficult to effects of acquired-at-birth differences between individuals and randomness of vital rates realization. Conversely, theoretical investigations of the role of each component of heterogeneity in evolutionary demographic models is a recent but thriving research field (see (Steiner et al., 2010; Tuljapurkar et al., 2009; Tuljapurkar and Steiner, 2010; Caswell, 2011, 2014) for instance). chapter 3 of this manuscript contains an example of such analysis of how these two components of heterogeneity combine to generate the observed diversity of life-trajectories. MPPM technology, and in particular Trait Level Analysis that we develop in chapter 2, allows to theoretically implement both sets of traits, and at the same time to generate the equivalent model where heterogeneity is undifferentiated and individuals are only characterized by age or another "basic" life-history-determining trait.

In the model for *physiological costs of reproduction with genetic basis*, the genotypic traits $\mathcal{G}$ corresponding to *hidden heterogeneity* and the *dynamic heterogeneity* traits $\mathcal{D}$ corresponding to *physiological costs of reproduction* thus make explicit the all-important heterogeneity in trajectories within the population. In order to do this, however, $\mathcal{G}$ and $\mathcal{D}$ families of traits rely first on common denominators of all organisms in the population. We call basic traits, $\mathcal{B}$ such characteristics that allow to define the general, central, life-history of the studied population.

Basic trait(s)

Because age is an inherent parameter to any projection model, the basic element of our model is thus an age-structured model, a Leslie matrix, corresponding to a specific genotype in a specific environment (Leslie, 1945). Other "basic" traits may be added to *age* to constitute the basic traits suite $\mathcal{B}$. In particular, for populations in which demographic characteristics are related to biological stages (such a seed, rosette, flowering plant, etc.) it seems largely preferable to use stage as a basic trait (Werner and Caswell, 1977,

see for instance); or rather to add *stage* to *age* in $\mathcal{B}$ as any Lefkovitch (i.e., stage-structured (Lefkovitch, 1965)) matrix can be demonstrated to actually be a *age*-and-*stage* MPPM (Lebreton, 2005). Other basic traits, that can be strong drivers of life-history are, among others, size (Usher, 1966), sex (Pollak, 1990) and location (Rogers, 1966).

Genetic costs and Hidden heterogeneity

Genetic costs - and more generally any *hidden heterogeneity* - will be implemented by adding one or several genotypic traits in the population characterization. The genotypic trait family $\mathcal{G}$ may, for instance, consist in combinations of the *acquisition* × *allocation* on the genotypic map; that is, each individual will be characterized by coordinates (i, j) on the relevant genotypic map (see figure 1.2) corresponding to alleles $acquisition_i$ and $allocation_j$. More generally, $\mathcal{G}$ may contain positions on the *robustness* gradient and on the slow-fast or any other life-history genotypic continuum. For a particular combination of genotypes g of $\mathcal{G}$, we denote $\mathbf{M_g^e}$ the relevant component of model $\mathbf{M^e}$.

The implementation of the effects of the various genotypes in the model - and thus of the genetic costs among others - will then consist in defining, for each $g \in \mathcal{G}$ and each environment $e \in \mathcal{E}$, all transitions of $\mathbf{M_g^e}$, that is all fertility and survival rates defined on $\{\mathcal{B}, \mathcal{D}\}$ for that particular $g \times e$ combination. For instance, let us consider a simple model where $\mathcal{B} = \{age\}$, $\mathcal{D} = \varnothing$ and $\mathcal{G} = \{g\}$, with $age = 1, 2$ and g a genotypic trait that can be worth either $g_1 = slow$ or $g_2 = fast$. The genetic costs of reproduction, relative to the slow-fast gradient can be implemented by providing lower fertility rates (for instance $\mathrm{M}_{1,1}^{slow} = \mathrm{M}_{1,2}^{slow} < \mathrm{M}_{1,1}^{fast} = \mathrm{M}_{1,2}^{fast}$) and higher survival rate ($\mathrm{M}_{2,1}^{slow} > \mathrm{M}_{2,1}^{fast}$) whilst still remaining iso-fitness ($eigs_{max}(\mathbf{M_{slow}}) = eigs_{max}(\mathbf{M_{fast}})$).

In a matrix model with genotypic traits, offspring cannot be expected to have exactly, and in all cases, the same genotype that its parent. Otherwise that would imply that the various $\mathbf{M_g^e}$ are square matrices within $\mathbf{M^e}$, therefore modeling totally hermetic populations. In order for the general model to make any sense, offspring class must be able to differ from parental class. In that case, survival components of $\mathbf{M_g^e}$ would be contained in g (g is "fixed-at-birth"), however the fertility components will be connected to other genotypes of $\mathcal{G}$, whilst still retaining the property that $\sum_{g \in \mathcal{G}} \mathbf{M_g^e} = \mathbf{M^e}$.

Therefore, in a population projection matrix framework, characterizing a population with genotypic traits, or in general with *hidden heterogeneity* traits that are (only partially) heritable raises questions . First, with regards to the interpretation of the ergodic state of a matrix in which different genotypes can cohabit. Second, with respect to the relevance of extracting selection gradients from models incorporating heredity.

Population-genetics/population-dynamics equilibria consistency One of the main feature of all matrix models, whether one-trait or multitrait, is the asymptotic stable state towards which it leads almost all initial population distributions (see the asymptotic analysis of multitrait models in sec.2.6.1, p.55) for a discussion on the dynamics consequences of the general reducibility of multitrait matrices). Once that state is reached, the proportions of individuals in each category remain forever constant. This may seem antagonistic with the fact that, as mentioned above, amongst the various genotype sub-models $\mathbf{M_g^e}$, some may be fitter then others, and thus expected to invade the population. This apparent dilemma is resolved by relating the population dynamics stable-state to population-genetics equilibria and show their equivalence.

In the trivial case where $\mathcal{B} = \{sex\}$ and $sex = \{m, f\}$ and vital rates are the same for all $g \in \mathcal{G}$ (no selection), then diploidy itself leads to the Hardy-Weinberg equilibrium. Adding $age \in \mathcal{B}$ in such a model leads to an age-structured hyomozygous/heterozygous genotypes equilibrium by linearization of the two-sex model (Caswell, 2008).

In the general linear case where $sex \notin \mathcal{B}$, the offspring genotype can only differ from its mother's genotype if mutations between the various genotypes in $\mathcal{G}$ are enabled by the model. Selection will promote fitter genotypes, but mutation may assign frailer genotypes to the offspring of the most robust individuals, leading to mutation/selection balance.

Following in the footsteps of Charlesworth (1970, 1980, 2000) we shall try and understand the relationship between the stable state theory of population dynamics and the mutation/selection equilibrium of population genetics. To do this, let us consider a simple example, where $\mathcal{B} = \{age\}$ and forsaking $\mathcal{D}$ for the time being. Therefore, each genotype, in a given environment, can be represented by its own Leslie matrix of expected vital rates, differing in realization between clones only by chance. The presence of several genotypes $g \in \mathcal{G}$ in the population implies - modelwise - the "cohabitation" of their related Leslie matrices on the Frobenius form of a multitrait matrix (see chapter 2, section 2.6.1, p.55). In this simple model, modeling an asexual haploid population in an heterogeneous context, such genotype matrices would be

interconnected by the mutation genetic process making it possible for an offspring of a particular genotype to belong to another genotype than its mother. Let μ be the generation mutation rate that determines the part μ of the offspring (of any individual of any genotype) that will have mutated and the part $1-\mu$ that will be of the same genotype than its parent. For simplicity, and because the categories of a matrix model are fixed, we make the oversimplification that mutated genomes fall evenly into the different existing genotypes.

From there, we can now establish, in the particular framework of multitrait models with *hidden heterogeneity* traits, the relationship between the population genetics concepts of "Wrightian fitness weights" and gene frequencies and the evolutionary demography concepts of ergodic growth rates and abundances. In the particular case where $\mu = 0$, the Leslie matrices are not interconnected,and the genotype with the highest fitness (ergodic growth rate) will invade (if the environment remains constant). For all other possible values ($0 \leq \mu < 1$), all implemented genotypes will coexist at the stable state (with equilibrium frequencies deducible from ergodic abundances). Let us consider 2 age classes and 2 genotypes, A and B. Let f_A^* and f_B^* be the ergodic abundances of the genotypes relative to the offspring ($age = 1$) state. Let also $\bar{w}$, w_A and w_B be the sums of the characteristic equations of respectively the population modeled by $\mathbf{M}$ and the sub-models $\mathbf{M}_A$ and $\mathbf{M}_B$ within $\mathbf{M}$. These quantities correspond to relative growth rates with respect to the population overall growth rate (see appendix 1.5.1). Asymptotic analysis of the dynamics of such a population leads to the following system :

$$\begin{cases} w_A.f_A^* + w_B.f_B^* = \bar{w} & (1.7) \\ \frac{w_A}{\bar{w}}.(1-\mu).f_A^* + \frac{w_B}{\bar{w}}.\mu.f_B^* = f_A^* & (1.8) \end{cases}$$

Interpreting w_A and w_B as the 'Wrightean' relative fitness weights - see discussion by Charlesworth (2000) - with f_A^* and f_B^* the genotypes frequencies, then the equations in this system (eq. 1.7 and 1.8), derived from stable state population dynamics asymptotic analysis, *are* the population genetics equations for the rate of change of gene frequencies when mutation is taken into account (Kimura, 1958; Crow and Kimura, 1970) at frequency equilibrium (i.e. selection/mutation balance). Equation 1.7 equates unsurprisingly $\bar{w}$ with the mean relative fitness of the population. Equation 1.8 incorporates generation mutation rate μ. It equates the next generation frequency of genotype A due to selection (embedded in w_A and w_B) and mutation (some B individuals, $\mu.f_B^*$, generate A offspring, whilst some A individuals, $(1-\mu).f_A^*$, do not mutate and also generate A offspring) with f_A^* the current frequency of genotype A (as expected since we are at stable-state/frequency equilibrium). This reasoning and these equations can readily be extended to any number of time-steps and any number of genotypes.

Thus, we have just demonstrated that it is possible to incorporate several genotypes or, more generally, partially hereditary *hidden heterogeneity* traits in a multitrait matrix model (the notions of mutation/heritability are related, as the mutation rate of a genotype can be interpreted as the probability that an offspring does not inherit the genotype from its parent). This may be unexpected as all non-negative matrices project populations toward a stable state (Caswell, 1989), whereas one would expect the genotype with highest growth rate to invade the population. And indeed if the trait is fully transmitted to offspring ($\mu = 0$), matrix $\mathbf{M}$ is just a block-diagonal matrix and only the highest-yielding of those blocks will have non-zero asymptotic abundances. On the contrary, if the trait is not heritable at all, with status of offspring drawn at random for the k genotypes (i.e. in the mutation framework of this section, $\mu = \frac{k-1}{k}$) then the study of the multigenotype model will also be pretty simple, with the various $\mathbf{M_g^e}$ only affecting within-generation genotype frequencies. However, for any other value of the heritability/mutation parameter, it will have strong effects on the dynamics and stable-state of the population; effects that can be measured with the tools developed in Chapter 2.

Growth rate sensitivities and selection gradients in models embedding *hidden heterogeneity* Matrix models are models of choice for evolutionary demography as they allow, among other tools, to generate *selection gradients* quantifying the force of selection on a particular life-history trait embedded in the model, as discussed in section 1.3.1. As we have just seen, implementing *hidden heterogeneity* as a family of traits, implies to input the fertility transitions between adults of a certain genotype and offspring of another and therefore to make assumptions about heredity/mutation.

In quantitative genetics models from which selection gradients stem, however, heredity and force of selection are components of the two different components which product yields the response to selection. In heritability and selection differentials in the Breeder's equation (Lush, 1937). In $\mathbf{G}$ (the additive genetic variance/covariance matrix) and the selection gradients in Lande's equation (Lande, 1982). In consequence, in a multigenotypic matrix model, the equation between sensitivity of population growth rate to vital rates of any genotype and selection gradient has to be treated with care. The inadequacy of such a model with quantitative genetics is not surprising since the latter is about quantitative traits described with by their variances, whilst the former incorporates, from the outset, all possible (discretized) genetic variant. The selection gradients, in such a model, then have to be calculated genotype by genotype, i.e. if $\mathcal{B} = \{age\}$, Leslie matrix by Leslie matrix, and not as the sensitivity of the MPPM growth rate to matrix entries.

These sensitivities then can be interpreted as selection gradients of the specific genotype's vital rates. The Lande's equation would then provide the expected change in these vital rates in a range allowed by $\mathbf{G}$ and paced by the selection gradient. If need be, new genotypes may then have to be added to $\mathcal{G}$. In multitrait models incorporating $\mathcal{G}$ as a family of traits, because of the possible confusion between the implemented genetic variance, and the $\mathbf{G}$ matrix, we think it preferable to not refer to sensitivities as selection gradient but simply as growth rate sensitivities of matrix entries, as we shall do in chapter 3.

Physiological costs and Dynamic heterogeneity

Capitals In order to implement the *physiological costs of reproduction*, our model $\mathbf{M}$ needs to account for the levels of the capital(s) of resources. In the framework defined in section 1.2.1, these are the Ratchet Capital and the Fluctuating Capital. This is done by adding these levels as traits in $\mathcal{D}$ the family of all *dynamic heterogeneity* traits of the model. The RC would represented by trait $rc \in \mathcal{D}$ which value at birth $RC(0)$ is itself determined by the $g \times e$ specific combination of $(\mathcal{G}, \mathcal{E})$, and thus a component of $\mathbf{M}^{\mathbf{e}}_{\mathbf{g}}$. In the particular case where $stor = 0$, the FC needs not be added as a trait. Indeed it starts at 0 and would then be reset at 0 at the beginning of each time-step. The effect of the environment $Env(t)$, occurring mid-period as individuals acquire FC resources before producing any reproductive effort, would already be implemented in $\mathbf{M}^{\mathbf{e}}_{\mathbf{g}}$ as it is determined by $g \times e$, in particular through the effect of allele acq_i in particular environment e. In general though and if one wishes to incorporate both capitals in the model, FC will be added as a trait $fc \in \mathcal{D}$.

Allocation process From environment $e \in \mathcal{E}$ and individual state $i = (b, g, fc \times rc) \in (\mathcal{B}, \mathcal{G}, \mathcal{D})$ we can compute the reproductive effort at that time t, $aRE(t)$, that the organism is permitted to produce in that particular $i \times e$ combination. And from $aRE(t)$, we can generate the distribution of $RE(t)$, the realized reproductive effort (a random variable, distributed between 0 and $aRE(t)$ in steps of size bre, see section 1.2.1).

survival transitions As discussed in section 1.2.1 both capitals needs to be provided for in order for the individual to survive, and thus survival transitions will be direct functions of state i. Indirectly survival will thus be a function of both very recent and older reproductive efforts. The actual reproductive effort produced $RE(t)$ will, in turn, determine the following state i' towards which the individual will survive as both fc and rc will change according to $RE(t)$ (g remaining by definition invariant). In simple projection matrix, implemented transition rates only provide the expectation of the random variable that is the vital rate. By extension this is also the case for most transitions in a multitrait matrix. The large distribution of traits allowed by multitrait matrices (see chapter 2), however, make it possible to detail the entire distribution of such a random variable by multiplying the output states of a transition. Nevertheless, even in that case, the model will not "know" that these several transitions from the same state are mutually exclusive and not independent. This implies that in all cases, whenever a random variable - as simple as a fertility rate, or as complex as $RE(t)$ - is used to build a multitrait matrix , its characteristics need to be precisely defined in order to be able to generate demographic moments which can not be directly produced by ergodic analysis of the MPPM. Chiefly among such quantities we find the demographic and environmental variances which account for the variance of fitness itself. Methods to do so are provided in chapter 3.

fertility transitions The fertility transitions, for their part, depend directly on the realized reproductive effort and thus of course, indirectly, on the levels of the capitals. In a simple model where reproductive efforts consist solely in giving birth, the distribution of fertility realizations by age $f(a)$ is equated to that of the realized reproductive effort $re(a)$ (see section 1.1).However, when res, the reproductive effort schedule, is not such a simple spike at time of birth (i.e. when $res \neq \delta_0$ where δ is the Dirac delta function), one needs to to implement the way res convolves with f to produce re (see sections 1.2.1 and 1.2.1) in order to implement fertility transitions in the MPPM from re. This can be done by splitting res into *reproductive efforts buckets* and adding corresponding traits $re_1, re_2, \ldots, re_m$ (where m is the number of periods of efforts necessary to produce an independent offspring) to $\mathcal{D}$. Then, provided the allocation function of $re(t)$ towards the various buckets $re = re_1 + re_2 + \cdots + re_m$ for a given state i, in a given environment e, one can track over time the accumulated efforts in each component. At each time-step, before a new reproductive effort is made, the level of re_1 is transferred into re_2 and resets at 0, the level of re_2 is transferred into re_3, etc. , and the level of re_m becomes, taking bre into account, the "fertility" rate $f(i)$. Or rather, it becomes the expected number of independent offspring produced by a parent in state i.

Such a mechanism allows to refine the evolutionary outputs of the model by relating the population's demography with reproductive efforts instead of fertility rates directly. When basing an evolutionary model solely on fertility rates (in a Leslie matrix for instance), one underestimates the efforts sustained before the time-step in which the birth occurs and forsakes entirely the importance of all efforts produced post-birth.

Such a common oversimplification turns perfectly understandable and modelable phenomena, like post-menopause longevity in humans and killer whales, into apparent life-history enigmas. This mechanism plays the same role as the tracking of age and aliveness of mother allowing Pavard and Branger (2012) to implement the effect of maternal care on juvenile survival. Tracking the status of the mother allows them to incorporate the effect of post-birth reproductive efforts, an inter-generational transfer called maternal care, on the beneficiaries of the cost of reproduction. The reproductive effort components method, described above, models the equivalent and symmetrical effect on the payer of the costs of reproduction. In both models, the cohorts of menopaused women with dependent children will have a positive survival selection gradient.

And thus, we have demonstrated the implementability of all aspects of the *physiological* (the two capitals, the stochastic allocative process, the complexity of the reproductive effort schedule per offspring) and *genetic* (their effects on rates, their heritability) *costs of reproduction* of section 1.2. However, a fuller comprehension of the mutitrait matrix framework is now necessary (chapter 2), in order to generate an actual matrix model for *physiological costs of reproduction with genetic basis* (chapter 3).

1.4 Discussion

Costs of reproduction have been investigated within three main research fields, physiology, ecology and genetics, with divergent concepts, vocabulary, methods and approaches. In this Chapter we have attempted to clarify and harmonize these different visions. This mainly lead us to reconcile terms and concepts (e.g., reproductive value and evolved reproductive effort schedule) and bring others nearer together, like the quantity-quality continuum and the stochasticity of reproductive effort, or the genetic variance along the slow-fast continuum and the costs themselves. Such reconciliations sometimes required extensions and porosity of some concepts, that may be frowned upon, but epitomize, we think, the pertinence of convoking various theories stemming from different fields to understand the costs of reproduction in particular, and trade-offs in general. Among others, we have freely extended the slow-fast concepts to any organism (but others have done so before (see for instance Nilsen et al., 2009)), and have, in general, applied the vocabulary of interspecific life-history continua to describe within-population genetic variance. They are, however, two categories that we have been enable to merge, despite their common name stemming from the common (mis)understanding that they *are*, indeed, the same thing : physiological and genetic costs of reproduction. They are, we show, different mechanisms that can act, jointly and simultaneously, at the level of any population. From this irreconcilable difference, we draw our entire theoretical model for costs of reproduction, decomposing and analyzing for each cost, in turn, the mechanism, its detectability and its relevant modeling.

1.4.1 Summary of mechanisms

Starting from Williams (1966)'s allocative definition of the costs of reproduction, we first try and develop a parsimonious theoretical model for *physiological costs of reproduction*, an allocation process at work at the individual level and at each time-step. The combination, in Williams (1966)'s definition, of different temporalities (the organism allocates resources that correspond to a capital, the reproductive value, which it has yet to acquire) prompts us to use, similarly, a combination of two resource capitals to account for an individual status : the Ratchet Capital and the Fluctuating Capital.

The former is related to Fisher (1930)'s reproductive value, and its initial level, for an organism, is the accumulation, from life expectancy of the organism until birth, of its evolved lifetime reproductive effort schedule. As it is built backwards, the corresponding resources cannot be acquired (time for instance), and the RC is maximum at birth. The FC develops forward as the organism acquires FC resources (say food) from the environment, over chronological time. By construction, both capitals determine the "state" of the organism. As such they are also the main determinants of the allocation process occurring at each time-step.

We show that the life-history strategy of an organism, and in particular its position on three life history continua are also important parameters of the allocation process. First, the storage capacity (of FC resources) – corresponding to the Income/Capital-Breeders spectrum – impacts the budgeting of the Fluctuating Capital. Second, the lifetime reproductive schedule – corresponding, among others (e.g., iteroparous/semelparous spectrum), to the position on the Slow-Fast Continuum – influencing the way the Ratchet Capital is managed. Third, the location on the Quantity-Quality spectrum drives the effect of demographic variance on reproductive effort, affecting, in turn, both capitals. From an individual's evolved life-history strategy, the time distribution per offspring of its reproductive effort (the reproductive effort schedule) and its personal state (levels of capitals, age, etc ...), this model allows us to predict the various characteristics of costs of reproduction with regards to the delay of occurrence, the component paying the cost (survival or fertility) and its detectability.

To *physiological costs of reproduction*, lying in the intermediate structure of Stearns (1989b)'s trade-off architecture, can one append a genetic basis (we call the combination *physiological costs of reproduction with genetic basis*). This genetic basis, located at Stearns (1989b)'s genotypic level, is a particular case of *genetic costs of reproduction*, a mechanism working at the level of the population that brings about negative correlations between fertility rates and later vital rates. These negative correlations can also be produced by *physiological costs of reproduction*, hence the common designation, but from a totally different mechanism. Moreover, we show that other *genetic costs of reproduction*, that do not require any allocative physiological mechanism to generate such correlations, are conceivable. In general, we show that *genetic costs of reproduction* are about variance in a gene that affects the position of the organism along the (intra-specific) Slow-Fast Continuum. For instance, variance in a gene affecting the allocation towards reproduction, in the case of genetic costs associated with a physiological allocative process.

We extend to the population and the inter-population levels, the analysis of detectability of costs both genetics and physiological. We analyze the effects of the main drivers of detectability : the absolute level of environmental conditions and the demographic and environmental variances. Demographic variance has contrasting effects at the level of the individual - where it fuels longitudinal detectability of *physiological costs of reproduction*– and the population – where it blurs the variance in allocation genotypes, i.e., the *genetic costs of reproduction*. Whilst poor absolute environmental levels improve detectability of *physiological costs of reproduction*, we show that environmental variance conceals both the costs at the individual level – as acquisition is an important component of the physiological process – and at the population level – where crossing reaction norms of the allocation gene would possibly hinder the emergence of negative correlations. Introducing *genetic costs of reproduction* into our model requires to incorporate genotypic variance. It thus makes sense to accompany this polymorphism in allocation, with polymorphism in acquisition genotype. This two-axis genetic variance is the main determinant of detectability of *genetic costs of reproduction*, each axis being itself dependent on environmental conditions. We show why negative correlations between early-life fertility rates and later vital rates are likelier to emerge from inter-population than intra-population studies.

1.4.2 Costs of reproduction and senescence theories

There is a striking parallel between the differences in mechanism and detectability between genetic and physiological costs of reproduction, and the differences between the two major senescence theories : the antagonistic pleiotropy and the disposable soma theories.

As already suggested by Orton (1929) for marine invertebrates and fishes, *physiological costs of reproduction* seem to lead to death via "accumulated senescence". Indeed, many recent studies (Lemaître et al., 2015; Boonekamp et al., 2014) bring to light the joint manifestation of both phenomena : cost of reproduction on fitness and actuarial senescence (taken as the increase in the force of mortality with age). That costs of reproduction constitute the keystone of senescence is also evidenced in the theory. First anecdotally as prominent theorists of the cost of reproduction happen to also be the main theorists of senescence; Bell (1980, 1984) and Williams (1957, 1966). Second and most importantly, because the mechanisms of both *genetic* and *physiological* mechanisms involved in the *costs of reproduction*, drawn from the general theories of trade-offs of Partridge et al. (1991), Roff (1992), Stearns (1989a) and Williams (1966) that we have tried to reconcile in our theoretical approach of section 1.2 (physiological costs in 1.2.1 and genetic costs in section 1.2.2), are also the building blocks of the two major evolutionary theories of senescence, the Antagonistic Pleiotropy Theory (APT) and the Disposable Soma Theory (DST).

Medawar (1952) drawing on earlier work by Fisher (1930) and Haldane (1941), devised the first major evolutionary theory of senescence based on individual selection : the mutation accumulation theory. As the force of selection, in age-structured populations, decreases with age, natural selection is much weaker and much slower to sweep deleterious mutations that only affect late-age fitness and are therefore likely to accumulate. From the same tenet, Williams (1957) constructed the APT, that describes how such late-age-acting deleterious alleles may invade via positive selection if the genes are pleiotropic with antagonistic effect, i.e. a positive effect on early-age fitness, a negative effect at late-age, and a net, overall, positive effect on fitness. Like the *genetic costs of reproduction* the APT mechanism lies at the gene (and thus population) level with effect blurred by individual and environmental stochasticity, and requires antagonistic gene pleiotropy. The main difference between APT and the genetic costs is that the latter describes a general variance life-history-driving genes in the population whilst the former predicts that this variance will be continuously increased at the "fast" end by new mutations and reduced at the "slow" end by selection.

The most recent evolutionary theory of senescence is the disposable soma theory developed by Kirkwood (Kirkwood and Holliday, 1979; Kirkwood, 1977). It explains senescence by the accumulation of damages at different physiological levels (for instance in proteins (Lindner and Demarez, 2009)) due to the divergence of some of the available energy from repair and maintenance towards other functions. DST has

the same mechanism as the *physiological costs of reproduction*, with a stronger focus on survival for the former and on reproduction for the latter. And like the physiological costs, DST is at work inside each individual trajectory and prone to chance (Finch and Kirkwood, 2000). Despite a large literature relating reproduction investments with senescence (see for instance Gustafsson and Pärt, 1990) and even with DST directly (Hammers et al., 2013), DST is not only about depletion of resources due to reproductive costs. It is an *evolutionary* theory of ageing. Indeed, it considers the way various organisms 'decide' to allocate to various functions at various ages as the product of evolution, mainly driven by extrinsic mortality : if an organism is consistently likely to die before a certain age, because of predators for instance, then it is expected that its allocation strategy will evolve to make sure enough energy is apportioned to reproduction before that age (Kirkwood and Rose, 1991). With the same evolutionary consideration in mind, we have constructed the Ratchet Capital, which implements the life-history strategy evolved by the species, and the genetic costs (i.e., the variance in allocation in the population) which incorporates the specific variant around this central strategy evolved by the specific lineage of the individual. In a nutshell, DST corresponds to *physiological costs of reproduction with genetic basis*.

In the same way, we have discussed, in this Chapter, the overlap and/or inclusion in one another, of genetic and physiological costs (see figure 1.3), many authors have tried to connect APT and DST. Some seem to think they may have found an overlap via the theoretical existence of genes acting on allocation (see Kirkwood and Rose, 1991; Partridge et al., 1991) but find the reconciliation to be forced, leading to "indirect reasoning" for the former and "a little perverse" for the latter. This intersection, in our models for costs, we call *physiological costs of reproduction with genetic basis*. To the contrary of such senescence theorists, we think it has fundamental importance. The only way to disentangle the roles of genetic and physiological trade-offs is actually, we think, to focus on the situations in which both trade-offs can jointly occur and to use the detectability patterns that emerge from the analysis of their mechanisms to paint the landscape of their effects and cross-effects.

More recently and generally however, authors have started to consider DST as a particular case of APT (Gavrilov and Gavrilova, 2002; Robins and Conneely, 2014; Rodríguez et al., 2017). This corresponds to our finding that *physiological costs of reproduction with genetic basis* are both physiological and genetic costs (the *allocation* gene playing the role of pleiotropic gene). Current research even seems to suggest that DST makes up a big part of APT. Is is notable, for instance, that several genes that have been identified in model organisms as affecting ageing rate are linked with the control of energy metabolism, e.g. via the insulin signaling pathway (Gems and Partridge, 2001). And therefore, the question whether this inclusion is a double inclusion is raised both in senescence studies and in this Chapter. We conclude, from the theoretical construction of *genetic* non-allocative *costs of reproduction* that DST is certainly only strictly included in APT (see figure 1.3). As long as one accepts the broad definition of genetic costs as the emergence of negative relationships between early-life and late-life vital rates, any feature that, for instance, attracts both mates and predators would generate *genetic costs of reproduction* providing that there is variance in genes promoting that feature.

There are many reasons why this equation of *physiological* (DST) and *genetic costs* (APT) is difficult to make. Whilst we identified situations where both costs cohabit, they differ so much in level of action (individual time-step vs individual trajectory), location in the architecture (intermediate structure vs genetic level), in individual stochasticity dependence (fueled by chance vs blurred by chance) and in levels of detectability that the comparison is hard to make.

Most fundamentally however, we think the difference between, physiological and genetic costs lie in the time window of their plasticity. Physiological costs compound over time (via their effects on the resource capitals) and, doing so, are plastic enough to buffer the environmental conditions over an horizon that can extend as far as the individual's lifetime. Genetic costs do not act at the individual level but at the level of the population. They will promote the life-history strategies best adapted to the recent evolutionary past. The variance in allocation strategies constitutes in itself the genetic plasticity available to the population as it faces uncertain environmental conditions over the near-evolutionary future. In a nutshell, these considerations characterize these two types of trade-offs as tools against environmental variance with different time windows of effects. Short-term (fraction of life expectancy) spikes are covered by physiological costs. Long term (evolutionary time) shifts by the genetic costs. In quantitative genetics terms, genetic costs (AP) draw the contours of the additive genetic variance-covariance matrix, $\mathbf{G}$, which combines with the selection gradient measuring the force of selection on multiple traits, in order to predict the evolutionary changes in both early-age fertility and late-age survival/fertility (Lande, 1982). For their part *physiological costs of reproduction*(DS) being mainly physiological processes have little effect on $\mathbf{G}$, but can have evolutionary implications via an effect on selection gradients that has to be investigated.

1.4.3 Modeling the costs of reproduction

Equipped with our design for both physiological and genetic costs, we set out to find the adequate model to implement *physiological costs of reproduction with genetic basis*. We show that each component of the costs corresponds to a typical and widely used projection model. *Physiological costs of reproduction* would be best projected over time by Individual Based Models, which, as their name indicates, work at the individual level. *Genetic costs of reproduction*, for their part, do not need such a level of scrutiny. However, as they are evolutionary models, they require specific tools to perform evolutionary analyses. Such tools are easily provided by population projection matrices, the model of choice for evolutionary demography. Combining both approaches, multitrait population projection matrices (developed and analyzed in chapter 2) allow to incorporate both individual/physiological and genetic/evolutionary aspects of *physiological costs of reproduction with genetic basis*.

Finally, we indicate directions for the implementation of the various components of our conception of the costs of reproduction. In particular, we show how to discriminate the various determinants of the costs into three different families of traits. One such family should contain the basic trait(s) best determining the organism life-history : age, potentially escorted by stage (or size, location, ...). The traits in the second and third compartments should be segregated according to the part of individual heterogeneity – either dynamic or hidden - they contribute to. Capital levels and reproductive efforts buckets track individual trajectories and correspond to the Dynamic Heterogeneity family of traits. The affiliations of an individual to particular genotypes are traits that, as they are fixed-at-birth, constitute the Hidden Heterogeneity family. Such genotypes may be the acquisition and allocation genotypes related to physiological costs in the *physiological costs of reproduction with genetic basis*framework. However they can also be unconnected to any allocative process, whilst still generating variance in robustness and in a gene with pleiotropic effects on different vital rates. We further show the implications of incorporating heritable traits (such as genotypes) in a matrix model in terms of stable-state vs. mutation/selection equilibria (we emulate Charlesworth's approach to show their correspondence via trait level analysis in appendix 1.5.1) and the interpretation of sensitivities as selection gradients.

In order to go any further, that is to actually generate a matrix incorporating all the components of the costs of reproduction (chapter 3), we first need to establish a building methodology and, most importantly, analysis tools for multitrait models (chapter 2).

1.4.4 Further developments

Looking back at the theoretical considerations on the topic of the costs of reproduction formulated in this chapter, it is clear that it is one amongst many ways to theorize the phenomenon. We draw our model from Williams (1966)' definition, but a different approach, say stemming from a physiologist per se, would lead to a different theory. Life historians, physiologists and ecologists may disagree with some aspects of the model, some components can be found to be overemphasized whilst others certainly would require more in-depth analysis.

This is in particular the case for the way organisms spread the reproductive efforts required to produce on (independent) offspring; a time distribution we call reproductive effort schedule or *res*. This repartition is certainly at least as important as the lifetime fertility schedule (akin to the slow-fast and iteroparous/semelparous continua) as a driver of reproductive efforts (the two distributions actually convolve to produce the lifetime reproductive effort schedule). In the next we mention a method that allows to account for the reproductive schedule in a matrix model. With that tool, it is possible to go beyond the simple fertility rates and to actually implement all aspects of reproductive efforts, including those, called parental care, produced long after the birth of offspring; this is of the upmost importance in order to correctly measure any selection gradient by age. Our model for costs of reproduction, even though wide enough to incorporate such parental care, does not include all inter-generational transfers. In social species, transfers of resources between individuals extend far beyond the care of a parent for its offspring. Grand-parents, aunts and other kin will also possibly provide help, as could brothers and sisters carrying out intra-generational transfers, as helpers-at-the-nest. We will in chapter 4 consider the literature and the current models for such kin selection, and provide the general idea on how to extend the concept of multitrait matrices implementing physiological trade-offs of this chapter towards kinship MPPMs incorporating several kin.

With regards to the prediction of detectability of the costs made in this chapter, these would need to be tested; ideally with a model organism for which genetic variance and environment are easy to control and measure. The same goes for our predictions of the effects of the life history continua (SFC, ICB, QQ) on the characteristics of the costs (delay, strength, component paying the costs, ...) that would benefit from a greater review of the relevant literature in order to be validated or overturned.

1.5 Appendices

1.5.1 Population genetics / population dynamics consistency

Let us consider a simplistic age & heterogeneity structured organism with 2 age classes and 2 genotypes, A and B, with respective Leslie matrices $\mathbf{M_A} = \begin{bmatrix} F_{1A} & F_{2A} \\ S_A & 0 \end{bmatrix}$ and $\mathbf{M_B} = \begin{bmatrix} F_{1B} & F_{2B} \\ S_B & 0 \end{bmatrix}$. Then the genotypes "internal" growth rates λ_A and λ_B are such that , from (Euler-Lotka),

$$\frac{F_{1A}}{\lambda_A} + \frac{F_{2,A}.S_A}{{\lambda_A}^2} = 1 \qquad \text{and} \qquad \frac{F_{1B}}{\lambda_B} + \frac{F_{2,B}.S_B}{{\lambda_B}^2} = 1$$

If the mutation rate per generation is μ then the entire population can be modeled by

$$\mathbf{M} = \begin{bmatrix} F_{1A}.(1-\mu) & F_{2A}.(1-\mu) & F_{1B}.\mu & F_{2B}.\mu \\ S_A & 0 & 0 & 0 \\ F_{1A}.\mu & F_{2A}.\mu & F_{1B}.(1-\mu) & F_{2B}.(1-\mu) \\ 0 & 0 & S_B & 0 \end{bmatrix}$$

Let λ be the population ergodic growth rate. Its associated right eigenvector $\boldsymbol{f} = \begin{pmatrix} f_{1A} & f_{2A} & f_{1B} & f_{2B} \end{pmatrix}'$, scaled to sum to 1, is the vector of ergodic relative abundances. Let $\bar{F}_1 = f_{1A}.F_{1A} + f_{1B}.F_{1B}$, $\bar{F}_2 = f_{2A}.F_{2A} + f_{2B}.F_{2B}$ and $\bar{S} = f_{1A}.S_A + f_{1B}.S_B$ be the mean ergodic-abundances-weighted vital rates (in chapter 2, we discuss such an operation, we call there Ergodic-Flow Preserving averaging or *folding*), and its properties). Then, by construction (see chapter 2 for generalization):

$$\frac{\bar{F}_1}{\lambda} + \frac{\bar{F}_2.\bar{S}}{\lambda^2} = 1 \tag{1.9}$$

Vector $\boldsymbol{f}$ is, by definition of an eigenvector, such that $\mathbf{M}.\boldsymbol{f} = \lambda.\boldsymbol{f}$ i.e.

$$\begin{cases} F_{1A}.(1-\mu).f_{1A} + F_{2A}.(1-\mu).f_{2A} + F_{1B}.\mu.f_{1B} + F_{2B}.\mu.f_{2B} = \lambda.f_{1A} & (1.10) \\ S_A.f_{1A} = \lambda.f_{2A} & (1.11) \\ F_{1A}.\mu.f_{1A} + F_{2A}.\mu.f_{2A} + F_{1B}.(1-\mu).f_{1B} + F_{2B}.(1-\mu).f_{2B} = \lambda.f_{1B} & (1.12) \\ S_B.f_{1B} = \lambda.f_{2B} & (1.13) \end{cases}$$

Incorporating equations 1.11 and 1.13 into equations 1.10 and 1.12 yields:

$$\begin{cases} F_{1A}.(1-\mu).f_{1A} + F_{2A}.(1-\mu).\frac{S_A.f_{1A}}{\lambda} + F_{1B}.\mu.f_{1B} + F_{2B}.\mu.\frac{S_B.f_{1B}}{\lambda} = \lambda.f_{1A} & (1.14) \\ F_{1A}.\mu.f_{1A} + F_{2A}.\mu.\frac{S_A.f_{1A}}{\lambda} + F_{1B}.(1-\mu).f_{1B} + F_{2B}.(1-\mu).\frac{S_B.f_{1B}}{\lambda} = \lambda.f_{1B} & (1.15) \end{cases}$$

Let $\bar{w} = \frac{\bar{F}_1}{\lambda} + \frac{\bar{F}_2.\bar{S}}{\lambda^2}$, $w_A = \frac{F_{1A}}{\lambda} + \frac{F_{2A}.S_A}{\lambda^2}$ and $w_B = \frac{F_{1B}}{\lambda} + \frac{F_{2B}.S_B}{\lambda^2}$, then the sum of equations 1.14 and 1.15 scaled by $1/\lambda$ gives :

$$w_A.f_{1A} + w_B f_{1B} = f_{1A} + f_{1B}$$

Letting $f_A^* = \frac{f_{1A}}{f_{1A}+f_{1B}}$ and $f_B^* = \frac{f_{1B}}{f_{1A}+f_{1B}}$ be the relative ergodic abundances of the two offspring states $(1, A)$ and $(1, B)$, the previous system of equations (1.14 and 1.15) can be rewritten, taking into account equation 1.9, as :

$$w_A.f_A^* + w_B.f_B^* = \bar{w}$$
$$\frac{w_A}{\bar{w}}.(1-\mu).f_A^* + \frac{w_B}{\bar{w}}.\mu.f_B^* = f_A^*$$

1.5.2 Effect of environmental variance on detectability of costs of reproduction

To illustrate the combination of effects of both the absolute level and the variance of the environment on the detectability of CB-FC costs, we construct the following simple model.

Let us consider, FC_i, the Fluctuating Capital for an organism at the beginning of period i, before feeding and then breeding and scaled by the reproductive effort it takes to produce 1 offspring. Putting aside individual stochasticity, let us make the assumption that, if after the acquisition period (feeding), the capital is higher than 1 (the level required to produce one offspring) then an offspring is produced. And therefore, $\forall i, 0 < FC_i < 1$.At each time-step, the resources acquired by the organism Env_i is drawn from an acquisition/environment random variable $\mathcal{E}$, uniformly distributed in $[\bar{Env} - v_{Env}, \bar{Env} + v_{Env}]$ where $1/2 < \bar{Env} < 1$ so that the expected amount of capital after acquisition $FC_i + Env_i$ is, in expectation, at least 1, the minimum amount to produce a child.

From these assumptions, the reproductive sucess" at time i, RE_i, simplified as the number of offspring produced (0 or 1) can therefore be written: $RE_i = \delta_{FC_i+Env_i-1}$ where δ is the Dirac function worth 1, in the positive domain, 0 in the negative. We set $FC_1 = 0$. From there, the sequence of capital values as a function of the environment is :

$$\forall i > 0 \quad FC_{i+1} = FC_i + Env_i - RE_i = FC_i + Env_i - \delta_{FC_i+Env_i-1}$$

In a simplistic model like this one, it is clear that realized reproductive effort has an effect on capital, and thus on fitness in general. However, as discussed in the chapter, such fitness/phenotypic costs may be reduced, or even canceled if the environment is very good for successive periods, and harder to detect if it has a large variance. To illustrate this, we compute, for this simple model, the correlation between successive reproductive efforts, according to different environmental distributions. These vary in expectation (absolute, or mean, environmental level) $\bar{Env}$ and in variance v_{Env} (the actual variance of uniformly distributed $\mathcal{E}$ is $\frac{v_{Env}^2}{3}$). As $\bar{Env}$ varies in its $[1/2, 1]$ range and v_{Env} between 0 and $\bar{Env}$, we plot - figure 1.5 - the succesive reproductive effort correlations $corr_i(RE_i, RE_{i+1})$. From this graph, we clearly see both detrimental effects of bad and varying environments:

When environmental variance is null, and all Env_i are at $1/2$, then depending on whether the capital is above or below $1/2$ itself, it will reproduce or not. Then, from equation above, we get $RE_i = \delta_{FC_i-1/2}$ (and therefore $RE_{i+1} = \delta_{FC_{i+1}-1/2}$), where $FC_{i+1} - 1/2 = FC_i - \delta FC_i - 1/2$. As $0 < FC_i < 1$, this implies $FC_i - 1/2$ will always be opposite sign than $FC_{i+1} - 1/2$ and thus $RC_{i+1} = 1 - RC_i$, therefore yielding a correlation of $corr_i(RE_i, RE_{i+1}) = -1$. As the environment worsen and varies, the correlation increases.

From this simple example, we see that "fitness costs", the emergence of the costs as correlation between fitness components, can be deceptive. Indeed, if the environment is too good, there is no fitness cost, as the environment immediately compensates for any reproductive effort. If it is too bad, then there is no fitness cost either, as the poor environment does not allow reproduction to occur. Fitness costs, in this example, are maximal when $Env_i = 1/2$, whereas the costs themselves at maximal at the other end of the range when $Env_i = 1$.

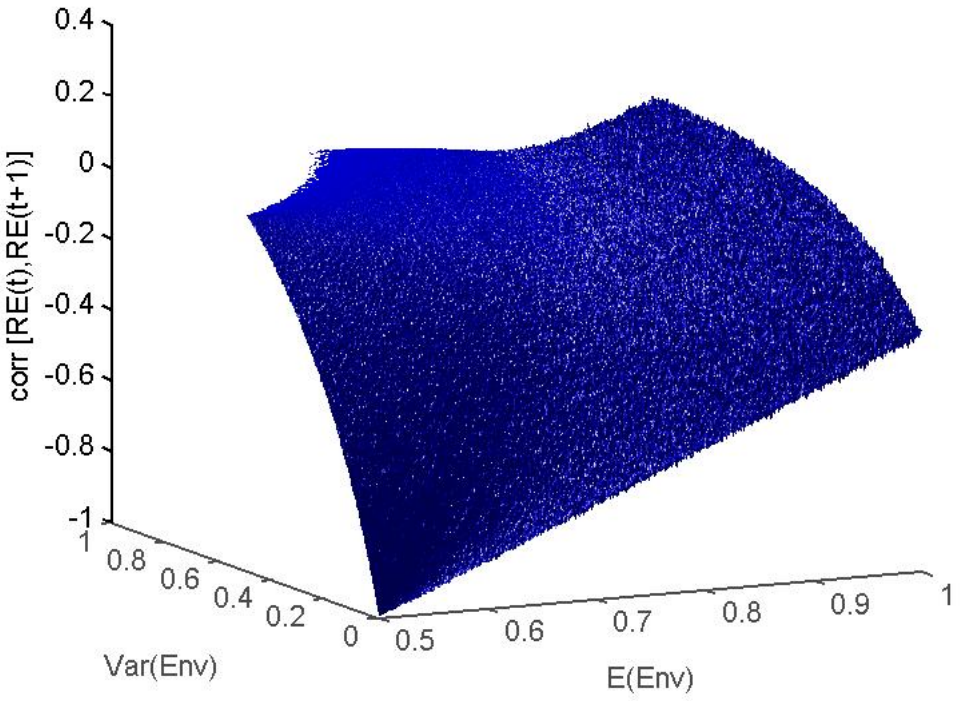


Figure 1.5: effects of absolute level and variance of environment on detectability of physiological costs

Chapter 2

Trait level analysis of multitrait population projection matrices

Contents

2.1 Presentation of the article

In this chapter, presented in its article form as published in Theoretical Population Biology, we develop a new method to generate a Multitrait Population Projection Matrix (MPPM) - a matrix projecting a population characterized by several traits - that is computationally efficient and we provide an analysis tool to measure the demographic effects of the multiple traits: the *Trait Level Analysis*.

First, in the introduction (section 2.2, p.43) we put the model in the general context of population projection matrices and review the early efforts to add a second trait to one-trait model. We discuss the topicality of MPPMs, with the advent of memory models (where extra traits are tracked over the life of individuals) and the general willingness to implement fixed-heterogeneity in evolutionary models.

In the method section (section 2.3, p.45), we first show how to build such an MPPM in three steps (section 2.3.1, p.45). First the vectorization of vital rates (fertility and survival) for all states. Second, the construction of the output vectors for each input state. Finally the combination of these vectors to generate the sparse definition of the MPPM. We then go on and show how to compute sensitivity analyses for such a model. This leads to a discussion regarding the primitivity of MPPMs and in particular with regards to the unicity of the maximum eigenvalue (in section 2.6.1, referred in the article as supplementary material 1).

Then, in section 2.3.2 (p.47), we describe the *Trait Level Analysis*Ẇe start by showing that when states of a graph are merged, the properties of the projection are altered (we demonstrate this in appendix 2.6.2, referred in the article as supplementary material 2). We choose to preserve ergodic flows. We show that it also preserves the asymptotic growth rate and ergodic abundances but at the cost of reproductive values (we demonstrate this in appendix 2.6.3, referred in the article as supplementary material 3). We extend this *EFP*-merging to entire traits, where we call it *folding*. We provide the equations allowing to implement *folding* in the article, as well as the code in annex 5.1.3 p.127 (this is referred to as supplementary material 5 in the article)

Finally, we illustrate this with a simple model (section 2.4 48) which parametrization is provided in annex 5.1.1 p.125 (this is referred to as supplementary material 4 in the article). We also illustrate a particular step of the methodology - the required nullification of near-zeroes in MPPM eigenvectors caused by the approximation inherent to convergence methods - in annex section 5.1.2 p.126 (this is referred to as supplementary material 4 in the article). The other supplementary materials of the article can be found via the link provided by the publisher (bottom of page 43). We conclude with a discussion (sec:2.5, p.51) about the further prospects for such a model, in terms of applications and in terms of theoretical improvements.

Theoretical Population Biology 116 (2017) 47–58

Contents lists available at ScienceDirect

Theoretical Population Biology

journal homepage: www.elsevier.com/locate/tpb

Trait level analysis of multitrait population projection matrices

Christophe F.D. Coste *, Frédéric Austerlitz, Samuel Pavard

UMR 7206 EcoAnthropologie et Ethnobiologie, MNHN, Université Paris Diderot, F-75016, Paris, France

HIGHLIGHTS

- A novel method for building multitrait population projection matrices is proposed.
- Asymptotic properties of multitrait matrices are explored.
- A new evolutionary demography tool, the trait level analysis, is proposed.
- Trait level analysis sheds light on effect of traits on multitrait dynamics.
- A parity–fertility–fecundity model reveals the potential of trait level analysis.

ARTICLE INFO

Article history:
Received 21 December 2016
Available online 27 July 2017

Keywords:
Evolutionary demography
Life-history theory
Matrix population model
MPPM
Multitrait model
Trait-level analysis

ABSTRACT

In most matrix population projection models, individuals are characterized according to, usually, one or two traits such as age, stage, size or location. A broad theory of multitrait population projection matrices (MPPMs) incorporating larger number of traits was long held back by time and space computational complexity issues. As a consequence, no study has yet focused on the influence of the structure of traits describing a life-cycle on population dynamics and life-history evolution.

We present here a novel vector-based MPPM building methodology that allows to computationally-efficiently model populations characterized by numerous traits with large distributions, and extend sensitivity analyses for these models. We then present a new method, the *trait level analysis* consisting in *folding* an MPPM on any of its traits to create a matrix with alternative trait structure (the number of traits and their characteristics) but similar asymptotic properties. Adding or removing one or several traits to/from the MPPM and analyzing the resulting changes in spectral properties, allows investigating the influence of the trait structure on the evolution of traits.

We illustrate this by modeling a 3-trait (age, parity and fecundity) population designed to investigate the implications of parity–fertility trade-offs in a context of fecundity heterogeneity in humans. The *trait level analysis*, comparing models of the same population differing in trait structures, demonstrates that fertility selection gradients differ between cases with or without parity–fertility trade-offs. Moreover it shows that age-specific fertility has seemingly very different evolutionary significance depending on whether heterogeneity is accounted for. This is because trade-offs can vary strongly in strength and even direction depending on the trait structure used to model the population.

1. Introduction

In early population projection models – those mathematical models used for the study of the dynamics and structure of populations projected over time – individuals were grouped according to one single trait (or i-state). This single trait was generally the age of the individuals (Euler, 1760; Lambert, 1772; Sharpe and Lotka, 1911). This was also the case for the original matrix models developed by Lewis (1942) and Leslie (1945). As ecologists started borrowing this powerful tool from classical demographers for species conservation and life-history evolution, one-trait models incorporating other traits than age, such as size or developmental stage were considered (Lefkovitch, 1965; Usher, 1969). Just as ecologists' interest in matrix population models prompted their development, evolutionary demographers' growing focus provided tools to understand the evolutionary processes at play. Demographic sensitivity analysis instruments (e.g., first and second level parameter sensitivities, life history graph and loop analysis) were early made available for one-trait models (see Caswell, 1978; de Kroon et al., 1986; Goodman, 1971).

* Correspondence to: UMR 7206, Laboratoire d'Eco-anthropologie et Ethnobiologie, Equipe "Anthropologie Evolutive", Musém National d'Histoire Naturelle, Site du Musée de l'Homme, 17 place du Trocadéro, F-75116, Paris, France.
E-mail address: christophe.coste@mnhn.fr (C.F.D. Coste).

http://dx.doi.org/10.1016/j.tpb.2017.07.002
0040-5809/

However, additional traits are often required in order to accurately study the dynamics of a population. In the literature so far, most multitrait population projection matrices (MPPM[1]), sometimes called metapopulation (Hanski, 1999), multidimensional (Van Imhoff, 1992), multistate (Rogers, 1980), multiregional (Rogers, 1966) or multisite (Lebreton, 1996) models, actually incorporate two traits. Rogers (1966) was the first to add a second dimension (location) to a one-trait (stage) matrix model and many important articles on stage-and-location modeling followed (e.g., Le Bras, 1970; Rogers, 1980, 1974). In a seminal paper, Goodman (1969) then introduced matrix models for both age-and-sex and age-and-parity structured populations. Such templates were later extended to model populations characterized by age and stage (e.g., Law, 1983). In parallel, perturbation and sensitivity analysis tools were extended from one-trait to multitrait models (Caswell, 2012; Willekens, 1977). Those instruments provide information on the impact on population dynamics of vital rates and other parameters. Other tools are however needed to investigate the behavior and properties of MPPMs. A one-trait model and a two-trait model of the same population indeed do not merely differ in level of scrutiny; they will also exhibit different population dynamics. The addition of a trait into a model therefore raises new questions, as for example, the extent to which this addition modifies the sensitivities of fitness to other traits. An analysis at the trait level is therefore required, and has yet to be developed.

Generalization to any number of trait was for a long time reined in by a lack of generalized building methodology – such models were mostly built, transition by transition, as concatenations of *ad-hoc* block matrices (e.g., Goodman, 1969; Le Bras, 1970; Lebreton, 2005; Rogers, 1966) – and by their space/time computational complexity (MPPMs increase in size and complexity with the number of traits). In 1969, Goodman hints at a three-trait model but does not build it (1969). And it would actually take another forty years before n-trait models (with $n > 2$) make their appearance. This emergence was due to efforts, first, from ecologists targeting a particular question (e.g., the mother hypothesis for Pavard and Branger, 2012; their made-to-measure model preventing oversize by only using biologically realistic combinations of traits as matrix entries). Second, from theorists: very recently, Roth and Caswell (2016) extended to any number of traits, the construction of MPPMs, which they denote as "hyperstate" matrices, via the vec-permutation approach previously developed for 2-trait models (Caswell and Salguero-Gómez, 2013; Caswell and Shyu, 2012; Hunter and Caswell, 2005). This approach formalizes the construction of an MPPM via the product of intermediary matrices, each representing the transitions between values taken by one of the traits when all others are fixed (thus decomposing an MPPM into a succession of independent processes).

Progress in the field of multitrait matrices is therefore at two levels. First, the growing focus on methodologies for building multitrait matrices has to be pursued. Computationally efficient methods are especially required to relax the compromise between number of traits and ability to build, analyze and perform perturbation analyses. Second, a theory of multitrait projection models is required to understand the impact of the traits themselves on population dynamics and life-history traits evolution. These developments are crucial for addressing emerging questions in evolutionary demography.

A recent developing field, for example, is that of *memory* models. Classical projection matrices – behaving like Markov chains – infer the entire future behavior of organisms from their current state. The fate of most natural organisms depends, however, on their whole life history trajectory (e.g., later life survival may be influenced by reproduction trajectories (Bell, 1980), or early life factors (Lemaître et al., 2015)). Adding traits is a solution to keep track of individual past events. This is the case, for instance, for models incorporating family structures where an individual's survival and reproduction depend on cooperation and/or competition relationships with its surviving kin. In such models, kin survival status and reproduction has to be recorded over time. For example, in order to understand the impact of maternal care on population dynamics, Pavard and Branger (2012) developed a one-sex projection model in which maternal and grand maternal survival status (along with age) impact juveniles survival rates. A woman's survival depends on her age and on the aliveness of her own mother, itself a function of the mother's age. This implies the use of three traits: age of individual, orphanhood, age of mother. Another example is the parity–fertility trade-off (also called cost of reproduction in ecology) whereby an individual fecundity or survival at a given age is compromised by its past reproductive effort (e.g., Boonekamp et al., 2014). As they develop, *memory* models will be increasingly demanding with regards to the number of traits.

In this context, individual heterogeneity, "*the variation observed in a trait among individuals*" (Plard et al., 2012) is more and more considered in population models (Vindenes, 2010). This heterogeneity can be split into dynamic observable heterogeneity and constant heterogeneity that is fixed-at-birth and cannot be observed directly, but can potentially be deduced from its impact on vital rates. The latter component was first called frailty in the context of survival models developed by Vaupel (1979) and collaborators. Models have been developed that implement both parts of heterogeneity (see the continuous time vitality-frailty model by Li and Anderson (2009)) and the dynamics of each component can be studied and its relative contribution to total heterogeneity analyzed (Caswell and Kluge, 2015; Tuljapurkar and Steiner, 2010). Multitrait models would allow for the incorporation of individual heterogeneity: accounting for observable dynamic heterogeneity component via the addition of (stage, spatial, social, etc.) traits and accounting for constant unobservable heterogeneity via the addition of fixed heterogeneity classes.

In this article, we first present an MPPM building methodology which is computational-efficiency-driven and alternative to the transition by transition building method and to the vec-permutation method of Roth and Caswell (2016). As in any MPPM, in our model, individuals are classified by multiple traits. There is no real limitation with respect to the nature of these traits: they can be categorical, discrete or discretized, observable (a measurable parameter) or unobservable (e.g., hidden heterogeneity). Those traits can be constant for an individual (inherited or acquired at birth) or varying throughout its life. In order to manage MPPMs increasing sizes and complexities with the number of traits, the matrix building methodology we develop here is vector-based and relies on sparse matrices (matrices in which most of the elements are zero). Because no loop is involved in the matrix building process – by contrast with the two alternative methods: the transition-by-transition and the vec-permutation approaches – the time computational cost associated with such an object is contained. The use of sparse matrices, for its part, drastically reduces space and thus time complexities. Through a sequential process, the method generates, in turns, (1) vital rates for each combination of traits, (2) output combinations of traits and corresponding distributions for each vital rate, and finally (3) all transitions between every pair of states. After a brief discussion of the existence and unicity of ergodic growth rates for MPPMs, we extend the computation methods of classical demographic measures, and most importantly, sensitivity analyses to our vector-based MPPM construction methodology.

We then develop a new type of evolutionary demography analysis, the *trait level analysis*, allowing the evaluation of the impact

[1] MPPM = multitrait population projection matrix.

of each trait on the whole population structure. We believe that this tool presents two main interests. First, for ecologists, to model one same population characterized by varying degrees of scrutiny (i.e., number of traits), and to assess the relative importance of traits on population dynamics. Second, for evolutionary demographers, to compare the strategies of two populations differing in their life-history but sharing the same dynamics (thus the same fitness as measured by growth rate). The *trait level analysis* is performed via a matrix operation we call *folding*. It applies to any MPPM and is independent from the way it was constructed. It consists in comparing an MPPM with an asymptotically equivalent *folded* version of it but where individuals are only characterized by a subset of the original traits. *Folding* process requires therefore the *merging* of states sharing the same values for some trait(s). We thus first analyze the various possible *merging* processes and describe the one we call Ergodic Flow Preserving *merging* (or EFP *merging*) that specifically preserves asymptotic properties (via the ergodic-abundances-weighted averaging of transitions). We then define the aforementioned *folding* process, as the extension of EFP *merging* to entire traits.

Finally, as an illustration, we extend Goodman's (1969) one-sex, age and parity model for humans, by adding a third trait categorizing women by fecundity. More precisely, this model incorporates parity–fertility trade-off (i.e., a *memory* model where past cumulative reproductive efforts compromise females' fertility rates) in a context of heterogeneity in fecundity. We then use this model in order to investigate the combined effects of heterogeneity and parity–fertility trade-off on trait-specific sensitivity of fitness. This example will illustrate the potential of the novel *trait level analysis* for investigating such interactions.

2. Model

2.1. Vector-based building of MPPM

2.1.1. Input: traits, states and transitions

The first step in building a multitrait population projection matrix $\mathbf{M}$ is to define traits for which dynamics over time will be projected. The second step is to relate all combinations of those traits to all entries of the matrix. Please relate to Table 1 for notations used throughout the Model and Illustration sections.

A population vector, even when representing multidimensional individuals, can only be projected via a 2-dimensional square matrix where entries are the transition rates between input states (columns) and output states (rows). Therefore, the multidimensional space of traits has to be vectorized: each of the q combinations of traits $\mathbf{t} = (t_1, t_2, \ldots, t_n)$, drawn from the "trait structure" vector $\mathbf{s} = (|t_1|, |t_2|, \ldots, |t_n|)$, where $|t_i|$ designates the number of different values trait t_i can take, will be given an index, or state: $\boldsymbol{state}_{\mathbf{s}}(t_1, t_2, \ldots, t_n)$ (see Appendix A).

Let us now implement the transition rates between every pair of states. Time is an inherent parameter to every population model as its elemental deed is to project population abundance over a defined time-interval. All transitions can then be categorized according to whether they stand for the persistence of individuals to the next time-step (i.e., through their survival) or to the production of new individuals (i.e., through fertility). This is true even when a transition incorporates changes in other traits that may be as crucial for life-history as survival and reproduction. For example, even in a size-structured population where age is not an explicit trait of the model, a transition rate does not infer whether an individual will grow from one stage to another stage, but rather whether it will both survive *and* grow, see Caswell (2012). It follows that survival and fertility rates for each state will form the building blocks of the construction methodology. Let us denote $\mathbf{vr}^{v}$ of size q the vector of vital rates (v stands for either survival or fertility) for each state i: an individual in state i will survive at a rate of $\mathrm{vr}_i^{survival}$ and reproduce at a rate of $\mathrm{vr}_i^{fertility}$.

Once $\mathbf{vr}^{fertility}$ and $\mathbf{vr}^{survival}$ have been computed, let us examine the way these vital rates are distributed over the various possible output states. We define $\mathbf{oi}^{v,i}$ the vector of indices of all biologically realistic potential output states from state i through vital process v. In other words, the combination of vectors $\mathbf{oi}^{fertility,i}$ and $\mathbf{oi}^{survival,i}$ is the set of all possible states j for which $M_{j,i}$ might not be zero. Let us also denote $\mathbf{op}^{v,i}$ the associated vector of probability distribution: through vital process v, an individual in state i may be projected in any of the states of $\mathbf{oi}^{v,i}$ and will be projected in specific state $\mathrm{oi}_j^{v,i}$ with probability $\mathrm{op}_j^{v,i}$. Being a probability distribution, $\forall v, i \quad \sum_j \mathrm{op}_j^{v,i} = 1$.

For example, let us imagine a 2-traits life-history structured by three age-classes and two size-classes. In this case, $n = 2$, $\mathbf{t} = (age, size)$, $t_1 \in [1, 3]$, $t_2 \in [1, 2]$, $\mathbf{s} = (3, 2)$ and $q = 6$ (Table 2). The state (i.e. the matrix entry number) corresponding to the pair of traits $(2, 2)$ is $\boldsymbol{state}_{\mathbf{s}}(2, 2) = 1 + (2 - 1) + (2 - 1) \times 3 = 5$. Conversely, the fifth entry of the matrix corresponds to trait couple $(2, 2)$ (given by the function $\boldsymbol{tuple}_{\mathbf{s}}(5)$, see Appendix A). Let us consider fertility and survival respectively (***fert*** and ***surv***) as functions of respectively *age* and (*age*, *size*). Over one time-step, through survival, individual age will increase by one. Assuming that this organism can only either grow or remain in the same size class with equal probabilities, an individual at $\boldsymbol{state}_{\mathbf{s}}(i) = (age, size)$ may survive as any of two possible combinations: $\mathbf{oi}^{survival,i} = (\boldsymbol{state}_{\mathbf{s}}(age + 1, size), \boldsymbol{state}_{\mathbf{s}}(age + 1, size + 1))$ and $\mathbf{op}^{survival,i} = \left(\frac{1}{2}, \frac{1}{2}\right)$.

2.1.2. Projection matrix

2.1.2.1. Projection matrix as a sparse matrix. Vectors $\mathbf{vr}^{v}$, $\mathbf{oi}^{v,i}$ and $\mathbf{op}^{v,i}$ contain all the information needed to build $\mathbf{M}$. For each state i, and each vital process v, let j be the m^{th} element of $\mathbf{oi}^{v,i}$, (i.e., $j = \mathrm{oi}_m^{v,i}$), then the transition rate from state i to state j, for vital process v, is $t_{i\to j}^{v} = \mathrm{vr}_i^{v}.\mathrm{op}_m^{v,i}$ (for a formal approach see Appendix B). For each state i, and each vital process v, we can then create a matrix of transitions $\mathbf{T}^{v,i}$, gathering all triplets of transitions $\left(i, j, t_{i\to j}^{v}\right)$ between state i and every state j reachable from i via v. Concatenating the $\mathbf{T}^{v,i}$ matrices for all states and for both vital rates gives us $\mathbf{T}$, the table of all transition triplets, representing the sparse formulation of matrix $\mathbf{M}$. They are equivalent, in the sense that both contain the same information, either as a $q \times q$ matrix containing every transition from any of the q states to any other, or as a table of all (bio)logically possible triplets of transitions $\left(i, j, t_{i\to j}^{v}\right)$ (see Appendix B for a formal definition of $\mathbf{T}$ and the $\mathbf{T}^{v,i}$). It must be stressed that several triplets in $\mathbf{T}$ may point to the same entry $M_{i,j}$: this is the case, for instance, in size-based models where small individuals can produce individuals in the same category via both fertility and survival. This possibility is taken into account when computing $\mathbf{T}$ as a sparse definition of $\mathbf{M}$. The transition rates (3rd element of the triplet) of all triplets in $\mathbf{T}$, sharing the same 1st and 2nd elements, will be summed.

2.1.2.2. Implementability and comparison with alternative building methods. This vector-based methodology allows the construction of any MPPM, with implementability, and in particular computational time and space complexity, in mind. We now compare this methodology with two alternative construction methods: (1) the transition by transition building of $\mathbf{M}$ looping through all q^2 matrix entries, and (2) the vec-permutation method and its sequenced vital processes.

The transition by transition building of $\mathbf{M}$ incurs a time and space complexity $O(q^2)$. (simply put this means the running time of the construction of $\mathbf{M}$ and its storage size are of the order of q^2) Computation efficiency of the basic method therefore dramatically

Table 1
Notations. In this article, we denote vectors in bold (e.g. **a**), matrices in bold capital (e.g. **A**), multidimensional matrices in gothic (e.g. $\mathbb{A}$) and functions in bold italic (e.g. ***a***). The transpose of a matrix is denoted with a prime (e.g. $\mathbf{A}'$) and the number of elements of a vector or matrix, or the number of values a determinist variable can take, by || (e.g. $|\mathbf{a}|$). Recurring names and denotations in the article are as follows:

Object or operation	Notation	Object or operation	Notation
Generic MPPM	$\mathbf{M}$	Vital process : fertility or survival	v
Number of traits	n	Vital rate vector	$\mathbf{vr}^v$
Combination/n-tuple of traits and equivalent state number	$\mathbf{t} = (t_1, t_2, \ldots, t_n)$ $\boldsymbol{state}_{\mathbf{s}}\ (t_1, t_2, \ldots, t_n)$	Set of output states from state i through vital process v	$\mathbf{oi}^{v,i}$
Trait structure = vector of traits sizes	$\mathbf{s} = (\lvert t_1\rvert, \lvert t_2\rvert, \ldots, \lvert t_n\rvert)$	Probability distributions for output states from state i through vital process v	$\mathbf{op}^{v,i}$
Total number of states	$q = \prod_{i=1}^{n} \lvert t_i\rvert$	Vector of parameters (size k)	$\mathbf{p} = (p_1, \ldots, p_k)$
Largest eigenvalue of $\mathbf{M}$	λ	Multidimensional matrix of sensitivities of $\mathbf{M}$ to $\mathbf{p}$	$\mathbb{S}$
Associated right eigenvector (scaled to sum to 1)	$\mathbf{w}$	(x, y, z) is a triplet of transitions of $\mathbf{M}$	$t_{x\to y} = \mathrm{z} = M_{y,x}$
Associated left eigenvector (scaled so that $\mathbf{v}'\mathbf{w} = 1$)	$\mathbf{v}$	Transition triplets matrix for state i, and vital process v	$\mathbf{T}^{v,i}$
Matrix products:		Matrix of all transition triplets	$\mathbf{T}$
Hadamard (entrywise) product	$\odot$		
Hadamard (entrywise) division	$\oslash$		
Kronecker product	$\otimes$		
$\mathcal{G}_{\mathbf{M}}$ (Graph of $\mathbf{M}$) vs $\mathbf{M}$ equivalences:	$\mathcal{G}_{\mathbf{M}}$ $\mathbf{M}$	Row concatenation of $\mathbf{A}$ and $\mathbf{B}$	$\mathbf{A} \cap \mathbf{B}$
Node vs State	node i state i	Identity Matrix of size b	$\mathbf{I}_b$
Transition vs Edge	$t_{i\to j}$ $M_{j,i}$	Vector of ones of size b	$\mathbf{1}_b$
Vec operator stacking elements of a multidimensional in one column vector	$\boldsymbol{vec}\ ()$	Operator permuting dimensions of a multidimensional matrix according to permutation of traits *sigma*	$\boldsymbol{perm}_{sigma}$
Operator reshaping a vector into a multidimensional matrix according to $\mathbf{s}$	$\boldsymbol{vec}_{\mathbf{s}}^{-1}\ ()$	n-dimensional expression of $\mathbf{w}$ according to $\mathbf{s}$	$\mathbb{W} = \boldsymbol{vec}_{\mathbf{s}}^{-1}\ (\mathbf{w})$

Table 2
Traits and states for a 2-ages 3-sizes organism.

	t_1 = age	t_2 = size	Space of traits combinations	Space of states
Bounds	[1,3]	[1,2]	no ordering	[1,6]
Number of elements	3	2	$q = 3 \times 2 = 6$	$q = 6$

slows with the number and size of traits. Creating, in a vector-based manner, $\mathbf{T}$ – as opposed to $\mathbf{M}$ – involves a finite number of operations on each state and has thus a much lower complexity $O(\text{ntrans})$ where ntrans is the total number of transitions implemented (a majorant of the number of non-zero transitions): $\text{ntrans} = \sum_v \sum_{i=1}^{q} |\mathbf{oi}^{v,i}|$ (where $|\mathbf{oi}^{v,i}|$ is the number of elements of the vector of output states for input states i, and vital rate v). Indeed, most input states have only a few potential output states, and we have $\text{ntrans} < q.\max_{v,i}|\mathbf{oi}^{v,i}| \ll q^2$. For the same reason, induced matrix $\mathbf{M}$ is sparse with density ntrans/q^2. Matrix $\mathbf{T}$ is mathematically equivalent to $\mathbf{M}$ but computationally more efficient, with a storage size of 3. ntrans ($\ll q^2$). As a consequence it accelerates further analyses performed on the MPPM (see Appendix C). Moreover, time and space complexities in creating $\mathbf{T}$ are not increasing functions of the number of traits. On the contrary, the proportion of non-zero transitions will decrease with every addition of a trait.

The vec-permutation approach (Roth and Caswell, 2016) is an elegant MPPM building technique. It was created with sensitivity analysis in mind, and for that reason prerequires a decomposition of the population dynamics into sequential and independent processes. Matrices modeling interdependent traits (for instance complex *memory* models where transition value is the product of functional relationships between several traits) may be difficult (and sometimes impossible) to build that way. This can be demonstrated by considering that vec-permutation technique can implement only $q\,(s_1 + s_2 + \cdots + s_n)$ transitions out of the q^2 possible transitions of the MPPM. Computerwise, the associated algorithm is much faster than the transition by transition technique, but the use of multiple loops on full (i.e., not sparse) matrices makes its algorithmic complexity $O(q^2)$. Moreover, its implementability requires the provision of a potentially large number, $q.\,(1/s_1 + 1/s_2 + \cdots + 1/s_n)$, of input matrices.

To summarize, most 2-trait MPPMs are simple enough to be actually faster built transition-by-transition. The vec-permutation approach produces flexible results where all basic parameters are easily identifiable making it an ideal method for a small number of independent traits. We feel that our vector-based approach is better suited for models encompassing larger number of traits (see Supplementary Material 5 where both methods are implemented for the MPPM of the illustration).

2.1.3. Sensitivity analysis of vector-based MPPMs

To perform sensitivity analysis is "*asking what would happen to some dependent variable if one or more independent variables were to change*" (Caswell, 2001, page 206). In population dynamics, the main dependent variable of interest is the population ergodic growth rate λ, mostly used by conservation ecologists and evolutionary demographers when studying organisms with overlapping generations (see Giske et al., 1993; Murray, 1992; Nur, 1984 for discussions). In the case of MPPMs, the unicity of λ is not guaranteed by the Perron–Frobenius theorem as MPPMs are reducible unlike most one-trait matrices (Caswell, 2001, page 81). We however show in Supplementary Material 1 how to reduce the problem to sub-models with unique λ in case of multiple maximum eigenvalues.

It is possible to assert the sensitivity of the ergodic growth rate to several layers of variables. First to matrix entries, the direct drivers of λ; this is the first level analysis. Second to parameters impacting, directly or indirectly, these matrix entries; this is the higher level sensitivity analysis.

First level sensitivity analysis of λ with respect to any entry of $\mathbf{M}$ is not more complicated for an MPPM than for a more simple matrix. The sensitivity matrix $\mathbf{S}$ is given by (Caswell, 1978):

$$\mathbf{S} = \frac{\partial \lambda}{\partial \mathbf{M}} = \mathbf{v}\mathbf{w}'. \tag{1}$$

However, in MPPMs, matrix entries aggregate several biological functions (e.g., growth and survival, fertility and migration). Therefore higher level sensitivity analysis is increasingly required as the number of traits increases. The relevant biological function parameters, along with factors driving potential functional relationships between trait values and transitions, will be stored in vector $\mathbf{p} = (p_1, p_2, \ldots, p_k)$

Vital rates $\mathbf{vr}^v$, the states toward which they are distributed $\mathbf{oi}^{v,i}$, with probabilities $\mathbf{op}^{v,i}$, are then implemented as functions of $\mathbf{p}$. Consecutively, all transitions and matrices can also be written as functions of $\mathbf{p}$: $t^v_{i\to j}(\mathbf{p})$, $\mathbf{T}(\mathbf{p})$, $\mathbf{M}(\mathbf{p})$. Expressing $\mathbf{M}$ as a function of $\mathbf{p}$ requires to apply the same construction steps as for the numerical version of $\mathbf{M}$ above to *formal/symbolic* versions of vectors $\mathbf{vr}^v(\mathbf{p})$, $\mathbf{oi}^{v,i}(\mathbf{p})$ and $\mathbf{op}^{v,i}(\mathbf{p})$. Once $\mathbf{M}(\mathbf{p})$ is computed, we can generate the *parameter sensitivity matrix* $\mathbb{S}$, a $q \times q \times k$ multidimensional matrix, displaying the sensitivities of every non-zero element of $\mathbf{M}$ to every element of $\mathbf{p}$, i.e., $\mathbb{S}(i,j,l) = \frac{\partial M_{i,j}}{\partial p_l}$. From Eq. (1) we can then express the sensitivity of λ to any parameter p_l:

$$\frac{\partial \lambda}{\partial p_l} = \sum_{i,j} \frac{\partial \lambda}{\partial M_{i,j}} \cdot \frac{\partial M_{i,j}}{\partial p_l} = \sum_{i,j} \mathrm{S}_{i,j} . \mathbb{S}_{i,j,l}. \tag{2}$$

(in matrix notation, for implementability purposes, this can be written as $\frac{\partial \lambda}{\partial p_l} = \mathbf{1}' \cdot (\mathbf{S} \odot \mathbb{S}(\cdot,\cdot,l)) \cdot \mathbf{1}$ where 1 is a vector of 1s, $\odot$ is the elementwise Hadamard product)

2.2. Trait level analysis

To perform *trait level analysis* is asking what would happen to some dependent variable if the traits describing the population were to change. In particular, to what extent the set of traits $\mathbf{t} = (t_1, t_2, \ldots, t_n)$ chosen to generate MPPM $\mathbf{M}$, impacts the population dynamics (spectral properties, net reproductive rate, generation time, etc.) and sensitivity of λ. It does so by comparing properties of $\mathbf{M}$ to those of *folded* matrix $\mathbf{M}^{\mathbf{fold}}_{\mathbf{t}\setminus\mathbf{st}}$, a MPPM modeling a population with the same ergodic flows than $\mathbf{M}$ but where only a subset $\mathbf{st}$ of those traits is used ($\mathbf{st} \subset \mathbf{t}$); the ergodic flow from state i to state j in matrix $\mathbf{M}$ being $\mathrm{w}_i . M_{j,i}$ (i.e., the transition rate weighed by the ergodic abundance). In order to perform *trait level analysis*, we first need to study the existence, unicity and properties of such *folded* matrices; a problem which reduces itself to understanding the associated merging process for states of a graph.

2.2.1. Merging of states

Reducing the set of traits from $\mathbf{t}$ to $\mathbf{st}$ requires to merge into one new node all the nodes sharing the same set of values for traits subset $\mathbf{t}\setminus\mathbf{st}$ (the complement of $\mathbf{st}$ in $\mathbf{t}$). In a state transition directed graph, each node is defined by the set of transitions from and to all other nodes and toward itself. What should those be for each new node derived from a merger, in order to preserve population dynamics? The answer is that there is no absolute way of perfectly merging states in a state transition directed multigraph such as $\mathfrak{G}_{\mathbf{M}}$ (the graph which transition matrix is $\mathbf{M}$) as demonstrated in Supplementary Material 2. There is a choice to be made with regards to the properties we want invariant under the operation of merging.

This dilemma has been discussed at length by mathematicians and economists under the expression of "Aggregation Problem in Input–Output analysis" (Ara, 1959; Fisher, 1958; Leontief, 1986; Morimoto, 1970; Simon and Ando, 1961) and first extended to the field of multitrait population dynamics by Rogers (1969). In all those studies, the issue was solved in favor of short term dynamics, by preserving flows at the following time-steps, with little to no consideration for stable state. One way to solve this dilemma from an asymptotical analysis of population dynamics viewpoint is to preserve relative asymptotic (ergodic) flows and abundances (relative to the sum of all flows/abundances) while merging states. Simply put, when concerned with preserving flows in the (life-) cycle, the aggregation of states of a population dynamics matrix model will always lead to abundance-weighted averaging of transitions. Concerned with short term consistency, economists used current abundances to weight transitions. Evolutionary ecologists are however more concerned with the population long term behavior and would weight transitions with ergodic abundances. This merging process, we call Ergodic Flow Preserving (or EFP), was first described by Enright et al. (1995), then formalized for age-structured populations by Hooley (2000) and for stage-structured organisms by Salguero-Gómez and Plotkin (2010). The resulting principles of this EFP *merging* are summarized in Fig. 1.

2.2.2. Folding traits of MPPM

The EFP *merger* process described above is the building block of the trait *folding* process of $\mathbf{M}$. We will now describe how to formally generate $\mathbf{M}^{\mathbf{fold}}_{\mathbf{t}\setminus\mathbf{st}}$, which is MPPM $\mathbf{M}$ *folded* upon a subset of its traits $\mathbf{st}$ by EFP *merging* of all the states sharing the same values for the traits in $\mathbf{st}$. In what follows, we formalize this using matrix notation.

We need to define two operators that will facilitate this operation. First an inverse morphism of the $\boldsymbol{vec}()$ operator vectorizing a matrix into one column vector (Henderson and Searle, 1981), $\boldsymbol{vec}^{-1}_{\mathbf{s}}()$, defined by $\forall \mathbb{A} \in M_{\mathbf{s}}(\mathbb{R})$, $\boldsymbol{vec}^{-1}_{\mathbf{s}}(\boldsymbol{vec}(\mathbb{A})) = \mathbb{A}$. We can then rewrite $\mathbf{w}$ as its $\mathbf{s}$-dimensional matrix expression $\mathbb{W} = \boldsymbol{vec}^{-1}_{\mathbf{s}}(\mathbf{w})$. Second the operator $\boldsymbol{perm}_{\boldsymbol{\sigma}}$ permuting the dimensions of a multidimensional matrix as a function of the corresponding permutation of traits: $\forall \mathbb{A} \in M_{\mathbf{s}}(\mathbb{R})$, $\boldsymbol{perm}_{\boldsymbol{\sigma}}(\mathbb{A})_{i_1,i_2..i_n} = \mathbb{A}_{\boldsymbol{\sigma}(i_1,i_2..i_n)}$. We can now project $\mathbf{w}$ onto one or more traits by regrouping the traits to be collapsed via trait permutation $\boldsymbol{\sigma}$ and then summing $\boldsymbol{perm}_{\boldsymbol{\sigma}}(\mathbb{W})$ on the relevant subset of its dimensions. Every permutation of traits $\boldsymbol{\sigma}$ can be extended to its corresponding permutation of states $\boldsymbol{\sigma}^*$:

$$\boldsymbol{\sigma}^*(1,\ldots,q) = \boldsymbol{vec}\left(\boldsymbol{perm}_{\boldsymbol{\sigma}}\left(\boldsymbol{vec}^{-1}_{\mathbf{s}}(1,\ldots,q)\right)\right).$$

With these tools, it is relatively easy to navigate through traits, project onto one or several of them and change the way they are ranked, in both population vectors (ergodic abundance, reproductive value, etc.) and multitrait projection matrices. They will enable us to *fold* $\mathbf{M}$ over the first m of its n-sized vector of traits ($m < n$), that is, to obtain a new matrix, $\mathbf{M}^{\mathbf{fold}}_{\mathbf{t}\setminus\mathbf{st}}$, projecting the population now "only" categorized by the $(n-m)$ remaining traits of $\mathbf{t} \setminus \mathbf{st}$. In $\mathbf{M}^{\mathbf{fold}}_{\mathbf{t}\setminus\mathbf{st}}$, all states sharing identical values for the m traits to be *folded* upon have been EFP *merged*. If the traits to be *folded* upon are not in the proper positions, we first need to generate a permutation-similar version of $\mathbf{M}$ where those m traits occupy the first m positions in $\mathbf{t}$. Then we contract and sum over, in $\mathbf{w}$, all dimensions to be *folded* upon in order to generate $\mathbb{W}^{\mathbf{fold}}_{\mathbf{t}\setminus\mathbf{st}}$, a $(n-m)$-dimensional matrix representing the ergodic abundance vector over the $(n-m)$ remaining traits:

$$\mathbb{W}^{\mathbf{fold}}_{\mathbf{t}\setminus\mathbf{st}}{}_{i_2,i_3,\ldots,i_{n-m+1}} = \sum_{i_1} \left[\boldsymbol{vec}^{-1}_{\left(\prod_{i=1}^{m}|t_i|,|t_{m+1}|,\ \ldots,|t_n|\right)}(\mathbf{w})\right]_{i_1,i_2,\ldots,i_{n-m+1}}. \tag{3}$$

With $\mathbb{W}^{\mathbf{fold}}_{\mathbf{t}\setminus\mathbf{st}}$ we can generate the relative ergodic abundances weights that have to be allocated to $M_{i,j}$, before summing all elements sharing the same trait values for $(t_{k+1},\ldots,t_n)$ in order to obtain $\mathbf{M}^{\mathbf{fold}}_{\mathbf{t}\setminus\mathbf{st}}$ according to the EFP *merging* process described in Fig. 1. Those weights take the form of matrix $\mathbf{Wght} = \left[\cap^q \mathbf{w} \oslash \left[\cap \prod_{i=1}^{m}|t_i| \boldsymbol{vec}\left(\mathbb{W}^{\mathbf{fold}}_{\mathbf{t}\setminus\mathbf{st}}\right)'\right]'\right]$ where $\oslash$ is the Hadamard, i.e., entrywise division and $\cap$ the symbol for row concatenation. $\mathbf{Wght}$ has the same size than $\mathbf{M}$. Let us now denote $\mathbf{P}^{\boldsymbol{BF}}$ the

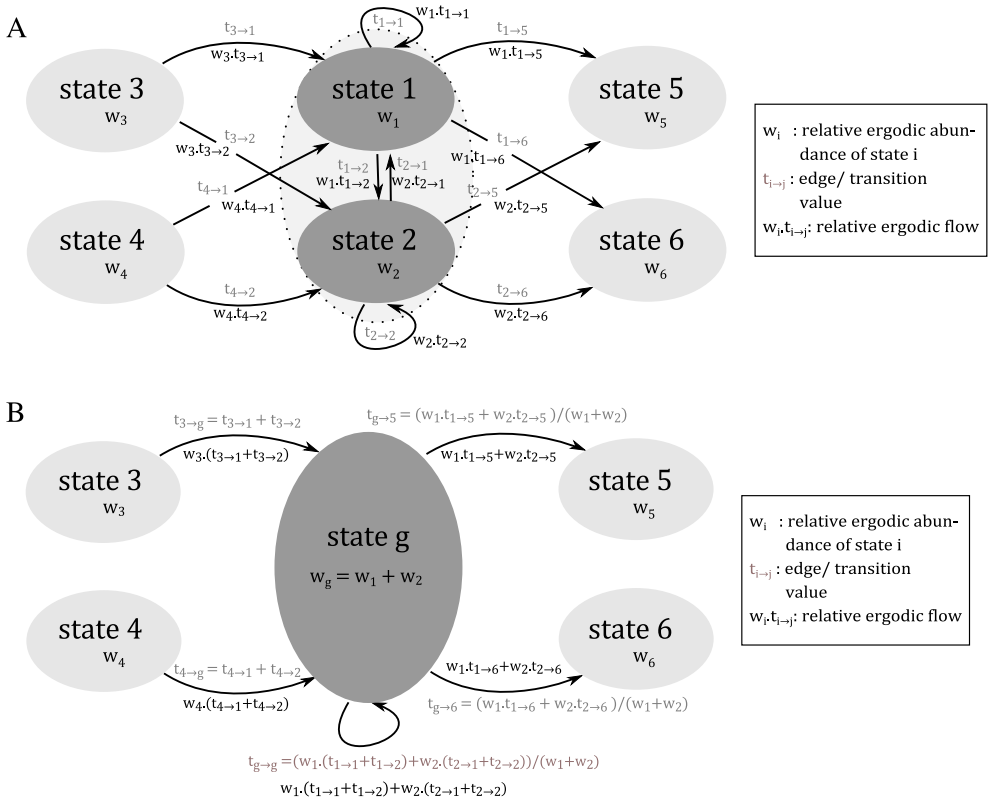


Fig. 1. EFP (Ergodic Flow Preserving) merging process illustrated by a generic six states system (**A**) where states 1 and 2 are merged into state g (**B**). Computing the new transitions inferred by the EFP merger of a group of states requires: (1) To perform a spectral analysis of the initial projection matrix of the entire graph of all transitions thus obtaining $\mathbf{w}$ (here $\mathbf{w} = [w_1, w_2, \ldots, w_6]$) associated with ergodic growth rate λ. (2) To sum incoming flows: edges formerly pointing towards a member of the group, now point to the group with identical transition value. Transitions from same state are summed. (3) To sum the ergodic-abundance-weighted outgoing flows: each edge leaving a member of the group will be replaced by a similar edge coming from the group but where the transition is scaled down by the relative ergodic-abundance of the former node to the abundance of the group, so that the asymptotic flow of individuals is preserved. Then, transitions towards the same state are summed. (4) To sum transitions internal to the group both ways and (5) To leave all transitions between states outside of the group unchanged.

Block-*Folding* "permutation" matrix, summing all elements sharing the same trait values for $\mathbf{t} \setminus \mathbf{st} = (t_{m+1}, \ldots, t_n)$, i.e., $\mathbf{P}^{\boldsymbol{BF}} = \left(\mathbf{I}_{\prod_{i=m+1}^{n}|t_i|} \otimes \mathbf{1}'_{\prod_{i=1}^{m}|t_i|}\right)$ where $\otimes$ is the Kronecker product, $\mathbf{I}_n$ the size n identity matrix product and $\mathbf{1}_n$ the size n column vector of 1s. Then the *folded* matrix can be written as follows:

$$\mathbf{M}^{\mathbf{fold}}_{\mathbf{t}\setminus\mathbf{st}} = \mathbf{P}^{\boldsymbol{BF}}.(\mathbf{M} \odot \mathbf{Wght}).\mathbf{P}'^{\boldsymbol{BF}}. \tag{4}$$

2.2.3. Implications of folding

This *folding* process preserves by construction the relative ergodic-abundances-weighted transitions. Preserving ergodic flows, it also preserves ergodic growth rate λ and relative abundance vector $\mathbf{w}$. This is however not the case for other demographic measures that may differ between MPPM $\mathbf{M}$ and *folded* matrices $\mathbf{M}^{\mathbf{fold}}_{\mathbf{t}\setminus\mathbf{st}}$ as for instance transient flows, net reproductive rate, generation time, life expectancy and most importantly reproductive values. Comparing these matrices outputs for $\mathbf{M}$ and $\mathbf{M}^{\mathbf{fold}}_{\mathbf{t}\setminus\mathbf{st}}$ provides crucial information on the role traits structure $\mathbf{s}$ plays on population dynamics.

We show in Supplementary Material 3 that the left eigenvector $\mathbf{v}$ is generally not preserved: the EFP *grouping* of several states preserves the relative reproductive values of all states only if contributions from all states (or future states) are equally broken down with regards to the soon-to-be-grouped states (all incoming flows to the future group coming from a single (future) state being a particular case). The non-conservation of reproductive value $\mathbf{v}$, "*the present value of the future offspring*" (Fisher, 1930), through EFP *merging* may at first appear rather surprising, as ergodic flows downstream of a state are preserved. However merging states implies simplifications i.e., loss of information, and by forcing ergodic flows and abundances to be constant, one allows relative reproductive value to vary.

In particular cases stated above and in Supplementary Material 3, some EFP *foldings* will preserve both $\mathbf{w}$ and $\mathbf{v}$; we shall call them *perfect foldings*. In the general case though, $\mathbf{v}$ will vary and this has fundamental implications: many important demographic measures depending on $\mathbf{v}$ will vary too, and most importantly the sensitivity matrix $\mathbf{S} = \frac{\partial\lambda}{\partial\mathbf{M}} = \mathbf{w}\mathbf{v}'$. This constitutes the core of the *trait level analysis*. A *folded* matrix $\mathbf{M}^{\mathbf{fold}}_{\mathbf{t}\setminus\mathbf{st}}$ models the exact same population than $\mathbf{M}$, but, without the presence of some traits, it may show different sensitivities. The interpretation of these differences may provide crucial information on the importance of the trait(s) *folded* upon for population dynamics and life-history evolution; and therefore on the potential cost of simplifying the trait structure of the life-cycle.

Some of the *folded* matrices resulting from such EFP *folding* along all-but-one traits will be of a more common use, especially at the beginning of a *trait level analysis*: the matrix *folded* over all traits but age $\mathbf{M}^{\mathbf{fold}}_{\mathbf{age}}$ (or stage $\mathbf{M}^{\mathbf{fold}}_{\mathbf{stage}}$) is one of them, and may be called the Reference Leslie (or Lefkovitch) Matrix.

3. Illustration

We illustrate this methodology with the analysis of a one-sex 3-trait population projection matrix incorporating age, parity (taken as the number of successful pregnancies) and fecundity classes (invariant during life, from sterile to very fertile). This is a *memory* model (where past-reproduction influences current reproduction) in a context of heterogeneity in female fecundity.

Table 3
Traits and states for the MPPM (multitrait population projection matrix) of the Illustration section incorporating age, parity and fecundity classes.

	t_1 = age	t_2 = parity	t_3 = fecundity class	Space of traits combinations	Space of states
Bounds	[1,99]	[0,11]	[1,10]	no ordering	[1, 11880]
Number of elements	99	12	10	$q = 99 \times 12 \times 10 = 11880$	$q = 11880$

More precisely, the model is designed to understand the evolutionary demographic implications of a negative effect of parity on fertility. The model is parameterized in the case of a human hunter-gatherer population (see Supplementary Material 4) and Matlab code to fully implement the following steps can be found in Supplementary Material 5.

3.1. Vector-based building of the age–parity–heterogeneity MPPM

3.1.1. Inputs

The model uses a time-step of one year. It incorporates 99 age classes, 12 parity categories and 10 fecundity classes; i.e., $n = 3$, $\mathbf{t} = (t_1, t_2, t_3) = (\text{age}, \text{par}, \text{fec})$, $\mathbf{s} = (99, 12, 10)$ and $q = 99 \times 12 \times 10 = 11{,}880$. The state-trait equivalence is given in (Table 3). Parametrization (Supplementary Material 4) provides us with the following vectors: **surv** the age-specific survival; **fert** the maximum age-specific fertility (thus for parity 0 and fecundity class 10); **parityeffect** the multiplying effect of parity on fertility per parity class; **classdistrib** the distribution of fecundity classes at birth; **classeffect** the multiplying effect of fecundity class on fertility per fecundity class. These parameter vectors are concatenated in $\mathbf{p} = [\mathbf{surv\ fert\ parityeffect\ classeffect\ classdistrib}]$ of size $k = 229$ (see Table in Supplementary Material 6).

Let us now implement both $\mathbf{vr}^v$ vectors; i.e., for each state $i = \boldsymbol{state}_{\mathbf{s}}(age, par, fec)$ its fertility rate $\text{vr}_i^{fertility}$ and its survival rate $\text{vr}_i^{survival}$. In this illustration, survival only depends on age and vector **surv** can then be replicated across parities and fecundity classes to generate $\mathbf{vr}^{survival}$. The fertility of an individual depends on the three traits such that: $\mathbf{vr}_i^{fertility} = \text{fert}_{age} \times \text{parityeffect}_{par} \times \text{classeffect}_{fec}$.

We can now generate for each state i the vectors $\mathbf{oi}^{v,i}$ and $\mathbf{op}^{v,i}$ of all potential output states and associated probability distribution, from i through vital process v over one time-step. Through survival, age increases by 1, parity increases by 1 or remains the same depending on reproductive success, and fecundity class is invariant. Thus the indices of states produced by input state i are $\mathbf{oi}^{survival,i} = (\boldsymbol{state}_{\mathbf{s}}(age + 1, par, fec), \boldsymbol{state}_{\mathbf{s}}(age + 1, par, fec))$, with distribution probability $\mathbf{op}^{survival,i} = \left(1 - \text{vr}_i^{fertility}, \text{vr}_i^{fertility}\right)$. Through fertility age becomes 1, parity becomes 0 and fecundity class is distributed according to **classdistrib**; i.e., $\mathbf{oi}^{fertility,i} = ((1, 0, 1), (1, 0, 2), \ldots, (1, 0, 10))$ and $\mathbf{op}^{fertility,i} = \mathbf{classdistrib}$.

3.1.2. Projection matrix

Vectors $\mathbf{vr}^v\mathbf{oi}^{v,i}$ and $\mathbf{op}^{v,i}$ allow us to generate **T**, the matrix of all transition triplets $\left(i, j, t^v_{i\to j}\right)$ and sparse definition of **M** (see Appendix B). In practice, we first generate $\mathbf{T}_1$, the matrix of all transition triplets through survival when parity remains constant, $\mathbf{T}_2$, the matrix of all transition triplets through survival when parity increases by 1 (together $\mathbf{T}_1$ and $\mathbf{T}_2$ form **TV** the matrix of survival transitions) and **TF** is the matrix of all transition triplets through fertility. We can then deduce **M** from **T** (see Appendix B).

3.1.3. Matrix properties

The number of implemented transitions (i.e., the number of rows of matrix **T**) is $\text{ntrans} = 141{,}340$, to be compared with the total number of entries in $\mathbf{M}, q^2 = 11{,}880^2 \approx 141.10^6$. The density of **M** is then $\text{ntrans}/q^2 \approx 0.001$ and **M** is definitely sparse. The gain in calculation time and storage space compared with the vec-permutation approach (Roth and Caswell, 2016) can be ascertained using the code allowing the construction of **M** with both techniques (Supplementary Material 5).

Matrix **M** contains numerous rows of zeroes, causing **M** to be reducible. The largest two eigenvalues of **M** are different $(1.01, 0.93 - 0.18i)$. Matrix **M** has therefore a unique and positive maximum real eigenvalue $\lambda = 1.01$ with non-negative right and left eigenvectors **w** and **v** which zeroes are only approximated by eigenvalue convergence algorithms and thus have to be nullified (see Supplementary Material 7). It must also be stressed that zeroes of **v** correspond to states having no impact on ergodic growth: menopaused and sterile women, and women having reached maximum parity. The eigentriad $(\lambda, \mathbf{w}, \mathbf{v})$ describes therefore the unique stable state towards which the population will tend asymptotically.

Having segregated fertility transitions (**TF**) and survival transitions (**TV**), we can construct the matrix of expected lifetime production **R** (see Supplementary Material 8). A fast extraction of **R**'s largest eigenvalue gives us $R_0 = 1.35$ daughters and generation time 26,6 years.

3.1.4. Stable-state analysis

The three-dimensional expression $\mathbb{W}$ of vector **w** provides the ergodic abundances by age, parity and fecundity class. The distribution of births by maternal classes of age, parity and fecundity is given by the product of state-wise abundances and fertility rates (written in matrix notation: $\mathbb{W} \odot \boldsymbol{vec}_{\mathbf{s}}^{-1}\left(\mathbf{vr}^{\boldsymbol{fertility}}\right)$). These distributions are depicted in absolute values in Supplementary Material 9 (A and B) and in relative contribution to each age class in Fig. 2A and 2B.

Comparing Fig. 2A and B allows visualizing the combined effects of the fertility–parity trade-off and fecundity heterogeneity. First, due to heterogeneity in fecundity, high fecundity classes play a disproportionate role in annual births by contrast with their low abundances (blue areas in B relative to A). Second, due to fertility–parity trade-off, the number of births decreases with parity: in each fecundity class, for high parities, the proportion of births by age-class in Fig. 2B is lower than the proportion of women by age-class in Fig. 2A (i.e., dark-shaded areas take up more space, light-shaded take up less space, in Fig. 2A than in Fig. 2B). Finally, due to the combination of both phenomena, the proportion of births from high-fecundity women decreases with age as they reach maximum parity earlier in life (decrease of blue area with age in Fig. 2B).

3.1.5. Sensitivity to parameters

Instead of using the usual matrix of sensitivities of λ to all entries of **M** from Eq. (1), we use the matrix **S** of sensitivities to transitions listed in **T** as it contains only biologically relevant transitions. Its sparse triplet definition can be built from **T** directly, where the 1st and 2nd columns of transitions input and output are the same and where the 3rd column (of transition rates in **T**) is replaced by the sensitivity of the transition rates calculated from **w** and **v** (i.e., in matrix notation, the sparse triplet definition of **S** is then $\left[\mathbf{T}_{.,1} \quad \mathbf{T}_{.,2} \quad \mathbf{w}_{\mathbf{T}_{.,1}} \odot \mathbf{v}_{\mathbf{T}_{.,2}}\right]$).

To study the influence of a particular set of parameters $\hat{\mathbf{p}}$ on our model, we build a symbolic version of $\hat{\mathbf{p}}$, denoted ${}_s\mathbf{p}$, where values are replaced by unknown variables using the relevant symbolic package of the computation programme used. Performing the three-step building method described in the methodology, we generate the formal matrix of all non-zero triplets of transitions ${}_s\mathbf{T}$ as function of ${}_s\mathbf{p}$ (the square matrix counterpart of ${}_s\mathbf{T}$, ${}_s\mathbf{M}$, is

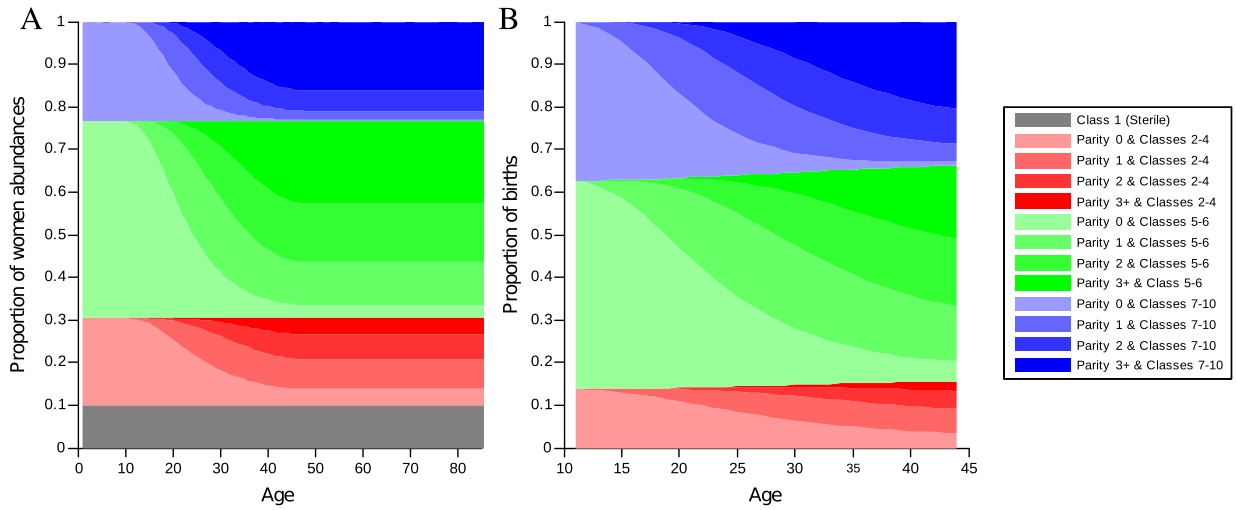


Fig. 2. Stable state distributions for each age class, grouped by parity and fecundity classes, of (**A**) women abundances and (**B**) annual births. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

too costly to build). We can now generate the multidimensional matrix of sensitivities of the MPPM to parameters, $\mathbb{S} = \left(\frac{\partial \mathbf{sM}}{\partial \mathbf{sp}}\right)_{|\mathbf{sp}=\hat{\mathbf{p}}}$. Then, we draw from Eq. (2) the sensitivity and elasticity to the l^{th} parameter; the elasticity $\mathbf{e}_x$ of λ to parameter x being defined as $\mathbf{e}_x = \frac{x}{\lambda}.\frac{\partial \lambda}{\partial x}$.

3.1.6. Sensitivity to fertility rates

The sensitivity of λ to other intermediary parameters like $\mathbf{vr}^{fertility}$ can be computed in two ways: either by adding the elements of $\mathbf{vr}^{fertility}$ to $\mathbf{p}$ or, more rapidly, by performing the 2nd level analysis directly. Indeed, in this illustration, $\text{vr}_i^{fertility}$ appears only a few times in $\mathbf{M}$, either in fertility transitions multiplied by **classeffect**, or in survival transitions multiplied either by **surv** or its complement to 1. Identifying and isolating these transitions yields all elements of $\frac{\partial \mathbf{M}}{\partial \mathbf{vr}^{\text{fertility}}}$ and thereby $\mathbf{e_M} = \left(\frac{\text{vr}_k^{\text{fertility}}}{\lambda} \sum_{i,j} \left(\frac{\partial \lambda}{\partial M_{i,j}} \times \frac{\partial M_{i,j}}{\partial \text{vr}_k^{\text{fertility}}}\right)\right)_{k=1,2..q}$, the elasticity of λ to the fertility rate of each state in $\mathbf{M}$. Fig. 3 displays these elasticities summed over various parity and fecundity classes. Indeed, elasticities are relevant sensitivity measures here since all fertilities are proportional to fert_{age} (with other factors structurally fixed) and thus they can be summed (the sum representing the relative effect on λ of the relative parallel change in fertilities).

Second level sensitivity analyses, alone, do not provide information on the cross-mechanisms between parity–fertility trade-off and fecundity heterogeneity. In order to understand these mechanisms, we need to perform a *trait level analysis*, i.e., to *fold* $\mathbf{M}$ over its traits and compare population dynamics inferred by resulting models.

3.2. Trait level analysis

We can sum $\mathbb{W}$ over its dimensions (Eq. (3)) to get expressions of $\mathbf{w}$ on only a subset of the traits: $\mathbb{W}^{\mathbf{fold}}_{\text{age}}$ (the ergodic abundance vector of the population characterized only by age), $\mathbb{W}^{\mathbf{fold}}_{\text{age,fec}}$ (the ergodic abundance vector of the population characterized by age and fecundity class), $\mathbb{W}^{\mathbf{fold}}_{\text{par,fec}}$, $\mathbb{W}^{\mathbf{fold}}_{\text{par}}$, ...

We replicate Eqs. (3) and (4), in order to build a **fold** function (see code in Supplementary Material 5) that *folds* $\mathbf{M}$ over any subset of its traits, providing $\mathbf{M}^{\mathbf{fold}}_{\text{age}}$, $\mathbf{M}^{\mathbf{fold}}_{\text{age,fec}}$, $\mathbf{M}^{\mathbf{fold}}_{\text{fec}}$, $\mathbf{M}^{\mathbf{fold}}_{\text{age,par}}$,.... For instance, the Reference Leslie Matrix $\mathbf{M}^{\mathbf{fold}}_{\text{age}}$ is $\mathbf{M}$ *folded* on all traits but age. Being a Leslie matrix, fertilities arise only on the first row (i.e., in matrix notation: $\mathbf{fert}_{\mathbf{M}^{\mathbf{fold}}_{\text{age}}} = \mathbf{M}^{\mathbf{fold}}_{\text{age}}(1, *)$). An eigenanalysis

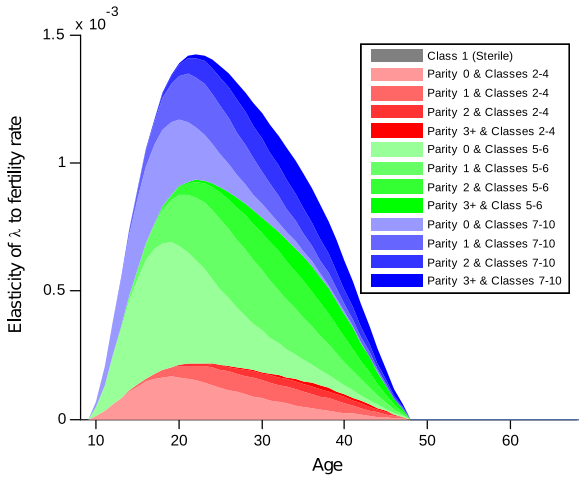


Fig. 3. Elasticity of the asymptotic growth rate λ to the fertility rate at each age according to parity and heterogeneity classes, for matrix $\mathbf{M}$. $\mathbf{M}$ is the matrix modeling the population characterized by traits age, parity and fecundity heterogeneity.

of $\mathbf{M}^{\mathbf{fold}}_{\text{age}}$, provides $\mathbf{v}_{\mathbf{M}^{\mathbf{fold}}_{\text{age}}}$, the right eigenvector associated with λ ($\mathbf{M}$ and $\mathbf{M}^{\mathbf{fold}}_{\text{age}}$ have, by construction, the same growth rate). The matrix of sensitivity of λ to $\mathbf{M}^{\mathbf{fold}}_{\text{age}}$ is then given by: $\mathbf{S}_{\mathbf{M}^{\mathbf{fold}}_{\text{age}}} = \mathbf{w}_{\mathbf{M}^{\mathbf{fold}}_{\text{age}}}.\mathbf{v}_{\mathbf{M}^{\mathbf{fold}'}_{\text{age}}}$ (which can be implemented as $\boldsymbol{vec}\left(\mathbb{W}^{\mathbf{fold}}_{\text{age}}\right).\mathbf{v}_{\mathbf{M}^{\mathbf{fold}'}_{\text{age}}}$). Finally, from $\mathbf{S}_{\mathbf{M}^{\mathbf{fold}}_{\text{age}}}$ we can generate the elasticity of λ to $\mathbf{fert}_{\mathbf{M}^{\mathbf{fold}}_{\text{age}}}$: $\mathbf{e}_{\mathbf{M}^{\mathbf{fold}}_{\text{age}}}$ (implemented by: $\mathbf{e}_{\mathbf{M}^{\mathbf{fold}}_{\text{age}}} = \left(\mathbf{fert}_{\mathbf{M}^{\mathbf{fold}}_{\text{age}}}/\lambda\right) \odot \mathbf{S}_{\mathbf{M}^{\mathbf{fold}}_{\text{age}}}(1, *)$).

As we did above for $\mathbf{M}$, we can extract implicit fertilities from $\mathbf{M}^{\mathbf{fold}}_{\text{age, par}}$ and $\mathbf{M}^{\mathbf{fold}}_{\text{age, fec}}$ as well as their associated elasticities $\mathbf{e}_{\mathbf{M}^{\mathbf{fold}}_{\text{age, fec}}}$ and $\mathbf{e}_{\mathbf{M}^{\mathbf{fold}}_{\text{age,par}}}$. As for $\mathbf{M}$, elasticities can be summed over one or several traits, and compared, as we do for $\mathbf{e}_{\mathbf{M}^{\mathbf{fold}}_{\text{age, fec}}}$, $\mathbf{e}_{\mathbf{M}^{\mathbf{fold}}_{\text{age, par}}}$ and $\mathbf{e}_{\mathbf{M}}$, summed on age so they can be compared with one another and with $\mathbf{e}_{\mathbf{M}^{\mathbf{fold}}_{\text{age}}}$ (Fig. 4).

Two matrices where one is derived from the other by *perfect* EFP *merger* will have equal sum of elasticities to fertilities. Indeed, as perfect EFP *mergers* preserve right and left eigenvectors, they equally preserve elasticities. In other cases, sums will differ.

Here, it can be shown that the *folding* over fecundity from $\mathbf{M}^{\mathbf{fold}}_{\text{age,fec}}$ to $\mathbf{M}^{\mathbf{fold}}_{\text{age}}$ is a *perfect* EFP *folding*. Indeed, the only "new" state with several direct ancestors is age $= 1$ and all its former

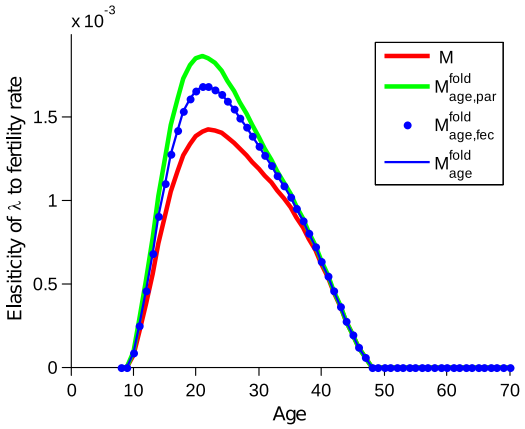


Fig. 4. Elasticity, summed by age, of the asymptotic growth rate λ to the fertility rates for matrices $\mathbf{M}$, $\mathbf{M}^{\mathbf{fold}}_{\mathbf{age,par}}$, $\mathbf{M}^{\mathbf{fold}}_{\mathbf{age,fec}}$ and $\mathbf{M}^{\mathbf{fold}}_{\mathbf{age}}$. $\mathbf{M}$ is the matrix modeling the population characterized by traits age, parity and fecundity heterogeneity. $\mathbf{M}^{\mathbf{fold}}_{\mathbf{age,par}}$ is $\mathbf{M}$ folded over fecundity. $\mathbf{M}^{\mathbf{fold}}_{\mathbf{age,par}}$ is $\mathbf{M}$ folded over parity. $\mathbf{M}^{\mathbf{fold}}_{\mathbf{age}}$ is $\mathbf{M}$ folded over parity and fecundity. See Eq. (4) for *folding* formalization.

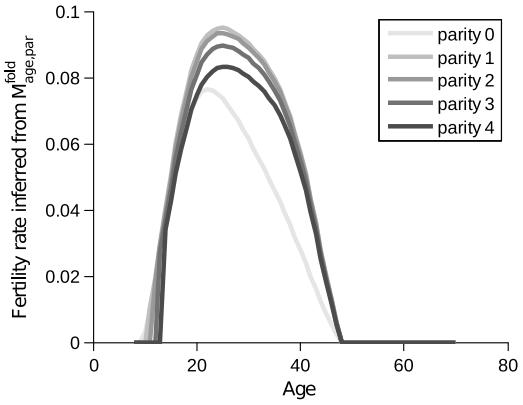


Fig. 5. Fertility rates inferred from the matrix *folded* upon fecundity, $\mathbf{M}^{\mathbf{fold}}_{\mathbf{age,par}}$, for some age and parity classes.

states were contributed proportionally along **classdistrib** (see Supplementary Material 3). Similar analyses show that the other discussed EFP *foldings* are not *perfect*. Consequently, for $\mathbf{M}$ and $\mathbf{M}^{\mathbf{fold}}_{\mathrm{age,par}}$ the sum of elasticities to fertilities and survivals will differ from 1, its value for the Reference Leslie Matrix $\mathbf{M}^{\mathbf{fold}}_{\mathrm{age}}$ where fertilities and survivals are precisely the entries of the matrix, and thus also for $\mathbf{M}^{\mathbf{fold}}_{\mathrm{age,fec}}$.

This simple *trait level analysis* yields two important results. First the sum on age of elements of $\mathbf{e}_{\mathbf{M}}$ is below that of $\mathbf{e}_{\mathbf{M}^{\mathbf{fold}}_{\mathrm{age}}}$ (Fig. 4). This is due to the fact that $\mathbf{M}$, contrary to $\mathbf{M}^{\mathbf{fold}}_{\mathrm{age}}$ has a trade-off implemented between parity and fertility: in $\mathbf{M}$, a successful fertility event for an individual decreases its expected fertility rates for all future fertility events whereas these events are independent in $\mathbf{M}^{\mathbf{fold}}_{\mathrm{age}}$. Consequently, this trade-off minimizes the influence of each specific annual fertility rate to the benefit of the lifetime offspring production and thus reduces the variance of the reproductive success. In a population with a parity–fertility trade-off, not implementing the relevant trait (parity) and trade-off leads to an overvaluation of the impact of annual fertility rates on the population growth rate and of the variance of the reproductive success.

Second, for each age class $\sum_{\mathrm{par}}\mathbf{e}_{\mathbf{M}^{\mathbf{fold}}_{\mathrm{age,par}}}$ is larger than both $\sum_{\mathrm{par,fec}}\mathbf{e}_{\mathbf{M}}$ and $\mathbf{e}_{\mathbf{M}^{\mathbf{fold}}_{\mathrm{age}}}$ (on Fig. 4). In symmetry with the preceding result, this may signal positive correlation between parity and fertility inferred by $\mathbf{M}^{\mathbf{fold}}_{\mathrm{age,par}}$. This correlation can indeed be detected in the fertility rates for all states, $\mathbf{fert}_{\mathbf{M}^{\mathbf{fold}}_{\mathrm{age,par}}}$, inferred by $\mathbf{M}^{\mathbf{fold}}_{\mathrm{age,par}}$ (Fig. 5): In $\mathbf{M}^{\mathbf{fold}}_{\mathrm{age,par}}$, where the *unobserved* heterogeneity class trait is *folded* upon, fertility does increase with parity for low parities, then plateaus, and finally, for high parities, decreases with parity. Since low parities are more abundant in the population, on average, $\mathbf{M}^{\mathbf{fold}}_{\mathrm{age,par}}$ exhibits a positive correlation between fertility and parity. Thus, in a heterogeneous population with parity–fertility trade-off, not modeling individual heterogeneity leads to misunderstanding the direction of the parity–fertility correlation. Implementing heterogeneity – even in the simplest manner – would prevent this misinterpretation.

4. Discussion

In this article, we propose a novel methodology to build multitrait population projection matrices (MPPMs) in a computationally efficient manner. We then discuss the implications of the general reducibility of MPPMs. Then we broaden stable-state demographic measures calculations and perturbation analysis tools (Caswell, 2012; Willekens, 1977) for such matrices. Finally, and most importantly, we propose a new tool, the *trait level analysis*, which investigates the impact of the trait structure (the number of traits and their characteristics) on population dynamics via the *folding* of MPPMs over any subset of their traits. As an illustration, we apply this methodology to the construction and analysis of a one-sex 3-traits (age, parity and fecundity class) model designed to study the evolutionary implications of both cost of reproduction and heterogeneity in fecundity.

4.1. Construction of MPPM and stable state theory tools

The methodology provides a step-by-step technique to generate any multitrait projection population matrix in a computationally inexpensive manner. This technique applies to any possible multitrait model, extending its reach far beyond the original multiregional models (Le Bras, 1970; Rogers, 1980, 1969, 1966). As such it will be useful to a wide range of researchers interested in population dynamics and evolution of Life History. The three step building method involves the computation of vital rates for all states, output states and output distributions for all states. Obviously this building technique is only a recommendation and the analysis tools provided in the second part of the methodology apply to any MPPM, whether built according to this method or not. However we think this method has the significant advantages, compared to the basic transition-by-transition approach and the vec-permutation method developed by Caswell and colleagues (Caswell, 2014; Caswell and Salguero-Gómez, 2013; Caswell and Shyu, 2012; Hunter and Caswell, 2005; Roth and Caswell, 2016) of (1) providing a framework applicable to all kind of traits, (2) avoiding loops and hence optimizing calculation time, and (3) only generating the meaningful transitions and thus making full use of the inherent sparsity of multitrait matrices.

We have implemented the computations needed for the illustration using Matlab (Matlab 2012a, The MathWorks Inc., Natick, MA, United States), a commercial software package. The mathematical functions presented could however be coded by other softwares, as long as they handle sparse matrices, multidimensional matrices and formal/symbolic matrices.

4.2. Eigen- and graph-properties of MPPMs

MPPMs, as structurally complex as they may be, preserve the one-trait matrices' ability to quickly generate growth rate,

stable-state abundances and reproductive values, as well as all other demographic functions that classical transient and stable-state theory provides. However, contrary to most one-trait projection matrices, they are not reducible. This has many implications with regards to the asymptotic dynamics of the population modeled by an MPPM. In Supplementary Material 1, we use the Frobenius normal form a matrix to show that the graph associated with an MPPM may contain several components each generating its own growth rate. Studies of the spectrum and normal form of MPPMs are required to improve our understanding of the relative importance of those components on transient and ergodic dynamics (e.g., the study from Li and Schreiber, 2006 of the graph properties of age-and-location MPPMs). Such interdisciplinary works would involve theoretical ecologists, graph theorists and linear algebraists. We denote split-population, these MPPMs which graph contains at least two strongly connected components, each with real positive maximal eigenvalues, and that are not bilaterally connected with one another. These special cases of MPPMs, in which, for some initial conditions, the asymptotical growth rate will be lower than the maximal eigenvalue, would also deserve in-depth studies.

4.3. Evolutionary demography and trait level analysis

In this article, we extend the classical calculation tools for perturbation analysis to multitrait matrices. For instance, we make it possible to generate at once, all lower-level sensitivities of the population's growth rate, through the construction of the multidimensional sensitivity matrix $\mathbb{S}$. Evolutionary ecologists and demographers will also benefit from the *trait level analysis* we devised, i.e., the ability to compare matrices (and the sensitivities of their growth rates) derived from various *foldings* of an initial MPPM over any number of its traits.

This *folding* process is the application of the Ergodic Flow Preserving (EFP) *merging* process to all states of an MPPM that share the same value for the trait to be *folded* upon. Increasing the number of traits in a model allows zooming in the dynamics of a population and increasing the granularity of the analyses and the forecasts. This will help refining the understanding of the trait-related processes driving the evolution of the studied population. Decreasing the number of traits by *folding* an MPPM allows zooming out and synthesizing the data: it will allow the ecologist to understand how each specific trait and the relationships amongst traits impact the dynamics and fitness of the population and its multitrait cohorts.

The EFP *merging* process, one in many ways of merging two states of a state transition graph, has the characteristic of preserving ergodic abundances, but at the expense of reproductive values, except in the particular cases of *perfect* EFP *folding*, where both relative ergodic abundances and reproductive values are preserved. This loss of information has far-reaching significance on what adding or subtracting (i.e., *folding on*) a trait means. In particular, this means that the sensitivities of fitness to parameters will vary as a function of the trait structure for matrices however modeling the same ergodic population. Consequently, the *trait level analysis* has the potential to investigate the effect of trait structure on the force of selection.

In the case of the age–parity–heterogeneity model developed in the Illustration section, the *trait level analysis* provides two main results. (i) In the case of the *unfolded* MPPM incorporating fertility–parity trade-off, the effect of natural selection on age-specific fertility is decreased compared to its Reference Leslie Matrix (an MPPM *folded* over all traits but age). Indeed the trade-off favors overall realized fertility over annual fertility and thus reduces the variance of reproductive success and the sensitivity to each single yearly fertility rate. (ii) *Folding* the 3-trait MPPM over heterogeneity increases the impact of natural selection on fertility even beyond the level of the Reference Leslie Matrix. We have shown that it is related to the fact that the age–parity model resulting from the full-trait MPPM *folded* over heterogeneity, exhibits a positive correlation between fertility and parity at low parities. In other words, a heterogeneous population in which reproduction is costly for fertility may however exhibit a positive correlation between fertility and parity when heterogeneity is not modeled as a trait in the MPPM. This is the evolutionary demography counterpart of the seminal results of Van Noordwijk and de Jong (1986), proving how large variance in heterogeneity (acquisition) can make the manifestations of the costs of reproduction "invisible".

4.4. Further prospects for trait level analysis

Adding a trait to a model by extending or *unfolding* an MPPM may generate a different second real maximal eigenvalue. In population dynamics terms, this means that the addition of traits alters the damping ratio. There is potential in analyzing the impact of trait structures on the transient dynamics of populations by performing spectral analysis of the *folded* and *unfolded* MPPMs. For instance, the possible (de)stabilizing properties of specific traits, group of traits and trait patterns in periodic (extending Tuljapurkar, 1985), stochastic (extending, among others, Tuljapurkar, 1986a, 1986b, Tuljapurkar and Orzack, 1980) or density-dependent environments (extending Caswell, 2008; Caswell et al., 2004) could be investigated via *trait level analysis*.

Transient analysis of populations modeled by MPPMs could make use of tools developed by Caswell (2007) to understand whether the main traits shaping the population structure make similarly important contribution to short term dynamics. Economists have developed a transient-dynamics merging analysis theory (Fisher, 1969, 1958; Simon and Ando, 1961). It was later partly used in a study by Rogers (1969) of *perfect* aggregation, when merging implies no loss of information, i.e., special cases of quasi identical states (see Supplementary Material 2). It would be interesting to extend the measuring of transient merging efficiency, as done for example with the linear aggregation coefficient introduced by Ijiri (1968), towards EFP *merging*. In particular, *trait level analysis* may be useful to study the effect on the distance between initial and stable-state population distributions and on 'population inertia' of adding or hiding one of several traits; thus allowing to better understand the varying roles of certain traits on 'real word' dynamics (see a review of transient analyses tools by Stott et al., 2011).

Generation time and net reproductive rate can be produced on any subset of the space of traits, and will be particularly meaningful with the Reference Leslie (or Lefkovitch) Matrix. The impact of multiple traits on the calculation of generation time and net reproductive rates for the MPPM and its various *folded* versions could be studied, extending the work by Steiner et al. (2014).

We have proved that the general merging of states implies a loss of information and demonstrated why a merging process preserving ergodic flows and abundances made sense from an evolutionary demography perspective. However alternative *folding* techniques should be considered. In a recent article published in this journal, Bienvenu et al. (2017) draw on earlier work on genealogical Markov chains associated with matrix population models (Bienvenu and Legendre, 2015; Demetrius, 1975, 1974; Tuljapurkar, 1993, 1982) in order to describe an alternative state-merging process. This process, they call "genealogical collapsing", has the property to preserve ergodic growth rate, abundances *and* reproductive values, but at the expense of ergodic flows. It would thus preserve sensitivities, and could constitute, as such, a 'neutral' comparative tool when extended to entire traits. It would also be worth investigating a *folding* process, preserving ergodic flows, growth and reproductive values at the cost of abundances. The

Table A.1
Traits and states for generic MPPM (multitrait population projection matrix).

	trait t_1	…	trait t_n	Space of traits combinations	Space of states
Bounds	$[min_1, max_1]$	…	$[min_n, max_n]$	no ordering	$[1, q]$
Number of elements	$\lvert t_1\rvert = max_1 - min_1 + 1$	…	$\lvert t_n\rvert = max_n - min_n + 1$	$q = \prod_{i=1}^{n} \lvert t_i\rvert$	q

potential value for multitrait analysis of such alternative merging methods deserves investigation. If, in general, reproductive values are not preserved by Ergodic Flow Preserving merging, it means that the weight of transitions, from an evolutionary demography viewpoint, is also not preserved both within and without a future grouping. In other words, the contribution of a group of states to fitness is not the sum of the contributions of its components. This is for example demonstrated in our age–parity–heterogeneity illustration where fertility rates by age class have dramatically different influences on fitness when parity is implemented or not. The implication is important: grouping states, adding traits, removing traits whilst preserving ergodic flows and abundances, will generally lead to changes in the influence of all states on the ergodic growth rate. Future theoretical and empirical works are needed to investigate these issues.

Acknowledgments

This work benefited from the support of the Chair "Modélisation Mathématique et Biodiversité of Veolia –Ecole Polytechnique–MNHN–Fondation X", in particular through a Ph.D. Grant to CFD.C. We thank Vincent Bansaye, Minus van Baalen and Pierre-Henri Gouyon for helpful comments and suggestions. We also wish to thank Roberto Salguero-Gomez, Uli Steiner and one anonymous reviewer for comments that greatly improved the manuscript.

Appendix A. Vectorization of the multidimensional space of traits

Suppose a life-history described by a combination of n discretized traits $\mathbf{t} = (t_1, t_2, \ldots, t_n)$, where the value of trait i, t_i can be any positive integer between min_i and max_i. Then the vector of trait sizes (also called trait structure) is $\mathbf{s} = (\max_1 - \min_1 + 1, \max_2 - \min_2 + 1, \ldots, \max_n - \min_n + 1)$. Each n-tuple of trait values$(t_1, t_2, \ldots, t_n)$, hereafter called an individual *tuple*, has to be converted into a single *state* corresponding to a specific row or column of $\mathbf{M}$ where $M_{i,j}$ corresponds to the transition rate $t_{j\to i}$ from state j to state i. This requires the vectorization of the space of *tuples* into the space of *states*. As those two spaces are isomorphic, they have the same cardinality q (Table A.1). The isomorphism converting any *tuple* $(t_1, t_2, \ldots, t_n)$ into its *state* can be written as a closed-form expression:

$$\boldsymbol{state}_{\mathbf{s}}(t_1, t_2, \ldots, t_n) = 1 + \sum_{i=1}^{n} (t_i - min_i) \prod_{j=0}^{i-1} \left(max_j - min_j + 1\right),$$

where $min_0 = max_0 = 1$.

By contrast, the reverse isomorphism, $\boldsymbol{tuple}_{\mathbf{s}}$, identifying the n-tuple of trait values corresponding to a given *state* i requires a loop (of complexity $O(n)$).

As they are equivalent, in the article, we often refer to both *tuples* and *states*, as states.

Appendix B. Relationships between transitions and vectors of vital rates, output states indices and distribution

Each entry of an MPPM, $M_{k,l}$, can be decomposed into its survival and fertility parts: $M_{k,l} = M_{k,l}^{survival} + M_{k,l}^{fertility}$, or if written in terms of graph transitions: $M_{k,l} = t_{l\to k}^{survival} + t_{l\to k}^{fertility}$. Each of these $t_{l\to k}^{v}$ can then be decomposed as a vital rate and a probability distribution, or using the notations of the article:

$$\forall i \in [1, q], \forall v \in \{fertility, \text{survival}\}, \forall m \in \left[1, \lvert \mathbf{oi}^{v,i}\rvert\right],$$
$$t^{v}_{i\to \mathrm{oi}^{v,i}_{m}} = \mathrm{vr}_i^v \times \mathrm{op}_m^{v,i}.$$

For each state i, and each vital process v, we can then create a matrix of transitions $\mathbf{T}^{v,i}$, gathering all triplets of transitions $\left(i, j, t_{i\to j}^{v}\right)$ between state i and every state j reachable from i. For instance:

$$\mathbf{T}^{v,i} = \left[i 1_{\lvert \mathbf{oi}^{v,i}\rvert} \quad \mathbf{oi}^{v,i} \quad \mathrm{vr}_i^v \mathbf{op}^{v,i}\right],$$

where $1_{\lvert \mathbf{oi}^{v,i}\rvert}$ is a vector of 1s of size, the size of vector $\mathbf{oi}^{v,i}$.

Concatenating the $\mathbf{T}^{v,i}$ matrices for all states and for both vital rates gives us:

$$\mathbf{T} = \left[\cap_i \mathbf{T}^{fertility,i}\right] \cap \left[\cap_i \mathbf{T}^{survival,i}\right],$$

where $\cap$ denotes the row concatenation of matrices.

$\mathbf{M}$ and $\mathbf{T}$ are equivalent, in the sense that they both contain the same information: from $\mathbf{T}$ we can generate $\mathbf{M}$:

$$\forall i, j \in [1, q]^2 \quad \forall l \in [1, \text{ntrans}] \mid \left(\mathrm{T}_{l,2}, \mathrm{T}_{l,1}\right) = (i, j)$$
$$M_{\mathrm{i,j}} = \sum_{l} t_{\mathrm{T}_{l,1}\to \mathrm{T}_{l,2}} = \sum_{l} \mathrm{T}_{l,3}.$$

Appendix C. Computational complexity of common operations on sparse matrices

Two of the most important computations for large projection matrices are accelerated when operated on a $q \times q$ sparse matrix. Fast sparse matrix multiplication (see Yuster and Zwick, 2005) allows to reduce complexity from $O\left(q^3\right)$ to $O\left(q^2\right)$. Fast eigenanalysis convergence processes, like the Arnoldi Iteration where finding the k largest eigenvalues only costs $O\left(k^2 q\right)$ (Arnoldi, 1951) massively improve naïve eigenvalues algorithms on a non-sparse matrix which have time complexity of $O\left(q^3\right)$.

Appendix D. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.tpb.2017.07.002.

References

Ara, K., 1959. The aggregation problem in input-output analysis. Econometrica 27, 257–262. http://dx.doi.org/10.2307/1909446.

Arnoldi, W.E., 1951. The principle of minimized iteration in the solution of the matrix eigenvalue problem. Q. Appl. Math. 9, 17–29.

Bell, G., 1980. The costs of reproduction and their consequences. Am. Nat. 116 (1), 45–76. http://dx.doi.org/10.2307/2460709.

Bienvenu, F., Akçay, E., Legendre, S., McCandlish, D.M., 2017. The genealogical decomposition of a matrix population model with applications to the aggregation of stages. Theor. Popul. Biol. 115, 69–80. http://dx.doi.org/10.1016/j.tpb.2017.04.002.

Bienvenu, F., Legendre, S., 2015. A new approach to the generation time in matrix population models. Am. Nat. 185, 834–843. http://dx.doi.org/10.1086/681104.

Boonekamp, J.J., Salomons, M., Bouwhuis, S., Dijkstra, C., Verhulst, S., 2014. Reproductive effort accelerates actuarial senescence in wild birds: an experimental study. Ecol. Lett. 17, 599–605. http://dx.doi.org/10.1111/ele.12263.

Caswell, H., 1978. A general formula for the sensitivity of population growth rate to changes in life history parameters. Theor. Popul. Biol. 14, 215–230. http://dx.doi.org/10.1016/0040-5809(78)90025-4.

Caswell, H., 2001. Matrix Population Models. Sinauer Associates Inc., Sunderland, MA.
Caswell, H., 2007. Sensitivity analysis of transient population dynamics. Ecol. Lett. 10, 1–15. http://dx.doi.org/10.1111/j.1461-0248.2006.01001.x.
Caswell, H., 2008. Perturbation analysis of nonlinear matrix population models. Demogr. Res. 18, 59–116. http://dx.doi.org/10.4054/DemRes.2008.18.3.
Caswell, H., 2012. Matrix models and sensitivity analysis of populations classified by age and stage: a vec-permutation matrix approach. Theor. Ecol. 5, 403–417. http://dx.doi.org/10.1007/s12080-011-0132-2.
Caswell, H., 2014. A matrix approach to the statistics of longevity in heterogeneous frailty models. Demogr. Res. 31, 553–592. http://dx.doi.org/10.4054/DemRes.2014.31.19.
Caswell, H., Kluge, F.A., 2015. Demography and the statistics of lifetime economic transfers under individual stochasticity. Demogr. Res. 32, 563–588. http://dx.doi.org/10.4054/DemRes.2015.32.19.
Caswell, H., Salguero-Gómez, R., 2013. Age, stage and senescence in plants. J. Ecol. 101, 585–595. http://dx.doi.org/10.1111/1365-2745.12088.
Caswell, H., Shyu, E., 2012. Sensitivity analysis of periodic matrix population models. Theor. Popul. Biol. 82, 329–339. http://dx.doi.org/10.1016/j.tpb.2012.03.008.
Caswell, H., Takada, T., Hunter, C.M., 2004. Sensitivity analysis of equilibrium in density-dependent matrix population models. Ecol. Lett. 7, 380–387. http://dx.doi.org/10.1111/j.1461-0248.2004.00595.x.
de Kroon, H., Plaisier, A., van Groenendael, J., Caswell, H., 1986. Elasticity: The relative contribution of demographic parameters to population growth rate. Ecology 67, 1427–1431. http://dx.doi.org/10.2307/1938700.
Demetrius, L., 1974. Demographic parameters and natural selection. Proc. Natl. Acad. Sci. 71, 4645–4647.
Demetrius, L., 1975. Natural selection and age structured populations. Genetics 79, 535–544.
Enright, N.J., Franco, M., Silvertown, J., 1995. Comparing plant life histories using elasticity analysis: theimportance of life span and the number of life - cycle stages. Oecologia 104, 79–84.
Euler, L., 1760. Recherches générales sur la mortalité et la multiplication. Mem. Acad. R. Sci. Belles Lett. 16, 144–164.
Fisher, R.A., 1930. The Genetical Theory of Natural Selection, Genetics. Oxford University Press. http://dx.doi.org/10.1038/158453a0.
Fisher, W.D., 1958. Criteria for aggregation in input-output analysis. Rev. Econ. Stat. 40, 250. http://dx.doi.org/10.2307/1927416.
Fisher, W.D., 1969. Clustering and Aggregation in Economics. The Johns Hopkins University Press.
Giske, J., Aksnes, D.L., Førland, B., 1993. Variable generation times and Darwinian fitness measures. Evol. Ecol. 7, 233–239. http://dx.doi.org/10.1007/BF01237741.
Goodman, L.A., 1969. The analysis of population growth when the birth and death rates depend upon several factors. Biometrics 25, 659–681. http://dx.doi.org/10.2307/2528566.
Goodman, L.A., 1971. On the sensitivity of the intrinsic growth rate to changes in the age-specific birth and death rates. Theor. Popul. Biol. 2, 339–354. http://dx.doi.org/10.1016/0040-5809(71)90025-6.
Hanski, I., 1999. Habitat connectivity, habitat continuity, and metapopulations in dynamic landscapes. Oikos 87, 209–219. http://dx.doi.org/10.2307/3546736.
Henderson, H.V., Searle, S.R., 1981. The vec-permutation matrix, the vec operator and kronecker products: a review. Linear Multilinear Algebr. 9, 271–288. http://dx.doi.org/10.1080/03081088108817379.
Hooley, D.E., 2000. Collapsed matrices with (almost) the same eigenstuff. College Math. J. 31, 297. http://dx.doi.org/10.2307/2687420.
Hunter, C.M., Caswell, H., 2005. The use of the vec-permutation matrix in spatial matrix population models. Ecol. Model. 188, 15–21. http://dx.doi.org/10.1016/j.ecolmodel.2005.05.002.
Ijiri, Y., 1968. The linear aggregation coefficient as the dual of the linear correlation coefficient. Econometrica 36, 252.
Lambert, J.H., 1772. Beyträge Zum Gebrauche Der Mathematik Und Deren Anwendung, Volume 3.
Law, R., 1983. A model for the dynamics of a plant population containing individuals clasified by age and size. Ecology. http://dx.doi.org/10.2307/1937069.
Le Bras, H., 1970. Equilibre et croissance de populations soumises a des migrations. Theor. Popul. Biol. 121, 100–121.
Lebreton, J.D., 1996. Demographic models for subdivided populations: the renewal equation approach. Theor. Popul. Biol. 49, 291–313. http://dx.doi.org/10.1006/tpbi.1996.0015.
Lebreton, J.D., 2005. Age, stages, and the role of generation time in matrix models. Ecol. Model. 188, 22–29. http://dx.doi.org/10.1016/j.ecolmodel.2005.05.003.
Lefkovitch, L.P., 1965. The study of population growth in organisms grouped by stages. Biometrics 21, 1–18. http://dx.doi.org/10.2307/2528348.
Lemaître, J.-F., Berger, V., Bonenfant, C., Douhard, M., Gamelon, M., Plard, F., Gaillard, J.-M., 2015. Early–late life trade-offs and the evolution of ageing in the wild. Proc. Biol. Sci. 282, 20150209. http://dx.doi.org/10.1098/rspb.2015.0209.
Leontief, W., 1986. An alternative to aggregation in input–output analysis and national accounts. In: Input-Output Economics. pp. 41–54.
Leslie, P.H., 1945. On the use of matrices in certain population mathematics. Biometrika 33, 183–212. http://dx.doi.org/10.2307/2332297.
Lewis, E.G., 1942. On the generation and growth of a population. Sankyhā 6 (part 1), 93–96. http://dx.doi.org/10.1007/978-3-642-81046-6_25.
Li, C., Schreiber, S., 2006. On dispersal and population growth for multistate matrix models. Linear Algebra Appl. 418, 900–912. http://dx.doi.org/10.1016/j.laa.2006.03.025.
Li, T., Anderson, J.J., 2009. The vitality model: A way to understand population survival and demographic heterogeneity. Theor. Popul. Biol. 76, 118–131. http://dx.doi.org/10.1016/j.tpb.2009.05.004.
Morimoto, Y., 1970. On aggregation problems in input-output analysis. Rev. Econ. Stud. 37, 119–126. http://dx.doi.org/10.2307/2296502.
Murray, B.G., 1992. The evolutionary significance of lifetime reproductive success. Auk 109, 167–172.
Nur, N., 1984. Fitness, population growth rate and natural selection. Oikos 42, 413. http://dx.doi.org/10.2307/3544416.
Pavard, S., Branger, F., 2012. Effect of maternal and grandmaternal care on population dynamics and human life-history evolution: A matrix projection model. Theor. Popul. Biol. 82, 364–376. http://dx.doi.org/10.1016/j.tpb.2012.01.007.
Plard, F., Bonenfant, C., Delorme, D., Gaillard, J.M., 2012. Modeling reproductive trajectories of roe deer females: Fixed or dynamic heterogeneity? Theor. Popul. Biol. 82, 317–328. http://dx.doi.org/10.1016/j.tpb.2012.03.006.
Rogers, A., 1966. The multiregional matrix growth operator and the stable interregional age structure. Demography 3, 537–544. http://dx.doi.org/10.2307/2060178.
Rogers, A., 1969. On perfect aggregation in the matrix-survival model of interregional population growth. J. Reg. Sci. 9, 417–424.
Rogers, A., 1974. The multiregional net maternity function and multiregional stable growth. Demography 11, 473–481. http://dx.doi.org/10.2307/2060439.
Rogers, A., 1980. Introduction to multistate mathematical demography. Environ. Plan. A 12, 489–498. http://dx.doi.org/10.1068/a120489.
Roth, G., Caswell, H., 2016. Hyperstate matrix models: extending demographic state spaces to higher dimensions. Methods Ecol. Evol. http://dx.doi.org/10.1111/2041-210X.12622.
Salguero-Gómez, R., Plotkin, J.B., 2010. Matrix dimensions bias demographic inferences: implications for comparative plant demography. Am. Nat. 176, 710–722. http://dx.doi.org/10.1086/657044.
Sharpe, F.R., Lotka, A.J., 1911. A problem in age distribution. Phil. Mag. 21, 435–438.
Simon, H.A., Ando, A., 1961. Aggregation of variables in dynamic systems, Econometrica 29, 111–138. http://dx.doi.org/10.2307/1909285.
Steiner, U.K., Tuljapurkar, S., Coulson, T., 2014. Generation time, net reproductive rate, and growth in stage-age-structured populations. Am. Nat. 183, 771–783. http://dx.doi.org/10.1086/675894.
Stott, I., Townley, S., Hodgson, D.J., 2011. A framework for studying transient dynamics of population projection matrix models. Ecol. Lett. 14, 959–970. http://dx.doi.org/10.1111/j.1461-0248.2011.01659.x.
Tuljapurkar, S.D., 1982. Why use population entropy? it determines the rate of convergence. J. Math. Biol. 13, 325–337. http://dx.doi.org/10.1007/BF00276067.
Tuljapurkar, S., 1985. Population dynamics in variable environments. VI. Cyclical environments. Theor. Popul. Biol. 28, 1–17. http://dx.doi.org/10.1016/0040-5809(85)90019-X.
Tuljapurkar, S., 1986a. Demography in stochastic environments II. Growth and convergence rates. J. Math. Biol. 24, 569–581. http://dx.doi.org/10.1007/BF00275685.
Tuljapurkar, S., 1986b. Demography in stochastic environments I. Exact distributions of age structure. J. Math. Biol. 19, 335–350. http://dx.doi.org/10.1007/BF00277104.
Tuljapurkar, S., 1993. Entropy and convergence in dynamics and demography, J. Math. Biol. 31, 253–271. http://dx.doi.org/10.1007/BF00166145.
Tuljapurkar, S., Orzack, S.H., 1980. Population dynamics in variable environments I. Long-run growth rates and extinction. Theor. Popul. Biol. 18, 314–342. http://dx.doi.org/10.1016/0040-5809(80)90057-X.
Tuljapurkar, S., Steiner, U.K., 2010. Dynamic heterogeneity and life histories. Ann. New York Acad. Sci. 1204, 65–72. http://dx.doi.org/10.1111/j.1749-6632.2010.05519.x.
Usher, M.B., 1969. A matrix model for forest management. Biometrics 25, 309–315.
Vaupel, J.W., Manton, K.G., Stallard, E., 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography 16, 439–454. http://dx.doi.org/10.2307/2061224.
Van Imhoff, E., 1992. A general characterization of consistency algorithms in multidimensional demographic projection models. Popul. Stud. (NY). 46, 159–169. http://dx.doi.org/10.1080/0032472031000146066.
van Noordwijk, A.J., de Jong, G., 1986. Acquisition and allocation of resources: their influence on variation in life history tactics. Am. Nat. 128 (1), 137–142.
Vindenes, Y., 2010. Stochastic modeling of finite populations with individual heterogeneity in vital parameters.
Willekens, F.J., 1977. Sensitivity analysis in multiregional demographic models. Environ. Plan. A 9, 653–674. http://dx.doi.org/10.1068/a090653.
Yuster, R., Zwick, U., 2005. Fast sparse matrix multiplication. ACM Trans. Algorithms 1, 2–13. http://dx.doi.org/10.1145/1077464.1077466.

2.6 Appendices

2.6.1 Asymptotic properties of MPPMs

The asymptotic properties of the dynamics of a structured population are determined by the eigenanalysis of the associated projection matrix. Is the maximum eigenvalue real and positive? Unique? Can the associated right- and left-eigenvectors be made (strictly) positive? The Perron-Frobenius theorem (Frobenius, 1912; Perron, 1907) asserts that a real square matrix with positive entries has a unique largest real eigenvalue and that the corresponding eigenvector can be chosen to have strictly positive components. Every non-negative matrix – and MPPMs belong to that category – is the limit of a series of positive matrices. As such, it has an eigenvector with non-negative components which corresponding eigenvalue is non-negative and greater than or equal, in absolute value, to all other eigenvalues (see Gantmacher, 1959). However this property cannot establish its unicity.

To tackle the issue, Frobenius (1912) studied the special case of non-negative matrices that are irreducible: $\mathbf{B}$ is irreducible iff $\mathbf{B}$'s associated graph $\mathcal{G}_{\mathbf{B}}$ is strongly connected, i.e. every vertex is reachable from every other vertex. He proved that all maximum eigenvalues of $\mathbf{B}$ lie, in the complex plane, on the circle of radius λ; with λ itself a simple (and thus unique) real maximum eigenvalue with both right and left eigenvectors strictly positive (Frobenius, 1912).

Most MPPMs will however be reducible because the presence of several traits leads to the existence of (dynamically) unreachable, (biologically) irrelevant combinations of those traits. Non-negative (reducible) square matrix, $\mathbf{M}$ may be written in upper-triangular block form, known as the Frobenius normal form of a reducible matrix (Varga, 1962):

$$\mathbf{P}.\mathbf{M}.\mathbf{P}^{-1} = \begin{bmatrix} \mathbf{B_1} & * & * \\ \mathbf{0} & \dots & * \\ \mathbf{0} & \mathbf{0} & \mathbf{B_m} \end{bmatrix}$$

where $\mathbf{P}$ is a permutation matrix and each $\mathbf{B_i}$ is a square matrix that is either irreducible or null. The spectrum of $\mathbf{M}$ is the union of the spectra of the $\mathbf{B_i}$, which represent strongly connected components of $\mathcal{G}_{\mathbf{M}}$. Each non-zero $\mathbf{B_i}$, being non-negative irreducible, will have a simple and unique real maximum eigenvalue λ_i with (strictly) positive corresponding right- and left-eigenvectors $\boldsymbol{w_i}$ and $\boldsymbol{v_i}$.

These (real) subcomponents maximum eigenvalues can be ordered, and in most cases, only one of them, say λ_i , maximises all the others. This means $\lambda_j = \lambda = \max_i(\lambda_i)$ is unique, representing the fact that $\mathcal{G}_{\mathbf{B_j}}$ is the ergodic-growth-generating component of $\mathcal{G}_{\mathbf{M}}$. The associated eigenvectors $\boldsymbol{w}$ and $\boldsymbol{v}$ will then be non-negative: positive on subindices belonging to $\mathcal{G}_{\mathbf{B_j}}$ and non-negative elsewhere. Zeroes of $\boldsymbol{v}$ – to be found only outside of the subgraph $\mathcal{G}_{\mathbf{B_j}}$ generating λ – are of particular importance. Indeed λ, $\boldsymbol{w}$ and $\boldsymbol{v}$ (the ergodic eigentriad of $\mathbf{M}$) will be the asymptotic growth rate and associated eigenvectors of a population projected by $\mathbf{M}$ iff the initial population vector $\boldsymbol{x}_0$ is such that $\boldsymbol{v}'.\boldsymbol{x}_0 \neq 0$ (i.e. at least one non-zero state of $\boldsymbol{x}_0$ is part of the non-zero elements of $\boldsymbol{v}$, or in other terms, $\boldsymbol{x}_0$ has a strictly positive total reproductive value for eigentriad λ, $\boldsymbol{w}$ and $\boldsymbol{v}$). If a realistic initial population $\boldsymbol{x}_0$ (i.e. where trait values of individuals are not mutually exclusive from a (bio)logical perspective) does not pass this test, it means that the population asymptotic growth rate $\lambda_k < \lambda$ will be suboptimal, governed by a subgraph $\mathcal{G}_{\mathbf{B_k}}$ (with $(k \neq j)$. In this case $\mathbf{M}$ and $\boldsymbol{x}_0$ model a population that cannot deterministically reach the maximal potential growth rate that transitions allow. The MPPM analysis, described in the body of the article, would then have to be performed on the submatrix of $\mathbf{M}$ containing $\mathbf{B_k}$ and all states reachable from $\mathbf{B_k}$.

If the largest two (or more) real eigenvalues of $\mathbf{M}$ are equal, it means that several components of $\mathcal{G}_{\mathbf{M}}$ share the same ergodic growth rate, and have asymptotic importance for $\mathbf{M}$. The MPPM analysis will then have to be carried out on each of these components and their descendant states. In the article we consider the general case where $\mathbf{M}$'s largest and real eigenvalue λ is unique, and where the support of $\boldsymbol{v}$ contains our initial population. In that case, a simple eigenanalysis provides all needed ergodic demographic measures (ergodic growth rate and abundances, reproductive value, damping ratio...).

2.6.2 Demonstration of the general impossibility of perfectly merging 2 states in a directed state transition graph

In a directed state transition graph, e.g. the life-cycle of a population which dynamics is governed by a projection matrix, there is no absolute way of *perfectly* merging 2 states. We call *perfect*, the merging of several states into 1 that has no impact on the population dynamics. Let us consider a given population characterized by a life cycle graph of $n+1 (n \in \mathbb{N}^{+*})$ nodes (states), corresponding to the $n+1$ rows and columns of its associated state transition matrix $\mathbf{A}$ (representing the $\mathbb{R}^{n+1}$ to $\mathbb{R}^{n+1}$ linear mapping

projecting the population over one time-step). In this case, there is, in general, no *perfect* way of merging two states; i.e. there exist no linear mapping $\mathbf{B}$, from $\mathbb{R}^n$ to $\mathbb{R}^n$, that would allow to identically project this population now distributed on n states.

Indeed, we demonstrate here that, in order for the *perfect* merging of specific states to be possible, all contributions (transition rates) from the states to be merged have to be equal as well as net inner transitions (i.e. the net contributions towards the future group from each element of the group). Apart from this particular case, states cannot be *perfectly* merged:

Theorem:

$$\mathbf{A} \in M_{n+1,n+1}(\mathbb{R}^+), \exists\ \mathbf{B} \in M_{n,n}(R^+), \forall x \in \mathbb{R}^{+^{n+1}} \quad \mathbf{B}.x^{\sim} = (\mathbf{A}.x)^{\sim}$$
$$\Rightarrow \begin{cases} \forall i \in [1, n-1], A_{i,n} = A_{i,n+1} \\ A_{n,n} + A_{n+1,n} = A_{n,n+1} + A_{n+1,n+1} \end{cases}$$

where $x^{\sim} = (x_1, x_2, \ldots, x_{n-1}, (x_n + x_{n+1}))$

Proof :

$$\mathbf{B}.x^{\sim} = \begin{bmatrix} (\sum_{j=1}^{n-1} B_{1,j}x_j) + B_{1,n}(x_n + x_{n+1}) \\ (\sum_{j=1}^{n-1} B_{i,j}x_j) + B_{i,n}(x_n + x_{n+1}) \\ (\sum_{j=1}^{n-1} B_{n,j}x_j) + B_{n,n}(x_n + x_{n+1}) \end{bmatrix} \quad (\mathbf{A}.x)^{\sim} = \begin{bmatrix} \sum_{j=1}^{n+1} A_{1,j}x_j \\ \sum_{j=1}^{n+1} A_{i,j}x_j \\ \sum_{j=1}^{n+1} (A_{n,j} + A_{n+1,j})x_j \end{bmatrix}$$

$$\mathbf{B}.x^{\sim} = (\mathbf{A}.x)^{\sim} \Rightarrow \begin{cases} \forall i \in [1, n-1], (\sum_{j=1}^{n-1} B_{i,j}x_j) + B_{i,n}(x_n + x_{n+1}) = \sum_{j=1}^{n+1} A_{i,j}x_j \\ (\sum_{j=1}^{n-1} B_{n,j}x_j) + B_{n,n}(x_n + x_{n+1}) = \sum_{j=1}^{n+1} (A_{n,j} + A_{n+1,j})x_j \end{cases}$$
$$\Rightarrow \begin{cases} \forall i \in [1, n-1] \forall j \in [1, n-1] B_{i,j} = A_{i,j} \\ \forall i \in [1, n-1], B_{i,n} = A_{i,n} = A_{i,n+1} \\ \forall j \in [1, n-1], B_{n,j} = A_{n,j} + A_{n+1,j} \\ B_{n,n} = A_{n,n} + A_{n+1,n} = A_{n,n+1} + A_{n+1,n+1} \end{cases}$$
$$\Rightarrow \begin{cases} \forall i \in [1, n-1], A_{i,n} = A_{i,n+1} \\ A_{n,n} + A_{n+1,n} = A_{n,n+1} + A_{n+1,n+1} \end{cases}$$

2.6.3 Preservation of reproductive value in *ergodic flows preserving* (EFP)-merging

Let $\mathbf{A} \in M_{n,n}(\mathbb{R}^+)$ be a population projection matrix, and $\boldsymbol{n}(t)$ an associated population vector. We regroup the first s states of $\mathbf{A}$ (out of n) into 1 state called g. The post-merging projection matrix $\mathbf{A}^*$ will then have $n - s + 1$ states labeled $g, s+1, s+2, \ldots, n$.

The elements of $\mathbf{A}$ represent transition rates: $A_{i,j} = t_{j \to i}$ is the transition rate between state j and state i. The population flow at time t between state $i$and state j is $n_i(t).t_{i \to j}$. We assume $\mathbf{A}$ has a unique maximal real eigenvalue λ with associated right eigenvector $\boldsymbol{w}$, scaled so $\sum_i w_i = 1$ and left eigenvector$\boldsymbol{v}$, scaled so $\sum_i v_i w_i = 1$ Then, when time tends to infinity, the relative ergodic flow from state i to state j (relative to all other flows occurring at that time step) is $w_i.t_{i \to j}$

These relative ergodic flows are unchanged through EFP-merging. This means that every state j outside the group retain the same w_j and that, within the group, $w_g = \sum_{i=1}^{s} w_i$. Transitionwise, this can be written as: $t_{j \to g} = \sum_{i=1}^{s} t_{j \to i}$ and $w_g.t_{g \to j} = \sum_{i=1}^{s} w_i.t_{i \to j}$ i.e. $t_{g \to j} = \sum_{i=1}^{s} \frac{w_i}{w_g} t_{i \to j}$. Hence, the EFP-merging of several states in a directed state transition graph consists in summing all transitions towards the future group and to weight transitions out of the group by the group's states relative ergodic abundances. By definition EFP-merging preserves ergodic flows and by construction, as ergodic flows are transitions-weighted ergodic abundances, and relative ergodic abundances.

Theorem : The EFP-merging of several states preserves the relative reproductive values of states iff contributions from all states (or future states) are equally broken down with regards to the soon-to-be-grouped states (all incoming flows towards the future group coming from a single (future) state being a particular case). The group's class reproductive value $c_g = w_g.v_g$ as (see Taylor et al., 2007; Taylor, 1990) is then the sum of the class reproductive value of its constituents.

Proof:

Downstream from the future group, we have, by definition of a reproductive value, $\forall i \leq s$, $\lambda.v_i = \sum_{j=1}^{n} t_{i \to j}.v_j$. If reproductive values are preserved by grouping we have also $\lambda.v_g = \sum_{j=1}^{n} t_{g \to j}.v_j$. Then $\lambda.v_g = \sum_{j=1}^{n} (\sum_{i=1}^{s} w_i/w_g.t_{i \to j}).v_j$ i.e.$\lambda.v_g.w_g = \sum_{i=1}^{s} w_i.(\sum_{j=1}^{n} t_{i \to j}.v_j) = \lambda.\sum_{i=1}^{s} w_i.v_i$. Downstream from the grouping, reproductive values preservation through an *EFP* -grouping implies that $c_g = v_g.w_g = \sum_{i=1}^{s} w_i.v_i = \sum_{i=1}^{s} c_i$.

Upstream from the grouping, we have$\forall i > s$, $\lambda.v_i = \sum_{j=1}^{n} t_{i\to j}.v_j = \sum_{j=1}^{s} t_{i\to j}.v_j + \sum_{j=s+1}^{n} t_{i\to j}.v_j$and after grouping, if reproductive values are preserved $\lambda.v_i = \sum t_{i\to j}.v_j = t_{i\to g}.v_g + \sum_{j=s+1}^{n} t_{i\to j}.v_j$ implying : $t_{i\to g}.v_g = \sum_{j=1}^{s} t_{i\to g}.w_j/w_g.v_j = \sum_{j=1}^{s} t_{i\to j}.v_j$ and thus $\forall i > s, \sum_{j=1}^{s} .w_j/w_g.v_j = \sum_{j=1}^{s} t_{i\to j}/t_{i\to g}.v_j$. Upstream from the grouping, reproductive values preservation through an *EFP*-merging implies then that the submatrix of all transitions towards the future group has rank one and therefore that all states have the same relative distribution of transitions towards the states of the future group.

In conclusion, when regrouping states so that ergodic flows and abundances are preserved through the grouping process, relative reproductive values of states will be not be preserved , unless all states having a transition to any member of the future group, allocate those contributions the same way (and then $v_g = \sum_{j=1}^{s} \frac{w_j}{w_g}.v_j$).

Chapter 3

The demographic and evolutionary consequences of physiological costs of reproduction

Contents

3.1 Introduction

In the first chapter of this thesis, we have developed a theoretical design of the costs of reproduction (see section 1.2 page 11) in which we have analyzed the different components of these costs and devised a method to model these in a unified framework (see 1.3.2 page 30). Among the most differentiating characteristics within the costs of reproduction lies the utmost segregation between *physiological* and *genetic costs of reproduction*. This dividing line is not specific to the costs of reproduction but can be found in each and every trade-off and is pervasive in Life History Theory (Braendle et al., 2011).

Physiological trade-offs (sometimes called mechanistic) and genetic trade-offs (also known as evolutionary) differ in many aspects. In the Stearnsian triptych – decomposing the architecture of all trade-offs in a genotypic level, an intermediate structure and a phenotypic level (Stearns, 1989b)– genetic trade-offs are part of the first component, whilst physiological ones lie in the second. Physiological trade-offs consist in a mechanism at work within every individual at every time-step, whilst Genetic trade-offs, being "evolutionary" work at the level of the population over the individuals' entire life trajectories.

Chiefs among life history trade-offs are the costs of reproduction as their trade-off function directly connects the primordial fitness components that are fertility and survival rates (Lessells, 1991). In the framework of these costs, we have shown in Chapter 1 that the differences in mechanisms between *physiological* and *genetic costs of reproduction* cause differences in behaviour and detectability. Individual stochasticity, for instance, fuels detectability of physiological costs whilst concealing genetic costs. More importantly, we have shown that these costs correspond to different time windows of effects : physiological costs buffer environmental changes and individual variance at the level of the season up to the entire lifetime of the organism. Genetic costs on their side – that is, the genetic variance in allocation strategy or in general, in life-history fertility schedule as evolved by the organism – are a picture of the recent evolutionary past and a buffer against environmental shifts over the near evolutionary future.

However, with all they differences in structure and effects, physiological and genetic trade-offs can, theoretically, co-habit. We have exhibited this for costs of reproduction by the construction of *physiological costs of reproduction with genetic basis* which contain both the intermediate structure of physiological costs and the genotypic level of genetic costs. In senescence theory, the physiological/genetic costs divide is mirrored exactly by the disposable soma/antagonistic pleiotropy theories split (see section 1.4.2 page 36). And here as well, senescence theoreticians have shown that both theories can intersect (Gavrilov and Gavrilova, 2002) but not be equated (Kirkwood and Rose, 1991).

As of today, it is still unclear whether genetic and physiological trade-offs are different manifestations of a single underlying mechanism, two sides of the same coins acting at different levels, or even completely different, even antagonistic, phenomena. A comprehensive theory encompassing genetic and physiological trade-offs and disentangling their respective role is necessary. As Braendle et al. (2011) put it "... it remains to be determined to what extent presumptive trade-offs are conclusively due to actual competition for limited resources or caused by alternative mechanisms, such as hormonal signaling independent of resource allocation [...]. The very limited knowledge on the mechanistic underpinnings of trade-offs therefore represents a current key problem in our understanding of life history evolution".

In order to advance towards answers to this central question to life history, one needs to have at hand both a suitable model framework and adapted analysis tools that could tackle the complexity of trade-offs. As mentioned in chapter 1 section 1.3.1 page 29, *physiological costs* are naturally suited to a family of models called Individual Based Models, whilst population projection matrices are the model of choice for the study of genetic costs. The multitrait population projection matrix (MPPM) framework developed in chapter 2 is such a method that allows to incorporate, via the addition of physiological and evolutionary traits to a standard matrix model, all aspects of the costs mentioned so far. Most importantly, it comes supplied together with the Trait Level Analysis tool that we have developed in Chapter 2, which - by folding the matrix over any subset of its traits - provides critical information on the evolutionary weight of the traits, and the trade-offs that connect them.

The most important strength of the MPPM building technology is that it allows to increase scrutiny enough so as to be able to incorporate processes that are almost individual based, like physiological trade-offs, whilst still retaining the ergodic analysis capabilities of all matrix models (see section 1.3.2 page 30). It does so by adding *dynamic heterogeneity* traits (a family of traits we called $\mathcal{D}$ in chapter 1) , to the basic life history traits $\mathcal{B}$ containing the characteristic best determining vital rates of the organism (like age, stage, size, etc.). Such *dynamic heterogeneity* traits may be resource capitals (the Ratchet and Fluctuating Capitals of chapter 1 section 1.2.1 for instance), gauges tracking the accumulation of the reproductive efforts required to produce one independent offspring, or any other tracker of individual life trajectory that is affected by the studied trade-off. By allowing the addition of $\mathcal{D}$ to $\mathcal{B}$, MPPMs allow to implement

physiological trade-offs in an evolutionary model framework.

This framework can be enhanced by the addition of *hidden heterogeneity* traits - regrouped in family $\mathcal{G}$ – like genotypes or any heritable trait that is "fixed-at-birth". Such an addition allows, for instance, to embed polymorphism in a gene that is antagonistically pleiotropic with respect to the fitness components linked by the studied trade-off. In the case of costs of reproduction, such a feature enables to implement variance in genes driving the allocation towards reproduction; i.e. to implement genetic costs. However, many other polymorphisms can be considered, and in particular in gene conferring, contrary to genetic trade-offs, fitness advantages to the bearers of certain alleles. Implementing a polymorphism in a gene driving acquisition of resources would generate such a fitness gradient in the population.

The ability to add several traits (potentially many and with large distributions thanks to the use of sparse matrices and vectorization methods (see chapter 2 (p.41))) in a matrix model does not, in itself, provide very useful evolutionary information. It certainly allows segmentation of ergodic abundances and a refined calculation of resilience (adding traits in a model reduces damping ratio; see discussion of chapter 2, sec.2.5, p.51), but it does not provide any information with respect to the evolutionary importance of the different traits and trade-offs implemented. This feature is made possible by the *Trait Level Analysis* described in sec.2.3.2 (p.47). By *folding* an MPPM along the traits implementing a physiological trade-off, one can measure its evolutionary implications by comparing the asymptotically equivalent pre-*folding* and post-*folding* matrices. By *folding* an MPPM along genotypic traits, hard to measure and thus rarely incorporated in matrix models, one can quantify the variation in demographic response such an addition entails. This therefore allows to quantify the price to pay, in terms of understanding the life-history of an organism, when forsaking genetic trade-offs (and genotypic polymorphism in general).

In most evolutionary mathematical models trade-offs are generally not implemented. This is quite astonishing considering that these are at the heart of life history theory. Deemed non-evolutionary since their window of action limits itself to lifetime of individuals, *physiological* trade-offs are altogether absent from evolutionary models. They are also mostly inexistent in evolutionary theories with a notable exception of Kirkwood and Holliday (1979)'s Disposable Soma Theory of senescence. *Genetic* trade-offs, for their part are not implemented in (as an input of) evolutionary models either (apart from non-generation overlapping population genetics models). To the contrary, they are actually mostly *derived from* such matrices via the Optimality Theory (see section 1.3.1 page 30): from the assumption that the populations modeled are at ESS (i.e. in Lande (1982)'s equation $0 = \mathbf{G}\nabla\boldsymbol{y}$), the genetic constrains between vital rates (in $\mathbf{G}$) stem directly from the selection gradient $\nabla\boldsymbol{y}$ corresponding to the vector of growth rates sensitivities $\frac{\partial\lambda}{\partial M_{i,j}}$. This approach can highlight genetic trade-offs and their mechanisms, from empirical data, when in long-term constant environments. It has however the drawback of defining the constraints as functions of vital rates (the entries of a projection matrix). Vital rates are the primary drivers of demography with other parameters affecting these deemed secondary. The dependency is reverse with respects to trade-offs. Even if a trade-off relate specific traits, its effect on population dynamics will be implemented via the effects of each of these traits on the vital rates. This makes the trade-off hard to read at the level of the matrix. As a matter of fact, considering a population modeled by a $q \times q$ matrix is at ESS, such an optimality analysis will generate q^2 pairwise genetic constrains, many of which are suspected of being by-products (of demographic and other genetic relationships) of little life-history and evolutionary significance.

Moreover and most importantly, the study of trade-offs, both physiological and genetic, is about environmental changes that may be buffered by the costs with different time scales and that invites genetic variance in the population. Understanding the role of genetic costs, and in particular their cross-effects with physiological costs, prompts us *not* to use a constant environment (ESS) optimality theory method to analyze genetic trade-offs. They thus need to be incorporated as traits in the matrix.

In this chapter, we wish to model physiological costs of reproduction in an evolutionary framework, and to be able to incorporate genetic costs as well. To do this, we start by relating the concepts of chapter 1, with the method of chapter 2, in order to design, in an MPPM framework, a model implementing *physiological costs of reproduction*. This model will actually be a family of models, that can incorporate varying life history strategies. We show how to implement the cohabitation of such varying strategies in a population with physiological costs, via a three trait *age-parity-heterogeneity*-MPPM. We then extend for such models, the classical calculation methods for selection gradient and for demographic and environmental variances, and the, not so classical, method to compute the variance of lifetime reproductive success. We then use these methods and the *Trait Level Analysis* to compute the deterministic and stochastic effects of costs of reproduction on (i) the shape of fertility and survival curves by age (ii) the fertility selection gradient (iii) the variance in lifetime reproductive output (and, consequently, the effective size of the population) and finally, (iv) the environmental variance and, in general, the stochastic growth rate.

3.2 Model for *physiological costs of reproduction with genetic basis*

The model we build in this chapter aims not to apply to any particular kind of organism, but should relate to most (aspects of) costs of reproduction for most species. For simplicity, we reduce the traits embedded in the model to one for each of the three families of traits ($\mathcal{B}$,$\mathcal{G}$ and $\mathcal{D}$) described in chapter 1 (sec.1.3.2, p.30) and that are required to implement both physiological and genetic trade-offs:

In the basic family of traits that best implement life-history, $\mathcal{B}$, age is the obvious choice. It is indeed the only trait common to all matrix models (explicitly or implicitly with infinite distribution).

Via the *hidden heterogeneity* family of traits, $\mathcal{G}$, we will allow polymorphism in life history strategy in the population; more precisely polymorphism on the position on the Slow-Fast Continuum with some genotypes with higher fertility and lower survival than others. Thus the different genotypes will mainly lie along an iso-fitness line (see figure 1.2 page 21). It is however also possible to incorporate fitter (respectively frailer) genotypes which have both higher (resp. lower) fertility and higher (resp. lower) survival.

Finally, in the *dynamic heterogeneity* family of traits, $\mathcal{D}$, we will also consider one unique trait. For simplicity, we do not consider storage facilities (i.e., $stor = 0$, in the context of chapter 1) and thus relate our trait to the Ratchet capital. Whilst Fluctuating Capital is not implemented as a trait *per se*, the (intra-periodic) effects of the environment on FC can still be embedded via the effects of the environment e on vital rates. In that case, instead of one MPPM $\mathbf{M}$ modeling the population in a constant environment, the model will consist in a family of matrices, each defined for a specific environment $\mathbf{e} : \{\mathbf{M_e}\}$.

We shall further simplify the analysis by equating, in this Chapter, reproductive effort and fertility (in the context of chapter 1 section 1.2.1 page 12, this means that $res = \delta_0$ which implies $re = f$ from eq.1.1 page 12). Since demographic variance is an important driver of costs of reproduction, which are therefore proportional to the organism reproductive granularity (See chapter 1 section 1.2.1 page 18), we shall consider that the organism can produce at maximum 1 offspring per period (thus at the quality end of the quality quantity spectrum, i.e., in the framework of chapter 1, $gr \approx 1$). And therefore, we shall implement the Ratchet Capital by tracking the parity (number of offspring ever born) of individuals as the $\mathcal{D}$ trait. It starts at 0 and is maximized by the length of the organism's reproductive period. Therefore trait *parity* corresponds to (the opposite of) a RC starting at the species' evolved duration of the reproductive period reduced by a realized reproductive effort ($RE(t)$ in Chapter 1) of 1 at each successful fertility period.

The particular effect of the lineage, i.e. of the genotype, on RC is embedded by assigning different values to the zero-parity fertility and survival rates for the different genotypes. In general, in order to help disentangling the roles of genetic and physiological costs, we shall consider that the former, together with the environment, only affect these zero-parity vital rates, whilst the latter consist in the mechanical reduction of vital rates with parity, from the zero-parity vital rate, at birth, to 0, when parity is maximum (at the reproductive period length). This physiological trade-off mechanism is thus the same for all individuals which will only differ in vital rates depending on their parity and their genotype (as a consequence, in this simple model, an individual which does not reproduce does not senesce). Indeed, as highlighted in chapter 1 section 1.2.1 page 16, we expect, for a given genotype, the capitals (i.e; parity here) to be the main drivers of the allocation process. In line with this consideration, we shall mostly consider, in this chapter, that *age* - albeit necessary as a basic trait to account for the "arrow of time" - has no direct effect on allocation and therefore on vital rates. These then only depend on *parity* and *heterogeneity*.

3.2.1 The three traits of (*age,parity,heterogeneity*)-MPPM

Trait *age* models life-history

In our one-sex model (with a time-step we consider to be one year without any loss of generality) we use a simple age structure with a maximum age of ω, an age at first reproduction of α and an age at last reproduction of β . As discussed before, we simplify our model by considering no storage capacity and therefore primarily focus our study on RC physiological costs of reproduction (related to the allocation of resources that cannot be acquired, like time), with the effect of the environment (on FC costs) possibly implemented in the general environmentally dependent model $\{\mathbf{M_e}\}$ model by the differences between environments e of the vital rates in the various $\mathbf{M_e}$. We also focus on organisms with high reproductive granularity, that can only produce a few offspring per breeding season, and simplify our model by considering an organism that produces, in its reproductive years, at most one offspring. Therefore, in the model of this chapter,

fertility *and* survival rates are between 0 and 1.

We will consider both the general case where fertility and survival rates at different ages a - f_a and s_a - can vary. More often that not, however and as discussed above, we shall consider that age does not directly drive vital rates (parity, genotypes and environment do), in which case the zero-parity vital rates are simply denoted f and s.

In the latter case, for $\omega = \beta = 3$ and $\alpha = 2$, this simple *age*-structured model can be represented by its Leslie matrix equivalent : $\mathbf{L} = \begin{bmatrix} 0 & f & f \\ s & 0 & 0 \\ 0 & s & 0 \end{bmatrix}$.

Trait *parity* models *physiological costs of reproduction*

In our model, we equate reproductive effort with the effort it takes to produce recruits; and thus the cumulative reproductive effort is commensurable with parity, the number of offspring ever born to an individual. Because, for a given environment, a given genotype, and a given parity, the fertility rate f corresponds to the expectation of the random variable $\mathcal{F}$ "having one offspring" at that time-step, each reproductive age now constitutes a fork in the life cycle where, in expectation, a portion f of individuals will move on to a higher parity class, and $1 - \mathrm{f}$ will remain in the same parity class.

We implement the costs themselves, i.e. the negative effect on fitness of having allocated towards reproductive efforts, by indexing vital rates on parity. To remain general, we consider that fertility *and* survival rates are affected by the costs (see chap.I sec.1.2.1 p.18), thus modeling both survival and fertility costs of reproduction as chategorized by Bell (1984). For simplicity, the indexation is linear and both vital rates become 0 when parity has reached the length of the species' reproductive period $\beta - \alpha + 1$.

In practice, this means that, for a given genotype for which the zero-parity fertility and survival rates at age a are f_a and s_a the fertility and survival rates at parity p are $f_{a,p} = f_a.(1 - \frac{p}{(\beta-\alpha+1)})$ and $s_{a,p} = s_a.(1 - \frac{p}{(\beta-\alpha+1)})$.

In order to build an (*age-parity*)-MPPM (the MPPM construction for the 3-trait *age-parity-heterogeneity*)-model is detailed in section 3.2.2), it is required to generate for each state, i.e. for each pair of trait values (a, p) *first* the vital rates (the survival and fertility rates) which we just computed, $f_{a,p}$ and $s_{a,p}$ and *second* the output states and their associated probabilities. In a genotypically uniform framework, one state will be the output state of any fertility transition: $(1, 0)$ (as every newborn is aged 1 and has parity 0). Survival transitions, on the contrary, will have 2 output states: $(a + 1, p + 1)$ with probability $f_{a,p}$ if the fertility event is successful and $(a + 1, p)$ with probability $1 - f_{a,p}$ otherwise. The resulting MPPM, in the case where fertility and survival are independent of age with zero-parity vital rates f and s , in an *age* structure where $\omega = \beta = 3$ and $\alpha = 2$ (that is, for an age-structured population with 3 age classes and maturity at 2 years of age) is then the following matrix:

$$\mathbf{C} = \begin{bmatrix} 0 & f & f & 0 & 0 & f/2 \\ s & 0 & 0 & 0 & 0 & 0 \\ 0 & s.(1-f) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & s.f & 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.1}$$

Trait *heterogeneity* models *genetic costs of reproduction*

In chapter 1, we introduced a setting where two genes act on the *physiological costs* of reproduction: 1/ an acquisition gene acting on the overall amount of resources the organism can gather and 2/ an allocation gene acting on the portion of acquired resources allocated to reproductive effort (see figure 1.2 page 21).

With regards to an MPPM with *heterogeneity*, such genes and their effects would be implemented by different categories in the *heterogeneity* trait. In an (*age, parity, heterogeneity*)-MPPM where the *physiological costs of reproduction* are mechanistically drawn from parity, these different genotypes are materialized by differing zero-parity Leslie matrices for the varying *heterogeneity* values. This way, the mechanism modeling the physiological costs is unchanged (the relationship between vital rates and their related zero-parity rates) but the genotypes differ via differing zero-*parity* vital rates between different genotypes (*heterogeneity* categories).

However, as we saw in chapter 1, section 1.2.2 (page 20) and as represented on figure 1.3, *genetic costs of reproduction* can be decomposed into two non-overlapping categories. First *physiological costs of repro-*

duction with genetic basis, where the genetic costs lie in the variance in the allocation and acquisition genes discussed above (sec.1.2.2 p.20). Second *genetic non-allocative costs* for which the genetic costs consist in variances in genes we call *slow-fast* and *robustness* with the same overall effects than respectively *allocation* and *acquisition* but where the action does not require any physiological allocative process (sec.1.2.2 p.22). Simply put, the *allocation* gene acts on the allocation towards fertility, which via the underlying allocative physiological costs, implies a cost on survival, whereas the *slow-fast* gene has the same effects, but because it has antagonistic effects on the fertility and survival "functions". In other words, the *allocation gene* is a *slow-fast* gene if the existence of *physiological costs of reproduction* is ignored. Similarly the action of the *acquisition* gene necessitates the presence of the acquisition-allocation mechanism of the physiological costs, on which it acts. To the contrary, the *robustness* gene, just segregates genotypes by fitness, without the necessity of referring to an acquisition process.

***Allocation* gene heterogeneity** Genotypes that differ only for the *allocation* gene correspond to lineages, of similar fitness, which have developed over time different strategies of allocation towards reproductive efforts. This categorization is only valid in a context of physiological costs. At both ends of a simplified spectrum would we find 1/genotypes allocating consistently more towards reproduction, at the cost of long-term fitness and 2/genotypes preserving the capital, with low fertility, and thus improved late fitness. Those genotypes may be be able to coexist in the population under the current environment as they have equivalent fitness (see fig.1.2, p.21 of chapter 1). Here are the (*age,parity*)-MPPM of 2 such genotypes for a 2-year organism; organism $all1$ has $f = 0.9$ and $s = 0.3$; organism $all2$ has $f = 0,7$ and $s = 0.8$:
$\mathbf{M_{all1}} = \begin{bmatrix} 0.9 & 0.9 & 0 & 0.45 \\ 0.4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.4 & 0 & 0 & 0 \end{bmatrix}$ $\mathbf{M_{all2}} = \begin{bmatrix} 0.7 & 0.7 & 0 & 0.35 \\ 0.24 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.56 & 0 & 0 & 0 \end{bmatrix}$, yielding $\lambda_{all1} = \lambda_{all2} = 1.04$.

***Acquisition* gene heterogeneity** In our model (in which *stor* is zero and where the capital, tracked by trait *parity*, corresponds to RC), both the environment and the capacity to acquire FC resource from it are still primordial. Indeed the combination of both affects the reproductive effort via FC (see eq.1.3 p.13 and eq.1.5 p.16 of chapter 1) and the reproductive effort produced, in turn, affects RC, i.e. *parity* in this model (eq.1.2 p.13 of chapter 1). In this context therefore (requiring the presence of physiological costs) genotypes that differ only for the *acquisition* gene would differ in fitness : some lineages would just have higher vital rates than others (fig 1.2 p. 21). Such genotypes would only be allowed to cohabit in the current environment if it is encountered by the population over short periods of time interrupted by environments in which the reaction norms have crossed (see discussion in chap. I sec.1.2.2 p.20) Here are the (*age,parity*)-MPPM of 2 such genotypes; organism $acq1$ has $f = 0.9$ and $s = 0.8$; organism $acq2$ has $f = 0.7$ and $s = 0.3$:
$\mathbf{M_{acq1}} = \begin{bmatrix} 0.9 & 0.9 & 0 & 0.45 \\ 0.08 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.72 & 0 & 0 & 0 \end{bmatrix}$ $\mathbf{M_{acq2}} = \begin{bmatrix} 0.7 & 0.7 & 0 & 0.35 \\ 0.09 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.21 & 0 & 0 & 0 \end{bmatrix}$, yielding $\lambda_{acq1} = 1.22, \lambda_{acq2} = 0.86$.

***Allocation* gene heterogeneity is particular case of *slow-fast* gene** Without the context of physiological costs, genotypes that differ only for the *slow-fast* gene correspond to lineages, of similar fitness, which have developed over time different strategies along the slow-fast continuum. At both ends of a simplified spectrum would we find 1/genotypes with higher early fertility rates at the cost of later fitness and 2/genotypes where early fertility is low but late fitness is improved. The *hidden heterogeneity* caused by the genetic polymorphism of the *allocation* gene is due to antagonistic pleiotropy : different alleles of the gene have opposing effects on early fertility and late fitness. Such a variance is however just a particular case of *genetic costs of reproduction* (fig. 1.3 p.23). The pleiotropy does not actually need *physiological costs of reproduction* to have effect, and we can conceive of non-allocative *genetic costs of reproduction* such as variance in a *slow-fast* gene. To illustrate this, here are the models for genotypes corresponding to allocation genotypes $all1$ and $all2$, but in a context with no *physiological costs of reproduction* they are now called $sf1$ and $sf2$ and correspond to variance in the (non-allocative but pleiotropic) *slow-fast* gene. They are thus structured by age only and obtained by *folding* over *parity* M_{all1} and M_{all2} over *parity* (see section 3.2.2 and chapter 2 (p.41)) :
$\mathbf{M_{sf1}} = \begin{bmatrix} 0.9 & 0.495 \\ 0.3 & 0 \end{bmatrix}$ $\mathbf{M_{sf2}} = \begin{bmatrix} 0.7 & 0.455 \\ 0.8 & 0 \end{bmatrix}$, yielding $\lambda_{sf1} = \lambda_{sf2} = 1.04$.
In this case, vital rates change with age as we want to show the relationship between *physiological costs of reproduction with genetic basis* and genetic non-allocative costs. In general however it is totally possible to consider *slow-fast* genotypes with vital rates independent from age, as we shall see later in this chapter.

Such a framework of (*age-heterogeneity*)-MPPMs, where different genotypes are iso-fitness and only differ in life history strategies, provides an illustration of the theoretical conclusions from chapter 1 with regards to detectability of genetic costs (section 1.2.3 p. 22). In a constant environment, for such a population with two genotypes (represented by $\mathbf{M_{sf1}}$ and $\mathbf{M_{sf2}}$) the costs of reproduction would not be observable at the level of the individual trajectories: there is no dynamic tracking trait on which individual trajectories can diversify. However, at the level of the population (in a constant environment still and

because there is no *robustness* polymorphism), negative correlations would emerge as $sf1$ individuals that are the most fertile at age 1 are less likely to have offspring at age 2 than individuals $sf2$.

***Acquisition* gene heterogeneity is a particular case of *robustness* gene heterogeneity**
The hidden heterogeneity caused by the genetic polymorphism of the *acquisition* gene, can, when forsaking underlying physiological processes, be called variance in *robustness* or frailty. With variance in *robustness*, some lineages are just overall fitter (in that environment) than others. To illustrate this, here are the models for genotypes, stemming by *folding* over *parity* from $acq1$ and $acq2$, but in a context with no *physiological costs of reproduction*, they are*robustness* $rob1$ and $rob2$:
$\mathbf{M_{rob1}} = \begin{bmatrix} 0.9 & 0.495 \\ 0.8 & 0 \end{bmatrix}$ $\mathbf{M_{rob2}} = \begin{bmatrix} 0.7 & 0.455 \\ 0.3 & 0 \end{bmatrix}$ yielding $\lambda_{rob1} = 1.22, \lambda_{rob2} = 0.86$.

In this short study, we display simple (*age-parity*)-models for variance in "orthogonal" genes we call *allocation* and *acquisition* in the context of physiological costs of reproduction. We also display the asymptotic-equivalent (*age*)-models for variance in "orthogonal" genes *slow-fast* and *robustness* in a context with no physiological costs. By doing so, and thanks to universality of the physiological trade-off (the general form of the equation is the same for all genotypes), gene *allocation* can be considered as a particular case of *slow-fast* genes, and gene *acquisition* as a particular case of *robustness* gene, both conceptually and computationally.

3.2.2 Construction and analysis of deterministic (*age,parity,heterogeneity*)-MPPM

MPPM construction

In chapter 2 (p.41), we describe the steps to follow in order to build an MPPM. The current model is a simple use of the tool as it implements 3 traits only. It is however important we replace this model in the framework of MPPMs in order to later use the analysis tools of sensitivity analysis and *trait level analysis*.

Trait structure. To summarize and agglomerate the above sections, this model requires age as a trait. This trait can take ω values. We implement *physiological costs of reproduction*, via a negative linear relationship, at any given age, between vital rates and trait *parity*. *Parity* needs to be added as a trait; it can take $par = \beta - \alpha + 1$ values. We also want to be able to implement *genetic costs of reproduction* i.e. a polymorphism for a gene that antagonistically affects early and late vital rates, and variance in *robustness* , i.e. a polymorphism for a gene that affects all vital rates in parallel. These different genotypes need to be added as a *heterogeneity* trait. In most cases, for simplicity, when accounting for heterogeneity in the population, we will implement $het = 2$ genotypes that can implement either *genetic costs of reproduction* (the 2 genotypes have similar fitness but different allocation strategies), robustness heterogeneity (all rates, and thus fitness, higher for one genotype) or a combination of both. Using the notations of chapter 2 (p.41)the n = 3 traits of our model are $\mathbf{t} = (age, parity, heterogeneity)$ and the trait structure (the n-tuple of trait sizes) is $\mathbf{s} = (\omega, par, het)$. The number of states is then $q = \omega.par.het$ and the MPPM $\mathbf{M}$ will be of size q × q.

Building-block vectors. The vital rates for each state $0 \leq i \leq q$, than can also be identified by its equivalent triplet of states $i \longleftrightarrow (a, p, h)$, are stored in vectors $\mathbf{vr}^v$ where v represents the vital process: fertility or survival. When given, for each genotype h, the zero-parity vital rates i.e. $\forall v \in \{fertility, survival\}, \forall\, 1 \leq h \leq \text{het}, \forall\, 1 \leq a \leq \omega,\ \text{vr}^v_{(a,0,h)}$ we can establish the vectors of all vital rates :

$$\text{vr}^v_{(a,p,h)} = \text{vr}^v_{(a,0,h)}.(1 - \frac{p}{par})$$

From there, we can proceed to implement vectors $\mathbf{oi}^v_i$ and $\mathbf{op}^v_i$ representing the vector of all possible output states and related distribution for vital process v from state i. In this model, via fertility, an individual in state $i \leftrightarrow (a, p, h)$ will generate offspring of age 1 and parity 0, that will be of the same genotype that the parent genotype if no mutation occurs. If mutation rate per generation is m, then

$$\mathbf{oi}^{fertility}_{i\leftrightarrow(a,p,h)} = \begin{bmatrix} (1,0,1) \\ (1,0,2) \end{bmatrix} \quad , \quad \mathbf{op}^{fertility}_{i\leftrightarrow(a,p,1)} = \begin{bmatrix} 1-m \\ m \end{bmatrix} \quad \text{and} \quad \mathbf{op}^{fertility}_{i\leftrightarrow(a,p,2)} = \begin{bmatrix} m \\ 1-m \end{bmatrix} \tag{3.2}$$

Via survival, an individual in state $i \leftrightarrow (a, p, h)$ will survive as itself, one time-step older.*Hidden heterogeneity* h will be, by definition, unchanged, and *parity* p either increased by 1 or unchanged depending on reproductive success. Thus: $\mathbf{oi}^{survival}_{i\leftrightarrow(a,p,h)} = \begin{bmatrix} (a+1,p,h) \\ (a+1,p+1,h) \end{bmatrix}$. The expectations of realization of the random fertility process being the fertility rates, we have

$$\mathbf{op}^{survival}_i = \begin{bmatrix} 1-\text{vr}^{fertility}_i \\ \text{vr}^{fertility}_i \end{bmatrix} \tag{3.3}$$

The emergence of a vital process in the output states distributions $\{\mathbf{op}^v, \mathbf{op}^v\}$, a general consequence of the implementation of physiological trade-offs via the addition of active (i.e. that affects vital rates) *dynamic heterogeneity* traits like *parity* in our (*age-parity-heterogeneity*)-MPPM, increases the complexity of computations as soon as one departs the simple deterministic framework (see section 3.3.2 for the computation of variance of reproductive success in such a model). More importantly, it raises questions with respect to the interpretation and analysis of *folded* matrices (see following section 3.2.2) that are structured with such traits (see appendix 3.0.4 for discussion).

The matrix. Vectors $\mathbf{vr}^v$, $\mathbf{op}_i^v$ and $\mathbf{oi}_i^v$ are the building blocks of any MPPM construction. From these, we can construct $\mathbf{T}$ the matrix of all transition triplets, which is the sparse definition of our MPPM $\mathbf{M}$ (see chapter 2 (p.41)). And thus, in $\mathbf{M}$, we have a model projecting over time, a population, where individuals are characterized but their *age*, *parity* and *genotype*, that implements *physiological costs of reproduction* in a context of heterogeneity, i.e. where genotypes characteristics of *genetic cost of reproduction* and/or variance in robustness can cohabit.

Folding and Trait Level Analysis

The Trail Level Analysis evoked in chapter 1 and formalized in chapter 2 (p.41), is an MPPM tool, setting an evolutionary-neutral framework, that allows to understand the evolutionary demographic importance of traits and underlying trade-offs. It is performed by comparing properties of the MPPM build with all traits, $\mathbf{M}$, with *folded* versions of $\mathbf{M}$. *Folding* an MPPM over a defined subset of its traits consists in Ergodic-Flow-Preserving of the transitions of $\mathbf{M}$, for states which share the same values for traits not *folded* upon.
Matrix $\mathbf{M_{a,p}} = \mathbf{M^{fold}_{age,parity}}$ for instance is $\mathbf{M}$ folded over *heterogeneity*. Matrix $\mathbf{M_{a,h}} = \mathbf{M^{fold}_{age,heterogeneity}}$ is $\mathbf{M}$ folded over *parity* and $\mathbf{M_a} = \mathbf{M^{fold}_{age}}$ is $\mathbf{M}$ folded over both *parity* and *heterogeneity*. As age is its only trait, $\mathbf{M_a}$ is a Leslie matrix we call the reference Leslie matrix.

By construction all these matrices share the same ergodic growth rate λ and abundances vector $\boldsymbol{w}$, the right eigen-vector of $\mathbf{M}$ associated with λ. Ergodic abundances are denoted $\boldsymbol{w_{a,h}}$, $\boldsymbol{w_{a,p}}$ and $\boldsymbol{w_a}$ when regrouped on certain traits as right-eigenvector of matrices $\mathbf{M_{a,h}}$,$\mathbf{M_{a,p}}$ and $\mathbf{M_a}$.

When traits "disappear" by *folding*, so do the trade-offs linking these traits. In model $\mathbf{M_{a,h}}$ vital rates only depend on age and heterogeneity. *Parity* has no effect any more, and thus *physiological costs of reproduction* are not implemented there. However $\mathbf{M_{a,h}}$ has the same ergodic properties as $\mathbf{M}$, and thus the same fitness. Comparing these evolutionary-equivalent models, differing only in the implementation of *physiological costs of reproduction*, can thus provide useful information on the consequences of this trade-off.

Similarly, in model $\mathbf{M_{a,p}}$ vital rates only depend on age and parity, with only one apparent class of heterogeneity. Models $\mathbf{M_{a,p}}$ and $\mathbf{M}$ have same ergodic properties. However it is important to keep in mind, that $\mathbf{M_{a,p}}$ is not constructed bottom-up from $\mathbf{vr}^v$, $\mathbf{oi}_i^v$ and $\mathbf{op}_i^v$, but is a top-down product of $\mathbf{M}$ via *folding*. Therefore, the relationship it embeds between vital rates and parity will generally not be linear or even negative any more. Comparing both matrices will provide information on the detectability of *physiological costs of reproduction* for various scenarios of *genetic costs of reproduction*.

3.2.3 Incorporation of stochasticity to (*age,parity,heterogeneity*)-MPPM

Matrix models, like the (*age,parity,heterogeneity*)-MPPM $\mathbf{M}$ we have just constructed, allow to project the population over time in a constant environment and according to the expectations of transition rates as incorporated in the model. Since the environment affects all individuals in a sub-population (we call here genotype, but could also be grouped by "patches"), environmental variance can be implemented by considering an environmental time distribution, where each environment e is assigned a specific matrix $\mathbf{M_e}$. The model then consist in a family of matrices, like the (*age,parity,heterogeneity*)-MPPM-environmental-suite $\{\mathbf{M_e}\}$.

As mentioned earlier this extension from the single MPPM (in constant environment), to an MPPM-environmental-suite allows to incorporate the intra-periodic FC costs (for our studied $stor = 0$ organism). For instance, as we will illustrate later in this chapter (section 3.4.4), by assigning a higher zero-parity fertility rate for all genotypes in good environments. The joint effects of RC and FC on reproductive effort - discussed in chapter 1 - are then incorporated: FC, via the environment, affects the zero-parity vital rates and RC impacts the actual fertility rate via *parity*. This extension also allows to implement reaction norms for the various genotypes of the population. For instance, by assigning, in some specific environments,to the "robust" genotype (in the constant mean environment) lower vital rates than the "frail" genotype, one can generate crossing reaction norms which effects can then be analyzed (with the tools from section 3.4.4).

Because individual stochasticity acts independently at the level of the individual time-step, there is unfortunately no equivalent way to directly extract its effects from a matrix or suite of matrices. We will later study ways to compute the variance in fitness generated by individual stochasticity (sections 3.3.2 and 3.3.3).

In the 19$^{\text{th}}$ century, demographers and mathematicians have started to investigate the effect of individual stochasticity on demography. In particular Bienaymé and later Watson and Galton have devised the first branching process - bearing their names - that investigates the effects of stochasticity in lifetime reproductive output (the single vital rate in a non-overlapping generation framework) on asymptotic demographic behaviors (Watson and Galton, 1875). Contrary to their predecessors who only considered the expectation of the reproductive rate to predict survival of a population, they showed that certain distributions of $\mathbf{R_0}$ will both imply ergodic growth (whenever $E(\mathbf{R_0}) > 1$) *and* a large probability of extinction. It would then take a century before branching processes modeling structured were developed (Crump and Mode, 1969; Jagers, 1982). More recently demographers have developed new tools (explicitly or implicitly derived from such branching processes) to understand and measure the effect of individual stochasticity on *dynamic heterogeneity* in age-structured populations, and in particular to disentangle the relative effects of individual and environmental stochasticity on the variance in life history trajectories in the population (see section 3.3.2). We will, in this chapter, detail and use such tools and develop new ones to adapt to the complexity of models implementing both *dynamic* and *hidden heterogeneities* as traits.

Environmental MPPM suite and vital rates random variables entirely define full stochastic model

We account, in this model, for different genotypes via the addition of a *heterogeneity* trait. As such, in a given environment e, fertility and survival rates, for a given age a, parity p and genotype h- $f_e(a,p,h)$ and $s_e(a,p,h)$ are only expectations. Indeed there is no such thing as $f_e(a,p,h)$ offspring in real life and the difference between producing 1 and 0 offspring, in our simple model where reproductive effort is equated to fertility, is individual stochasticity. Therefore $f_e(a,p,h)$ and $s_e(a,p,h)$ represent the probability for an individual in that state and that environment to respectively produce one offspring and to survive. In other terms, $f_e(a,p,h)$ and $s_e(a,p,h)$ are the parameters, and hence the expectancy, of Bernoulli random variables – i.e. r.v. which only possible outcomes are 0 and 1 - representing fertility $\mathcal{F}_e(a,p,h)$ and survival $\mathcal{S}_e(a,p,h)$.

Random variables $\mathcal{F}_e(a,p,h)$ and $\mathcal{S}_e(a,p,h)$ represent stochasticity at the level of, respectively, $\mathbf{vr}^{fertility}$ and $\mathbf{vr}^{survival}$. Whilst invisible at the level of the $\{\mathbf{M_e}\}$, these processes are essential components of the model. A third distribution needs also to be made explicit: the allocation, for each state i, of the output of vr_i^v towards the different states. Contrary to $\mathcal{F}$ and $\mathcal{S}$, this distribution is already fully embedded in $\mathbf{M}$ via vectors $\mathbf{oi}_i^v$ and $\mathbf{op}_i^v$. And by providing $\mathcal{F}_e(a,p,h)$ for all states i and environments e, it is clear that distributions $\mathbf{op}_i^{fertility}$ are independent from $\mathcal{F}(i)$ and $\mathcal{S}(i)$ (by MPPM construction, from equations 3.2 and 3.3). Therefore our (*age-parity-heterogeneity*)-model is fully described, and fully analyzable, both deterministically and stochastically when presented as the combination of the deterministic transitions, the random variables of the vital rates and the trait structure that relate them : $\{M_e, \mathbf{s}, \mathcal{F}_e, \mathcal{S}_e\}$ (where $\mathbf{s}$ is the trait structure, in our case $\mathbf{s} = (\omega, par, het)$ related to the trait vector $\mathbf{t} = (age, parity, heterogeneity)$).

From time-step to lifetime individual stochasticity

The effects of the individual stochasticity, stemming from the random variables of the vital rates, $\mathcal{F}$ and $\mathcal{S}$, compounded time-step after time-step over individuals' lifetimes generate diversity in lifetime trajectories. To illustrate this, let us simplify our model by forsaking traits *parity* and *heterogeneity*; the MPPM of this simple age-structured model is therefore a Leslie matrix. In this model, each state a has, at each time-step, only 2 possible outputs per vital rate. Via survival, because there is no *dynamic heredity* tracker to account for, an individual can only transition towards *death* or $a+1$. Via fertility, because there is no *hidden heredity* classes to assign to, an individual can only transition towards 0 or 1 offspring. There are however $2^{\omega+1} - 2$ different individual trajectories allowed by the model, as shown in appendix 3.0.1. It is not easy to make sense from such a large distribution, which size will increase with every addition of *dynamic heterogeneity* a trait. Therefore from this large diversity of trajectories generated by individual stochasticity, we extract moments with simpler distributions and higher significance from a demographic and evolutionary standpoint.

Among such moments, lifetime statistics are favored as they directly show the accumulated effect of time-step individual randomness at the level of the entire trajectories; like the commonly studied variance in longevity (corresponding to variance in age at death) (Tuljapurkar, 2011; Gillespie et al., 2014; Engelman et al., 2014; van Raalte and Caswell, 2013; Caswell, 2009). From an evolutionary perspective

however, two of these lifetime variances are of paramount importance as they measure the effects of individual stochasticity on the two main population fitness measures, respectively the asymptotic growth rate λ, and the lifetime reproductive output $\mathcal{LRO}$ (which mean realization is often called net reproductive rate or $\mathbf{R_0}$). These are known as, respectively, the "demographic variance" σ_d^2 which corresponds to the expected variance of an individual contribution (survival and reproduction) to the following year's population size, and the variance of lifetime reproductive output $\sigma^2_{\mathcal{LRO}}$.

In this article, whilst we use λ as a genotype and population measure of fitness, we prefer to use $\mathcal{LRO}$ as a measure of individual fitness. Indeed lifetime reproductive output has been shown to be less relevant a fitness measure than λ at the level of a population, as it does not account for life pacing and in particular reproductive rhythm (Giske et al., 1993; Murray, 1992; Nur, 1984). It has the drawback of losing track of chronological time - a real shortcoming when dealing with models with strongly overlapping generations - and only deal with generation time. However $\mathbf{R}_0$ has for itself the advantage that it is an individual measure (that is readily aggregable at the level of the population). It is the expected number of offspring, at birth, of an individual taken at *random* in the population. This is not the case for λ, despite efforts to conceive an "individual growth rate" however still difficult to fathom (McGraw and Caswell, 1997).

From an evolutionary standpoint, σ_d^2 and $\sigma^2_{\mathcal{LRO}}$ are key statistics that allow to quantify - beyond the variance of fitnesses, induced by vital rates variance at the level of each time-step - many primordial evolutionary measures for a population. On the own hand, these variances in fitness affect the sampling of alleles from one generation to the next and therefore have been shown to affect effective population size and thus the strength of natural selection (Barrowclough and Rockwell, 1993; Engen et al., 2005a; Felsenstein, 1971; Hedrick, 2005; Vindenes et al., 2010; Hill, 1979; Rockwell and Barrowclough, 1995), . On the other hand, since variance in fitness is an individual variance, its effect at the level of the population is inversely proportional to population size, with direct impact on extinction probabilities (Engen et al., 1998, 2005b). Therefore it will itself be under negative selection (Gillespie, 1974, 1975; Vindenes et al., 2010; Shpak, 2007), which in turns can affect senescence rates (Giaimo, 2014). Finally, and most importantly, extending Tuljapurkar (1982b)'s result for environment variance, Engen et al. (2005b) showed that demographic variance affects stochastic fitness itself.

Effect of active *dynamic heterogeneity* trait

There are two kinds of *dynamic heterogeneity* traits in an MPPM. *Neutral* or passive traits which act as trackers and have no effect on the model and *active* ones which determine the model and affect vital rates. A *neutral* trait has no effect on vital rates and consequently, it is also neutral with regards to all fitness and demographic measures. Such a trait acts as a tracker, that brings information useful to the ecologist. This is the case for instance of *age*, which can be added as a neutral trait to a *stage*-structured model, in order to compute longevity and its variance. This is also the case of trait *parity* which can also be added, in the absence of costs or reproduction, to any *age*- or *stage*-structured in order to track parity trajectories and, for instance, compute inter-birth intervals (see chapter 4 sec. 4.4.2 p.112). *Folding* an MPPM on a neutral *dynamic heterogeneity* trait brings back the original model to which the "tracker" trait was appended. Active *dynamic heterogeneity* traits, for their part, have a direct determining effect on vital rates. This is the case, in particular, for traits which are part of a physiological trade-off.

In the (*age-parity-heterogeneity*)-MPPM of this chapter, via its action of vital rates, *parity* is both a product and a producer of individual stochasticity. Indeed individual stochasticity, via the realizations of $\mathcal{F}$ and $\mathcal{S}$ at former time-steps, generates stochasticity in *parity*. And, in turn, this variance in *parity* affects all remaining steps of individual trajectories via the costs of the reproduction, represented by the p indexation of $\mathcal{F}_e(a,p,h)$ and $\mathcal{S}_e(a,p,h)$. Individual stochasticity creates a variance in trajectories, even between individuals of same genotype and in same environments, that allows the costs of reproduction to generate from this variety of inputs a plurality of outputs. It is for that reason that we called - in chapter 1 (sec.1.2.3 p.23) - individual stochasticity, the fuel of the detectability of *physiological costs of reproduction*.

Individual stochasticity in a context of *hidden heterogeneity*

If several genotypes cohabit in a population, the differences in expected vital rates between these, will further increase the variance of λ and $\mathcal{LRO}$. However, the subtle difference between both fitness measures resurfaces when investigating the effect of (hidden) heterogeneity on their variances.

Let us consider, the now classical split of the variance of the stochastic growth rate into demographic and environmental variance stated by Engen et al. (1998) : $\sigma^2 = \frac{\sigma_d^2}{N} + \sigma_e^2$ where N is the population size and σ_e^2 is the environmental variance (the between-year variance of the expected individual contribution to next year's population size). With regards to λ at the level of the population, whilst the effects of the variance around the mean evolved vital rates for each genotype will decrease with population size, the

effects of the different genotypes on the population will not. They are therefore not part of the demographic variance, *sensu* Engen et al. (1998). This is especially true when not considering variability in mutation/heritability (as is the case in our model). However, the differences in vital rates and heritability implementable *between* the different environments, hint at the key fact that σ_e^2 already includes all effects of the environment, including via its affecting of genotypes reaction norms.

At the individual $\mathbf{R_0}$ level however, no change in environment is required for the differences in vital rates between genotypes to impact lifetime reproductive output. This implies that, here, genetic variance will be included in $\sigma^2_{\mathcal{LRO}}$. And therefore, in this chapter, we will denote, in heterogeneous contexts, ${\sigma^{\text{sto}}_{\mathcal{LRO}}}^2$ and ${\sigma^{\text{het}}_{\mathcal{LRO}}}^2$, the components of $\sigma^2_{\mathcal{LRO}}$ arising, respectively, from individual stochasticity (i.e. dynamic heterogeneity) and from (hidden) heterogeneity. (notation alert: het is notation that we use to segregate the two components of $\sigma^2_{\mathcal{LRO}}$İt has nothing to do with *het*, the size of the heterogeneity class in our model)

In section 3.3.2, we shall demonstrate how to compute σ_{LRO} for the various types of MPPMs encountered in this analysis.

3.3 Methods

In this section, we apply, and when needed, extend computations of key measures for evolutionary demography to matrix models incorporating *hidden heterogeneity* and/or *dynamic heterogeneity* traits. These calculations stem 1/from the field of linear algebra - for eigen-analysis yielding λor $\mathbf{R_0}$ and related sensitivity analysis - 2/ from the field of stochastic processes (in particular diffusion processes and Markov processes considering the survival transitions of a projection matrix as a Markov chain) - for the calculation of the variance in lifetime reproductive output and the demographic and environmental variances - and 3/ from both these fields - for the calculation of the stochastic growth rate for instance. We first extend the calculation of selection gradients (the sensitivities of fitness to parameters) to MPPMs. Then we show how to compute the variance of reproductive output in such models where a *dynamic heterogeneity* trait (parity) and a *hidden heterogeneity* one (genotype number) both influence and reflect reproductive success. Finally we specify how to calculate the infinitesimal demographic and environmental variances for such models.

Some intermediary tools, are readily usable for MPPMs. Among others, from the split of any multitrait matrix into its fertility and survival component : $\mathbf{M} = \mathbf{F} + \mathbf{T}$, we can draw $\mathbf{N} = (\mathbf{I} - \mathbf{T})^{-1}$ the *fundamental* matrix, that contains the expectation of time spent in each state (columns), for an individual in each state (rows). The fundamental matrix provides the distribution and thus all the moments of age-at-death (and therefore the variance in longevity in the population). From the *fundamental* matrix, we can compute $\mathbf{R} = \mathbf{FN}$ the *next generation* matrix, that contains the expectation of offspring of each category (rows) expected to be produced, over its entire remaining lifetime, by an individual in a particular state (columns).

3.3.1 Computation of selection gradients for multitrait models

In evolutionary demography, for simple *age-* (or *stage-*) structured populations, the force of selection (or selection gradient) on survival or fertility is a concept equated with sensitivities or elasticities of ergodic growth rate to vital rates. In particular it has been shown that Hamilton's indicator of the force of selection on mortality is exactly the elasticities of λ to entries of the sub-diagonal of the Leslie matrix modeling the age-structured population (Baudisch, 2005). Indeed Leslie matrices' transitions rates are directly either fertility rates or survival rates and therefore elasticities can be directly calculated from the right and left eigen-vectors $\mathbf{w}$ and $\mathbf{v}$ associated with λ, the maximum eigenvalue or the ergodic growth rate.

Two issues however complicate the extraction of selection gradients on mortality or fertility from a multitrait model. First a conceptual one: the addition of *hidden heterogeneity* traits in an MPPM, and thus of heredity, raises questions with respect to the interpretation of growth rate sensitivities as selection gradients, as discussed in chapter 1 section 1.3.2 p.33. Second a computation issue. In MPPMs, because of the multiple traits, these vital rates do not appear directly in the $M_{i,j}$ matrix entries, which contain combinations of such vital rates (vr_i^v), probability distribution of output states $\text{op}_{i,j}^v$ and other parameters; they are lower level parameters. The sensitivity and elasticity of λ to any parameter p, including any vital rate vr_i^v, for a population modeled by MPPM $\mathbf{M}$ requires the calculus chain-rule to be generated: $\frac{\partial \lambda}{\partial p} = \sum_{i,j} \frac{\partial \lambda}{\partial M_{i,j}} . \frac{\partial M_{i,j}}{\partial p}$ (Caswell, 1989), which can computed directly from the multidimensional sensitivity matrix $\mathbb{S}$ defined in chapter 2 (p.41).

In the specific case of (*age-parity-heterogeneity*)-MPPMs, vital rates appear directly in the reference Leslie matrix $\mathbf{M_a}$. In matrices $\mathbf{M}$, $\mathbf{M_{a,p}}$ and $\mathbf{M_{a,h}}$ however, entries are combinations of different components. First, external parameters (i.e., not a trait or a vital rate). In $\mathbf{M}$ and $\mathbf{M_{a,h}}$, fertility transitions

depend on mutation rate (called m in $\mathbf{M}$). Second multiple vital rates. In $\mathbf{M}$ and $\mathbf{M_{a,p}}$, survival transitions are the product of survival rates with fertility or its complement to 1. And third these vital rates are themselves combinations of trait values and zero-parity vital rates. In $\mathbf{M}$ for instance, parity p and zero-parity vital rates generate all vital rates. Because these combinations are multiplicative, we measure the selection gradient with elasticities. And because the trade-off we are studying, the *physiological costs of reproduction*, is about the effects of the realization of fertility rates, we will specifically focus on elasticities of λ to fertility rates.

To achieve this, we need to obtain both components of the chain-rule equation in order to calculate $\mathbf{e}_{\mathbf{M}_t}$ the multidimensional vector of elasticities to fertility of all states i of λ_t the growth rate of $\mathbf{M}_t$ (with associated eigenvectors $\mathbf{w}_{\mathbf{M}_t}$ and $\mathbf{v}_{\mathbf{M}_t}$) that is $\mathbf{M}$ folded on set of traits t (which can be empty) :

$$\mathbf{e}_{\mathbf{M}_t} = \left\{ \frac{fert_i}{\lambda} . \frac{\partial \lambda}{\partial fert_i} \right\}_i = \left\{ \frac{fert_i}{\lambda} . \sum_{j,k} \frac{\partial \lambda}{\partial M_{j,k}} . \frac{\partial M_{j,k}}{\partial fert_i} \right\}_i \tag{3.4}$$

where $\left\{ \frac{\partial \lambda}{\partial M_{j,k}} \right\}_{j,k} = \{w_k . v_j\}_{j,k}$ is the sensitivity matrix (Caswell, 1989) and $\left\{ \frac{\partial M_{j,k}}{\partial fert_i} \right\}_{i,j,k}$ is the multidimensional parameter sensitivity matrices (of dependencies of matrix entries to fertility rates) deduced from the construction method. These are respectively denoted $\mathbf{S}$ and $\mathbb{S}$ in chapter 2 equation 2 (sec.2.3,p. 47), and the detailed steps to obtain these matrices are provided there.

In order to be able to compare $\mathbf{e}_{\mathbf{M}}$, $\mathbf{e}_{\mathbf{M}_{a,h}}$, $\mathbf{e}_{\mathbf{M}_{a,p}}$ and $\mathbf{e}_{\mathbf{M}_a}$, we need to fold these on their common denominator trait: *age*; i.e. to produce the elasticity of λ to fertility for states sharing the same *age* category. This is simply done by summing all elements of $\mathbf{e}_{\mathbf{M}_t}$ with the same age value, thus considering parallel moves in all fertility rates of the category. This makes all the more sense since, in the main matrix, all fertilities are proportional to zero-parity fertilities with other factors structurally fixed. Thus we obtain as a measure of the force of selection on fertility, for ergodic-equivalent models ($\mathbf{M}$, $\mathbf{M_{a,p}}$, $\mathbf{M_{a,h}}$ and $\mathbf{M_a}$) incorporating physiological costs of reproduction ($\mathbf{M}$ and $\mathbf{M_{a,p}}$) and heterogeneity (($\mathbf{M}$ and $\mathbf{M_{a,h}}$) or not, the elasticities of λ (by construction the same for all 4 models) to fertility rate by age classes $\mathbf{e}^{age}_{\mathbf{M}}$, $\mathbf{e}^{age}_{\mathbf{M}_{a,h}}$, $\mathbf{e}^{age}_{\mathbf{M}_{a,p}}$ and $\mathbf{e}_{\mathbf{M}_a}$.

Having extended the deterministic computation framework from matrices structured by one or several *basic* traits to multitrait matrices incorporating *dynamic* and/or *hidden heterogeneity* traits, we now wish to be able to compute the effect of stochasticity on such fitness measures as the variance in lifetime reproductive output (next section 3.3.2) and the ergodic (stochastic) growth rate (section 3.3.3).

3.3.2 Variance in lifetime reproductive output for multitrait models

The net reproductive rate, $\mathbf{R_0}$ is an individual measure. The fate of individuals depends on environmental conditions, the $\mathbf{e}$ in $\mathcal{F}_e(i)$ and $\mathcal{S}_e(i)$. However since these are shared by all individuals in the population (the reason why all individuals under e share the same stochastic model $\{\mathbf{M_e}, \mathbf{s}, \mathcal{F}_e, \mathcal{S}_e\}$), the effect of environmental variance is better suited to analysis at the level of the population. This will therefore be provided by the population fitness measure that is the stochastic growth rate λ (see next section 3.3.3). To the contrary, the level of the individual, and its fitness measure $\mathbf{R_0}$ (or $\mathcal{LRO}$ when designated as a random variable), befits the analysis of the effects of individual stochasticity. The entire distribution of $\mathcal{LRO}$ is informative, but most key initial conclusions can be drawn from its second central moment $\sigma^2_{\mathcal{LRO}}$.

As we have seen before, formulas of deterministic transition rates are complicated by the addition of *parity* and/or *heterogeneity* as traits (section 3.2.2). This is consequently also the case for stochastic transitions. If every transition rate in Leslie matrix $\mathbf{M_a}$correspond to a simple Bernoulli process (either $\mathcal{F}$ or $\mathcal{S}$), it is not the case for the other MPPMs. In $\mathbf{M}$, for instance, for each state i and each vital process v (fertility or survival), the set of all stochastic transitions towards every reachable state (i.e. in set $\mathbf{oi}^v_i$) constitute a categorical distribution (or generalized Bernoulli). For vital process fertility for instance, the categorical process is the product of Bernoulli processes $\mathcal{F}$ and categorical process represented by distribution $\mathbf{op}^{fertility}_i$. This increase in complexity in the stochastic processes causes necessarily an increase in complexity for calculation of $\sigma^2_{\mathcal{LRO}}$.

Age-structured populations

The random variable $\mathcal{LRO}$ can be considered to be an extension of r.v. age-at-death. Whilst the later is based solely on time-step stochastic process $\mathcal{S}$, the former combines both $\mathcal{S}$ and $\mathcal{F}$. This provides a pathway towards the calculation of $\sigma^2_{\mathcal{LRO}}$ in *age-* (or *stage-*) structured models: first generate the distribution of survival trajectories, which act as a backbone on which to graft the stochastic fertility process at each time-step. This was, we think the reasoning behind the approach of Tuljapurkar, Steiner and

colleagues, who seemingly start from Caswell (2006)'s formula for variance in longevity as the first element of $(2\mathbf{N}-\mathbf{I}).\mathbf{N}$, and write that the 2nd moment of $\mathcal{LRO}$ is the first element of $\mathbf{F}.(2\mathbf{N}-\mathbf{I}).\hat{\mathbf{F}}.\mathbf{N}$ (Tuljapurkar and Steiner, 2010; Steiner and Tuljapurkar, 2012), where $\hat{\mathbf{F}}$ is a matrix with only diagonal elements equal to the fertility rates. At the same time, Caswell and colleagues adapted to demography, the mathematical tool of *Markov chain with rewards* (MCwR) -its concept is detailed in next section 3.3.2 - which provides the various moments of $\mathcal{LRO}$ by matrix multiplicative convergence.(Caswell, 2011; Caswell and Salguero-Gómez, 2013; Caswell, 2014; van Daalen and Caswell, 2015; Caswell and Sánchez Gassen, 2015). Both approaches have in common to append on the stochasticity of survival trajectories (stemming from $\mathbf{N}=(\mathbf{I}-\mathbf{T})^{-1}$ or directly from $\mathbf{T}$ for MCwR), the stochasticity of fertility successes reaped along these trajectories (stemming $\mathbf{F}$ or the *reward matrix* of MCwR, see sec. 3.3.2).

However we find these two approaches unsatisfactory. The matrix closed-form formula we deem incorrect (we certainly are wrong, but it seems to append *deterministic* fertility behavior on top of survival stochasticity). The MCwR's complexity we deem unnecessary for such simple models. For these reasons, we provide here a closed-form formula -all the steps that lead to the formula are to be found in appendix section3.0.2- for the variance of reproductive output in an age-structured population:

$$Var(\mathcal{LRO}) = \alpha_1 = \sum_{i=1}^{n} P_i \left[Var(\mathcal{F}_i) + y_{i+1}^2 s_i(1-s_i) \right], \tag{3.5}$$

where $\mathcal{F}_i$ is the fertility process at age i of expectation f_i, s_i the survival rate at age i, $P_i = \prod_{k=1}^{i-1} s_i$ the probability to survive to age i and $y(i) = \frac{1}{P_i}\sum_{j=i}^{n} f_j P_j$ the expectation of $\mathcal{LRO}_\mathrm{i}$, the remaining reproductive output for an individual aged i. Survival is necessarily a Bernoulli process in an age-structured model. Fertility, for its part, can have any distribution as long as it is positive. In the framework of our (*age-parity-heterogeneity*)-MPPM with maximum 1 offspring per time step, the equation can be simplified by setting $Var(\mathcal{F}_i) = f_i(1-f_i)$ (see eq. 3.39 in section 3.0.2, p.92). From general equation 3.5, can one draw the variance in longevity by simply setting all f_i and $\mathcal{F}_i$ to 0. And it is also possible to disentangle the effects of survival stochasticity on $\mathcal{LRO}$, by setting all s_i at 1.

Multitrait models with *hidden heterogeneity* traits

The addition of *hidden heterogeneity* traits adds complexity to the calculation of the variance of reproductive success $\sigma^2_{\mathcal{LRO}}$. Indeed, the difference in the number of offspring between two individuals can now stem either from *dynamic heterogeneity* generated by individual stochasticity (as studied in the previous section 3.3.2), or from the *hidden heterogeneity* that assigns different mean vital rates to the different genotypes.

As a matter of fact, the addition of *hidden heterogeneity* traits already adds complexity to the calculation of the *expectation* of reproductive success $\mathbf{R_0}$. We provide in appendix section 3.0.3 (p.93), a discussion on the matter and the calculation steps towards a new formula for $\mathbf{R_0}$ in heterogeneous populations (equation 3.42). This equation provides $\mathbf{R_0}$ from the projection matrix of any structured population with several classes of offspring. In non-matrix notation, this formula can be written :

$$E(\mathcal{LRO}) = \sum_{h=1}^{het} \mathrm{e}_h^{\mathcal{LRO}}.\mathrm{w}_h^{\diamond} \quad , \tag{3.6}$$

where $\mathbf{e}_{\mathcal{LRO}}$ is the sum of all lines in Next-Generation Matrix $\mathbf{R}=\mathbf{F}.(\mathbf{I}-\mathbf{T})^{-1}$ (from the survival-fertility decomposition of $\mathbf{M}=\mathbf{T}+\mathbf{F}$) and $\mathrm{w}^{\diamond}$ is the vector of relative ergodic abundances of all states for which $age=1$.

We shall here use this result and extend this approach to calculate $\sigma^2_{\mathcal{LRO}}$ in a population structured by *age* and *heterogeneity*. In an heterogeneous population, this quantity corresponds to the variance of $\mathcal{LRO}$ for an individual taken at random in the population and therefore will also make use of the ergodic relative offspring abundances in $\mathbf{w}^{\diamond}$. To simplify our formulas, but without any loss of generality, let us consider $het=2$ *heterogeneity* classes, we call h_1 and h_2. Let $\mathcal{LRO}_1$ be the random variable representing $\mathcal{LRO}$ knowing the individual is of class h_1 and respectively $\mathcal{LRO}_2$ for class h_2. Then the vector $\mathbf{e}_{\mathcal{LRO}}$ defined above and in appendix 3.0.3 (the expectation of lifetime reproductive output per *heterogeneity* class is the couple $\{E(\mathcal{LRO}_1), E(\mathcal{LRO}_2)\}$.

Similarly, we can define vector $\boldsymbol{\sigma}^{\mathbf{2}}_{\mathcal{LRO}}$ providing the variance of lifetime reproductive output per *heterogeneity* class, i.e. the couple $\{Var(\mathcal{LRO}_1), Var(\mathcal{LRO}_2)\}$. Contrary to vector $\mathbf{e}_{\mathcal{LRO}}$ (the sum of lines of $\mathbf{R}$), $\boldsymbol{\sigma}^{\mathbf{2}}_{\mathcal{LRO}}$ cannot be vectorially extracted from $\mathbf{M}$. Rather, from the full model $\{\mathbf{M}, \boldsymbol{s}, \mathcal{F}\}$, one can derive the survival rates and fertility random processes for each genotype from which the equation 3.5 obtained in previous section 3.3.2 yields the desired variances.

From the usual decomposition of variance,

$$Var(\mathcal{LRO}) = E(\mathcal{LRO}^2) - E(\mathcal{LRO})^2 \quad , \tag{3.7}$$

emerge two quantities. First $E(\mathcal{LRO})$, which according to equation 3.6 corresponds to

$$E(\mathcal{LRO}) = \mathrm{w}_1^\diamond.\mathrm{e}_{\mathcal{LRO}_1} + \mathrm{w}_2^\diamond.\mathrm{e}_{\mathcal{LRO}_2} \tag{3.8}$$

Second $E(\mathcal{LRO}^2)$, the expectation of the square of reproductive success, which can be decomposed, at the stable state, like $E(\mathcal{LRO})$ (eq.3.6):

$$E(\mathcal{LRO}^2) = w_1^\diamond.E({\mathcal{LRO}_1}^2) + w_2^\diamond.E({\mathcal{LRO}_2}^2) \tag{3.9}$$

Equation 3.9 can be rewritten using the variance decomposition, already used in equation 3.7, for $\mathcal{LRO}_1$ and $\mathcal{LRO}_2$:

$$E(\mathcal{LRO}^2) = w_1^\diamond.(\mathrm{Var}(\mathcal{LRO}_1) + \mathrm{E}(\mathcal{LRO}_1)^2) + w_2^\diamond.(\mathrm{Var}(\mathcal{LRO}_2) + \mathrm{E}(\mathcal{LRO}_2)^2) \tag{3.10}$$

And thus, replacing equations 3.8 and 3.10 into equation 3.7, we can write $\sigma^2_{\mathcal{LRO}}$ as a function of vectors $\mathbf{e}_{\mathcal{LRO}}$, $\boldsymbol{\sigma}_{\mathcal{LRO}}$ and $\boldsymbol{w}^\diamond$, in a manner - hinted at in section 3.2.3 - that reveals its decomposition into ${\sigma^{\mathrm{sto}}_{\mathcal{LRO}}}^2$ the individual stochasticity component (related to *dynamic heterogeneity*) and ${\sigma^{\mathrm{het}}_{\mathcal{LRO}}}^2$, the (hidden) *heterogeneity* component:

$$\sigma^2_{\mathcal{LRO}} = \underbrace{w_1^\diamond.\sigma^2_{\mathcal{LRO}_1} + w_2^\diamond.\sigma^2_{\mathcal{LRO}_2}}_{{\sigma^{\mathrm{sto}}_{\mathcal{LRO}}}^2} + \underbrace{w_1^\diamond.{\mathrm{e}_{\mathcal{LRO}_1}}^2 + w_2^\diamond.{\mathrm{e}_{\mathcal{LRO}_2}}^2 - {\mathrm{e}_{\mathcal{LRO}}}^2}_{{\sigma^{\mathrm{het}}_{\mathcal{LRO}}}^2} \tag{3.11}$$

We know the heterogeneity component, ${\sigma^{\mathrm{het}}_{\mathcal{LRO}}}^2$, is indeed positive from the Jensen inequality applied to the convex square function and indeed we can rewrite, as $w_2^\diamond = 1 - w_1^\diamond$,

$${\sigma^{\mathrm{het}}_{\mathcal{LRO}}}^2 = w_1^\diamond.{\mathrm{e}_{\mathcal{LRO}_1}}^2 + (1 - w_1^\diamond).{\mathrm{e}_{\mathcal{LRO}_2}}^2 - (w_1^\diamond.\mathrm{e}_{\mathcal{LRO}_1} + (1 - w_1^\diamond).\mathrm{e}_{\mathcal{LRO}_2})^2$$

which yields

$${\sigma^{\mathrm{het}}_{\mathcal{LRO}}}^2 = w_1^\diamond(1 - w_1^\diamond)(\mathrm{e}_{\mathcal{LRO}_1} - \mathrm{e}_{\mathcal{LRO}_2})^2 \tag{3.12}$$

Multitrait models with *dynamic heterogeneity* traits

The principle of the above formulae, providing $\sigma^2_{\mathcal{LRO}}$ for *age*-structured populations (equation 3.5), and (*age-heterogeneity*)-MPPMs (equation 3.11) can be extended to any *basic* trait (in the $\mathcal{B}$ family), *hidden heterogeneity* trait (in the $\mathcal{G}$ family) trait and *neutral dynamic heterogeneity* trait (a trait in the $\mathcal{D}$ family that serves as a tracker and not as a determinant). It can be shown to be also the case for the approaches of Tuljapurkar and Steiner (2010) and Caswell (2011) mentioned above. However whenever an *active dynamic heterogeneity* trait is incorporated (a trait that influences vital rates), none of these approaches and formulae can provide $\sigma^2_{\mathcal{LRO}}$. This is because with such traits, and contrary to all other cases, the random processes behind the transitions cannot always be considered independent, a necessary condition for these computations (see appendix section 3.0.2).

Specifically, if all traits of a model are drawn from $\mathcal{B}$ or $\mathcal{G}$, or are passive members of $\mathcal{D}$, then the stochastic processes at play for state i, combining vital processes $\mathcal{F}_i$ and $\mathcal{S}_i$ with output states distributions $\{\mathbf{oi}_i^v, \mathbf{op}_i^v\}$ towards the various values of $\{\mathcal{B}, \mathcal{G}\}$ (when $v = fertility$) and $\{\mathcal{B}, \mathcal{D}\}$ (when $v = survival$) are all independent. An equivalent alternative way to consider this, is to say that output states distributions $\{\mathbf{oi}_i^v, \mathbf{op}_i^v\}$ do not depend on $\mathcal{F}_i$ and $\mathcal{S}_i$. Simply put, for such traits, the realization of vital rates only depend on input state i.

To the contrary, whenever a physiological trade-off is implemented in a matrix model, via the addition of an active trait in $\mathcal{D}$, like *parity*, this is not the case any more. In our *age-parity*-MPPM, for instance, the realization of $\mathcal{F}_i$ depends on the output state of i through $\mathbf{oi}_i^{survival}$. This may seem counter-intuitive at first, as $\mathcal{F}_i$ is clearly only a function of i. However, since the output state of i via survival depends on the realization of its fertility event at that time, $\mathcal{F}_i$ and the stochastic process behind $\{\mathbf{oi}_i^{survival}, \mathbf{op}_i^{survival}\}$ are dependent. Calculation of $\sigma^2_{\mathcal{LRO}}$ for such a model, therefore requires the use of a tool for which the fertility process is not implemented as *vectors* of moments per input state (like vector $\mathbf{vr}^{fertility}$ which provides the expectation of fertility rates of each state and, in the particular case where $\mathcal{F}$ is Bernoulli, all further moments as well), but as a *matrix* containing the different moments of fertility for an individual surviving from state i to state j. A Markov chain with rewards (MCwR) is such a tool, which "reward" matrix can implement the various moments of $\mathcal{F}$ as a function of both ends of an $i \leftrightarrow j$ transition. And thus we will use MCwR to compute the variance of lifetime reproductive output in our (*age-parity*)- and

(*age-parity-heterogeneity*)-MPPMs.

To describe the MCwR mathematical framework, we will use the approach of Caswell (2011) when he applied, first, these tools - originally described in (Howard, 1960; Hatori, 1966)- to demography. The MCwR framework requires two instruments. First, $\tilde{\mathbf{T}}$ the extended matrix of transitions which is $\mathbf{T}$, the usual matrix of transitions, where, ordinarily implicit absorbing state *death* is made explicit in the $q+1^{\text{th}}$ position of the matrix : $\tilde{\mathbf{T}} = \left[\begin{smallmatrix} \mathbf{T} & 0 \\ \mathbf{m} & 1 \end{smallmatrix}\right]$, where $\mathbf{m}$ is the vector of mortality rates, i.e. $\mathbf{m} = \mathbf{1}' - \mathbf{1}'.\mathbf{T}$. Matrix $\tilde{\mathbf{T}}$ is a stochastic matrix - columns sum to 1 - that fully describes the Markov chain of all possible survival trajectories that any individual in the population can take before being absorbed by *death*. Second, the family of "reward matrices" $\mathbf{Rw^k}$ where $\mathrm{Rw^k}_{ij}$ is the k^{th} moment of the random variable of the reward - i.e. the birth of 1 offspring - for an individual transitioning for state j to state i.

As discussed above, whenever no trade-off involving fertility is involved, then matrices $\mathbf{Rw}^k$ have rank 1, with all lines equal to the fertility rate vector $\mathbf{vr}^{fertility}$. In $\mathbf{M}$ -implementing *physiological costs of reproduction* current reproductive success does not only depend on the state $i = (a, p, h)$ of an individual, but also on the state j it is transitioning toward. In detail, if $j = (a+1, p+1, h)$ then reproduction is being successful and thus its expectancy is $\mathrm{Rw}^1_{j,i} = 1$. If individual survives but reproduction is not successful, i.e. when $j = (a+1, p, h)$, then $\mathrm{Rw}^1_{j,i} = 0$. Finally, if individual dies at the end of the period, i.e. $j = death$, then (as fertility and survival processes are independent) $\mathrm{Rw}^1_{death,i} = f_i = \mathrm{vr}^{\mathrm{fertility}}{}_i$. This completes the construction of the "reward matrix" $\mathbf{Rw}^1$. It is, therefore, is an empty matrix, bar one sub-diagonal made of 0s, one sub-diagonal made of 1s and the bottom row worth $\mathbf{vr^f}$. The reproductive rewards are (in our model) Bernoulli processes, and thus reward matrices for any other moment, $\mathbf{Rw}^k$, are equal to reward matrix of expectations $\mathbf{Rw}^1$. Let $\boldsymbol{\rho}_\mathbf{k}$ be the vector of the kth moment of $\mathcal{LRO}$, indexed on individuals "starting" states. These are calculated as the convergence of backwards accumulation of "remaining" rewards following individuals from *death* (where there is no remaining reward left) to birth (or age $a = 1$). From (Caswell et al., 2011), we draw the following convergence equations, for the first two moments:

$$\boldsymbol{\rho_1} = \lim_{t\to+\infty} \boldsymbol{\rho_1}(t) \text{ with } \boldsymbol{\rho_1}(t+1) = (\tilde{\mathbf{T}} \circ \mathbf{Rw^1})'.\mathbf{1} + \tilde{\mathbf{T}}.\boldsymbol{\rho_1}(t) \tag{3.13}$$

$$\boldsymbol{\rho_2} = \lim_{t\to+\infty} \boldsymbol{\rho_2}(t) \text{ with } \boldsymbol{\rho_2}(t+1) = (\tilde{\mathbf{T}} \circ \mathbf{Rw^2})'.\mathbf{1} + 2.(\tilde{\mathbf{T}} \circ \mathbf{Rw^1})'.\boldsymbol{\rho_1}(t) + \tilde{\mathbf{T}}.\boldsymbol{\rho_2}(t) \tag{3.14}$$

with initial conditions $\boldsymbol{\rho_1}(0) = \boldsymbol{\rho_2}(0) = \mathbf{0}$ ($\circ$ is the Hadamard, termwise, product)

Let us now reduce all $\mathcal{LRO}$ moments $\boldsymbol{\rho}_\mathbf{k}$ to states of age $a = 1$ (and thus $p = 0$), i.e. to the offspring states. Then $\boldsymbol{\rho_1}$ and $\boldsymbol{\rho_2}$ are of size *het* (the number of classes of *trait heterogeneity*).

Then the vector of expectancy of $\mathcal{LRO}$ for each offspring class is

$$\mathbf{e}_{\mathcal{LRO}} = \boldsymbol{\rho_1} \tag{3.15}$$

and corresponds - the addition of *heterogeneity* has no effect here - to $\mathbf{1}'.\mathbf{R}^*$ for any constructed matrix model (see appendix 3.0.3). This is not true for some *folded* matrices such as $\mathbf{M_{a,p}}$, which case we discuss at the end of this section and thoroughly in appendix 3.0.4.

And the vector of variance of $\mathcal{LRO}$ for each offspring class is

$$\boldsymbol{\sigma}^\mathbf{2}_{\mathcal{LRO}} = \boldsymbol{\rho_2} - \boldsymbol{\rho_1} \circ \boldsymbol{\rho_1} \tag{3.16}$$

Thus, in this section, we have provided, for the calculation of $\sigma^2_{\mathcal{LRO}}$, a closed-form formula for models structured by *age* only (equation 3.5), an eigen-analysis equation for variance at the stable-state for models incorporating *hidden heterogeneity* (equation 3.11) and finally a convergence procedure to reach that quantity via MCwR for models also embedding *active dynamic heterogeneity* traits (equation 3.16). Since these categories are nested (a model structured by *age* only can, for instance, be considered to be a model structured by *age* and *heterogeneity*, with only 1 *heterogeneity* class), the variance in lifetime reproductive output of a population modeled by a simple Leslie matrix - implemented directly or *folded* from a larger model like $\mathbf{M_a}$- can be calculated via the three formulas. Similarly from a constructed (age-heterogeneity)-MPPM $\mathbf{M_{age,heterogeneity}}$ or a folded one, like $\mathbf{M_{a,h}}$, one can equivalently compute $\sigma^2_{\mathcal{LRO}}$ from either eq.3.11 or eq.3.16. Finally, whilst the variance of $\mathcal{LRO}$ can only be calculated from eq.3.16 for models incorporating *parity*, we expect its expectation $\mathbf{R_0}$ to be equivalently calculated by the MCwR approach (eq. 3.15) or by the Next Generation Matrix directly (see appendix 3.0.3). However, because of the emergence of vital processes in the output states distributions due to the addition of active *dynamic heterogeneity* traits implementing physiological trade-offs (see eq. 3.3 and eq. 3.2), this is not the case for *folded* matrices incorporating *dynamic heterogeneity* traits. Matrix $\mathbf{M_{a,p}}$ in particular - $\mathbf{M}$ folded on *heterogeneity*- is such a model for which the discrepancy between the MCwR and $\mathbf{R}$ approaches for $E(\mathcal{LRO})$ calculation is just a revealer of larger interpretation and analysis issues. This is an important aspect of *folding*, and as we think we did not make this point clear enough in chapter 2 (p.41), we discuss this issue further in appendix section 3.0.4.

3.3.3 Computation of stochastic growth rate

So far, we have designed an evolutionary-demographic model that implements physiological costs of reproduction in a genotypically heterogeneous context via *age-parity-heterogeneity*-MPPM $\mathbf{M}$. In order to be able to understand the specific and combined consequences of costs of reproduction and heterogeneity, we have derived, from $\mathbf{M}$ via EFP-folding, the evolutionary-equivalent models $\mathbf{M_{a,h}}$, $\mathbf{M_{a,p}}$ and $\mathbf{M_a}$. Then we have designed tools to measure the consequences of costs and heterogeneity by comparing these models for certain evolutionary outputs; the force of selection via the computation of selection gradients, and the variance in lifetime reproductive output. These models implement variance in genotypes and we have shown they can be used to compute the effects of individual stochasticity. All the tools so far were however suitable for analyses in constant environment, by considering only the (*age-parity-heterogeneity*)-MPPM $\mathbf{M}$ constituted of the mean environmental vital rates; $\mathbf{M} = \mathbf{M^{\bar{e}}} = E_{\mathbf{e}}(\mathbf{M_e})$, where $\mathbf{M_e}$ is the MPPM in environment $\mathbf{e}$

We will now contemplate the effects of general stochasticity, and environmental stochasticity in particular. In other words, we will now consider the full stochastic model (see section 3.2.3) which consist of the environmental suite of MPPMs, as well as the individual stochastic processes : $\{M_e, \mathbf{s}, \mathcal{F}_e, \mathcal{S}_e\}$. There are two approaches to such a problem, the probabilistic approach and the simulation approach. The simulation approach consists in projecting over (long periods of time), over many runs, the evolution of a population where, at each time-step, environment e is drawn from the family of all possible environments $\mathcal{E}$, and the survival and reproductive fate of *each* individual is drawn from the individual vital rates processes associated with e, $\mathcal{F}_e$ and $\mathcal{S}_e$. This allows to picture the effects on both individual and environmental stochasticities on the fitness of a population, measured as the long term stochastic growth rate. We will display the results of such a simulation for our model in section 3.4.4.

The probabilistic approach consists in a direct calculation of the demographic and environmental variances, and therefore of the stochastic growth rate, using the sensitivity matrix of evolutionary demography. As a consequence, it is an approach that is appropriate for small environmental variations (the sensitivity matrix provides the *marginal* effect on λ of *marginal* changes in matrix entries). Drawing from earlier work by Cohen (Cohen, 1977, 1979) on stochasticity in age-structured populations and by Tuljapurkar on environmental stochasticity (Tuljapurkar, 1982a, 1990b), and his own earlier work on demographic variance (Engen et al., 1998), Engen et al. (2005b) showed, that the first-order approximation of the long-term stochastic growth rate of an age-structured population depends on three parameters. First, λ the deterministic ergodic growth rate drawn from the matrix $\mathbf{M}$ of mean (environmentally and demographically) vital rates (see section 3.2.2). Second, the environmental variance introduced by Tuljapurkar (Tuljapurkar, 1982a, 1990b) which measures the (infinitesimal) variance induced on the growth rate by (infinitesimal) changes on vital rates due to environmental stochasticity :

$$\sigma_{\mathrm{e}}^2 \approx \sum_{i,j}\sum_{k,l} \lambda^{-2} \frac{\partial \lambda}{\partial \mathrm{M}_{i,j}} \frac{\partial \lambda}{\partial \mathrm{M}_{k,l}} \mathrm{Cov_e}(\mathrm{M}_{i,j}, \mathrm{M}_{k,l}) \tag{3.17}$$

where $\mathrm{Cov_e}(\mathrm{M}_{i,j}, \mathrm{M}_{k,l})$ is the environmental component of the covariance between entries $\mathrm{M}_{i,j}$ and $\mathrm{M}_{k,l}$ (Tuljapurkar, 1982a). Third, the demographic variance, measuring the variance induced on the stochastic growth rate due to individual stochasticity, which in an MPPM is much simplified in :

$$\sigma_{\mathrm{d}}^2 \approx \sum_{i,j}\sum_{k,l} \lambda^{-2} \frac{\partial \lambda}{\partial \mathrm{M}_{i,j}} \frac{\partial \lambda}{\partial \mathrm{M}_{k,l}} .N.\mathrm{Cov_d}(\mathrm{M}_{i,j}, \mathrm{M}_{k,l}) \tag{3.18}$$

where N is population size and $\mathrm{Cov_d}(\mathrm{M}_{i,j}, \mathrm{M}_{k,l})$ is the demographic component of the covariance of realizations of transitions $\mathrm{M}_{i,j}$ and $\mathrm{M}_{k,l}$ (Engen et al., 2005b).

Let us quickly describe these two variances and how they combine to generate σ^2, the total variance in individual contribution to the population at next time step in growth rate:

$$\sigma^2 = \sigma_{\mathrm{e}}^2 + \frac{\sigma_{\mathrm{d}}^2}{N} \approx \sum_{i,j}\sum_{k,l} \frac{\partial ln(\lambda)}{\partial \mathrm{M}_{i,j}} \frac{\partial ln(\lambda)}{\partial \mathrm{M}_{k,l}} .\mathrm{Cov}(\mathrm{M}_{i,j}, \mathrm{M}_{k,l}) \quad , \tag{3.19}$$

which can be interpreted as the effect on $ln(\lambda)$ of the flexibility allowed by the model: all pairs of entries are "authorized" to draw nearer or pull away according to their covariance, and the effect on $ln(\lambda)$ is proportional to sensitivities $\frac{\partial ln(\lambda)}{\partial \mathrm{M}_{i,j}}$ and $\frac{\partial ln(\lambda)}{\partial \mathrm{M}_{k,l}}$. Then this total variance is decomposed between amongst time-steps variance (eq. 3.17) and within time-step variance (eq. 3.18). The emergence of N in eq. 3.18 comes from the original definition of overall variance σ^2 primarily environmental variance σ_e^2. When individual stochasticity is not accounted for, the individual contribution to population growth is: $Var(N_{t+1} \mid N_t = N) = N^2\sigma_e^2$. Individual stochasticity can be added into the mix, however, since the individual processes are all independent, we get : $Var(N_{t+1} \mid N_t = N) = N^2\sigma_e^2 + N\sigma_d^2$ and therefore the stochastic growth rate $\ln \lambda_{\mathrm{s}} = \lim_{t\to+\infty} \frac{\ln N}{t}$ has therefore the infinitesimal variance $\sigma_{\mathrm{e}}^2 + \frac{\sigma_{\mathrm{d}}^2}{N}$

found in eq. 3.19. The same difference would emerge if instead of analyzing $\mathcal{LRO}$, one would focus on variance of average reproductive success in the population ($pop\mathcal{LRO}$) At the level of the population (N individuals), the variance of $pop\mathcal{LRO}$ is then (as individuals reproduce independently from one another): $Var(pop\mathcal{LRO}) = Var(\frac{1}{N}\sum_{i=1}^{N}\mathcal{LRO}_i) = \frac{1}{N^2}\sum_{i=1}^{N}Var(\mathcal{LRO}_i) = \frac{Var(\mathcal{LRO})}{N}$ and thus we have the same ratio between variance of $\mathcal{LRO}$ at the level of the individual and the population than demographic variance (an individual measure).

From equation 3.19, Tuljapurkar (1982a) and Engen et al. (2005b) provided the approximation of the stochastic order at the first-order (stemming from $ln(1+x) = x - \frac{x^2}{2} + o(x^2)$):

$$\ln \lambda_{\mathrm{s}} \approx \ln \lambda - \frac{\sigma_{\mathrm{e}}^2}{2} - \frac{\sigma_{\mathrm{d}}^2}{2.N} \tag{3.20}$$

(Engen et al., 2005b).

Computation of demographic variance for MPPMs

In the general MPPM framework we have designed to implement physiological costs of reproduction, the addition of traits enlarges the model, but stochasticity analysis is made simpler. Indeed, contrary to the general probabilistic framework (e.g. equation 3.18), the individual scrutiny provided by the numerous traits allows to make all individual processes between different states independent. For instance, instead of rendering physiological costs by declaring that the 2-year and 3-year fertility rates have negative covariance (a method which would fail to implement the costs as defined in chapter 1), the indexation of states on *parity* in our (*age-parity*)-model ensures that for all states $i \leftrightarrow (a,p)$ and $j \neq i$, and any pair of states (k,l), $Cov_d(M_{k,i}, M_{l,j}) = 0$. This allows us to simplify equation 3.18, in the MPPM framework $\{\mathbf{M_e}, \mathbf{s}, \mathcal{F}_e, \mathcal{S}_e\}$ where all determining traits are in trait structure $\mathbf{s}$ and all constrains implemented in the $\mathbf{M_e}$, $\mathcal{F}_e$ and $\mathcal{S}_e$:

$$\sigma_{\mathrm{d}}^2 \approx \lambda^{-2}\sum_{i=1}^{\mathrm{q}} w_i \sum_{j,k} v_j v_k \mathrm{Cov_d}(\mathrm{M}_{j,i}, \mathrm{M}_{k,i}) \tag{3.21}$$

where $\boldsymbol{w}$ and $\boldsymbol{v}$ are the right- and left-eigenvector of the MPPM, corresponding to λ. This simplification - not as radical as for a Leslie matrix, see equation 3 in (Engen et al., 2005a) - stems from the consideration that the only non-negative covariances are those between transitions sharing same input state, e.g. $Cov_d(M_{j,i}, M_{k,i})$. Then from the sensitivity matrix formula $\frac{\partial \lambda}{\partial \mathrm{M}_{i,j}} = v_i.w_j$ (Caswell, 1978), one gets equation 3.21.

In the specific case of model $\mathbf{M}$ of this chapter (simplified by considering only $het = 1$ genotype), the knowledge of the specific stochastic processes driving all transitions from $i = (a,p)$ allows to further develop the σ_{d}^2 formula. 3 transitions are possible from i. The fertility transition is towards $(1,1)$ via Bernoulli process $\mathcal{F}_i$ of parameter f_i. The first possible survival transition is towards $(a+1, p+1)$ via the product of Bernoulli processes $\mathcal{F}_i$ and $\mathcal{S}_i$ of parameter s_i. Because these processes are independent (the realization of $\mathcal{F}_i$ only affects later survival, via *parity*), the r.v. product $\mathcal{F}_i.\mathcal{S}_i$ is itself a Bernoulli process of parameter $f_i.s_i$. The second survival transition is towards $(a+1, p)$ via Bernoulli process $\mathcal{F}_i(1-\mathcal{S}_i)$ of parameter $(1-f_i).s_i$. And thus we can compute the covariances between all 3 process using the property that all moments of a Bernoulli process are equal to its parameter and that $\mathcal{F}_i$ and $\mathcal{S}_i$ are independent :

$$\begin{cases}
\mathrm{Cov}_d(M_{(1,1),(a,p)}, M_{(1,1),(a,p)}) = \mathrm{Var}(\mathcal{F}_i) = f_i(1-f_i) \\
\mathrm{Cov}_d(M_{(a+1,p+1),(a,p)}, M_{(a+1,p+1),(a,p)}) = \mathrm{Var}(\mathcal{F}_i\mathcal{S}_i) = f_i s_i(1-f_i s_i) \\
\mathrm{Cov_d}(M_{(a+1,p),(a,p)}, M_{(a+1,p),(a,p)}) = \mathrm{Var}(\mathcal{F}_i(1-\mathcal{S}_i)) = (1-f_i)s_i(1-s_i-f_i s_i) \\
\mathrm{Cov_d}(M_{(1,1),(a,p)}, M_{(a+1,p+1),(a,p)}) = \mathrm{Cov}(\mathcal{F}_i, \mathcal{F}_i\mathcal{S}_i) \\
\qquad = \mathrm{E}(\mathcal{F}_i^2\mathcal{S}_i) - \mathrm{E}(\mathcal{F}_i).\mathrm{E}(\mathcal{F}_i\mathcal{S}_i) = f_i s_i - f_i f_i s_i = (1-f_i)f_i s_i \\
\mathrm{Cov_d}(M_{(1,1),(a,p)}, M_{(a+1,p),(a,p)}) = \mathrm{Cov}(\mathcal{F}_i, (1-\mathcal{F}_i)\mathcal{S}_i) \\
\qquad = \mathrm{E}(\mathcal{F}_i(1-\mathcal{F}_i)\mathcal{S}_i) - \mathrm{E}(\mathcal{F}_i).\mathrm{E}((1-\mathcal{F}_i)\mathcal{S}_i) = -(1-f_i)f_i s_i \\
\mathrm{Cov_d}(M_{(a+1,p),(a,p)}, M_{(a+1,p+1),(a,p)}) = \mathrm{Cov}(\mathcal{F}_i\mathcal{S}_i, (1-\mathcal{F}_i)\mathcal{S}_i) \\
\qquad = \mathrm{E}(\mathcal{F}_i(1-\mathcal{F}_i)\mathcal{S}_i^2) - \mathrm{E}(\mathcal{F}_i\mathcal{S}_i).\mathrm{E}((1-\mathcal{F}_i)\mathcal{S}_i) = -(1-f_i)f_i s_i^2
\end{cases} \tag{3.22}$$

Integrating these covariances (equations 3.22) into the general MPPM formula for demographic variance (eq. 3.21) yields σ_{d}^2 for each genotype of (*age-parity-environment*)-MPPM $\mathbf{M}$: :

$$\sigma_{\mathrm{d}}^2 \approx \lambda^{-2}\sum_{(a,p)} w_{a,p}[v_{1,1}^2 f_i(1-f_i) + v_{a+1,p+1}^2 f_i s_i(1-f_i s_i) + v_{a+1,p}^2(1-f_i)s_i(1-s_i-f_i s_i) + \\ v_{1,1}v_{a+1,p+1}(1-f_i)f_i s_i - v_{1,1}v_{a+1,p}(1-f_i)f_i s_i - v_{a+1,p+1}v_{a+1,p}(1-f_i)f_i s_i^2] \tag{3.23}$$

The averaging of all such quantities over all genotypes $g \in \mathcal{G}$, weighted by the offspring classes abundances $\boldsymbol{w}^{\diamond}$ and all environments $e \in \mathcal{E}$, weighted by their time distribution, yields the demographic variance of our model. This demographic variance formula, even though analytic and thus easier to analyze than a simulation result, is still an approximation. It reduces, for instance, all stochastic processes to their covariances notwithstanding fundamental differences like the mutual exclusivity of the two survival transitions from a given state. Still it allows to compute and ponder the effect of individual stochasticity on the stochastic growth rate and to compare it to the effect of environmental variance.

Computation of environmental variance for MPPMs

Environmental variance can be computed in a similar fashion. We shall explicitly here consider the model (see section 3.2.1) for which vital rates are independent from age: in the mean environment, $f(a,p) = f_p = f(1 - p/\omega)$ and $s(a,p) = s_p = s(1 - p/\omega)$. For simplicity, let us reduce $\mathcal{E}$ to two environments, $\mathcal{E} = \{g, b\}$. Environment g is frequent and good (and therefore close, in its effects on vital rates, to the average environment). Environment b is very rare - appearing randomly with probability ϵ ($\epsilon \ll 1$) - and is so bad that the Fluctuating Capital is empty and therefore reproductive effort is impossible. We have $f_g = f(1+\epsilon)$, $s_g = s$, $f_b = 0$ and $s_b = s$.

In both matrices of the full model - $\{\mathbf{M^g}, \mathbf{M^b}\}$ - physiological costs of reproduction are at play, and thus

$$\begin{cases} f_p^g = f_g(1 - p/\omega) = f(1+\epsilon)(1 - p/\omega) \\ s_p^g = s_g(1 - p/\omega) = s_p \\ f_p^b = 0 \\ s_p^b = s_b(1 - p/\omega) = s_p \end{cases}$$

Matrixwise this means:

$$\begin{cases} \mathrm{M}^g_{(a,p)\to(1,1)} = f_p^g \\ \mathrm{M}^g_{(a,p)\to(a+1,p+1)} = s_p^g f_p^g = s_p f_p^g \\ \mathrm{M}^g_{(a,p)\to(a+1,p)} = s_p(1 - f_p^g) \\ \mathrm{M}^b_{(a,p)\to(1,1)} = 0 \\ \mathrm{M}^b_{(a,p)\to(a+1,p+1)} = 0 \\ \mathrm{M}^b_{(a,p)\to(a+1,p)} = s_p \end{cases} \tag{3.24}$$

Let us now check, that the assumption, made earlier, that $\mathbf{M}$ is the mean environment matrix is correct. By definition, the mean environment model is $\mathbf{M^{\bar{e}}} = \epsilon\mathbf{M^b} + (1-\epsilon)\mathbf{M^g}$, which implies

$$\begin{cases} \mathrm{M}^{\bar{e}}_{(a,p)\to(1,1)} = (1-\epsilon)f_p^g = (1-\epsilon)f(1+\epsilon)(1 - p/\omega) \approx f_p \\ \mathrm{M}^{\bar{e}}_{(a,p)\to(a+1,p+1)} = (1-\epsilon)s_p f_p^g \approx f_p s_p \\ \mathrm{M}^{\bar{e}}_{(a,p)\to(a+1,p)} = (1-\epsilon)s_p(1 - f_p^g) + \epsilon s_p = s_p(1 - (1-\epsilon)f_g^p) \approx s_p(1 - f_p) \end{cases}$$

And thus we have demonstrated that $\bar{\mathbf{M}} \approx \mathbf{M}$

From equation system 3.24, we can now calculate, for $\mathcal{E} = \{\}, \lfloor\}$, all the environmental covariances of matrix entries :

$$\begin{cases} \mathrm{Var_e}(\mathrm{M}_{(a,p)\to(1,1)}) = (1-\epsilon)(f_p^g.f_p^g) - [(1-\epsilon)f_p^g]^2 \approx \epsilon f_p^2 \\ \mathrm{Var_e}(\mathrm{M}_{(a,p)\to(a+1,p+1)}) \approx \epsilon(f_p s_p)^2 \\ \mathrm{Var_e}(\mathrm{M}_{(a,p)\to(a+1,p)}) \approx \epsilon(f_p s_p)^2 \\ \mathrm{Cov_e}(\mathrm{M}_{(a,p)\to(1,1)}, \mathrm{M}_{(a',p')\to(1,1)}) \approx \epsilon f_p f_{p'} \\ \mathrm{Cov_e}(\mathrm{M}_{(a,p)\to(1,1)}, \mathrm{M}_{(a',p')\to(a'+1,p'+1)}) \approx \epsilon f_p f_{p'} s_{p'} \\ \mathrm{Cov_e}(\mathrm{M}_{(a,p)\to(a+1,p+1)}, \mathrm{M}_{(a',p')\to(a'+1,p'+1)}) \approx \epsilon f_p f_{p'} s_{p'} s_{p'} \\ \mathrm{Cov_e}(\mathrm{M}_{(a,p)\to(a+1,p)}, \mathrm{M}_{(a',p')\to(a'+1,p')}) \approx \epsilon f_p f_{p'} s_p s_{p'} \\ \mathrm{Cov_e}(\mathrm{M}_{(a,p)\to(1,1)}, \mathrm{M}_{(a',p')\to(a'+1,p')}) \approx -\epsilon f_p f_{p'} s_{p'} \\ \mathrm{Cov_e}(\mathrm{M}_{(a,p)\to(a+1,p+1)}, \mathrm{M}_{(a',p')\to(a'+1,p')}) \approx -\epsilon f_p f_{p'} s_p s_{p'} \end{cases} \tag{3.25}$$

Integrating these covariances (equation 3.25) in the general formula for environmental variance (equation 3.17) whilst using eigenvectors to measure sensitivies ($\frac{\partial\lambda}{\partial \mathrm{M}_{i,j}}$) allows to generate a closed-form analytic formula for σ_e^2. The method can obviously be extended to any number of environments, genotypes and to age-dependent vital rates.

3.4 Results

Equipped with a battery of tools, we can investigate the effect of *physiological costs of reproduction* on the dynamics and the evolutionary demography of a (possibly heterogeneous) population. With regards to research investigation methods, this entails to combine two toolboxes. First, the various methods to yield, for the two main fitness measures, their expectancy and variances. Some of these methods are generic and well-known. Some, especially when related to multitrait models, are new tools develloped in section 3.3. Second, the *Trait Level Analysis* which provides an evolutionary-demography neutral framework as it generates, by *folding*, asymptotically-equivalent matrices. Thanks to both tools, it is now possible to compare these various fitness measures between equivalent models either implementing the costs or not, either embedding heterogeneity others not.

First we describe the mechanical effect of these costs on the mean vital rates of a population, showing they certainly influence the shapes of survival and fertility curves in nature. Second, we investigate the effects of costs of reproduction on selection gradients in order to understand how such a physiological trade-off can have evolutionary consequences through its effects on the force of selection for certain vital rates. We also study the effect of heterogeneity on the detectability of the phenomenon. Third, we investigate the effects of the *physiological costs of reproduction* on the variance of reproductive success and contemplate their consequences on effective size. And finally we both study, both formally in the probabilistic framework, and via a simulation, the effects of the costs on environmental and demographic variance.

3.4.1 Mechanical effects of costs and heterogeneity on aggregated vital rates

The most immediate repercussion of *physiological costs of reproduction*concerns the vital rates by age of an age-structured population. In most models for structured populations used by empiricists, the individuals are categorized by one trait only, mainly *age*, sometimes *stage*. There are interesting arguments regarding the benefits of each kind of model, but in reality one single trait, even if clearly explanatory, cannot generally capture more than the simplest characteristics of a population dynamics nor segregate the organism into groups of individuals with very similar vital rates. Conversely, the particular shapes of fertility and survival rates curves by age, that are the building blocks of age-structured models are likely to be influenced by these other determining traits that are not implemented. The simplest way to illustrate this is to consider models where vital rates are actually independent from age as we have done in section 3.2.

To illustrate this, let us consider the (*age-parity*)-MPPM $\mathbf{C}$ (formula 3.1) built in section 3.2.1 that corresponds to a 3-year model with zero-parity fertility and survival rates f and s, with implementation of *physiological costs of reproduction*. Folding $\mathbf{C}$ over *parity* yields $\mathbf{C}^{\mathbf{fold}}_{\mathbf{age}}$, the reference Leslie matrix, representing the ergodic-equivalent population model with only *age* as a trait :

$$\mathbf{C}^{\mathbf{fold}}_{\mathbf{age}} = \begin{bmatrix} 0 & f & f.(1-\frac{f}{2}) \\ s & 0 & 0 \\ 0 & s & 0 \end{bmatrix}$$

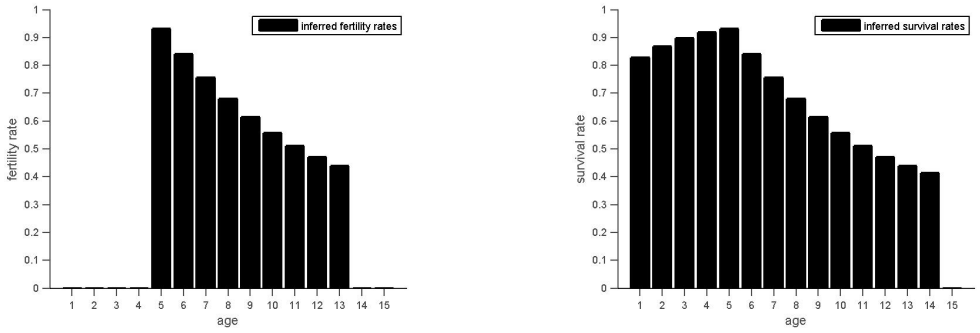
Though originally *age*-independent, fertility now decreases with age, in the population now only characterized by age. This the work of *physiological costs of reproduction*, happening in the population but not implemented any more.

Obviously, when implemented, *heterogeneity* would also have effects, and combine with implemented trade-offs to shape the aggregated vital rates for populations studied by *age* only. To understand this, we generate an (*age, parity, heterogeneity*)-MPPM $\mathbf{M}$ for a population with 2 genotypes, one robust with zero-parity vital rates of 0,9 and one frail with vital rates at 0,55. Vital rates are independent of age, and only depend on the genotype and on parity (model of section 3.2.2). Then we generate the reference Leslie matrix $\mathbf{M_a}$ by folding $\mathbf{M}$ on *parity* and *heterogeneity*, and we observe its (inferred) fertility and survival rates by age depicted in figure 3.1.

Here again, once aggregated by *age*, vital rates fluctuate with age. The familiar shape of these curves are due to the cross-effects of *physiological costs of reproduction* and *heterogeneity*. First, once age at maturity is reached, the proportion, in each genotype, of low parity individuals diminishes, and thus so do the mean vital rates with age after maturity. We can observe this phenomenon on both fertility rates (figure 3.1a) and survival rates (figure 3.1b). Second, heterogeneity generates an opposite effect. As frail individuals survive less, and therefore die sooner, their proportion is the population decreases with time. And thus the mean vital rates in the population correspond increasingly to the vital rates of the robust genotype as age increases. The heterogeneity effect is smaller than the parity effect, and is therefore easier

observable in survival rates at young ages, before cost of reproduction kicks in, where they increase until maturity (left part of figure 3.1b) and at old ages where heterogeneity seem to generate a mortality plateau (right part of figure 3.1b). The link between heterogeneity and mortality plateaus is actually a topical issue in the field of senescence both for theorists (Charlesworth and Partridge, 1997; Missov and Vaupel, 2015) and empiricists (Drapeau et al., 2000; Vaupel et al., 1998; Chen et al., 2013).

The argument, here, is not to claim that shapes of vital rates curves in nature are single-handedly produced by *physiological costs of reproduction* in a context of heterogeneity. Actually, the effects on potential but unaccounted for physiological and genetic trade-offs on vital rates are well known. We however wish to stress the importance of MPPMs and *folding* process, in order to measure the pressures exerted by these unimplemented traits on the *age* (or *stage*) demographics of population.



(a) (b)

Figure 3.1: fertility and survival rates of $\mathbf{M_a}$, the *folded* reference Leslie matrix of $\mathbf{M}$ the (*age,parity,heterogeneity*)-model with one genotype with zero-parity fertility and survival rates of 0.95, and a second genotype with zero-parity fertility and survival rates of 0.55. Mutation rate m is 0.3, maximum age $\omega = 15$ age-at-maturity $\alpha = 5$ and last reproductive age $\beta = 13$

3.4.2 Effects of *physiological costs of reproduction* on selection gradients

We can now use the tools developed and discussed in section 3.3.1 in order to quantify the effect of *physiological costs of reproduction* on selection gradients, measured as elasticity of ergodic growth rate to a fertility rate. As we incorporate physiological costs of reproduction, only the realization of fertility events has relative effects on fitness, not the realization of survival events; see appendix section 5.2.1 for a discussion on that subject. As hinted at in chapter 2 (p.41), the first important evolutionary consequence of *physiological costs of reproduction* is the drastic reduction in selection gradients, especially at maturity. Let us illustrate this, with a simple model implementing the costs of reproduction in an homogeneous population. The selection gradient for such a population modeled by an (*age,parity*)-MPPM are provided in figure 3.2.

Mechanical explanation

The contraction of selection gradients at all ages between $\mathbf{M}$, the model with *physiological costs of reproduction* and $\mathbf{M_a}$, its asymptotically-equivalent model, where the costs have disappeared by *folding*, results from a mechanical *buffering* effect. To understand this, let us consider the emergence of a new allele in the population, which bearers have, everything else being equal, a slightly higher fertility rate at maturity. These individuals will be obviously be fitter than the host, but their increased fertility at maturity causes them, on average, to reach higher parities faster and thus have weaker late reproduction than other individuals. As *physiological costs of reproduction* buffer the impact of successful reproduction by promoting those individuals which have so far not been able to recruit efficiently, the force of selection on fertility is much reduced when these costs are present.

Since this compensating effect needs time to act, the effect of the costs on selection gradients is maximum at maturity and decreases with age. At the last reproductive age, a failed reproductive event will have no more beneficial impact than a successful one, as the parity effect has run out of reproductive time to be able mitigate the damage. This can be illustrated by the distribution of the time an individual is expected to spend in the different parity classes during its entire lifetime. In figure 3.3, we represent the difference in these distributions between the model $\mathbf{M}$ with *physiological costs of reproduction* and $\mathbf{M_a}$ its ergodic-equivalent Leslie matrix with no trade-off. It is clear that *physiological costs of reproduction* concentrate

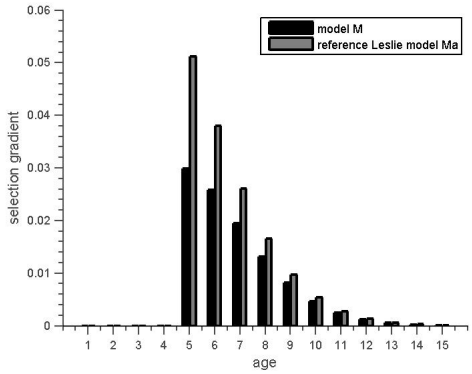


Figure 3.2: selection gradient measured by the elasticity of ergodic growth rate to fertility rates, summed by age, for $\mathbf{M}$ modeling an (*age-parity*) population with physiological costs of reproduction and $\mathbf{M_a}$ its reference Leslie matrix, which is $\mathbf{M}$ folded on *parity*, modeling the same population but characterized only by age. The population has maximum age $\omega = 15$ and age-at-maturity $\alpha = 5$. The zero-parity fertility and survival rates are 0.85. Cost of reproduction is modeled by relatively decreasing each vital rate by $1/(1+\omega-\alpha)$ per parity.

trajectories towards central parities through relative improvement of rates for low parity-individuals and deterioration for high-parity ones. In a model embedding the costs, individuals thus spend less time in 'extreme parities' than in the population without the costs.

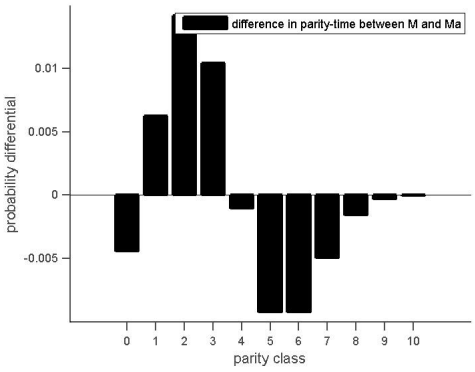


Figure 3.3: Difference between the probability distributions of the expected time spent in each parity over its entire lifetime (for an individual of age 1), for the model incorporating costs of reproduction $\mathbf{M}$ and its ergodic equivalent-Reference Leslie matrix (with no cost of reproduction) $\mathbf{M_a}$. The population has maximum age $\omega = 15$ and age-at-maturity $\alpha = 5$. The zero-parity fertility and survival rates are 0.85. Cost of reproduction is modeled by relatively decreasing each vital rate by $1/(1+\omega-\alpha)$ per parity.

Evolutionary consequences

From an evolutionary standpoint, the difference in selection gradients between a population with *physiological costs of reproduction* and its ergodic-equivalent population devoid of any trade-off, depicted in figure 3.2 leads to contemplate this result in the light of the general question about the relative role of physiological and genetic costs that we discussed theoretically in chapter 1 section 1.2. A question akin to the relative roles of the two main theories of senescence (see chapter 1 section 1.4.2 page 36).

In the age-structured and homogeneous population of a fast organism (higher fertility and mortality rates), physiological costs of reproduction have a weaker effect than for a slow one. Indeed, the physiological costs lack the temporal room for manoeuvre that allows to efficiently buffer variations in reproductive effort (see chapter 1 section 1.2.1). From a sensitivity analysis perspective, this weakness of physiological costs is reflected in the steepness of the selection gradient curve by age (the decrease of the sensitivity of λ to fertility rates by age of an age-structured organism is universal (Hamilton, 1966; Baudisch, 2005)). In a

fast organism, this curve is steeper than for slower ones, as illustrated in figure 3.4. Indeed, we know from section 3.4.2 and in particular figure 3.2 that physiological costs reduce that incline. This steepness of selection gradients will then, according to the antagonistic pleiotropy theory devised by (Williams, 1957), invite alleles with faster strategy. This mechanism will therefore (temporarily) increase the genetic variance in allocation strategy in the population: genetics costs of population would emerge. This increase in variance would be mainly unidirectional however - genotypes allocating more towards reproduction would be easily accepted in the population. This trend towards a faster mean strategy in this - now heterogeneous - population, would in turn make the physiological even weaker. As a matter of fact, the compounding over evolutionary time of this process - where steep gradients invite faster alleles with even steeper gradients - would surely lead, after after a period of augmented genetic variance, to the collapse of heterogeneity: all individuals are semelparous, a strategy on which physiological costs do not exist (see Bell, 1984).

Conversely in a slow organism, physiological costs would have a strong effect on the gradients (figure 3.4). This influence may lead to a selection gradient that is flat enough, with regard to genetic drift/selection balance, to prevent the proliferation of alleles of alternative strategies in the population. Simply put, the strength of physiological costs, caused by the slow original strategy in the population, prevents the emergence of genetic costs.

This dichotomy - when physiological costs are weak, genetic costs can emerge and vice versa - is akin to Williams (1957)'s implication that the life-history scenarios favoring the emergence of senescence via AP (steepness of gradient) may be different from the scenarios actually showing (actuarial) senescence (the negative correlations caused by physiological costs).Our analysis actually goes one step further, as it seems to imply, at first sight, that genetic and physiological costs of reproduction are mutually exclusive. *Physiological costs* prevent the emergence of genetic costs. The framework of genetic costs -the steep selection gradient - favors in time the emergence of organisms on which physiological costs have no effect.

This saddle point is however unsatisfactory from an ecological point of view. All fast organisms are not doomed to become semelparous, and within slow organisms there is inter-specific variance in slow-fast strategies that one would expect to also find intra-specifically (see Nilsen et al., 2009). The deficiency of the above approach stems, we think, from the absence, in a deterministic selection gradient analysis, of one crucial factor : stochasticity and within stochasticity, chiefly, environmental variance. As we have discussed in chapter 1 (see, for instance, section 1.4.2), we expect genetic and *physiological costs of reproduction* to be key components of the adaptability of organisms to their varying environment. They buffer the environment with differing time horizons of effects: generation time for physiological costs, evolutionary time for genetic costs. It is the environment, for instance, that can positively select slower genotypes in a population end thus, acting together with AP, enlarge the genetic variance in strategies bidirectionally. And the strength of physiological costs for slow organisms can then be interpreted as a way to compensate the absence of effect of genetic variance within individuals' long life-trajectories; in other words, as a spare environmental buffering mechanism.

As a consequence, in order to progress in our interpretation of the evolutionary consequences of selection gradient reduction by the physiological costs, we need to add stochasticity to the deterministic tool that is the selection gradient - yielding the *effective selection gradient* concept of section 3.4.3 - or to use approaches readily integrating fitness stochasticity, like the stochastic growth rate (section 3.4.4)

Effect of heterogeneity

In this section, we aim at extending the analysis of the effects of heterogeneity on selection gradients, initiated in chapter 2 (p.41). These effects are not evolutionary *per se* as selection works at the level of genotype. However, the selection gradient approach combined with the *trait level analysis* generate an adequate framework to study the potentially deleterious effect of genotypic polymorphism on the detectability of *physiological costs of reproduction* that we have theoretical analyzed in chapter 1 (section 1.2.3 p. 22). Via *Trait Level Analysis* we can *fold* $\mathbf{M}$, the(*age-parity-heterogeneity*)-MPPM, over *heterogeneity* to generate $\mathbf{M_{a,p}}$, and then over *parity* to generate $\mathbf{M_a}$.

Overall fitness of the population is preserved by *folding* (a key property of *trait level analysis*, see chapter 2 (p.41)), and therefore we can compute the selection gradients of population modeled by $\mathbf{M_{a,p}}$ and $\mathbf{M_a}$(section 3.3.1). From these *folded* matrices also, we can infer vital rates as we did in section 3.4.1. Aware of the ambiguity of the interpretation of fertility rates in $\mathbf{M_{a,p}}$(see section 3.0.4), we decide that they are to be found on the first line of the matrix, where they are categorized by *age* and *parity*.

In order to understand the effect of heterogeneity on these gradients and on these inferred fertility rates, and therefore on the detection of the costs, we compare two heterogeneous populations (fig. 3.5). Population **A** (fig.3.5a,fig.3.5c) is constituted of genotypes $\mathbf{G_1}$ and $\mathbf{G_2}$, and population **B** (fig.3.5b and 3.5d) is constituted of genotypes $\mathbf{G_1}$ and $\mathbf{G_3}$. All three genotypes are iso-fitness and differ by their posi-

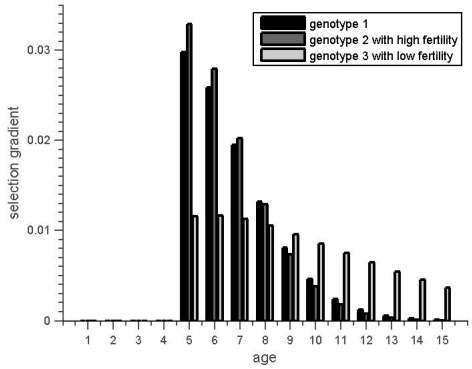


Figure 3.4: selection gradient measured by the elasticity of ergodic growth rate to fertility rates,summed by age, for three genotype models $\mathbf{G}_1$ $\mathbf{G}_2$ and $\mathbf{G}_3$ modeling an (*age-parity*) population with physiological costs of reproduction. The population has maximum age $\omega = 15$ and age-at-maturity $\alpha = 5$. The zero-parity fertility and survival rates of $\mathbf{M}_1$ are 0.85. Fertility rate is 0.98 and survival rate 0.83 for $\mathbf{G}_2$ modeling an alternative genotype with higher early fertility but reduced late fitness.Fertility rate is 0.23 and survival rate 0.98 for $\mathbf{G}_3$ modeling an alternative genotype with lower early fertility but improved survival. All three models have same fitness $\lambda = 1.055$. Cost of reproduction is modeled by relatively decreasing each vital rate by $1/(1+\omega-\alpha)$ per parity.

tion on the Slow Fast Continuum. Genotype $\mathbf{G_1}$ is central (zero-parity fertility and survival rates of 0,85 and 0,85) . $\mathbf{G_2}$ is a little faster (0,98 and 0,83) and $\mathbf{G_3}$ much slower (0,23 and 0,98). Therefore these two populations can be considered to display *physiological costs of reproduction with genetic basis*. The physiological costs are implemented by *parity*, the genetic costs by the cohabitation of different allocation strategies. With respect to that variance in allocation however, population **B** is much more heterogeneous than population **A**. For these two populations, we thus display the selection gradient (of $\mathbf{M}$,$\mathbf{M_{a,p}}$ et $\mathbf{M_a}$) on fig.3.5a-3.5b and the inferred fertility rates (for $\mathbf{M_{a,p}}$) on fig.3.5c-3.5d.

We can see on fig. 3.5a-3.5b that selection gradient for the models *folded* over *heterogeneity* ($\mathbf{M_{a,p}}$and $\mathbf{M_a}$) are higher than those of the full models, $\mathbf{M}$; with differences varying massively between less heterogeneous population **A** and more heterogeneous population **B**. This hints at the fact that, when *heterogeneity* is not taken into account, the costs of reproduction detected are smaller (population **A**) than reality (selection gradient for$\mathbf{M_{a,p}}$is in between those for $\mathbf{M}$ and $\mathbf{M_{a,p}}$) but can also be reversed for very heterogeneous populations (population **B**) as can be seen from the selection gradient for $\mathbf{M_{a,p}}$ being even higher than for $\mathbf{M_a}$.

This is confirmed by comparing inferred fertility rates by parity of $\mathbf{M_{a,p}}$ for population **A** (fig. 3.5c) and population **B** (fig. 3.5d). Whilst at the level of each genotype the fertility-parity relationship is, by construction, exactly linear: fertility rates are independent from age and decrease linearly from the zero-parity rate to zero. At the level of the populations, this relationship generally seems to hold for population **A** with however slightly weaker observed costs of reproduction (fertility rates decrease less for age classes 7 and 11 than for age class 15, which is almost exclusively occupied by $\mathbf{G_1}$. This is not at all the case for very heterogeneous population **B**, where observed costs of reproduction are, for some parity classes, even reversed : 11-years old have increasingly higher fertility rates as parity increases from 2 to 4. This phenomenon is due to a gradual shift in genotype distributions by age classes. The slower genotypes "realize" their (equivalent) fitness later than the fast ones, thus progressively invading the age classes.

This phenomenon is related to the deceptive effect of heterogeneity on perceived trade-offs famously brought to light by van Noordwijk and de Jong (1986), but our structured model allows to deepen this general detectability analysis. The natural extension of van Noordwijk and de Jong (1986)'s analysis to an heterogeneous population corresponds to genetic variance in acquisition strategy (Houle, 1991). In supplementary material section 5.2.2 (page 140) we display the selection gradient (fig. 5.3a) and the inferred fertility rate (fig. 5.3b) for such a population with acquisition heterogeneity. Unsurprisingly, as the fitter lineages, the super-flies of (Reznick et al., 2000), survive better, and thus progressively invade the age classes, positive correlations emerge between fertility and parity, at low and mid parities.

Our structured model allows to go further as it shows (fig. 3.5) that a similar phenomenon can derive from an heterogeneity that is purely allocative. We have indeed shown that the cohabitation of several iso-fitness genotypes (that have therefore similar acquisition capabilities) - i.e., that *genetic costs*

of reproduction - can also mask the underlying *physiological costs of reproduction*.

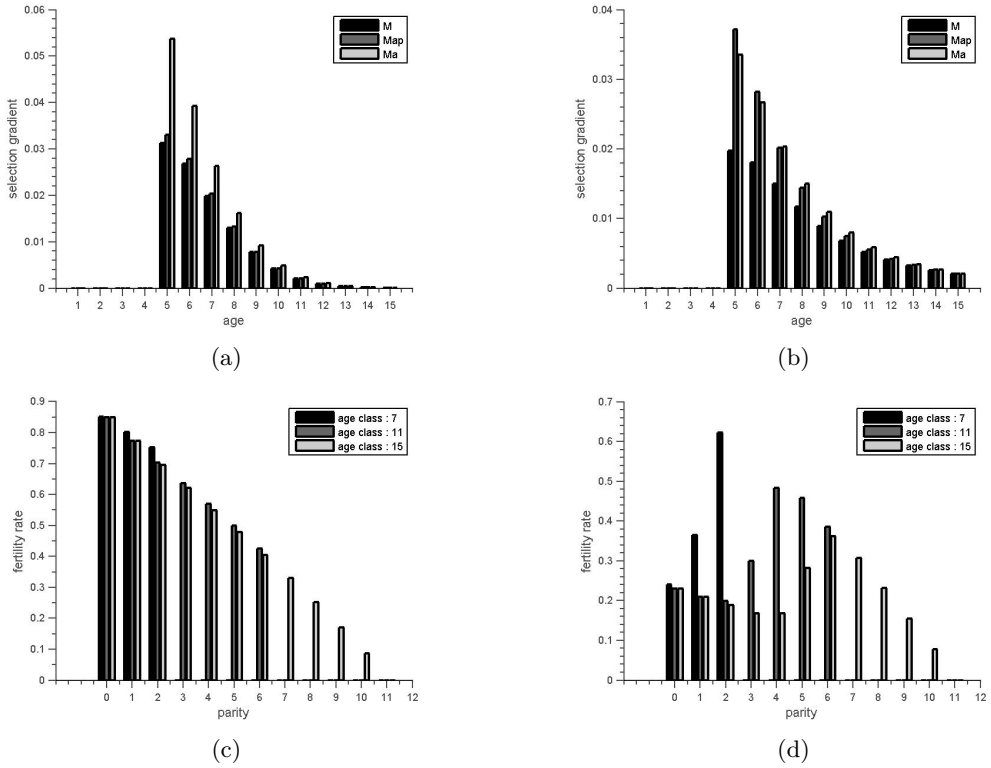


Figure 3.5: Selection gradient measured by the elasticity of ergodic growth rate to fertility rates, summed by age for population **A** (fig. 3.5a) and population **B** (fig.3.5b) and inferred fertility rates by parity for population **A** (fig.3.5c) and **B** (fig.3.5d). Population **A** is constituted of genotypes $\mathbf{G}_1$ and $\mathbf{G}_2$ and population **B** of genotypes $\mathbf{G}_1$ and $\mathbf{G}_3$. Mutation rate is $m = 0.3$. The inferred fertility rates are obtained from each model *folded* on *heterogeneity*. Both populations are modeled by (*age-parity-heterogeneity*)-MPPMs with physiological costs of reproduction, with maximum age $\omega = 15$ and age-at-maturity $\alpha = 5$. The zero-parity fertility and survival rates of $\mathbf{G}_1$ are 0.85. Fertility rate is 0.98 and survival rate 0.83 for $\mathbf{G}_2$ modeling an alternative genotype with higher early fertility but reduced late fitness. Fertility rate is 0.23 and survival rate 0.98 for $\mathbf{G}_3$ modeling an alternative genotype with lower early fertility but improved survival. All three genotypes have same fitness $\lambda = 1.055$. Cost of reproduction is modeled by relatively decreasing each vital rate by $1/(1+\omega-\alpha)$ per parity.

3.4.3 Effects of *physiological costs of reproduction* on $\sigma^2_{\mathcal{LRO}}$ and effective size

In previous section 3.4.2, we have studied the effects of *physiological costs of reproduction* on the force of selection on expected vital rates. However, the stochastic realizations of these vital rates, as they combine with the fertility-*parity* trade-off, also have repercussions on fitness, and in particular its variance. This fitness variance can either be measured by the exact calculation of the variance in lifetime reproductive output $\sigma^2_{\mathcal{LRO}}$ or by demographic variance σ^2_d the infinitesimal variance of individual contributions to growth rate. These two measures are projections in the dimension of individual variance, of the two main fitness measures $\mathbf{R_0}$ and λ and, for the same reasons mentioned in section 3.2.3, often correspond to different research areas.

In appendix section 3.0.5 section we try and reconcile these two concepts in the case of age-structured populations, and disclose their equivalence for stationary populations.

In this section, we shall focus on $\sigma^2_{\mathcal{LRO}}$ as it was until the recent advent of stochastic growth rate studies, the key measure of individual variance in evolutionary demography, and in particular with respect to its evolutionary consequences on effective population size. We will briefly discuss σ^2_d in following section 3.4.4.

However, investigating the impact of the costs on the variance of reproductive success only makes sense if the effect on the expectancy of that quantity - $\mathbf{R_0}$ - is known. In other words, if we know $\mathbf{R_0}[\mathbf{M_a}]$, $\mathbf{R_0}[\mathbf{M_{a,p}}]$ and $\mathbf{R_0}[\mathbf{M_{a,h}}]$ as functions of $\mathbf{R_0}[\mathbf{M}]$. Because, by construction λ is preserved by *folding*, this step is not necessary when investigating σ_d^2 and σ_e^2 (as we will do in next section 3.4.4), but necessary for $\sigma^2_{\mathcal{LRO}}$. Intuitively we expect $\mathbf{R_0}$ to be preserved by *folding*. This is however, in general, not the case (see supplementary material 5.2.3 page 140). However we show, in appendix 3.0.6, that, for multitrait models implementing *age*, *folding* over other traits than age does preserve the net reproductive rate, thus encompassing, in the evolutionary neutral framework of *Trait Level Analysis* both λ and $\mathbf{R_0}$. As we now know that $\mathbf{R_0}[\mathbf{M_a}] = \mathbf{R_0}[\mathbf{M_{a,p}}] = \mathbf{R_0}[\mathbf{M_{a,h}}] = \mathbf{R_0}[\mathbf{M}]$, we can now turn ourselves to the study of the effects of the costs on $\sigma^2_{\mathcal{LRO}}$.

Physiological costs of reproduction lessen $\sigma^2_{\mathcal{LRO}}$

We have just analyzed the buffering effect of *physiological costs of reproduction* with respect to the deterministic mean vital rates at the population level (in previous section 3.4.2). We have already observed that this buffering effect is also at work at the level of each individual when comparing the distribution of time spent in the different parity classes between (i) the full (*age-parity*)-model implementing the costs and (ii) the *folded* ergodic-equivalent model $\mathbf{M}^{\mathbf{fold}}_{\mathbf{parity}} = \mathbf{M}^*$ in which they are absent (figure 3.3 page 78). In this section, we will focus on this distribution at age-at-death, i.e. on the lifetime reproductive output $\mathcal{LRO}$. In particular, we will aim our attention at the second moment of this distribution, $\sigma^2_{\mathcal{LRO}}$, in order to bring to light the patterns of the effects of *physiological costs of reproduction* on individual trajectories.

In appendix section 3.0.7, we formally demonstrate (eq. 3.53) that costs of reproduction reduce $\sigma^2_{\mathcal{LRO}}$. Indeed we show that $\sigma^2_{\mathcal{LRO}}[\mathbf{M}] - \sigma^2_{\mathcal{LRO}}[\mathbf{M}^*] \leq 0$, by focusing on the parity distributions, at stable-state, in the successive age-classes in both models. If $\mathbf{M}$'s transitions are known, this result can be easier obtained by using the sensitivity matrix-based calculations of the demographic variance (when given a specific model, see supplementary material section 5.2.4) and the $\sigma_d^2\sigma^2_{\mathcal{LRO}}$ equivalence (see appendix section 3.0.5).

With this equivalence in mind, the easiest way to confirm the intuition that costs of reproduction lower the variance in reproductive output, is however, simply, to consider the general formula of σ_d^2 (equation 3.18, p.73) for both $\mathbf{M}$ and $\mathbf{M}^*$, but in a way where $\mathbf{M}$ is only characterized by *age*. In other words, this means, that instead of using the general MPPM framework, where in the full model $\{\mathbf{M}, \mathbf{s}, \mathcal{F}, \mathcal{S}\}$ where all traits are incorporated in $\mathbf{M}$, and thus the stochastic processes for different states are independent (which allowed to simply equation 3.18, p.73, into equation 3.21, p.74), the complexity of the other traits than age is tranfered from the matrix to the stochastic processes. In that framework, by properties of *Trait Level Analysis*, $\mathbf{M} = \mathbf{M}^*$. By applying eq. 3.18 to $\mathbf{M}$, and eq. 3.21 tp $\mathbf{M}^*$, we get :

$$\sigma^2_{\mathrm{d}}[\mathbf{M}] - \sigma^2_{\mathrm{d}}[\mathbf{M}^*] \approx \lambda^{-2}.N.\sum_{a_1<a_2}\frac{\partial\lambda}{\partial \mathrm{M}_{1,a_1}}\sum_{l=\{1,a_2+1\}}\frac{\partial\lambda}{\partial \mathrm{M}_{l,a_2}}.\mathrm{Cov_d}(\mathrm{M}_{1,a_1},\mathrm{M}_{l,a_2})$$

As the realization of fertility at age a_1 decreases the probability to reproduce (M_{1,a_2}) and survive (M_{a_2+1,a_2}) at age $a_2 > a_1$ (as it increases p, in both thf $f(a,p)$ and $s(a,p)$ formulas, see section 3.2.1), we get : $\sigma^2_{\mathrm{d}}[\mathbf{M}] - \sigma^2_{\mathrm{d}}[\mathbf{M}^*] < 0$. And therefore, as predicted in chapter 1, *physiological costs of reproduction* buffer individual stochasticity.

To better understand the effects of *physiological costs of reproduction* on $\sigma^2_{\mathcal{LRO}}$ we plot in figure 3.6, for a range of zero-parity fertility and survival rates (and maximum age $\omega = 5$), the difference in variance for the model with ($\mathbf{M}$) and without ($\mathbf{M}^*$) the costs. Figure 3.6a depicts the difference in variance - $\sigma^2_{\mathcal{LRO}}[\mathbf{M}] - \sigma^2_{\mathcal{LRO}}[\mathbf{M}^*]$ - and figure 3.6b the difference in coefficient of variation, $\frac{\sqrt{\sigma^2_{\mathcal{LRO}}}}{\mathbf{R_0}}[\mathbf{M}] - \frac{\sqrt{\sigma^2_{\mathcal{LRO}}}}{\mathbf{R_0}}[\mathbf{M}^*]$. For references we also plot the $\sigma^2_{\mathcal{LRO}}[\mathbf{M}^*]$ in fig. 3.6c and the iso-fitness curves (for both λ and $\mathbf{R_0}$) on the zero-parity rates map (fig 3.6d).

The first observation, is that the costs reduce variance (fig.3.6a) and coefficient of variation (fig. 3.6b) of $\mathcal{LRO}$. The first, we have discussed above and formally demonstrated in appendix 3.0.7. The second ensues from the first, as $\mathbf{R_0}$ is preserved by folding in such models (see section 3.0.6). The second observation, is that the effects of the costs on $\sigma^2_{\mathcal{LRO}}$ (fig.3.6a), follows the general shape of $\sigma^2_{\mathcal{LRO}}$ itself (fig. 3.6c), on which we shall now take a closer look.

The shape of $\sigma^2_{\mathcal{LRO}}$ as a function of zero-parity rates f and s - which exact formula, equation 3.5 (page 70), we have computed in section 3.3.2) - reveals that it results from the combined effects of three parameters. First, the variance in fertility rates at each age, $Var(\mathcal{F}^*{}_a) = f^*_a(1-f^*_a)$, which is the engine of the variance in $\mathcal{LRO}$ and confers to the latter, the $x(1-x)$ shape of the former, along the zero-parity fertility rate axis. The importance of $Var(\mathcal{F}^*{}_a)$ for late ages however requires survival (P_i in eq. 3.5), and thus the increase in $\sigma^2_{\mathcal{LRO}}$ as survival increases. At the same time, as age increases, on average f^*_a

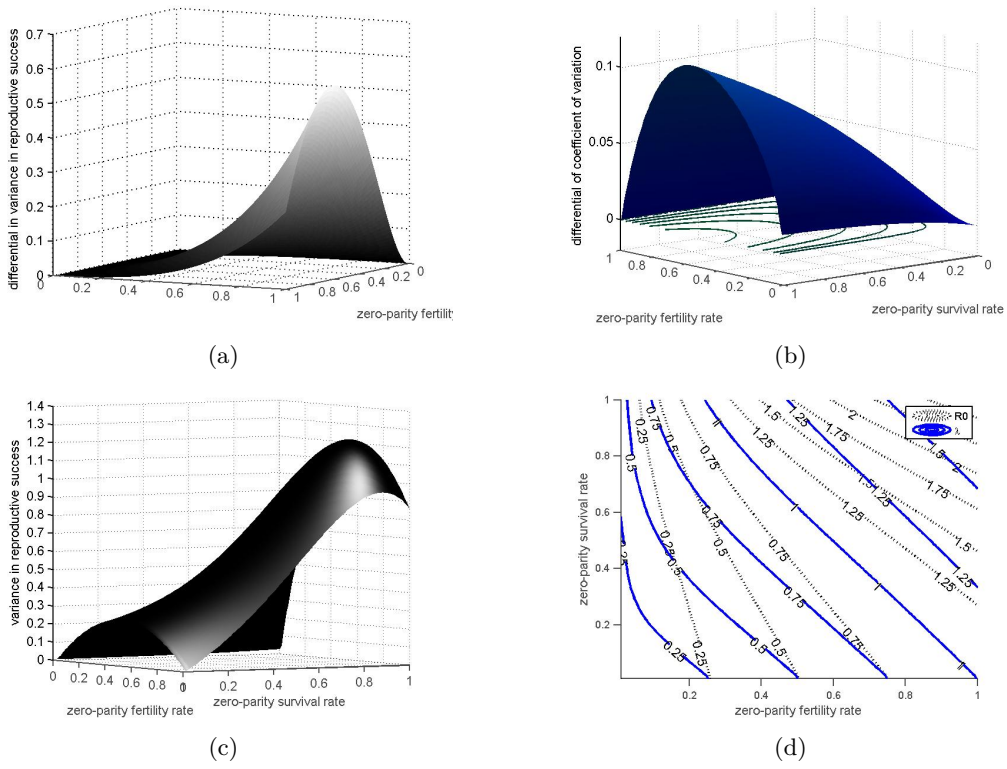


Figure 3.6: For (age-*parity*) models (differing only by their zero-parity vital rates) and their related *reference leslie matrices*, we plot the difference in variance 3.6a and coefficient of variation 3.6b of reproductive output between the model (implementing *physiological costs of reproduction*) and its its reference leslie model with no trade-off implemented. The variance in reproductive output for the reference leslie model is also displaid fig. 3.6c. The value for each combination of zero-parity vital rates of fitness measures, $\mathbf{R_0}$ and λ are represented on fig 3.6d. The population has maximum age $\omega = 5$ and age-at-maturity $\alpha = 1$. Cost of reproduction is modeled by relatively decreasing each vital rate by $1/(1+\omega-\alpha)$ per parity.

decreases because of the costs. Therefore even very high zero-parity fertility rates (conferring no variance in $\mathcal{F}_a$ at early ages) will at late ages generate variance. Hence, the asymmetrical $x(1-x)$ shape at high survival rates, across fertility rates. Finally, survival does not only act as a promoter of variance in fertility, but as a stochastic process itself - $s_i(1-s_i)$ in eq. 3.5 - which explains the decrease in the variance as survival reaches its maximum levels. These last two effects, explain why in our model, the maximum $\sigma^2_{\mathcal{LRO}}$ is attained by organisms with zero-parity (survival, fertility) coordinates of $(s = 0.94, f = 0.64)$.

This general shape - the $x(1-x)$ pattern on zero-parity fertility axis and general increase with survival- is preserved when switching from $\sigma^2_{\mathcal{LRO}}$ (fig.3.6c) to $\sigma^2_{\mathcal{LRO}}[\mathbf{M}] - \sigma^2_{\mathcal{LRO}}[\mathbf{M}^*]$ (fig.3.6a). This is due to the difference being a linear function of variance itself, as shown in equation 3.52 (appendix 3.0.7 page 3.52). However this equation also shows the difference in variance to also linearly depend on survival and fertility. This explains why the difference in variance between models with and without the costs (fig.3.6a) increases with survival even at high survival rates, and is flat at very low survival rates. These patterns are preserved when correcting for $\mathbf{R_0}$, i.e. for high fertility and survival rates, as can be observed from the differential in the coefficient of variation (fig.3.6b). Logically both the exponential increase with survival and the the asymmetrical effect for high fertilities disappear. These observations demonstrate the prediction made in chapter 1, than even though very short-lived or semelparous organisms (i.e., with $s \approx 0$) exhibit variance in reproductive success (fig 3.6c), the costs of reproduction does not affect it (figs 3.6a and (fig 3.6b)) and that the effects of the costs will increase with iteroparity/longevity.

From an evolutionary life history perspective however, considering organisms with very different fitness (the iso fitness curves for $\mathbf{R_0}$ and λ are represented on figure 3.6d) does not make a lot of sense. However, from the statistics of figure 3.6 for all possible zero-parity vital rates, we can extract the combinations that are iso-fitness. In figure 3.6d, we represent, for each possible zero-parity fertility rate, first, the corresponding zero-parity survival rate for a fitness of $\lambda \approx \mathbf{R_0} \approx 1$ (grey curve, right y-axis). Second, for each such pair of coordinates, we extract the variances in $\mathcal{LRO}$ for each model (blue curves, left y-axis), and the difference (red curve, left y-axis). This specifies the general conclusions drawn above when considering organisms that are iso-fitness (here all organisms have stationary growth rate, and therefore $\mathbf{R_0}$ worth unity). The variance in reproductive success requires both variance in fertility and survival and is therefore maximal for intermediary values of f and s, but survival is also required to promote late fertility, and this pushes s_{max} higher and therefore f_{max} lower than the point of equal coordinates ($f = 0.54$, $s = 0.54$). As expected, because of the costs of reproduction, this is less true for $\mathbf{M}$, for which s_{max} is lower and therefore f_{max} higher than for $\mathbf{M}^*$. To the contrary, the differential in variances between the two models (red curve), is maximal for the maximum possible survival rate $s = 1$ and its related zero-parity fertility-rate $f \approx 0.22$. In other words, this result shows that whilst the effect of individual stochasticity is not a monotonous function of the pace of organisms as measured by their position on the slow-fest continuum, the effects of costs of reproduction on such individual stochasticity increase with pace and are maximum for slow organisms.

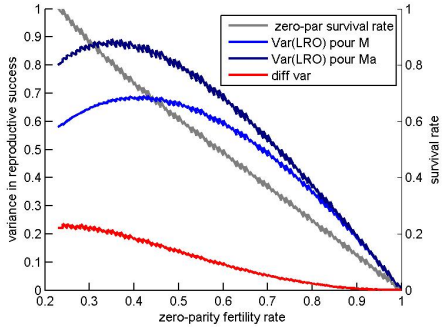


Figure 3.7: For all combinations of zero-parity vital rates yielding an ergodic growth rate $1-\epsilon \leq \lambda \leq 1+\epsilon$ (with $\epsilon = 0.01$),for each zero-parity fertility rate, the related zero-parity survival rate, and the variance of reproductive output for the model (implementing *physiological costs of reproduction*) and its its reference leslie model with no trade-off implemented and their difference;The population has maximum age $\omega = 5$ and age-at-maturity $\alpha = 1$. Cost of reproduction is modeled by relatively decreasing each vital rate by $1/(1+\omega-\alpha)$ per parity.

Adding heterogeneity into the mix raises questions with regard to the cross-effects of the costs and heterogeneity on the variance in reproductive success of the overall population and of the relative importance of heterogeneity and stochasticity as components of $\sigma^2_{\mathcal{LRO}}$ for populations with and without the costs. In appendix 3.0.8, we demonstrate that costs of reproduction and heterogeneity act independently on $\sigma^2_{\mathcal{LRO}}$ and in particular that the heterogeneity component of the variance in lifetime reproductive success is un-

affected by costs : ${\sigma^{\text{het}}_{\mathcal{LRO}}}^2[\mathbf{M}] = {\sigma^{\text{het}}_{\mathcal{LRO}}}^2[\mathbf{M}^*]$. We also show that, whilst the costs of reproduction increase the heterogeneity portion of $\sigma^2_{\mathcal{LRO}}$ (as it remains constant whilst stochasticity is reduced), it remains very low with comparison to the principal generator of demographic variance: individual stochasticity.

Demographic variance and effective size

In the field of population genetics, the effective size, N_e, corresponds to the size of an 'ideal' population yielding the same rate of genetic drift (Wright, 1931) than the 'real' population of total size N. The 'ideal' population of population genetics is a stationary population of diploid individuals with non-overlapping generations and where all individuals in the current generation have the same probability of being the parent of each individual in the next one. This characteristic implies that, in the ideal population, each of the N_e adults has probability $p = \frac{1}{N_\text{e}}$ of being parent of each of the N_e offspring. The expected $\mathcal{LRO}$ of a parent is therefore the sum of N_e independent Bernoulli processes of parameter p, that is, by the Poisson paradigm, a Poisson law of parameter $\sum_{i=1}^{N_\text{e}} p = 1$. Therefore, a variance larger (respectively lower) than 1 - or, more generally for non-stationary populations, than the mean - implies conversely a non-Poisson family size, and thus a non-random distribution of parents. This implies an increase (resp. decrease) in genetic drift rate, and therefore a smaller (resp. larger) effective size $N_\text{e} < N$.

By definition, the effective size of a population thus determines the strength of genetic drift. It also determines the increase in inbreeding coefficients and the effectiveness of selection: the fate of an allele, of selection coefficient s in a population of effective size N_e, is provided by the effective selection coefficient $s.N_\text{e}$ which combines both the effect of selection (s) and the effect of variance of reproductive success on genetic drift (through N_e). The effect of demographic variance on effective selection has also been studied directly. First, by Gillespie (Gillespie, 1974, 1975), showing the selective advantage of reduced variance in offspring number; and later many others (see, for instance, Shpak, 2005, 2007; Giaimo, 2014).

An age-structured population model - represented matrixwise by a Leslie matrix - infringe many laws of Wrightean 'ideal' populations. The population it models is made of haploid individuals, which number of offspring is not a Poisson and which generations overlap. Moreover the growth rate of such models is allowed to deviate from the stationarity of the population genetics 'ideal' population. The effective size of such age-structured populations has been extensively studied by Felsenstein (1971), Hill (1979, 1972) and Nomura (1996) but our study is not so much about the absolute effect of trait *age* on effective size as about the relative effect of trait *parity*, embedding the *physiological costs of reproduction*, and its *folding* upon, in an age-structured population.

The allele frequency variance - the engine of genetic drift - is, in an haploid population, $V_p \approx p(1-p)\sigma^2/N$. From this, Engen et al. (2005a) demonstrates that for age-structured populations the effective size can be approximated by $N_\text{e} = \frac{N}{\sigma^2_\mathbf{d}.T}$ where T is generation time. From the equivalence between $\sigma^2_{\mathcal{LRO}}$ and σ^2_d established in appendix 3.0.5, we can consider that Engen et al. (2005a)'s result is a generalization of Hill (1972)'s effective size formula for age-structured stationary populations $N_\text{e} = \frac{N\bar{b}T}{\sigma^2_{\mathcal{LRO}}}$.

Thus, in all cases, the reduction of demographic variance caused to the fertility buffering effect of *physiological costs of reproduction* implies that costs of reproduction increase effective size. In other words, the N_e of an age-structured population is underestimated when the costs are not accounted for. Compared to an asymptotically equivalent population without costs of reproduction, the population with the costs is therefore less prone to genetic drift and more to selection. This is all the more important to consider when modeling age-structured populations of small sizes. This result also implies that inbreeding coefficient will increase less in the population with *physiological costs of reproduction*. From a *kinship demography* point of view, finally, this difference between the two models in variance of family sizes hints at the fact that the entire kinship distribution in the stable-state population will be vary.

The reducing effect of the costs on selection gradient (section 3.4.2) now needs to be revisited in the light of their concurrent positive effect on selection effectiveness. To do that we devise the following effective selection measure, we call *variance-effective selection gradient*, $\frac{1}{\sigma^2_{\mathcal{LRO}}}\frac{\partial\lambda}{\partial f_a}$, that allows to compare selection gradients between models with differing $\sigma^2_{\mathcal{LRO}}$. As we can see in figure 3.8, representing the *variance-effective selection gradient* for the same ergodic-equivalent models with and without for which the absolute selection gradient were plotted in fig. 3.2, the reducing effects of the costs on demographic variance does not necessarily obliterate their weakening effects on the force of selection.

In some cases however the variance-effective selection gradient for fertility at age a is actually increased by the implementation of *physiological costs of reproduction*. To measure this effect, we have calculated for a range of models, varying in maximum age ω in $\{2, 7, 12, 17\}$, age-at-maturity α in $\{1, 5, 9\}$, in zero-parity-fertility $f(\alpha)$ in $\{0.2, 0.4, 0.6, 0.8, 1\}$ and in zero-parity-survival $s(\alpha)$ in $\{0.2, 0.4, 0.6, 0.8, 1\}$, the ratio, at α of variance-effective selection gradient between $\mathbf{M}^*$ the ergodic-equivalent model without any cost and the full model $\mathbf{M}$ with the costs, and : $\frac{\sigma^2_{\mathcal{LRO}}}{\sigma^{*2}_{\mathcal{LRO}}}\frac{\partial\lambda}{\partial f^*_a}\frac{\partial f_a}{\partial\lambda}$. Among these $4\times3\times5\times5 = 300$ models, 24 had

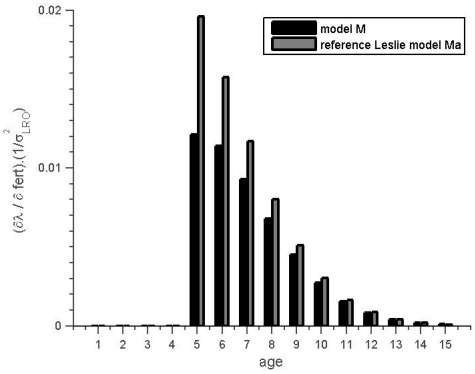


Figure 3.8: Effective size-scaled selection gradient measured by the sensitivity of ergodic growth rate to fertility rates,summed by age divided by variance of reproductive success, for $\mathbf{M}$ modeling an (*age-parity*) population with physiological costs of reproduction and $\mathbf{M_a}$its reference Leslie matrix, which is $\mathbf{M}$ folded on *parity*, modeling the same population but characterized only by age. The population has maximum age $\omega = 15$ and age-at-maturity $\alpha = 5$. The zero-parity fertility and survival rates are 0.85. Cost of reproduction is modeled by relatively decreasing each vital rate by $1/(1+\omega-\alpha)$ per parity.

a variance-effective selection gradient ratio below 1, and the 12 models with the lowest ratio values are presented (by increasing value of this ratio) in table 3.1. As we can see, all models for which the effective selection gradient is increased by the costs of reproduction correspond to long-lived organisms (high ω and maxmial zero-parity survival), with long reproductive periods (low α), and among this group to those with central zero-parity fertility (which we know, since previous section 3.4.3, maximizes the reduction of variance in reproductive success).

The mechanical explanation for such a phenomenon is simple. The buffering effect of *physiological costs of reproduction* needs individual stochasticity at each time-step, i.e. a fertility rate with maximum variance ($f(\alpha) \approx 0.5$). It also needs time (high ω, low α) and life ($s(\alpha)$) to efficiently buffer this individual stochasticity. These characteristics are reflected in the difference in demographic variance, and therefore in N_e, but is not compensated by the ratio of selection gradients. This is because, for such organisms, fertility selection gradients for the model $\mathbf{M}$ with costs are definitely very low, but not much lower that the age only model $\mathbf{M}^*$ with no costs. Indeed, the sheer longevity of these organisms already strongly buffers fertility (Morris et al., 2008). Unrealized fertility events are not postponed, via promoted future rates like in the case of the costs of reproduction, but irrevocably lost for the individual. However, the large number of these fertility events means that failure at a particular time-step is much less costly than it would be for a short lived or semelparous organism. Therefore, whilst further reducing the absolute selection gradients, the *physiological costs of reproduction* will actually overall increase *variance effective selection gradients* in long-lived organisms.

This results hints at a stabilizing role on AP - i.e., on *genetic costs* - for *physiological costs of reproduction*. Whilst in age-structured populations without costs, the selection gradient will vary massively between fast organisms - with steep gradients fast inviting even faster alleles in the population - and slow ones, which almost flat gradient incline seems to prevent any AP, we see that *physiological costs of reproduction* will smooth these differences, straightening the gradients of fast organisms and curve the gradients of slow ones. This seems therefore to put away any question surrounding the disparity of selection gradients among organisms' life pace as ill-founded. Thanks to the physiological costs of reproduction, antagonistic pleiotropy on fertility seems to work at a similar rate on all organisms, fast and slow, to which, in a constant environment, it continuously provides faster alleles better suited to benefit from this constancy.

3.4.4 Effects of *physiological costs of reproduction* on environmental stochasticity and stochastic growth rate

In this section, we shall further extend our analysis of the effects of *physiological costs of reproduction* considering how they affect the stochastic growth rate of a population. Like the *variance-effective selection gradient* of the previous section, the stochastic growth rate $\ln \lambda_s$ of a population, has deterministic component related to the mean vital rates, and stochastic components. This appears clearly in Engen's equation eq. 3.20 : $\ln \lambda_s \approx \ln \lambda - \frac{\sigma_e^2}{2} - \frac{\sigma_d^2}{2.N}$.

Table 3.1: Models with the 12 smallest values, in increasing order, for variance-effective selection gradient - $\frac{\sigma^2_{\mathcal{LRO}}}{\sigma^{*2}_{\mathcal{LRO}}}\frac{\partial\lambda}{\partial f^*_a}\frac{\partial f_a}{\partial\lambda}$ - among 300 models with zero-parity vital rates ranging from 0.2 to 1, ω ranging from 2 to 17 and α from 1 to 9

$f(\alpha)$	0,4	0,4	0,2	0,6	0,6	0,6	0,4	0,8	0,2	0,8	0,8	0,4
$s(\alpha)$	1	1	1	1	1	1	1	1	1	1	1	1
α	1	1	1	1	1	1	1	1	1	1	1	5
ω	17	12	17	12	7	17	7	7	12	12	17	17
$\frac{\sigma^2_{\mathcal{LRO}}}{\sigma^{*2}_{\mathcal{LRO}}}\frac{\partial\lambda}{\partial f^*_a}\frac{\partial f_a}{\partial\lambda}$	0,64	0,65	0,69	0,70	0,71	0,72	0,76	0,79	0,80	0,81	0,83	0,87

Remaining in the general framework of the preceding section, with an (*age-parity*)-MPPM $\mathbf{M}$ implementing the costs, and its related asymptotically-equivalent matrix, *folded* on *parity*, $\mathbf{M}^*$ in which the costs are absent, we wish to finalize the analysis by comparing $\ln\lambda^*_{\mathrm{s}}$ and $\ln\lambda_{\mathrm{s}}$.

By the principles of *Trait Level Analysis* we know that the deterministic growth rate is preserved by *folding*, $\ln\lambda^* = \ln\lambda$. Moreover from the previous section we know, that $\sigma^*_{\mathcal{LRO}} = \sigma_{\mathcal{LRO}}$, hence, from the equivalence in appendix 3.0.5, $\sigma^*_d = \sigma_d$ (the specific calculations of this difference, and that of σ^2_e for a given full model is provided in supplementary material section 5.2.4). The comparison of $\ln\lambda^*_{\mathrm{s}}$ with $\ln\lambda_{\mathrm{s}}$ therefore comes down to calculating $\sigma^*_d - \sigma_d$.

Demonstration of $\sigma^2_e[\ \mathbf{M_{age,parity}}] < \sigma^2_e[\ \mathbf{M^{fold}_{age}}]$

In this section we shall formally express the environmental variance for model with and without *physiological costs of reproduction* to calculate their difference. We will provide the calculations for an (*age-parity*)-model with 2 age classes, and 2 environments. The computations can then be extended to any number of age-classes and any number of environments.

Let p be the probability of the 'bad' environment occurring at any time, and $\mathbf{M_g}$ and $\mathbf{M_b}$ the vital rate expressions of the full model in each environment. Let them be $\mathbf{M^*_g}$ and $\mathbf{M^*_b}$ for the *folded* model. For the full model, the mean-environment matrix $\mathbf{M}$ is worth $\mathbf{M} = (1-p).\mathbf{Mg} + p.\mathbf{Mb}$ which we can also write vertically :$\mathbf{vec(M)} = (1-p).\mathbf{vec(Mg)} + p.\mathbf{vec(Mb)}$.

Now, from section 3.3.3, let us rewrite the environmental variance in matrix notation :

$$\sigma^2_{\mathrm{e}}\lambda^2 = \mathbf{1}'(\hat{\mathbf{S}} \circ \hat{\mathbf{V}})\mathbf{1} \quad , \tag{3.26}$$

where $\hat{\mathbf{S}}$ is the symmetric matrix (of size $\mathrm{q}^2 \times \mathrm{q}^2$,) of products of all pairs of sensitivities of entries of $\mathbf{M}$: $\hat{\mathrm{S}}_{(i,j),(k,l)} = \frac{\partial\lambda}{\partial \mathrm{M}_{i,j}}\frac{\partial\lambda}{\partial \mathrm{M}_{k,l}} = \mathrm{S}_{i,j}\mathrm{S}_{k,l} = v_i w_j v_k w_l$, i.e.,

$$\hat{\mathbf{S}} = \mathbf{vec(S)}.\mathbf{vec(S)}' = \mathbf{vec(v.w')}.\mathbf{vec(v.w')}' \quad , \tag{3.27}$$

and where $\hat{\mathbf{V}}$ is the environmental variance-covariance matrix of size $\mathrm{q}^2 \times \mathrm{q}^2$,

$$\hat{\mathbf{V}} = p.\mathbf{vec(M_b)}.\mathbf{vec(M_b)}' + (1-p).\mathbf{vec(M_g)}.\mathbf{vec(M_g)}' - \mathbf{vec(M)}.\mathbf{vec(M)}' \tag{3.28}$$

Then $\mathbf{M_g} = \begin{bmatrix} f & f & 0 & f/2 \\ (1-f)s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ fs & 0 & 0 & 0 \end{bmatrix}$ and $\mathbf{M^*_g} = \begin{bmatrix} f & f(1-\frac{f}{2}) \\ s & 0 \end{bmatrix}$ computed, for the first one, thanks to the parity-fertility formula of the costs $\mathrm{M}_{\mathrm{g}1,4} = f(1-\frac{1}{1+\omega-\alpha})$ and, for the second one, from the principles of *Trait Level Analysis* $\mathbf{M^*_g} = \mathbf{M^{fold}_{gage}}$ and in particular, the mean 2-year fertility is the transition-weighted 2-year fertility for both parities : $\mathrm{M}^*_{\mathrm{g}1,2} = (1-f)f + f.\frac{f}{2}$

For the sake of simplicity, bad environment is here an environment that cancels any fertility event. Then : $\mathbf{M_b} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ and $\mathbf{M^*_b} = \begin{bmatrix} 0 & 0 \\ s & 0 \end{bmatrix}$

Then from 3.28, and with factorization we get :

$$\begin{cases} \hat{\mathbf{V}} = p(1-p)f^2.\mathbf{B} \\ \hat{\mathbf{V}}^* = p(1-p)f^2\mathbf{B}^* \end{cases} \tag{3.29}$$

, with intermediary matrices $\mathbf{B}$ (eq. 3.54) and $\mathbf{B}^*$ (eq. 3.55) displayed in appendix 3.0.9.

The $\hat{\mathbf{S}}$ matrices can be expressed via the eigenvectors and eigenvectors themselves can be expressed as functions of w_1 and v_1 which are the relative abundance and reproductive value of the first *age* class of $\mathbf{M}^*$.

By construction and because there is only one parity class at *age* 1, w_1 is also the relative abundance of *age* 1, *parity* 0 class of $\mathbf{M}$. And because this particular folding is a 'perfect aggregation' (see chapter 2 (p.41)), then $w_i v_i$ is preserved for each *age* class i. In particular this means v_1 is also the scaled reproductive value of *age* 1, *parity* 0 class of $\mathbf{M1}$. Because *age* 2 individuals all come from the same state, we have $:w_{2,0} = (1-f)w_2$ and $w_{2,1} = fw_2$. and because of the perfect aggregation $:w_{2,0}v_{2,0} + w_{2,1}v_{2,1} = w_2 v_2$. But the ratio of $v_{2,0}$ to $v_{2,1}$ is the ratio of their fertilities $:v_{2,0} = 2v_{2,1}$. Thus, the perfect aggregation equation becomes, $v_{2,0} = v_2/(1-f/2)$ or $v_{2,1} = v_2/(2-f)$. Moreover, the eigen.equations yield $w_2 = sw_1/\lambda$ and $v_2 = (1 - v_1 w_1)/w_2$. And thus all 4 vectors can be expressed as functions of v_1 and w_1.

$$\begin{cases} \boldsymbol{w}^* = \begin{pmatrix} w_1 & w_1/\lambda \end{pmatrix} \\ \boldsymbol{v}^* = \begin{pmatrix} v_1 & \lambda(1 - v_1 w_1)/s w_1 \end{pmatrix} \\ \boldsymbol{w} = \begin{pmatrix} w_1 & (s/lam) w_1 (1-f) & 0 & (s/lam) w_1 f \end{pmatrix} \\ \boldsymbol{v} = \begin{pmatrix} v_1 & (2\lambda(1 - w_1 v_1)/(s.w_1(2-f)) & 0 & (\lambda(1 - w_1 v_1)/(s w_1 (2-f)) \end{pmatrix} \end{cases} \tag{3.30}$$

Then implementing these elements (eq 3.30) into eq.3.27, gives us $\hat{\mathbf{S}}$ and $\hat{\mathbf{S}^*}$ which we can combine with the expressions of $\hat{\mathbf{V}}$and $\hat{\mathbf{V}}*$ (eq 3.34) we get from overall equation 3.26:

$$\begin{cases} \sigma_{\mathrm{e}}^{*2}\lambda^2 = \mathbf{1}'(\hat{\mathbf{S}}^* \circ \hat{\mathbf{V}}^*)\mathbf{1} = p(1-p) f^2 w_1^2 v_1^2 (1 + \frac{s(1-f/2)}{\lambda}) \\ \sigma_{\mathrm{e}}^{2}\lambda^2 = \mathbf{1}'(\hat{\mathbf{S}} \circ \hat{\mathbf{V}})\mathbf{1} = p(1-p) f^2 w_1^2 v_1^2 (1 + \frac{s(1-f/2)}{\lambda} + \frac{\lambda(v_1 w_1 - 1)}{(2-f) v_1 w_1}) \end{cases} \tag{3.31}$$

This system of equation provides us with important information. First, as expected, the environmental component of the growth rate variance is directly related with the environmental variance of the time series : $p(1-p)$. Second it is also directly related to fertility itself (unsurprisingly for such an environment which effect is to cancel reproduction) and to generation time $\frac{1}{v_1 w_1}$. Most importantly, it provides

$$\sigma_e^2 - \sigma_e^{2*} = \mathbf{1}'(\hat{\mathbf{S}} \circ \hat{\mathbf{V}} - \hat{\mathbf{S}}^* \circ \hat{\mathbf{V}}^*)\mathbf{1} = p(1-p) f^2 w_1^2 v_1^2 (\frac{\lambda(v_1 w_1 - 1)}{(2-f) v_1 w_1}) \tag{3.32}$$

As $v_1 w_1 = 1 - v_2 w_2 \leq 1$, in *all* cases, the environment variance of the model with *physiological costs of reproduction* implemented is lower than the one of its equivalent Leslie matrix. Therefore, we have demonstrated that *physiological costs of reproduction* also buffer environmental variance. This effect on σ_e^2 combined with the effect on σ_d^2 and the preservation of deterministic λ, yields :

$$\ln \lambda_{\mathrm{s}}^* > \ln \lambda_{\mathrm{s}} \tag{3.33}$$

Through its effects on both σ_d^2 and σ_e^2 the *physiological costs of reproduction* improve the stochastic fitness of the population. We shall now illustrate this with a simulation.

Simulation of stochastic growth for model with and without costs

the model

In this section we will simulate the dynamics of the same initial population both via the model implementing *physiological costs of reproduction* and via its asymptotically-equivalent folded model where the costs are absent. We track a 5 year organism ($\omega = 5$) that can produce one offspring every year, from the first year of its life ($\alpha = 5$).

We will use the environmental model setup in section 3.3.3 and used in previous section 3.4.4, with a bad environment that prevents reproduction (nullifying all fertility rates but leaving survival rates identical to these of the good environment) and occurring $\epsilon = 10\%$ of the time. We want our mean zero-parity vital rates to be 0.49 for fertility (at all ages for $\bar{\mathbf{M}}$) and 0.65 for survival (at all ages for $\bar{\mathbf{M}}$), implying an asymptotic growth rate for $\bar{\mathbf{M}}$ of $\lambda \approx 1.02$.

As described in section 3.3.3, for that purpose, we set the zero-parity fertility rate of $\mathbf{M} = \mathbf{M_g}$, the "good" environment matrix at $0.49 \times (1 + \epsilon) = 0.539$. This ensures that the relative difference in all fertility rates between the desired mean matrix and the actual mean matrix is less than $\epsilon^2 = 0.01$, and thus containing also the relative difference in ergodic growth rate between the two matrices (which actually is $max_{eigs}(\mathbf{M}) - max_{eigs}(\bar{\mathbf{M}}) - = 0.004$).

In order to preserve at best the mean growth rates between the models with and without costs (an arbitrary choice), the folded models are obtained by folding their full versions according to the ergodic abundance vector of the mean matrix, incurring slight differences between the growth rates and abundances of the "good" environment matrices (0.008), but limiting the differences in mean matrices growth rates (0.0008).

Finally, we pick an initial population of 250 individuals which is, for our 5-year organism, large enough so that individual stochasticity is not too strong a driver of the behavior of each population, obliterating the influence of the environment, but small enough for its effects to be observable. The detectability is enhanced by the use of one, common, environmental time series (where the environment at each time-step is drawn at random with probabilities ϵ and $1-\epsilon$) for all populations, in which each model (with costs and without costs) will be run 5 times.

results The resulting graph is pictured in figure 3.9. It illustrates the result we have formally proven in previous section 3.4.4: the stochastic growth rate is improved for the population with the costs (blue curves) with respect to the populations without the costs (red curves). The *physiological costs of reproduction*, through they buffering effect on lifetime reproduction, limit the variance in fitness of a population, and thus improves the stochastic growth rate, increasing the population's resilience. This is due to the combined effect of the costs on demographic and on environmental variance.

The effect on individual stochasticity can be visualized by the differences of trajectories for the various seeds for each model (i.e. between the different lines of same color, as they correspond to the same model in the same environmental series). The *folded* model $\mathbf{M}^*$ (red curves) quickly generates a larger variance in population numbers between the different runs (experiencing the same environment) than $\mathbf{M}$ (blue curves); see at $t \approx 300$.

The effects on environmental stochasticity, observable over time as environments unfold (dotted line), are obvious at $t \approx 200$ where a series of bad environments has stronger effects on the population with *physiological costs of reproduction*, and then again at $t \approx 400$ where a poor and sustained environment pushes down numbers of all populations.

The combined effect of environmental and demographic variance brings about the extinction of 4 of the 5 populations with no costs implemented, between $t \approx 380$ and $t \approx 410$. The unique remaining population goes extinct later ($t \approx 760$), whereas all 5 populations with *physiological costs of reproduction* implemented survive until the end of the simulation of 1000 years. These populations obviously suffer also from environmental and demographic stochasticity: they stochastic growth rates are lower than the deterministic growth rate resulting from the mean-environment matrices (black lines; the slight difference between them due to the 0.0008 difference in growth rate discussed above accumulated over 1000 time-steps). However, they have higher resilience than the populations from the ergodic-equivalent model missing the physiological costs.

3.5 Discussion

This chapter consists in hinting at the first answers to the questions asked in chapter 1 thanks to the MPPM framework developed in chapter II. In chapter 1, we saw that physiological and genetic trade-offs, whilst having similar phenotypic manifestations (negative correlations between vital rates at different ages for costs of reproduction, for instance) have very different mechanisms with differences at many levels : level of action (individual vs population), location of mechanism (intermediate structure vs genotypic level), time window of effect (generation time vs evolutionary time). However, despite these differences, we noticed they were not mutually exclusive, and devised the mixed mechanism of *physiological costs of reproduction with genetic basis*. This possible cohabitation of genetic and physiological trade-offs implies, inter alia, that the life history scope of each of the two kind of trade-offs is not limited to the location of its mechanism. Whilst we all know that genetic trade-offs have physiological effects, promoting certain functions in certain lineages in the population and not in others, it should be equally obvious that physiological trade-offs have evolutionary consequences.

The MPPM framework allows to implement - with the addition of dynamic and hidden heterogeneity traits – both physiological and genetic trade-offs. Via the *Trait Level Analysis*, it permits to analyse the evolutionary demographic repercussions of these trade-offs by *folding* the model on the trait incorporating it. In the particular - but essential - case of costs of reproduction, such a model therefore enables the study the cross effects of genetic and physiological costs and, above all, to measure the evolutionary consequences of *physiological costs of reproduction*.

In the first section of this chapter, we start by devising a family of MPPMs able to model various populations differing in their early-life fertility and survival rates, but all encountering the same physiological trade-offs. The zero-parity vital rates, defining the life history strategy of the population, is reduced as parity increases. This simple setup allows to incorporate the key elements of *physiological costs of reproduction* as discussed in chapter 1, with trait *parity* in the role of Ratchet Capital, and Fluctuating Capital reduced to environmental effects. In case of heterogeneity in the population, i.e. if different genotypes with different life-history strategies- iso-fitness or not – cohabit in the population, we show how to integrate

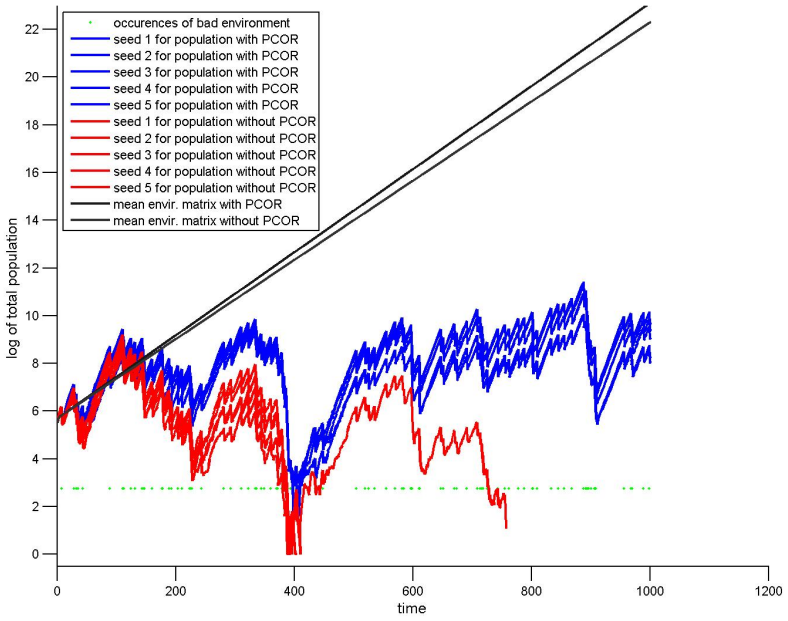


Figure 3.9: Simulations on 1000 time-steps for populations modeled by good/bad environment matrix pair ($\mathbf{M_g}/\mathbf{M_b}$) for populations with *physiological costs of reproduction* and ($\mathbf{M_g^*}/\mathbf{M_b^*}$) for populations without *physiological costs of reproduction*. Environmental series has been computed once for all simulations (environment is drawn at random each time-step, with probability $\epsilon = 0.1$ to be bad). Then the simulation has been run 5 times for each matrix pair. In all runs, initial population is 250 newborns. All matrices model populations with maximum age $\omega = 5$ and age-at-maturity $\alpha = 1$. $\mathbf{M_b}$ (model used in good environment) and $\mathbf{M_g}$ (model used in bad environment) are (*age-parity*)-MPPM with *physiological costs of reproduction*. Their zero-parity fertility are 0.5390 and 0 and the zero-parity survival rate is 0.65. To these models correspond, for the populations incurring no costs of reproduction, $\mathbf{M_b^*}$ (model used in good environment) and $\mathbf{M_g^*}$ (model used in bad environment), which are $\mathbf{M_b}$ and $\mathbf{M_g}$ *folded* over *parity*, along the ergodic abundance vector of $\bar{\mathbf{M}} = \epsilon\mathbf{M_g} + (1-\epsilon)\mathbf{M_b}$

these elements to generate an (*age-parity-heterogeneity*)-MPPM that can incorporate *physiological costs of reproduction*, *genetic costs of reproduction* and heterogeneity in fitness in the population.

In the second section, we describe current tools -for one trait model - and come up with new ones – for multitrait models – to yield the key fitness measures of a population. First, whilst λ is preserved by *folding*, this is not the case for its elasticity to fertility rates, the selection gradients, that we show how to compute for a multitrait model. Second, as $\mathbf{R_0}$ is also preserved by *folding* such models (this is shown in appendix 3.0.6), we show how to compute its second moment, the variance in lifetime reproductive out $\sigma^2_{\mathcal{LRO}}$ for age-structured populations (closed-form formula). We then formulate the use of *Markov chains with rewards* in order to compute $\sigma^2_{\mathcal{LRO}}$ for memory models and in particular for models embedding an active *dynamic heterogeneity* trait (to implement a trade-off for instance) like *parity* in our family of models. Finally we show the steps to compute demographic and environmental variance for multitrait models.

We can now combine these fitness measures tools with the properties of *Trait Level Analysis* that allow to *fold* a model implementing the costs over *parity* to generate an asymptotic-equivalent model from which the costs are absent, to measure the demographic and evolutionary effects of *physiological costs of reproduction*. First, by extracting the vital rates from the Leslie reference matrix of an (*age-parity-heterogeneity*)-MPPM, we show the mechanical role played by physiological and genetic costs of reproduction on the shape of mortality and fertility curves represented by *age* only.

Second, we show that *physiological costs of reproduction* reduce the fertility selection gradient. This result highlights the buffering effect of the costs on fertility stochasticity (described in chapter 1) whereby the increase of realization probability of a particular fertility event would have limited positive effect on λİndeed, it would set individuals on higher parity trajectories with therefore reduced future vital rates. By considering the selection gradients of the matrices folded on *parity*, *heterogeneity* and both, we can

extend the results from van Noordwijk and de Jong (1986) and Houle (1991) and show that the variance in allocation strategy itself might prevent detectability of the allocative costs.

Surprisingly, the overall decrease in selection gradients caused by *physiological costs of reproduction*, seems to imply, contrary to our expectations from chapter 1, that physiological and genetic costs might not cohabit in organisms. Physiological costs, if present, would flatten the selection gradient, preventing the invasion of different life-history strategies in the population, whilst if absent the steep curve of the gradients would, according to Williams (1957)'s antagonistic pleiotropy theory, invite faster alleles in the population.

This puzzling result is contrasted by the analysis of the effects of *physiological costs of reproduction* on the variance in reproductive success, which strongly reduced by the costs as expected from chapter 1. We show moreover that this effect is stronger for long-lived individuals, for which, even along an iso-fitness continuum, the decrease in variance caused by the costs is maximal. Via the role played by the variance in reproductive success on the effective size of a population, this results has important consequences on efficient selection. This prompts us to revisit the results stemming from the calculation of the absolute selection gradient and introduce the *variance-effective-selection gradient* that scales the absolute selection gradient by the inverse of $\sigma^2_{\mathcal{LRO}}$. When considering this efficient selection measure, we realize that, whilst the net effect of the costs is still a reduction for most models used in this chapter, for very slow organisms the *variance-effective-selection gradient* is increased by the *physiological costs of reproduction*. This hints at the fact that the seemingly very large differences in gradients between short (very steep gradients) and long-lived (much flatter gradients) organism, when structured and analysed by *age* (only), with its corollary exponential invasion of faster alleles for already fast organisms, may be an artifact of models forsaking the implementation of trade-offs.

In the following section, we consider another measure that, like the *variance-effective-selection gradient*, combines both the deterministic and the stochastic effects of the costs: the stochastic growth rate. We know that *Trait Level Analysis* preserves λ. We showed that costs reduce $\sigma^2_{\mathcal{LRO}}$ it therefore reduces demographic variance by their equivalency (appendix 3.0.5). In that section, by formally proving that *physiological costs of reproduction* also buffer environmental variance, we establish that they, in all cases, increase the stochastic growth rate. To illustrate this, we have simulated the fate of populations modelled by the two ergodic-equivalent matrices with and without the costs, in the same environments. The effect of the costs on environmental variance can be observed by comparing the populations between the two models. The effect of the costs on demographic variance can be seen by comparing the variance in trajectories within the populations (sharing the exact same mean parameters) of the same model. The combination of both effects provides a better resilience to the model with *physiological costs of reproduction*.

This relatively simple MPPM model therefore yields several important results, in the light of senescence theories. It highlights the general buffering effects of the Disposable Soma Theory on environmental and demographic variance and its beneficial impact on stochastic fitness. However it also shows that the allocation process at the core of DST is not the only mechanism able to buffer life-history. The sheer longevity of certain organisms, and by then the spreading of their reproductive schedule, already reduces the importance of single fertility events. And therefore the Disposable Soma theory seems to act as a regulator of the strength of the Antagonistic Pleiotropy theory, which, in turn, may hint at a combined role in inviting and filtering the variance in life history strategies in a population.

From a methodological point of view, this chapter also highlights the strong limitation of life-history models structured by age (only). It is often argued that these models are ideal to study populations for which age is the main determinant of life history. However this paradigm, we have shown here, has to be moved beyond. First because trade-offs are a key component of life history and that, to be implemented, a trade-off needs at least two traits to lean on. Second, because age is only the best predictor of vital rates as it already encompasses the trade-offs. We showed, in section 3.4.1, how a model where vital rates do not depend on age, can seem to be strongly age-driven when folded on the physiological and genetic trade-offs. Whilst using only one trait, partially incorporating, by linkage the underlying trade-offs (and thus seemingly life history determinant) will generally provide appropriate results with respect to population dynamics and demography, it will generate poor results from an evolutionary demography viewpoint. This can be readily ascertained by any empiricist when comparing the variance in reproductive success inferred by the Leslie matrix of her/his studied organism with the actual number. We show, in this chapter, that this discrepancy is only one in many consequences of interpreting age-structured models without understanding the effects of the simplification. The effect on selection gradients is another. It could therefore be argued that the addition of a 2nd trait on an age-structured model is as key to understand the life history of an organism from an evolutionary perspective, than the addition of the 1st trait (age) on a non-overlapping generation model is from a demographic viewpoint. In general, this prompts us to revisit general results stemming from one trait analyses. For instance, Charlesworth (1980) demonstrated that the age-structure of population has little impact on their population genetics. Would that result hold when a 2nd trait, implementing a constraint, is added to the age-structured model ?

In this chapter, we implement key elements of costs of reproduction as identified in chapter 1. However, for simplicity, many specific components, the combination of which make up the vast diversity of costs of reproduction in nature, were left aside. This is the case, for instance of the storage capacity, *stor*, which positions organisms on the income-capital breeding continuum according to their ability to save some of the resources they acquire from the environment (forsaking *stor* in our model, allows to implement FC from the environmental matrices directly). This the case also of the reproductive effort schedule *res* which represent the time distribution of effort required to produce an independent offspring. In the model of this chapter, we equate reproductive effort and fertility event. However, in most organisms, reproductive efforts start before birth (e.g., mating, gestation) and continue after (e.g., lactation). Post-natal care can be protracted, especially in social species where it takes the name of parental care. We hinted, in chapter 1 (section 1.3.2 page 34), at a way to implement an extended reproductive effort schedule in general, and parental care in particular, by adding extra *dynamic heterogeneity* traits to the model trait structure. These "buckets" would segment the reproductive effort schedule, and only when the last "bucket" is filled would a new independent offspring be deemed to appear in the population. Further extending parental care to kinship care is a natural step. When an adult male human takes time and spends money to care for his grand-children, his nephew or his pregnant sister, he can be considered to be promoting the future reproductive value of his genotype, at the cost of current reproduction. Stretching Williams (1966)'s definition of the costs of reproduction to include kinship care may be frowned upon, but the analogy of principles implies an analogy in potential models that we shall evoke in the next chapter.

Whilst it seems that kinship may be added as an input to the evolutionary model framework of MPPMs, section 3.4.3 shows us that it is an (overlooked) output of any structured model. We showed here the effect of the costs on variance of reproductive success, or family size as Hill (1972) calls it, but the kinship consequences of the structure of a population extend far beyond that measure to include all distributions of kin. The combined study of the kinship input and output of structured populations, we call kinship demography, will be the topic of the next short chapter.

3.6 Appendices

3.0.1 Note on the number of life trajectories inferred by a Leslie matrix

For an age-structured population, with maximum age ω, and all fertility f and survival rates s strictly comprised between 0 and 1 with the maximum of number of offspring per time-step at 1 (i.e. $\mathcal{F}$ is Bernoulli) we can compute different individual lifetime trajectories. Survival trajectories are blind to fertility and therefore only differ by age-at-death. Reproductive trajectories consist of the sequence of reproductive realization over life, but are blind to longevity. Finally lifetime trajectories, *per se*, consider both survival and fertility events and therefore individuals of the same lifetime trajectory have both the same survival *and* reproductive trajectories. For such a Leslie matrix, there are therefore ω possible survival trajectories corresponding to all longevities between 1 and ω. For the survival trajectory of length i, $\quad 1 \leq i \leq \omega$, there are as much as i possible fertility events, hence 2^i possible different sequences. In total therefore, there are $\sum_{i=1}^{\omega} 2^i = 2^{\omega+1} - 2$ different individual lifetime trajectories allowed by the model.

3.0.2 Computation of $\sigma^2_{\mathcal{LRO}}$ for an age-structured population

In an *age*-structured population, with n age classes (rendered by a $n \times n$ Leslie matrix), let fertility process at age i be $\mathcal{F}_i$ of expectation f_i and survival process $\mathcal{S}_i$ of expectation s_i (then $s_n = 0$). Let $P_i = \prod_{k=1}^{i-1} s_i$ be the probability to survive to age i (then $P_{n+1} = 0$ and we let $P_1 = 1$). Then let us define

$$y_i = \frac{1}{P_i} \sum_{j=i}^{n} f_j P_j \tag{3.34}$$

which represents the expectation of $\mathcal{LRO}_\text{i}$, the remaining reproductive output for an individual aged i (and alive at that age): $y_i = E(\mathcal{LRO}_\text{i})$. Therefore we have $y_1 = E_{\mathcal{LRO}} = \mathbf{R_0}$. Similarly let us define v_i the variance of $\mathcal{LRO}_\text{i}$: $v_i = Var(\mathcal{LRO}_\text{i})$. Then $v_i = \mathcal{LRO} = \sigma^2_{\mathcal{LRO}}$ is the quantity we are looking to compute. The series of v_i can be computed backwards :

$$\begin{cases} v(n) = Var(\mathcal{F}_n) \\ \forall\ 1 \leq i < n \quad v(i) = Var(\mathcal{F}_i) + Var(\mathcal{S}_i \times \mathcal{LRO}_{i+1}) \end{cases} \tag{3.35}$$

Now, if the population is structured *only* by *age*, then all vital processes can be considered independent, and thus eq. 3.35 is equivalent to :

$$\begin{cases} v(n) = Var(\mathcal{F}_n) \\ \forall\ 1 \leq i < n \quad v_i = Var(\mathcal{F}_i) + y_{i+1}^2 s_i(1-s_i) + v_{i+1}.s_i \end{cases} \tag{3.36}$$

We can multiply all four sides of equation system 3.36 by P_i and, by letting $\alpha_i = v_i P_i$, we get :

$$\begin{cases} \alpha_n = P_n Var(\mathcal{F}_n) \\ \forall\ 1 \leq i < n \quad \alpha_i = P_i \left[Var(\mathcal{F}_i) + y_{i+1}^2 s_i(1-s_i) \right] + \alpha_{i+1} \end{cases} \tag{3.37}$$

As $\alpha_{n+1} = v_{n+1} P_{n+1} = 0$ and $s_n = 0$, we can rewrite eq. 3.37 :

$$\forall\ 1 \leq i \leq n \quad \alpha_i - \alpha_{i+1} = P_i \left[Var(\mathcal{F}_i) + y_{i+1}^2 s_i(1-s_i) \right] \tag{3.38}$$

The sum of the n equations of eq. system 3.38 yields

$$\alpha_1 = v_1 P_1 = Var(\mathcal{LRO}) = \sum_{i=1}^{n} P_i \left[Var(\mathcal{F}_i) + y_{i+1}^2 s_i(1-s_i) \right]$$

which, in the case where the fertility process consist in producing either 1 or 0 offspring per period (as in the framework of this chapter), gives:

$$\alpha_1 = Var(\mathcal{LRO}) = \sum_{i=1}^{n} P_i \left[f_i(1-f_i) + y_{i+1}^2 s_i(1-s_i) \right] \tag{3.39}$$

3.0.3 Note on $\mathbf{R_0}$ in matrix models with *hidden heterogeneity* trait

The net reproductive rate, the expectation of lifetime reproductive output noted $\mathbf{R_0}$ by Dublin and Lotka (1925) - a concept extensively used in epidemiology (see review by Heesterbeek, 2002)- is a key demographic measure. The extension of the concept and its calculation towards structured population models is not however without difficulty. This is in particular the case for models incorporating both a basic trait (say *age*)and a *hidden heterogeneity* trait (i.e. several classes of offspring), despite numerous works on the subject (De-Camino-Beck et al., 2008; Caswell, 2011; Cushing and Zhou, 1994).

Such a (*age-heterogeneity*)-population model can be represented by two matrices. The first one, we have fully developed in section 3.2.2, is the MPPM $\mathbf{M}$, which we can call the next time-step matrix. It provides, in $M_{i,j}$, the expected number of individuals in state $i \leftrightarrow (a_i, h_i)$ an individual in state $j \leftrightarrow (a_j, h_j)$ generates at each time-step. The second one is the Next Generation Matrix $\mathbf{R}$. It provides, in $R_{i,j}$, the expected number of individuals in state $i \leftrightarrow (a_i, h_i)$ an individual in state $j \leftrightarrow (a_j, h_j)$ generates *over its remaining lifetime*. As all offspring produced have age 1, $R_{i,j}$ will only be strictly positive if $a_i = 1$; that is, the only non-zero lines of $\mathbf{R}$ lines will correspond to the *het* offspring states $(1,1), (1,2), ..., (1, het)$. Therefore we can reduce, without any loss of ergodic information, $\mathbf{R}$ to $\mathbf{R}^*$ the sub-matrix of $\mathbf{R}$ defined only in these *het* offspring states. We can further define horizontal vector $\mathbf{e}_{\mathcal{LRO}}$ as the sum of all lines in $\mathbf{R}^*$; in matrix notation $\mathbf{e}_{\mathcal{LRO}} = \mathbf{1}'.\mathbf{R}^*$. This way, $e_h^{\mathcal{LRO}}$ represent the expected number of offspring (of all classes) for an individual in heterogeneity class h during its entire lifetime, hence the notation.

In the case where $\mathbf{R}$ is the only matrix provided - i.e. intra-generational dynamics are completely unknown - then the population ergodic generation growth rate (whatever the interpretation of such a measure when generations massively overlap) represent the maximum of likelihood for the mean lifetime reproductive success : $E(\mathcal{LR}\prime) = \mathbf{R_0} = \max eig(\mathbf{R}*) = \max eig(\mathbf{R})$ (De-Camino-Beck et al., 2008; Caswell, 2011; Cushing and Zhou, 1994; Jones, 2007). Indeed, such a population will tend toward a stable-state with regard to generation time, of growth rate the maximum eigenvalue $\mathbf{R_0}$ and relative abundance the vector of offspring categories $\mathbf{w}*$ the associated right-eigenvector :

$$\mathbf{R}^*\mathbf{w}^* = \mathbf{R_0}\mathbf{w}^*, \tag{3.40}$$

with $\mathbf{w}*$ scaled to sum to 1. Then, by summing all lines on both sides of eq. 3.40, one gets :

$$\mathbf{e}_{\mathcal{LRO}}.\mathbf{w}^* = \mathbf{R_0}, \tag{3.41}$$

This Net Reproductive Rate is thus the $\mathbf{w}^*$-weighted sum of all the lifetime reproductive output per genotype, i.e. the expected lifetime reproductive output by an offspring taken at random in the generation ergodic heterogeneity distribution. As such, it is only the "best guess" when intra-generational dynamics are ignored, and a potentially acceptable approximation of $E(\mathcal{LRO})$ when the vital rates vary little with *heterogeneity*.

However, in the general case where $\mathbf{M}$ and therefore intra-generational dynamics are known, this is not the case, despite opposite statements in the literature aforementioned. This is because, in that case, the stable-state distribution of offspring classes is only poorly approximated by distribution $\mathbf{w}^*$ of eq.3.41. Indeed the eigen-equation at the time-step level then yields $\mathbf{Mw} = \lambda\mathbf{w}$, with $\mathbf{w}$ scaled to sum to 1, which provides vector $\mathbf{w}$ representing ergodic abundances of all (a, h) states.
Vector $\mathbf{w}^\diamond$ is the sub-vector of $\mathbf{w}$ containing only offspring states (*age* 1):

$$\mathbf{w}^\diamond = \frac{1}{w_{1,1} + w_{1,2} + ... + w_{1,h}}.(w_{1,1}, w_{1,2}, ..., w_{1,h}).$$

It represents the exact ergodic distribution of *heterogeneity* classes at birth, and should therefore replace $\mathbf{w}^*$ in eq. 3.41. Indeed, from the decomposition of the MPPM into survival and fertility matrices $\mathbf{M} = \mathbf{T} + \mathbf{F}$, can one easily generate the fundamental matrix $\mathbf{N} = (\mathbf{I} - \mathbf{T})^{-1}$, and therefore the Next-Generation Matrix $\mathbf{R} = \mathbf{F}.\mathbf{N}$, and thus $\mathbf{R}^*$ and $\mathbf{e}_{\mathcal{LRO}}$.

Therefore, we get, for a population structured by *age* and *heterogeneity* classes, for which the time-step transition matrix $\mathbf{M}$ is known, the following formula for $\mathbf{R_0}$:

$$\mathbf{R_0} = \mathbf{e}_{\mathcal{LRO}}.\mathbf{w}^\diamond \tag{3.42}$$

3.0.4 Note on matrix $\mathbf{M_{a,p}}$: calculation of $\mathbf{R_0}$ and interpretation

Matrix $\mathbf{M_{a,p}}$ is matrix $\mathbf{M}$, the *age-parity-heterogeneity* MPPM, *folded* over *heterogeneity*. In $\mathbf{M_{a,p}}$, heterogeneity is not implemented as a trait any more, but since it was implemented in the full-traited matrix $\mathbf{M}$ from which it is derived, it has effects on $\mathbf{M_{a,p}}$, making its interpretation both challenging and interesting.

Discrepancies in calculation of $\mathbf{R_0}$

In $\mathbf{M}$, because of the implementation of *physiological costs of reproduction* via trait *parity* (see section 3.2.2), survival transitions output states depend from fertility (see equation 3.3). Simply put, survival transitions from any state $i = (a, p, h)$ to either $j_1 = (a + 1, p, h)$ or $j_2 = (a + 1, p + 1, h)$ are $\mathrm{M}_{j_1,i} = s_i * (1 - f_i)$ and $\mathrm{M}_{j_2,i} = s_i * f_i$, with a 3rd, implicit transition towards *death* worth $\tilde{\mathrm{M}}_{death,i} = 1 - s_i$. These three transitions sum to 1. In the MCwR tool, the corresponding fertility rewards expectations for these three transitions (in $\mathbf{Rw^1}$) are respectively 0, 1 and f_i. Thus the mean expected reward is $0.(s_i.(1 - f_i)) + 1.(s_i.f_i) + f_i.(1 - s_i) = f_i$ and therefore both MCwR and $\mathbf{R}$ approaches ($\mathbf{e}_{\mathcal{LRO}} = \mathbf{1}'.\mathbf{R}^*$, see appendix 3.0.3) provide the same results for $E(\mathcal{LRO})$. In $\mathbf{M}_a$, the Leslie reference matrix, survival and fertility transitions are completely separated, with only one output per survival transition. Thus in this case also, both MCwR and $\mathbf{R}$ provide the same results.

However for the intermediary matrix, $\mathbf{M}_{a,p}$ - $\mathbf{M}$ *folded* on *heterogeneity*- both measures differ. Indeed, through *folding*, the survival transitions from state $i = (a, p)$ towards either $j_1 = (a + 1, p)$ or $j_2 = (a + 1, p + 1)$ will, in general, not be distributed according to the transition value between i and the (unique) offspring state 1 which we interpret as fertility rate ($\mathrm{M}_{1,i \leftrightarrow (a,p)} = f_i$). This is caused by the *heterogeneity* modeled in $\mathbf{M}$ expressing itself through EFP-*merged* vital rates, now that *heterogeneity* is not a trait any more (see chapter 2 (p.41)).

Let us illustrate this, seemingly paradoxical, situation with a simple example : Let us imagine a population structured by 2 *age* - and therefore 2 *parity* classes - and 2 *heterogeneity* classes (A and B) produced in equal measures ($m = 0.5$) at each fertility event. A individuals have all vital rates at 1 and B individuals at 0.5. Then all A newborns (half the population of newborns), will produce 1 offspring and become individuals of age 2 and parity 1. Half of B newborns will produce 1 offspring and half of B individuals will survive. Those halves are independent, and thus a quarter of B individuals will survive *and* become adults of *parity* 1, and another quarter will become adults of *parity* 0. Thus for the population *folded* on *heterogeneity*, i.e. where individuals are only characterized by *age* and *parity*, the newborn fertility rate is $0.5 \times 1 + 0.5 \times 0.5 = 0.75$. Similarly the survival rate for newborns is $0.5 \times 1 + 0.5 \times 0.5 = 0.75$. However, for an average newborn in the population, the probability of transitioning towards a *parity* 1 adult is $0.5 \times 1 + 0.5 \times 0.25 = 0.625$ and to a *parity* 0 adult is $0.5 \times 0 + 0.5 \times 0.25 = 0.125$. As we can see here, the sum of the survival transitions is (by construction) equal to the survival rate, but the distribution towards higher parity $\frac{0.625}{0.625+0.125} \approx 0.83$ is not equal to the fertility rate 0.85 as one does not make the distinction between individuals A and B any more.

Interpretation of $\mathbf{M_{a,p}}$

These considerations have important consequences with regards to the interpretation of $\mathbf{M_{a,p}}$. It basically comes down to deciding whether fertility rates are to be found on the first line of the matrix, on in the distribution rates towards classes of higher *parity*. In the first case, $E(\mathcal{LRO})$ should be calculated using

$\mathbf{R_0} = R_{1,1}$, in the second case, via MCwR. Because, as we just illustrated, the 'inferred' fertility rates are, in general, different between fertility transition and distribution of survival transition, these two methods provide different results for $\mathbf{R_0}$. Considering that the folding operation does not alter the fact that $\mathbf{M_{a,p}}$ is an age-based MPPM with no *heterogeneity* implemented and thus only 1 offspring class, it makes sense to resolve the dispute in favor of considering the first line of the matrix as fertility rates for all states. Then, however, it implies that, the second trait of the model is abusively called *parity*. The categories it generates still correspond to states with decreasing vital rates as the category number increases (i.e. to physiological costs of reproduction on survival and fertility), but an increment in the category number does not imply 1 exact additional offspring.The relationship is not linear any more, and the trait *parity* rather becomes a general "measure of overall reproductive success" than parity exactly, though we will still use that name for the trait itself.

That the second trait of $\mathbf{M_{a,p}}$ cannot be interpreted as parity *stricto sensu* also has repercussions in terms of measures for the variance of $\mathcal{LRO}$. We just saw, that MCwR cannot be used to measure $E(\mathcal{LRO})$ and for the same reason, this framework cannot be used for precisely calculating $\sigma^2_{\mathcal{LRO}}$ either. Moreover, even if *parity* does not account for the exactly reproductive success any more, there is still interdependence between fertility rates and survival transitions, making the formulas stemming from $\mathbf{R}$ equally unsatisfactory. We shall therefore use both approaches, as proxies, keeping in mind that none can provide an exact result, which reflects the fact that matrix $\mathbf{M_{a,p}}$ is not a constructed model, but the product of *folding* from a model embedding a physiological trade-off, implying a shift in the interpretation of its *dynamic heterogeneity* trait.

3.0.5 Two individual variances : $\sigma^2_{\mathcal{LRO}}$ and σ^2_d

Variance in lifetime reproductive output $\sigma^2_{\mathcal{LRO}}$ (self explanatory) and demographic variance σ^2_d (defined in section 3.3.2) are two measures of the effect of individual (or demographic) stochasticity on fitness either taken as the ergodic growth rate λ in the case of σ^2_d or as the lifetime reproductive output ($\mathbf{R_0}$) in the case of $\sigma^2_{\mathcal{LRO}}$.

Similar concepts lead to similar usage

Corresponding to similar concepts, these two measures have a lot in common, and in particular they are used in analogous computations. For instance, the effect of individual stochastiocity of effective population size (we discuss in section 3.4.3 is studied via σ^2_d by (Engen et al., 2005a) and via $\sigma^2_{\mathcal{LRO}}$ by Crow and Kimura (see equation 7.6.2.17 page 351 of Crow and Kimura, 1970) and later refined by Rockwell and Barrowclough (1995) and Hill (1979). Extinction probabilities also can be approached either via the populationwise and infinitesimal approach (σ^2_d) or by the individual and exact method ($\sigma^2_{\mathcal{LRO}}$). The former framework is used by Lande and Orzack (1988) and Tuljapurkar (1982b) who show that the distribution of extinction time follows an inverse Gaussian distribution of variance proportional to overall individual variance in contribution to growth rate σ^2. The latter framework, extended to the entire distribution of lifetime reproductive success as a field known as branching process theory, was used to calculate extinction times and probabilities from Galton-Watson processes (modeling non-overlapping generations, see (Keyfitz and Caswell, 2005; Ellison, 1994)), Birth-death processes (implementing variability in longevity, see (Goodman, 1967)) and Crump-Mode-Jaggers processes (modeling age-structured populations, see (Crump and Mode, 1968, 1969)).

Similar concepts but different approaches

However similar, stemming from different approaches, $\sigma^2_{\mathcal{LRO}}$ and σ^2_d correspond to different concepts. Lifetime reproductive output, the random variable $\mathcal{LRO}$ which expectation is $\mathbf{R_0}$ is an individual concept. The exact distribution of $\mathcal{LRO}$ can be computed, provided the stochastic processes of vital rates. Compounded over the different (categories of) individuals in the population, $\mathbf{R_0}$ becomes a population fitness measure. But, in age-structured populations, $\mathbf{R_0}$ is a poor fitness measure compared to λ, as the latter also takes into account the life history pace of the organism, whilst $\mathcal{LRO}$ does not 'care' whether the offspring are produced early or late in reproductive life (see discussion in appendix 3.0.3).

From its individualistic roots, $\mathcal{LRO}$ retains two important properties. First, it is a random variable that each individual (within a genotype × environment configuration) will realize differently and independently, with expectancy, variance and higher moments exactly describing the variety of all possible reproductive life trajectories in the population. Second, it does not depend on the population being at stable-state regime and is meaningful even for a population which distribution is very different from $\boldsymbol{w}$. These properties are reminiscent of those of Fisher's reproductive value (Fisher, 1930), $\boldsymbol{v}/v(1)$ which, in age-structured population, only depends on the age of the individual and is the same in transient or stable-state regime, a fact clouded by its computation as left-eigenvector of the matrix model. The 'individuality'

of $\boldsymbol{v}$ is also obvious from $v(a)$ being frequently described as "individual reproductive value" (for an individual in age class a) by contrast with the age a class reproductive value : $c(a) = v(a).w(a)$ (Taylor, 1990; Taylor et al., 2007). This reminiscence is not surprising as, in age-structured populations, expressions of reproductive value at birth and $\mathbf{R_0}$ only differ in the discounting or not by λ: $\mathbf{R_0} = \sum_i f(i) \prod_{j=1}^{i-1} s(j)$ and $v(1) = \sum_i \lambda^{-i} f(i) v(1) \prod_{j=1}^{i-1} s(j)$.

By contrast, and despite efforts from (McGraw and Caswell, 1996) and others, λ, as fitness, is a population concept. It is possible to compute λ for a population with heterogeneity in expected vital rates or for a given genotype within that population (Charlesworth, 2000) but it is hard to fathom at the level of the individual. However, since the asymptotic growth rate is the population fitness measure of choice for populations with overlapping generations, efforts have been made to implement the individual contributions to λ in order to take individual stochasticity into account. Such efforts, produced by Engen and collaborators (Engen et al., 1998, 2005b,c) yielded the demographic variance σ_d^2 which measures the variance in growth rate inferred by stochasticity in vital rates. In age-structured populations, the calculation is based on the sensitivities of λ to infinitesimal changes in vital rates (see eq. 3.18 in section 3.3.3). As such, the demographic variance calculations can only provide an approximation (sensitivities are only valid for infinitesimal changes in rates) and is only valid at the level of the stable-state.

Actually, as the individual stochasticity of vital rates are generally larger than their environmental stochasticity, equation 3.18 is expected to be a worth predictor of demographic variance, than equation 3.17 is of environmental variance. Let us mention here an interesting approach by (Engen et al., 2007, 2009) who circumvent the pitfall of stable-state dependency of σ_d^2 by measuring the population, not by its numbers in each age class $n(a)$ (regrouped in population vector $\boldsymbol{n}$ of sum N), but by its class-reproductive value $n^*(a) = v(a).n(a)$ (regrouped in vector $\boldsymbol{n^*}$ of sum N^*); taking advantage of the fact, contrary to N, N^* does not require stable-state to grow at λ; indeed, $N^*_{t+1} = \boldsymbol{v}.\boldsymbol{n}_{t+1} = \boldsymbol{v}.\mathbf{M}.\boldsymbol{n}_t = \lambda.\boldsymbol{v}.\boldsymbol{n}_t = \lambda.N^*_t$.

Comparison of formulas for age-structured populations

Let us now compare the specific formulas for σ_d^2 and $\sigma^2_{\mathcal{LRO}}$ for age-structured MPPMs (i.e., populations for which vital rates only depend on age and are their realizations are independent). From section 3.3.2, we have the formula for $\sigma^2_{\mathcal{LRO}}$(eq. 3.5 and 3.34):

$$\begin{cases} \sigma^2_{\mathcal{LRO}} = \sum_{i=1}^{n} P_i \left[Var(\mathcal{F}_i) + y_{i+1}^2 s_i(1-s_i) \right] \\ y_i = \frac{1}{P_i} \sum_{j=i}^{n} f_j P_j \end{cases} \tag{3.43}$$

From section 3.3.3, we can adapt the generic formula for σ_d^2 in age-structured populations (eq. 3.21) to implement independence of vital rates, i.e. $\mathrm{Cov_d}(\mathrm{M}_{1,i}, \mathrm{M}_{1,i}) = Var(\mathcal{F}^i)$ and $\mathrm{Cov_d}(\mathrm{M}_{i+1,i}, \mathrm{M}_{i+1,i}) = Var(\mathcal{S}^i)$ and all other covariances at 0:

$$\sigma_d^2 = \sum_{i=1}^{n} \frac{\lambda^{-2}}{w_i} \left[(\frac{\partial \lambda}{\partial f_i})^2 Var(\mathcal{F}_i) + (\frac{\partial \lambda}{\partial s_i})^2 Var(\mathcal{S}_i) \right]$$

Using the entries of the sensitivity matrix of the model as we did for (*age-parity*)-MPPM in section 3.3.3, i.e. $\frac{\partial \lambda}{\partial f_i} = v_1 w_i$ and $\frac{\partial \lambda}{\partial v_i} = v_{i+1} w_i$, and introducing $\bar{b} = (\sum_1^n P_i \lambda^{-i})^{-1}$ the birth rate, and $T = \sum_1^n i f_i P_i \lambda^{-i}$ the generation time (mean age of parents), we can write σ_d^2 as a function of $\boldsymbol{v^*}$ (with $v_1^* = 1$) the Fisherian reproductive value :

$$\begin{cases} \sigma_d^2 = \frac{1}{\bar{b} T^2} \sum_{i=1}^{n} \frac{P_i}{\lambda^i} \left[Var(\mathcal{F}_i) + v_{i+1}^{*2} s_i(1-s_i) \right] \\ v_i^* = \frac{\lambda^{i-1}}{P_i} \sum_{j=i}^{n} f_j P_j \lambda^{-j} \end{cases} \tag{3.44}$$

(Engen et al., 2005a)

Comparing eq. systems 3.43 and 3.44 is very informative with respect to the similarities and differences between the exact formula for $\sigma^2_{\mathcal{LRO}}$ and the approximation of σ_d^2. The structures are strikingly similar, with the major differences arising from the relation with λ of σ_d^2 and from the time-step building block of each approach (chronological time for σ_d^2 and generation time for $\sigma^2_{\mathcal{LRO}}$) from which emerges T and $\bar{b}$.

These connections are made even clearer but reducing the problem to stationary populations:

$$\lambda = 1 \implies \sigma_d^2 = \frac{\sigma^2_{\mathcal{LRO}}}{\bar{b} T^2} \tag{3.45}$$

The relationship in that particular case (eq. 3.45), allows incidentally to reconcile Hill (1972)'s equation for effective size in populations with overlapping generations with constant size as a function of variance in reproductive success, with Engen et al. (2005a)'s formula using the infinitesimal approach.

Particular case of 2 age classes. In the particular case of two age classes only.The model is $\mathbf{H} = \left[\begin{smallmatrix} f_1 & f_2 \\ s_1 & 0 \end{smallmatrix}\right]$ and we consider $\lambda = 1$. From the Euler-Lotka equation (Euler, 1760; Lotka, 1939) we know this implies $\lambda = f_1 + s_1 f_2 = 1 = \mathbf{R_0}$. From $\mathbf{R_0} = f_1 + s_1 f_2$ we can easily calculate the variance of reproductive output : $\sigma^2_{\mathcal{LRO}} = f_1(1-f_1) + s_1 f_2(1 - s_1 f_2)$. To compute demographic variance, we observe that $\boldsymbol{w}$ is the right-eigenvector, summing to 1, i.e. $\mathbf{H}.\boldsymbol{w} = \lambda.\boldsymbol{w}$ which implies $s_1 w_1 = \lambda w_2$ thus $w_1 = \frac{1}{1+s_1}$ and $w_2 = \frac{s_1}{1+s_1}$. And $\boldsymbol{v}$ is the right-eigenvector, such that $\boldsymbol{v}'.\boldsymbol{w} = 1$, i.e. $\boldsymbol{v}'.\mathbf{H} = \lambda.\boldsymbol{v}'$ which implies $v_1 f_2 = \lambda v_2$, but $w_1 v_1 + w_2 v_2 = 1 = w_1 v_1 + w_2 v_1 f_2$ and thus $v_1 = \frac{1+s_1}{1+f_2 s_1}$ and $v_2 = \frac{f_2(1+s_1)}{1+f_2 s_1}$. And thus $\sigma^2_{\mathrm{e}} = w_1 v_1^2[f_1(1-f_1) + f_2^2 s_1(1-s_1) + s_1 f_2(1-f_2)]$. Here, $\mathrm{T} = (w_1 v_1)^{-1}$, therefore we have $\sigma^2_{\mathrm{e}} = \frac{1}{\mathrm{T}^2}\frac{\sigma^2_{\mathcal{LRO}}}{w_1}$.

3.0.6 Preservation of $\mathbf{R_0}$ by *folding* for models with *age*

We shall demonstrate here that for MPPMs with *age* as a trait, $\mathbf{R_0}$ is preserved by *folding* (over any combination of any other trait than *age*). This may may seem self-evident, but it really is not. *Trait Level Analysis* - developed in chapter 2 (p.41)- allows to draw conclusion between matrices considered equivalent because their share the asymptotic stable state properties; and chiefly among them, λ. However in general it does not preserve $\mathbf{R_0}$. to see that, one needs only to contemplate the folding of a simple non-*age* based MPPM, as we do in Supplementary Material 5.2.3 page 140.

To prove the preservation of $\mathbf{R_0}$ by *folding* in the specific case where *age* is a trait and is not *folded* upon, let us consider a model $\mathbf{M}$, that is re-organized (if need be) so that *age* is the last trait of the trait structure $\mathbf{s}$. Let us regroup all other traits as one unique trait t which can take values from $t = 1$ to $t = tmax$, representing the $tmax$ combinations of other (than *age*) traits. Trait vector is thus $\mathbf{t} = \{t, age\}$ and trait structure $\mathbf{s} = (tmax, \omega)$ (there are ω age classes). With no loss of generality therefore, we shall study the effect of folding $\mathbf{M}$ over t on $\mathbf{R_0}$. The operation produces $\mathbf{M^{fold}_{age}} = \mathbf{M_a}$ only characterized by age: $\mathbf{t_{M_a}} = \{age\}$ and trait structure $\mathbf{s_{M_a}} = (\omega)$.For simplicity, we shall use a block-matrix approach for the demonstration.

Matrix $\mathbf{M_a}$ is a Leslie matrix: $\mathbf{M_a} = \left[\begin{smallmatrix} f_1 & f_2 & \cdots & f_\omega \\ s_1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & s_{\omega-1} & 0 \end{smallmatrix}\right]$ with well-know net reproductive rate, we denote $\mathbf{R^a_0}$: $\mathbf{R^a_0} = \sum_{i=1}^{\omega} f_i(\prod_{j=1}^{i-1} s_j)$.

Matrix $\mathbf{M}$ can be written a block-Leslie matrix :

$$\mathbf{M} = \left[\begin{array}{c|c|c|c|c} \mathbf{F}_1 & \mathbf{F}_2 & \dots & \mathbf{F}_{\omega-1} & \mathbf{F}_\omega \\ \hline \mathbf{S}_1 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \hline \dots & \dots & \dots & \dots & \dots \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \dots & \mathbf{S}_{\omega-1} & \mathbf{0} \end{array}\right] ,$$

where each submatrix is a square matrix of size $tmax \times tmax$. Specifically they are such that for a vector $\boldsymbol{n_i}$ of abundances of individuals at age i, $\mathbf{F}_i.\boldsymbol{n_i}$ is the vector of abundances of offspring produced by these individuals at a given time step and $\mathbf{S}_i.\boldsymbol{n_i}$ is the vector of abundances of their survived selves. By construction $\mathbf{M}$ and $\mathbf{M_a}$ share the same growth rate λ. Their related right eigenvectors, both summing to 1, $\boldsymbol{w} = \begin{bmatrix} \boldsymbol{w}_1 & \boldsymbol{w}_2 & \dots & \boldsymbol{w}_\omega \end{bmatrix}$ (this formula displays $\boldsymbol{w}$ as a vector of vectors) and $\boldsymbol{w^*} = \begin{bmatrix} w_1^* & w_2^* & \dots & w_\omega^* \end{bmatrix}$ are such that $\mathbf{1}'.\boldsymbol{w}_i = w_i^*$.

Then, from appendix section 3.0.3 p. 93 providing the general formula for $\mathbf{R_0}$ when the model has several classes of offspring and a known *time-step* projection matrix, we get (we allow ourselves to equate matrices of different sizes whenever they have equal non-zero diagonal block-matrices on their Frobenius normal form):

$$\mathbf{R_0} = \frac{\mathbf{1}'.\mathbf{R}.\boldsymbol{w_1}}{w_1^*} \tag{3.46}$$

Writing out $\mathbf{R}$, we get :

$$\mathbf{R} = \mathbf{F}.(\mathbf{I} + \mathbf{T} + \mathbf{T^2} + \cdots + \mathbf{T}^\omega) = \sum_{i=1}^{\omega} \mathbf{F}_i \mathbf{P}_i \tag{3.47}$$

where $\mathbf{P}_i = \prod_{j=i-1}^{j=1} \mathbf{S}_i$ (order of multiplicands is important here) and $\mathbf{P}_1 = \mathbf{I}$. Then from equations 3.47 and 3.46 we get :

$$\mathbf{R_0} = \mathbf{1}'.\sum_{i=1}^{\omega} \mathbf{F}_i \mathbf{P}_i.\frac{\boldsymbol{w_1}}{w_1^*} = \sum_{i=1}^{\omega} \mathbf{1}'.\mathbf{F}_i \mathbf{P}_i.\frac{\boldsymbol{w_1}}{w_1^*} \tag{3.48}$$

Considering the eigen.equation $\mathbf{M}.\boldsymbol{w} = \lambda \boldsymbol{w}$ by blocks, we immediately get $\mathbf{S_i}.\boldsymbol{w_i} = \lambda \boldsymbol{w_{i+1}}$ and $\sum_{i=1}^{\omega} \mathbf{F}_i.\boldsymbol{w_i} = \lambda \boldsymbol{w_1}$. The eigen.equation at the level of $\mathbf{M_a}$- $\mathbf{M_a}.\boldsymbol{w^*} = \lambda \boldsymbol{w^*}$ - implies that $s_i.w_i^* = \lambda w_{i+1}^*$. Thus $\mathbf{S_i}.\dfrac{\boldsymbol{w_i}}{w_i^*} = s_i \dfrac{\boldsymbol{w_{i+1}}}{w_{i+1}^*}$. From there, we infer

$$\mathbf{P}_i.\frac{\boldsymbol{w_1}}{w_1^*} = (\prod_{j=1}^{i-1} s_j)\frac{\boldsymbol{w_i}}{w_i^*} \tag{3.49}$$

By definition of the EFP-*folding* (see chapter 2 (p.41)), the *folding* of matrices by ergodic abundance weighted average of transitions

$$f_i = \mathbf{1}'.\mathbf{F}_i.\frac{\boldsymbol{w_i}}{w_i^*} \tag{3.50}$$

Multiplying both sides of equation 3.49 $\mathbf{1}'.\mathbf{F}_i$, and simplifying the result thanks to equation 3.50, we can rewrite equation 3.48 in a way that provides the proof:

$$\mathbf{R_0} = \sum_{i=1}^{\omega} f_i(\prod_{j=1}^{i-1} s_j) = \mathbf{R_0^*} \tag{3.51}$$

3.0.7 Demonstration of $\sigma^2_{\mathcal{LRO}}[\ \mathbf{M_{age,parity}}] < \sigma^2_{\mathcal{LRO}}\left[\ \mathbf{M_{age}^{fold}}\right]$

Let us consider an (*age-parity*)-model, $\mathbf{M_{age,parity}}$ we denote here $\mathbf{M}$, implementing *physiological costs of reproduction*, as described in section 3.2.1. Without loss of generality, for simplification, we shall consider that only fertility rates are affected by the costs. Let us also consider, $\mathbf{M_{age}^{fold}}$, which $\mathbf{M}$ folded on *parity*, that we denote here with an asterisk: M^*. In order to demonstrate that $\sigma^2_{\mathcal{LRO}}[\ \mathbf{M_{age,parity}}] < \sigma^2_{\mathcal{LRO}}\left[\ \mathbf{M_{age}^{fold}}\right]$, we shall set our investigation at the sable-state. By the properties of *Trait Level Analysis*, at the stable state, $\mathbf{M}$ and $\mathbf{M}^*$ have the same growth rate and the associated right-eigen vector on *age* $\boldsymbol{w_a}$.

Let us denote $\mathcal{P}_a$ and ${\mathcal{P}^*}_a$ the random variables giving the parity of a random individual in the a age-class, in the stable state population, for respectively $\mathbf{M}$ and $\mathbf{M}^*$. The parity r.v. at age $(a+1)$ are worth $\mathcal{P}_{a+1} = \mathcal{S}_a.\mathcal{F}_{a,p} + \mathcal{P}_a$ and ${\mathcal{P}^*}_{a+1} = {\mathcal{S}^*}_a.{\mathcal{F}^*}_a + {\mathcal{P}^*}_a$.

We can get expectations for the r.v. of the multitrait model, according to parity. $E_p(\mathcal{P}_a) = \sum_{p=1}^{a} \frac{p.w_{a,p}}{w_a}$ is $\bar{p_a}$ the average parity at that age a. $E_p(\mathcal{F}_{a,p}) = \sum_{p=1}^{a} \frac{f(a,p).w_{a,p}}{w_a}$. Since $f(a,p) = f_a(1-\frac{p}{\omega})$, and from the *Trait Level Analysis* principles, we have $E_p(\mathcal{F}_{a,p}) = \frac{f_a}{w_a}\sum_{p=1}^{a}(1-\frac{p}{\omega}).w_{a,p} = \frac{f_a}{w_a}\left[\sum_{p=1}^{a} w_{a,p} - \frac{1}{\omega}\sum_{p=1}^{a} p\right] = f_a(1-\frac{\bar{p_a}}{\omega}) = E({\mathcal{F}^*}_a) = \bar{f_a} = f_a^*$. From the summation of $E(\mathcal{P}_{a+1}) = E(\mathcal{S}_a).E(\mathcal{F}_{a,p}) + E(\mathcal{P}_a)$ over a, we therefore get $\forall a \ \ E_p(\mathcal{P}_a) = E({\mathcal{P}^*}_a)$. A result related to the preservation of $\mathbf{R_0}$ we demonstrate in section 3.0.6. To simplify further calculations, as we base our analysis at the time-step level, we shall now consider only one process projecting an individual from age a to $a+1$, which combines survival and survival: $\mathcal{Q}_{a,p} = \mathcal{S}_a.\mathcal{F}_{a,p}$. As a product of Bernoulli processes, $\mathcal{Q}$ is itself Bernoulli, of parameter $q_{a,p} = f_a.s_a.(1-\frac{p}{\omega})$.

Let us now turn ourselves to the variances of these r.v., since in a population structured by age only, the vital rates are independent from parity, we have $Var_p({\mathcal{P}^*}_{a+1}) = Var_p({\mathcal{Q}^*}_a) + Var_p({\mathcal{P}^*}_a)$ and thus $Var_p({\mathcal{P}^*}_{a+1}) - Var_p({\mathcal{P}^*}_a) = q_a^*(1-q_a^*)$. We also have : $Var_p(\mathcal{P}_{a+1}) = Var_p(\mathcal{Q}_a) + Var_p(\mathcal{P}_a) + 2.Cov_p(\mathcal{Q}_a, \mathcal{P}_a)$. As $\mathcal{Q}$ is Bernoulli, we have $Var_p(\mathcal{Q}_a) = \sum_{p=1}^{a} 1^2.q(p).w_{a,p} - (q_a^*)^2 = q_a.(1-q_a) = Var_p({\mathcal{Q}^*}_a)$. And therefore, the difference in change in variances, by age, lies in the covariance $Cov(\mathcal{F}_a, \mathcal{P}_a)$ component:

$$[Var_p(\mathcal{P}_{a+1}) - Var_p(\mathcal{P}_a] - [Var_p(\mathcal{P}_{a+1}^*) - Var_p(\mathcal{P}_a^*] = 2.Cov_p(\mathcal{Q}_a, \mathcal{P}_a)$$

We can explicit this component :

$$Cov(\mathcal{Q}_a, \mathcal{P}_a) = \sum_{p=1}^{a} p.q(a,p).w_{a,p} - \bar{p_a}.q_a^*$$

where

$$\sum_{p=1}^{a} p.q(a,p).w_{a,p} = q(a)\sum_{p=1}^{a} p.(1-\frac{p}{\omega}).w_{a,p} = q(a)\sum_{p=1}^{a} p.w_{a,p} - \frac{1}{\omega}q(a)\sum_{p=1}^{a} p^2.w_{a,p}$$

$$= q(a)\bar{p_a} - \frac{1}{\omega}q(a)\sum_{p=1}^{a} p^2.w_{a,p} = [q(\bar{a}) + q(a)\frac{p_a}{\omega}].\bar{p_a} - \frac{1}{\omega}q(a)\sum_{p=1}^{a} p^2.w_{a,p}$$

and thus

$$Cov_p(\mathcal{Q}_a, \mathcal{P}_a) = \frac{q_a}{\omega}\left[\bar{p_a}^2 - \sum_{p=1}^{a} p^2.w_{a,p}\right] = \frac{q_a}{\omega}.(-Var_p(\mathcal{P}_a))$$

And therefore,

$$[Var_p(\mathcal{P}_{a+1}) - Var_p(\mathcal{P}_a] - [Var_p(\mathcal{P}^*_{a+1}) - Var_p(\mathcal{P}^*_a] = 2.\frac{f_a.s_a}{\omega}.(-Var_p(\mathcal{P}_a)) \tag{3.52}$$

And thus $\forall a, Var(\mathcal{P}_a) < Var(\mathcal{P}^*_a)$: in each age-class, the variance of parity is lower for the population modeled by $\mathbf{M}$ than for the population modeled by $\mathbf{M}^*$. At a given time, in age class a of each population, some individuals will be removed, their lifetime trajectory stopped and therefore their parity at that time will be their $\mathcal{LRO}$. As survival is the same for both populations (and here independent from parity), we get $\sigma^2_{\mathcal{LRO}} = \sum_{a=1}^{\omega} Var(\mathcal{P}_a).\prod_{i=1}^{a-1} s(i)$ and therefore we get

$$\sigma^2_{\mathcal{LRO}}[\mathbf{M}] - \sigma^2_{\mathcal{LRO}}[\mathbf{M}^*] = \sum_{a=1}^{\omega}(Var(\mathcal{P}_a) - Var(\mathcal{P}^*_a)).\prod_{i=1}^{a-1} s(i) \leq 0 \tag{3.53}$$

3.0.8 Effect of heterogeneity on $\sigma^2_{\mathcal{LRO}}$

We analyze here the cross-effects of the costs and heterogeneity on the variance in reproductive success of the overall population. And more we try and hint at the relative effects of heterogeneity and individual stochasticity in the making of $\sigma^2_{\mathcal{LRO}}$.

Heterogeneity, costs and $\sigma^2_{\mathcal{LRO}}$

Let us now imagine that the population modeled by $\mathbf{M}$ contains two genotypes. These two genotypes will therefore also be found in M^* which is an (*age-heterogeneity*)-MPPM, corresponding to $\mathbf{M}$ *folded* over *parity*. By the principles of *Trait Level Analysis* $\boldsymbol{w}^*$ the (*age,heterogeneity*)-right-eigen.vector, associated to λ for $\mathbf{M}^*$, correspond to (*age-parity-heterogeneity*)-right-eigen of $\mathbf{M}$ when summed on parity. Since offspring are all of parity 0, this implies the offspring abundances are the same for both models: $\boldsymbol{w}^{*\diamond} = \boldsymbol{w}^{\diamond}$. Put simply, this means that the effects of the costs and heterogeneity are independent. This can be further understood by considering the heterogeneity component of $\sigma^2_{\mathcal{LRO}}$ (introduced in section 3.3.2). From the equality between offspring abundances and between $\mathbf{R_0}$ (see section 3.0.6) between the two models, equation 3.12 yields ${\sigma^{\text{het}}_{\mathcal{LRO}}}^2[\mathbf{M}] = {\sigma^{\text{het}}_{\mathcal{LRO}}}^2[\mathbf{M}^*]$, and therefore,

$$\begin{aligned}\sigma^2_{\mathcal{LRO}}[\mathbf{M}] - \sigma^2_{\mathcal{LRO}}[\mathbf{M}^*] &= {\sigma^{\text{sto}}_{\mathcal{LRO}}}^2[\mathbf{M}] - {\sigma^{\text{sto}}_{\mathcal{LRO}}}^2[\mathbf{M}^*] \\ &= w_1^{\diamond}.(\sigma^2_{\mathcal{LRO}_1}[\mathbf{M}] - \sigma^2_{\mathcal{LRO}_1}[\mathbf{M}^*]) + w_2^{\diamond}.(\sigma^2_{\mathcal{LRO}_2}[\mathbf{M}] - \sigma^2_{\mathcal{LRO}_2}[\mathbf{M}^*])\end{aligned}$$

is independent from *heterogeneity*.

Order of magnitude of heterogeneity component of $\sigma^2_{\mathcal{LRO}}$

If the difference in variance in $\mathcal{LRO}$ between the models with and without the costs only depends on the stochastic difference - i.e. on the differences at the level of each genotype - the variance itself can be strongly impacted by heterogeneity, and specifically by differences in $\mathbf{R_0}$. As we can see from equation 3.12 (page 71), the effect of heterogeneity on the variance of $\mathcal{LRO}$ is exactly proportional to both the square of the difference in $\mathcal{R}$, and to the variance of the offspring distribution. These two components are not independent (high difference in reproductive rates causes high difference in genotypic $\lambda$and therefore large discrepancy in offspring abundances) but for small variations, the heterogeneity component of $\sigma^2_{\mathcal{LRO}}$ is maximal, for two genotypes cohabiting in the population, when $w_1 = w_2 = \frac{1}{2}$ and the difference in $\mathbf{R_0}$ between the genotypes is maximum. This implies that they are located - in the zero-parity vital rate map of figure 3.6d - on a line orthogonal to the iso-$\mathbf{R_0}$ curve. For the 5-year models figured in fig. 3.6, moving away from a stationary mean genotype located at $(f, s) = (.60, 49)$ in a direction (roughly $(1, 1)$) orthogonal to the stationary line, towards coordinates $(f_1, s_1) = (.70, 59)$ on one side and $(f_2, s_2) = (.50, 39)$ on the other side. For the mean genotype (i.e. for the Reference Leslie matrix of the model) $\mathbf{R_0} \approx 1$, $\lambda \approx 1$, $\sigma^2_{\mathcal{LRO}}[\mathbf{M}] = .594$ (for the full model) $\sigma^2_{\mathcal{LRO}}[\mathbf{M}^*] = .6862$ (for the model folded on *parity*). For the fit genotype (numbered 1), $\mathbf{R_0}_1 = 1.2947$, $\lambda = 1.1986$ $\sigma^2_{\mathcal{LRO}}[\mathbf{M_1}] = .7071$ and $\sigma^2_{\mathcal{LRO}}[\mathbf{M^*_1}] = .8461$, whereas for the frail genotype (numbered 2), $\mathbf{R_0}_2 = 0.7528$, $\lambda = 0.8310$ $\sigma^2_{\mathcal{LRO}}[\mathbf{M_2}] = .4776$ and $\sigma^2_{\mathcal{LRO}}[\mathbf{M^*_2}] = .5293$. For this heterogeneous population, we can therefore compute the *heterogeneity* component of $\sigma^2_{\mathcal{LRO}}$:${\sigma^{\text{het}}_{\mathcal{LRO}}}^2 = w_1(1-w_1)*(\mathbf{R_0}_1 - \mathbf{R_0}_2)^2 = (0.5)^2(1.2947 - 0.7528)^2 = 0.0734$. And the stochastic component component for the model with the costs ${\sigma^{\text{sto}}_{\mathcal{LRO}}}^2 = w_1.\sigma^2_{\mathcal{LRO}}[\mathbf{M_1}] + w_2.\sigma^2_{\mathcal{LRO}}[\mathbf{M_2}] = 0.5 \times 0.7071 + 0.5.4776 = 0.5923$ and without the costs ${\sigma^{*\text{sto}}_{\mathcal{LRO}}}^2 = w_1.\sigma^2_{\mathcal{LRO}}[\mathbf{M^*_1}] + w_2.\sigma^2_{\mathcal{LRO}}[\mathbf{M*_2}] = 0.5 \times 0.8461 + 0.5.5293 = 0.6877$. We can see here, that even though the costs raise the heterogeneity component of $\sigma^2_{\mathcal{LRO}}$ in the population from $\frac{0.0734}{0.6862} = 0.107$ to $\frac{0.0734}{0.594} = 0.123$ (as they keep the heterogeneity component unchanged), that the

demographic variance of a population is more driven by stochasticity than heterogeneity even for genotypes with differences in fitness.

3.0.9 Intermediary matrices for section 3.4.4

$$\mathbf{B} = \begin{bmatrix} 1 & -s & 0 & s & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ -s & s^2 & 0 & -s^2 & -s & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -s/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ s & -s^2 & 0 & s^2 & s & 0 & 0 & 0 & 0 & 0 & 0 & 0 & s/2 & 0 & 0 & 0 \\ 1 & -s & 0 & -s & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & -s/2 & 0 & s/2 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.54}$$

$$\mathbf{B}^* = \begin{bmatrix} 1 & 0 & 1-f/2 & 0 \\ 0 & 0 & 0 & 0 \\ 1-f/2 & 0 & (1-f/2)^2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.55}$$

Chapter 4

Kinship demography

Contents

4.1 Introduction

In this short chapter, we discuss kin transfers - the transfers of resources between related individuals of a population- in the light of the theoretical (chapter 1) and modeling (chapter 3) considerations of this thesis regarding costs of reproduction. We set our discussion in a framework that can study the coevolution of kinship and life history - i.e., both the effects of kin transfers on vital rates and the reciprocal influence of life history on kinship distribution - that we call *kinship demography*.

From the costs of reproductions to *kinship demography*

Kin transfers can, on the one hand, be considered to be a component of the costs of reproduction. Indeed we showed in chapter 1 (sec.1.2.1,p.19) that parental care, the postnatal transfer of time, food, energy and care from a parent (let us say, for simplicity, the mother) to its offspring, is a part of *physiological costs of reproduction*. Specifically, it emerged as the post-birth segment of the *reproductive effort schedule* (we denoted it *res*) representing, for one average adult in the population, the average time distribution of efforts required to produce one independent offspring, see figure 4.1. In organisms with protracted altricial juvenile period with respect to the mean interbirth interval, this (per birth) reproductive effort distribution will have a long tail and therefore the *reproductive effort schedule* corresponding to successive offspring will overlap, see figure 4.2. In chapter 1 (section 1.3.2 p.34) we devised a method to account for this characteristic by the segmentation of *res* into *reproductive efforts buckets* incorporated as *dynamic heterogeneity* traits in the model. Since the two effects of parental transfers - a cost to the mother and a benefit to the offspring - are included in the costs of reproduction, the most fundamental of kin transfers - parental transfer can be considered to be component of *physiological costs of reproduction*. This hints at the possibility to extend the conceptual models of chapter 1 (sec. 1.3.2, p.30) with the method of chapter 2 (p.41), and the tools of chapter 3 (sec. 3.3, p.68) to model *kinship demography* in an evolutionary framework.

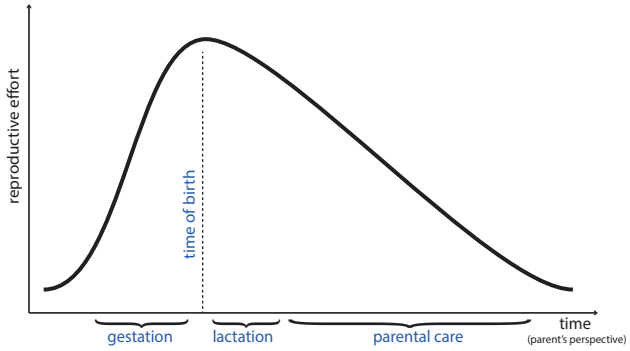


Figure 4.1: *Reproductive effort schedule* for altricial mammal. The figure display a typical *reproductive effort schedule* for an altricial mammal, with efforts on both sides of birth, but a long post-natal tail corresponding to parental care.

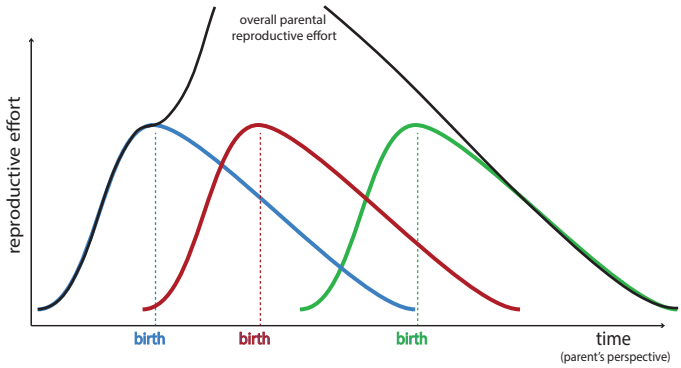


Figure 4.2: *Overlap* of *reproductive effort schedules* for altricial mammal. The figure shows that an inter-birth interval shorter than the duration of parental care implies the overlap of reproductive effort schedules

On the other hand however, the scope of kin transfers extends beyond the concepts of costs of reproduction for two reasons. First, because it is possible to contemplate kin transfers from the point of view of only one of the two actors: the payer or the beneficiary. This is often the case, as we shall see in the forthcoming review, but frowned upon by some theoreticians who consider it important to control for the net nullification of transfers in a population (i.e. to focus on the transfers between all pair of individuals in a population and not on the individuals themselves) (Lee, 2003). Second, and most importantly, because kin transfers, *apart* from parental care, cannot be implemented in a life-cycle: they connect individuals that are more than than a fertility event away from one another.

As a matter of fat, even kinship models that would implement both sides of transfers within a life cycle - i.e., the benefit (for the offspring) *and* the costs of maternal care - differ conceptually from the cost of reproduction model of chapter 1 (sec. 1.3.2, p.30). As the costs of reproduction are individual-based, there are only two ways to represent them in a life cycle. Either by implementing the full life history from birth until death and therefore forsaking post-natal maternal care (in a life cycle, the only connection between a parent and its offspring is at time of birth, via the fertility rate). Or by having the life cycle start at maturity and to transfer the premature period into the *reproductive effort schedule* (a parent then "gives birth" to a fully independent offspring, see discussion in sec.1.3.2 p.34). Whilst in the latter case, parental care is implemented, its effects on juveniles cannot be extracted from the model, as it only starts at the first age of reproduction. Simply put, in that case, juveniles are not individuals in the model, but mere demographic properties (components of the reproductive value) of the the mother. The simplest of kinship models - a maternal transfer model implementing both the costs of the mother and the benefits to the child - would therefore already be a significant step forward from the costs of reproduction model. This implies that the extension we call for, will need to incorporate traits that allow to embed either (allo)parental care or cost of care or both, and tools to account for the asymmetry of only implementing on side of transfers.

Therefore kinship transfers can both be considered as a special case (when considering the costs of maternal care) and, more generally, an extension (as transfers are about pairwise relationships and not individuals and extend beyond the mother-offspring connection to any alloparental care) of the costs of reproduction. In other words, kin transfers can be considered, in the light of (Williams, 1966)'s definition of the costs of reproduction, as individual allocative processes between current *inclusive* reproductive effort and future *inclusive* reproductive value (Hamilton, 1964b,a). This implies that similar models can be used, but that alterations are required. This also hints at the complexity of the conjoint study of kin transfers, life history and demography, the intricacy of which make up the field of *kinship demography.*

Kinship demography as the study of the coevolution of kinship and demography

In a social species, demography is a function of cooperative and competitive interactions between individuals. Growth, reproduction, survival of an individual depend on growth, reproduction and survival of other individuals in the population. On of the main factors structuring competitive and cooperative behavior is kinship. *Kinship demography* studies the interactions between demographic traits and kinship. First, by asking how kin relationships influence demographic traits, population dynamics and evolution of Life History traits. Second, in return, by studying how demographic traits structure kinship in a population.

First level of kinship demography would consist in the study of maternal/parental care/investment, common to many species, and thoroughly studied either directly or in the context of the costs of reproduction, of which it is an important component (as mentioned in previous section see sec. 4.1 and chapter 1 section qdsjdhsqhd). However in social species, and in particular in humans, family structures assign a caring function to other kin : grandparents, sibling which can act as helpers at the nest, uncles, aunts, cousins etc. In such cases, growth, survival and reproduction are both a function and a determinant of the family structure. Non-kin obviously also interact, raising the questions of the differing effects of kin and non-kin in alloparental care. In humans - see box table 4.1 - the sociality of the species, combined with its longevity, renders the study of *kinship demography* both complex and primordial to the understanding of our evolution.

In this chapter, we will therefore first describe the scope and the main characteristics of *kinship demography* in humans, a species both social and long-lived, see box table 4.1. The interactions between evolutionary demography and sociality imply that human *kinship demography* lies at the cross-roads of many fields, each considering specific aspects of the effect of kinship on demography, of demography on kinship, and of both on adaption, genetics, and human culture. We will discuss some of these fields and then go on to specifically review two specific but key components of human kinship transfers. First the determinant of child survival, and second, the costs of parental care.

In the second part, we will then hint at the initial steps to extend the MPPM framework for the costs of reproduction (chapter 1, sec. 1.3.2, p. 30) to implement kinship transfers.

Table 4.1: **Life history , parental investment and sociality in humans**

Primates have long average adult lifespans and few offspring relative to other mammals (Charnov and Berrigan, 1993). They are at the slow end of the *slow–fast* continuum observed in mammals. At this slow end of the continuum, juvenile survival is the most important fitness component and drives the evolution of other life-history traits (Heppell et al., 2000). Humans are among the "slowest" primates with respect to life history traits, with the exception of fertility. Relative to other primates, human females can bear a large number of children over a short reproductive period due to short birth spacing (Mace, 2000). The reproductive window is narrowed at one end because of protracted periods of infancy and childhood and, at the other end, because of reproductive senescence and eventual cessation at menopause. As a consequence, females may care for several children of different ages at the same time. This care includes feeding, protection, affection and education and is needed for the physiological and behavioral development of offspring. The relationship between child development and post-natal care exhibits two particularities in humans.
First, the psychomotricity and sensori-motricity of human neonates are immature relative to those of neonatal great apes (Parker, 1977), reflecting the fact that human gestation length is shorter than expected based on body size (Little, 1989). As a consequence, juveniles are thought to have become more dependent on post-natal care for survival during human evolution (Martin, 2007). Second, post-natal care is crucial for the development of the cognitive, linguistic and social capabilities particular to our species (Geary and Flinn, 2001) and juveniles have likely evolved to become increasingly dependent on post-natal care for education and socialization.

In humans, the primary caregiver is the mother. Her survival is a major predictor of child survival in many populations (Sear and Mace, 2008). While the role played by fathers, grandparents and older siblings in caring for infants and children have recently been the focus of a few studies (Bentley and Mace, 2009; Derosas, 2002), these have failed to show common patterns among human populations. It is however increasingly recognized that non-maternal care is nonetheless crucial for children survival because, in any human population, dying is not the unavoidable fate of a motherless newborn, which it is in other mammal species. Because post-natal care is protracted and occurs within complex social and generational structures, child survival depends on both the age and kinship structures of human populations. Going beyond age-trajectories of fertility and survival to incorporate these linkages within and across generations is therefore fundamental for understanding human population dynamics and the coevolution of life-history traits and sociality (Metcalf and Pavard, 2007); and is the subject of increasingly numerous theoretical and empirical research. Theoretical modeling predicts that human population growth is limited, not only by resources, but also because of the importance of age structure for resource acquisition (Lee, 2003). For example, the proportion of juveniles relative to adults increases in a growing population, up to a point where adults cannot produce enough resources to care for more immatures (Cyrus Chu et al., 2006; Lee and Tuljapurkar, 2008). From an evolutionary point of view, Pavard and Branger (2012) showed that an increase in newborn altriciality may have favored the emergence of human specific life history traits such as extended longevity and females' physiological capacity to give birth to a large amount of children and that maternal and grand maternal care may explain the emergence of menopause and post-reproductive life. However, the coevolution of maternal care with traits specific to humans, such as altricial infancy, long childhood, extended lifespan, short but intense reproductive period, menopause and post-reproductive life, is still poorly understood. More generally, the role played by familial investments on individual survival and reproduction throughout the individual life is still largely unknown.

In the third and last part, we will discuss approaches to extract, from a multitrait demographic model, the inferred kinship distribution of a population. And we illustrate this by calculating the effects of *physiological costs of reproduction* (as implemented in chapter 3, sec. 3.2, p.61) on kinship moments such as the mean inter-birth interval or the expected number of older sisters.

4.2 *Kinship demography* in humans : a short review

4.2.1 The large scope of human kinship demography

Economic anthropology has demonstrated the existence of a complex network of resources transfers between kin and non-kin in human societies. A large body of literature explores transfers among and between age classes and kin categories in terms of foraging, food-sharing, childcare or transfers of wealth. For example, while foraging trip composition is not based on kinship in Ache, food is primarily shared within the family (Gurven, 2004), although this also depends on the nature of the food (Gurven et al., 2004). This type of study gave birth to a function summarizing the net balance between resources acquisition and consumption at each age, which has then been incorporated into life-history models (Gurven et al., 2012; Kaplan and Robson, 2002; Lee, 2003).

These studies were the first to demonstrate the importance of inter-generational transfers in the evolution of aging in humans. Although remarkable, they do not analyze demographic data and they suffer therefore from strong assumptions linking the balance of resources production/consumption to demographic traits. Population dynamics and evolution studies consist mainly in counting births and deaths, and a demographic approach is required to measure to what extent kin investments influence individuals' survival and reproduction. In this context, most studies have focused on the effect of maternal care on infant and child survival (qualitatively reviewed in (Sear and Mace, 2008)). The fact that motherless children at birth are not doomed to a certain death in humans has however led to the recognition of the crucial role played by other caregivers than the mother: father, grandparents and older siblings (Bentley and Mace, 2009; Derosas and Oris, 2002; Sear and Mace, 2008). What are the implications and ramifications of kinship demography in humans?

History of the family, kinship ethnology and cultural anthropology Human is a species where kin are organized in families and family structure is organized by cultural rules. In recent books (Bentley and Mace, 2009; Derosas and Oris, 2002), the editors have gathered articles providing a demonstration of the complexity of allomaternal care and of the socio-demographic consequences of orphanhood across cultures and through time. For example in Mayas, Kramer (2009) shows the complex interactions between family dynamics and individual life-cycles with respect to childcare: mother, father and older siblings invest differently in direct child care or food/wealth production according to their age, but also according to family structure (mainly in terms of survival status of the parents and grandparents and siblings' number, age and sex). More generally, the study of familial determinants of demographic behaviors has allowed an "ecologically founded comprehension of cultural kinship" (Leonetti 2008) and, by allowing objective quantifications and tests of the qualitative theories, "strengthened and reinvigorate" the field (Shenk and Mattison 2011).

Statistical demography and population dynamics When adult mortality is high, the proportion of orphans is far from negligible (eg up to 30% of children are motherless when $e(15) = 15$ (Pavard and Branger, 2012)). As a consequence, recording data on orphans' survival is crucial for demographers; especially in population without registers where orphans are more difficult to record because of adoption, institutionalisation, or more often omission in retrospective data. For example, in populations experiencing HIV pandemics, accounting for the increased mortality of motherless children has proved crucial in estimating child survival (Mahy and Zaba, 2003) and HIV related mortality (Zaba et al., 2005). This has also important consequences in modeling population dynamics. For example, lower levels of allomaternal care may largely compromise population growth because motherless children are not properly taken care of (Pavard et al., 2007a) and kin transfers of resources between and within generations may modify long term population growth in constant (Lee and Tuljapurkar, 2008), finite (Puleston and Tuljapurkar, 2008) or fluctuating (Lee et al., 2009) environments.

Population genetics, quantitative genetics and medicine Cultural traits influence the fate of the different alleles in a population. This factor has therefore to be accounted for by population and quantitative geneticists and can have important medical consequences. This was exemplified by Austerlitz and Heyer (1998) who demonstrated that the socially driven heritability of reproductive success in historical Quebec had increased the frequency of inherited disorders in the population. The effect of social structures and kinship networks on genetic diversity is an integral part of *kinship demography* as it considers the cross influence of kinship distribution and demographic parameters on the drivers of

population genetics. In chapter 3 (sec.3.4.3, p.81), for instance, we showed that the basic principle at the core of parental care - the costs of reproduction - had consequences on the variance in reproductive success $\sigma^2_{\mathcal{LRO}}$. And whilst a portion of this variance (we denoted ${\sigma^{\text{het}}_{\mathcal{LRO}}}^2$) is due to genetic variance in the population, a larger component (${\sigma^{\text{sto}}_{\mathcal{LRO}}}^2$) is caused by stochastic variance intrinsic to heritable genotypes. This heritable variance affects, in turn, the effective size of the population, with consequences on the force of selection with respect to genetic drift. We also showed in chapter 3 (sec 3.4.2, p.77) that kinship care also affects the force of selection via the spreading of reproductive efforts over protracted periods of time. This overall effect of kin transfers on *both* components of effective selection has therefore important impact on human population and quantitative genetics and therefore on medicine. Age-structured population genetics allows, for instance, to disentangle the puzzle of the allelic spectrum of late onset diseases (see an article in prep. "Are old ages useless? Shedding light on the allelic spectrum of late-onset diseases" and presented in Annex 5.3.1, 142).

Life History theory and evolution of sociality Other domains are concerned with both the reciprocal effects of demography on kinship. The formation, the stability, the organizational structure, and the social dynamics of biological families has been theorized upon three conceptual evolutionary pillars: the ecological constraints theory (which focuses on offspring dispersal and kinship spatial structure), inclusive fitness theory (which analyzes how social interactions among family members optimized individuals fitness; Hamilton, 1964b,a), and reproductive skew theory (which analyze how reproductive conflicts may be solved in group-selection modeling) (Emlen, 1995). In humans, the coevolution of life-history and sociality has mainly been analyzed through the perspective of inclusive fitness theory (Hamilton, 1964b,a). Allomaternal care and cooperative breeding are seen as important reproductive strategies, as a buffer to child mortality in a context of rapid reproduction (Mace, 2000). Reproductive cessation at menopause in females has been explained by trade-off between giving up their own reproduction in exchange of greater survival of depending children (ie “the good mother theory”; eg Pavard et al., 2008; Peccei, 1995) or grand children (ie "the grandmother theory"; eg Hawkes et al., 1997); as well as by reduced conflict in reproductive competition with daughters and daughters in law (Cant and Johnstone, 2008). Maternal or grand-maternal investments, or, more generally, all transfers of resources from adults towards juveniles, have also been proven to increase tremendously the strength of natural selection on late survival and be a major factor explaining long life-expectancy in humans (Gurven et al., 2012; Lee, 2003; Pavard et al., 2007a). In Annex 5.3.1, 142 (an article in preparation), we demonstrate, via a population genetics model, how features of human social relationships - maternal care, grandmaternal care late male reproduction - influence the selection gradient with age and therefore the evolution of senescence.

4.2.2 Analysis of a receiver of kin investment : child survival

Maternal death compromises children survival in all demographic studies, most of the time to a large extent, although the magnitude of these effects varies in a large amount. The higher risk of dying of motherless children remains significant throughout childhood, although a decrease in dependency from maternal care as the child grows older is clearly visible (Pavard et al., 2007a). Results are largely contrasted in the case of fathers. In half the studies using appropriate statistical techniques, the father’s death makes no difference to child survival, and, when it does, it has substantially a lesser impact than the mother’s death (Sear, 2008). This may be because paternal care may vary with the sex of the child. For example, in a patrilineal Ethiopian community, father’s absence doubles a son’s risk of dying in infancy but has a positive influence on daughter survival (Gibson, 2008). It may also be that paternal care is more important at a later child age, a transfer little covered by these studies.

A meta-analysis of grandmaternal care by Strassmann and Garrard (2011) showed that maternal rather than paternal grandmother tend to have the larger effect on grand-child survival, even in patrilineal societies. No clear pattern can however be generalized across populations and no age-dependent effect is demonstrated. Little is known on the importance of siblings for infant and child survival despite the widespread observation that the labor of older children is used by parents both for domestic work (including child care) and productive activities (eg, Borgerhoff Mulder and Milton, 1985, , and for a review (Kramer, 2005)). A few contradictory studies show that older siblings have either a protector effect on younger siblings survival (ie “helper at the nest”; eg Sear 2008) or a deleterious effect (due to competitive relationships; eg Sparks et al., 2013). It is also acknowledged that these estimations are especially difficult to carry out because of the cofounding effect of clusters of sibling deaths due to shared environmental conditions (Ronsmans, 1995), and because siblings interactions are sex-specific (Sear, 2008). Moreover, it is thought that siblings may switch from helping to competing according to abundance of familial resources (Sear, 2008). Overall, whether or not these variations reflect differences in population behaviors with respect to child care is far from being clear for two reasons.

Statistical problems in estimating intrafamilial correlation in survival

First because of statistical problems in estimating intrafamilial correlation in survival. The study of child mortality following parental death presents many difficulties with regards to the distinction between causes and confounders. Three main factors may indeed increase both parents and children mortality, leading to clusters of parents–child deaths (Pavard et al., 2005; Ronsmans, 1995; Sear, 2008).

Between-family heterogeneity: All families are not equal with respect to mortality levels. It is well known by demographers that heterogeneity in mortality between families arise from socioeconomic conditions and behavioral characteristics of the parents (Sastry, 1997), as well as biological endogenous factors shared by the members of the same family (ie genetic and obstetric factors; Yerushalmy et al., 1956). Parents and children deaths may therefore cluster within a given family and, in this case, children deaths do not result from the loss of parental care following parental deaths.

Within-family heterogeneity. Children of the same family may have a differential risk of death linked to parental reproductive history. The most obvious is parental age at child birth. Children born to old parents have a larger risk of death due to decrease in gametes quality and increased birth defects. The risk of mortality varies also from one child to another with respect to birth interval, survival of the preceding child, sibship size and/or birth order (eg Ikamari, 2000; Kuate Defo, 1997). For example, birth intervals resulting from the death of the preceding child increases the mortality of both the index child and the mother because of maternal depletion syndrome (Jelliffe and Maddocks, 1964).

Sporadic increase of mmortality. Pavard et al. (2005) have first demonstrated that the mortality of children whose mother will die start increasing before maternal death. This is due to factors leading to a sudden rise in familial mortality such as accidents, epidemics, cross-infections or any familial reversal of fortune. Since then, this has specifically been analyzed in two articles confirming a bell-shaped increased risk centered on maternal death (Clark et al., 2013; Ronsmans et al., 2010).

Child dependency as a dynamic multi-actors process

Second because child dependency to parents and grandparents care is a dynamic multi-actors process involving child physiological and psychological development and mourning process. Indeed, most studies analyzed the increased risk of dying of a child whose parent dies into a given period. However, Beekink et al. (2002) have first shown that the time elapsed since the parent's death matters. For example, a child whose parent has died a long time ago has a higher survival rate than a child who has recently lost its parent (Pavard et al., 2005; Willführ and Gagnon, 2013). The increased mortality due to the loss of parental care is therefore a dynamic process depending on both the current child age and the child age at the death of the parent(s). Moreover, the effect of a parent's death depends also on the family structure both before and after the death of the parents. Curiously, only one study demonstrates the obvious: the effect of a parent's death varies according to whether and which other parents are still alive (Derosas, 2002). Step-parents also exhibit various effects. For example, the mother's remarriage has no effect in historical Germany, while it is beneficial for children survival in historical Québec. In contrast, the father's remarriage dramatically reduces children survival in historical Germany, while such an effect is not seen for historical Québec (Willführ and Gagnon, 2013). This is fundamental because it gives indication on who is compensating, or not, for the loss of parental care after the death of a parent. This type of analysis holds the only way to untangle the multi-actor network of investments that allows children to reach maturity in humans.

4.2.3 Analysis of a payer of kin investment: human costs of parental care

Optimal age-trajectories of adult survival and reproduction have been extensively formalized by life-history theory based on two main trades-off: (i) the costs of reproduction : the trade-off between present reproduction and future survival/reproduction; (ii) the trade-off between quantity and quality of offspring. This theoretical framework has proved efficient in explaining evolution of life-history across a large range of species, including species where post-natal parental investment is required for improving offspring quality at the cost of parental future reproduction (Clutton-Brock, 1991; Gross, 2005). In humans however, the empirical demonstration of trade-off between survival and reproduction has proved difficult (Gagnon et al., 2009) and reproductive success has been reported to vary both positively and negatively with women adult survival (reviewed by Larke and Crews, 2006). One of the reasons is that individual heterogeneity may mask this trade-off because robust individuals that survive better may also be those that reproduce the most. This heterogeneity may also be familial resulting from socio-economic differences. For example, Lycett et al. (2000) showed that trade-off between reproductive success and survival was only detectable in the lowest economic class in historical Germany (see also Gillespie et al., 2008, for historical Finland).

Another reason is because family investments have rarely been incorporated in these analyses (Larke and Crews, 2006). Reproduction does not ensure reproductive success in humans and parental investment

is required for children survival until maturity . Giving birth to a large number of children does not mean paying large costs in terms of parental investment if most of them die. As a consequence, intergenerational transfers has proved crucial in optimizing quantity/quality trade-offs faced by human females (Borgerhoff Mulder, 2000). Understanding familial strategies to enhance the competitive success of offspring is therefore needed to understand the balance between fitness components throughout the males and females life course (Lawson and Mace, 2011) and thus the peculiar feature of human reproductive life history. Reproductive schedule in human females is strikingly faster than that of other great apes, even when body size is controlled for (for example, interbirth intervals is 3 yrs in humans against 4.5 years in chimpanzees and 8 years in orang-utans; Mace, 2000). This rapid pace of child bearing could be even physiologically faster, demographers and anthropologists having long recognized that all human societies limit birth rates to some extent, ensuring that few women reach the biological maximum, even under the most favorable conditions (Bongaarts, 1975). As a consequence, parents have an increasing number of offspring at different stages of dependency, needing to be cared for simultaneously.

To achieve this, mothers need help with childcare and nutrition, whether it is coming from male provisioning or post-reproductive females. However, if there is increasing empirical evidence that a mother gets help in caring for her children (from the studies associating child survival with the presence or absence of family members discussed previously), the fact that this help allows a mother to reproduce more is less clear; although there is an extended body of literature on the topic Lawson and Mace (2011). Effect of kin investments on child survival and females' fertility may even be opposite. For example, in rural Gambia, Sear et al. (2003) found that parents in law increased a female's fertility while her parents and her elder sisters had no effect. This is the opposite of what had been found regarding children survival on the same population (Sear et al., 2002): maternal grandmother or elder sisters had a significant positive effect on the survival probabilities of children, whereas paternal grandparents had no effect. The answer to this apparent paradox may be that relationships between females in this population may be structured by age and dispersal: females compete only when they are reproductively active and co-residing in the same compound (Mace and Alvergne, 2012). To our knowledge, the only study that demonstrates a clear and coherent effect of grandmothers is that of Lahdenpera et al. (2004) for two historical populations, where post-reproductive mothers allow their children to breed earlier, more frequently and more successfully.

The role of siblings is even more ambiguous. In historical Finland, the presence of elder siblings improved the chances of younger siblings to survive until sexual maturity. However, after reaching sexual maturity, same-sex elder siblings' presence was associated with reduced adult reproductive success (Nitsch et al., 2013). In Historical Germany, socioeconomic condition matters: children's probabilities of marrying or emigrating unmarried are affected by the number of living same-sexed sibs in farmers' families but not in the landless laborers (Voland and Dunbar, 1995). To the contrary, the number of older siblings are a strong predictors of fertility in !Kung of Botswana, especially for males (Draper and Hames, 2000). More generally, Mathews and Sear (2013) analyzed results from a large body of literature and showed that, the extent to which the presence of kin (parents, parents-in-law, siblings) is correlated with fertility, is mainly a matter of environment and culture. For example, levels of intergenerational transfers of wealth depend on subsistence mode (substantial among pastoral and agricultural societies and limited among horticultural and foraging peoples; Mulder et al., 2009). In return, it has also been suggested that kin investment is a key factor in explaining emergence of cultural kinship practices. For example sex-specific competition for household resource, in matrilines between mothers/daughters/sisters (Mace, 2013) or in patrilines between fathers/sons/brothers (Ji et al., 2014) may have underpinned the cultural evolution of marriage, residence, and inheritance norms (such as late male marriage or primogeniture). Unfortunately, to date, no cross-cultural study succeeds to demonstrate the precise role of social organization (in particular according to descent, marital and residence rules) in determining who help mothers to bear children, and at which cost.

4.3 Implementing *kinship demography* in evolutionary models

In order to flourish, *kinship demography* requires models that can implement kin transfers - whether accounting for both sides of the transfer or only the payer/receiver side - for a wide range of kin relationships.

4.3.1 Precursor models

To achieve this, our concepts will be derived from the costs of reproduction model of chapter 1, but will also stem from two precursors: Lee (2003)'s intergenerational tranfers model and the ad-hoc evolutioanry kinship models from (Tuljapurkar et al., 2007; Pavard et al., 2008; Pavard and Branger, 2012).

Lee's models of age-structured intergenerational tranfers In 2003, Ronald Lee proposed a theoretical model integrating transfers in an aged-structure population (Lee, 2003). In this model,

great importance is attached in implementing both sides of transfers: the net balance between resources production and consumption of individuals of age a is given by a transfer function. If positive (i.e., individuals are producing more resources than they consume), excess of resources is transferred to younger-age classes. Among others results, assuming a link between resources consumed and survival, makes juveniles (who consume more than they produce) dependent on resources produced by adults for their survival. As a consequence, this also models positive and negative correlations of survival and reproduction between age-classes. This theoretical framework has led to emphasize the fundamental role of transfers in evolution of aging (Cyrus Chu et al., 2006) but also to better understand what may have been human population dynamics in constant (Lee and Tuljapurkar, 2008), fluctuating (Lee et al., 2009) and finite (Puleston and Tuljapurkar, 2008) environments. We may however address two critics to these models. First, these models become rapidly unsolvable when complexity is increased by addition of other traits than age. For example, when adding sharing groups in his transfer of resources structures, Lee (2008) had to switch to agent-based models because non-linear equations were not solvable. Second, these models assume these transfers are unrelated to kinship, whilst we know that, in humans, they are in reality mainly (but not only) structured by kinship. We may wonder to what extent the authors' results still hold when kinship is incorporated, together with age, in the model.

Early multitrait kinship models Recently, the implementation of kinship in evolutionary demography models has become increasingly topical. For instance, Samuel Pavard has shown, by implementing different vital rates for orphans and non-orphans, that maternal care is not only crucial for explaining emergence of menopause in humans (Pavard et al., 2008) but also for population dynamics (Pavard et al., 2007b) and evolution of life history traits throughout an individuals life (Pavard and Branger, 2012). For example, Pavard et al. (2007a)) incorporated maternal care into an Hamiltonian framework of evolution of senescence (Hamilton, 1966). The authors showed that the whole gradient of selection on age-specific survival and reproduction changes dramatically in magnitudes and shape by age compared with a species where maternal care is not (or less) required for children survival. Using a numerical equivalent of the *Trait Level Analysis*, this has allowed the authors to provide a new coherent scenario for the emergence of traits peculiar to humans, such as the extended juvenile period, the intense but short reproductive period for females, menopause and long post-reproductive life (Pavard and Branger, 2012). In parallel another multitrait model was used to implement kinship: the sex and age model by Tuljapurkar et al. (2007). This model allowed the authors to implement husband-wife transfers and provided interesting evolutionary results on the effect of mating patterns on human longevity.

4.3.2 A new general template for Kinship Models

From these early efforts, we know that a general kinship model therefore requires the incorporation of kinship traits. In the first chapter of this manuscript (sec. 1.3.2 p.30), we have discussed a general method to incorporate all aspects of a trade-off in a multitrait population projection matrix (MPPM, see chapter 2 (p.41)). We have shown that it meant incorporating three families of traits in the model. The basic traits $\mathcal{B}$ that best determine the life-history of the organism, the *dynamic heterogeneity* traits $\mathcal{D}$ modeling the physiological (mechanistic) side of the trade-off and the *hidden heterogeneity* family of traits $\mathcal{G}$ modeling the genetic trade-off (the genotypic polymorphism). In chapter 1, this was applied to the most prominent of all trade-offs : the costs of reproduction. In the wide acceptance of the term, these costs include all post-birth investments provided by a parent, a cost known as parental care. In this section, we wish to indicate modeling paths allowing to extend the trade-off MPPM framework towards an evolutionary kinship models incorporating kin transfers.

The general model.

To do this, we first simplify the framework by focusing on the mechanistic constrains - the dynamic traits that track resources capitalized by individuals via $\mathcal{B}$ and $\mathcal{D}$ - forsaking $\mathcal{G}$ the genetic variance in the population (but keeping in mind a genotypic polymorphism can be later added as theorized in chapter 1 and performed in chapter 3). When modeling transfers, the resources provided to another individual or received from it can be considered to be acquirable (for instance food transfers as in (Lee, 2008)) or not (if for instance transfers are about time spent educating kin). In the first case, transfers could be modeled by the Fluctuating Capital (FC) described in chapter 1, and the second case by the Ratchet Capital (RC). In practice these resources transferred may be modeled by the addition of a special trait NTC that would account for the net transfer received by an individual since birth (Lee, 2003). Capital NTC may be merged with FC/RC in order for vital rates to depend on the total amount of resources, acquired and net received. NTC could also be kept segregated from FC/RC if total transfers at the level of the population are required to cancel out (Lee, 2003). The effect of kin transfers can then be implemented by implementing vital rates as functions of capitals NTC/FC/RC, themselves function of the environment, the life trajectory of the individual and its kinship distribution. In a simplified version of the model focusing on kinship only, vital

rates can just be direct functions of the distribution of kin alive. In all cases therefore, this requires us to track kinship distribution.

In order to be able to analyze the demographic and evolutionary consequences of kinship, a kinship model requires first to identify the kin providing care to an individual and to add these as active *dynamic heterogeneity* traits : the $\mathcal{K}$ kinship traits. This turns a kinship model into a *memory* model where the entire kinship distribution is tracked time-step after time-step. Via a survival transition, this distribution will be affected by basic trait $\mathcal{B}$: all kin alive will survive one more time-step or die according to their basic trait in $\mathcal{B}$. Via a fertility transitions, the kinship distribution is transferred and shifted, according the basic genealogical rules, from mother to child.

Illustration

Let us illustrate this with the example of an aged-structured population model ($\mathcal{B} = \{age\}$) that implements the (positive) effect of older sisters and aunts on fitness. Let $r = \beta - \alpha + 1$ be the number of fertility events of the age-structured population. Then kinship family of trait $\mathcal{K}$ would consist in the position k of individuals' birth in their mother's reproductive period, m the aliveness of the mother of individuals, position k^* of the mother's birth in the grand-mother's reproductive period and 3 r-tuples of traits : $\{s_i\}_{i\in(1,r)}$, $\{d_i\}_{i\in(1,r)}$ and $\{t_i\}_{i\in(1,r)}$ providing the aliveness of the i^{th} sister, daughter and aunt (in their respective sisterhood) of the individuals. This therefore implies the model has $3r + 4$ traits and its trait vector and trait structure are $\mathbf{t} = (age, k, m, k^*, s_1, \ldots, s_r, d_1, \ldots, d_r, t_1, \ldots, t_r)$ and $\mathbf{s} = (\omega, r, 2, r, 2, \ldots, \ldots, 2)$.

As the individual - we call *ego* - survives to the next age-class (as a function of $\mathcal{K}$), *age* increases by 1. k is constant but provides the age of the mother $age + k$, which in turn yields, from a baseline survival function $s^\dagger$ by *age* defined for a mean individual in the population, the state m of the mother. Similarly, the position of each of *ego*'s i^{th} aunts (respectively daughters, sisters) is constant, but its age $age+k+k^*-i$ (respectively $age - \alpha + 1 - i$ and $age - i$) which yields her aliveness at the next time-step t_i (respectively d_i and s_i). Equivalently, the baseline fertility function$f^\dagger$ by *age* defined for a mean individual in the population, will provide new values for the younger sisters and aunts, unborn at ego's birth $\{t_i\}_{age+k+k^*\leq i\leq r}$ and $\{s_i\}_{age+k\leq i\leq r}$.

As the mother produces an offspring, her, yet unexisting, $d_{age-\alpha}$ daughter becomes alive. Moreover, her sisters become the child's aunts and her daughters the child's older sisters. she transfers her $\{d_i\}_{i\in(1,r)}$ (respectively $\{s_i\}_{i\in(1,r)}$) into her daughter's $\{s_i\}_{i\in(1,r)}$ (respectively $\{t_i\}_{i\in(1,r)}$). The daughter is also assigned the following values: m is set at 1 (mother is alive) , $r = age - \alpha + 1$.

And therefore, thanks to trait family $\mathcal{K}$ and basic family $\mathcal{B}$ (on which baseline vital rate function $f^\dagger$ and $s^\dagger$ are defined) , the distribution of kin is tracked throughout the life history trajectory of individuals and can therefore be made to affect vital rates : $f(age, m, s_1, \ldots, s_r, d_1, \ldots, d_r, t_1, \ldots, t_r)$ and $s(age, m, s_1, \ldots, s_r, d_1, \ldots, d_r, t_1, \ldots, t_r)$. This can be done by considering, for instance, an additive impact (each kin alive has a positive effect, which effect can be weighted, for instance, by the coefficient of relationship), or a multiplicative impact (to implement, for example, the fact that the role of aunts is only positive if the mother is dead).

Solving the transfer asymmetry issue

The model, as it stands, suffers however from one major inconsistency. The model implements vital rates of individuals of a population -f and s - as affected by basic traits $\mathcal{B}$ and by the distribution of their kin. However the fate of their kin only depends on $\mathcal{B}$ -via $f^\dagger$ and $s^\dagger$ - whilst they are obviously also part of the modeled population. This is due to the fact that we implement kinship only *one-way*: the receivers benefit from transfers provided by they kin alive, whilst the payers bear no cost. In order to correct the erroneous results such a model with non-zero net transfer would yield (Lee, 2003), there are two possibilities.

The first possibility is to implement kinship transfers *two-way*. That is incorporating the positive effect of aunts on nieces, together with the negative effect of nieces on aunts. Embedding both the benefits of having older sisters, and the costs of caring for younger ones. By netting the transfers and turning the model into a symmetrical design, the side effects of the baseline vital rates will also be (partially) netted as they will be used to implement both costs and benefits. In a *two-way* kinship model, the net transfer capital NTC aforementioned - increased (or decreased) at each time-step by the net kinship-coefficient-weighted number of kin - should be added as a trait. It allows to segregate the effects of kinship on net transfers from its effect on vital rates. By construction, the NTC whilst evolving for each individual, will be constant at the population level (Lee, 2003): there are as many individuals aged x with and aunt aged y, than individuals aged y with a niece aged x. This does not mean that the population is necessarily stationary. Indeed, if vital rates are asymmetrically affected by kinship (for instance if the transfer benefits

more the niece than it costs its aunt) the population can grow.

The second possibility, if *two-way* kinship tracking is too tedious or impossible to implement, is to make the *one-way* model consistent by convergence. Indeed, implementing the status of kin requires to add potentially numerous traits to $\mathcal{K}$. Implementing the status of kin of kin would require the square of these numbers and so on and therefore in a kinship model the fate of kin do not depend on $\mathcal{K}$ but only on $\mathcal{B}$. We called these vital rates, baseline fertility and survival rates $f^\dagger$ and $s^\dagger$. However the kin of individuals are part of the same population and should therefore share the same vital rates on the characteristics in which they are defined. Simply put, the baseline vital rates should be the expected vital rates in the population for kin's $\mathcal{B}$ characteristics, i.e. the mean vital rates, we denote f° and s°.

This consideration is however circular - the mean vital f° and s° result from the model, itself a function of the baseline vital rates $f^\dagger$ and $s^\dagger$ - and the model cannot therefore be implemented directly. In order to ensure that $f^\dagger$ and $s^\dagger$ *are* the mean vital rates f° and s°, a convergence process is required. Let us call $\mathbf{M}(f^\dagger, s^\dagger)$ the kinship model, where the population is characterized by $\{\mathcal{B}, \mathcal{K}\}$, implementing kin benefits - via functions f and s defined on $\{\mathcal{B}, \mathcal{K}\}$ - and where kin vital rates are represented by $f^\dagger$ and $s^\dagger$ defined on $\mathcal{B}$. Then the mean vital rates of the population defined on $\mathcal{B}$ can be extracted from $\mathbf{M}_\mathcal{B}^{\mathbf{fold}}(f^\dagger, s^\dagger)$, which is $\mathbf{M}(f^\dagger, s^\dagger)$ *folded* on all traits in $\mathcal{K}$ (see chapter 2 (p.41)). These two matrices are ergodically equivalent, and share the same asymptotic abundances when characterized by traits in $\mathcal{B}$. The kinship transitions of $\mathbf{M}(f^\dagger, s^\dagger)$ are averaged (weighted by their asymptotic relative abundances) to yield the transitions of $\mathbf{M}_\mathcal{B}^{\mathbf{fold}}(f^\dagger, s^\dagger)$. In other words, the mean vital rates of the populations are those emerging from the *folded* matrix, such that: $f^\circ = f(\mathbf{M}_\mathcal{B}^{\mathbf{fold}}(f^\dagger, s^\dagger))$ and $s^\circ = s(\mathbf{M}_\mathcal{B}^{\mathbf{fold}}(f^\dagger, s^\dagger))$.

By considering $\mathbf{f}_t$ and $\mathbf{s}_t$ as series of vectors of size, the size of $\mathcal{B}$, then - if converging - the following iterating *folding* process

$$\begin{cases} \mathbf{f}_0 = f^\dagger \\ \mathbf{s}_0 = s^\dagger \\ \mathbf{f}_{t+1} = f(\mathbf{M}_\mathcal{B}^{\mathbf{fold}}(\mathbf{f}_t, \mathbf{s}_t)) \\ \mathbf{s}_{t+1} = s\mathbf{M}_\mathcal{B}^{\mathbf{fold}}(\mathbf{f}_t, \mathbf{s}_t)) \end{cases} \tag{4.1}$$

would tend towards $\mathbf{M}(\mathbf{f}_{+\infty}, \mathbf{s}_{+\infty})$, which is the desired kinship model. This is another application of the *Trait Level Analysis* described in chapter 2 (chapter 2 (p.41)).

These theoretical and practical considerations are only first steps towards a general theory of kinship MPPMs. Further analysis will benefit from comparing the outputs of Lee (2003)'s transfers models by age in an evolutionary demography framework that segregates age and transfer structures, with ours. The economics field of input-output analysis, put in matrix form by Leontief (1951) - and which 20 years before demographers had started to investigate the addition of traits on a projection matrix (see chapter 2 (p.41)) - is likely to provide help here, once again, especially as it integrates transfers in the framework of input-output matrices (Leontief, 1936; Chichilnisky, 1983; Lopes and Neder, 2017).

4.4 Extracting *kinship demography* from evolutionary models

In the previous section, we showed how to implement kinship in an MPPM to build what we call a *kinship model*. But *kinship demography* requires also to be able to measure the effect of demography on kinship. We will discuss here, methods - first particular for particular kinship then in the general case - to measure the general effects of vital rates on kin distribution. We will illustrate these by drawing on the model implementing costs of reproduction of chapter 3.

As we saw in that chapter, the simplifying choices made when building a population model are not neutral. They need to be kept in mind when asserting the results provided by the model. Some of the implications of these simplifications are obvious. Characterizing a population only by age, for instance, implies that all vital rates depend only on age, and are therefore, implicitly, independent from one another, which leads to further implications rarely accounted for. For instance, we saw (sec.3.4.3, p.85) that the independence of successive fertility processes (as Bernoulli r.v.) violates the simplified model of population genetics, which requires that, on average, every individual in the population has the same probability of being the parent to any offspring of the next generation (Kimura, 1958; Hill, 1972; Felsenstein, 1971). The effect of age (only)-structuring can then be measured, for instance, by comparing the variance in lifetime reproductive success stemming from the Leslie model (sec.3.4.3, p.85) with the variance assigned to it by the Wright-Fisher population genetics model (that will be higher, the variance of a Poisson distribution being its mean , whilst the variance of a Bernoulli/Binomial distribution is lower than the mean), and the actual

number in the natural population (that we expect to be lower due to the underlying trade-offs, and in particular the *physiological costs of reproduction* which we studied the effects of variance in $\mathcal{LRO}$ in chapter 3).

But the consequences of the simplifying choices of model do not stop at family size (as $\sigma^2_{\mathcal{LRO}}$ is often called), they impact the entire kinship distribution. The aliveness (and number) of an individual's mother, sisters, aunts, grand-mother, daughters, etc. are functions of the vital rates of the population. A population structured only by age (i.e., modeled by a Leslie matrix), for instance, automatically generates a kinship distribution in the asymptotic population. We shall now consider historical approaches to measure the kinship distribution inferred by a population model. Then we will provide examples for specific kin like the expected number of older sisters at birth. This quantity is important as these older sisters are, in some social species and especially in humans, helpers at the nest, providing care to the individual, and along with other kin of close relatedness with the individual, replacing the mother in this role when she dies. And finally we hint at a generalization of such methods to any kin.

4.4.1 Precursor methods

Different fields have long been looking for mathematical predictions of kinship structure as a result of population demography. The first demographers to attack the problem based their computations on age-structured populations and resorted to stable-state theory (Le Bras, 1973; Goodman et al., 1974b,a). From the relative ergodic abundances by age-class, they would draw for an individual *ego* taken at random in the sable-state population, the distribution of the age and (alive) status of its mother, grandmother, etc. and from these the distributions of sisters, aunts, etc. The approach is very interesting but the computations are analytic, strenuous and adhoc. These draw-backs meant that these methods were quickly abandoned, with a new focus on a simplified structure where the multiple random processes corresponding to fertility and survival rates at each age were replaced by two random variable only : longevity and net reproduction (Pullum, 1982). Deemed too complex to implement still, it prompted demographers to turn to individual-based modeling (called microsimulation in that field) to simulate the projection of kinship networks over time (Pullum and Wolf, 1991; Reeves, 1987; Smith, 1987; Wachter et al., 1997).

In parallel mathematicians attacked the problem from the field of Branching Processes. First by considering life history as a Birth and Death process (the BP equivalent to the two vital processes - longevity and net reproduction - just described) (Waugh, 1955), by setting the life cycle aside completely and considering non-overlapping generations (this corresponds to the Galton-Watson BP process Waugh, 1981; Joffe and Waugh, 1982), and finally by considering age as trait (corresponding to multitype BP Joffe and Waugh, 1985).

Therefore whilst mathematicians augmented the complexity of their object, to finally start studying the BP equivalent of a Leslie matrix, the complexity of the probabilistic approach meant demographers could not extract the desired quantities from such models, and proceeded on their side to simplify their framework from an evolutionary age-structured model to IBM simulations. As of today, the distributions provided by Goodman et al. (1974b) are therefore still the most accomplished kinship inference method in a structured population, and constantly referred to (Lahdenperä et al., 2012). Our aim here is to take up the torch and to reinvigorate this field by hinting at the first steps towards a general kinship extraction method.

4.4.2 Neutral trait approach: application to inference of the inter-birth interval from matrix model

For simple kinship-related measures, the computations and the analysis can be made even simpler by the addition of a neutral *dynamic heterogeneity* trait, i.e. a trait that acts as a tracker and indicator of an individual's state but has no effect in itself on vital rates. Among such neutral traits that can be readily appended to any structured model is *age* (tracking chronological time in a *stage*-structured population for instance) and *parity*. *Parity*, as a neutral trait, could be implemented to produce ergodic abundances or calculate reproductive values as functions of it. It can also benefits the calculation of kinship related measures in the population, like the inter-birth interval.

As a matter of fact, in order to compute the expected inter-birth interval that is determined by a population characterized by trait structure $\boldsymbol{t}$, we add to it two *neutral* traits: *parity* and *newparity*, that indicates if the individual just had a successful fertility event at the last time-step. From the initial model (characterized by, say, age, denoted $\mathbf{a}$, we build $\mathbf{M_{a,p,np}}$ the model with *parity* (we denote $\mathbf{p}$) and *newparity* (denoted $\mathbf{np}$) counters added as traits. In order to do so, it required to split each survival transitions into two transitions according to the success of the fertility event which parameter is the fertility rate, as we did for the model of chapter 3 (sec.3.2.2, p.64). Fertility successful survival transitions increment *parity* and assign a 1 to *newparity*. Fertility unsuccessful survival transitions keep *parity* constant and assign a 0

to *newparity*.

The full matrix can be split into fertility and survival components: $\mathbf{M_{a,p,np}} = \mathbf{F_{a,p,np}} + \mathbf{T_{a,p,np}}$. It can also be *folded* on *newparity*, yielding $\mathbf{M^{fold}_{a,p}} = \mathbf{F^{fold}_{a,p}} + \mathbf{T^{fold}_{a,p}}$. Eigen analysis of $\mathbf{M_{a,p,np}}$ allows to extract the vector of relative abundances $\mathbf{w_{a,p,np}}$ from which we extract the second half, scaled to one : $\mathbf{w}^* = \frac{\mathbf{w_{a,p,np}}_{np=1}}{\mathbf{1}'.\mathbf{w_{a,p,np}}_{np=1}}$. Vector $\mathbf{w}^*$ provides the distributions, at the stable state, of the *age* and *parity* of the individuals which just had a successful fertility event at the previous time step.

We are now required to compute, for each of these *age* × *parity* combinations, the expected inter-birth interval, i.e. the expected time spent in the current parity. From $\mathbf{T^{fold}_{a,p}}$, we can generate $\mathbf{T}^\diamond$ - such that $T^\diamond_{i,j} = \frac{{T^{fold}_{a,p}}_{i,j}}{\sum_i {T^{fold}_{a,p}}_{i,j}}$ - which is the matrix of survival transitions of the model characterized by *age* and *parity* and where *death* is not accounted for. This is implemented so that mortality does not interfere with the computation of interbirth interval (if one wishes to take it into account in the computation, then the computation of $\mathbf{T}^\diamond$ is unnecessary and $\mathbf{T^{fold}_{a,p}}$ can be used directly in what follows). From $\mathbf{T}^\diamond$, we can compute fundamental matrix $\mathbf{N}^\diamond = (\mathbf{I} - \mathbf{T}^\diamond)^{-1}$ which provides the required expected time that an individual (characterized by *age* and *stage*) is expected to spend in subsequent age and parity states as its life trajectory unfolds. From these states we are only interested in those sharing the same parity than the input state, matrixwise, this means we are only interested in the diagonal blocks of

$$\mathbf{N}^\diamond = \left[\begin{array}{c|c|c|c} \mathbf{N_{p_1,p_1}} & \mathbf{N_{p_1,p_2}} & \cdots & \mathbf{N_{p_1,p_{pmax}}} \\ \hline \mathbf{N_{p_2,p_1}} & \mathbf{N_{p_2,p_2}} & \cdots & \mathbf{N_{p_2,p_{pmax}}} \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline \mathbf{N_{p_{pmax},p_1}} & \mathbf{N_{p_{pmax},p_2}} & \cdots & \mathbf{N_{p_{pmax},p_{pmax}}} \end{array}\right]$$

, of which we sum the columns : $\mathbf{t} = \left[\begin{array}{c|c|c|c} \mathbf{1}'.\mathbf{N_{p_1,p_1}} & \mathbf{1}'.\mathbf{N_{p_2,p_2}} & \cdots & \mathbf{1}'.\mathbf{N_{p_{pmax},p_{pmax}}} \end{array}\right]$. Vector $\mathbf{t}$ represents the expected time an individual, in a given *age* × *parity* combination will remain in the current parity. From these we yield the desired formula for the mean birth interval at stable state:

$$ibi = \mathbf{t}'.\mathbf{w}^*$$

4.4.3 Genealogical Markov chain approach

For most kinship inference calculations however, the addition of *neutral* traits is not sufficient to provide the expected number, and we have to use the properties of genealogical Markov chains, as we shall illustrate, here, for the computation - in a one-sex one-offspring-class model with independent vital processes - of the expected number, at birth, of older sisters. A simplified version of this quantity - the expected number of *ever-born* older sisters - can be computed via the addition of *neutral* trait parity: the *parity* of the *ego*'s mother at its birth is *ego*'s number of ever-born older sisters. We illustrate this, by providing, in appendix section 4.6.1, the effects of *physiological costs of reproduction*, as modeled in chapter 2, on that kinship measure. In order to compute the expected number of older sisters alive however - an arguably more meaningful measure from a *kinship demography* standpoint - we have resort to the genealogical Markov chains.

Application to inference of the expected number of older sisters from matrix model

Let us call $\mathbf{M}$ such a model with ergodic growth rate λ and related abundance vector $\mathbf{w}$. From these, we can derive its related backward genealogical Markov chain $\mathbf{P}$, a concept that was introduced by (Demetrius, 1974, 1975) and later used by (Tuljapurkar, 1982c, 1993) and Bienvenu (Bienvenu and Legendre, 2015; Bienvenu et al., 2017). Matrix $\mathbf{P}$ represents the asymptotic distribution of input states for each possible output state. Simply put, $P_{i,j}$ is the probability that an individual observed in class i at time t comes from class j : $P_{i,j} = \frac{M_{i,j}}{\lambda}.\frac{w_j}{w_i}$.

Let us split $\mathbf{M}$ into the matrix of fertility transitions $\mathbf{F}$ and the matrix of fertility transitions $\mathbf{T}$, and let us split $\mathbf{P}$ into its first line $\mathbf{d}$ ($d_i = \mathrm{P}_{1,i}$) and the rest of the matrix $\tilde{\mathbf{P}}$. In a 1-sex 1-offspring-class model, like $\mathbf{M}$, $\mathbf{d}$ is thus the asymptotic distribution of maternal state. For each maternal state i at birth of individual (ego), the expected number of daughters produced by the mother t time-steps before birth of ego, is then the first and only element of $\mathbf{F}\tilde{\mathbf{P}}'^t\boldsymbol{e}_i$ ($\boldsymbol{e}_i$ is a vector of zeroes with 1 in i^{th} position). Therefore the number of ever-born older sisters to ego is $\sum_t \mathbf{F}\tilde{\mathbf{P}}'^t\boldsymbol{e}_i = \mathbf{F}(\sum_t \tilde{\mathbf{P}}'^t)\boldsymbol{e}_i$. Thus, the vector of expected number of older sisters ever born - characterized by state of mother - is $\boldsymbol{e}'_1.\mathbf{F}\sum_t \tilde{\mathbf{P}}'^t$ (the first line of $\mathbf{F}\sum_t \tilde{\mathbf{P}}'^t$). Similarly, the expected number of older sisters alive at birth of ego, distributed by state of mother i, is $\sum_t \mathbf{T}^t\mathbf{F}\tilde{\mathbf{P}}'^t\boldsymbol{e}_i$ and thus, the vector of expected number of big sisters alive , distributed by age of mother, is the first line of $\sum_t \mathbf{T}^t\mathbf{F}\tilde{\mathbf{P}}'^t$.

We can extend this approach to other kin. The distribution of combination of states at birth for mother and grandmother is given by $\mathbf{D} = \boldsymbol{d}.\boldsymbol{d}'$, where $\mathrm{D}_{i,j} = d_i.d_j$ the probability that mother give birth to ego at state $i = (i_a, ..)$ (i_a is the age component of state i) and that her own mother had her at j (this general approach stemming from (Goodman et al., 1974b) is used, in an equivalent but non matrix form in Annex 5.3.1, p. 142. The distribution of state of the grandmother at birth is given by $\mathbf{T}^{i_a}\boldsymbol{e}_j$

Aunts distribution can also be computed this way. Aunts are either born between at $(i_a+j_a) \geq t \geq (i_a+1)$ and the vector of expected number of aunts older than mother, alive, distributed by state of grandmother j, is $\sum_t \mathbf{T}^{t-i_a-j_a}\mathbf{F}\tilde{\mathbf{P}}'^{t-i_a}\boldsymbol{e}_j$. Or they born between at $(i_a-1) \geq t \geq 1$ then the vector of expected number of aunts younger than mother, alive, distributed by state of grandmother j, $\sum_t \mathbf{T}^{t-i_a-j_a}\mathbf{F}\mathbf{T}^{t+j_a}\boldsymbol{e}_j$. Such calculations (applicable to any matrix model, as soon as the independence is respected, as is the case for most matrix modes), can be extended to further and further kin, thus, in practice drawing the outline of a *kinship matrix* providing the distribution of kinship that can be expected at stable state for an individual picked at random.

Word categorization of individuals in a structured-population's genealogy

To understand how one could extend this calculation to other kin, let us first consider the case of an *age*(only)-structured population, where fertility and survival rates at age i are F_i and S_i, with maximum age n and where mothers can have either 0 or 1 offspring at each time-step. The related survival probability is thus $P_i = \prod_{j=1}^{i-1} S_j$ and Euler-Lotka provides us with λ: $\sum_{i=1}^{n} F_i P_i \lambda^{-i} = 1$.

Let us now characterize all potential individuals in the population by a *word* system relating them to one common ancestor which is represented by the one-letter *word* "1". Each individual's word would consist of a series of letters drawn from the n letters of $[1, n]$. Any individual named with a *word* of length $k+1$, "$1l_1l_2..l_k$", would have a daughter at age j, with probability F_j (providing it is alive then, and that it ever existed). We call her, its j^{th} potential daughter and she would be named "$1l_1l_2..l_kj$". Thus "132" is a potential grand-daughter of the ancestor, and more precisely the offspring given birth to at age 2, by her mother herself produced when her mother (the ancestor) was 3.

Such a *word* system would help to relate genealogies and demographic processes. Let us contemplate individual "$1l_1l_2..l_k$". From its *word*, we can draw its generation number (since the ancestor): k. We call also get its age: it was born in the year $\sum_{i=1}^{k} l_i$ (after birth of the ancestor). Its *word* also provides the probability it was ever born: $\prod_{i=1}^{k} P_{l_i}F_{l_i}$, and the probability it is alive at time t (taken as the number of years after the birth year of the ancestor) : $(\prod_{i=1}^{k} P_{l_i}F_{l_i})P_{t-\sum_{i=1}^{k} l_i}$ (unless for cases where $t < \sum_{i=1}^{k} l_i$ or $t > n + \sum_{i=1}^{k} l_i$ where the probability is 0).

Let us now contemplate the genealogy and demography of *ego*, a newborn in the population issued from the ancestor, but alive at a time long enough from the ancestor's birth that the population is at stable-state. Let us denote *ego*'s word "$1l_1l_2..l_{k-1}l_k$" (with k therefore very large). Then the only individuals that are known to have ever existed are the ancestors of *ego* : its mother "$1l_1l_2..l_{k-1}$", its grand-mother "$1l_1l_2..l_{k-2}$", up until its most ancient ancestors "$1l_1$" and "1", the population's ancestor. Given the name of *ego* we know therefore its entire lineage from which the genealogy and the fate of all other potential individuals can be derived. Let us consider potential individual m "$1m_1m_2..m_{k^*}$". The Most Recent Common Ancestor (MRCA) of *ego* and m is thus "$1l_1..l_{k'}$"="$1m_1..m_{k'}$" such that k' is the largest number for which that property holds. Simply put, the *word* of *ego*'s and m's MRCA is the longest beginning of *ego*'s *word* that is also the beginning of m's *word*. This MRCA being in the lineage of *ego*, we know it existed, and was alive at the time it gave birth to "$1l_1..l_{k'}l_{k'+1}$". And thus the probability that m was ever born (ever existed) is, if $m_{k'+1} < l_{k'+1}$, provided by $e_m = {}^{P_{m_{k'+1}}}/_{P_{l_{k'+1}}} F_{m_{k'+1}} \prod_{i=k'+2}^{k^*} P_{m_i}F_{m_i}$. The age of m at *ego*'s birth is $a_m = \sum_{i=k'+1}^{k} l_i - \sum_{i=k'+1}^{k^*} m_i$. This, in turn, provides us with the probability of m's aliveness at *ego*'s birth : $e_m P_{a_m}$ (where P_i is extended to be worth 0 outside its natural domain of definition).

Therefore, for a random individual in the population, *ego* and for any other potential individual m, we know the likelihood of m's aliveness at *ego*'s birth - $e_m P_{a_m}$ - from m's word. Moreover we can also infer the kinship relatedness *ego* and m. It is the sum of their distances to the MRCA :

$$r(ego, m) = (k - k') + (k^* - k') = k + k^* - 2k'$$

Vectorizing the *words* - vectorization of individual "$1l_1..l_k$" is $1 + \sum_{i=1}^{k} l_i(n)^{i-1}$ - of any two potential individuals m_1 and m_2 in the population into "kinship states" i and j, and considering that m_1 is *ego* and m_2 is m, we can compute

$$K_{i,j} = r(ego, m).e_m P_{a_m}$$

, the relatedness weighted abundance. Together, these $K_{i,j}$ constitute the symmetrical and infinite *kinship matrix* $\mathbf{K}$.

The distribution, for an existing *ego*, of its possible lineages is provided by the stable-state distribution of maternal age at birth (we called $\mathbf{d}$ in the general case) which is, in an age-structured population: $d(i) = P_i F_i \lambda^{-i}$. Indeed, from this, the probability distribution for the lineage of newborn *ego* "$1l_1l_2..l_{k-1}l_k$" is $\prod_{i=1}^{k} d(l_i)$. From there one can compute the expected number of individuals in the population that are related with related coefficient r^* :

$$k(r^*) = \prod_{ego=1l_1l_2...l_{k-1}l_k} d(l_i) \times (\sum_{m|r(ego,m)=m^*} e_m P_{a_m})$$

, thus generating a $\boldsymbol{k}$ a (infinite) vector of *kinship distribution*, providing for an individual taken at random in the stable state population, the distribution of kin for each relatedness as inferred by the demographic parameters of the model.

4.5 Discussion

In this chapter we have hinted at the theory, the concepts and models that constitute the new field of *kinship demography* we hope to promote. With regards to the theory, a lot needs to be done in order to relate to the different approaches of the different fields connected, one way or another, to *kinship demography*. As we have shown, the inputs from cultural anthropology and kinship ethnology affect the evolutionary demography of a population and therefore its life history and the dynamics of its genes with further implications in many fields. Conversely, the demography of a population and its life history, as collected by human demographers and conservationists has effects on its kinship distribution which itself has effect on sociality. These two, coevolutive, sides of *kinship demography* however require new methods in order to be implemented and inferred from/to the models of evolutionary demography.

Stemming from the conceptual model for *physiological costs of reproduction* over chapter 1 and its application in chapter 3, we propose the initial steps of a new kind of *kinship models* that allow to implement the effect of transfers between various kin in the population. It extends the works of Tuljapurkar et al. (2007) - who proposed a husband and wife model - and Samuel Pavard - who has built daughter-mother and daughter-mother-grandmother models (Pavard et al., 2008; Pavard and Branger, 2012). Because the effect of kinship on the demography of a population is about transfers of resources from one individual - paying the cost of the transfer - and another - benefiting from it - we wish for our model to be able to implement both sides of transfers, i.e. both effects of kinship care. We also devise a way to correct for the asymetry of only incorporating costs or benefits, and ensure the implemented vital rates of kin correspond to that of the population they belong to. However, this approach requires further analysis, and in particular with regards to (Lee, 2003)'s work on transfers in age-structured population. As a first step, we would need to study the similarities and the differences between the kinship model, *folded* on kinship and the (Lee, 2003)'s transfer model.

From the initial work from Le Bras (1973); Goodman et al. (1974b,a) - that still constitutes the state of the art, we built tools that enable to extract the kinship distribution inferred by a population model. We first show the ability of MPPMs to implement further *neutral* traits makes it possible to compute for simple kinship distributions. We then show how genealogical Markov chains - introduced by (Demetrius, 1974, 1975) and later used by (Tuljapurkar, 1982c, 1993; Bienvenu and Legendre, 2015; Bienvenu et al., 2017) - can provide, for a slightly more complex kinship distribution - the expected number of older sisters *alive* - the calculation steps to extract it from a general structured population model. Finally we extend the concepts to include the distribution of all kin relatedness distribution in the population, in the case of age-structured models.

A further step will consist in emulating this approach towards computation of *kinship distribution* vector $\boldsymbol{k}$ and *kinship matrix* $\mathbf{K}$ for any matrix model. The structure of the ancestry lineage from *ego* would then be provided by genealogical matrices $\mathbf{P}$ and its two components : $d_i = \mathrm{P}_{1,i}$ across generation $\tilde{\mathbf{P}}$ within generation. Then, the branching out forward towards all potential descendants of the ancestors of *ego* could be performed by matrices $\mathbf{F}$ (equivalent to the F_i of the age-structured model) and $\mathbf{T}$ (equivalent to the S_i and P_i of the age-structured model).

Most importantly however, we think further progress in the kinship inference from demographic models will require the incorporation of multitype branching process, in particular the so-called Crump-Jaggers-Mode multitype BP (Crump and Mode, 1968, 1969; Jagers, 1982; Jagers and Nerman, 1984). The connection between kinship, demography and branching processes was made early by (Waugh, 1955), who,

together with Joffré went on to use Birth and Death and Galton-Watson processes to compute kinship distributions (Waugh, 1981; Joffe and Waugh, 1982). The multitype model they proposed in 1985, will be the base from which to improve our kinship extraction approach (Joffe and Waugh, 1985).

4.6 Appendix

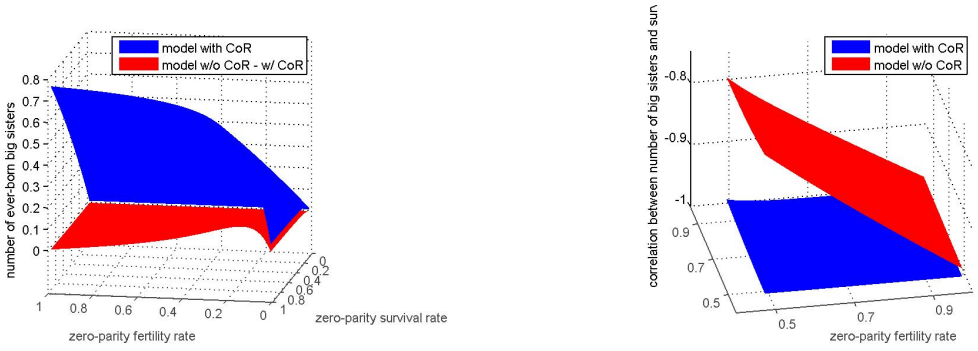
4.6.1 Illustration: effect of costs of reproduction on the expected number of older sisters

In this section, we will illustrate the calculations of section 4.4 by displaying the result of the computing of the expected number of older sisters for the models of chapter 3. By using the trait *parity* on the matrix implementing the costs and by adding it back, but as a *neutral* trait, on the asymptotically-equivalent matrix *folded* on parity from which the costs are absent, we can compute the distribution of older sisters ever-born (by state of mother) for both models and for a range of zero-parity vital rates.

In all cases, we find the expected number of big sisters to be higher for the model without the costs (see 4.3a). This may seem surprising since the number of sisters is obviously linked to lifetime reproductive output, the expectation of which is equal between the two models as we demonstrated (sec.3.0.6, p.97). The difference can be explained however, by the fact that, in the model with the costs, fertility is reduced by parity. And thus, whilst a 1st offspring is "easier" to have compared to the model without the costs, a 2nd child (that would be the first one to have an older sister) - let alone a 3rd child etc. - will be "harder" to obtain.

However most importantly for our study, we want the distribution of older sisters with respect to the aliveness of the mother. We characterize this relation by plotting the coefficient of correlation between the two distributions for both models (figure 4.3b). Unsurprisingly the correlation between survival of mother and number of older sisters is extremely low for the model with fertility parity trade-off. This is because, in our model, maternal survival declines with the number of offspring : $s(a,p) = s.(1 - p/n)$. Because the survival at the last age-class is always zero whatever the parity, this correlation is not exactly -1 and can rise for large survival rates and small fertility rates (as then almost all children are firstborns). In all cases however, it is much lower than the correlation for the model without the costs, even though the latter still display negative correlations. This is because in any model, parity is positively correlated with age (as it never decreases) whilst the correlation of survival with age is established (see sec.3.4.1, p.76).

This implies that in the models with the costs, the expected number of older sisters at birth helps compensate for the potential loss of the mother. If the care provided by a sister is a portion c of the care provided by a mother, then the expected care received by a newborn which mother is at stage (a,p), when giving birth to it, is $c.p + s.(1 - p/n)$. This total transfer from kin a newborn receives is thus independent from the state (age or parity) of the mother whenever $c = \frac{s}{n}$. In the model without the costs, survival of the mother is a function of age, independent from parity, and there will be no such compensation effect.



(a) (b)

Figure 4.3: For (age-*parity*) models (differing only by their zero-parity vital rates) and their related *reference leslie matrices*, we plot the and the difference in expected number of big sisters at birth between the model (implementing *physiological costs of reproduction*) and its its reference leslie model with no trade-off implemented and that quantity for the models with the costs. We also plot, for each model, the correlation between the expected number of big sisters at birth and survival of the mother. The population has maximum age $\omega = 15$ and age-at-maturity $\alpha = 1$. Cost of reproduction is modeled by relatively decreasing each vital rate by $1/(1 + \omega - \alpha)$ per parity

Discussion

In the introduction of this thesis, I have set out the questions ecologists are currently asking with respect to the cross, relative and conjoint evolutionary consequences of genetic and physiological trade-offs. In chapter 1, I proposed a new life-history conceptualization of physiological costs of reproduction, and its integration in a framework of *genetic costs of reproduction* has allowed us to disentangle the specific mechanisms and detectability patterns of each of these two families of trade-offs. In chapter 2 we developed a new family of evolutionary models - the Multitrait Population Projection Matrix (MPPM) - and its related analysis toolbox – the *Trait Level Analysis* that allows to investigate the demographic and life history consequences of traits. It therefore enables to implement and analyze physiological trade-offs in a context of genetic variance. In chapter 3, we built such a model for costs of reproduction, and analyze it to better understand the joint evolutionary roles of *physiological and genetic costs of reproduction.*

Intermediary results and related further investigations

This study in three steps – (life-history) conceptual, (evolutionary model) theoretical, and (costs of reproduction) analytical – brought results along the way that do not pertain directly to the subject of the thesis. Moreover, we have indicated, in chapter 4, an extension of the general model template for trade-offs to incorporate kinship transfers. Many of these intermediary results pave the way to further investigations, either because additional research is required to circumscribe the question asked, or because our research hints at new angles that we are urged to consider. We shall now summarize these intermediary results and discuss related further investigations.

Chapter 1

In chapter 1, we go back to Williams (1966)'s initial definition of the cost, and infer, from it, that physiological costs of reproduction combine a physiological allocative process that affect resource capitals as chronological times passes "forward" and an evolutionary mechanism that controls "backward" the resource capital expenditures according to the species evolved life-history strategy. We argue that these two processes correspond to two types of resources making up, respectively, the Fluctuating Capital related to acquirable resources (like food) in the environment and the Ratchet Capital associated with resources that cannot be acquired (like time) and that is related to the organism's evolved lifetime reproductive schedule. From this differentiation, we make predictions on the strength and therefore on the patterns of detectability of physiological costs depending on an organism's evolved position on three key life-history continua: the Slow-Fast continuum, the Income-Capital Breeding spectrum, and the Quantity-Quality continuum.

First, the position on the SFC controls the pace at which reproductive efforts have evolved to be made and therefore the strength of the long term RC costs. Long-lived organisms have large RC in early-life which makes their RC costs long-term and only observable at the back-end of the organism life expectancy. Second, the position on the ICB controls the capacity to build reserves and therefore the strength of the short-/mid-term FC costs. Income breeders cannot store FC resources, and their FC costs are therefore confined to survival costs in the current season. Capital breeders can, with therefore mid-term FC costs. Finally the position on the Quantity-Quality continuum drives the effect of demographic variance on reproductive effort, and therefore the strength of both costs.

The differing compositions of both capitals, connected by the time-step reproductive effort that depends on both capitals being supplied, prompts us to forecast templates for the detectability of physiological costs as a function of the environment, its variance and individual stochasticity. We show that these observability patterns differ from those of the genetic costs of reproduction, for which we also discuss the effects of genetic variance.

Going forward, these detectability patterns need to be tested, both empirically and theoretically. In particular, our approach that considers physiological costs of reproduction as functions of the position of an organism on multiple life history continua - we also discussed the effect of the organism on the

semelparous-iteroparous spectrum - will be required to be further examined in the light of Charnov (1990)'s dimensionless numbers (see also Charnov, 2002).

In the last part of the chapter, we discuss the adequate population models for each family of costs of reproduction – Agent-Based Models for physiological costs and matrix models for genetic costs – and suggest that the addition of traits in a matrix models could make it possible to study both kinds of trade-off in a single evolutionary model. In particular, adding genotypic polymorphism in a matrix model prompts us to revisit Charlesworth (2000)'s reconciliation of the allele-frequency equilibrium of population genetics and the stable-state of evolutionary demography, from a matrix model perspective. This hints at the role that MPPMs, and in particular *Trait Level Analysis* (see below), can play in simplifying and deepening – via the incorporation of other traits - the study of age-structured population genetics (Charlesworth, 1980).

Chapter 2

In chapter 2, the pivotal methodological chapter of this thesis – presented in its article version as published in Theoretical Population Biology in July 2017 - we first describe a method to build a Multitrait Population Projection Matrix (MPPM). The methodological efforts towards extending matrix models to populations characterized by age and stage and location etc. is topical as evidenced by the publication in Methods in Ecology and Evolution in September 2016 of an alternative building technique (Roth and Caswell, 2016). Because vector-based (avoiding loops and making use of sparse matrices), our method allows to build the model in a computationally inexpensive manner. Both time and space complexities are reined in compared to the alternative method, allowing for more traits to be implemented. We extend the sensitivity analysis of a one-trait matrix to an MPPM, but first discuss the latter's primitivity. Contrary to most one-trait models, MPPMs will not be primitive, but we show how to extend Perron-Frobenius results on ergodicity by considering the Frobenius normal form of an MPPM.

The key result of the article lies in the *Trait Level Analysis* which aims at being an asymptotically neutral framework for MPPMs. We define the merging of nodes with weights equal to their ergodic abundances, as Ergodic-Flow-Preserving merging of a graph. We extend EFP-merging to entire traits via an operation we call EFP-folding or simply folding. By construction, a matrix and its folded versions – among which the Reference Leslie Matrix, where only age remains as a trait – share the same asymptotic growth rate and relative abundances. This versatile tool has many properties and allow, among other things, to compare the modeling effects of different levels of scrutiny on a population (e.g., for conservationists), and to compare populations with various underlying processes (as for instance physiological and genetic trade-offs) in an ergodic-neutral framework (e.g., for evolutionary demographers).

The advent of multitrait matrices and their trait-based analysis tools open up multiple theoretical and empirical fields of research. Among these, we wish to isolate two key angles of investigation, which if fruitful will further enhance the interest of MPPMs and promote the development of a multitrait-structured population theory.

First, the application of *Trait Level Analysis* to transient dynamics, which may help answer questions such as the relative effect of a specific trait on the damping ratio, and the stabilizing/destabilizing effects of some traits on perturbation analysis. It could possibly be approached by comparing the spectra of folded matrices, or by comparing two folded versions of the same full model, either via EFP-folding, or via the transient-dynamics merging tool of economical input-output theory (Fisher, 1969, 1958; Simon and Ando, 1961).

Second, the incorporation into the theory of alternative folding methods that preserve other ergodic properties of models. We show, in the article, that the EFP-folding is only one specific choice of merging and that, even when focusing on the ergodic state, alternative forms of merging are conceivable. In a recent article published in Theoretical Population Biology, Bienvenu et al. (2017) use the genealogical Markov chains stemming from the work of Demetrius (1975, 1974) and Tuljapurkar (1982a) to describe an alternative state-merging process they call "genealogical collapsing". Life EFP-merging, it preserves the asymptotic growth rate and abundances. Contrary to EFP-merging however, it also preserves reproductive value, but at the cost of ergodic flows. It remains now to be investigated whether such a tool – once extended from states to traits – could be embedded in *Trait Level Analysis* and what it would mean to the field of evolutionary biology.

Chapter 3

The exploration of the cross effects of physiological and genetic costs, in chapter 3, prompted some methodological results.

First, it drove us to reconsider the generally accepted interpretation and computation of the net reproductive rate for (matrix) models with several classes of offspring. We also produced a closed-form formula of the variance in reproductive success for age-structured populations and combined these two results to yield a formula for variance in reproductive success in a population characterized by age and fixed-heterogeneity, in a manner that segregates the portion due to stochasticity and the portion due to hidden heterogeneity.

We also showed how to compute the variance in reproductive output for a population characterized with a dynamic heterogeneity trait implementing a trade-off, via the use of Markov chains with rewards. And finally we extended to multitrait models the age-based formula for vital rates selection gradients, demographic variance and environmental variance.

This research also encouraged us to initiate a study of the conceptual and formal similarities and differences between the variance in reproductive success related to the net reproductive rate and the demographic variance related to the stochastic growth rate. We provide an equation between these for stationary populations. This study will need to be further investigated in order to yield the long-needed reconciliation of two demographic approaches: the one focusing on the stochastic growth rate and its infinitesimal variations and the other considering the net reproductive rate $\mathbf{R_0}$ and its variance.

Chapter 4

In chapter 4, we hint at an extension of the theoretical model developed in the second part of chapter 1 towards kinship models, which can implement and infer the distribution of kin in a population and the transfers of resources between kin. We discuss important features required of such models and in particular show how to resolve the "transfer asymmetry issue" stemming from the implementation of only on side of kin transfers (only vital rates of beneficiaries or of payers are impacted by kinship distribution).

Across this new field of *kinship demography* - that aims at studying the coevolution of demographic and kinship parameters in a population – we also demonstrate how to infer, from an evolutionary model, simple kinship distributions and indicate the first steps to extend these computations to the general kinship distribution of a population, we call the *kinship matrix*.

The new field constitutes a vast area of future research in itself. In particular the kinship model framework presented here will need to be tested and compared with the evolutionary models focusing on intergenerational transfers, but forsaking kin, stemming from (Lee, 2003). Progress is also needed on the other facet of kinship demography where we wish to push forward the quest for tools generating the *kinship matrix*. The methods presented in chapter 4 are only first steps and an immense amount of work still remains to be done. Help towards that goal will stem from the initial works of demographers (Goodman et al., 1974b; Pullum, 1982), the early use of Galton-Watson branching process for inferring kinship (Joffe and Waugh, 1985, 1982; Waugh, 1981) and the mathematical field of multitrait branching processes (Crump and Mode, 1969, 1968; Jagers, 1982; Jagers and Nerman, 1984; Mode et al., 1987).

Main results regarding genetic and physiological trade-offs

The theoretical and mathematical formulations of the costs of reproduction - the "most prominent life-history trade-offs" (Stearns, 1989b) – of chapters I and III, bring initial answers to the fundamental questions asked by Braendle et al. (2011) and Edward and Chapman (2011) (Introduction pages 4 and 5) on the nature of physiological and genetic trade-offs and on their relative, cross and joint effects on the ecology, demography and evolution of populations.

Physiological and genetic trade-offs are different mechanisms

In the first chapter of this manuscript, we describe the mechanisms of genetic and physiological trade-offs by stepping back to the fundamental considerations of the founding fathers of theoretical life history trade-off theory (Partridge et al., 1991; Roff, 1992; Stearns, 1989b; Stearns and Koella, 1986). Both trade-offs are evolutionary mechanisms. As a matter of fact, genetic costs of reproduction stem from a purely genetic apparatus, the variance in a portion of a genome (one - at least - pleiotropic gene or several linked genes) that is antagonistic for two life history features. To the contrary, physiological costs of reproduction combine, as we saw in chapter 1, two components: a physiological one (we model via the Fluctuating Capital) and an evolutionary one (represented by the evolved Ratchet Capital). However, as we can see from the differences in mechanisms, understanding the physiological costs of reproduction as an evolutionary mechanism does not make them genetic costs. This can be seen, first, by the fact that they can co-habit in a framework we call *physiological costs of reproduction with genetic basis*.

The profound differences between genetic and physiological costs stem from the distinct position of their mechanism in the Stearnsian trade-off architectural triptych – genotypic level, intermediate structure and phenotypic level (Stearns, 1989b). The related divergence in their level of action – individual for physiological costs, population for genetic costs – has also important consequences, often disregarded. First, whilst, on the surface, both costs seem to have the same life history consequences (negative correlations between early fertility and late fitness) and environmental variance buffering effects - a similarity which historically bestowed them the same name of costs of reproduction - we show that they actually have different detectability patterns. Second, their different time levels of action hints at different windows of evolutionary effects. Physiological costs act at the time-step level and buffer environmental and individual stochasticity on a time scale of the order of the species' life expectancy. Genetic costs act at the level of the whole individual life trajectory and buffer environmental variance on an evolutionary time scale.

These results regarding the physiological and genetic costs of reproduction can be readily projected on their related senescence theories, respectively (Kirkwood and Holliday, 1979)'s Disposable Soma Theory and Williams (1957)'s Antagonistic Pleiotropy Theory. In particular, they bring to light the fact that the DST is most likely *not* a special case of the APT.

The joint role of physiological and genetic costs of reproduction

Thanks to the Multitrait Population Projection Matrix tools of chapter 2, and in particular the *Trait Level Analysis*, we were able to go further in the analysis of physiological and genetic costs. In particular in our understanding of their joint demographic and evolutionary effects as well as of the evolutionary consequences of physiological costs of reproduction. Indeed, this enabled us to build an (*age-parity-heterogeneity*)-structured model in which the physiological costs were implemented by having the vital rates dependent on parity, and the genetic costs by embedding several classes of genotypes differing with respect to their life history strategies.

We first showed the mechanical effects that the joint trade-offs have on an age-structured population by considering a model where vital rates only depend on parity and the heterogeneity but do not depend on age. Folding this model over parity and heterogeneity yields the Reference Leslie Matrix, its asymptotically equivalent model now only characterized by age. The analysis of this Reference Leslie Matrix shows us that, when the population is only characterized by age, the population's vital rates – i.e. the vital rates by age for an average individual in the population – now vary with age, yielding the familiar shapes of survival and fertility curves often found in the literature (Jones et al., 2014). Both costs of reproduction affect therefore age-structured demographics.

Second, we show that physiological costs have a strong impact on the selection gradient by age. As they spread out the reproductive effort of individuals over their lifetime, the force of selection on fertility at a specific age decreases strongly when passing from the model implementing the costs to its folded asymptotically equivalent model without costs. This has profound implications on the validity of estimations of the selection gradient as extracted from a population structured by age only, especially if one considers physiological costs to be ubiquitous in nature.

Moreover, this result provides a theoretical illustration and an extension to the corollary of Williams (1957)'s second prediction. This prediction indicates that, for a given population, the senescence-to-come (as anticipated by APT, given the shape of its age-structured selection gradient) does not depend on the current observed (actuarial) senescence in the population (but only on the absolute level of vital rates). Indeed physiological costs generate - by ensuring that high early fertilities lead up to shorten longevity and weaker late-life reproduction - actuarial senescence in the population. They also flatten the slope of the age-structured selection gradient, and in doing so, prevent further senescence propagating in the population. By contrast, the absence of physiological costs will limit actuarial senescence in the population. However, the steep gradient, promoted by this absence, ensures that senescence will be increase in the near evolutionary future.

The analysis of physiological costs of reproduction in the light of selection gradient also provides an interesting extension of the famous Houle (1991); van Noordwijk and de Jong (1986)'s results. van Noordwijk and de Jong (1986) showed that physiological costs of reproduction can be made undetectable by a large variance in acquisition capabilities among individuals in the population. Houle (1991) demonstrated that genetic costs of reproduction could be blurred by a larger genetic variance in robustness (both these results are re-established and discussed in chapter 1). We show, in chapter 3, that the genetic costs *themselves* conceal the underlying physiological costs of reproductions.

The last key results from chapter 3 regarding the costs of reproduction, consist in the (formal and via simulation) demonstration that physiological costs of reproduction indeed buffer individual stochasticity

(accounted via the variance in lifetime reproductive output or the demographic variance) and environmental variance. This therefore confirms the theoretical inferences of chapter 1. Moreover, the reducing effect of physiological costs on the variance in reproductive success has further consequences. Via its increasing impact on the effective population size, it affects *effective* selection. We measure *effective* selection with the ratio of fertility sensitivities by age by the demographic variance. We denote this ratio *variance-effective selection gradient*.

The antagonistic effects of physiological costs on both components of *effective* selection hints at their stabilizing role on APT with respect to life history strategy. Indeed when considering population structured by age only - the common framework for the study of senescence - fast organisms on the slow-fast continuum have steeper selection gradients that slow ones. This is because longevity itself buffers the effects of individual age-specific fertility events and spread out the reproductive effort of individuals over their lifetime. This result, although widely accepted as a good approximation of reality, seem however to point towards an evolutionary dichotomy, segregating organisms along their life history strategies, that has no empirical support. Indeed, on the slow side of a saddle point on the SFC - corresponding to a gradient steepness that is not strong enough to counter genetic drift – "faster" alleles have no significant selective advantages and the organism remains slow. On the fast side of this point however, "faster" alleles are allowed to invade the population, steepening the selection gradient, and therefore inviting even "faster" alleles and so on and so forth until the organism has evolved into a very short-lived living thing.

Instead of conjuring (very strong) environmental forces to explain why this effect is unobserved in nature, one needs to turn to the physiological costs of reproduction. When accounted for, fast organisms have flatter gradients and slow organisms steeper gradients than when omitted. Physiological costs therefore stabilize the strength of APT along the slow-fast continuum.

The costs of age-structured populations models

These key results from chapters I and III, considered with a step back, prompt us to ask a more general question about the pertinence of age-structured models.

Age-structured population studies are ubiquitous in evolutionary demography. Age-structured models are deemed, from a life historian's perspective, to be a major progress from models with non-overlapping generations, as they allow to incorporate the intra-generational transitions related to a life cycle. For that reason, as we saw in the introduction, age-structured models are even deemed to incorporate genetic life history trade-offs (in the sensitivities of their transitions).

However, as we saw in chapter 1, age-structured models do not incorporate genetic trade-offs, but only the position of a specific genotype on a life-history trade-off. To the contrary, at this genotypic level, they should be able to incorporate physiological trade-offs. But a physiological trade-off, a bivariate function of traits, requires a model with at least two traits. From our point of view, therefore, an age-structured model is merely the Reference Leslie Matrix of "nature's multitrait-model with trade-offs". In other words, age-structured populations are massively studied in life history, whereas they are only projections of life cycles on a single dimension which therefore cannot account for the keystones of life history: physiological trade-offs.

Obviously, we are not claiming that when basing their analysis on age-structured populations, Williams (1957),Hamilton (1966) or Charlesworth (1980) deemed it an exact representation of natural behavior and were not aware of some implications of the simplification. However, the focus on age structures, may have induced a shift from a paradigm of age-structured population to a paradigm of populations structured by age *only*. Therefore, many results put forward by these inspiring authors on subjects such as selection gradients, effective selection, fitness measures and their variance in the population do not hold, as we saw throughout this thesis, when accounting for the most basic of all trade-offs: the costs of reproduction.

The use of age as the *trait* of study – for many but not all organisms – is a product of statistical analysis (PCA for instance) that suggests it is the *single* trait that best determines the vital rates in the population. From a physiological trade-off perspective this is not a surprise as their effects are compounded over time, and therefore over age. As we saw in chapter 3, vital rates independent from age became functions of age when the model implementing the costs of reproduction was *folded* over them. However trade-offs relate life history *features*, not *traits*, and considering age as the best single determining *trait* does not diminish its weak signification as a *sole* parameter for life history. In other words, age may be the most explanatory of *traits* and the best sole determinant of the demography of a population, but – unable to implement trade-offs - it is a poor determinant of its life-history evolutionary demography.

We claim that the addition of a second trait in a population model - allowing to implement the trade-offs at the core of life history, in an evolutionary framework – finally bridges the gap between the two

branches of life history – trade-off theory and evolutionary demography – discussed in the introduction. We also allege that the addition of a trade-off *trait* has more effect on the accuracy of evolutionary interpretations with respect to a model structured by age only, than an age-structured model has with respect to a model with non-overlapping generations. The multitrait model (the MPPM), described and analyzed (the *Trait Level Analysis*) in this thesis, will go a long way, we hope, towards rebuilding the field of structured populations studies upon stronger life history foundations. Armed with MPPMs, other important evolutionary results stemming from age-structured models may be challenged in the light of life history. It would be worthwhile, for instance, to extend the study of age-structured population genetics towards a multitrait-structured population genetics theory and test whether, in that framework too, the addition of a structure to a non-overlapping model is of little consequence for population genetics (Charlesworth, 1980).

On the 5th of October 2015, at the Evolutionary Demography Society Annual Meeting, Hal Caswell, discussing the implementation of heterogeneity into evolutionary models pronounced this phrase "Accounting for heterogeneity [...] is the oldest problem in demography. When the heterogeneity is not accounted for, it can distort the conclusions based on demographic models. Heterogeneity is tamed by accounting for it !", that we can paraphrase as such:

Accounting for the physiological trade-offs of life history theory is an important but only recently acknowledged problem in evolutionary demography. When trade-offs are not accounted for, it can distort the conclusions based on demographic models. The evolutionary consequence of physiological trade-offs is tamed by accounting for them via MPPMs.

Chapter 5

Annexes

Contents

Parametrization of Model of Illustration of Chapter II

Yearly survival P_x is fitted using a Siler model (Siler, 1979) parameterized by Gurven and Kaplan (2007) for an average hunter-gatherer population (see fig. A below). Distribution of females' heterogeneity in quality at birth p(h) is fitted with a discretized lognormal distribution (mean=0.5, var=0.3) whose class 1 is set to 10%.(see fig. B below). Effect of heterogeneity e(h) is set linear, with fertility of class 1 being 0 (i.e. for sterile women) to 1 for class 10. Baseline fertility density with age $\widehat{F_x}$ is first fitted with a 4-parameters Brass polynomial (Brass, 1960) from Ache fertility table (Hill et al., 1996) multiplied by the mean effect of heterogeneity $\sum_h p(h)e(h)$ and rescaled such that $\sum_x \widehat{F_x} = \widehat{TFR}$ (Baseline fertility $\widehat{F_x}$ is then defined as the fertility by age of women whose fertility is not compromised by parity and $\widehat{TFR}$ is the Total Fertility Rate (i.e. the mean number of daughters born to women surviving the end of the reproductive period) of these women (see fig. C below).

Now assuming that parity p compromised linearly fertility by a factor $\sigma(p) \in [0,1]$ (from 1 for parity p=0 to 0 for p>10; 10 daughters meaning an averaged parity of 20 children). Mean population fertility F_x can be calculated at each age x such that $F_x = \widehat{F_x}\sigma(\sum_0^{x-1} F_x)$ where $\sum_0^{x-1} F_x$ is the mean parity of women at age x.

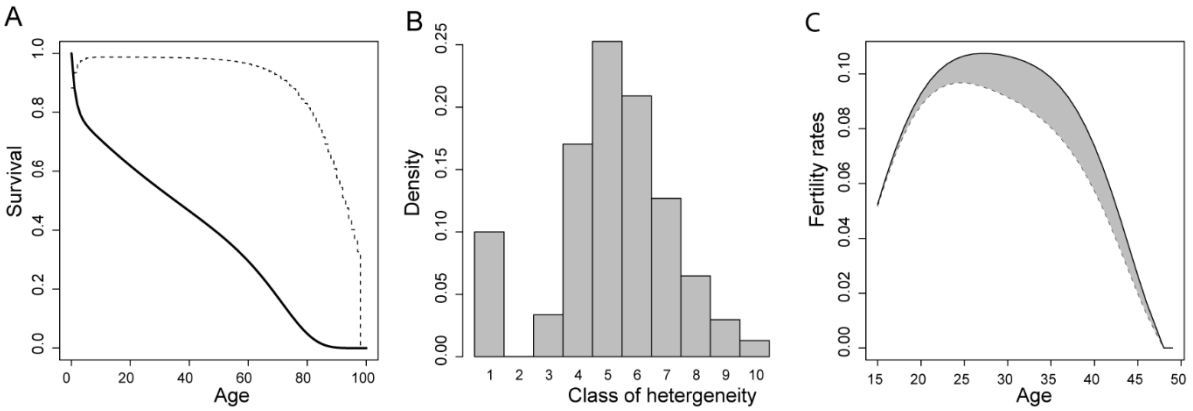


Figure - (**A**) Survival S(x) from birth to age x (plain line) and yearly survival P(x) = S(x+1)/ S(x) (dotted line), fitted with a Siler model (Siler, 1979) such that $S(x) = \exp(\frac{a_1}{b_1}(e^{-b_1 x} - 1)).\, e^{-a_2 x}.\exp(\frac{a_3}{b_3}(1 - e^{b_3 x}))$ with [a1=0.422,b1= 1.131, a2=0.013, a3=1.47E-04 and b3=0.086]. (**B**) Probability distribution of being in heterogeneity class h, with class 1 corresponding to sterile women with a coefficient $\sigma(1) = 0$ and class 10 corresponding to most robust women with coefficient $\sigma(10) = 1$. (**C**) Fertility rates from first to last age at reproduction (respectively α=15, β=49) in the case where parity has no compromising effect on fertility rates ($\widehat{F_x}$, plain line). Fertility by age is fitted by a Brass polynomial (Brass, 1960) $y = (14 - x).(49 - x)^3.(x^2 - 49x + 700)$ scaled so $\widehat{TFR} = 4$ and in the case where mean parity p at age x compromise fertility by a factor $\sigma(p) \in [0,1]$ declining linearly from $\sigma(0) = 1$ to $\sigma(p > 10) = 0$ (F(x), dotted line).

Bibliography

Brass, W., 1960. The Graduation of Fertility Distributions by Polynomial Functions. Popul. Stud. (NY). 14, 148–162. doi:10.2307/2172011

Gurven, M., Kaplan, H., 2007. Longevity Among Hunter-Gatherers Examination. Popul. Dev. Rev. 33, 321–365.

Hill, K., Hurtado, A.M., Low, B.S., 1996. Aché Life History: The Ecology and Demography of a Foraging People, Evolution and Human Behavior. doi:10.1016/S1090-5138(97)00035-4

Siler, W., 1979. A Competing-Risk Model for Animal Mortality. Ecology 60, 750–757. doi:10.2307/1936612

Nullification of approximated zeroes in MPPM eigenvector.

Finding all eigenpairs $(\lambda, \boldsymbol{w})$ of a square matrix $\mathbf{M}$ means solving $(\mathbf{M} - \lambda\mathbf{I})\boldsymbol{w} = 0$. This can be done by first finding the eigenvalues λ as the solutions of the characteristic equation $det(\mathbf{M} - \lambda\mathbf{I}) = 0$, and then finding the associated eigenvectors. The left member of the characteristic equation is a polynomial in λ of degree the size of $\mathbf{M}$.

Abel's impossibility theorem (Abel, 1824) states that there is no algebraic solution (i.e. solution in radicals) to the general polynomial equations of degree five or higher with arbitrary coefficients. This implies that algorithms that calculate exact eigenvalues in a finite number of steps only exist for general square matrices of size four or below. For larger matrices, and MPPMs will be amongst them, these solutions can be computed to any desired degree of accuracy using numerical methods ranging from naïve iteration algorithms to fast sparse convergence algorithms. In all cases however, the resulting eigenvectors will be proxies of the solutions. Consequently each element of $\boldsymbol{w}$ and $\boldsymbol{v}$ will itself be approximated and this is generally acceptable except for zeroes. Indeed the proxy of a zero (even if computed with a great level of accuracy) will lack many properties of a real zero.

The importance of zeroes in abundances and reproductive values in population-dynamics-related eigenanalysis means those near-zeroes have to be replaced by actual zeroes. This can be done by analysing the distribution of the logarithm of absolute eigenvector entries (indeed some near-zeros may be negative); this allows discriminating between non-zeros and real zeroes (see figure below). Indeed, the distribution of small value entries decreases with the log of their value. The sudden peak in the distribution after a certain threshold (corresponding to the accuracy of the algorithm used) is due to the approximations of the zeroes of the eigenvector. These near-zeroes have to be nullified (see code in supplementary material 7).

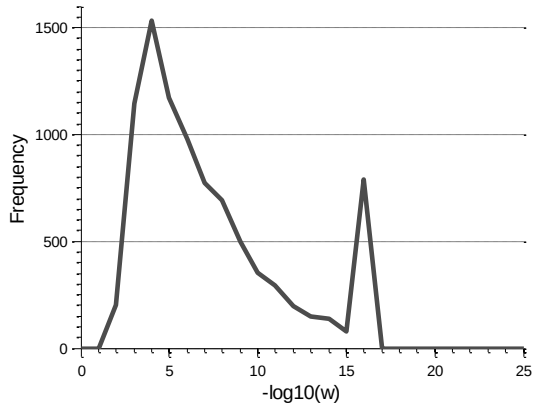


Figure: distribution of log values of maximal-eigenvalue-related right-eigenvector of the parity-fecundity MPPM constructed in the illustration, obtained by sparse matrix fast convergence algorithm. The figure shows that the majority of entries of $\boldsymbol{w}$ below 10^{-15} are actual zeroes. As such they have to be nullified.

Bibliography:

Abel, N.H., 1824. Mémoire sur les équations algébriques, où on démontre l'impossiblité de la résolution de l'equation générale du cinquième dégré. Librarian, Faculty of Science, University of Oslo.

Code for Chapter II

Functions and script published with MATLAB® R2012a

In this section, we provide the code for the main computations described in the article, mainly in the vector-based construction of the MPPM (in I) and the Fold function (in II) that allows the specific trait-level analysis of the illustration (in III). We also provide the code for sensitivity analysis to parameters for MPPMs constructed via the vector based technique (IV), and the code for the vec() function (only necessary for older versions of Matlab) (V) and for the generation of the input parameters of the illustration (VI). Finally we provide the code for the vec-permutation construction of the (same) model described in the illustration (VII) that necessitate the adaptations of subfunctions from (Roth and Caswell, 2016) (VIII).

I. VECTOR-BASED CONSTRUCTION OF MPPM

```
%% We construct the 3-trait MPPM (age, parity, heterogeneity class) described in
the illustration from the following inputs (than can be obtained from section VI of
this Supplementary Material):
% 1/ baseline survival and fertility rates
surv %the baseline yoy survival
fert %the baseline fertility
% 2/ trade-off coefficients for the parity effect
parityeffect %containing the multiplying factors affecting fertility for each
parity class
% 3/ heredity coefficients (distribution of offspring in the different
heterogeneity classes)
classdistrib %containing the probability distribution of heterogeneity class at
birth
% 4/ frailty coefficients (effect of frailty on baseline fertility)
classeffect  %containing the multiplying factors affecting fertility, for each
heterogeneity class

%%
maxparity=length(parityeffect)-1; % maximum parity is inferred from the fertility-
parity trade-off coefficients

%% FIRST STEP OF MPPM CONSTRUCTION  : define set of traits and trait structure
% we order the 3 traits as such : (age,parity,heterogenity), this leads to
% the following trait structure s :
s=[length(surv)+1 (maxparity+1) length(classeffect)] % trait structure i.e.
(ordered) trait sizes vector
n=size(s,2) %number of traits
q=prod(s) %number of states

%% SECOND STEP OF MPPM CONSTRUCTION : generate the vector of survival rate for each
state and the vector of fertility rate for each state

% in our illustration, survival only depends on age, so we replicate
% the baseline survival vector surv for all parity and heterogeneity classes :
Vrsurvival=repmat([surv' 0]', [1 (maxparity+1) length(classeffect)]) ; %repmat
matlab function replicates a vector (first argument) along the dimension(s) in the
second argument

% in our illustration, fertility rate for state (a,b,c) is
fert(a)*parityeffect(b)*classeffect(c)
% this is translated vector-wise as :
Vrfertility=reshape(vec(fert*parityeffect')*classeffect',s);
% explanation : fert*parityeffect'  generates the matrix providing fertility rate
at a given age (row) and parity (column) for the robust class.
% vec() then turns this matrix into a vector by stacking the columns on top of one
another
% we then repeat the process by multiplying this vector by  classeffect', which
provides fertility rate at a given (age,parity) state (row) and heterogeneity class
(column)
```

```
% We then "reshape" this 1-dimensional vector into a 3-dimensional one (of
dimensions s) in order to be able to read fertility for all ages (dimension 1),
parity classes (dimension 2) and heterogeneity class (dimension 3)

%% THIRD STEP OF MPPM CONSTRUCTION  : generate the transition matrix T (of width 3)
containing all transitions (one per row)
% where, for each transition, the element in the first column is the input
% state, the element in the 2nd column the output state, and the element in
% the 3rd column the transition rate = vital rate * output probability
% distribution

order=reshape(1:q,s); % state indices redistributed over the 3 dimensions of s
% this multidimensional vector will be very useful to relate state number and
triplet of states in an efficient manner.

% we will first generate transitions for vital process = survival (and
% store them in T1 and T2)

% survival transitions : for each input state, there are 2 potential output
% states, as age is deterministic, heterogeneity is invariant and parity
% can either remain the same or increase by one : (age+1,parity,classfert) and
(age+1,parity+1,classfert)
% We store the transitions towards (parity) in T1 and (parity+1) in T2.

% transitions towards (parity)
% transition rate : the output probability is (1-fertility rate), hence the
transition rate (= vital rate * output probability) for each input state is
tS1=Vrsurvival.*(1-Vrfertility);
% input states : we have to limit input states to those which are not at maximum
age, i.e. those which state numbers are order(1:end-1,:,:)
%output states: to each such input state, the related output state is
%located in the following "row" of the 3-dimensional vector "order".
% Then, the matrix of transitions - (input,output,transition) - for all transitions
towards (parity) is
T1=[vec(order(1:end-1,:,:)) vec(order(2:end,:,:)) vec(tS1(1:end-1,:,:))] ;

% transitions towards (parity+1)
% transition rate : the output probability is fertility rate, hence the transition
rate is :
tS2=Vrsurvival.*Vrfertility;
% input states : we have to limit input states to those which are not at maximum
age and parity , i.e. those which state numbers are order(1:end-1,1:end-1,:)
%output states: to each such input state, the related output state is located in
the following "row" and "column" of the 3-dimensional vector "order".
% Then, the matrix of transitions is
T2=[vec(order(1:end-1,1:end-1,:)) vec(order(2:end,2:end,:)) vec(tS2(1:end-1,1:end-
1,:))] ;

% now we turn to vital process = fertility (and store those transitions in T3)
% fertility transitions : for each input state, there are 10 potential
% outputs corresponding to the 10 heterogeneity classes (age will always be
% 1 and parity 0). We will treat those 10 outputs simultaneously.
% transition rate : the output probability is by construction classdistrib, hence
the transition rate (= vital rate * output probability) for
% each input state AND each outputstate, is classdistrib * Vrfertility
% we can generate T3, containing all fertility transitions, where these fertility
transitions are sorted, first by input state, then by heterogeneity class of the
% output sate :
T3=[ repmat(vec(order),length(classdistrib),1)
vec(repmat(vec(order(1,1,1:length(classdistrib)))',q,1))
vec(Vrfertility(:)*classdistrib) ]; %transition triplet for fertility towards 1, 1,
1:10
% 1st column of T3 repeats all input states 10 times (for each output state
heterogeneity class)
% 2nd column of T3 repeats all 10 outputstates [(1,1,1), (1,1,2) (1,1,3)
% ...(1,1,10)] for each inputstate
% 3rd column of T3 provides the relevant transition rate, dependent of the
```

```
% fertility rate of each input state and the distribution probability
% toward the output state

% we can now generate the matrix T containing all transitions :
T=vertcat(T1,T2,T3);

% we can also generate the matrices TV (resp.  TF) containing all survival (resp.
fertility) transitions
TF=vertcat(T3);
TV=vertcat(T1,T2);


%% from the matrix of transitions T to the MPPM M

% T is the sparse definition of M, i.e.
% instead of determining the value for each (i,j) entry, we have determined
% the transition triplet : ( j , i , M(i,j) ) for each relevant transition
% We can then generate the (square) projection matrix for our population,
% the MPPM M :
M=sparse(T(:,2),T(:,1),T(:,3),q,q); %

% we can also generate the projection matrix for fertility F and survival V
% :
F=sparse(TF(:,2),TF(:,1),TF(:,3),q,q); % fertility transitions
V=sparse(TV(:,2),TV(:,1),TV(:,3),q,q); % survival transitions


%% density of M
ntrans=size(T,1) %number of transitions implemented
density=ntrans/q^2 % density of M=sparse(T)
```

II. FOLD FUNCTION

```
%% generates the Ergodic Flow Preserving -folding of an MPPM M of trait structure
traitsize with ergodic abundance vector w on the traits to be found in positions
apos of the trait structure.

function [ Mfolded ] = fold(M,w,traitsize, apos) a=length(apos); % those a traits
will "disappear"
n=length(traitsize); % (n-a) traits will remain
sign=1:n; [Lia,Locb] = ismember(apos,sign);
sign(Locb)=[];sign=[apos sign] ; % reordering trait order so traits to disappear
appear in the first a positions, remaining traits in the (n-a) last positions
order=reshape(1:prod(traitsize),traitsize); %all state-indices of M, in M
multidimensional notation
I = eye(n,n);P = I(sign,:); traitsize2=(P*traitsize')';% traitsize vector for new
ordering of traits
order=permute(order,sign); signstate=order(:);I =
eye(prod(traitsize),prod(traitsize));PermutStates = sparse(I(signstate,:));       %
permutation matrix reordering traits in M to match traitsize2
W=reshape(w,traitsize);% multidimensional version of w
W2=permute(W,sign); % permutation of dimensions in w to match new ordering
 wnewstates=sum(reshape(W2,[prod(traitsize2(1:a)) traitsize2(a+1:n)]) ,1); % w for
remaining traits after folding over the other traits
wnewstates2=repmat(wnewstates(:),1,prod(traitsize2(1:a)))'; % replication of this
new w (a row vector) over folded traits
weightmat=sparse(repmat( (W2(:)./wnewstates2(:))' ,
prod(traitsize2),1));weightmat(isnan(weightmat))=0;weightmat=sparse(weightmat);
%matrix of all weights to be applied to M : for each state w(state)/w(all states
sharing same remaining traits)
Mweighted=sparse( (PermutStates*M*PermutStates').*weightmat); %M, reordered and
weighted by the appropriate weights
Pr=sparse(kron(eye(prod(traitsize2(a+1:n))),ones(1,prod(traitsize2(1:a))))); %trait
reduction "permutation" matrix
Mfolded= Pr *Mweighted* Pr';
```

```
end
```

III. TRAIT LEVEL ANALYSIS OF THE ILLUSTRATION

```
%% spectral analysis
find(sum(M,1)==0); %M is reducible
d = eigs(M,2); % 2 largest eigenvalues absolute value are not equal
[vcpd,vlp] = eigs(M,1); %convergence of calculation of lambda and associated right
eigenvector of M
[vcpg,vlpg] = eigs(M',1);%convergence of calculation of lambda and associated right
eigenvector of transpose of M (thus left eigenvector of M)
lam=vlp %lambda
w=vcpd/sum(vcpd); %w is scaled to sum to 1
v=vcpg/(vcpg'*w); %v is scaled so v'*w=1

% analyse and zeroing of close-to-zero values of v and w
figure;plot(histc(-log10(w),0:1:25));
i=1e-17; wcut=w ; wcut(find(abs(w)<i))=0;
w=wcut;
figure;plot(histc(-log10(v),0:1:25));
i=1e-13;vcut=v ; vcut(find(abs(v)<i))=0;
v=vcut;


%% trait-level analysis 1 : folding w and M
%multidimensional w and its folded versions
W=reshape(w,s); % 3-dimensionnal version of w
wage=sum(sum(W,2),3); % w  folded on all traits but age
wparity=squeeze(sum(sum(W,1),3))'; % w projected on parity
wfertclass=squeeze(sum(sum(W,1),2));% w projected on fecundity class
wageparity=squeeze(sum(W,3)); % w folded on fecundity class
wagefertclass=squeeze(sum(W,2)); % w folded on parity
%folded versions of M
Mage=fold(M,w,s, [2 3]); % reference Leslie matrix
Mparity=fold(M,w,s, [1 3]); % M folded on age and fecundity class
Mageparity=fold(M,w,s, [3]); % M folded on fecundity class
Magefertclass=fold(M,w,s, [2]); % M folded on parity

%% trait-level analysis 2 : number of births and abundances for age and parity and
full-traited-model
Mageparityreshape=reshape(Mageparity(1,:),s(1:end-1));
birthperyearforageandparityofmother=Mageparityreshape.*wageparity; % births by age
and parity
rspwageparity=reshape(wageparity,s(1),s(2)); % abundances by age and parity

birthperyearforageparityandfertofmother=fertstates.*wM; % number of births and
abundances for age and parity and hetero

%% trait-level analysis 3 : sensitivities to fertilities of various folded forms of
M

% 1/ M
statenum=prod(s); % number of states
% fertilities for all states are transitions towards triplets (1,1,class)
fertstates=M(sub2ind(s,ones(1,nclasses),ones(1,nclasses),(1:nclasses)),:);
fertstates=sum(fertstates,1);
fertstates=reshape(full(fertstates),s); % fertility for all states
 % all other transitions are survival
MM=M;MM(sub2ind(s,ones(1,nclasses),ones(1,nclasses),(1:nclasses)),:)=sparse(nclasse
s,statenum);survstates=sum(MM,1); % survival for all states
survstates=reshape(full(survstates),s); clear MM;
% eigenanalysis
[vcpd,vlp] = eigs(M,1); %convergence of calculation of lambda and associated right
eigenvector of M
```

```
[vcpg,vlpg] = eigs(M',1);%convergence of calculation of lambda and associated right
eigenvector of transpose of M (thus left eigenvector of M)
lamM=vlp
wM=vcpd/sum(vcpd); %w is scaled to sum to 1
vM=vcpg/(vcpg'*wM); %v is scaled so v'*w=1
wM=reshape(wM,s);
vM=reshape(vM,s);
% sensitivities of fertility
% we build K the matrix where each column i represent the sensitivity of M to
fert(i)
K=sparse(size(M,1),size(M,2));
% we use classdistrib for fertility transitions output
K(sub2ind(s,ones(1,nclasses),ones(1,nclasses),(1:nclasses)),:)=repmat(classdistrib'
,1,statenum);
diago=survstates;diago(s(1),:,:)=zeros(1,s(2),s(3));diago(:,s(2),:)=zeros(s(1),1,s(
3));
diago=vec(diago);
K1=-spdiags(diago,-1,size(M,1),size(M,2));
K2=spdiags(diago,-1-s(1),size(M,1),size(M,2));
K=K+K1(1:statenum,1:statenum)+K2(1:statenum,1:statenum);
% now we can compute sensitivity to fert
sensifertstatesM1=ones(1,statenum)*(S.*K);
sensifertstatesM1=reshape(sensifertstatesM1,s);
elastifertstatesM=(1/lam)*(sensifertstatesM1.*fertstates);

% 2/M folded on fecundity class
%eigenanalysis
[vcpd,vlp] = eigs(Mageparity,1); [vcpg,vlpg] = eigs(Mageparity',1);
lamMageparity=vlp
wMageparity=vcpd/sum(vcpd); %w is scaled to sum to 1
vMageparity=vcpg/(vcpg'*wMageparity); %v is scaled so v'*w=1
wMageparity=reshape(wMageparity,[s(1) s(2)]);
vMageparity=reshape(vMageparity,[s(1) s(2)]);
%states
nts=s(1:2); % new trait structure
statenum=prod(nts) % number of states
%fertilities and survivals
nfertstates=Mageparity(1,:); % all fertilities are on line 1
nfertstates=reshape(full(nfertstates),nts);
nfertstates=reshape(full(nfertstates),nts);
MM=Mageparity; MM(1,:)=sparse(1,statenum);
nsurvstates=sum(MM,1); % survival transitions are all other transitions
nsurvstates=reshape(full(nsurvstates),nts); clear MM;
% sensitivity and elasticity matrices
Sageparity=vMageparity(:)*wMageparity(:)'; %sensitivity of lambda to each entry
Eageparity= (1/lamMageparity)*Sageparity.*Mageparity;  % elasticity
% sensitivities of elasticities
% we build K …
K=sparse(size(Mageparity,1),size(Mageparity,2));
K(1,:)=ones(1,statenum);
diago=nsurvstates;diago(nts(1),:)=zeros(1,nts(2));diago(:,nts(2))=zeros(s(1),1);dia
go=vec(diago);K1=-diag(diago,-1);K2=diag(diago,-1-nts(1));
K=K+K1(1:statenum,1:statenum)+K2(1:statenum,1:statenum);
% now we can compute sensitivity to fert
sensifertstatesMageparity=ones(1,statenum)*(Sageparity.*K);
sensifertstatesMageparity=reshape(sensifertstatesMageparity,nts);
elastifertstatesMageparity=(1/lamMageparity)*(sensifertstatesMageparity.*nfertstate
s);

% 3/M folded on parity
%eigenanalysis
[vcpd,vlp] = eigs(Magefec,1);[vcpg,vlpg] = eigs(Magefec',1);
lamMagefec=vlp
wMagefec=vcpd/sum(vcpd); vMagefec=vcpg/(vcpg'*wMagefec);
wMagefec=reshape(wMagefec,[s(1) s(3)]);
vMagefec=reshape(vMagefec,[s(1) s(3)]);
%states
nts=s;nts(2)=[]; % new trait structure
```

```
statenum=prod(nts) % number of states
%fertilities and survivals
% fertilities are all transitions towards (1,1,class) states to sum
nfertstates=Magefec(sub2ind(nts,ones(1,nclasses),(1:nclasses)),:);
nfertstates=sum(nfertstates,1);nfertstates=reshape(full(nfertstates),nts);
nsurvstates=sum(Magefec,1)-nfertstates(:)';
nsurvstates=reshape(full(nsurvstates),nts);
% sensitivity and elasticity matrices
Sagefec=vMagefec(:)*wMagefec(:)'; %sensitivity of lambda to each entry of M
Eagefec= (1/lamMagefec)*Sagefec.*Magefec;  % elasticity
% sensitivities of elasticities
% we build K …
K=sparse(size(Magefec,1),size(Magefec,2));
K(sub2ind(nts,ones(1,nclasses),(1:nclasses)),:)=repmat(classdistrib',1,statenum);
% now we can compute sensitivity to fert
sensifertstatesMagefec=ones(1,statenum)*(Sagefec.*K);
sensifertstatesMagefec=reshape(sensifertstatesMagefec,nts);
elastifertstatesMagefec=(1/lamMagefec)*(sensifertstatesMagefec.*nfertstates);

% 4/Reference Leslie Matrix : M folded on parity and fecundity class
%eigenanalysis
[vcpd,vlp] = eigs(Mage,1); [vcpg,vlpg] = eigs(Mage',1);lamMage=vlp
wMage=vcpd/sum(vcpd); vMage=vcpg/(vcpg'*wMage);
%states
nts=s(1); % new trait structure
statenum=prod(nts)
%fertilities and survivals
nfertstates=Mage(1,:);
MM=Mage; MM(1,:)=sparse(1,statenum);nsurvstates=sum(MM,1);
% sensitivity and elasticity matrices
Sage=vMage*wMage'; Eage= (1/lamMage)*Sage.*Mage;
% sensitivities of elasticities
sensifertstatesMage=Sage(1,:);
elastifertstatesMage=(1/lamMage)*(sensifertstatesMage.*nfertstates);

%% multidimensional elasticities are summed

elastifertstatesM_sumonage=sum(sum(elastifertstatesM,2),3);
elastifertstatesMageparity_sumonage=sum(elastifertstatesMageparity,2);
elastifertstatesMagefec_sumonage=sum(elastifertstatesMagefec,2);

%% table with elasticities of lambda to fertilities of each folded matrix of M. If
multidimensional, elasticity is summed on age
elasfert=[elastifertstatesM_sumonage elastifertstatesMageparity_sumonage
elastifertstatesMagefec_sumonage elastifertstatesMage' Elastibaselinefert']

figure;
plot(1:99,elasfert(:,1),'-ro',1:99,elasfert(:,2),'-bx',1:99,elasfert(:,3),'-
m.',1:99,elasfert(:,4),'k^',1:99,elasfert(:,5),'y')
legend('M','Mageparity','Magefec', 'Mage','baseline')
%% end
```

IV. SENSITIVITY ANALYSIS TO PARAMETERS (HIGHER LEVEL SENSITIVITY ANALYSIS) FOR THE ILLUSTRATION

```
%% vectors and matrices functions of p
ssurv = sym(zeros(1,length(surv))); %
for k=1:length(surv) ssurv(k) = sym(sprintf('ssurv%d', k)); end
assume(ssurv,'real');

sfert = sym(zeros(1,length(fert)));
for k=1:length(fert) sfert(k) = sym(sprintf('sfert%d', k)); end
assume(sfert,'real');
```

```
sparityeffect = sym(zeros(1,length(parityeffect) ));
for k=1:length(parityeffect) sparityeffect(k)=sym(sprintf('sparityeffect%d',k));
end
assume(sparityeffect,'real');

sclasseffect = sym(zeros(1,length(classeffect)));
for k=1:length(classeffect) sclasseffect(k)=sym(sprintf('sclasseffect%d', k));end
assume(sclasseffect,'real');

sclassdistrib = sym(zeros(1,length(classdistrib)));
for k = 1:length(classdistrib) sclassdistrib(k) = sym(sprintf('sclassdistrib%d',
k)); end
assume(sclassdistrib,'real');

%formal vector of all parameters
sp = [ssurv sfert sparityeffect sclasseffect sclassdistrib ];

% vital rate vectors as function of p
sVrsurvival=repmat([ssurv 0]', [1 (maxparity+1) length(sclasseffect)]) ;%
sVrfertility=reshape(vec(sfert'*sparityeffect)*sclasseffect,s); %

%transition triplets matrices as function of p
stS1=sVrsurvival.*(1-sVrfertility);
sT1=[vec(order(1:end-1,:,:)) vec(order(2:end,:,:)) vec(stS1(1:end-1,:,:))];
stS2=sVrsurvival.*sVrfertility;
sT2=[vec(order(1:end-1,1:end-1,:)) vec(order(2:end,2:end,:)) vec(stS2(1:end-
1,1:end-1,:))];
sT3=[repmat(vec(order),length(sclassdistrib),1)
vec(repmat(vec(order(1,1,1:length(sclassdistrib)))',q,1))
vec(sVrfertility(:)*sclassdistrib) ];

% T as function of p
sT=vertcat(sT1,sT2,sT3);

%% 2nd level analysis : sensitivity to  p
deriv=zeros(ntrans,psize);% deriv is the gothic S of the method : the matrix of
sensitivity of M's entries to p
spar=zeros(1,psize); % is (will be) sensitivity of lambda to all parameters
T2=T(:,2);T1=T(:,1);sT3=sT(:,3);
for i=1:psize % loop on all parameters
    param=sp(i);%
    deriv=double(subs(diff(sT3,param),sp,p)); %
    spar(1,i)=full(sum(sum((S).*sparse(T(:,2),T(:,1),deriv,q,q))));
end
epar= (1/lam)*spar.*p; % elasticity of lambda to all parameters
Sensibaselinefert=spar(parameterstable(2,2):parameterstable(2,3));
Elastibaselinefert=(1/lam)*Sensibaselinefert.*fert'; % elasticity to all components
of fert
```

V. VEC FUNCTION

```
function x = vec(X)
x = X(:);
```

VI. INPUT PARAMETERS FOR ILLUSTRATION

```
% this part generates for our future MPPM where maximum age is 99 :
% 1 : the vector "surv", the baseline yoy survival from a Siler with parameters
% 2 : the vector "fert", the baseline fertility from a Brass polynomial with
parameters
% 3 : the vector "parityeffect",  containing the multiplying factors affecting
fertility for each parity class
```

```
% 4 : the vector "classdistrib",  containing the probability distribution of
heterogeneity class at birth
% 5 : the vector "classeffect",  containing the multiplying factors affecting
fertility, for each heterogeneity class

agemax=99;

%% 1/ baseline yoy survival vector : "surv"
% we draw this vector from a Siler hazard function

syms x a1 b1 a2 a3 b3 t % classic variables of Siler hazard function
h(x,a1,b1,a2,a3,b3) = a1.*exp(-b1.*x)+a2+a3.*exp(b3.*x); % mortality hazard
S(t,a1,b1,a2,a3,b3) = exp(-int(h,x,0,t)); % Probability of surviving to age t
a1=0.157;b1=0.721;a2=0.013;a3=4.80E-05; b3=0.103;% Siler hazard function parameters
values

surv=double(S(1:1:agemax,a1,b1,a2,a3,b3)./S(0:1:agemax-1,a1,b1,a2,a3,b3))';%yoy
survival, birth pulse - prebreeding
surv(end)=0; % yoy survival has to be zero for last age class
firstyearsurvival=surv(1);surv=surv(2:end); % 1st element represents 1st year
survival and is actually part of fertility vector
surv; % our surv vector, of size 98, integrating yoy survival  from (age 1 to age 2
survival) to (age 98 to age 99 survival)
clear x a1 a2 a3 b1 b3 t h S

%% 2/ baseline fertility vector : "fert"
% we draw this vector from a  brass polynomial
brassfert=@(age, p1,p2,p3,p4,alphaplusone, beta )  (age>=(alphaplusone -
1)).*(age<beta-1).*((age-(alphaplusone -1)).*((beta-
age).^2).*(p1+(p2*age)+(p3*(age.^2))+(p4*(age.^3))));
brassfert2=@(x,xdata) brassfert(xdata,x(1),x(2),x(3),x(4),10,49);
% we fit from data
a=[0.008 0.151 0.275 0.298 0.318 0.279 0.219 0.069 0];ydata=a/sum(a);
xdata=10:5:50;
x =
lsqcurvefit(brassfert2,[0,0,0,0],xdata,ydata);fert0=brassfert2(x,0:1:100)';fert0=fe
rt0/sum(fert0);fert0=fert0*6; % fit and rescaled for TFR=6
fert0=fert0(2:100);
fert=fert0*firstyearsurvival; % survival to year 1 factored in

clear brassfert brassfert2 a fert0 x xdata ydata firstyearsurvival
%% 3/ effect of parity on fertility
parityeffect=zeros(36,1); parityeffect(1:11)=1:-.1:0; % multiplying effect of
parity on fertility
maxparity=max(find(parityeffect)+1); % we deduce maximum parity from multiplying
effect of parity on fertility
parityeffect= parityeffect(1:maxparity+1); % adjustment of size parityeffect with
maxparity

%% 4/ fertility class heterogeneity
classdistrib =
.9*(lognpdf(0.05:.1:.95,log(0.5),0.3))/sum(lognpdf(0.05:.1:.95,log(0.5),0.3));
classdistrib(1)=0.1; %
classdistrib; % distribution of heterogeneity class at birth

%% 5/ effect of heterogeneity on fertility
classeffect=(0:1/9:1)'; % multiplying effect of fertility class on fertility

%% we combine all those parameter vector of parameters
psizes= [length(surv) length(fert) length(parityeffect) length(classeffect)
length(classdistrib)] ;% sizes of parameter vectors
p= [surv' fert' parityeffect' classeffect' classdistrib ]; %concatenation of
parameter vectors into one large parameter vector p
psize=sum(psizes); % total size of parameter vector
pnames = char('surv','fert','parityeffect','classeffect','classdistrib'); % name of
parameter vectors
```

```
parameterstable=[  1:length(psizes); 1 1+cumsum(psizes(1:end-
1));cumsum(psizes(1:end))]';
parametercell=[cellstr(pnames) num2cell(parameterstable) num2cell(psizes') ]
%parameter table with names, parameter numbers, indices in p and sizes
clear pnames

%% end
```

VII. VEC-PERMUTATION CONSTRUCTION OF MPPM

```
% in order to generate our MPPM the vec-permutation way, we first need to
% decide for a chronological order for the various "processes" at play :
% demography (survival,fertility), "change in parity class" "change in
% heterogeneity class"

% First, as we can see, when the other traits are not related to physical
% properties : (size, stage, location) but hereditary, or kinship
% properties, the concept of sequential processes is hard to fathom

% Second, because those processes are sequential, not all MPPMs (taken in
% the broad sense of a matrix representing the projection from any n-tuple
% of traits to itself) can be represented/constructed via vec-permutation
% example of age/location where the mother's pattern for the location of
% its eggs depends also on the mother's age.

% Here we decide, the processes are
% demographic (individual survives and/or gives birth) then the parity of
% the outcome of this demographic process changes, then its heredity class


%% we use the same input than our construction :
% from it we generate the same  traits/states measure and the same vectors
% of survival and fertility for all states :
clear all
% load parameters (from part VI)

maxparity=length(parityeffect)-1; % maximum parity is inferred from the fertility-
parity trade-off coefficients
s=[length(surv)+1 (maxparity+1) length(classeffect)]; % (ordered) trait sizes
vector
n=size(s,2); %number of traits
q=prod(s); %number of states
Vrsurvival=repmat([surv' 0]', [1 (maxparity+1) length(classeffect)]) ; %repmat
matlab function replicates a vector (first argument) along the dimension(s) in the
second argument
Vrfertility=reshape(vec(fert*parityeffect')*classeffect',s);
%% then following caswell presentation ..
% we are supposed to have the q*(1/s1 + 1/s2 + ...) matrices representing the
transitions, for each trait, between the different values of that trait when all
other traits are fixed
% they are called the transition matrices and are stored in a :
% a{1}(:,:,par,het) for age , a{2}(:,:,age,het) for parity, and  a{3}(:,:,age,par)
for heterogenity
m=n; siz=s ;
% we dont have them so have to construct them. this is equivalent to
% constructing oi and op in our vector-bases approach.

%a{1}
% a{1}(:,:,par,het) is a leslie matrix which fertility rates are
Vrfertility(:,par,het) and survival Vrfertility(:,par,het)
for par=1:s(2)
for het=1:s(3)
a{1}(:,:,par,het)=diag(Vrsurvival(1:(s(1)-1),par,het),-1);
a{1}(1,:,par,het)=Vrfertility(:,par,het);end
end
```

```
%a{2}
%a{2}(:,:,age,het) is the 12*12 matrix (parities) for individuals of age age and
heterogeneity het
% if age=1 it means, we deal with a newborn, its parity has to become 1
% otherwise parity remains or increments depending on the fertility rate of (age-1)
% indeed the demographic process, ageing the individuals, has already passed
for het=1:s(3)
    for age=1:1
a{2}(:,:,age,het)= zeros(s(2),s(2));
a{2}(1,:,age,het)= ones(1,s(2));
    end
    for age=2:s(1)
a{2}(:,:,age,het)= diag(1-Vrfertility(age-1,1:s(2),het)) + diag(Vrfertility(age-
1,1:(s(2)-1),het),-1);
    end
end

%%a{3}
% a{3}(:,:,age,par) is the 10*10 matrix (heterogeneity classes) for
% individuals of age age and parity par
% if age 1, new born, then the distribution is 'classdistrib'
%otherwise it is an invariant
for par=1:s(2)
    for age=1:1
a{3}(:,:,age,par)=repmat(classdistrib',1,s(3)) ;
    end
    for age=2:s(1)
a{3}(:,:,age,par)=eye(s(3)) ;
    end
end
%% CONSTRUCTION
tic
%% Create the block diagonal matrices A1...Am with the function BD proj mat and
% store them in a cell array A of size m as follows:

% for k=1:m
%     A{k}=BD_proj_mat(a{k});
% end

% 'out of memory' so we have tweaked the code to make it sparse
for k=1:m
    A{k}=BD_proj_mat_sp(a{k});
end

%% Construct the total projection matrix MM, according to equation
%MM=hyper_state_matrix(siz,A);

% 'out of memory' so we have tweaked the code to make it sparse
MM=hyper_state_matrix_sp(siz,A);
vecpermutation_construction_time=toc

%% M and MM are the same

% M construction
tic
order=reshape(1:q,s); % state indices redistributed over the 3 dimensions of ts
tS1=Vrsurvival.*(1-Vrfertility);
T1=[vec(order(1:end-1,:,:)) vec(order(2:end,:,:)) vec(tS1(1:end-1,:,:))] ;
tS2=Vrsurvival.*Vrfertility;
T2=[vec(order(1:end-1,1:end-1,:)) vec(order(2:end,2:end,:)) vec(tS2(1:end-1,1:end-
1,:))] ;
T3=[ repmat(vec(order),length(classdistrib),1)
vec(repmat(vec(order(1,1,1:length(classdistrib)))',q,1))
vec(Vrfertility(:)*classdistrib) ]; %transition triplet for fert towards 1, 1, 1:10
T=vertcat(T1,T2,T3);
```

```
M=sparse(T(:,2),T(:,1),T(:,3),q,q); %
vectorbased_construction_time=toc

% comparison 2 methods
M-MM
```

VIII. VEC-PERMUTATION SUBFUNCTIONS FROM AND ADAPTED FROM (ROTH AND CASWELL 2016)

```
% Those functions are taken from (Roth and Caswell 2016)
% some are adapted by us to use sparsity (function names ending in _sp)
function A = BD_proj_mat(B)
siz=size(B);
siz=siz(2:end);
s=prod(siz);                    %size of the expected block-diagonal matrix
sk=s/siz(1);                    %number of block on the diagonal
siz=siz(2:end);                 %maximal value of each index i_1,...,i_r
A=zeros(s,s);
 for i=1:sk
 A=A+kron(Emat(i,i,sk),B(:,:,ind2sub(siz,i)));
end
end

function A = BD_proj_mat_sp(B)
siz=size(B);
siz=siz(2:end);
s=prod(siz);                    %size of the expected block-diagonal matrix
sk=s/siz(1);                    %number of block on the diagonal
siz=siz(2:end);                 %maximal value of each index i_1,...,i_r
A=sparse(zeros(s,s));
 for i=1:sk
 A=A+sparse(kron(sparse(Emat(i,i,sk)),sparse(B(:,:,ind2sub(siz,i)))));
end
end

function k= Qmat(u,v)
 k=zeros(u*v);
a=zeros(u,v);
for i=1:u
    for j = 1:v
        e=a;
        e(i,j)=1;
        k = k+ kron(e,e');
    end
end


function k= Qmat_sp(u,v)
k=sparse(zeros(u*v));
a=sparse(zeros(u,v));
for i=1:u
    for j = 1:v
        e=a;
        e(i,j)=1;
        k = k+ sparse(kron(e,e'));
    end
end

function p = vecperm_hyp(k,s)
[~,m]=size(s);
if k==1
    p=eye(prod(s));
elseif k == m
    p = Qmat(prod(s(1:m-1)),s(m))*kron(eye(prod(s(m))),Qmat(prod(s(1:m-2)),s(m-
1)))';
```

```
else
    p=kron(eye(prod(s(k+1:m))),Qmat(prod(s(1:k-
1)),s(k)))*kron(eye(prod(s(k:m))),Qmat(prod(s(1:k-2)),s(k-1)))';
end

function p = vecperm_hyp_sp(k,s)
[~,m]=size(s);
if k==1
    p=sparse(eye(prod(s)));
elseif k == m
    p = Qmat_sp(prod(s(1:m-
1)),s(m))*sparse(kron(sparse(eye(prod(s(m)))),Qmat_sp(prod(s(1:m-2)),s(m-1))))';
else
    p=sparse(kron(sparse(eye(prod(s(k+1:m)))),Qmat_sp(prod(s(1:k-
1)),s(k))))*sparse(kron(sparse(eye(prod(s(k:m)))),Qmat_sp(prod(s(1:k-2)),s(k-
1))))';
end

function E= Emat(i,j,u)
    E=zeros(u);
    E(i,j)=1;
End

function hsm=hyper_state_matrix(siz,A)
m=length(siz);          %number of dimensions
s=prod(siz);
K=eye(s);
tildA=A{1};
for k=2:m
    tildA=A{k}*vecperm_hyp(k,siz)*tildA;
    K=vecperm_hyp(k,siz)*K;
end
hsm=K'*tildA;
end

function hsm=hyper_state_matrix_sp(siz,A)
m=length(siz);          %number of dimensions
s=prod(siz);
K=sparse(eye(s));
tildA=sparse(A{1});
for k=2:m
    tildA=sparse(A{k})*sparse(vecperm_hyp_sp(k,siz))*sparse(tildA);
    K=sparse(vecperm_hyp_sp(k,siz))*sparse(K);
end
hsm=sparse(K)'*sparse(tildA);
end
```

IX. BIBLIOGRAPHY

Roth, G., Caswell, H., 2016. Hyperstate matrix models: extending demographic state spaces to higher dimensions. Methods Ecol. Evol. doi:10.1111/2041-210X.12622

5.1 Annexes from chapter 3

5.1.1 Survival selection gradients

In figure 5.1, we illustrate the fact that the (*age-parity-heterogeneity*)-MPPM of this chapter does not exhibit any variation in survival selection gradients between the full model implementing *physiological costs of reproduction* and *heterogeneity* and the *folded* models with no trade-off and/or no heterogeneity. Indeed, whilst aiming to model the 'Stearnian' intermediate structure (Stearns, 1989a) of such costs, our model, because a population model and not a physiological one, can only implement its immediate effects at the phenotypic level, i.e. the vital rates. A matrix cannot model consequences of the realization of allocations towards current reproductive effort and future reproductive value, only the consequences of the resulting realization of vital rates (on other vital rates).

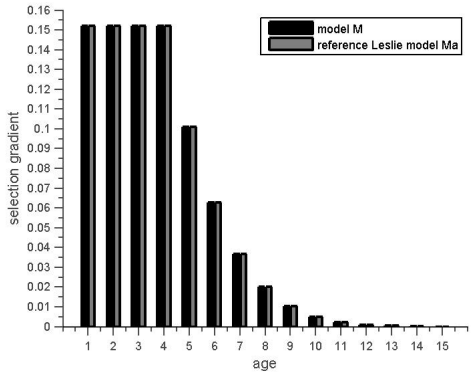


Figure 5.1: selection gradient measured by the elasticity of ergodic growth rate to survival rates, summed by age, for $\mathbf{M}$ modeling an (*age-parity*) population with physiological costs of reproduction and $\mathbf{M_a}$ its reference Leslie matrix, which is $\mathbf{M}$ folded on *parity*, modeling the same population but characterized only by age. The population has maximum age $\omega = 15$ and age-at-maturity $\alpha = 5$. The zero-parity fertility and survival rates are 0.85. Cost of reproduction is modeled by relatively decreasing each vital rate by $1/(1+\omega-\alpha)$ per parity.

Simply put, in a matrix model, it is possible to implement the effect of realized and unrealized fertility events - this is what this chapter is about - but not the effect of realization of survival process, because of the rather definitive consequence of the latter on an individual. However the physiological mechanism for survival is finer that the all-or-nothing phenotypic result: an individual may allocate more or less towards survival at a given time-step. Let us hint at a way to measure the force of such "survival effort" on λ(by contrast with the meaningless survival selection gradient we just considered)

Let us consider that "survival effort" at age a, $se(a)$ is constrained by reproductive effort $re(a) = f(a)$, i.e. fertility, at that age. This constraint can be best approximated by using an *optimality theory* approach on the model devoid of trade-off on either fertility or survival rate : $\mathbf{M_a}$. By considering that in this model, that does not contain any asymmetry in treatment of vital rates, λ is locally maximized - $d\ln f(a).\frac{\partial \ln \lambda}{\partial \ln f(a)} + d\ln s(a).\frac{\partial \ln \lambda}{\partial \ln s(a)} = d\ln\lambda = 0$ - we get the ratio of survival to reproductive effort at age a, $r(a)$: $r(a) = \frac{\partial \ln \lambda / \partial \ln f_{\mathbf{M}_a}(a)}{\partial \ln \lambda / \partial \ln s_{\mathbf{M}_a}(a)} = \frac{\left\{\mathbf{e}^f_{\mathbf{M}a}\right\}_a}{\left\{\mathbf{e}^s_{\mathbf{M}a}\right\}_a}$. This equation provides the iso-fitness (i.e. λ is constant) possible distribution of allocation towards current reproductive effort and survival effort : $-d\ln f(a).r(a) + d\ln s(a) = 0$. Making the assumption, this ratio holds for $\mathbf{M}$, we have $r(a) = \frac{\partial \ln \lambda / \partial \ln f_{\mathbf{M}}(a)}{\partial \ln \lambda / \partial \ln s_{\mathbf{M}}(a)}$ and therefore from there, we have the "survival effort" selection gradient :

$$\frac{\partial \ln \lambda}{\partial \ln s_{\mathbf{M}}(a)} = \frac{\left\{\mathbf{e}^f_{\mathbf{M}}\right\}_a}{r(a)}$$

We depict "survival effort" selection gradients in figure 5.2.

This approach shows a way to try and quantify the selection gradient on survival effort that takes into account the allocation process that selection gradient on survival rates does not. However, this is still not satisfactory as it does consider that the allocation only starts when sexual maturity is reached, hence a jump between selection gradient before *alpha* where every increase in survival effort bears no cost, and after *alpha* where the energy allocated towards survival effort is effectively taken from reproductive effort.

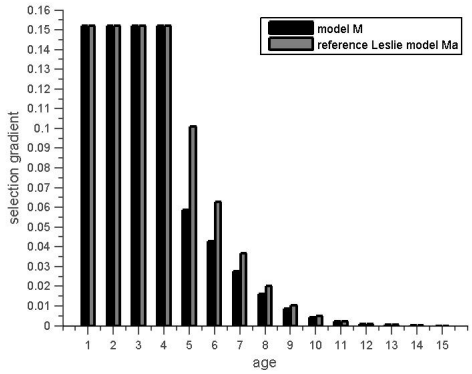
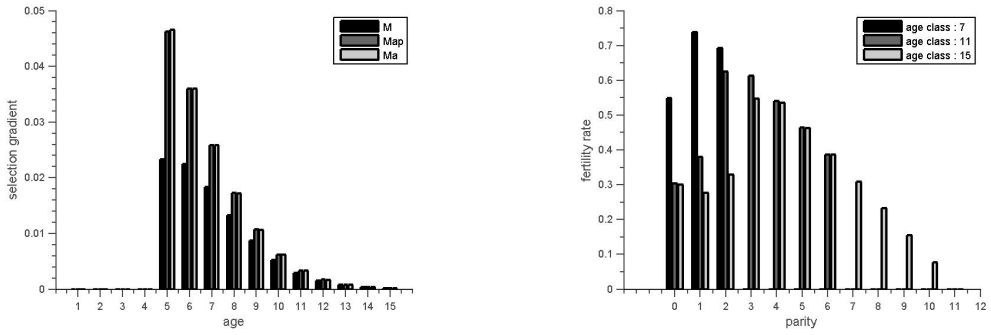


Figure 5.2: selection gradient measured by the elasticity of ergodic growth rate to "survival effort", summed by age, for $\mathbf{M}$ modeling an (*age-parity*) population with physiological costs of reproduction and $\mathbf{M_a}$ its reference Leslie matrix, which is $\mathbf{M}$ folded on *parity*, modeling the same population but characterized only by age. For $\mathbf{M_a}$ this 'survival effort' selection gradient is equated to elasticity to survival rates. For $\mathbf{M}$, it is calculated by multiplying the ratio of survival over fertility selection gradient for $\mathbf{M_a}$ by $\mathbf{M}$'s fertility elasticity . The population has maximum age $\omega = 15$ and age-at-maturity $\alpha = 5$. The zero-parity fertility and survival rates are 0.85. Cost of reproduction is modeled by relatively decreasing each vital rate by $1/(1+\omega-\alpha)$ per parity.

5.1.2 Heterogeneity effect on detectability of costs when heterogeneity is related to variance in acquisition capability

We illustrate the effect of genotypic heterogeneity characterized by variance in acquisition strategies (overall differences in fitness between the genotypes), by two figures. Figure 5.3a displays the selection gradients and figure 5.3b the inferred fertility rates by parity, for such an heterogeneous population



(a) (b)

Figure 5.3: Plot of selection gradient (fig.5.3a) measured by the elasticity of ergodic growth rate to fertility rates,summed by age and inferred fertility rates by parity (fig. 5.3b) for heterogeneous population $\mathbf{C}$ with variance in acquisition strategy. Population $\mathbf{C}$ is constituted of genotypes $\mathbf{G}_1$ and $\mathbf{G}_4$. Mutation rate is $m = 0.3$. The inferred fertility rates are obtained from each model *folded* on *heterogeneity*. Both populations are modeled by (*age-parity-heterogeneity*)-MPPMS with physiological costs of reproduction. The population has maximum age $\omega = 15$ and age-at-maturity $\alpha = 5$. The zero-parity fertility and survival rates of $\mathbf{G}_1$ are 0.85. Fertility rate is 0.3 and survival rate 0.6 for $\mathbf{G}_4$ modeling a less fit genotype with lower fertility and lower survival rates. Fitness of $\mathbf{G}_1$ is 1.05 and fitness of $\mathbf{G}_4$ is 0.7. Cost of reproduction is modeled by relatively decreasing each vital rate by $1/(1+\omega-\alpha)$ per parity.

5.1.3 Illustration of general non-preservation of $\mathbf{R_0}$ by *folding*

Consider 2 traits, t_1 and t_2, with 2 trait values each, and where there is only 1 offspring state: $(1,1)$. Let us further consider that the (t_1-t_2)-MPPM for this population is : $\mathbf{M} = \begin{bmatrix} 0.6 & 0.6 & 0.6 & 0.6 \\ 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0 \end{bmatrix}$. Then we get $\lambda_{\mathbf{M}} = 1.0317$ and - by letting $\mathbf{F}$ be the first line of $\mathbf{M}$, $\mathbf{T}$ the complement of $\mathbf{F}$ in $\mathbf{M}$, and $\mathbf{R_0}$

the first element of $\mathbf{R} = \mathbf{F}.\mathbf{N}$ (see appendix 3.0.3 of chapter III, page 93) - we also get $\mathbf{R_0} = 1.05$. The eigen-analysis of $\mathbf{M}$ allows us to fold it over t_1 (see chapter 2 (p.41)), and we get $\mathbf{M}_{t_2}^{fold} = \begin{bmatrix} 0.9368 & 0.6 \\ 0.1632 & 0 \end{bmatrix}$.

Eigen-analysis of $\mathbf{M}_{t_2}^{fold}$ yields $\lambda_{\mathbf{M}_{t_2}^{fold}} = 1.0317 = \lambda_{\mathbf{M}}$. By construction, in this matrix also, offspring are only to be found on the first line. And therefore, we can proceed as we just did for $\mathbf{M}$, to generate the net reproductive rate, and we get $\mathbf{R_0} = 1.035$. Therefore this simple model illustrates the general non-preservation of $\mathbf{R_0}$ by EFP-*folding*.

5.1.4 demographic and environmental variances calculations for M and $\mathbf{M_a}$

In this section, we wish show calculation steps to calculate σ_d^2 and σ_e^2 for model $\mathbf{M}$ implementing the costs and $\mathbf{M_a}$, which is $\mathbf{M}$ folded over *parity*, in which the costs are absent, using the eigen-attributes of $\mathbf{M}$.

σ_d^2

Let us first restate the generic formulation of σ_d^2 : $\sigma_d^2 = \sum_{i,j}\sum_{k,l} \lambda^{-2} \frac{\partial \lambda}{\partial \mathrm{M}_{i,j}} \frac{\partial \lambda}{\partial \mathrm{M}_{k,l}} .N.\mathrm{Cov_d}(\mathrm{M}_{i,j}, \mathrm{M}_{k,l})$. We also know the demographic covariances of matrix elements for $\mathbf{M}$ from section 3.3.3. For $\mathbf{M_a}$ the demographic covariances are simpler : $\mathrm{Var}(\mathcal{F}_i) = f_i(1-f_i)$ and $\mathrm{Var}(\mathcal{S}_i) = f_i(1-s_i)$, and all other covariances are zero (each vital process is independent from all others). Then for $\mathbf{M_a}$, $\sigma_d^2 = \lambda^{-2} N \sum_i \left[\mathrm{S}_{1,i}^2 f_i(1-f_i) + \mathrm{S}_{i+1,i}^2 s_i(1-s_i)\right]$ where $\mathbf{S}$ is the sensitivity matrix; specifically in that case $\mathbf{S} = \{v_i w_j\}_{i,j}$. And for $\mathbf{M}$ (processes among ages are dependent but not between ages) for which we denote i^+, for each state i, the state corresponding to the next age class and next parity class, i.e. the state reached, from i, through successful reproduction, and i^- the other reachable state.

$$\begin{cases} \sigma_d^2 = \lambda^{-2} N \sum_i \left[\mathrm{S}_{1,i}^2 f_i(1-f_i) + \mathrm{S}_{1,i}\mathrm{S}_{i^+,i}(1-f_i)f_i s_i - \mathrm{S}_{1,i}\mathrm{S}_{i^-,i}(1-f_i)f_i s_i + A\right] \\ A = (\mathrm{S}_{i^+,i}^2 - \mathrm{S}_{i^-,i}^2) f_i s_i (1 - f_i s_i) + \mathrm{S}_{i^-,i}^2 s_i(1-s_i) - \mathrm{S}_{i^+,i}\mathrm{S}_{i^-,i}((1-f_i)f_i s_i^2) \end{cases}$$

which we can rewrite:

$$\begin{cases} \sigma_d^2 = \lambda^{-2} N \sum_i (C + D) \\ C = \mathrm{S}_{1,i}^2 f_i(1-f_i) + \mathrm{S}_{i^-,i}^2 s_i(1-s_i) \\ D = s_i \mathrm{S}_{1,i}(\mathrm{S}_{i^+,i} - \mathrm{S}_{i^-,i}) - (\mathrm{S}_{i^+,i}^2 - \mathrm{S}_{i^-,i}^2) f_i s_i (1 - f_i s_i) - \mathrm{S}_{i^+,i}\mathrm{S}_{i^-,i}((1-f_i)f_i s_i^2) \end{cases}$$

C contains elements equivalent to these of the demographic variance of the Leslie matrix and D contains the sum of three negative elements. Indeed $\mathrm{S}_{i^+,i} < -\mathrm{S}_{i^-,i}$ as higher parity individuals have a smaller reproductive value than lower parity ones. This is because the difference of sensitivities of λ to elements from the same state is proportional to the difference in the reproductive values of the destination states.

σ_e^2

Let us first restate the generic formulation of σ_d^2 :$\sigma_e^2 = \sum_{i,j}\sum_{k,l} \lambda^{-2} \frac{\partial \lambda}{\partial \mathrm{M}_{i,j}} \frac{\partial \lambda}{\partial \mathrm{M}_{k,l}} .\mathrm{Cov_e}(\mathrm{M}_{i,j}, \mathrm{M}_{k,l})$ We know the demographic covariances of matrix elements for $\mathbf{M}$ from section 3.3.3. The demographic covariances of matrix elements for $\mathbf{M}^*$ are $\mathrm{Cov}(\mathcal{F}_i, \mathcal{F}_j) = \epsilon f_i f_j$ and all other covariances are zero. Then for $\mathbf{M}^*$, $\sigma_e^2 = \epsilon \lambda^{-2} \sum_{i,j} \mathrm{S}_{1,i}\mathrm{S}_{1,j} f_i f_j$, where $\mathbf{S}$ is the sensitivity matrix; specifically in that case $\mathbf{S} = \{v_i w_j\}_{i,j}$. For $\mathbf{M}$ (all processes are interdependent with regards to the environment) (we denote i^+, for each state i, the state corresponding to the next age class and next parity class, i.e. the state reached, from i, through successful reproduction, and i^- the other reachable state).

$$\sigma_e^2 = \epsilon \lambda^{-2} \sum_{i,j} \left[\mathrm{S}_{1,i}\mathrm{S}_{1,j} f_i f_j + (\mathrm{S}_{i^+,i} - \mathrm{S}_{i^-,i})\mathrm{S}_{1,i} f_i f_j s_j + (\mathrm{S}_{i^+,i} - \mathrm{S}_{i^-,i})(\mathrm{S}_{j^+,i} + \mathrm{S}_{j^-,i}) f_i f_j s_i s_j\right]$$

Like for demographic variance, we find, inside the sum, the first element, akin the one in the Leslie case, and then a sum of negative elements (same reason as for demographic variance)

5.2 Annex from chapter 4

Manuscript in preparation

Are old ages useless? Shedding light on the allelic spectrum of late-onset diseases

Samuel Pavard[1*], Christophe F.D. Coste[1]

[1] UMR 7206 EcoAnthropologie et Ethnobiologie, MNHN, Université Paris Diderot, F-75016, Paris, France

* Corresponding author. Email address: Samuel.pavard@mnhn.fr. Postal address : UMR 7206, Laboratoire d'Eco-anthropologie et Ethnobiologie, Equipe "Anthropologie Evolutive", Muséum National d'Histoire Naturelle, Site du Musée de l'Homme, 17 place du Trocadéro, F-75116 Paris, France

Introduction

The role played by selection in shaping the allelic spectrum of common late-onset genetic diseases is yet unknown. It is however at the core of the long lasting debate between the tenants of the Common Disease Common Variant ('CDCV') and those of the Common Diseases Rare Variant ('CDRV') hypotheses (reviewed in 1, 2). In the 'CDCV', Susceptibility Alleles to Late-Onset Diseases (hereafter denoted 'SALOD') are considered as close to neutrality. Genetic drift is the dominant evolutionary force shaping the diseases' allelic spectrum and it thereby predicts that most diseases prevalence results from a limited number of alleles at moderate/high frequency (theorized by 3; see box 1). By contrast, the 'CDRV' emphasizes the importance of numerous rare variants in diseases prevalence resulting from a balance between mutations and purifying selection at a large number of loci (theorized by (4) and compared to (3) in (5)).

Several lines of evidence can be discussed on whether purifying selection occurs on SALOD. The first line of evidence comes from the fact that the allelic spectrum of common late-onset diseases is mostly characterized by a multitude of rare and recent variants: a pattern more compatible with the CDRV than with the CDCV hypothesis. This is notably the cases of familial forms of cancer, coronary artery and Alzheimer dementia (6, 7). The example the most frequently cited is that of familial form of breast and ovarian cancers. More than 2000 mutations in the *BRCA1* and *BRCA2* genes have been associated with larger risk of breast and ovarian cancer; most of them are found at very low frequencies and most of those reaching larger frequencies do so due to founder effects and populations history (8). The CDRV hypothesis also receives support from the crossing between genetic and epidemiological data demonstrating that alleles frequencies decline as a function of alleles' effect size in the case of late-onset diseases (9-11) which suggests a differential of level of selection between SALODs. These studies also find unquestionable signature of antagonistic pleiotropy: the fact that many alleles frequency is also determined by positive selection from a beneficial effect at young age even to the detriment of a deleterious effect at later age. Finally selection test found no difference in magnitude of selection between mutations causing early- *vs* late-onset disorders which may suggests that SALODs remains under significant selection (12). Although evidence for selection on SALOD accumulates, the dominant idea is still that late-onset diseases did not compromised reproductive success. When discussed, evidence for selection is rather explained by antagonistic pleiotropy or by selective effects due to recent environmental changes (12-14).

Box 1 – Theory behind the 'Common (late-onset) Diseases Common Variant' hypothesis.

In the late nineties, the quest to unravel the allelic spectrum of common genetic diseases took a new turn with the development of DNA sequencing and microarrays. Massive effort has been invested in genetic and epidemiological research in the hope to eventually 'map' the individual genetic susceptibility to common genetic diseases; especially for late-onset diseases in a context of ageing populations. Together with these technological developments, this hope has been supported by the theoretical prediction that most of the prevalence of common late-onset diseases is associated with a small number of allelic variants coexisting at large frequencies in a given population.

This 'Common Diseases Common Variant' (CDCV) hypothesis postulates that common diseases are either complex and/or late onset (15). Complex diseases, in which susceptibility alleles modify only a single element of a complex molecular circuitry regulated by epistatic relationships between many loci, are opposed to diseases which transmission obeys Mendelian segregation rule (reviewed in (1)). Susceptibility alleles to complex diseases are thought to be leading to low disease penetrance that does not significantly compromise carriers reproductive success and are therefore consider under negligible purifying selection. Late-onset diseases are - generally defined as diseases whose mean age at onset is later than age 40 (12) – are incorporated into the CDCV hypothesis, whether they are complex (as Alzheimer) or quasi-Mendelian (as *BRCA1* and *BRCA2* forms of breast cancer) (14). Magnitude of selection is indeed predicted to decrease drastically with age: the later in life an allele is deleterious, the lower the proportion of carriers surviving at these ages, and the smaller the fraction of reproductive success (RS) compromised by this deleterious effect (16, 17). Because of menopause in humans, this fraction is zero in post-reproductive years during which the death of a woman no longer reduces her already achieved RS. The biomedical community has also long believed that the males and females survivors past 45 years were anyway too few in pre-industrial humans for alleles deleterious past these ages to compromise carriers' fitness.

Being under negligible negative selection, the fate of alleles involved in complex and/or late-onset diseases is therefore mainly driven by genetic drift, which tends to randomly eliminate most of the genetic variants while the remaining ones reach higher frequencies (Reich and Lander 2001). More precisely, the authors predict that, in a growing population, the disappearance of susceptibility alleles with slightly deleterious (corresponding to a selection coefficient about $s=10^{-6}$) effect is sufficiently slowed down by genetic drift to facilitate the emergence of common variants. As a consequence, the CDCV predicts that a small number of alleles at high frequency are expected to account for most of the prevalence of familial form of common diseases. By contrast, the Common Diseases Rare Variants (CDRV) hypothesis a mild purifying selection (of an order of magnitude of $s=10^{-4}$) which, combined with genetic drift into a multilocus Wright-Fischer model, can lead to the accumulation of rare variants and the maintaining of few variants at intermediate frequencies (4).

Estimating magnitude of selection on SALOD is therefore crucial for understanding the allelic spectrum of late-onset diseases. Even more, we need to estimate how fast purifying selection declines with SALOD-specific mean age at diseases onset to assess the potential for positive selection due to antagonistic pleiotropic effects: a faster decline means that a smaller beneficial effect a young age can be positively selected at the expanse of a lager cost at older age. Both magnitude and decline are thus the missing parameters to population genetic models essential to predict the respective share of late-onset diseases prevalence originating from common and rare variants. To do this, one has to link an allele age-specific

phenotypic expression to its carrier's fitness. But was this link can it be? Three phenomena may explain the persistence of selection on SALOD:

Variance in disease onset – A particular familial form of a late-onset disease never occurs exactly at one specific age. Rather, onset probability spread across a range of ages according to an age-specific risk function characterized *a minima* by a mean and a variance (as illustrated in figure 1). It was first suggested by (18) that late-onset disease may therefore occur in a non-negligible proportion of young individuals even though mean age at onset is late in life. Such a proportion could significantly compromise the average reproductive success of SALOD carriers. This has been proved true in the case of BRCA1 alleles (19): although the mean age of onset occurs after menopause (around 55 years old), between 15 to 27% of cases occur before age 45, allowing negative selection to operate. To our knowledge, variance in diseases onset was not accounted for in previous tests of selection (10, 12). Of course this variance reflect in part the fact that age at onset is influenced by the type or the location of mutations on susceptibility genes (e.g, in the case of polycystic kidney diseases linked to *PKD1* and *PKD2, (20)*) or by epistatic relationships between mutations and risk-modifying loci (e.g., (21) for Alzeimer Dementia linked to *PSEN2*, (22) for breast and ovarian cancer linked to *BRCA1,* or (23) for Huntington disease linked to *HTT*). However aggregated data at the gene level gives a fair suspicion that large variance at onset is an ubiquitous characteristic of familial form of late-onset diseases. Even more, it has been shown in the case of neurodegenerative disorders that genetically-driven diseases occur on average sooner than sporadic not-familial diseases (24). There is also fair suspicion that variance in age at onset increases with mean age at onset (see SM2 for empirical data) as predicted in (25)**.** Huntington disease, a unique case where mutation-specific age at onset distributions is known, variance increases with mean age at onset (26).

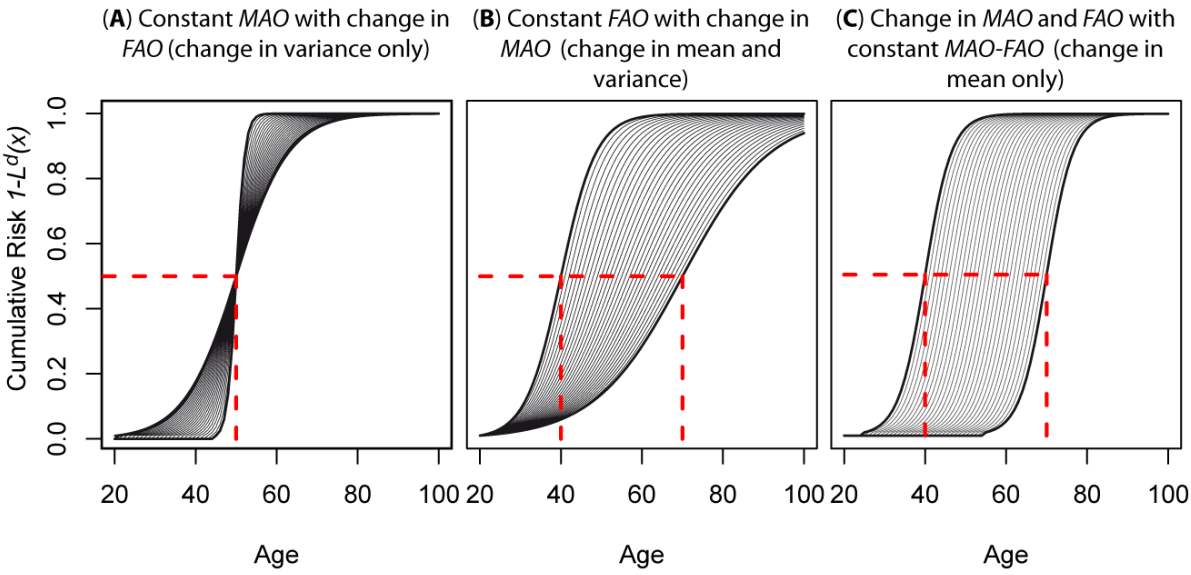


Figure 1 - Cumulative risk of disease onset with age according to change in Mean Age at disease Onset (*MAO*) and First Age at disease Onset (*FAO*). If disease onset is confounded with morbidity or full incapacitation, cumulative risk is $1-L^d(x)$ where $L^d(x)$ is the disease specific survival at age x. Age at onset distribution if fitted with a two parameters logistic function (see SM2). *FAO* is defined as the age at which one 1% of SALOD carrier have develop the disease ($1-L^d(x)=0.01$ and setting $1-L^d(x<FAO)=0$). This modeling approach allows the exploration of a parameter space encompassing distribution of constant *MAO* and varying *FAO* (Left panel); constant *FAO* with varying *MAO* (middle panel) and constant *MAO-FAO* but varying in *MAO* (right panel). An incomplete penetrance at age 100 can be modeled by multiplying these functions by a factor lower than 1.

Reproduction of men at old ages - Although women stop reproducing at menopause, men can pursue reproduction until late in life. Tuljapurkar et al. (2007) (27) showed that male fertility provides a non-negligible selective force against autosomal deleterious mutations at ages far past female menopause. The proportion of children born to old men depends on population survival and matrimonial behaviors. To the opposite of great apes where males seems to favor reproduction with older females (28), human males are prone to reproduce with younger women. First, men are on average older than women at first marriage and reproduction (e.g., men are six years older in average (29); and see SM3), even in societies were sexuality is not prohibited before marriage (30) and in matrilineal/matrilocal societies (31, 32). Second, because widowed or divorced men can remarry even at old ages. In absence of modern medicine, around 20% of men surviving at age 50 years old have lost their first wife (see SM3). Divorce, nearly absent in western Christian historical populations, are allowed in most ethnological populations (33) and may reach very large rates (e.g., in Ache, 61% of marriage end up during the first year (34)). Third, most past-human

populations were likely mildly polygynous (35). Polygyny is often associated with large difference of age between husband and wife; for example because men has to queue for access to matrimonial market (ie, “wealth increasing” poygyny) or because they marry the younger sisters of their first wife (ie, sororeal polygyny, SM3).

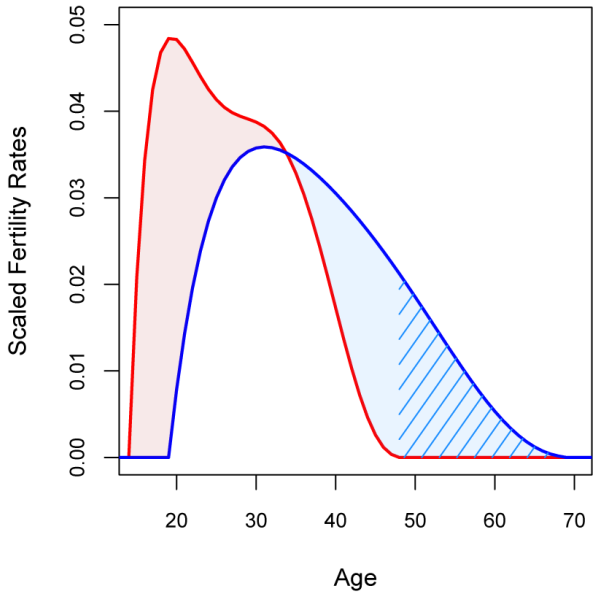


Figure 2 - Scaled age-specific fertility rates for women ($F(x)$, red) and men ($F^m(x)$, blue) fitted with a Brass-Polynomial Function from published estimates for non-industrialized populations (i.e., Kung, Ache, Tsimane, rural Gambia, Peul Bandé, and French canadians of the XVII-XVIIIth centuries). Data and modeling are detailed in SM3. Blue area emphasizes ages at which males fertility is larger than that of females, and dashed area the part of this surface where female fertility is down to zero. Parameters for males are fine tuned in order for lifetime reproductive success of males to be equal to that of female in the case of a particular population mortality (here that of an averaged Hunter-Gatherer population (36)).

Parental and grandparental care – Humans are mammals. As such, neonates are altricial: a newborn that is not taken care for is doomed to death. But humans are also cooperative breeders where the mother is the primary caregiver but not the only one. Maternal care but also in a lesser extent paternal and grandmaternal care has been proved keeping children alive (37, 38). In demographic terms, this means than an infant or a child having loss its mother, father or maternal and/or paternal grandmother is less cared for, and have a lower chance of surviving until adulthood (see figure 3 in SM4). Form the (grand)parental perspective this means that a individuals may continue to increase their reproductive success by increasing their children chance to survive, even if they do not reproduce. Diseases morbidity or incapacitations have therefore an additional cost than prohibiting future reproduction by compromising the chance of already born children to reach adulthood and to reproduce. This effect may be large. For example, maternal care has been proved

sufficient to enhance selection maintain purifying selection in the ten yrs following menopause (39).

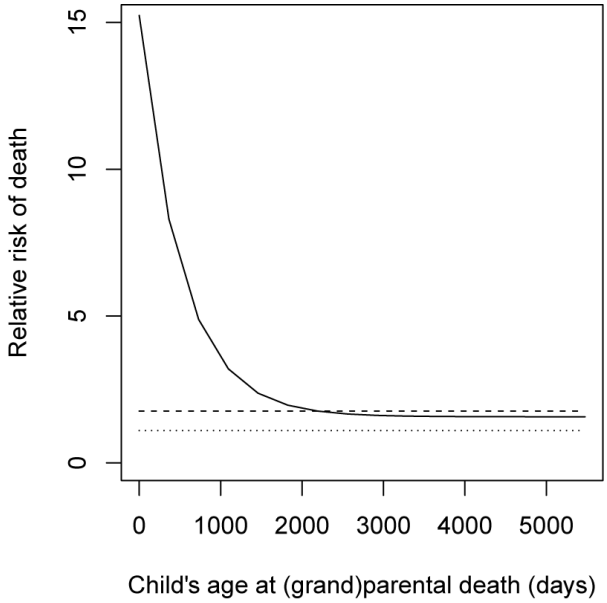


Figure 3 – Relative risk of death of a child having lost is mother (plain line), grandmother (dash line) or father (dot line) compared to those who have not, according to the child's age at maternal, (paternal or maternal) grandmaternal or paternal death (denoted respectively y_1, y_2 and y_3 in model). Data are fitted from published estimates from 31 articles detailed in SM4. All further calculations consider only the care provided by the maternal grandmother (see Box 2). Assuming that these risks are multiplicatives (see SM4), a child's age-specific relative risk of death $RR(t \mid y_1, y_2, y_3)$ can be calculated at any age t for any combination $[y_1, y_2, y_3]$; then its age-specific mortality hazard $h_0(t)RR(t \mid y_1, y_2, y_3)$ for any baseline mortality $h_0(t)$, and thus its mean survival until age at maturity α, $S(y_1, y_2, y_3) = \exp\left[-\int_0^{\alpha} h_0(t)RR(t \mid y_1, y_2, y_3)dt\right]$.

In this study we aim at calculating selection coefficients (i) for SALODs leading to diseases whose onset probability varies across ages; (ii) in a two-sex model where men can potentially reproduce until old ages and; (iii) in a three-generations model where child survival depends on maternal, paternal and grandmaternal care. The predictive value of such calculations lies in the careful discussion of these input parameters (hence detailed in SM2, SM3 and SM4). Assumptions made by the model and caveats are also fundamental (and listed in Box 2) because it emphasizes what we do not know; what we do not model to limit model's complexity or because, even if modeled, empirical data would not support enough careful parameterization; what may vary vastly between human population entailing results' generalization; and what we considered as negligible.

Box 2 – Main assumptions and caveats

Mutation is rare – Required by the model. Incorporating frequency-dependence is vastly complex and data for parameterization are missing. This hypothesis is however reasonable since most SALODs are rare or *de novo*. Our model will however fail to predict the fate of mutations having reached large frequency (as due to founder effects).

Mutation is autosomal, dominant, disease occurs in equally in both sexes, penetrance is complete – Not required by the model and results for recessive alleles, alleles located in mitochondrial or sexual chromosome, for disease occurring in females or males only, and for incomplete penetrance are respectively provided in Supp. Fig1 and Supp. Fig2.

No antagonistic pleiotropy – Could be incorporated. However, if antagonistic pleiotropic (AP) have been demonstrated at the genome level (10) and for particular couple gene-disease (as for *BRCA1* (40)), data on allele-specific effect sizes are still missing for parameterization. Decline of selection coefficients with mean age at onset provides nevertheless the potential for AP to occur.

No epistasis, no epigenetic and no environment-gene interactions – These mechanisms may modulate late-onset gene expressions; e.g., (21, 23, 41) for epistasis or (13) for environmental changes. However, their incorporation would be complex and data for parameterization are mostly missing. For example, our model allows to calculate selection coefficients for a given haplotypic combinations but would fail to predict its fate when recombination occurs.

Disease is lethal or fully incapacitating at a given age – Many late-onset diseases are slowly degenerative and/or not lethal (e.g., age-related macular degeneration or neurodegenerative diseases). Data are however missing on how the disease compromises individual's age-specific fitness through decreased fertility or the care towards offspring. Disease can also lead children to invest time and energy towards their parents' care (a phenomenon referred to as 'filial piety' by (42)) and as such entails their fitness.

Population is stable – Required by the model to access to population age-structure. Incorporating transient dynamics (through density-dependence or stochasticity) would make the model vastly more complicated.

No cost of reproduction, effect of care is similar on sons and daughters, no effect of parental care on children adult life – Required by the model. The incorporation of these phenomena to the model would make the Euler-Lotka equation unsolvable or requiring the addition of traits defining individuals. Recent development of Multitraits Population Projection Matrix model (Coste et al. 2017) may however make it possible. This is also promising to incorporate importance of transfer of behavior or knowledge for offspring survival and reproduction throughout life.

No difference in survival between males and females – Not a required by the model. Sex-specific survival data are however missing for not-industrial populations.

No effect of maternal and paternal age on offspring fitness – Not a required by the model. How parental age compromises children fitness because of increased gametes' mutation load is yet unknown. Moreover, many social factors increasing with parental age (as parental experience, influence and wealth) are by contrast correlated with better offspring survival.

Only one-side grandmaternal care (here the maternal one) is considered. Not a required by the model. However data suggest that most of the time, only one of the two grandmothers is primarily caring for grandchildren, and more frequently the maternal one (see SM4).

No structure of matrimony and fertility (beyond the fact that women reproduce with men of the same age or older). Required by the model. Demographic models for age-structured matrimony are available. However, matrimony varies widely across human population (see SM3) and incorporating it would add a new layer of complexity to the detriment of results' generalization. Moreover, this would also require to model and parameterize extramarital reproduction, remarriage and polygamy, which are also known to interact with maternal and paternal care.

Results and discussion

Old ages in the selection light - In the absence of pleiotropic effects, most Susceptibility Alleles to Late-Onset Disease (SALOD) are unlikely neutral, but rather purified by negative selection. First this is due to the epidemiology of the late-onset disease. Variance in disease onset means that a proportion of cases occurs at reproductive age even when Mean Age of Onset (*MAO*) is well beyond the end of reproduction. SALODs would be purified up to MAO between 60-65 years old if First Age at Onset (*FAO*) is 20 years earlier (corresponding to a standard deviation around the mean of 6.32 years, Fig. 4B) while selection would not occur beyond age 50 if *MAO* and *FAO* were confounded (Fig. 4A).

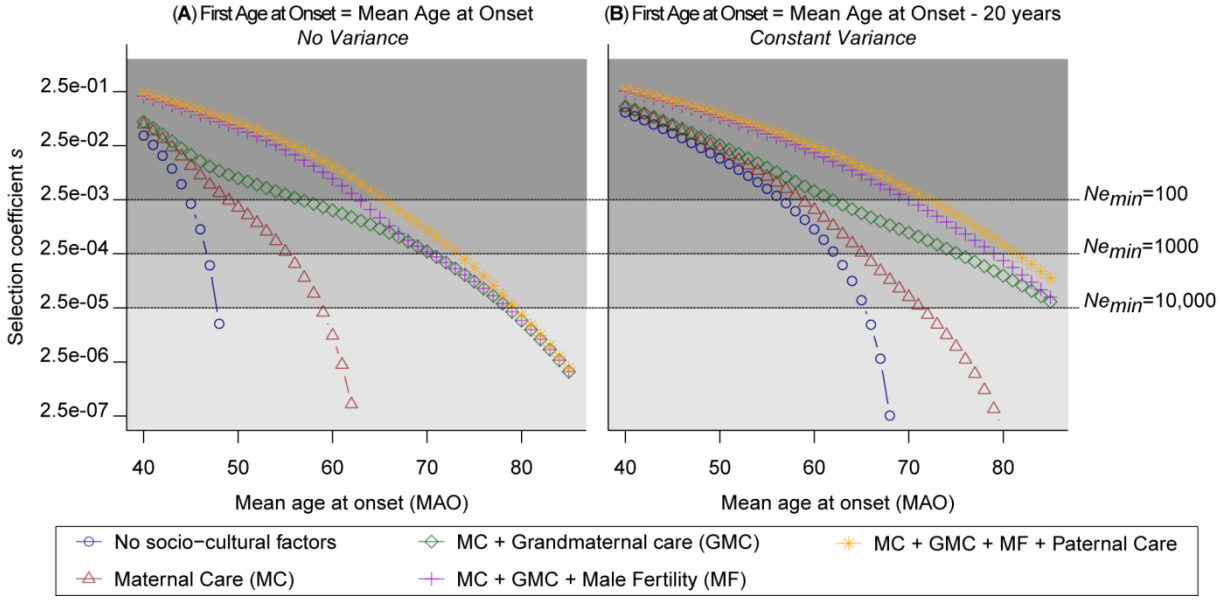


Figure 4 – Coefficient of selection *s* as a function of Mean Age at Disease Onset (*MAO*) in the case where disease onset has no variance (i.e., disease occurs at a unique age, **A**) and where age at first onset (*FAO*) is observed 20 years before the mean (constant Standard deviation equaled to 6.32 years; **B**). Calculations are calculated for various socio-cultural scenarios. Ne_{min} indicates, for one given selection coefficient, the minimum effective size for which $4N_e s > 1$, holding that, in humans, *Ne* is rarely lower than 100 and larger than 10,000. Mortality is that of a mean hunter-gatherer population (36). Allele is rare, autosomal, dominant, disease in both sexes and penetrance is complete.

Second selection on SALOD also arises from socio-cultural factors. Deaths occurring past 45 do not fall under a "selection shadow" (16) where selection has no leverages to operate. The fuel to negative selection comes primarily from the fact that, even in hunter-gatherers, individuals surviving at first reproductive age have large chance to survive past age 45 (here 63% (36)). [Calculations with survival of Sweden 1765 in SR4- **STILL TO DO**]. Then, individuals may enhance their fitness past this age. Maternal care (MC) has a large effect in magnitude but limited in time to 10-15 years after menopause (as in (39)). Grandmaternal care (GMC) has a large effect up to old ages (60-70 in Fig.1A and 60-85 in Fig.1B). This is

because, although GMC is much less important for child survival than that of MC (see Fig.3), increased risk of disease onset is much larger at grandmaternal age than at maternal age. Finally, if paternal care has little effect, males' ability to reproduce at ages where women are menopaused increases selection coefficients in a large extent.

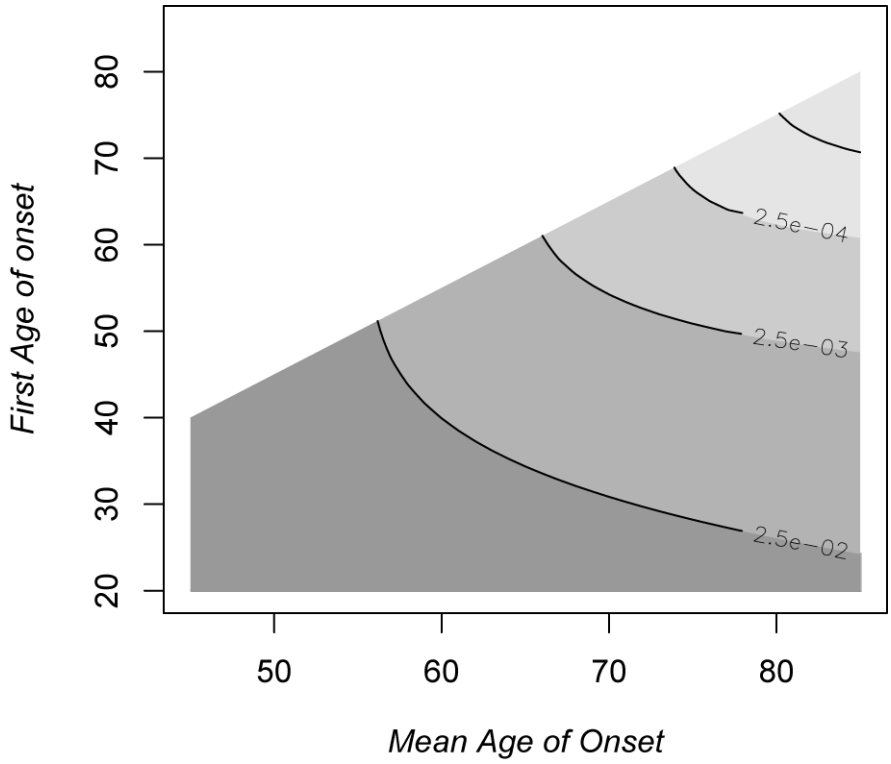


Figure 5 – Coefficient of selection *s* as a function of Mean Age at Disease Onset (*MAO*) and First Age at Onset (*FAO*). Above *s*=2.5e-02, Ne_{min}<100 and selection is expected in all human populations (darker grey). Below *s*=2.5e-04, Ne_{min}>10,000 and alleles are neutral in most human populations. In between selection levels of selection will vary of 'small' (dark grey) or large 'Ne' (light grey). Allele is rare, autosomal, dominant, disease in both sexes and penetrance is complete

When all socio-cultural factors are incorporated, an allele of *MAO-FAO*>20 will be under significant purifying selection in most human population, even if the *MAO* is at 80 years old of age. (Fig. 4B and Fig. 5). Levels of selection considered by (4), about 10x-04 is found only for very late-onset diseases (*MAO*>75) with low variance and those considered by (3), about 10x-06, never occur. Even more, an allele leading to *FAO* earlier than 25 will always be strongly negatively selected (s>10x-03, not shown), whatever the *MAO* is. This argues for a revision of the prediction done by (3, 4). Levels of selection is found in a similar range of magnitude for automosal alleles leading to sex-specific diseases or supported by Y or Mitochondrial chromosomes because of the dual effects of (grand)maternal care operating throughout the females lineages and late males' reproduction operating through the male lineage (see Supp. Fig1). By contrast, levels of selection are a linear function of diseases penetrance (when the age at onset distribution is conserved; see Supp. Fig2). A "rule of

thumbs" table of coefficients of selection as a function of mean and variance age at onset and cumulative risk at age 100 is provided in Supp. Table1. **[STILL TO DO]**

All these factors also change the shape of the decline of selection coefficient with *MAO,* making it less steep (from a more-than- to a less-than-exponential decline). Does this make antagonistic pleiotropy less likely to occur? Late-life deleterious effects (after age 60-70) will still entail 100 to 10,000 times less fitness than very early effect (before age 45). Early positive effect will still trade easily to very late negative effects. But this may be less easy against aged 40-60 deleterious effects than entails only 10 to 100 folds more fitness than early effects. This is important for example in the case of *BRCA1*. (40) find large beneficial effect of *BRCA1* mutations on carriers fertility which led (43) to propose that grandmother effects, maintaining deleterious effects under significant selection, was the best explanation to explain why alleles have not reached fixation. Here we confirm this and go beyond showing that grandmaternal care is an important but not the only factor maintaining SALODs in the "selection light".

Prediction for specific late-onset diseases – We predict selection coefficients for different couple gene-diseases (preliminary results in Figure 6 not accounting for uncomplete penetrance of the disease: Results are overestimated by a factor equaled to cumulative risk at age 100).

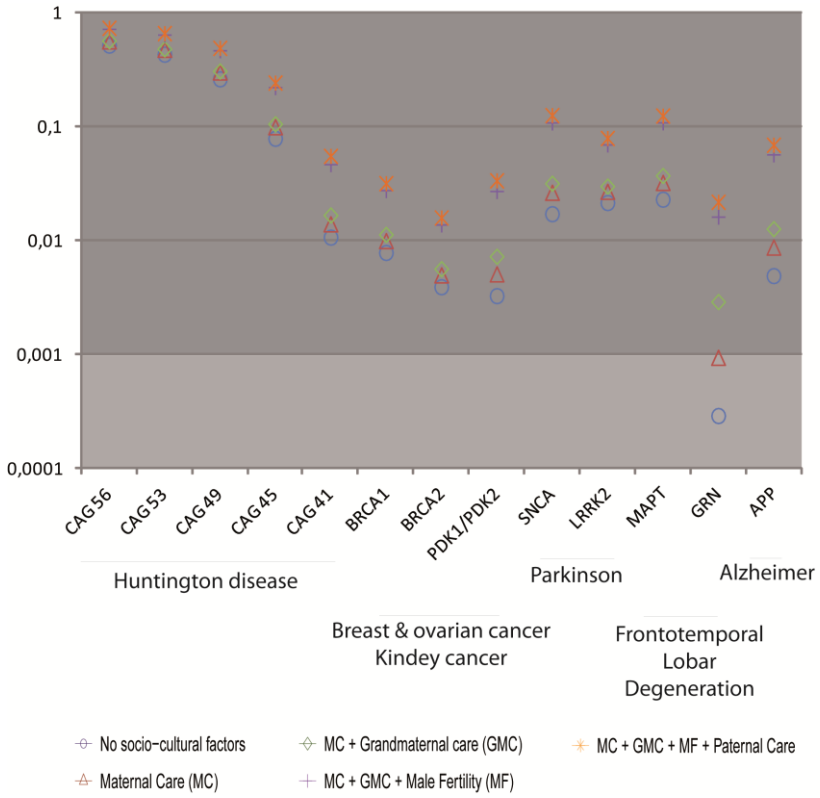


Figure 6 – Coefficient of selection *s* for mutation in specific genes. *MAO* and *FAO* for Huntington diseases comes from (26). Other *MAO* and *FAO* from Table 1 in SM2.

For most diseases *s* never falls under the mild selection threshold at $s=10^{-4}$ from Pritchard (5) and the slight selection at $s=10^{-6}$ from Reich (3), due to variance in disease onset and whatever the socio-cultural scenario considered, and in the case of survival taken for an hunter-gatherer population. As a consequence we think that the CDCV is largely wrong. Selection do occurs on SALOD and its level is as key than Ne or mutation rate to understand the allelic spectrum of late-onset diseases.

Importance of evolution of aging in humans – Selection gradient with age is the founding ground of the two major evolutionary hypotheses of ageing: antagonistic pleiotropy and mutation accumulation. But nobody has ever seen a selection gradient estimated from empirical data. Rather, it is conceptual tools, mainly calculated from life tables, considering theoretical allele with an infinitesimal deleterious effect at a specific age. We show here that the steepness of selection gradients is likely massively overestimated in humans but likely in most long lived species. This is because, from the allele perspective, selection lies not only into a one-space age distribution but into a multiple axes epidemiological-space encompassing mean and variance at onset as well as cumulative penetrance. From the disease perspective, diseases specific selection gradient should moreover encompass the multiplicity of loci from potentially different genes having different epidemiological output. At the species level, ageing encompass the increased incidence with age of many diseases aggregated into one unique age-specific mortality trajectories. Inferring senescence pattern from such an aggregated mortality trajectory sweeps this complexity under the carpet and one has yet to proof the predictive power of such a conceptual tool.

This is even truer in humans because selection is not only a matter of number of children produced in a live time but that of those taken care for and educated over a protracted infancy and childhood by several kin. By flattening the “true” allele-specific selection gradient all these socio-cultural factors tends to make accumulation of deleterious mutation and antagonistic pleiotropy less efficient in our species promoting extended lifespan. This emphasizes the fact that evolution of life-history and sociality are intermingled in our species and that genomic, genetic, epidemiology, demography and anthropology are all required to make sense of the allelic spectrum of late-onset diseases.

Methods

We aim at calculating the selection coefficient *s* for a SALOD for (i) diseases onset characterized by a distribution of onset probability, in (ii) a two-sex model were men can reproduce at old ages, and where (iii) child survival depends on maternal, paternal and grandmaternal care. To do this we need to calculate the reproductive value of men and women carrying or not the SALOD; carriers exhibiting an excess of mortality due to allele-specific disease morbidity. It must be stressed that any incapacitating disease fully compromising reproductive success will have an identical effect on *s* than a lethal disease and the following model applies to both.

1. Population genetics frameworks

Assuming that the mutation is autosomal, dominant and rare (Box 2), the selection coefficient associated with the mutation is given by $s = 1 - W^C / W^{NC}$, where reproductive values of respectively non-carriers and carriers W^{NC} and W^C are defined as the mean number of children produced during their lifetime. If the mutation is rare enough, the proportion of carriers relative to that of non-carriers is negligible at all ages and the population's age-specific mortality hazard (i.e., the instantaneous rate of mortality) is that of non-carriers $h^{NC}(t)$ (Pavard and Metcalf, 2007). Assuming that carriers' age-specific morbidity is independent from any other causes of death existing into the population (i.e. the corresponding mortality hazards are independent), then morbidity is an additive excess of mortality of carriers relative to non-carriers such that $h^C(t) = h^{NC}(t) + h^d(t)$, where $h^d(t)$ is the mortality hazard resulting from disease's morbidity. Assuming that mutation has no antagonistic pleiotropic effect (Box2), then reproductive success of non-carriers and carriers are the sum overall ages x of the product between the survival probability $l(x)$ for non carriers and $l(x)l^d(x)$ for carriers; with $l(x) = e^{-\int_0^x h^{NC}(t)dt}$ and $l^d(x) = e^{-\int_0^x h^d(t)dt}$) and the fertility $F(x)$ at these ages:

$$W^{NC} = \int_x l(x)F(x)dx \qquad (0.1)$$

And

$$W^C = \int_x l(x)l^d(x)F(x)dx \qquad (0.2)$$

2. Reproductive success of non-carriers W^{NC}

We aim at incorporating maternal, grand maternal and paternal cares through the extent to which a child survival until maturity is compromised by the death of his mother, and/or grandmother and/or father. To do this, let us first write W^{NC} as a function of the population's mean children survival $\overline{S}$ at maturity α such that equation (1.1) becomes:

$$W^{NC} = \int_{x=\alpha}^{\varpi} L(x)F(x)\overline{S}dx \qquad (0.3)$$

where $L(x) = l(x)/\overline{S}$ and $L(x<\alpha) = 1$.

In what follows, maternal, grand maternal and paternal care are integrated via the function $S(y_1, y_2, y_3)$ which gives the survival probability at age α of a child who has lost its mother, grandmother and father at ages y_1, y_2 and y_3 respectively (see SM4). A child with y_1, y_2 and y_3 larger than α is fully cared for and exhibit a maximal chance of surviving until maturity. By contrast, a child with any one of y_1, y_2 or y_3 lower than α has lost one of its parents before reaching maturity and exhibits a compromised survival.

In this setting, survival varies among children according to their mother's, grandmother's and father's survival probabilities within the years following the child's birth. As a consequence, a child survival is a function the age of its mother, grandmother and father at its childbirth (denoted respectively x_1, x_2 or x_3): the older its (grand)parents at its childbirth, the higher their risk of dying (or being already dead for grandmothers) within the following years (a demographic relationship first explored by Goodman 1974), and the lower the child survival until maturity. This leads to a negative covariance between parents' ages at the birth of a child and this later survival (Pavard et al. 2007a). As a consequence, the population's mean children survival $\overline{S}$ will depends on the population distribution of x_1, x_2 or x_3. In an infinite and constant environment these distributions are stable and depend on the male and female age-structure of the population. Distribution of x_1 and x_2 are independent: a mother's age at the birth of a child does not depend on the age of her own mother at her birth. Distributions of mother's age x_1 and father's age x_3 at child's birth are not independent when matrimony is age-structured.

The aim of the following derivations is to calculate $\overline{S}$ as a function of the survival space $S(y_1, y_2, y_3)$, knowing adult survival and fertility for female ($L(x)$ and $F(x)$), and males ($L^m(x)$ and $F^m(x)$), and holding that $\overline{S}$ determines and is determined by males and females age-structures. To do this, we express the probability that $[y_1, y_2, y_3]$ occurs as a function of maternal age at childbirth x_1.These derivations are detailed in SM1 Child and are summarized below.

Let us define $S(x_1)$ as the probability for the daughter to survive until maturity as a function of the female's age at child's birth. The probability $S(x_1)$ can be calculated as the sum over all possible child's age at mother's, grandmother's and father's death (y_1, y_2 and y_3 respectively) of the product between (**a**) the probability for a child born to a mother at age x_1 to lose his/her mother at age y_1, his/her maternal grandmother at age y_2 and his/her father at age y_3 and (**b**) the survival $S(y_1, y_2, y_3)$:

(0.4) $$S(x_1) = \iiint p(y_1 \mid x_1)\, p(y_2 \mid x_1)\, p(y_3 \mid x_1) S(y_1, y_2, y_3)\, dy_1 dy_2 dy_3$$

Equation (1.3) can then be rewritten as:

(0.5) $$W^{NC} = \int_{x_1=\alpha}^{\varpi} L(x_1) F(x_1) S(x_1) dx_1$$

Population mean children survival depends on population parental age-structure and can be calculated as follows,

(0.6) $$\overline{S} = \frac{\int_{x_1=\alpha}^{\varpi} L(x_1) F(x_1) S(x_1) dx_1}{\int_{x_1=\alpha}^{\varpi} L(x_1) F(x_1) dx_1}$$

It must be stressed that $\overline{S}$ could also have been calculated by expressing equation (1.4) as a function of father's age at childbirth x_3 and the W^{NC} of men (equaled to that of women) calculated.

3. Reproductive success of carriers W^C

The SALOD (rare, autosomal, dominant, and increasing the risk of disease onset in both sex equally) can be carried out by a man or by a woman with a probability equals to the sex-ratio (hereafter considered equaled to ½). If the mutation is carried by a man, and because the mutation is rare, we can consider that his wife(s) and his mother(s) in law are not carrying the mutation. If the mutation is carried by a woman, her husband is also free of mutation. In this case however, the woman may have inherited her mutation from her mother or her father with a probability equaled to ½. In mean the reproductive success of carriers can be written as:

$$W^C = \frac{1}{2}W\left[\text{man}^C, \text{wife(s)}^{NC}, \text{wife(s)' mother}^{NC}\right] + \frac{1}{2}\left(\begin{array}{l}\frac{1}{2}W\left[\text{woman}^C, \text{husband(s)}^{NC}, \text{woman's mother}^{NC}\right] \\ +\frac{1}{2}W\left[\text{woman}^C, \text{husband(s)}^{NC}, \text{woman's mother}^{C}\right]\end{array}\right) \quad (0.7)$$

where the superscripts ^{NC} and ^{C} define whether or not the individual carries or not the SALOD; individuals carrying the SALOD being at risk of developing the diseases and exhibits an excess of mortality do to disease morbidity $L^d(x)$. Equation (1.7) makes it also easy to calculate *s* in the case of other genetic compartments or for sex-specific disease (Box 2).

4. Parameterization

Age at onset distribution, age-specific fertility and effect of maternal, grandmaternal and paternal care on child survival are respectively detailed in Fig. 1 and SMX, Fig. 2 and SMX and Fig. 3 and SMX. All results are provided for a population mortality corresponding to that of an averaged hunter-gatherer population fitted with a Siler mortality model (36). In this case $l(15) = 0.57$ and $l(45) = 0.36$ which means that $L(45) = l(45)/l(15) = 0.63$: 63% of survivors at age 15 will survive at age 45. While life expectancy at birth is 31 years, the remaining life expectancy at age 45 is $e_{45} = 20.7$ years. Parameter A^m of the Brass polynomial is fine tuned to 4.97e-05 according to this mortality. Mean number of children equal to 2.29 (for both male and female) while TFR is respectively 4.79 for females and 5.66 for TFR males. The proportion of children born to father older than 50 years old is 14% (around 1 child over 5.66) which is about 11% of mean males reproductive success. Results for larger survival pre-industrial population, i.e., 1765 Sweden, is provided in Fig.Sx **[STILL TO DO]**.

5. Scaling selection to genetic drift in humans

In population genetics, selection coefficients are scaled by a factor $4Ne$, where *Ne* is the effective population size, to account for the stochastic effect of genetic drift in determining the fate of an allele. This is especially crucial in the case of alleles under weak selection or close to neutrality. In humans, population sizes have to refer to a carefully defined population, over a given time-scale and with a given demographic history, for which the allelic spectrum is observed. For example, (3-5) aimed a explaining the allelic spectrum defined at the level of the humanity since its origin and considered that $N_e = 10,000$. By contrast, if one wants to explain the allelic spectrum of the descendants of the French Canadian settlers, one would consider a founder population having undergone a large population growth during 300 years (i.e., 1000 from (44)). Finally, human populations' Ne rarely falls below 100, which is the range found for very isolated insular populations (45). To discuss our results let us, let us consider them forward, focusing on one given allele. The probability of fixation approaches the neutral value at $4N_e s = 1$ (i.e., the process of genetic drift and selection are equal, reviewed in (46)). Here we consider that selection is weak when $1 < 4N_e s < 10$ and strong when $4N_e s > 10$.

References

1. Schork NJ, Murray SS, Frazer KA, & Topol EJ (2009) Common vs. rare allele hypotheses for complex diseases. *Curr Opin Genet Dev* 19(3):212-219 .
2. Gibson G (2012) Rare and common variants: twenty arguments. *Nat Rev Genet* 13(2):135-145 .
3. Reich DE & Lander ES (2001) On the allelic spectrum of human disease. *Trends in Genetics* 17(9):502-510 .
4. Pritchard JK (2001) Are rare variants responsible for susceptibility to complex diseases? *Am J Hum Genet* 69(1):124-137 .
5. Pritchard JK & Cox NJ (2002) The allelic architecture of human disease genes: common disease-common variant...or not? *Human Molecular Genetics* 11(20):2417-2423 .
6. Wright AF & Hastie ND (2001) Complex genetic diseases: controversy over the Croesus code. *Genome Biology* 2(8):COMMENT2007 .
7. Wright A, Charlesworth B, Rudan I, Carothers A, & Campbell H (2003) A polygenic basis for late-onset disease. *Trends Genet* 19(2):97-106 (in English).
8. Karami F & Mehdipour P (2013) A Comprehensive Focus on Global Spectrum of and Mutations in Breast Cancer. *Biomed Res Int* 2013:928562 .
9. Park J-H*, et al.* (2011) Distribution of allele frequencies and effect sizes and their interrelationships for common genetic susceptibility variants. *Proceedings of the National Academy of Sciences* 108(44):18026-18031.
10. Rodríguez JA*, et al.* (2017) Antagonistic pleiotropy and mutation accumulation influence human senescence and disease. 1:0055.
11. Harris TJR & McCormick F (2010) The molecular pathology of cancer. *Nat Rev Clin Oncol* 7(5):251-265.
12. Blekhman R*, et al.* (2008) Natural selection on genes that underlie human disease susceptibility. *Curr Biol* 18(12):883-889 .
13. Di Rienzo A & Hudson RR (2005) An evolutionary framework for common diseases: the ancestral-susceptibility model. *Trends in Genetics* 21(11):596-601 .
14. Di Rienzo A (2006) Population genetics models of common diseases. *Curr Opin Genet Dev* 16(6):630-636 .
15. Lander ES (1996) The new genomics: global views of biology. *Science* 274(5287):536-539 .
16. Medawar PB (1952) An Unsolved Problem of Biology. *Uniqueness of the Individual*, (H.K. Lewis.), pp 44-70.
17. Williams GC (1957) Pleiotropy, Natural-Selection, and the Evolution of Senescence. *Evolution* 11(4):398-411 (in English).
18. Pritchard JK (2001) Are rare variants responsible for susceptibility to complex diseases? *Am J Hum Genet* 69(1):124-137 .
19. Pavard S & Metcalf CJE (2007) Negative Selection on BRCA1 Susceptibility Alleles Sheds Light on the Population Genetics of Late-Onset Diseases and Aging Theory. *PLoS ONE* 2(11):e1206.
20. Cornec-Le Gall E*, et al.* (2013) Type of PKD1 mutation influences renal outcome in ADPKD. *J Am Soc Nephrol* 24(6):1006-1013 .
21. Marchani EE*, et al.* (2010) Evidence for three loci modifying age-at-onset of Alzheimer's disease in early-onset PSEN2 families. *Am J Med Genet B Neuropsychiatr Genet* 153B(5):1031-1041 .
22. Couch FJ*, et al.* (2013) Genome-wide association study in BRCA1 mutation carriers identifies novel loci associated with breast and ovarian cancer risk. *PLoS Genet* 9(3):e1003212 .
23. Berger F*, et al.* (2013) The impact of single-nucleotide polymorphisms (SNPs) in OGG1 and XPC on the age at onset of Huntington disease. *Mutat Res* 755(2):115-119 .
24. Lindner AB & Demarez A (2009) Protein aggregation as a paradigm of aging. *Biochim Biophys Acta* 1790(10):980-996 .
25. Moorad JA & Promislow DEL (2008) A Theory of Age-Dependent Mutation and Senescence. *Genetics* 179(4):2061-2073.
26. Langbehn DR, Brinkman RR, Falush D, Paulsen JS, & Hayden MR (2004) A new model for prediction of the age of onset and penetrance for Huntington's disease based on CAG length. *Clin Genet* 65(4):267-277 .
27. Tuljapurkar SD, Puleston CO, & Gurven MD (2007) Why Men Matter: Mating Patterns Drive Evolution of Human Lifespan. *PLoS ONE* 2(8):e785.
28. Muller MN, Thompson ME, & Wrangham RW (2006) Male chimpanzees prefer mating with old females. *Curr. Biol.* 16(22):2234-2238 (in English).

29. Fenner JN (2005) Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *American Journal of Physical Anthropology* 128(2):415-423 .
30. Marlowe F (2004) Mate preferences among Hadza hunter-gatherers. *Hum Nat* 15(4):365-376 (in English).
31. Leonetti DL, Nath DC, & Hemam NS (2007) In-law Conflict : Women's Reproductive Lives and the Roles of Their Mothers and Husbands among the Matrilineal Khasi. (Translated from 0011-3204) *Anglais* 48(6):861-890 (in 0011-3204).
32. Sear R (2008) Kin and Child Survival in Rural Malawi. *Hum Nat* 19(3):277-293.
33. Ackerman C (1963) Affiliations: Structural Determinants of Differential Divorce Rates. *American Journal of Sociology* 69(1):13-20.
34. Hill K & Hurtado AM (1996) *Ache life history: the ecology and demography of a foraging people* (De Gruyter, New York, USA).
35. Walker RS, Hill KR, Flinn MV, & Ellsworth RM (2011) Evolutionary History of Hunter-Gatherer Marriage Practices. *PLoS ONE* 6(4):e19066.
36. Gurven M & Kaplan H (2007) Longevity among hunter-gatherers: A cross-cultural examination. *Population and Development Review* 33(2):321-365 (in English).
37. Sear R & Mace R (2008) Who keeps children alive? A review of the effects of kin on child survival. *Evolution and Human Behavior* 29:1-18.
38. Bentley G & Mace R (2009) *Substitute parents: biological and social perspectives on alloparenting in human societies* (Berghahn Books, Oxford).
39. Pavard S, Sibert A, & Heyer E (2007) The effect of maternal care on child survival: A demographic, genetic, and evolutionary perspective. *Evolution* 61(5):1153-1161 (in English).
40. Smith KR, Hanson HA, Mineau GP, & Buys SS (2012) Effects of BRCA1 and BRCA2 mutations on female fertility. *Proceedings of the Royal Society B: Biological Sciences* 279(1732):1389-1395.
41. Couch FJ*, et al.* (2013) Genome-Wide Association Study in BRCA1 Mutation Carriers Identifies Novel Loci Associated with Breast and Ovarian Cancer Risk. *PLoS Genet* 9(3):e1003212.
42. Cyrus CC & Lee RD (2013) On the evolution of intergenerational division of labor, menopause and transfers among adults and offspring. *Journal of Theoretical Biology* 332:171-180 .
43. da Silva J (2012) BRCA1/2 mutations, fertility and the grandmother effect. *Proceedings of the Royal Society B: Biological Sciences* 279(1740):2926-2929.
44. Austerlitz F & Heyer E (1998) Social transmission of reproductive behavior increases frequency of inherited disorders in a young-expanding population. *Proceedings of the National Academy of Sciences of the United States of America* 95(25):15140-15144 (in English).
45. Morton NE (1982) *Outline of Genetic Epidemiology* (S. Karger., Basel, Switzerland).
46. Hamilton MB (2009) *Population Genetics* (Wiley-Blackwell, Chichester, UK).

SUPLEMENTARY MATERIAL 1 - Child survival as a function of **mother**'s age at childbirth

1. Objectives and notation

Maternal care is easily incorporated into a non-overlapping population genetics framework [3] because it involves only one generation (the scale upon which is built population genetics model) and because population age-structure is not required. As described into the main text, in a two-sexes/two-generations model, the population's age-structure is required to estimate the frequencies of the mother-grandmother's age or husband-wife's age at th birth of a child. To present this model, we first denote:

- y_1, y_2 and y_3 the daughter age at the death of her mother, grandmother and father respectively.
- x_1, x_2 and x_3 the ages of the mother, the grandmother and the father at the birth of their child (ie the daughter for mother and father, the mother for the maternal grandmother).
- α is the age at maturity (the age at which maternal, grandmaternal and paternal care is not anymore needed for daughter survival). It is also the age at first observed reproduction for female.
- β is the age at last reproduction for female.
- β^m is the age at last reproduction for men.
- ω is the maximum lifespan.
- $S(x)$ is the daughter survival probability until age α as a function of a parent's age at childbirth x.
- $S(y_1, y_2, y_3)$ is the daughter survival probability until age α as a function of her age at mother's, grandmother's and father's death.
- $L(x)$ and $L^m(x)$ are the survival of indivividual surviving at age α, in the following adult years (ie for $x \geq \alpha$ and with $L(\alpha)$and $L^m(\alpha)$ equalled to 1), for women and men.
- $F(x)$ and $F^m(x)$ are the fertility at age x of women and men respecively.
- λ and λ^m are the finite rates of women and men population growth.

In order to calculate the female non-carrier reproductive success W^{NC}[Female] (see SM4, equation (6)), we need to express $S(x_1)$ as a function of $S(y_1, y_2, y_3)$. The probability $S(x_1)$ can then be defined as the sum over all possible child's age at mother's, grandmother's and father's death y_1, y_2 and y_3, of the product between (i) the probability for a child born to a mother at age x_1 to lose his/her mother at age y_1, his/her maternal grandmother at age y_2 and his/her father at age y_3 and (ii) the survival $S(y_1, y_2, y_3)$. The probability $S(x_1)$ can then be written:

$$S(x_1) = \int_{y_1} \int_{y_2} \int_{y_3} p(y_1|x_1)\, p(y_2|x_1)\, p(y_3|x_1)\, S(y_1, y_2, y_3)\, dy_3 dy_2 dy_1 \quad (1)$$

2. Probability for a daughter to lose her mother at age y_1

The probability $p(y_1|x_1)$ is the probability for a child born to a mother at age x_1 to lose its mother at age y_1. Because the mother is obviously alive at the child's birth this probability is (see [2, 3] for more details):

$$p(y_1|x_1) = \frac{L(x_1+y_1)}{L(x_1)} h(x_1+y_1) \tag{2}$$

The child's age at mother death y_1 is defined between 0 (the mother dies at child's delivery) and $\varpi - \alpha$ (the mother is at the youngest possible age α at childbirth and dies at the maximum lifespan ϖ). For any given age x_1, we have $\min(y_1|x_1) = 0$ and $\max(y_1|x_1) = \varpi - x_1$. The sum over all possible age y_1 of the probability $p(y_1|x_1)$ equals one such that:

$$\int_{y_1} p(y_1|x_1)\, dy_1 = \int_0^{\varpi - x_1} p(y_1|x_1)\, dy_1 = 1 \tag{3}$$

3. Probability for a daughter to lose her grandmother at age y_2

The probability $p(y_2|x_1)$ is the probability for a child born from a mother at age x_1 to lose its maternal grandmother at age y_2. This probability depends therefore of the grandmother's age x_2 at the birth of the mother. This age is independent of the mother's age at child birth x_1 and can range from α to β. The probability $p(y_2|x_1)$ can then be written as follows:

$$p(y_2|x_1) = \int_\alpha^\beta p(x_2)\, p(y_2|x_1, x_2)\, dx_2 \tag{4}$$

- The probability $p(x_2)$ is the probability that the mother is born to the grandmother at age x_2. When the age-structure of the population is stable, the probability $p(x_2)$ equals the proportion of children born to a mother at age x_2 such that:

$$p(x_2) = (\lambda)^{-x_2} L(x_2) F(x_2) \overline{S} \tag{5}$$

  with the sum $\int_\alpha^\beta (\lambda)^{-x_2} L(x_2) F(x_2) \overline{S} dx_2 = 1$ being the Euler-Lotka equation.

- The probability $p(y_2|x_1, x_2)$ is the probability for a child born to a mother at age x_1, herself born to a mother at age x_2, to lose its grandmother at age y_2.The grandmother being obviously alive at age x_2 when she gave birth to her daughter, this probability is:

$$p(y_2|x_1, x_2) = \frac{L(x_2+x_1+y_1)}{L(x_2)} h(x_2+x_1+y_1) \tag{6}$$

For any given x_1 and x_2, we have $\min(y_2|x_1, x_2) = -x_1$ (the grandmother died at the birth of the mother, thus x_1 years before the child's birth) and $\max(y_2|x_1, x_2) = \omega - x_1 - x_2$ (the grandmother died at age ϖ). The sum over all possible y_2 of the probabilities $p(y_2|x_1)$ is therefore:

$$\int_{y_2} p(y_2|x_1)\, dy_2 = \int_\alpha^\beta p(x_2) \int_{-x_1}^{\varpi - x_1 - x_2} p(y_2|x_1, x_2)\, dy_2 dx_2 \tag{7}$$

Setting $p(y_2|x_1, x_2) = 0$ for impossible occurences such that $y_2 < -x_1$ and $y_2 > \omega - x_1 - x_2$, equation (7) can be written as:

$$\int_{y_2} p(y_2|x_1)\, dy_2 = \int_{-x_1}^{\varpi - x_1} \int_\alpha^\beta p(x_2)\, p(y_2|x_1, x_2)\, dx_2 dy_2 \tag{8}$$

4. Probability for a daughter to lose her father at age y_3

The probability $p(y_3|x_1)$ is the probability for a child born from a mother at age x_1 to lose its father at age y_3. This probability depends therefore of the father's age at the birth of the child x_3. For simplicity we assume that fathers are always of the same age or older than the mother without further matrimonial structuring (structuring matrimony would indeed imply extending the model for incorporating widowing and divorced probability, remarriage probability and polygamy). The probability $p(y_3|x_1)$ can then be written :

$$p(y_3|x_1) = \int_{\alpha}^{\beta} p(x_3|x_1)\, p(y_2|x_3)\, dx_3 \tag{9}$$

The probability $p(x_3|x_1)$ is the probability for the father of being x_3 years old at daughter's birth, knowing that his wife is x_1 years old at that time and was younger than him. When the stable males distribution is reached the distribution of fathers' age x_3 is given by $(\lambda^m)^{-x_3} L^m(x_3) F^m(x_3) \overline{S}$, with $\int_{\alpha}^{\beta^m} (\lambda^m)^{-x_3} L^m(x_3) F^m(x_3) \overline{S} dx_3 = 1$ being the Euler-Lotka equation. Knowing that $x_3 \geq x_1$, the probability $p(x_3|x_1)$ can be written:

$$p(x_3|x_1) = \frac{(\lambda^m)^{-x_3} l^m(x_3) F^m(x_3) \overline{S}}{\int_{x_1}^{\beta^m} (\lambda^m)^{-x_3} l^m(x_3) F^m(x_3) \overline{S} dx_3} \tag{10}$$

The probability $p(y_3|x_3)$ is the probability for a daughter born to a father at age x_3 to lose her father at age y_3. Assuming that the father is alive at child's birth, this probability is:

$$p(y_3|x_3) = \frac{L(x_3 + y_3)}{L(x_3)} h(x_3 + y_3) \tag{11}$$

For any given x_3 we have $\min(y_3|x_3) = 0$ and $\max(y_3|x_3) = \varpi - x_3$. The sum over all possible y_3 of the probabilities $p(y_3|x_1)$ is therefore:

$$\int_{y_3} p(y_3|x_1) = \int_{x_1}^{\beta^m} p(x_3|x_1) \int_0^{\varpi - x_3} p(y_3|x_3)\, dy_3 dx_3 \tag{12}$$

Setting $p(y_3|x_3) = 0$ for impossible occurrence such that $y_3 > \varpi - x_3$, equation (12) can be written:

$$\int_{y_3} p(y_3|x_1) = \int_0^{\varpi - x_3} \int_{x_1}^{\beta^m} p(x_3|x_1)\, p(y_3|x_3)\, dx_3 dy_3 \tag{13}$$

5. Solving the Euler-Lotka Equation and calculation of $S(x_1)$

To calculate $S(x_1)$ from equation (1) we first need to find the two finite growth rates of the female and the male populations λ and λ^m in order to calculate the age-structure of females and males in equation (5) and (10). To do this, we solve the Euler-Lotka equation:

$$1 = \int_{\alpha}^{\omega} (\lambda)^{-x_1} L(x_1) F(x_1) S(x_1)\, dx_1, \tag{14}$$

replacing $S(x_1)$ by its expression from equation (1). To solve equation (14) we use the approximation $\log \lambda \approx \log R_0/T$, discussed in [1], in order to estimate λ from λ^m, with R_0 being the mean number of children produced by individuals druing a lifetime and T being the generation time. Indeed, because males and females produce in mean the same number of children over a lifetime, males' and females' R_0 are identical. Therefore $\log(\lambda^m) = \log(\lambda)(T/T^m)$. It must be stressed that T depends only on adult survival and fecundity and is independent of juvenile survival; this relationship holds therefore for any level of maternal, grand maternal and paternal care implemented into the model.

Finite population rate of increase λ can therefore be estimate by iteration: a given λ corresponds to a given approximated λ^m, which together allow the calculation of a given distribution $S(x_1)$. We iterate λ until , λ, λ^m and $S(x_1)$ satisfies Euler-Lotka equation (14).

References

[1] H. Caswell. *Matrix population models*. Sinauer Associates, Inc., Sunderland, MA., 2nd edition, 2001.

[2] Leo A. Goodman, Nathan Keyfitz, and Thomas W. Pullum. Family formation and the frequency of various kinship relationships. *Theoretical Population Biology*, 5(1):1–27, 1974.

[3] S. Pavard, A. Sibert, and E. Heyer. The effect of maternal care on child survival: A demographic, genetic, and evolutionary perspective. *Evolution*, 61(5):1153–1161, 2007.

SUPLEMENTARY MATERIALS 2 - Age at Onset Distribution

1 Data

Mutation-specific distribution of age at disease onset is unknown for most of familial form of late-onset diseases. Table SM2 gathers published data (ploted in Figure SM2.1) for mean and minimum age at onset (hereafter respectively denoted *MAO* and *FAO*) for a range of couple gene-diseases. Of course, these are crude data (i) differing in sample size between studies and therefore in the probability of detecting early onset, (ii) aggregated at the level of the gene across mutations, (iii) not accounting for differences in populations survival between studies, (iv) differing in the definition of diagnosis, (v) not accounting for epistatic relationships modifying age at onset patterns (as for example in [7] in the case of *BRCA1*). It gives however a crude idea of the extent to which individuals sharing mutations on susceptibility genes may developp disease at very different ages due to the physiopathology of the disease, epistatic effects or gene-environment interactions. For example, the two-hit Knudson rule [11]applies in the case of germinal mutations in *BRCA1* and *BRCA2* susceptibility gene to breast and ovarian. A first germinal mutation compromises one copy of the gene, but this tumor suppressor gene is fully inactivated only when a second somatic mutation compromises the second copy of the gene. Variance in disease onset arises therefore, at least in part, from the random occurence of the second mutation.

Figure SM2.1 shows that *MAO* varies roughly between 30 and 65 and *FAO* between 25 and 55. A linear regression weighted with sample sizes show a significant linear relationship between MAO and FAO (p=1.27e-05). If further confirmed, this indicate that variance in disease onset, as much as the mean, should be considered as a age-related phenotype of late-onset diseases.

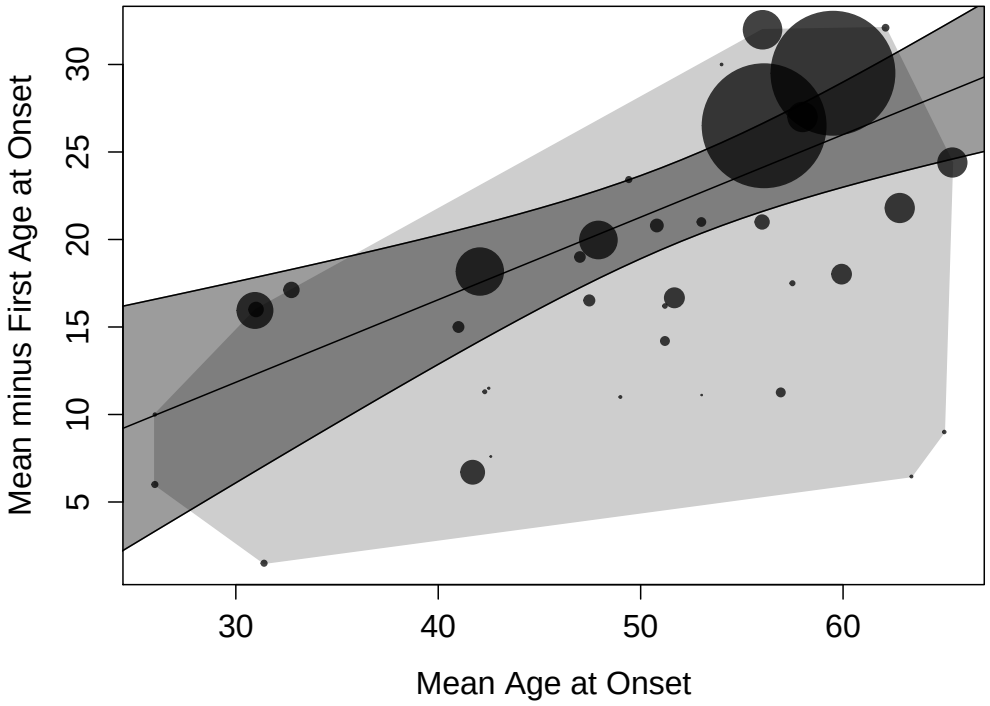


Figure 1: Difference *MAO-FAO* as a function of *MAO* in the case of the couple gene-diseases listed in Table SM2 (to the exeption of *PARK2* and *PINK1* linked to juvenile cases of parlinson diseases). Size of the points is proportional to the coresponding study sample size. Light grey area is the convex hull of the points. Dark grey area is the 99% confidence interval of predicted linear relationship between *MAO* and *MAO-FAO* (β=0.45, p=1.27e-05)

2 Model and Parametrization

The aim here is not to estimate parameters of disease onset distibution from data, but to explore a large parameter space encompassing most of its epidemiological variabillity amongst late-onset diseases. To do this, we model the distribution of the age at disease onset for SALOD carriers in a population absent of any other causes of death than that resulting of the disease. This can be described by three equivalent functions: the distribution of age at onset $f(x)$, the cumulative risk $F(x)$,of developing the disease before age x, $p(x \leq X)$ and the onset hazard $h(x) = f(x)/(1 - F(x))$. This distribution can be defined by three parameters: its mean age at onset, its variance and the cumulative penetrance at old ages (hereafter considered at age 100 years old). This last parameter fit the fact that, for some disease with incomplete penetrance, SALOD carriers may not developp the disease.

Although complex models have been proposed to predict age-specific risks (e.g., [2] for mutations in *BRCA1* and *BRCA2*), logistic distribution has been proved a good fit for Hungtington disease [12] and provides a flexible distribution that would encompass differences in distribution between diseases as well as modification of distribution due to epistatic relationship within diseases (as for example the logistic-shaped epistatic changes demonstrated for *BRCA1* mutation in [7]). Assuming a logistic distribution of mean equals to *MAO* and variance $\frac{1}{3}s^2\pi^2$, the cumulative risk function $F(x)$ is:

$$F(x) = \frac{1}{1 + e^{-\frac{x - MAO}{s}}}$$

Because we find more intuitive to describe such distribution in terms of mean and first age at onset (*MAO* and *FAO*) rather than in termes of mean and variance, let us define the disease first age at onset *FAO* as the age at which the proportion ϕ of SALOD carriers have developped the disease such that $F(\text{FAO}) = \phi$. Then we can calculate the coresponding parameter s sur that:

$$s = -\frac{FAO - MAO}{\log\left(\frac{1}{\phi} - 1\right)}$$

Setting $F(x < FAO) = 0$, disease occurs at a rate ϕ at age FAO. Using this setting, age at onset distributions can be modelled accoring to a MAO- and FAO- space despicted in Figure 1 of the main texte.

Table 1: Data on mean and minimum age at onset for different pathologies and associated genes

Phenotype	Gene	Age at onset		n	Reference	Note
		Mean	Min			
	APP	61.2	35	19	[14]	Families with at least one early onset (<60 yrs) in three generation
Alzheimer Dementia	APP	52	35	90	[8]	Data extracted from figure
	APP promoter	65	56	14	[10]	
	PSEN1	41.7	35	108	[14]	
	PSEN1	42.6	35	6	[9]	
	PSEN1	NA	24	120	[5]	Families with at least one early onset (<61 yrs) in three generation
	PSEN1	42	24	217	[8]	Data extracted from figure
	PSEN2	57	46	39	[8]	Data extracted from figure
Frontotemporal Lobar Degeneration	MAPT	48	28	172	[8]	Data extracted from figure
	GRN	60	42	89	[8]	Data extracted from figure
	VCP	49	26	26	[8]	Data extracted from figure
	CHMP2B	63	57	11	[8]	Data extracted from figure
Parkinson Disease	PARK7	31	30	26	[8]	Data extracted from figure
	PARK2	31	3	164	[8]	Data extracted from figure
	PINK1	33	16	69	[8]	Data extracted from figure
	SNCA	47	31	49	[8]	Data extracted from figure
	LRRK2	56	24	176	[8]	Data extracted from figure
Hungtington (adult onset)	HTT, 40 CGA rep	53	32	38	[16]	Data extracted from figure
	HTT, 40 CGA rep	56	35	64	[4]	
	HTT, 40 CGA rep	58	31	134	[12]	
	HTT, 50 CGA rep	26	20	26	[16]	Data extracted from figure
	HTT, 50 CGA rep	26	16	13	[4]	
	HTT, 50 CGA rep	31	12	65	[12]	
Breast and/or ovarian cancer	BRCA1	42.5	31	8	[1]	Breast cancer only
	BRCA1	51,2	37	39	[15]	Ovarian cancer only
	BRCA1	50.8	30	57	[13]	Ovarian cancer only, onset range taken from 30-40 and 80-90 yrs categories
	BRCA1	41	26	48	[3]	Breast cancer
	BRCA1	49	38	11	[3]	Ovarian cancer
	BRCA2	42	31	16	[1]	Breast cancer only
	BRCA2	57,5	40	21	[15]	Ovarian cancer only
	BRCA2	62.1	30	29	[13]	Ovarian cancer only, onset range taken from 30-40 and 80-90 yrs categories
	BRCA2	47	28	47	[3]	Breast cancer
	BRCA2	54	24	10	[3]	Ovarian cancer
Autosomal Dominant polycystic kidney disease	PKD1	56.1 for males 59.5 for females	30	571	[6]	Age at End Stage Renal Disease, minimum age from personal communication
	PKD2	62.8 for males 65.4 for females	41	133	[6]	Age at End Stage Renal Disease, minimum age from personal communication

References

[1] Anglian-Breast-Cancer-Study-Group. Prevalence and penetrance of brca1 and brca2 mutations in a population-based series of breast cancer cases. anglian breast cancer study group. *Br J Cancer*, 83(10):1301–8, 2000.

[2] A. C. Antoniou, P. D. Pharoah, G. McMullan, N. E. Day, B. A. Ponder, and D. Easton. Evidence for further breast cancer susceptibility genes in addition to brca1 and brca2 in a population-based study. *Genet Epidemiol*, 21(1):1–18, 2001.

[3] R.D. Brandao. *Improving the risk assessment of inherited breast and ovarian cancer: clinical significance of BRCA1/2 variants and risk modifiers.* PhD thesis, 2012.

[4] R. R. Brinkman, M. M. Mezei, J. Theilmann, E. Almqvist, and M. R. Hayden. The likelihood of being affected with huntington disease by a particular age, for a specific cag size. *Am J Hum Genet*, 60(5):1202–10, 1997.

[5] D. Campion, C. Dumanchin, D. Hannequin, B. Dubois, S. Belliard, M. Puel, C. Thomas-Anterion, A. Michon, C. Martin, F. Charbonnier, G. Raux, A. Camuzat, C. Penet, V. Mesnage, M. Martinez, F. Clerget-Darpoux, A. Brice, and T. Frebourg. Early-onset autosomal dominant alzheimer disease: prevalence, genetic heterogeneity, and mutation spectrum. *Am J Hum Genet*, 65(3):664–70, 1999.

[6] E. Cornec-Le Gall, M. P. Audrezet, J. M. Chen, M. Hourmant, M. P. Morin, R. Perrichot, C. Charasse, B. Whebe, E. Renaudineau, P. Jousset, M. P. Guillodo, A. Grall-Jezequel, P. Saliou, C. Ferec, and Y. Le Meur. Type of pkd1 mutation influences renal outcome in adpkd. *J Am Soc Nephrol*, 24(6):1006–13, 2013.

[7] F. J. Couch, X. Wang, L. McGuffog, A. Lee, C. Olswold, K. B. Kuchenbaecker, P. Soucy, Z. Fredericksen, D. Barrowdale, J. Dennis, M. M. Gaudet, E. Dicks, M. Kosel, S. Healey, O. M. Sinilnikova, F. Bacot, D. Vincent, F. B. Hogervorst, S. Peock, D. Stoppa-Lyonnet, A. Jakubowska, P. Radice, R. K. Schmutzler, S. M. Domchek, M. Piedmonte, C. F. Singer, E. Friedman, M. Thomassen, T. V. Hansen, S. L. Neuhausen, C. I. Szabo, I. Blanco, M. H. Greene, B. Y. Karlan, J. Garber, C. M. Phelan, J. N. Weitzel, M. Montagna, E. Olah, I. L. Andrulis, A. K. Godwin, D. Yannoukakos, D. E. Goldgar, T. Caldes, H. Nevanlinna, A. Osorio, M. B. Terry, M. B. Daly, E. J. van Rensburg, U. Hamann, S. J. Ramus, A. E. Toland, M. A. Caligo, O. I. Olopade, N. Tung, K. Claes, M. S. Beattie, M. C. Southey, E. N. Imyanitov, M. Tischkowitz, R. Janavicius, E. M. John, A. Kwong, O. Diez, J. Balmana, R. B. Barkardottir, B. K. Arun, G. Rennert, S. H. Teo, P. A. Ganz, I. Campbell, A. H. van der Hout, C. H. van Deurzen, C. Seynaeve, E. B. Gomez Garcia, F. E. van Leeuwen, H. E. Meijers-Heijboer, J. J. Gille, M. G. Ausems, M. J. Blok, M. J. Ligtenberg, M. A. Rookus, P. Devilee, S. Verhoef, T. A. van Os, J. T. Wijnen, D. Frost, S. Ellis, E. Fineberg, R. Platte, D. G. Evans, L. Izatt, R. A. Eeles, J. Adlard, D. M. Eccles, J. Cook, C. Brewer, F. Douglas, S. Hodgson, P. J. Morrison, et al. Genome-wide association study in brca1 mutation carriers identifies novel loci associated with breast and ovarian cancer risk. *PLoS Genet*, 9(3):e1003212, 2013.

[8] M. Cruts, J. Theuns, and C. Van Broeckhoven. Locus-specific mutation databases for neurodegenerative brain diseases. *Hum Mutat*, 33(9):1340–4, 2012.

[9] L. Gustafson, A. Brun, E. Englund, O. Hagnell, K. Nilsson, M. Stensmyr, A. K. Ohlin, and M. Abrahamson. A 50-year perspective of a family with chromosome-14-linked alzheimer's disease. *Hum Genet*, 102(3):253–7, 1998.

[10] L. Guyant-Marechal, A. Rovelet-Lecrux, L. Goumidi, E. Cousin, D. Hannequin, G. Raux, C. Penet, S. Ricard, S. Mace, P. Amouyel, J. F. Deleuze, T. Frebourg, A. Brice, J. C. Lambert, and D. Campion. Variations in the app gene promoter region and risk of alzheimer disease. *Neurology*, 68(9):684–7, 2007.

[11] Jr. Knudson, A. G. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A*, 68(4):820–3, 1971. Knudson, A G Jr United states Proceedings of the National Academy of Sciences of the United States of America Proc Natl Acad Sci U S A. 1971 Apr;68(4):820-3.

[12] D. R. Langbehn, R. R. Brinkman, D. Falush, J. S. Paulsen, and M. R. Hayden. A new model for prediction of the age of onset and penetrance for huntington's disease based on cag length. *Clin Genet*, 65(4):267–77, 2004.

[13] R. Moslehi, W. Chu, B. Karlan, D. Fishman, H. Risch, A. Fields, D. Smotkin, Y. Ben-David, J. Rosenblatt, D. Russo, P. Schwartz, N. Tung, E. Warner, B. Rosen, J. Friedman, J. S. Brunet, and S. A. Narod. Brca1 and brca2 mutation analysis of 208 ashkenazi jewish women with ovarian cancer. *Am J Hum Genet*, 66(4):1259–72, 2000.

[14] G. Raux, L. Guyant-Marechal, C. Martin, J. Bou, C. Penet, A. Brice, D. Hannequin, T. Frebourg, and D. Campion. Molecular diagnosis of autosomal dominant early onset alzheimer's disease: an update. *Journal of Medical Genetics*, 42(10):793–5, 2005.

[15] Harvey A. Risch, John R. McLaughlin, David E. C. Cole, Barry Rosen, Linda Bradley, Elaine Kwan, Elaine Jack, Danny J. Vesprini, Graciela Kuperstein, John L. A. Abrahamson, Isabel Fan, Betty Wong, and Steven A. Narod. Prevalence and penetrance of germline brca1 and brca2 mutations in a population series of 649 women with ovarian cancer. *The American Journal of Human Genetics*, 68(3):700–710, 2001.

[16] N. S. Wexler, J. Lorimer, J. Porter, F. Gomez, C. Moskowitz, E. Shackell, K. Marder, G. Penchaszadeh, S. A. Roberts, J. Gayan, D. Brocklebank, S. S. Cherny, L. R. Cardon, J. Gray, S. R. Dlouhy, S. Wiktorski, M. E. Hodes, P. M. Conneally, J. B. Penney, J. Gusella, J. H. Cha, M. Irizarry, D. Rosas, S. Hersch, Z. Hollingsworth, M. MacDonald, A. B. Young, J. M. Andresen, D. E. Housman, M. M. De Young, E. Bonilla, T. Stillings, A. Negrette, S. R. Snodgrass, M. D. Martinez-Jaurrieta, M. A. Ramos-Arroyo, J. Bickham, J. S. Ramos, F. Marshall, I. Shoulson, G. J. Rey, A. Feigin, N. Arnheim, A. Acevedo-Cruz, L. Acosta, J. Alvir, K. Fischbeck, L. M. Thompson, A. Young, L. Dure, C. J. O'Brien, J. Paulsen, A. Brickman, D. Krch, S. Peery, P. Hogarth, Jr. Higgins, D. S., and B. Landwehrmeyer. Venezuelan kindreds reveal that genetic and environmental factors modulate huntington's disease age of onset. *Proc Natl Acad Sci U S A*, 101(10):3498–503, 2004.

SUPLEMENTARY MATERIALS 3 - Fertility of men and women

1. Data

Data on both men and women age-specific fertility ratesare rarily avalaible in the case of pre- and not-industrialized populations. Some of these data are represented in Fig.SM3.1 using the same color code than [1].

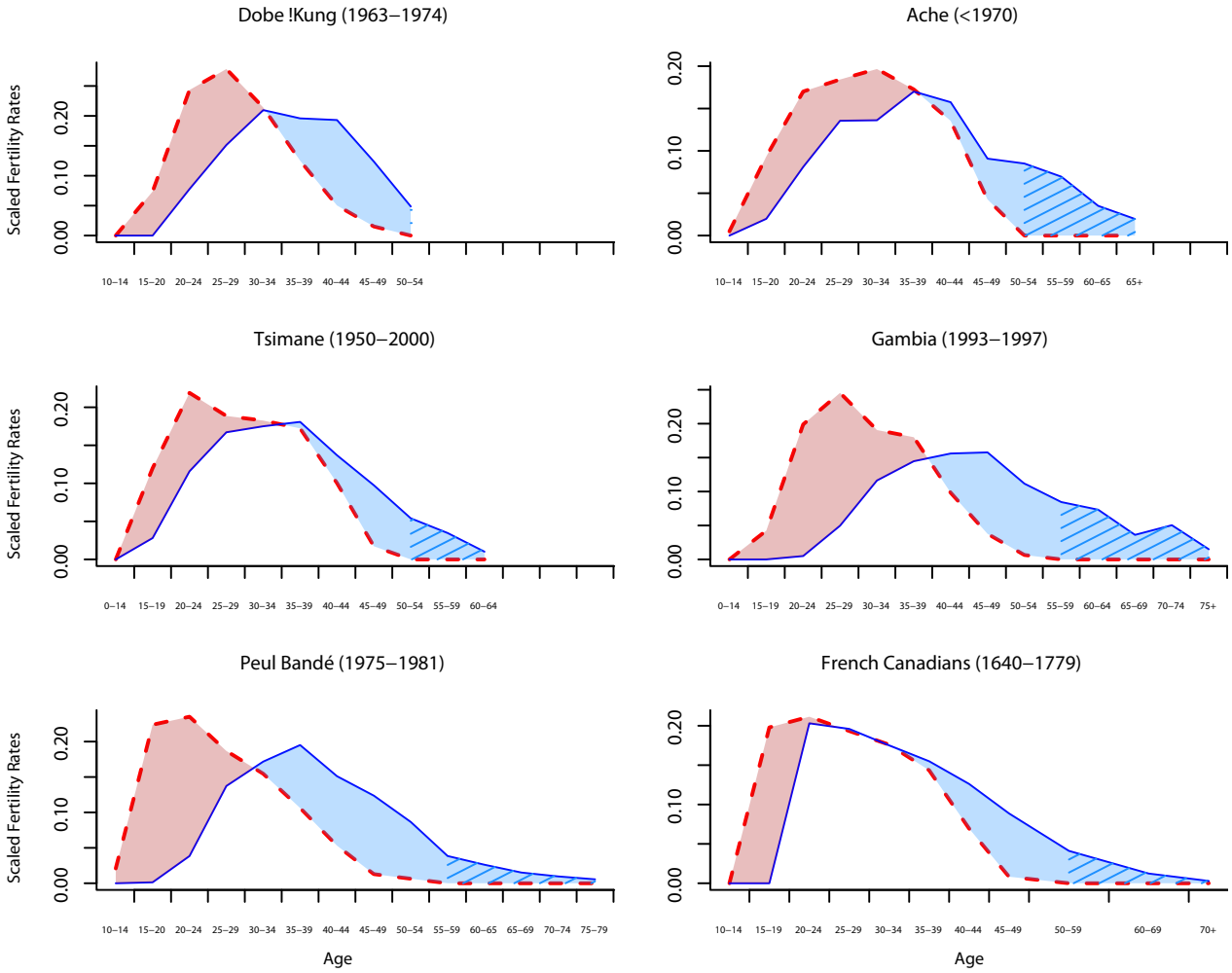


Figure 1: **Scaled age-specific fertility rates for women** (in red) **and men** (in blue) in the case of the Dobe !Kung from [2], Ache from [3], Tsimane from [4] (extracted from figure), population of rural Gambia from [5] (extracted from figure), Peul Bandé from [6] and pre-industrial population of French canadians (1640-1779) from [7] (these estimates includes only married men and women). Red areas show that the fertility of women is larger than that of men at young ages. Blue areas show that fertility of men overcome that of women at older ages. Dashed blue area emphasizes the ages were men carry on reproduction while women cease due to menopause.

For all of these populations, females starts reproducing earlier than men. As a consequence, women fertility is larger than that of men during the first part of their reprodutive life, up to a age comprised between 25 and 40. After this age, men fertility exceeds than that of women and men may continue to reproduce up to older ages. In females, the decline of fertility prior to menopause is explained by decreasing fecundity but also by the increasing proportion of widowed or divorced women who do not remary. Men fertility peak and subsequent decline is shifted to the older ages compared to that of women due to later age at first reproduction. Subsequent decline is however slower in men because of polygyny and because of the possibility for widowed or divorced men to remarry. This allows men to reproduce with younger women at ages where women have cessed reproduction. To convice the reader that this pattern is wide spread accross human populations and therefore relevant for modeling of human evolution, we briefly review the anthropological demographic components that generate these male-female differences in fertility trajectory.

Age at first reproduction - Men are as average older than women at first birth. Holding the fact that marriage encompassed a large range of social and economical aspects [8], it is nevertheless common to all culture [9], and data on age at first marriage gives a good idea the extent to which this pattern is widespread across societies. Figure SM3.2 shows difference in men and women age at first marriage in the case of 177 not-industrialized societies, 27 historical populations and 39 western contemporary nations. Except for one population of British columbia, men are older that women at first marriage in all populations. This holds even in societies were sexuality is not prohibited before marriage (eg in Hadza [10]), in matrilineal/matrilocal societies (eg [11] in Malawi and [12] in Khasi where men are respectively five and three years older than women), as well as in western developed population nowadays. Examining large amount of data in the case of 157 not-industrialized populations, [13] estimates that the mean age at first marriage is 14 for females and 21 for males in the case of 157 not-industrialized populations.

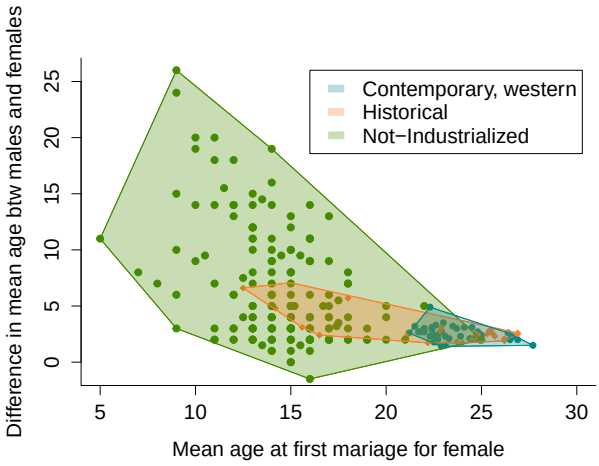


Figure 2: **Difference between mean male and female age at first marriage as a function of female age at first marriage** for **(i) 177 not-industralized populations** (data for 170 populations from [14], and [3] for the Ache, [10] for the Hazda of Tanzania [median], [15] for the Tsimane of Bolivia, [11] for population of rural Malawi, [6] for the Peul Bandé of Senegal, [12] for Khasi of Magalaya [median], and [5] for rural population of Gambia); **(ii) 27 historical populations** (data for England 1550-1599 from [16] and 1600-49, 1650-99, 1700-49, 1750-99 from [17], for XVIIIth Mexico [not nobility] from [18], for Saint-Hyacinthe, Quebec, 1852, 1861, 1871, 1881 and 1891 from [19], for Germany, 1700-49, 1750-99, 1800-24, 1825-49, 1850-74 and 1875-99 from [20], for Japan 1716-1759, 1760-1799, 1800-1839, 1840-1870 from [21], for chinese nobility of 1700 from [22], for Estonia 1835, 1840, 1845 and 1850 from [23]; for **39 contemporary western populations** from the seventies-eighties from [24].

Widowhood - Widowhood is little studied in anthropological demography and widow rates are rarely available. Widowhood exists however in all human populations since its depends mainly on adults mortality but in what extent? Supplementary material SM5 gives estimates of adult survival in the case one averaged hunter-gatherer population and one historical western population (Sweden 1751). Assuming that 20 years old men first marry 15 years old women and that divorce rate is negligible, the proportion of men surviving

at age 50 years old and having lost their first wife is $\left(1-\frac{l(45)}{l(15)}\right)\frac{l(50)}{l(20)}$. These proportion are 22% and 19% for the hunter-gatherer and the historical populations respectively: around one man over five loses its first wife before she is reaching age 45 years old. In western historical population where divorce were prohibited by religion, death of the husband or wife was the most important factor in marriage dissolution and remarriage.

Divorce -Ethnological studies showed that divorce is a common practice in human societies. Although measure of divorce rates is arduous because it depends on both the population age and nuptiality structures [25], data show that divorce rates vary in a large amount between populations[26]. Western christian populations where divorce were negligible (eg. in historical Germany [20], Quebec [27] or Finland [28]) have however been an exception in human history and divorces are found, even at lower levels, in most non-European societies [29]. In some populations permissiveness is loose and divorce may results from any grounds of discord between husband and wife. In most populations repeted infidelity, sterility, sexual unwillingness, laziness or mistreatments are accepted reasons for divorce [29, 30]. Divorce rates also vary with social organisation and level of exogmay [26]. In some populations, divorces are generating more remarriage than widowing. For example, 25% of Aka and 35% of Paliyan marriages end up in divorce [31]. In Ache, the absolute divorce rate is about 61% during the first year of a marriage and explains why ache men have an average of 10.8 mates during a lifetime [32]. In matrilineal Khasis, up to 50% of women aged 40+ have divorced at least once [12].

Polygamy - Measuments and categorisation of polygamy has proved difficult and has been largely discussed since the precursor work of Murdock [33] (eg, by [34], [35] or [36]]; for instance according to the rates of men and women married polygynously, which categories of men or women are polygynously married (as for example according to rank or stratification amongs men and women), and how they are distributed within families and with residence. Using an dataset on 190 populations, mainly from the Standard Cross-cultural Sample [37], [38] showed that polygyny is found in 73.7% of societies, while it is reported as common in 21.5% and rare in 78.5% societies. Polyandry is also present and is reported in 12.1% of cases, always at low levels. The use of a subsample of 54 societies equally distributed across geographical areas does not modify substancially these results. Using written historical data on 16 civilisations from the past 5000 years, [39]showed that polygyny was accepted in 57% and monogmay in 43% of cases. Categorisation of common and rare polygyny have been associated with two main types of polygyny [34]. In 'wealth-increasing' polygyny, women's labor generates wealth and most men are able to become polygynists with age (which is sustainable in even slightly growing populations where young age classes are always found at larger proportions than older age classes [40]). In 'exceptional men' polygyny, most of the provisioning is generated by the husband and polygyny is usually associated with the exceptional productivity and /or social status of particular men. Balance between males and females provisioning appears being the main determinant of polygyny although other factors, such as pathogen stress, mating decision, filiation and residence rules also have an influence on level of polygyny (eg [36, 38, 40]). Polygyny is often associated with large difference of age between wife and husband. For example, in 'wealth-increasing' polygyny, men queu for access to women and marry therefore to older ages. In sororeal polygyny, the eldest girl in a family marries first and that as they come of age her younger sisters join her as cowives. It must be stressed that polyandry may be largely underestimated and may occur at a large rate over short period of time in case of desequilibrium in sex-ratio resulting for large male adult mortality [41]. Finaly, phylogenetic reconstructions using both anthropological and mtDNA data suggest that levels of polygyny was the most likely scenarios for most of ancestral human societies [8].

2. Model and parametrization

We aim at modelling average fertility functions that described differences between females and males fertility described above. To do this we chose to model females fertility rates $F(x)$ and $F^m(x)$ by the Brass-Polynomial function [42] known being flexible in fitting fertility data [43], as follows:

$$F(x) = (x-\alpha)\,(\beta-x)^2\left[A + Bx + Cx^2 + Dx^3\right], \tag{1}$$

where we set α and β (or α^m and β^m for males) are the first and last year of reproduction and are respectively set to 15 and 49 for females and 20 and 70 for males. We then fit age-specific fertility rates from the six populations presented in Fig.SM3.1 using a least-square methods (using 'nls' function in 'R' software). Results are shown respectively in Fig.3.3A and Fig.3.3B.

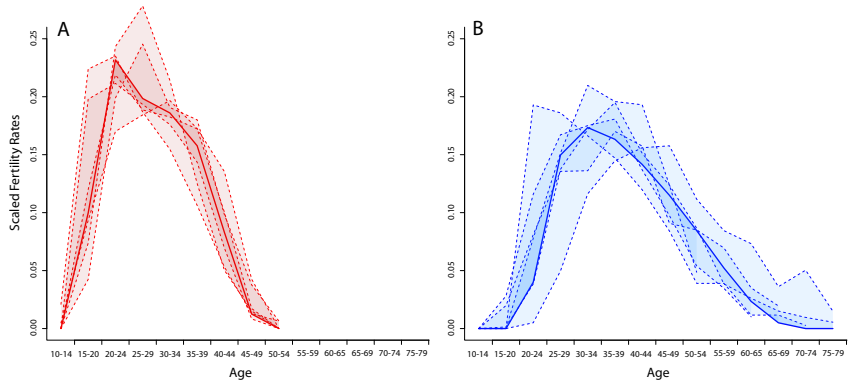


Figure 3: Scaled age-specific fertility rates for women (**A**) and men (**B**). Dashed lines are published estimates for the Kung, Ache, Tsimane, Gambia, Peul Bandé and French canadians. Plain lines are the coresponding the fitted brass polynomial. Parameters are respectively $A = 4.64e-04$, $B = -4.26e-05$, $C = 1.34e-06$ and $D = -1.37e-08$ for females and $A^m = 4.80e-04$, $B^m = -2.47e-06$, $C^m = 5.05e-08$ and $D^m = -3.5e-10$ for males.

We need to incorporate males and females age-specific fertility rates into a two-sex model where children are born to mother and father at a specific age. Because males and females produce in mean the same number of children over a lifetime, $F(x)$ and $F^m(x)$ must satisfy:

$$\int_x l(x)F(x)dx = \int_x l^m(x)F^m(x)dx, \tag{2}$$

where $l(x)$ and $l^m(x)$ are the survival of females and males into a given population. Because survival and fertility are usually estimated with two disctinct datasets respectively for females and for males, and because of statistical errors, this equation never holds with estimated data. This is why we fine-tuned the parameter A^m of the Brass polynomial function for males in order to satisfy equation (2) such that:

$$A^m = \frac{\sum_x L(x)F(x) - \sum_x L^m(x)\,(x-\alpha^m)\,(\beta^m - x)^2\,[Bx + Cx^2 + Dx^3]}{\sum_x L^m(x)\,(x-\alpha^m)\,(\beta^m - x)^2} \tag{3}$$

Resulting age-specific men and women fertility rates are shown in Fig.2A of the main text.

References

[1] Shripad D. Tuljapurkar, Cedric O. Puleston, and Michael D. Gurven. Why men matter: Mating patterns drive evolution of human lifespan. *PLoS ONE*, 2(8):e785, 2007.

[2] N. Howell. *Demography of the Dobe !Kung.* Aldine De Gruyter, New York, 2nd edition, 2000.

[3] K. Hill and A.M. Hurtado. *Ache life history: the ecology and demography of a foraging people.* De Gruyter, New York, USA, 1996.

[4] H. Kaplan, M. Gurven, J. Winking, P. L. Hooper, and J. Stieglitz. Learning, menopause, and the human adaptive complex. *Ann N Y Acad Sci*, 1204:30–42, 2010.

[5] A. A. Ratcliffe, A. G. Hill, and G. Walraven. Separate lives, different interests: male and female reproduction in the gambia. *Bull World Health Organ*, 78(5):570–579, 2000.

[6] G. Pison. *Dynamique d'une population traditionnelle: les Peul BandÃ© (SÃ©nÃ©gal oriental)*, volume 99 of *Edition de l'INED, Travaux et documents.* Presses Universitaires de France, 1982.

[7] F.F. Payeur. *Age et fertilitÃ© masculine : une analyse biodÃ©mographique.* PhD thesis, 2008.

[8] Robert S. Walker, Kim R. Hill, Mark V. Flinn, and Ryan M. Ellsworth. Evolutionary history of hunter-gatherer marriage practices. *PLoS ONE*, 6(4):e19066, 2011.

[9] G. P. Murdock. *Social Structure*. The MacMillan Company, New York, 1949.

[10] FrankW Marlowe. Mate preferences among hadza hunter-gatherers. *Human nature*, 15(4):365–376, 2004.

[11] R. Sear. Kin and child survival in malawi: are matrilineal kin always beneficial in a matrilineal society? *Human nature*, 19(3):277–293, 2008.

[12] Donna L. Leonetti, Dilip C. Nath, and Natabar S. Hemam. In-law conflict: Women's reproductive lives and the roles of their mothers and husbands among the matrilineal khasi. *Current Anthropology*, 48(6):861–890, 2007.

[13] J. N. Fenner. Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *Am J Phys Anthropol*, 128(2):415–23, 2005.

[14] L.R. Binford. *Constructing Frames of Reference: An Analytical Method for Archaeological Theory Building Using Hunter-Gatherer and Environmental Data Sets*. University of California Press, Berkeley, 2001.

[15] J. Winking, J. Stieglitz, J. Kurten, H. Kaplan, and M. Gurven. Polygyny among the tsimane of bolivia: an improved method for testing the polygyny-fertility hypothesis. *Proc Biol Sci*, 280(1756):20123078, 2013.

[16] R.D. Lee and R.S. Schofield. *British population in the eighteen century*, volume 1, pages 17–35. Cambridge University Press, Cambridge, 1981.

[17] E. A. Wrigley and R. S. Schofield. English population history from family reconstitution: Summary results 1600-1799. *Population Studies*, 37(2):157–184, 1983.

[18] R. Gutierrez. *When Jesus Came, the Corn Mothers Went Away: Marriage, Sexuality, and Power in New Mexico, 1500-1846*. Stanford University Press, Stanford, 1991.

[19] P. Gossage. Family formation and age at marriage in saint-hyacinthe quebec, 1854-1891. *Social history*, 24(47):61–84, 1991.

[20] J.E. Knodel. *Demographic Behavior in the Past: A Study of Fourteen German Village Populations in the Eighteenth and Nineteenth Centuries*. Cambridge University Press, Cambridge, 1988.

[21] N.O. Tsuya and S. Kurosu. Economic and household factors of first marriage in early modern japan: Evidence from two northeastern villages, 1716-1870, 2000.

[22] J. Lee, Feng W., and D. Ruan. *Nuptiality among the Qing Nobility, 1640-1900*, pages 353–373. Oxford University Press, Oxford, 2001.

[23] E. Waris. Historical-anthropological family research. a case study on marriage and family in estonia (1750-1850), 15-16 april 2004.

[24] Generations and Gender Programme. Generations and gender contextual database, 2013.

[25] J. A. Barnes. Measures of divorce frequency in simple societies. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, 79(1/2):37–62, 1949.

[26] Charles Ackerman. Affiliations: Structural determinants of differential divorce rates. *American Journal of Sociology*, 69(1):13–20, 1963.

[27] H. Charbonneau, A. A. Guillemette, J. LÃ©garÃ©, B. Desjardin, Y. Landry, and F. Nault. *Naissance d'une population. Les FranÃ§ais Ã©tablis au Canada au XVIIe siÃšcle*. Presses de l'universitÃ© de MontrÃ©al, MontrÃ©al, 1987.

[28] Pekka Kaar, Jukka Jokela, Juha Merila, Timo Helle, and Ilpo Kojola. Sexual conf lict and remarriage in preindustrial human populations: Causes and fitness consequences. *Evolution and Human Behavior*, 19(3):139–151, 1998.

[29] George Peter Murdock. Family stability in non-european cultures. *Annals of the American Academy of Political and Social Science*, 272(ArticleType: research-article / Issue Title: Toward Family Stability / Full publication date: Nov., 1950 / Copyright Â© 1950 American Academy of Political and Social Science):195–201, 1950.

[30] Laura Betzig. Causes of conjugal dissolution: A cross-cultural study. *Current Anthropology*, 30(5):654–676, 1989.

[31] P.M. Gardner and T. Ingold. *Pressures for Tamil propriety in Paliyan social organization.*, volume 1 of *Hunters and gatherers*, pages 91–106. Berg, Oxford, 1991.

[32] K. Hill and A.M. Hurtado. *Ache life history: the ecology and demography of a foraging people.* De Gruyter, New York, USA, 1996.

[33] G.P. Murdock. *Ethnographic Atlas.* University of Pittsburgh Press, Pittsburgh, Pennsylvania., 1967.

[34] Douglas R. White, Laura Betzig, Monique Borgerhoff Mulder, Garry Chick, John Hartung, William Irons, Bobbi S. Low, Keith F. Otterbein, Paul C. Rosenblatt, and Paul Spencer. Rethinking polygyny: Co-wives, codes, and cultural systems [and comments and reply]. *Current Anthropology*, 29(4):529–572, 1988.

[35] Bobbi S. Low. Measures of polygyny in humans. *Current Anthropology*, 29(1):189–194, 1988.

[36] Frank W. Marlowe. The mating system of foragers in the standard cross-cultural sample. *Cross-Cultural Research*, 37(3):282–306, 2003.

[37] George P. Murdock and Douglas R. White. Standard cross-cultural sample. *Ethnology*, 8(4):329–369, 1969.

[38] Menelaos Apostolou. Sexual selection under parental choice: the role of parents in the evolution of human mating. *Evolution and Human Behavior*, 28(6):403–409, 2007.

[39] M. Apostolou. Sexual selection under parental choice: evidence from sixteen historical societies. *Evol Psychol*, 10(3):504–18, 2012.

[40] Gilles Pison. La dÃ©mographie de la polygamie. *Population (French Edition)*, 41(1):93–122, 1986.

[41] K. E. Starkweather and R. Hames. A survey of non-classical polyandry. *Hum Nat*, 23(2):149–72, 2012.

[42] W. Brass. The graduation of fertility distributions by polynomial functions. *Population Studies*, 14(2):148–162, 1960.

[43] J. M. Hoem, D. Madsen, J. L. Nielsen, E. M. Ohlsen, H. O. Hansen, and B. Rennermalm. Experiments in modeling recent danish fertility curves. *Demography*, 18(2):231–244, 1981. Lq706 Times Cited:21 Cited References Count:44.

SUPLEMENTARY MATERIALS 4 - Child survival as a function of (grand)parental care

1. Data

The importance of (grand)parental care for child survival (also a measure of child altriciality) is classicaly estimated in demography as the increased risk of death of children whose parent is dead (or absent) compared to those whose parent is alive (and present). Table SM4 gathers 31 articles having estimated this increase in the case of the absence or death of the mother, the father and maternal or paternal grandmother; in various pré- or not-industrialized populations [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31]. Overall, we gathered 194 estimators of children relative risk of death in response to the death of the mother (79), of the father (40), of the maternal grandmother (38) and paternal grand mother (37). These data are also plotted in figure SM4.1. Results are largely contrasted between studies. Whether or not these variations reflect differences in population behaviors with respect to child care is however far from being clear.

Indeed these variations amy also be due to statistical problems in estimating intrafamilial correlation in survival. This type of analysis is far from being statiscally easy because a child's death may clutser with its parent's death, not only due to the subsequent lack of care, but also due to many other cofounding factors [20]: 1) *Between-family heterogeneity* means that all families are not equal with respect to mortality levels due to variance in socioeconomic conditions and behavioral characteristics of the parents [32], as well as biological endogenous factors shared by the members of the same family (i.e., genetic and obstetric factors; [33]). 2) *Within-family heterogeneity* means that children of the same family may have a differential risk of death linked to maternal reproductive history. The most obvious ismaternal age at child birth. Children born to old mother have a larger risk of death due to decrease in gametes quality and increased birth defects. The risk of mortality varies also from one child to another with respect to birth interval, survival of the preceding child, sibship size and/or birth order [e.g., 34, 35]. 3) *Sporadic increase in familial mortality* means that deaths can suddenly cluster within a family due to accidents, epidemics, cross-infections or any familial reversal of fortune [20, 36, 22].

Despite these statistical difficulties, it is clear from figure SM1.1A that maternal death compromises children survival in most studies (only 3 estimates out of 79 are lower or equal to one) and most of the time to a large

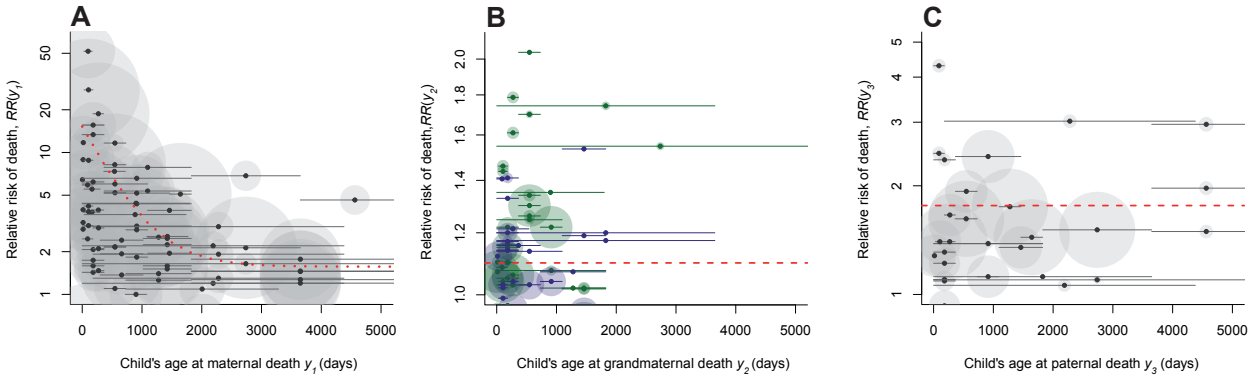


Figure 1: Cpation

extent (53 estimates out of 79 are larger than 2). The higher risk of dying of motherless children remains significant throughout childhood, although a decrease in dependency from maternal care as the child grows older is clearly visible.

Results are largely contrasted in the case of fathers. Twelve estimates out of 40 are lower or equal than 1 meaning that the father's death makes no difference to child survival. When it does, it has a substantially lesser impact than the mother's death [37] and only 7 estimates out of 40 are larger than 2. This may be due to statistical difficulties but also to the fact that father's care may be more directed to providing ressources and education o children than nursing at young age. As a consequence, paternal care may be more important at a later child age, a period of ages little covered by these studies. This could also be due to the fact that paternal care may vary with the sex of the child while most relative risk are estimated for both sexes. For example, in a patrilineal Ethiopian community, father's absence doubles a son's risk of dying in infancy but has a positive influence on daughter survival [12]. Paternal effect may also depends on whether mother remarry or not in pre-industrial Québec [38].

A previous meta-analysis of grandmaternal care, showed that maternal rather than paternal grandmother tend to have the larger effect on grand-child survival, even in patrilineal societies [39]. This is confirmed by our results over a larger sample of estimates. Maternal grandmother estimates are larger than 1 in 30 cases out of 38 while in only 20 cases out of 37 for the paternal grandmother. The effect remains however of small magnitude (all estimates but one are lower than 2). Finally, estimates favour the exhistence of a preferential garndmother. Over the 12 studies analyzing the effect of both maternal and paternal grandmother, the mean estimate of maximum value (whether this is the maternal or paternal grandmother) is 1.279 while the mean estimate of minimum value is 0.91; both being significantly different ($p_{Wilcoxon} = 1.86\text{x}10^{-8}$).

A more detailed meta-analysis of these data directed to kinship demographers is in preparation.

2. Model and parametrization

The aim is to fit functions for the relative risk of death $RR(y)$ as a function of child's age at maternal, grandmaternal and paternal death y using data pictured in figure SM4.1. To do this, y is condidered as the mid-point of the period of child's age at parental death upon which the mortality was analysed. Five functions were then tested by least square methods (using the 'nls' function of the package 'stats' in 'R'): $ae^{-by} + 1$, $ay^{-b} + 1$, $ae^{-by} + 1 + c$, $ay^{-b} + 1 + c$, a; where a, b and c the parameters of the fitted function. the best model was selected according to F-statistics using ANOVA (function 'anova' of the package 'stats' in 'R).

A composite weighing factor was also incorporated into the regression as the addition of three independent weights. First, because standard error of relative risk were not provided in all articles, we used study sample size (including all children, orphans and not) as a proxy of the study statistical power. Studies' estimates were then ranked according to the quartiles of sample size by a categorial weight $w_1 = [0, 3]$.

Second, if the increased risk of child's death declines with the child's age at (grand)parental loss, the hypothesis that the fitted function passes at the mid-point of the interval may not always be reasonable. If the decline is fast enought, relative risk estimated over large periods of age means that the real risk is much larger than the estimate at the beginning of the period, meets the estimate at an age younger than the mid-point of the period and is lower than the estimate afterwards. We aim therefore to penalize study analyzing young children risk of death over a long period of age (e.g., examining risk of death for children from birth to 5 years old). However, we do not want to penalize estimates from periods of two kinds: if the period is short enought at young ages (e.g., estimated over weeks after birth) or if the period is large but starting when children are older and therefore during which the risk is declining at a lower pace (e.g., examining the risk of death of children older than 5 years of age during 5 years).. To do this we define a second weight $w_2 = [0, 3]$categorizing the quartile of the metric $\log(y_i/(y_f - y_i))$, where y_i and y_f are respectively the children age at the beginning and the end of the period. Of course, this weight is especially relevant for fitting $RR(y)$ as a function of child's age at maternal death y_1 because figure SM4.1A exhibits a clear decline in $RR(y_3)$ but is removed from analyses fitting a uniform relative risk of death as a function of child's age at grandmaternal death y_2 and paternal death y_3.

Third, as discuss above, studies differs in their efficiency in controlling for counfonding factors, i.e., *within-* and *between-family heterogeneity* as well as *sporadic increase of familial mortality*. To account for this, a

third weight $w_3 = [0, 3]$ is calculated adding respectively 0.5 or 1 points to study partially or satisfactorily controling for these factors following entries of table SM4. It must be stressed that controlling for *within-family heterogeneity* is relevant only for controlling for cluster of deaths between mother and chidren and is not accounted for when analysing grandmothernal and paternal deaths. Similarily, controlling for cluster of deaths between grandmothers and children due to sporadic increase of familial mortality makes less sense because grandmothers may be dead long before the children birth.

Results are despecited in figure SM4.1 and in figure X of the main text. Children relative risk of death as a function of child age at maternal death y_1 were best fitted by function $ay^{-b}+1+c$ with $a = 13.67$, $b = 0.7086$ and $c = 0.5678$. Children relative risk of death as a function of child age at grandmaternal and paternal death were best fitted by uniform distribution of respective parameters $a = 1.098$ and $a = 1.763$.

Assuming that these risks are multiplicatives, a child's age-specific relative risk of death $RR(t)$ can be calculated for any combination $[y_1, y_2, y_3]$ as:

$$RR\left(t|y_1, y_2, y_3\right) = RR\left(t|y_1\right) RR\left(t|y_2\right) RR\left(t|y_3\right)$$

Holding the fact that relative risk of death is 1 before the death of the parent and $RR(t|y)$ after its death, then:

$$RR\left(t|y_1, y_2, y_3\right) = \left(\begin{array}{ll} RR\left(y_1\right) & \text{for} \quad \text{t} > y_1 \\ 1 & \text{for} \quad \text{t} < y_1 \end{array} \right) \left(\begin{array}{ll} RR\left(y_2\right) & \text{for} \quad \text{t} > y_2 \\ 1 & \text{for} \quad \text{t} < y_2 \end{array} \right) \left(\begin{array}{ll} RR\left(y_3\right) & \text{for} \quad \text{t} > y_3 \\ 1 & \text{for} \quad \text{t} < y_3 \end{array} \right)$$

Table SM4 - Demographic studies analyzing the effect of parental or grandparental death on infant and child increased risk of death

Study	Data	Kin	Sample Size [a]	Age period	Children age at parental death	Methods	Dependent variable/ Reference group	Within family heterogeneity [b]	Betwen family heterogeneity [c]	Sporadic increase in mortality [d]
Anderson et al., 2007	Rural Haiti, 1997-1999	Mother	232	0-12 yrs	Soon after birth	Calculation of Odd Ratios	Mortality/ Kin is alive.	No	Yes	Yes
Becher et al., 2004	Burkina Faso, 1993- 1999	Mother	10122	2 periods from birth to 2 yrs	On period or before	Proportional hazard model	Mortality/ Kin is alive.	Yes	Partially	No
Beise & Voland 2002	Krümmhorn, Germany, 1720-1870	MGM, PGM	3550	6 periods from birth to age 60 mo	On period or before	Proportional hazard	Mortality/ Kin is alive.	Partially	Yes	No
Beise 2005	French Canada, 1680-1750	Mother Father MGM MGF PGM PGF	26449	6 periods from birth to age 60 mo	On period or before	Proportional hazard	Mortality/ Kin is alive.	Yes	No	No
Beekink et al., 1999	Woerden, Netherlands, 1850-1930	Father Mother	3936	2 periods from births to 12 yrs	On period or before	Proportional hazard	Mortality/ Kin is alive.	Partially	Partially	Partially
Beekink et al., 2002	Woerden, Netherlands, 1850-1930	Father Mother	3936	2 periods from births to 12 yrs	On period or before, time since parental death	Proportional hazard	Mortality/ Kin is alive.	Partially	Partially	Partially
Bishai et al. 2003	Rural Uganda, 1994-1999	Mother Father	2332	0-5 yrs	On Period	Bivariate analysis	Mortality/ Kin is alive.	Partially	Yes	No
Borgerhoff-Mulder 2007	Kipsigis of Kenya, 1945-1990	MGM MGF PGM PGF	785	0-60 mo	Before fifth birthday	Proportional hazard model	Survival/ Kin is dead	Yes	Yes	No
Breschi and Manfredini 2002	Tuscany, Italia, 1819-1959	Mother Father	~3744	0-12 yrs	Within period	Event-histoy analysis	Mortality/ Kin is alive.	Partially	Partially	Partially
Brittain 1992	St Bathélemy island, 1878-1976	Mother	4286	0-1 yrs	within a year of child's birth	Stepwise regression	Mortality/ Kin is alive.	Partially	No	No
Gibson & Mace 2005	Oromo of Ethiopia, 1999-2003	MGM MGF PGM PGF	2746	0-3 yrs	On period or before	Logistic Regression	Mortality/ Kin is dead	Partially	Yes	No
Gibson 2008	Oromo of Ethiopia, 1999-2003	Father	3720	0-1 yrs	On period	Proportional hazard	Mortality/ Kin is dead or absent	Partially	Partially	No
Jamison et al., 2002	Nagano, Japan , 1671-1871	Mother Father MGM MGF PGM PGF	2381	0-15 yrs	On period or before	Logistic regression	Mortality Kin is dead	Partially	No	No
Katz et al 2003	Sarlahi district, Nepal 1994-1997	Mother	15469	3 periods from birth to 24 wks	On period or before	Logistic regression	Mortality/ Kin is alive.	Yes	Yes	No
Kemkes-Grottenthaler 2005	Germany, 1704-1899	MGM MGF PGM PGF	1590	4 periods from birth to 2 yrs	On period or before	Binary logistic regression	Survival/ Kin is dead	Partially	No	No
Leonetti et al., 2005	Bengalis of Assam and Khasis of Megalaya, India, 1980-2000	MGM PGM	2069 and 2545	0-10 yrs	On period or before	Proportional hazard	Mortality/ Kin is alive	Yes	Yes	No
Masmas et al 2004	Rural Guinea-Bissau, 1990-1996	Mother	11447	6 periods from birth to 5yrs	On period or before	Proportional Hazard	Mortality/ Kin is alive	No	Yes	No
Nakiyingi et al., 2003	Uganda, 1989-2000	Mother	3727	0-5 yrs	On period	Piecewise exponential hazards	Mortality/ Kin is alive	Partially	No	No
Newell et al., 2004	West, South and East Africa	Mother	3468	1mo-58mo	Recently dead	Proportional hazard	Mortality/ Kin is alive	No	No	No
Pavard et al., 2005	French Canadia, 1625-1759	Mother	58365	4 periods from birth to 15 yrs	On period and before	Proportional hazard	Mortality/ Kin is alive	Yes	Yes	Yes
Reher et al 2003	Spain, 1870-1910 and 1911-1950	Mother	8049 and 9796	5 periods from birth to 2 yrs	<24 mo after birth	Ratio of mortality coefficient	Mortality/ Kin is alive	No	No	No

Table SM4 (suite)

Ronsmans et al 2010	**Bangladesh, 1982-2005**	**Mother Father**	**143473**	**7 periods from birth to 10 yrs**	**On period or before**	**Poisson regression on person-time**	**Mortality/ Kin is alive**	**No**	**Yes**	**Yes**
Sartorius et al., 2012	South Africa, 1992-2007	Mother Father	46675	1-4 yrs	On period or before	Proportional hazard	Mortality/ Kin is alive	Yes	Partially	Partially
Sear et al 2000	Rural Gambia, 1950-1974	Mother Father MGM MGF PGM PGF	1691	3 periods from birth to 3	On period or before	Event-history models	Mortality/ Kin is alive	Yes	Yes	No
Sear et al 2002	Rural Gambia, 1950-1974	Mother Father MGM MGF PGM PGF	2294	3 periods from birth to 3	On period or before	Event-history models	Mortality/ Kin is alive	Yes	Yes	Partially
Sear 2008	Malawi, 1997	Father MGM MGF PGM PGF	1635	0-10 yrs.	On period	Logistic regression	Mortality/ Kin is alive	Yes	No	No
Strassman 2011	Dogon of Mali 1998-2011	Mother Father MGM MGF PGM PGF	2933	0-5 yrs	On period	Proportional hazard model	Mortality/ Kin is dead	Partially	Yes	No
Tymicki 2009	Bejsce , Poland 1737-1819 and 1820-1917	Mother Father MGM MGF PGM PGF	6569 and 13680	0-12 mo 0-60 mo	On period or before	Parametric gompertz hazard model	Mortality/ Kin is alive	Partially	No	No
Voland & Beise 2002	Krümmhorn, Germany 1720-1874	MGM PGM	3095	6 periods from birth to 60 mo	On period or before	Piece-wise exponential hazard	Mortality/ Kin is alive	Partially	No	No
Winking et al., 2011	Tsimane of Bolivia 2002-2005	Mother Father	6795	3 periods from birth to 10 yrs	On period or before	Logistic regression models	Mortality/ Kin is alive	Partially	Yes	No
Zabba et al 2005	Tanzania,Uganda,Malawi, 1980s-2000	Mother	10849	2 periods from birth to 5 yrs	Less than one year ago	Piece-wise exponential hazards model	Mortality/ Kin is alive	Partially	No	No

a Total number of children incorporated into the study. We distinguished sample sizes when different populations or time-period are considered. However, we did not incorporate the sizes of the sub-samples of children used in each age-period analysis.

b Considered as fully (or partially) controlled for when at least two (or one) of the following factors are accounted for into the analysis, either by the incorporation of the variable into multivariate analysis or by selection of cases: maternal age and/or birth order, interbirth interval and/or survival status of previous child, twinning.

c Considered as fully accounted for when only children of the same family are compared , when a random variable on the mother's idi is incorporated into the analysis, when ego's mortality is corrected for that of its siblings or when at least three fixed variables describing family socio-economic or geographical conditions are incorporated. Considered as partially if less than three of these variables are incorporated or when incorporated random variable stands for groups of larger granularity (e.g., at the village level).

d Considered as controlled for when clusters of death between the considered kin and child have been controlled or removed. Considered as partially accounted for when variables describing fluctuation of temporal condition are incorporated (e.g., period effect, epidemics). Many studies are categorized as "no" while some of the estimates could be considered as "yes" when the kin is obviously dead before the considered children's period of age.

References

[1] F. W. Anderson, S. U. Morton, S. Naik, and B. Gebrian. Maternal mortality and the consequences on infant and child survival in rural haiti. *Matern Child Health J*, 11(4):395–401, 2007. Anderson, Frank W J Morton, Sarah U Naik, Sujata Gebrian, Bette Research Support, Non-U.S. Gov't United States Maternal and child health journal Matern Child Health J. 2007 Jul;11(4):395-401. Epub 2007 Jan 31.

[2] H. Becher, O. Muller, A. Jahn, A. Gbangou, G. Kynast-Wolf, and B. Kouyate. Risk factors of infant and child mortality in rural burkina faso. *Bull World Health Organ*, 82(4):265–73, 2004. Becher, Heiko Muller, Olaf Jahn, Albrecht Gbangou, Adjima Kynast-Wolf, Gisela Kouyate, Bocar Research Support, Non-U.S. Gov't Switzerland Bulletin of the World Health Organization Bull World Health Organ. 2004 Apr;82(4):265-73.

[3] Erik Beekink, Frans van Poppel, and Aart C. Liefbroer. Surviving the loss of the parent in a nineteenth-century dutch provincial town. *Journal of Social History*, 32(3):641–669, 1999.

[4] E. Beekink, F. van Poppel, and A.C. Liefbroer. *Parental death and death of the child: Common causes or direct effects?* Bern, Peter Lang, New-York, 2002.

[5] J. Beise. *The helping grandmother and the helpful grandmother: The role of maternal and paternal grandmothers in child mortality in the 17th and 18th century population of French settlers in Quebec, Canada*, page 215â238. Rutgers University Press, New Brunswick, 2005.

[6] Jan Beise and Eckart Voland. A multilevel event history analysis of the effects of grandmothers on child mortality in a historical german population: KrummhÃ¶rn, ostfriesland, 1720-1874. *Demographic Research*, 7(13):469–498, 2002. 10.4054/DemRes.2002.7.13.

[7] David Bishai, Heena Brahmbhatt, Ron Gray, Godfrey Kigozi, David Serwadda, Nelson Sewankambo, El Daw Suliman, Fred Wabwire-Mangen, and Maria Wawer. Does biological relatedness affect child survival? *Demographic Research*, 8(9):261–278, 2003. 10.4054/DemRes.2003.8.9.

[8] Monique Borgerhoff Mulder. Hamilton's rule and kin competition: the kipsigis case. *Evolution and Human Behavior*, 28(5):299–312, 2007.

[9] M. Breschi and M. Manfredini. *Parental loss and kin networks: Demographic repercussions in a rural Italian village*. Bern, Peter Lang, New-York, 2002.

[10] A. W. Brittain. Birth spacing and child mortality in a caribbean population. *Hum Biol*, 64(2):223–41, 1992. Brittain, A W United states Human biology Hum Biol. 1992 Apr;64(2):223-41.

[11] Mhairi A. Gibson and Ruth Mace. Helpful grandmothers in rural ethiopia: A study of the effect of kin on child survival and growth. *Evolution and Human Behavior*, 26(6):469–482, 2005.

[12] MhairiA Gibson. Does investment in the sexes differ when fathers are absent? *Human Nature*, 19(3):263–276, 2008.

[13] C. S. Jamison, L. L. Cornell, P. L. Jamison, and H. Nakazato. Are all grandmothers equal? a review and a preliminary test of the "grandmother hypothesis" in tokugawa japan. *Am J Phys Anthropol*, 119(1):67–76, 2002. Jamison, Cheryl Sorenson Cornell, Laurel L Jamison, Paul L Nakazato, Hideki Historical Article United States American journal of physical anthropology Am J Phys Anthropol. 2002 Sep;119(1):67-76.

[14] J. Katz, Jr. West, K. P., S. K. Khatry, P. Christian, S. C. LeClerq, E. K. Pradhan, and S. R. Shrestha. Risk factors for early infant mortality in sarlahi district, nepal. *Bull World Health Organ*, 81(10):717–25, 2003. Katz, Joanne West, Keith P Jr Khatry, Subarna K Christian, Parul LeClerq, Steven C Pradhan, Elizabeth Kimbrough Shrestha, Sharada Ram Clinical Trial Randomized Controlled Trial Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S. Switzerland Bulletin of the World Health Organization Bull World Health Organ. 2003;81(10):717-25. Epub 2003 Nov 25.

[15] Ariane Kemkes-Grottenthaler. Of grandmothers, grandfathers and wicked step-grandparents. differential impact of paternal grandparents on grandoffspring survival. *Historical Social Research / Historische Sozialforschung*, 30(3 (113)):219–239, 2005.

[16] D. L. Leonetti, D. C. Nath, N. S. Hemam, and D. B. Neill. *Kinship organisation and the impact of grandmothers on reproductive success among the matrilineal Khasi and patrilineal Bengali of Northeast India*, page 194â214. Rutgers University Press, New Brunswick, 2005.

[17] T. N. Masmas, H. Jensen, D. da Silva, L. Hoj, A. Sandstrom, and P. Aaby. Survival among motherless children in rural and urban areas in guinea-bissau. *Acta Paediatr*, 93(1):99–105, 2004. Masmas, T N Jensen, H da Silva, D Hoj, L Sandstrom, A Aaby, P Research Support, Non-U.S. Gov't Norway Acta paediatrica (Oslo, Norway : 1992) Acta Paediatr. 2004 Jan;93(1):99-105.

[18] J. S. Nakiyingi, M. Bracher, J. A. Whitworth, A. Ruberantwari, J. Busingye, S. M. Mbulaiteye, and B. Zaba. Child survival in relation to mother's hiv infection and survival: evidence from a ugandan cohort study. *AIDS*, 17(12):1827–34, 2003. Nakiyingi, Jessica S Bracher, Michael Whitworth, James A Ruberantwari, Anthony Busingye, June Mbulaiteye, Sam M Zaba, Basia Research Support, Non-U.S. Gov't England AIDS (London, England) AIDS. 2003 Aug 15;17(12):1827-34.

[19] M. L. Newell, H. Coovadia, M. Cortina-Borja, N. Rollins, P. Gaillard, and F. Dabis. Mortality of infected and uninfected infants born to hiv-infected mothers in africa: a pooled analysis. *Lancet*, 364(9441):1236–43, 2004. Newell, Marie-Louise Coovadia, Hoosen Cortina-Borja, Marjo Rollins, Nigel Gaillard, Philippe Dabis, Francois Ghent International AIDS Society (IAS) Working Group on HIV Infection in Women and Children N01 HD 33163/HD/NICHD NIH HHS/United States N01 HD 33345/HD/NICHD NIH HHS/United States Research Support, U.S. Gov't, P.H.S. England Lancet Lancet. 2004 Oct 2-8;364(9441):1236-43.

[20] S. Pavard, A. Gagnon, B. Desjardins, and E. Heyer. Mother's death and child survival: The case of early quebec. *Journal of Biosocial Science*, 37(2):209–227, 2005. 921HC Times Cited:1 Cited References Count:54.

[21] D. S. Reher and F. Gonzalez-Quinones. Do parents really matter? child health and development in spain during the demographic transition. *Popul Stud (Camb)*, 57(1):63–75, 2003. Reher, David Sven Gonzalez-Quinones, Fernando Historical Article England Population studies Popul Stud (Camb). 2003;57(1):63-75.

[22] Carine Ronsmans, Mahbub E. Chowdhury, Sushil K. Dasgupta, Anisuddin Ahmed, and Marge Koblinsky. Effect of parent's death on child survival in rural bangladesh: a cohort study. *The Lancet*, 375(9730):2024–2031, 2010.

[23] B. Sartorius, K. Kahn, M. A. Collinson, P. Vounatsou, and S. M. Tollman. Survived infancy but still vulnerable: spatial-temporal trends and risk factors for child mortality in the agincourt rural subdistrict, south africa, 1992-2007. *Geospat Health*, 5(2):285–95, 2011. Sartorius, Benn Kahn, Kathleen Collinson, Mark A Vounatsou, Penelope Tollman, Stephen M 058893/Z/99/A/Wellcome Trust/United Kingdom 069683/Wellcome Trust/United Kingdom 069683/Z/02/Z/Wellcome Trust/United Kingdom 069683/Z/08/Z/Wellcome Trust/United Kingdom 085477/Wellcome Trust/United Kingdom Research Support, Non-U.S. Gov't Italy Geospatial health Ems36790 Geospat Health. 2011 May;5(2):285-95.

[24] Rebecca Sear, Ruth Mace, and Ian A. McGregor. Maternal grandmothers improve nutritional status and survival of children in rural gambia. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 267(1453):1641–1647, 2000.

[25] Rebecca Sear, Fiona Steele, Ian A. McGregor, and Ruth Mace. The effects of kin on child mortality in rural gambia. *Demography*, 39(1):43–63, 2002.

[26] R. Sear. Kin and child survival in rural malawi : Are matrilineal kin always beneficial in a matrilineal society? *Human Nature*, 19(3):277–93, 2008. Sear, Rebecca United States Human nature (Hawthorne, N.Y.) Hum Nat. 2008 Sep;19(3):277-93. doi: 10.1007/s12110-008-9042-4.

[27] Beverly I. Strassmann. Cooperation and competition in a cliff-dwelling people. *Proceedings of the National Academy of Sciences*, 108(Supplement 2):10894–10901, 2011.

[28] Krzysztof Tymicki. The correlates of infant and childhood mortality: A theoretical overview and new evidence from the analysis of longitudinal data of the bejsce (poland) parish register reconstitution study of the 18th-20th centuries. *Demographic Research*, 20(23):559–594, 2009. 10.4054/DemRes.2009.20.23.

[29] Eckart Voland and Jan Beise. Opposite effects of maternal and paternal grandmothers on infant survival in historical krummhÃ¶rn. *Behavioral Ecology and Sociobiology*, 52(6):435–443, 2002.

[30] Jeffrey Gurven Michael Kaplan Hillard Winking. The impact of parents and self-selection on child survival among the tsimane of bolivia. *Current anthropology*, 52(2):277–284, 2011.

[31] B. Zaba, J. Whitworth, M. Marston, J. Nakiyingi, A. Ruberantwari, M. Urassa, R. Issingo, G. Mwaluko, S. Floyd, A. Nyondo, and A. Crampin. Hiv and mortality of mothers and children: evidence from cohort studies in uganda, tanzania, and malawi. *Epidemiology*, 16(3):275–80, 2005. Zaba, Basia Whitworth, Jimmy Marston, Milly Nakiyingi, Jessica Ruberantwari, Anthony Urassa, Mark Issingo, Raphaeli Mwaluko, Gabriel Floyd, Sian Nyondo, Andrew Crampin, Amelia 075886/Wellcome Trust/United Kingdom Research Support, Non-U.S. Gov't United States Epidemiology (Cambridge, Mass.) Epidemiology. 2005 May;16(3):275-80.

[32] Narayan Sastry. Family-level clustering of childhood mortality risk in northeast brazil. *Population Studies*, 51(3):245–261, 1997. 0032-4728 Article type: Full Length Article / Full publication date: Nov., 1997 (199711). / Copyright 1997 Population Investigation Committee.

[33] j. Yerushalmy, j.m. Bierman, d.h. Kemp, a. Connor, and f.e. French. Longitudinal studies of pregnancy on the island-of-kauai, territory-of-hawaii. *Am J Obst Gynecol*, 71(1):80–96, 1956. Article WJ763 English Times Cited:72 Cited References Count:23.

[34] L. D. E. Ikamari. Sibling mortality correlation in kenya. *Journal of Biosocial Science*, 32:265–278, 2000.

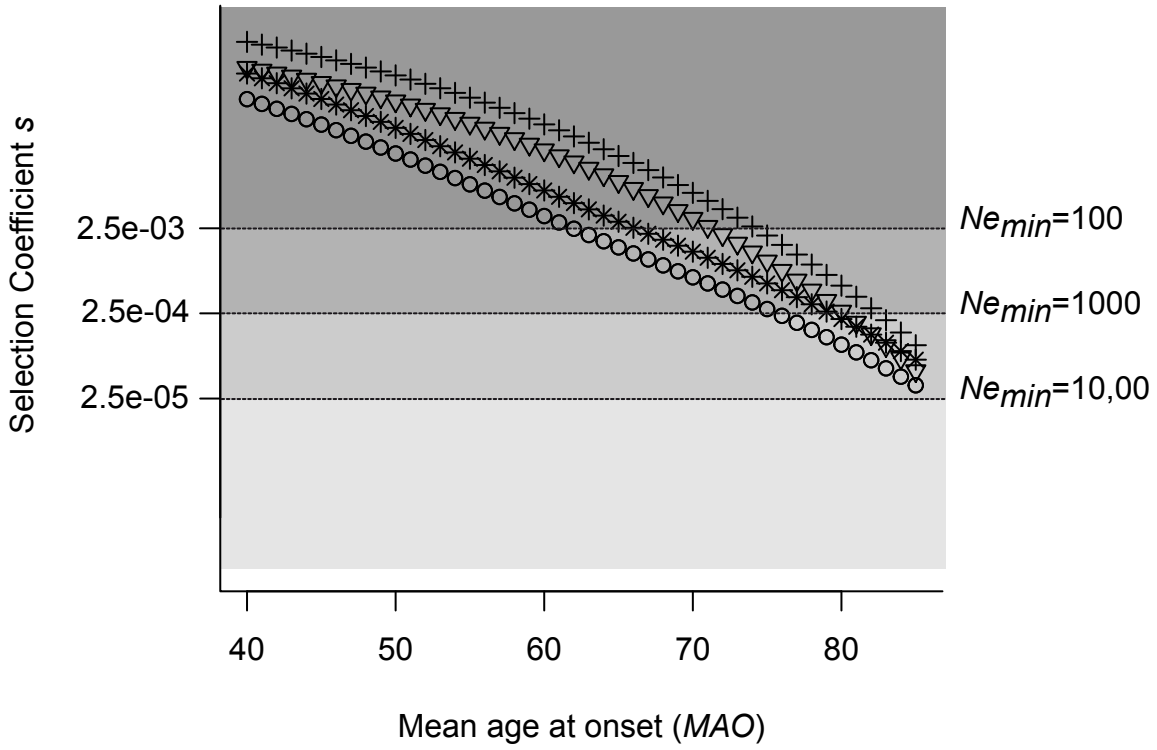
[35] B. Kuate Defo. Effect of infant feeding practices and birth spacing on infant and child survival: a reassessement from retrospective and prospective data. *Journal of Biosocial Science*, 29(303-326), 1997.

[36] Samuel J. Clark, Kathleen Kahn, Brian Houle, Adriane Arteche, Mark A. Collinson, Stephen M. Tollman, and Alan Stein. Young children's probability of dying before and after their mother's death: A rural south african population-based surveillance study. *PLoS Med*, 10(3):e1001409, 2013.

[37] R. Sear and R. Mace. Who keeps children alive? a review of the effects of kin on child survival. *Evolution and Human Behavior*, 29:1–18, 2008.

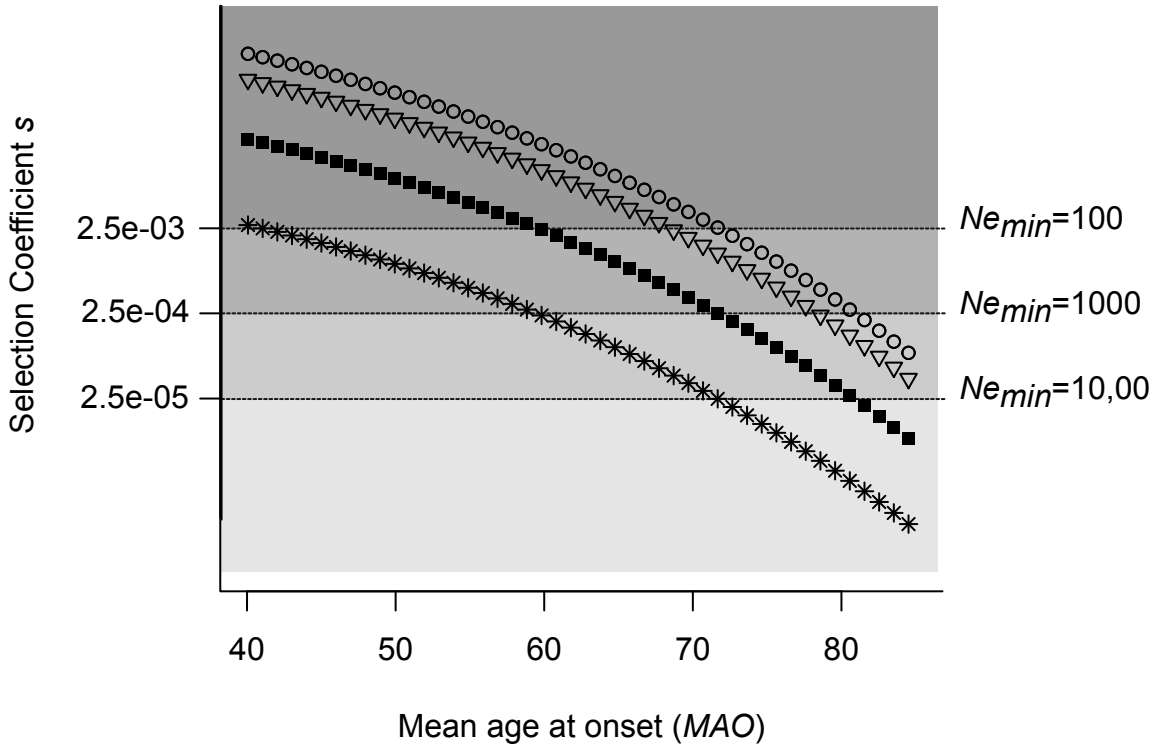
[38] Kai P. WillfÃŒhr and Alain Gagnon. Are stepparents always evil? parental death, remarriage, and child survival in demographically saturated krummhÃ¶rn (1720â1859) and expanding quÃ©bec (1670â1750). *Biodemography and Social Biology*, 59(2):191–211, 2013.

[39] B. I. Strassmann and W. M. Garrard. Alternatives to the grandmother hypothesis: a meta-analysis of the association between grandparental and grandchild survival in patrilineal populations. *Hum Nat*, 22(1-2):201–22, 2011. Strassmann, Beverly I Garrard, Wendy M HD 07480-02/HD/NICHD NIH HHS/United States Meta-Analysis Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S. United States Human nature (Hawthorne, N.Y.) Hum Nat. 2011 Jul;22(1-2):201-22. doi: 10.1007/s12110-011-9114-8.



Supplementary Figure 1 - Selection coefficients as a function of Mean Age at Disease Onset (*MAO*), First Age at Onset (*FAO*) being 20 years earlier, in the case of (**i**) an autosomal allele leading to disease in females only (circles), W^C of male carriers being replaced by W^{NC} in equation (7); (**ii**) an autosomal allele leading to disease in males only (triangles), W^C of female carriers being relaced by W^{NC} in equation (7); (**iii**) an allele carried by the mitochondrial chromosome (stars), canceling the male element of equation (7); and (**iv**) an allele carried by the Y-Chrom (crosses), canceling the male element of equation (7).

When variance in disease onset and all socio-cultural factors are accounted for there are no large difference in magnitude of selection on these alleles: the cross the s=2.5e-04 lines between MAO of ages 75-85. A differential of selection may however be expected for small population (of Ne between 100 and 1000) between alleles in the Y-Chromsome or leading to disease in males only (more puryfied) and alleles in the Mt-Chromsome or leading to disease in females only (less puryfied). This is because coefficient ofs eection s is a little less sensitive to (grand)maternal care than to male reproduction.



Supplementary Figure 2 - Selection coefficients as a function of Mean Age at Disease Onset (*MAO*), First Age at Onset (*FAO*) being 20 years earlier, in the case of an autosomal allele leading to disease cumuative penetrance at age 100 of 100% (circles), 50% (triangles), 10% (squares) and 1% (stars). In this case *FAO* is defined as the age at which of the cumulative distribution respectively reaches 1, 0.5, 0.1 and 0.01; the risk of disease onset being zero before this age.
Selection coefficient is a linear function of penetrance (estimated here by the cumuative risk at age 100), making of this later an obvious fundemmental parameter for estimating levels of selection.

Bibliography

Alfred J. Lotka (1925). Elements of Physical Biology. *Williams and Wilkins Company*, later call:435.

Allee, W. C. (1934). Recent studies in mass physiology. *Biological Reviews*, 9(1):1–48.

Angert, A. L., Kimball, S., Peterson, M., Huxman, T. E., and Venable, D. L. (2014). Phenotypic constraints and community structure: Linking trade-offs within and among species. *Evolution*, 68(11):3149–3165.

Ara, K. (1959). The Aggregation Problem in Input-Output Analysis. *Econometrica*, 27(2):257–262.

Arnoldi, W. E. (1951). The principle of minimized iteration in the solution of the matrix eigenvalue problem. *Quarterly of applied mathematics*, 9(1):17–29.

Aubry, L. M., Cam, E., Koons, D. N., Monnat, J. Y., and Pavard, S. (2011). Drivers of age-specific survival in a long-lived seabird: Contributions of observed and hidden sources of heterogeneity. *Journal of Animal Ecology*, 80(2):375–383.

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

Barrowclough, G. F. and Rockwell, R. F. (1993). Variance of Lifetime Reproductive Success: Estimation Based on Demographic Data. *American Naturalist*, 141(2):281–295.

Baudisch, A. (2005). Hamilton's indicators of the force of selection. *Proceedings of the National Academy of Sciences of the United States of America*, 102(23):8263–8268.

Beekink, E., van Poppel, F., and Liefbroer, A. C. (2002). Parental death and death of the child: common causes or direct effects? *When Dad Died: Individuals and Families Coping with Distress in Past Societies*, pages 234–260.

Bell, G. (1980). The costs of reproduction and their consequences. *The American Naturalist*, 116(1):45–76.

Bell, G. (1984). Evolutionary and Nonevolutionary Theories of Senescence. *The American Naturalist*, 124(4):600–603.

Bell, G. (2011). Foreword: Harvey's legacy. In Flatt, T. and Heyland, A., editors, *Mechanisms of Life History Evolution*, pages vi–xii. Oxford University Press, oxford bio edition.

Bell, G., Handford, P., and Dietz, C. (1977). Dynamics of An Exploited Population of Lake Whitefish (Coregonus clupeaformis). *Journal of the Fisheries Board of Canada*, 34(82):942–953.

Bentley, G. and Mace, R. (2009). *Substitute Parents: Biological and Social Perspectives on Alloparenting in Human Societies*, volume 3. Berghahn Books, ned - new edition.

Bernardo, J. (1996). The particular maternal effect of propagule size, especially egg size: patterns, models, quality of evidence and interpretations. *American Zoologist*, 36(2):216–236.

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.

Bjørnstad, O. N. and Hansen, T. F. (1994). Individual Variation and Population Dynamics. *Oikos*, 69(1):167.

Bongaarts, J. (1975). Why high birth rates are so low. *Popul. Dev. Rev.*, 1:289–296.

Boonekamp, J. J., Salomons, M., Bouwhuis, S., Dijkstra, C., and Verhulst, S. (2014). Reproductive effort accelerates actuarial senescence in wild birds: an experimental study. *Ecology Letters*, 17(5):599–605.

Borgerhoff Mulder, M. (2000). Optimizing offspring: the quantity-quality tradeoff in agropastoral Kipsigis. *Evol Hum Behav*, 21(6):391–410.

Borgerhoff Mulder, M. and Milton, M. (1985). Factors Affecting Infant Care in the Kipsigis. *J Anthrop Res*, 41(3):231–262.

Braendle, C., Heyland, A., and Flatt, T. (2011). *Integrating mechanistic and evolutionary analysis of life history variation*.

Butte, N. F. and King, J. C. (2005). Energy requirements during pregnancy and lactation. *Public Health Nutrition*, 8(7A):1010–1027.

Calow, P. (1979). The cost of reproduction - a physiological approach. *Biological Review*, 54:23–40.

Cameron, R. D. (1994). Reproductive pauses by female caribou. *Journal of Mammalogy*, 75(1):10–13.

Cant, M. A. and Johnstone, R. A. (2008). Reproductive conflict and the separation of reproductive generations in humans. *PNAS*, 105(14):5332–5336.

Caswell, H. (1978). A general formula for the sensitivity of population growth rate to changes in life history parameters. *Theoretical Population Biology*, 14(2):215–230.

Caswell, H. (1982a). Life history theory and the equilibrium status of population. *American Naturalist*, 120:317–339.

Caswell, H. (1982b). Optimal life histories and the age-specific costs of reproduction. *Journal of theoretical biology*.

Caswell, H. (1982c). Optimal life histories and the maximization of reproductive value: a general theorem for complex life cycles. *Ecology*, 63(5):1218–1222.

Caswell, H. (1984). Optimal life histories and age-specific cost of reproduction: two extensions. *Journal of Theoretical Biology*, 107:169–172.

Caswell, H. (1989). Analysis of life table response experiments I. Decomposition of effects on population growth rate. *Ecological Modelling*, 46(3-4):221–237.

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

Caswell, H. (2006). Applications of Markov chains in demography*. *MAM2006: Markov Anniversary Meeting. Raleigh, ...*, 02543:319–335.

Caswell, H. (2007). Sensitivity analysis of transient population dynamics. *Ecology Letters*, 10(1):1–15.

Caswell, H. (2008). Perturbation analysis of nonlinear matrix population models. *Demographic Research*, 18(3):59–116.

Caswell, H. (2009). Stage, age and individual stochasticity in demography. *Oikos*, 118(12):1763–1782.

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

Caswell, H. (2012). Matrix models and sensitivity analysis of populations classified by age and stage: a vec-permutation matrix approach. *Theoretical Ecology*, 5(3):403–417.

Caswell, H. (2014). A matrix approach to the statistics of longevity in heterogeneous frailty models. *Demographic Research*, 31(1):553–592.

Caswell, H. and John, A. M. (1992). From the individual to the population in demographic models. *Individual-based models and approaches in ecology. Springer US, 1992. 36-61*, pages 36–61.

Caswell, H., Neubert, M. G., and Hunter, C. M. (2011). Demography and dispersal: Invasion speeds and sensitivity analysis in periodic and stochastic environments. *Theoretical Ecology*, 4(4):407–421.

Caswell, H. and Real, L. (1987). An approach to the perturbation analysis of optimal life histories. *Ecology*, 68(4):1045–1050.

Caswell, H. and Salguero-Gómez, R. (2013). Age, stage and senescence in plants. *Journal of Ecology*, 101(3):585–595.

Caswell, H. and Sánchez Gassen, N. (2015). The sensitivity analysis of population projections. *Demographic Research*, 33(October):801–840.

Caswell, H., Takada, T., and Hunter, C. M. (2004). Sensitivity analysis of equilibrium in density-dependent matrix population models. *Ecology Letters*, 7(5):380–387.

Charlesworth, B. (1970). Selection in populations with overlapping generations. I. The use of Malthusian parameters in population genetics. *Theoretical Population Biology*, 1(3):352–370.

Charlesworth, B. (1980). *Evolution in age structured populations*.

Charlesworth, B. (1990). Optimization Models, Quantitative Genetics, and Mutation. *Evolution*, 44(3):520–538.

Charlesworth, B. (2000). Fisher, Medawar, Hamilton and the evolution of aging. *Genetics*, 156(3):927–931.

Charlesworth, B. and Leon, J. A. (1976). The relation of reproductive effort to age. *The American Naturalist*, 110(973):449–459.

Charlesworth, B. and Partridge, L. (1997). Ageing: Levelling of the grim reaper. *Current Biology*, 7(7):R440–R442.

Charnov, E. L. (1989). Phenotypic evolution under Fisher's Fundamental Theorem of Natural Selection. *Heredity*, 62(1):113–116.

Charnov, E. L. (1990). On evolution of age of maturity and the adult lifespan. *Journal of Evolutionary Biology*, 3(1-2):139–144.

Charnov, E. L. (2002). Reproductive effort , offspring size and benefit – cost ratios in the classification of life histories. *Evolutionary Ecology Research*, 4:749–758.

Charnov, E. L. and Berrigan, D. (1993). Why do female primates have such long lifespans and so few babies? or Life in the slow lane. *Evolutionary Anthropology*, 1(6):191–194.

Chen, H.-y., Zajitschek, F., and Maklakov, A. A. (2013). Why ageing stops: heterogeneity explains late-life mortality deceleration in nematodes. *Biology letters*, 9(5):20130217.

Chichilnisky, G. (1983). The transfer problem with three agents once again. *Journal of Development Economics*, 13(1-2):237–247.

Christians, J. K. (2000). Trade-offs between egg size and number in waterfowl: An interspecific test of the van Noordwijk and de Jong model. *Functional Ecology*, 14(4):497–501.

Clark, C. W. and Mangel, M. (2000). *Dynamic State Variable Models in Ecology*.

Clark, S. J., Kahn, K., Houle, B., Arteche, A., Collinson, M. A., Tollman, S. M., and Stein, A. (2013). Young Children's Probability of Dying Before and After Their Mother's Death: A Rural South African Population-Based Surveillance Study. *PLoS Med*, 10(3):e1001409.

Clutton-Brock, T. H. (1984). Reproductive Effort and Terminal Investment in Iteroparous Animals. *The American Naturalist*, 123(2):212.

Clutton-Brock, T. H. (1991). *The evolution of parental care*. Univ. of Princeton Press, Princeton, NJ.

Clutton-Brock, T. H., Albon, S., and Guinness, F. E. (1985). Parental invetment and sex differences in juvenile mortality in birds and mammals. *Nature*, 313:131–133.

Clutton-Brock, T. H., Albon, S. D., and Guinness, F. E. (1989). Fitness costs of gestation and lactation in wild mammals.

Clutton-Brock, T. H., Guinness, F. E., and Albon, S. D. (1983). The costs of reproduction to red deer hinds. *Journal of Animal Ecology*, 52:367–383.

Cochran, M. E. and Ellner, S. (1992). Simple methods for Calculating Age-Based Life History Parameters for Stage-Structured Populations. *Ecological Monographs*, 63(3):345–364.

Cody, M. L. (1966). A general theory of clutch size. *Evolution, 20, 174-194.*, 20(2):174–184.

Cohen, J. E. (1977). Ergodicity of age structure in populations with Markovian vital rates, III: Finite-state moments and growth rate; an illustration. *Advances in Applied Probability*, 9(03):462–475.

Cohen, J. E. (1979). Ergodic theorems in demography. 1(2).

Cole, L. C. (1954). The population consequences of life history phenomena. *The Quarterly review of biology*, 29(2):103–137.

Constantz, G. D. (1974). Reproductive Effort in Poeciliopsis occidentalis (Poeciliidae). *The Southwestern Naturalist*, 19(1):47.

Crow, J. F. and Kimura, M. (1970). *An Introduction to Population Genetics Theory.*

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

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

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

Cyrus Chu, C., Lee, R. D., Chu, C. Y., and Lee, R. D. (2006). The co-evolution of intergenerational transfers and longevity: An optimal life history approach. *Theoretical Population Biology*, 69(2):193–201.

Darwin, C. (1859). *On the Origin of the Species*, volume 5.

De-Camino-Beck, T., Lewis, M. A., De-Camino-Beck, T., and Lewis, M. A. (2008). On Net Reproductive Rate and the Timing of Reproductive Output. *The American Naturalist*, 172(1):128–139.

de Kroon, H., Plaisier, A., van Groenendael, J., and Caswell, H. (1986). Elasticity: The Relative Contribution of Demographic Parameters to Population Growth Rate. *Ecology*, 67(5):1427–1431.

Demetrius, L. (1969). The sensitivity of population growth rate to pertubations in the life cycle components. *Mathematical Biosciences*, 4(1-2):129–136.

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.

Demetrius, L. (1981). the Malthusian Parameter and the Effective Rate of Increase. *Journal Of Theoretical Biology*, 92(2):141–161.

Derosas, R. (2002). Fatherless families in 19th century Venice. *When Dad Died: Individuals and Families Coping with Distress in Past Societies*, pages 421–452.

Derosas, R. and Oris, M. (2002). *When Dad Died. Individuals and Families Coping with Family Stress in Past Societies.* Peter Lang, Bern.

Descamps, S., Gaillard, J. M., Hamel, S., and Yoccoz, N. G. (2016). When relative allocation depends on total resource acquisition: implication for the analysis of trade-offs. *Journal of Evolutionary Biology*, 29(9):1860–1866.

Dijkstra, C., Bult, A., Bijlsma, S., Daan, S., Meijer, T., and Zijlstra, M. (1990). Brood Size Manipulations in the Kestrel (Falco tinnunculus): Effects on Offspring and Parent Survival. *Journal of Animal Ecology*, 59(1):269–285.

Dobson, F. S. and Oli, M. K. (2007). Fast and slow life histories of mammals. *Ecoscience*, 14(3):292.

Drapeau, M. D., Gass, E. K., Simison, M. D., Mueller, L. D., and Rose, M. R. (2000). Testing the heterogeneity theory of late-life mortality plateaus by using cohorts of Drosophila melanogaster. *Experimental Gerontology*, 35(1):71–84.

Draper, P. and Hames, R. (2000). Birth order, sibling investment, and fertility among Ju/'Hoansi (!Kung). *Human Nature*, 11(2):117–156.

Drent, R. H. and Daan, S. (1980). The Prudent Parent: Energetic Adjustments in Avian Breeding 1). *Ardea*, 68:225–252.

Dublin, L. I. and Lotka, A. J. (1925). On the True Rate of Natural Increase. *Journal of the American Statistical Association*, 20(151):305–339.

Edward, D. A. and Chapman, T. (2011). Mechanisms underlying reproductive trade-offs: Costs of reproduction. In Flatt, T. and Heyland, A., editors, *Mechanisms of Life History Evolution*, chapter 11, pages 137–152. Oxford University Press, oxford bio edition.

Ellison, P. T. (1994). Extinction and descent. *Human Nature*, 5(2):155–165.

Emlen, J. M. (1970). Age Specificity and Ecological Theory. *Ecology*, 51(4):588.

Emlen, S. T. (1995). An evolutionary theory of the family. *Proceedings of the National Academy of Sciences*, 92(18):8092–8099.

Engelman, M., Caswell, H., and Agree, E. M. (2014). Why do lifespan variability trends for the young and old diverge? A perturbation analysis. *Demographic Research*, 30(1):1367–1396.

Engen, S., Bakke, O., and Islam, A. (1998). Demographic and Environmental Stochasticity-Concepts and Definitions. *Biometrics*, 54(3):840–846.

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

Engen, S., Lande, R., Saether, B.-E., and Dobson, F. S. (2009). Reproductive value and the stochastic demography of age-structured populations. *The American naturalist*, 174(6):795–804.

Engen, S., Lande, R., Sæther, B. E., and Festa-Bianchet, M. (2007). Using reproductive value to estimate key parameters in density-independent age-structured populations. *Journal of Theoretical Biology*, 244(2):308–317.

Engen, S., Lande, R., Sæther, B. E., and Weimerskirch, H. (2005b). Extinction in relation to demographic and environmental stochasticity in age-structured models. *Mathematical Biosciences*, 195(2):210–227.

Engen, S., Lande, R., Seæther, B. E., and Bregnballe, T. (2005c). Estimating the pattern of synchrony in fluctuating populations. *Journal of Animal Ecology*, 74(4):601–611.

Enright, N. J., Franco, M., and Silvertown, J. (1995). Comparing plant life histories using elasticity analysis : theimportance of life span and the number of life - cycle stages. *Oecologia*, 104(1):79–84.

Erikstad, K. E., Fauchald, P., Tveraa, T., and Steen, H. (1998). On the cost of reproduction in long-lived birds: The influence of environmental variability. *Ecology*, 79(5):1781–1788.

Euler, L. (1760). Recherches générales sur la mortalité et la multiplication. *Mémoires de l'Académie Royal des Sciences et Belles Lettres*, 16:144–164.

Felsenstein, J. (1971). Inbreeding and variance effective numbers in populations with overlapping generations. *Genetics*, 68(4):581–597.

Finch, C. E. and Kirkwood, T. B. L. (2000). *Chance, Development, and Aging*.

Fischer, B., Taborsky, B., and Dieckmann, U. (2009). Unexpected patterns of plastic energy allocation in stochastic environments. *The American naturalist*, 173(3):E108–20.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford University Press.

Fisher, W. D. (1958). Criteria for Aggregation in Input-Output Analysis. *The Review of Economics and Statistics*, 40(3):250.

Fisher, W. D. (1969). *Clustering and Aggregation in Economics*. The Johns Hopkins University Press.

Flatt, T. and Heyland, A. (2011). *Mechanisms of Life History Evolution*.

Frobenius, G. (1912). *Über Matrizen aus nicht negativen Elementen*. S.-B. Preuss Akad. Wiss. Berlin.

Gadgil, M. and Bossert, W. H. (1970). Life Historical Consequences of Natural Selection. *The American Naturalist*, 104(935):1–24.

Gagnon, A., Smith, K. R., Tremblay, M., Vézina, H., Paré, P.-P. P., and Desjardins, B. (2009). Is there a trade-off between fertility and longevity? A comparative study of women from three large historical databases accounting for mortality selection. *Am J Hum Biol*, 21(4):533–540.

Gaillard, J.-M., Pontier, D., Allainé, D., Lebreton, J.-D., Trouvilliez, J., and Clobert, J. (1989). An analysis of demographic tactics in birds and mammals. *Oikos*, 56(1):59–76.

Gaillard, J.-M., Yoccoz, N. G., Lebreton, J.-D., Bonenfant, C., Devillard, S., Loison, A., Pontier, D., and Allaine, D. (2005). Generation time: a reliable metric to measure life-history variation among mammalian populations. *The American naturalist*, 166(1):119–123; discussion 124–128.

Gantmacher, F. (1959). *The Theory of Matrices, Volume 2*. American Mathematical Soc.

Garnier, A., Gaillard, J.-M., Gauthier, D., and Besnard, A. (2015). What shapes fitness costs of reproduction in long-lived iteroparous species? A case study on the Alpine ibex. *Ecology*, 97(1):150806103644006.

Gavrilov, L. A. and Gavrilova, N. S. (2002). Evolutionary theories of aging and longevity. *TheScientificWorldJournal*, 2:339–356.

Geary, D. C. and Flinn, M. V. (2001). Evolution of Human Parental Behavior and the Human Family. *Parenting: Science and Practice*, 1-2(June):5–61.

Gems, D. and Partridge, L. (2001). Insulin/IGF signalling and ageing: Seeing the bigger picture.

Giaimo, S. (2014). Evolution of aging through reduced demographic stochasticity - an extension of the pleiotropy theory to finite populations. *Ecology and evolution*, 4(2):167–73.

Gibson, M. (2008). Does Investment in the Sexes Differ When Fathers Are Absent? *Human Nature*, 19(3):263–276.

Gillespie, D. O., Russell, A. F., and Lummaa, V. (2008). When fecundity does not equal fitness: evidence of an offspring quantity versus quality trade-off in pre-industrial humans. *Proc Biol Sci*, 275(1635):713–722.

Gillespie, D. O. S., Trotter, M. V., Krishna-Kumar, S., and Tuljapurkar, S. D. (2014). Birth-Order Differences Can Drive Natural Selection on Aging. *Evolution*, 68(3):886–892.

Gillespie, J. H. (1974). Nautural selection for within-generation variance in offspring number. *Genetics*, 76(3):601–6.

Gillespie, J. H. (1975). Natural selection for within-generations variance in offspring number II. Discrete haploid models. *Genetics*, 81(11):403–413.

Giske, J., Aksnes, D. L., and Førland, B. (1993). Variable generation times and Darwinian fitness measures. *Evolutionary Ecology*, 7(3):233–239.

Goodman, L. A. (1967). The probabilities of extinction for birth-and-death processes that are age-dependent or phase-dependent. *Biometrika*, 54(3-4):579–596.

Goodman, L. A. (1969). The analysis of population growth when the birth and death rates depend upon several factors. *Biometrics*, 25(4):659–681.

Goodman, L. A. (1971). On the sensitivity of the intrinsic growth rate to changes in the age-specific birth and death rates. *Theoretical Population Biology*, 2(3):339–354.

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

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

Gosling, L. M. (1986). Selective Abortion of Entire Litters in the Coypu: Adaptive Control of Offspring Production in Relation to Quality and Sex. *The American Naturalist*, 127(6):772–795.

Gould, S. J. and Lewontin, R. C. (1979). The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society B: Biological Sciences*, 205(1161):581–598.

Gross, M. R. (2005). The evolution of parental care. *Quarterly Review of Biology*, 80(1):37–45.

Gurven, M. (2004). Reciprocal altruism and food sharing decisions among Hiwi and Ache hunter-gatherers. *Behavioral Ecology and Sociobiology*, 56(4):366–380.

Gurven, M., Costa, M., Trumble, B., Stieglitz, J., Beheim, B., Rodreiquez, D. E., Hooper, P. L., and Kaplan, H. S. (2016). Costs of reproduction are minimal despite high fertility, mortality and active lifestyle. *PNAS submitted*, (April):1–18.

Gurven, M., Hill, K., and Jakugi, F. (2004). Why do foragers share and sharers forage? Explorations of social dimensions of foraging. In *Research in Economic Anthropology*, volume 23, pages 19–43.

Gurven, M., Stieglitz, J., Hooper, P. L., Gomes, C., and Kaplan, H. (2012). From the womb to the tomb: the role of transfers in shaping the evolved human life history. *Exp Gerontol*, 47(10):807–813.

Gustafsson, L. and Pärt, T. (1990). Acceleration of senescence in the collared flycatcher Ficedula albicollis by reproductive costs. *Nature*, 347:279–281.

Haldane, J. B. S. (1941). *New paths in genetics*.

Haldane, J. B. S. (1957). The cost of natural selection. *Journal of Genetics*, 55(3):511–524.

Hamel, S., Côté, S. D., Gaillard, J.-M., and Festa-Bianchet, M. (2009). Individual variation in reproductive costs of reproduction: high-quality females always do better. *Journal of Animal Ecology*, 78(1):143–151.

Hamel, S., Gaillard, J. M., Yoccoz, N. G., Loison, A., Bonenfant, C., and Descamps, S. (2010). Fitness costs of reproduction depend on life speed: Empirical evidence from mammalian populations.

Hamilton, W. (1964a). The genetical Evolution of Social Behavior II.

Hamilton, W. D. (1964b). The genetical evolution of social behavior. *J. Theoret. Biol.*, 7:1–16.

Hamilton, W. D. (1966). The moulding of senescence by natural selection. *Journal of theoretical biology*, 12(1):12–45.

Hammers, M., Richardson, D. S., Burke, T., and Komdeur, J. (2013). The impact of reproductive investment and early-life environmental conditions on senescence: Support for the disposable soma hypothesis. *Journal of Evolutionary Biology*, 26(9):1999–2007.

Hanski, I. (1999). Habitat Connectivity, Habitat Continuity, and Metapopulations in Dynamic Landscapes. *Oikos*, 87:209–219.

Harvey, G. (1651). *Exercitationes de generatione animalium*.

Hatori, H. (1966). On Markov chains with rewards. *Kodai Mathematical Seminar Reports*, 18(2):184–192.

Hawkes, K., OConnell, J. F., and Jones, N. G. B. (1997). Hadza women's time allocation, offspring provisioning, and the evolution of long postmenopausal life spans. *Current Anthropology*, 38(4):551–577.

Hayward, A. D., Nenko, I., and Lummaa, V. (2015). Early-life reproduction is associated with increased mortality risk but enhanced lifetime fitness in pre-industrial humans. *Proceedings of the Royal Society B: Biological Sciences*, 282(1804):20143053–20143053.

Hedrick, P. (2005). Large variance in reproductive success and the Ne/N ratio. *Evolution*, 59(7):1596–1599.

Heesterbeek, J. A. P. (2002). A brief history of R0 and a recipe for its calculation. *Acta Biotheorica*, 50(3):189–204.

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

Hendry, A. P. and Berg, O. K. (1999). Secondary sexual characters, energy use, senescence, and the cost of reproduction in sockeye salmon. *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, 77(11):1663–1675.

Heppell, S., Pfister, C., and de Kroon, H. (2000). Elasticity analysis in population biology: Methods and applications. *Ecology*, 81(3):605–606.

Hill, W. G. (1972). Effective size of populations with overlapping generations. *Theoretical Population Biology*, 3(3):278–289.

Hill, W. G. (1979). A note on effective population size with overlapping generations. *Genetics*, 92:317–322.

Hooley, D. E. (2000). Collapsed Matrices with (Almost) the Same Eigenstuff. *The College Mathematics Journal*, 31(4):297.

Houle, D. (1991). Genetic covariance of fitness correlates: what genetic correlations are made of and why it matters.

Houston, A., Clark, C., McNamara, J., and Mangel, M. (1988). Dynamic models in behavioural and evolutionary ecology.

Howard, R. A. (1960). *Dynamic Programming and Markov Processes*, volume 3.

Hunter, C. M. and Caswell, H. (2005). The use of the vec-permutation matrix in spatial matrix population models. *Ecological Modelling*, 188(1):15–21.

Hutchinson, G. E. (1948). Circular causal systems in ecology. *Annals of the New York Academy of Sciences*, 50(4):221–246.

Huxley, J. (1942). *Evolution: The Modern Synthesis*, volume 1943.

Ijiri, Y. (1968). The linear aggregation coefficient as the dual of the linear correlation coefficient. *Econometrica*, 36(2):252.

Ikamari, L. D. E. (2000). Sibling mortality correlation in Kenya. *Journal of Biosocial Science*, 32:265–278.

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

Jagers, P. and Nerman, O. (1984). The Growth and Composition of Branching Populations. *Advances in applied probability*, 16(2):221–259.

Jasienska, G. (2009). Reproduction and lifespan: Tradeoffs, overall energy budgets, intergenerational costs, and costs neglected by research. *American Journal of Human Biology*, 21(4):524–532.

Jelliffe, D. B. and Maddocks, I. (1964). Notes on ecologic malnutrition in the new guinea highlands. *Clin Pediatr (Phila)*, 3:432–438.

Ji, T., Xu, J.-J., and Mace, R. (2014). Intergenerational and Sibling Conflict Under Patrilocality. *Human Nature*, 25(1):66–79.

Joffe, A. and Waugh, W. A. O. (1982). Exact distributions of kin numbers in a Galton-Watson process. *Journal of Applied Probability*, 19(4):767–775.

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.

Jones, J. H. (2007). Notes on R0. *Stanford University, Department of Anthropological Sciences*, page 19.

Jørgensen, C., Enberg, K., and Mangel, M. (2016). Modelling and interpreting fish bioenergetics: A role for behaviour, life-history traits and survival trade-offs. *Journal of Fish Biology*, 88(1):389–402.

Kaplan, H. S. and Robson, A. J. (2002). The emergence of humans: The coevolution of intelligence and longevity with intergenerational transfers. *PNAS*, 99(15):10221–10226.

Kengeri, S. S., Maras, A. H., Suckow, C. L., Chiang, E. C., and Waters, D. J. (2013). Exceptional longevity in female rottweiler dogs is not encumbered by investment in reproduction. *Age*, 35(6):2503–2513.

Kent, E. B. (1981). Life history responses to resource variation in a sessile predator, the ciliate protozoan Tokophrya lemnarum Stein. *Ecology (Washington D C)*, 62(2):296–302.

Keyfitz, N. and Caswell, H. (2005). *Applied Mathematical Demography*. Statistics for Biology and Health. Springer-Verlag New York, New York.

Keyfitz, N. and Murphy, E. M. (1967). Matrix and multiple decrement in population analysis. *Biometrics*, 23(3):485–503.

Kimura, M. (1958). On the change of population fitness by natural selection. *Heredity*, 12(240):143–167.

King, E. G., Roff, D. A., and Fairbairn, D. J. (2011). Trade-off acquisition and allocation in Gryllus firmus: A test of the Y model. *Journal of Evolutionary Biology*, 24(2):256–264.

Kirkwood, T. B. and Holliday, R. (1979). The evolution of ageing and longevity. *Proceedings of the Royal Society of London. Series B, Biological sciences*, 205(1161):531–46.

Kirkwood, T. B. L. (1977). Evolution of ageing. *Nature*, 170:201–4.

Kirkwood, T. B. L. and Rose, M. R. (1991). Evolution of Senescence: Late Survival Sacrificed for Reproduction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 332(1262):15–24.

Koivula, M., Koskela, E., Mappes, T., and Oksanen, T. A. (2003). Cost of Reproduction in the Wild : Manipulation of Reproductive Effort in the Bank Vole. *Ecology*, 84(2):398–405.

Kramer, K. L. (2005). Children's Help and the Pace of Reproduction: Cooperative Breeding in Humans. *Evolutionary Anthropology*, 14(6):224–237.

Kramer, K. L. (2009). Does it take a family to raise a child? Cooperative breeding and the contributions of Maya siblings, parents and older adults in raising children. In *Substitute Parents: Biological and Social Perspective on Alloparenting Across Human Societies*, pages 77–99.

Kuate Defo, B. (1997). Effect of infant feeding practices and birth spacing on infant and child survival: a reassessement from retrospective and prospective data. . *Journal of Biosocial Science*, 29(303-326).

Lack, D. (1947). The Significance of Clutch-size. *Ibis*, 89(2):302–352.

Lack, D. (1954). The evolution of reproductive rates. In et al. Huxley, J., editor, *Evolution as a Process*, page p. 367. Allen and Unwin.

Lahdenperä, M., Gillespie, D. O. S., Lummaa, V., and Russell, A. F. (2012). Severe intergenerational reproductive conflict and the evolution of menopause. *Ecology Letters*, 15(11):1283–1290.

Lahdenpera, M., Lummaa, V., Helle, S., Tremblay, M., and Russell, A. F. (2004). Fitness benefits of prolonged post-reproductive lifespan in women. *Nature*, 428(6979):178–181.

Lambert, J. H. (1772). *Beyträge zum Gebrauche der Mathematik und deren Anwendung, Volume 3*.

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

Lande, R., Engen, S., and Saether, B.-E. (2003). *Stochastic population dynamics in ecology and conservation*.

Lande, R. and Orzack, S. H. (1988). Extinction dynamics of age-structured populations in a fluctuating environment. *Proceedings of the National Academy of Sciences of the United States of America*, 85(19):7418–7421.

Larke, A. and Crews, D. E. (2006). Parental investment, late reproduction, and increased reserve capacity are associated with longevity in humans. *J Physiol Anthropol*, 25(1):119–131.

Law, R. (1979). Optimal Life Histories Under Age-Specific Predation. *The American Naturalist*, 114(3):399–417.

Law, R. (1983). A model for the dynamics of a plant population containing individuals clasified by age and size.

Lawson, D. W. and Mace, R. (2011). Parental investment and the optimization of human family size. *Phil T Royal Soc B*, 366(1563):333–343.

Le Bras, H. (1970). Equilibre et Croissance de Populations Soumises a des Migrations. *Theoretical Population Biology*, 121(t 971):100–121.

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

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

Lebreton, J. D. (1996). Demographic models for subdivided populations: the renewal equation approach. *Theoretical Population Biology*, 49(3):291–313.

Lee, C. T., Puleston, C. O., and Tuljapurkar, S. (2009). Population and prehistory III: Food-dependent demography in variable environments. *Theoretical Population Biology*, 76(3):179–188.

Lee, C. T. and Tuljapurkar, S. (2008). Population and prehistory I: Food-dependent population growth in constant environments. *Theoretical Population Biology*, 73(4):473–482.

Lee, R. D. (2003). Rethinking the evolutionary theory of aging: transfers, not births, shape senescence in social species. *Proceedings of the National Academy of Sciences of the United States of America*, 100(16):9637–9642.

Lee, R. D. (2008). Sociality, selection, and survival: Simulated evolution of mortality with intergenerational transfers and food sharing. *PNAS*, 105(20):7124–7128.

Lefkovitch, L. P. (1965). The Study of Population Growth in Organisms Grouped by Stages. *Biometrics*, 21(1):1–18.

Lemaître, J.-F., Berger, V., Bonenfant, C., Douhard, M., Gamelon, M., Plard, F., and Gaillard, J.-M. (2015). Early-late life trade-offs and the evolution of ageing in the wild. *Proceedings. Biological sciences / The Royal Society*, 282(1806):20150209.

Leontief, W. (1936). *Note on the Pure Theory of Capital Transfer*, chapter 8, pages 84–91. McGraw-Hill.

Leontief, W. (1986). An alternative to aggregation in input-output analysis and national accounts. In *Input-Output Economics*, volume 49, pages 41–54.

Leontief, W. W. (1951). *Input-output economics*, volume 2. JSTOR.

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

Lessells, C. M. (1991). The evolution of life-histories. In Krebs, J. R. and Davies, N. B., editors, *Behavioural ecology: an evolutionary approach*, pages 32–68. Blackwell, Oxford, UK.

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

Lewontin, R. C. (1965). Selection for colonizing ability. In *The genetics of colonizing species*, pages 77–94.

Li, C. and Schreiber, S. (2006). On dispersal and population growth for multistate matrix models. *Linear Algebra and its Applications*, 418(2-3):900–912.

Li, T. and Anderson, J. J. (2009). The vitality model: A way to understand population survival and demographic heterogeneity. *Theoretical Population Biology*, 76(2):118–131.

Lindner, A. B. and Demarez, A. (2009). Protein aggregation as a paradigm of aging. *Biochim Biophys Acta*, 1790(10):980–996.

Little, B. B. (1989). Gestation Length, Metabolic-Rate, and Body and Brain Weights in Primates - Epigenetic Effects. *American Journal of Physical Anthropology*, 80(2):213–218.

Lopes, T. C. and Neder, H. D. (2017). Sraffa, Leontief, Lange: The political economy of input–output economics. *EconomiA*, 18(2):192–211.

Lorenzini, A., Stamato, T., and Sell, C. (2011). The disposable soma theory revisited: Time as a resource in the theories of aging. *Cell Cycle*, 10(22):3853–3856.

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

Lush, J. L. (1937). *Animal breeding plans*.

Lycett, J. E., Dunbar, R. I., and Voland, E. (2000). Longevity and the costs of reproduction in a historical human population. *Proc Biol Sci*, 267(1438):31–35.

Mace, R. (2000). Evolutionary ecology of human life history. *Animal Behavior*, 59(1):1–10.

Mace, R. (2013). Cooperation and conflict between women in the family. *Evolutionary Anthropology: Issues, News, and Reviews*, 22(5):251–258.

Mace, R. and Alvergne, A. (2012). Female reproductive competition within families in rural Gambia. *Proceedings of the Royal Society B: Biological Sciences*.

Mahy, M. and Zaba, B. (2003). Measuring Child Mortality in AIDS-Affected Countries. Technical report, New York.

Malthus, T. R. (1798). An Essay on the Principle of Population. page 273.

Mangel, M. and Clark, C. W. (1986). Towards a unified foraging theory. *Ecology*, 67(5):1127–1138.

Martin, R. D. (2007). The evolution of human reproduction: A primatological perspective. *American Journal of Physical Anthropology*, (45):59–84.

Mathews, P. and Sear, R. (2013). Family and Fertility: Kin Influence on the Progression to a Second Birth in the British Household Panel Study. *PLoS ONE*, 8(3).

Maynard Smith, J. (1958). The effects of temperature and of egg-laying on the longevity of Drosophila subobscura. *Journal of Experimental Biology*, 35(4):832–842.

McGraw, J. B. and Caswell, H. (1996). Estimation of Individual Fitness from Life-History Data. *The American Naturalist*, 147(1):47.

McGraw, J. B. and Caswell, H. (1997). Estimation of individual fitness from life history data.

McNamara, J. M. and Houston, A. I. (1986). The Common Currency for Behavioral Decisions. *The American Naturalist*, 127(3):358.

McNamara, J. M. and Houston, A. I. (1996). State-dependent life histories. *Nature*, 380(6571):215–221.

Medawar, P. B. (1952). An unsolved problem of biology.

Mendel, G. (1865). Versuche uber Pflanzenhybriden. *Verh. natf. ver. Brunn. abh.*, 3:S—-3.

Metcalf, C. J. E. (2016). Invisible Trade-offs: Van Noordwijk and de Jong and Life-History Evolution. *American Naturalist*, pages 000–000.

Metcalf, C. J. E. and Pavard, S. (2007). Why evolutionary biologists should be demographers. *Trends in Ecology and Evolution*, 22(4):205–12.

Min, K.-J., Lee, C.-K., and Park, H.-N. (2012). The lifespan of Korean eunuchs. *Current Biology*, 22(18):R792–R793.

Missov, T. I. and Vaupel, J. W. (2015). Mortality Implications of Mortality Plateaus. *SIAM Review*, 57(1):61–70.

Mode, C. J., Jacobson, M. E., and Pickens, G. T. (1987). A Stochastic Population Projection System Based on General Age-Dependent Branching-Processes. *JOURNAL OF APPLIED PROBABILITY*, 24(1):1–13.

Morimoto, Y. (1970). On Aggregation Problems in Input-Output Analysis. *The Review of Economic Studies*, 37(1):119–126.

Morris, W., Pfister, C., Tuljapurkar, S., Haridas, C., Boggs, C., Boyce, M., Bruna, E., Church, D., Coulson, T., Doak, D., Forsyth, S., Gaillard, J.-M., Horvitz, C., Kalisz, S., Kendall, B., Knight, T., Lee, C. T., and Menges, E. (2008). Longevity can buffer plant and animal populations against changing climate variability. *Ecology*, 89(1):19–25.

Mulder, M. B., Bowles, S., Hertz, T., Bell, A., Beise, J., Clark, G., Fazzio, I., Gurven, M., Hill, K., Hooper, P. L., Irons, W., Kaplan, H., Leonetti, D., Low, B., Marlowe, F., McElreath, R., Naidu, S., Nolin, D., Piraino, P., Quinlan, R., Schniter, E., Sear, R., Shenk, M., Smith, E. A., von Rueden, C., Wiessner, P., Borgerhoff-Mulder, M., Bowles, S., Hertz, T., Bell, A., Beise, J., Clark, G., Fazzio, I., Gurven, M., Hill, K., Hooper, P. L., Irons, W., Kaplan, H., Leonetti, D., Low, B., Marlowe, F., McElreath, R., Naidu, S., Nolin, D., Piraino, P., Quinlan, R., Schniter, E., Sear, R., Shenk, M., Smith, E. A., von Rueden, C., and Wiessner, P. (2009). Intergenerational Wealth Transmission and the Dynamics of Inequality in Small-Scale Societies. *Science*, 326(5953):682–688.

Murray, B. G. (1992). the Evolutionary Significance of Lifetime Reproductive Success. *The Auk*, 109(1):167–172.

Nilsen, E. B., Gaillard, J. M., Andersen, R., Odden, J., Delorme, D., Van Laere, G., and Linnell, J. D. (2009). A slow life in hell or a fast life in heaven: Demographic analyses of contrasting roe deer populations. *Journal of Animal Ecology*, 78(3):585–594.

Nitsch, A., Faurie, C., and Lummaa, V. (2013). Are elder siblings helpers or competitors? Antagonistic fitness effects of sibling interactions in humans. *Proc Roy Soc B*, 280(1750).

Nomura, T. (1996). Effective size of selected populations with overlapping generations. *Journal Of Animal Breeding And Genetics-Zeitschrift Fur Tierzuchtung Und Zuchtungsbiologie*, 113(1-6):1–16.

Nur, N. (1984). The Consequences of Brood Size for Breeding Blue Tits II. Nestling Weight, Offspring Survival and Optimal Brood Size. *The Journal of Animal Ecology*, 53(2):497.

Oli, M. K. and Dobson, F. S. (2003). The Relative Importance of Life-History Variables to Population Growth Rate in Mammals: Cole's Prediction Revisited. *American Naturalist*, 161(161):422–440.

Orell, M. and Belda, E. J. (2002). Delayed cost of reproduction and senescence in the willow tit Parus montanus. *Journal of Animal Ecology*, 71(1):55–64.

Orton, J. H. (1929). Reproduction and death in invertebrates and fishes. *Nature*, 3088(123):14–15.

Parker, G. and Maynard Smith, J. (1990). Optimality theory in evolutionary biology.

Parker, S. T. (1977). Piaget's sensorimotor period series in an infant macaquea: model for comparing non stereotyped behavior and intelligence in human and nonhuman primates. In P., C.-S., editor, *Prim. Biosoc. Dev.* Garland, New York.

Partridge, L. (1992). Measuring reproductive costs.

Partridge, L., Sibly, R., Beverton, R. J. H., and Hill, W. G. (1991). Constraints in the Evolution of Life Histories. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 332(1262):3–13.

Pasteur, L. (1862). *Mémoire sur les corpuscules organisés qui existent dans l'atmosphère: examen de la doctrine des générations spontanées.*

Pavard, S. and Branger, F. (2012). Effect of maternal and grandmaternal care on population dynamics and human life-history evolution: A matrix projection model. *Theoretical Population Biology*, 82(4):364–376.

Pavard, S., Gagnon, A., Desjardins, B., and Heyer, E. (2005). Mother's death and child survival: The case of early Quebec. *Journal of Biosocial Science*, 37(2):209–227.

Pavard, S., Koons, D. N., and Heyer, E. (2007a). the Influence of Maternal Care in Shaping Human Survival and Fertility. *Evolution*, 61(12):2801–2810.

Pavard, S., Metcalf, C. J. E., and Heyer, E. (2008). Senescence of reproduction may explain adaptive menopause in humans: A test of the "Mother" hypothesis. *American Journal of Physical Anthropology*, 136(2):194–203.

Pavard, S., Sibert, A., and Heyer, E. (2007b). the Effect of Maternal Care on Child Survival: a Demographic, Genetic, and Evolutionary Perspective. *Evolution*, 61(5):1153–1161.

Pearl, R. and Reed, L. J. (1920). On the rate of growth of the population of the United States since 1790 and its mathematical representation. *Proceedings of the National Academy of Sciences*, 6(6):275–288.

Peccei, J. S. (1995). The origin and evolution of menopause: The altriciality-lifespan hypothesis. *Ethology and Sociobiology*, 16(5):425–449.

Peck, S. L. (2004). Simulation as experiment: a philosophical reassessment for biological modeling. *Trends in Ecology & Evolution*, 19(10):530–534.

Perron, O. (1907). Zur Theorie der Matrices. *Mathematische Annalen 1*, 64(1):248–263.

Pianka, E. R. and Parker, W. S. (1975). Age-Specific Reproductive Tactics. *American Naturalist*, 109(968):453–464.

Plard, F., Bonenfant, C., Delorme, D., and Gaillard, J. (2012). Modeling reproductive trajectories of roe deer females: Fixed or dynamic heterogeneity? *Theoretical Population Biology*, 82(4):317–328.

Pollak, R. A. (1990). Two-sex demographic models. *Journal of political economy*, 98(2):399–420.

Proaktor, G., Coulson, T., and Milner-Gulland, E. J. (2008). The demographic consequences of the cost of reproduction in ungulates. *Ecology*, 89(9):2604–2611.

Promislow, D. E. L. and Harvey, P. H. (1990). Living fast and dying young: A comparative analysis of life-history variation among mammals. *Journal of Zoology*, 220(3):417–437.

Puleston, C. O. and Tuljapurkar, S. (2008). Population and prehistory II: Space-limited human populations in constant environments. *Theoretical Population Biology*, 74(2):147–160.

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

Pullum, T. W. and Wolf, D. A. (1991). Correlations between frequencies of kin. *Demography*, 28(3):391–409.

Reeves, J. H. (1987). Projection of Number of Kin. In *Family Demography: Methods and their Application*, pages 228–248.

Reznick, D., Nunney, L., and Tessier, A. (2000). Big houses, big cars, superfleas and the costs of reproduction. *Trends in Ecology and Evolution*, 15(10):421–425.

Ricklefs, R. E. and Cadena, C. D. (2007). Lifespan is unrelated to investment in reproduction in populations of mammals and birds in captivity. *Ecology Letters*, 10(10):867–872.

Robins, C. and Conneely, K. N. (2014). Testing evolutionary models of senescence: traditional approaches and future directions.

Rockwell, R. and Barrowclough, G. (1995). Effective population size and lifetime reproductive success.

Rodríguez, J. A., Marigorta, U. M., Hughes, D. A., Spataro, N., Bosch, E., and Navarro, A. (2017). Antagonistic pleiotropy and mutation accumulation influence human senescence and disease. *Nature Ecology & Evolution*, 1:0055.

Roff, D. A. (1992). *The evolution of life histories: theory and analysis*, volume 3.

Roff, D. A. and Fairbairn, D. J. (2007). The evolution of trade-offs: Where are we?

Rogers, A. (1966). The Multiregional Matrix Growth Operator and the Stable Interregional Age Structure. *Demography*, 3(2):537–544.

Rogers, A. (1969). On Perfect Aggregation in the Matrix-Survival Model of Interregional Population Growth. *Journal of Regional Science*, 9(3):417–424.

Rogers, A. (1974). The Multiregional Net Maternity Function and Multiregional Stable Growth. *Demography*, 11(3):473–481.

Rogers, A. (1980). Introduction to multistate mathematical demography. *Environment and Planning A*, 12(5):489–498.

Ronsmans, C. (1995). Patterns of Clustering of Child Mortality in a Rural Area of Senegal. *Population Studies*, 49(3):443–461.

Ronsmans, C., Chowdhury, M. E., Dasgupta, S. K., Ahmed, A., and Koblinsky, M. (2010). Effect of parent's death on child survival in rural Bangladesh: a cohort study. *The Lancet*, 375(9730):2024–2031.

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

Salguero-Gómez, R. and de Kroon, H. (2010). Matrix projection models meet variation in the real world. *Journal of Ecology*, 98(2):250–254.

Salt, G. (1936). Experimental Studies in Insect Parasitism: IV. The Effect of Superparasitism on Populations of Trichogramma evanescens. *The Journal of Experimental Biology*, 13(3):363–362.

Santos, E. S. A. and Nakagawa, S. (2012). The costs of parental care: A meta-analysis of the trade-off between parental effort and survival in birds. *Journal of Evolutionary Biology*, 25(9):1911–1917.

Sastry, N. (1997). Family-Level Clustering of Childhood Mortality Risk in Northeast Brazil. *Population Studies*, 51(3):245–261.

Schaffer, W. M. (1974). Selection for optimal life histories: the effects of age structure. *Ecology*, 55(2):291–303.

Sear, R. (2008). Kin and child survival in rural Malawi : AAAAre matrilineal kin always beneficial in a matrilineal society? *Human Nature*, 19(3):277–293.

Sear, R. and Mace, R. (2008). Who keeps children alive? A review of the effects of kin on child survival. *Evolution and Human Behavior*, 29(1):1–18.

Sear, R., Mace, R., and McGregor, I. A. (2003). The effects of kin on female fertility in rural Gambia. *Evolution and Human Behavior*, 24(1):25–42.

Sear, R., Steele, F., McGregor, A. A., and Mace, R. (2002). The effects of kin on child mortality in rural Gambia. *Demography*, 39(1):43–63.

Sharpe, F. R. and Lotka, A. J. (1911). A problem in age distribution. *Philosophical Magazine*, 21:435–438.

Shefferson, R. P., Jones, O. R., and Salguero-Gómez, R. (2017). Introduction: Wilting Leaves and Rotting Branches: Reconciling Evolutionary Perspectives on Senescence. In Shefferson, R. P., Jones, O. R., and Salguero-Gomez, R., editors, *The Evolution of Senescence in the Tree of Life*, pages 1–20. Cambridge University Press, Cambridge.

Shpak, M. (2005). Evolution of variance in offspring number: The effects of population size and migration. *Theory in Biosciences*, 124(1):65–85.

Shpak, M. (2007). Selection against demographic stochasticity in age-structured populations. *Genetics*, 177(4):2181–2194.

Shpak, M., Orzack, S. H., and Barany, E. (2013). The influence of demographic stochasticity on evolutionary dynamics and stability. *Theoretical Population Biology*, 88:47–56.

Sibly, R. and Calow, P. (1984). Direct and absorption costing in the evolution of life cycles. *Journal of Theoretical Biology*, 111(3):463–473.

Sibly, R. M. and Calow, P. (1983). An integrated approach to life cycle evolution using selective landscapes. *Journal of Theoretical Biology*, 102:527–547.

Simon, H. A. and Ando, A. (1961). Aggregation of Variables in Dynamic Systems. *Econometrica*, 29(2):111–138.

Smith, C. C. and Fretwell, S. D. (1974). The optimal balance between size and number of offspring. *American Naturalist*, 108(962):499–506.

Smith, J. E. (1987). The Computer Simulation of Kin Sets and Kin Counts. In *Family Demography: Methods and their Application1*, pages 249–266.

Sparks, C. S., Wood, J. W., and Johnson, P. L. (2013). Infant mortality and intra-household competition in the Northern Islands of Orkney, Scotland, 1855-2001. *American Journal of Physical Anthropology*, 151(2):191–201.

Spitze, K. (1991). Chaoborus Predation and Life-History Evolution in Daphnia pulex: Temporal Pattern of Population Diversity, Fitness, and Mean Life History. *Evolution*, 45(1):82–92.

Stearns, S. C. (1983). The influence of size and phylogeny on patterns of cavariation among life-history traits in the mammals. *Oikos*, 41(2):173–187.

Stearns, S. C. (1989a). The evolutionary significance of phenotypic plasticity. *BioScience*, 39(7):436– 445.

Stearns, S. C. (1989b). Trade-offs in life-history evolution. *Functional Ecology*, 3(1974):259–268.

Stearns, S. C. (1992). *The evolution of life histories*, volume 1.

Stearns, S. C. and Koella, J. C. (1986). The Evolution of Phenotypic Plasticity in Life-History Traits: Predictions of Reaction Norms for Age and Size at Maturity. *Evolution*, 40(5):893–913.

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

Steiner, U. K., Tuljapurkar, S., and Coulson, T. (2014). Generation Time, Net Reproductive Rate, and Growth in Stage-Age-Structured Populations. *The American Naturalist*, 183(6):771–783.

Steiner, U. K., Tuljapurkar, S., and Orzack, S. H. (2010). Dynamic heterogeneity and life history variability in the kittiwake. *Journal of Animal Ecology*, 79(2):436–444.

Stenseth, N. C. (1984). Fitness, Population Growth Rate and Evolution in Plant-Grazer Systems: A Reply to Nur. *Oikos*, 42(3):414.

Stephenson, A. G. (1980). Fruit set, herbivory, fruit reduction, and the fruiting strategy of Catalpa speciosa (Bignoniaceae). *Ecology*, 61(1):57–64.

Stott, I., Townley, S., and Hodgson, D. J. (2011). A framework for studying transient dynamics of population projection matrix models. *Ecology Letters*, 14(9):959–970.

Strassmann, B. I. and Garrard, W. M. (2011). Alternatives to the grandmother hypothesis: a meta-analysis of the association between grandparental and grandchild survival in patrilineal populations. *Hum Nat*, 22(1-2):201–222.

Tait, D. E. N. (1980). Abandonment as a reproductive tactic - The example of grizzly bears. *The American Naturalist*, 115(6):800–808.

Tarín, J. J., Gómez-Piquer, V., García-Palomares, S., García-Pérez, M. a., and Cano, A. (2014). Absence of long-term effects of reproduction on longevity in the mouse model. *Reproductive Biology and Endocrinology*, 12:84.

Tavecchia, G., Coulson, T., Morgan, B. J. T., Pemberton, J. M., Pilkington, J. C., Gulland, F. M. D., and Clutton-Brock, T. H. (2005). Predictors of reproductive cost in female Soay sheep. *Journal of Animal Ecology*, 74(2):201–213.

Taylor, H. M., Gourley, R. S., Lawrence, C. E., and Kaplan, R. S. (1974). Natural selection of life history attributes: An analytical approach. *Theoretical Population Biology*, 5(1):104–122.

Taylor, P. D., Wild, G., and Gardner, A. (2007). Direct fitness or inclusive fitness: How shall we model kin selection? *Journal of Evolutionary Biology*, 20(1):301–309.

Taylor, P. P. D. (1990). Allele-frequency change in a class-structured population. *American Naturalist*, 135(1):95–106.

Thomas, F., Teriokhin, A. T., De Meeûs, F., Guegan, J. F., Renaud, F., De Meeus, T., and Guegan, J. F. (2000). Human longevity at the cost of reproductive success: evidence from global data. *Journal of Evolutionary Biology*, 13(3):409–414.

Tinkle, D. W. (1969). The concept of reproductive effort and its relation to the evolution of life histories of lizards. *The American Naturalist*, 103(933):501.

Tuljapurkar, S. (1985). Population dynamics in variable environments. VI. Cyclical environments. *Theoretical Population Biology*, 28(1):1–17.

Tuljapurkar, S. (1986a). Demography in stochastic environments I. Exact distributions of age structure. *Journal of mathematical biology*, 19(3):335–350.

Tuljapurkar, S. (1986b). Demography in stochastic environments II. Growth and convergence rates. *Journal of Mathematical Biology*, 24(5):569–581.

Tuljapurkar, S. (1989). An uncertain life: demography in random environments. *Theoretical population biology*, 35(3):227–294.

Tuljapurkar, S. (1990a). Delayed reproduction and fitness in variable environments. *Proceedings of the National Academy of Sciences of the United States of America*, 87:1139–1143.

Tuljapurkar, S. (1990b). *Population Dynamics in Variable Environments*, volume 85 of *Lecture Notes in Biomathematics*. Springer Berlin Heidelberg, Berlin, Heidelberg.

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

Tuljapurkar, S. (2011). Demography as the Human Story. *Population and development review*, 37(1):166–171.

Tuljapurkar, S., Gaillard, J.-M., and Coulson, T. (2009). From stochastic environments to life histories and back. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364(1523):1499–1509.

Tuljapurkar, S., Horvitz, C. C., and Pascarella, J. B. (2003). The many growth rates and elasticities of populations in random environments. *The American naturalist*, 162(4):489–502.

Tuljapurkar, S. and Orzack, S. H. (1980). Population dynamics in variable environments I. Long-run growth rates and extinction. *Theoretical Population Biology*, 18(3):314–342.

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

Tuljapurkar, S. D. (1982a). Population dynamics in variable environments II. Correlated environments, sensitivity analysis and dynamics. *Theoretical Population Biology*, 21:114–140.

Tuljapurkar, S. D. (1982b). Population dynamics in variable environments III. Evolutionary dynamics of r-selection. *Theoretical Population Biology*, 21(1):141–165.

Tuljapurkar, S. D. (1982c). Why use population entropy? It determines the rate of convergence. *Journal of Mathematical Biology*, 13(3):325–337.

Tuljapurkar, S. D., Puleston, C. O., and Gurven, M. D. (2007). Why Men Matter: Mating Patterns Drive Evolution of Human Lifespan. *PLoS ONE*, 2(8):e785.

Usher, M. (1969). A Matrix Model for Forest Management. *Biometrics*, 25(2):309–315.

Usher, M. B. (1966). A Matrix Approach to the Management of Renewable Resources, with Special Reference to Selection Forests. *The Journal of Applied Ecology*, 3(2):355–367.

van Daalen, S. and Caswell, H. (2015). Lifetime reproduction and the second demographic transition: Stochasticity and individual variation. *Demographic Research*, 33(1):561–588.

van den Berghe, E. P. (1992). Parental care and the cost of reproduction in a Mediterranean fish. *Behav.Ecol.Sociobiol.*, 30:373–378.

van Groenendael, J., de Kroon, H., Kalisz, S., and Tuljapurkar, S. (1994). Loop analysis: Evaluating life history pathways in population projection matrices. *Ecology*, 75(8):2410–2415.

Van Imhoff, E. (1992). A General Characterization of Consistency Algorithms in Multidimensional Demographic Projection Models. *Population Studies*, 46(1):159–169.

Van Imhoff, E. and Post, W. (1998). Microsimulation methods for population projection. *Population: An English Selection*, 10(1):97–138.

van Noordwijk, A. J. and de Jong, G. (1986). Acquisition and Allocation of Resources: Their Influence on Variation in Life History Tactics.

van Raalte, A. A. and Caswell, H. (2013). Perturbation Analysis of Indices of Lifespan Variability. *Demography*, 50(5):1615–1640.

Van Tienderen, P. H. (1995). Life cycle trade-offs in matrix population models. *Ecology*, 76(8):2482–2489.

van Tienderen, P. H. (2000). Elasticities and the link between demographic and evolutionary dynamics. *Ecology*, 81(3):666–679.

Varga, R. S. (1962). *Matrix Iterative Analysis*. Springer.

Vaupel, J. W., Carey, J. R., Christensen, K., Johnson, T. E., Yashin, A. I., Holm, N. V., Iachine, I. A., Kannisto, V., Khazaeli, A. A., Liedo, P., Longo, V. D., Zeng, Y., Manton, K. G., and Curtsinger, J. W. (1998). Biodemographic Trajectories of Longevity. *Science*, 280(5365):855–860.

Vaupel, J. W., Manton, K. G., and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16(3):439–454.

Venner, S., Siberchicot, A., Pélisson, P.-F., Schermer, E., Bel-Venner, M.-C., Nicolas, M., Débias, F., Miele, V., Sauzet, S., Boulanger, V., and Delzon, S. (2016). Fruiting Strategies of Perennial Plants: A Resource Budget Model to Couple Mast Seeding to Pollination Efficiency and Resource Allocation Strategies. *The American Naturalist*, 188(1):000–000.

Verhulst, P. (1845). Recherches mathématiques sur la loi d'accroissement de la population. *Nouveaux mémoires de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles*, pages 14–54.

Vindenes, Y. (2010). *Stochastic modeling of finite populations with individual heterogeneity in vital parameters*. PhD thesis.

Vindenes, Y., Engen, S., and Sæther, B. (2008). Individual Heterogeneity in Vital Parameters and Demographic Stochasticity. *The American Naturalist*, 171(4):455–467.

Vindenes, Y., Lee, A. M., Engen, S., and Sæther, B. E. (2010). Fixation of slightly beneficial mutations: Effects of life history. *Evolution*, 64(4):1063–1075.

Voland, E. and Dunbar, R. I. M. (1995). Resource competition and reproduction. *Human Nature*, 6(1):33–49.

Wachter, K. W., Blackwell, D., and Hammel, E. A. (1997). Testing the validity of kinship microsimulation. *Mathematical and Computer Modelling*, 26(6):89–104.

Waddington, C. (1953). Genetic assimilation of an acquired character. *Evolution*, 7(2):118–126.

Walker, R. S., Gurven, M., Burger, O., and Hamilton, M. J. (2008). The trade-off between number and size of offspring in humans and other primates. *Proceedings. Biological sciences / The Royal Society*, 275(1636):827–33.

Watson, H. W. and Galton, F. (1875). On the Probability of the Extinction of Families. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 4(1875):138–144.

Waugh, W. A. O. (1955). An Age-Dependent Birth and Death Process. *Biometrika*, 42(3):291–306.

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

Werner, P. A. and Caswell, H. (1977). Population growth rates and age versus stage-distribution models for teasel (Dipsacus sylvestris Huds.). *Ecology*, 58:1103–1111.

Westendorp, R. G. and Kirkwood, T. B. (1998). Human longevity at the cost of reproductive success. *Nature*, 396(6713):743–746.

Willekens, F. J. (1977). Sensitivity analysis in multiregional demographic models. *Environment and Planning A*, 9(6):653–674.

Willführ, K. and Gagnon, A. (2013). Are step-parents always evil? Parental death, remarriage , and child survival in demographically saturated Krummhörn (1720-1859) and expanding Québec (1670-1750). *Biodemography and Social Biology*, 58(2):149–161.

Williams, G. C. (1957). Pleiotropy, natural selection and the evolution of senescence. *Evolution*, 11(4):398–411.

Williams, G. C. (1966). Natural Selection, the Costs of Reproduction, and a Refinement of Lack's Principle. *The American Naturalist*, 100(916):687–690.

Woltereck, R. (1909). Weitere experimenelle Untersuchungen uber Artveranderung, speziell uber des Wesen quantitativer Artunterschiede bei Daphniden. *Verhandlungen der deutschen zoologischen Gesellschaft*, 19:110–172.

Wright, S. (1931). Evolution in mendelian populations. *Bulletin of Mathematical Biology*, 52(1-2):241–295.

Yerushalmy, j., j.m. Bierman, d.h. Kemp, Connor, a., and f.e. French (1956). Longitudinal studies of pregnancy on the island-of-Kauai, territory-of-Hawaii. *Am J Obst Gynecol*, 71(1):80–96.

Yoccoz, N. G., Mysterud, A., Langvatn, R., and Stenseth, N. C. (2002). Age- and density-dependent reproductive effort in male red deer. *Proceedings of the Royal Society B: Biological Sciences*, 269(1500):1523–1528.

Yuster, R. and Zwick, U. (2005). Fast sparse matrix multiplication. *ACM Transactions on Algorithms*, 1(1):2–13.

Zaba, B., Whitworth, J., Marston, M., Nakiyingi, J., Ruberantwari, A., Urassa, M., Issingo, R., Mwaluko, G., Floyd, S., Nyondo, A., and Crampin, A. (2005). HIV and mortality of mothers and children: evidence from cohort studies in Uganda, Tanzania, and Malawi. *Epidemiology*, 16(3):275–280.