Evolution of environmentally mediated social interactions under isolation by distance

Charles Mullon[*1], Jorge Peña[†2,3], and Laurent Lehmann[‡4]

[1]Department of Ecology and Evolution, University of Lausanne, Switzerland
[2]Institute for Advanced Study in Toulouse, University of Toulouse Capitole, France
[3]Department of Human Behavior, Ecology and Culture, Max Planck Institute for
Evolutionary Anthropology

Working paper

[*]charles.mullon@unil.ch
[†]jorge.pena@iast.fr, jorge_pena@eva.mpg.de
[‡]laurent.lehmann@unil.ch

Abstract

We derive the selection gradient on a quantitative trait affecting the dynamics of biotic and/or abiotic state variables that feedback on individual survival and reproduction ("environmentally mediated social interactions") in a finite population of constant size subject to isolation by distance. Our results show that this selection gradient depends on how an individual expressing the evolving trait influences (i) its own fitness and that of all current relatives living at the different spatial positions of the population and (ii) the fitness of its downstream relatives through modifications of the state variables at the different spatial positions. These state-variable changes are captured by a formalization of the extended phenotypic effect of an individual taking the form of press perturbations from theoretical ecology. When the effect of the evolving trait on reproduction and survival is mediated by some material payoff function, the selection gradient can be expressed in terms of extended phenotypic effects weighted by scaled-relatedness coefficients. These weights quantify how local competition and genealogy interact to determine the genetic value measured at the payoff level of recipients of the extended phenotypic effects, possibly living far into the future. All quantities are expressed in terms of inverse Fourier transforms, which provide a computable measure of the effect of selection on environmentally mediated social interactions. As an illustration of our results, we derive the selection gradient on a trait affecting the dynamics of a spatially diffusive common-pool resource in lattice structured populations and show that posthumous spite can readily evolve by natural selection if there is a negative covariance across the habitat in the dispersal of the focal species and the resource it affects.

Keywords: adaptive dynamics, metacommunity, niche construction, spite, relatedness

1 Introduction

It is essentially inevitable that organisms interact with one another and with their environment, thereby impacting their own as well as the fitness of others. Such social interactions are incredibly diverse in nature but can usefully be classified as to whether they are direct, such as grooming, alloparental care, or combats to control resources or breeding territories; or as to whether they are indirectly mediated by the environment, such as through the depletion or enrichment of resources, the release or degradation of pollutants, or the construction of nests, burrows, or other shelters (Estrela et al., 2019). Direct social interactions thus typically occur among contemporaries that are physically close to one another. Environmentally mediated interactions, in contrast, can extend further in time and space. In fact, when external modifications have lasting and long-ranging effects on the environment, indirect social interactions may occur between individuals whose lifetimes show little or even no overlap. This can lead to forms of trans-generational harming (e.g., when overconsumption of a slowly renewable resource leads to stock collapse and poor harvest for future generations) or helping (e.g., when underconsumption ensures healthy stock maintenance).

The theory devoted to understanding the evolution of quantitative genetic traits that influence direct social interactions is well established (Frank, 1998; Rousset, 2004; Lehmann and Rousset, 2014; Van Cleve, 2015; McNamara and Leimar, 2020b). One of the main contributions of this theory has been to highlight the importance of limited dispersal in determining how Darwinian evolution shapes social traits (Frank, 1998; Rousset, 2004). This is because under limited dispersal in spatially or family structured populations, individuals expressing the same genetic traits may be more or less likely to directly interact with one another than with individuals expressing alternative traits. The importance of such genetic associations is enshrined in the gradient (or marginal) form of Hamilton's rule (Frank, 1998; Rousset, 2004), which captures the first order effects of selection on quantitative traits and is sufficient to characterise trait values towards which a population converges under mutation limited evolution (Van Cleve, 2015, i.e., to characterise convergence stability, Eshel, 1983). The gradient form of Hamilton's rule is computationally attractive because all the necessary information about direct interactions among carriers of the same genes is summarized in pairwise relatedness coefficients that are evaluated under neutrality (i.e., in the absence of selection). This remarkable characteristic is also true of the extensions made to consider isolation by distance (Rousset, 2004), where populations are structured according to the stepping-stone model of population genetics, or more generally according to lattice or network models with arbitrary dispersal distribution (e.g., Comins et al., 1980; Rogers, 1990; Taylor, 1992a, 1994; Irwin and Taylor, 2001; Gandon and Rousset, 1999; Rousset and Billiard, 2000; Rousset and Gandon, 2002; Taylor et al., 2007; Lehmann et al., 2007; Grafen and Archetti, 2008; Ohtsuki, 2010). Relatedness coefficients can be easily computed for a given biological scenario using standard coalescent arguments and readily estimated in natural populations. Because it combines computational accessi-

bility and empirical connections, the gradient form of Hamilton's rule or one of its extensions has been applied to understand an untold number of scenarios involving direct social interactions (Frank, 1998).

In contrast, the mathematical theory on the evolution of environmentally mediated social interactions in spatially structured populations has been much less widely applied, even though there exists a gold-standard expression for the selection gradient on quantitative traits that takes into account all relevant genetic and demographic stochastic effects owing to local finite patch size (Rousset, 2004; Rousset and Ronce, 2004). Why there are so few applications may be explained, at least partly, by the computational complexity of solving such problems. The difficulty arises because one needs to keep track of inter-dependent evolutionary and environmental stochastic processes that feedback on one another. More specifically, local sampling effects cause local genetic fluctuations that in turn influence how the environment varies in time and space (e.g., a greater frequency of polluters locally tends to lead to a greater concentration of pollutants in the vicinity). This variation then feeds back onto the survival and reproduction of individuals thus eventually influencing local allele frequency changes. As a result of this feedback, most problems of environmentally mediated social interactions in spatially structured populations involve a prohibitively large evolutionary and environmental state space on which allele frequency and environmental variables change stochastically (Rousset and Ronce, 2004). Characterising this state space and transitions between states is even more laborious when local environmental effects can be transferred in space and time (e.g., when pollutants diffuse to other locations or last in the environment beyond the lifetime of the individuals who release them).

The challenge is apparent from evolutionary models that allow for trait-driven changes in local demography, which can be thought of as an environmental variable that mediates social interactions. Even in the island model of dispersal where spatial structure is only implicit (Wright, 1931), the selection gradient depends on the entire probability distribution of local demographic states in the absence of genetic variation (e.g., Rousset, 2004; Rousset and Ronce, 2004; Ohtsuki et al., 2020). Because there is often no analytical solutions to such distributions, analyses of selection and convergence stability typically rely heavily on numerical methods (e.g., Cadet et al., 2003; Parvinen et al., 2003; Rousset and Ronce, 2004; Lehmann et al., 2006; Wild et al., 2009). This "curse of dimensionality" becomes even more acute under isolation by distance, where the distribution of whole meta-population states in the absence of genetic variation need to be considered (Rousset and Ronce, 2004, eq. 23). To circumvent this challenge, two approximations have been suggested. One is the pair approximation for lattice-structured populations (Nakamaru et al., 1997; van Baalen and Rand, 1998; Le Galliard et al., 2003, 2005; Nakamaru and Iwasa, 2005; Ohtsuki et al., 2006; Lion and van Baalen, 2007; Peña et al., 2016). This approximation is based on moment equations of the state distribution and consists in ignoring third and higher order moments. Another approximation is to consider that the dynamics of environmental state variables are locally deterministic, so that there is no environmental fluctuation in the absence of genetic variation (Mullon and Lehmann, 2018). In this case, the selection gradient can be expressed as an extension to Hamilton's

rule, involving extended phenotypic effects mediated by the environment. So far, this approach has been applied to the island model, so in the absence of isolation by distance. General formulas for the selection gradient due to inter-temporal fitness effects under isolation by distance exist (Lehmann, 2008; Lehmann and Rousset, 2012), but those leave the effects of the environment implicit so that how environmentally mediated social interactions evolve remains to be specified.

Here, we fill this gap by characterising the selection gradient on a trait that impacts the deterministic dynamics of an environmental state variable which feeds back on survival and reproduction under isolation by distance. Using Fourier analysis, we express the selection gradient in terms of extended phenotypic effects and relatedness coefficients scaled to local competition, both of which provide biological insights about the nature of selection and are straightforward to compute for a wide range of classical models (e.g., Wright-Fisher and Cannings models). We use our results to investigate the evolution of environmentally-mediated helping and harming through time and space. In contrast to previous suggestions, our analyses indicate that indiscriminate spite where individuals suffer a cost to harm others living in the future can readily evolve.

2 Model

2.1 Spatial structure, life cycle, traits and environmental variables

We consider a population of homogeneous individuals (no class structure) that is distributed among D patches (or demes) connected by dispersal where each patch is occupied by N adult individuals. The population is censused at discrete demographic time steps during which the following events occur in cyclic order: (a) reproduction and adult survival; (b) dispersal; and (c) density-dependent regulation such that each patch contains exactly N adult individuals at the beginning of the next demographic time step. Through adult survival, individuals can go through multiple rounds of these life-cycle events so that our models allows for overlapping generations (but no age structure).

Patches are arranged homogeneously in d dimensions, with D_j patches in dimension $j \in 1, \ldots, d$. For example, under a lattice structure in a one dimensional habitat, $D = D_1$ patches are arranged on a circle, while in a two dimensional habitat, $D = D_1 \times D_2$ patches are arranged on a torus. More generally we denote by $\mathcal{G} = \{(i_1, i_2, ..., i_d) : 0 \leq i_j < D_j\}$ the set of all patches, which we endow with an abelian group structure (see Box 1).

Each patch is characterized by a quantitative state variable representing a biotic or abiotic environmental factor, which we refer to as an environmental state variable (e.g., density of a common-pool resource, of a pollutant, or quality of the habitat). Meanwhile, each individual in the population is characterised by a genetically determined quantitative trait that influences the environment and fitness

5

(e.g., consumption of a resource, release of a pollutant). We are interested in the evolution of this trait under biological scenarios that display the following three features.

(i) *Trait and environmentally mediated reproduction and survival*. By expressing the evolving trait, individuals can affect the survival and reproduction of any other individual in the population. For example, individuals may engage in costly combats for resources in other patches and return to their own to share these resources with patch neighbours. The effects of trait expression on others are assumed to be spatially homogeneous, i.e., the effect of an individual from patch $\mathbf{i} = (i_1, i_2, ..., i_d) \in \mathcal{G}$ on the survival and/or reproduction of an individual in patch $\mathbf{j} = (j_1, j_2, ..., j_d) \in \mathcal{G}$ only depends on the "distance" $\mathbf{j} - \mathbf{i}$ among the two patches (unless specified otherwise, distance throughout means relative position and thus can be negative, e.g., Malécot, 1975). The survival and reproduction of an individual may also depend on the environmental state variable of each patch in a spatially homogeneous way (i.e., the effect of the environmental state variable of a patch $\mathbf{i}$ on the survival and reproduction of an individual residing in patch $\mathbf{j}$ only depends on the distance $\mathbf{j} - \mathbf{i}$).

(ii) *Localized dispersal*. Each individual either stays in its natal patch or disperses to another patch. Dispersal occurs with non-zero probability so that patches are not completely isolated from one another. We assume that dispersal from one patch $\mathbf{i}$ to another $\mathbf{j}$ depends only on the distance $\mathbf{k} = \mathbf{j} - \mathbf{i}$ between patches so that dispersal is also spatially homogeneous. We can thus write $m_{\mathbf{k}}$ for the probability that an individual disperses to a patch at distance $\mathbf{k}$ from it natal patch (with $\sum_{\mathbf{k} \in \mathcal{G}} m_{\mathbf{k}} = 1$). This allows us to consider short and long range dispersal by assuming e.g., that $m_{\mathbf{k}}$ is based on a binomial distribution (Fig. 1AB and 2A).

(iii) *Trait and environmentally mediated environmental dynamics*. Through trait expression, individuals can also affect the dynamics of the environmental state variables over one demographic time step. For example, the environmental variable may be a common-pool resource that individuals absorb locally, or a pollutant produced by individuals which then diffuses in the environment. Such trait effects on the environment are spatially homogeneous (so that the effect of an individual from patch $\mathbf{i}$ on the environmental state variable of patch $\mathbf{j}$ only depends on $\mathbf{j} - \mathbf{i}$).

2.2 The focal individual, its fitness, and environmental dynamics

The assumption of spatial homogeneity means that patch $\mathbf{0} \in \mathcal{G}$ can be taken as a representative patch of the population and that any individual in this focal patch can be taken as a representative individual from the population, henceforth called the focal individual. Let us then denote by $z_{\bullet}$ the realized value of the trait of a focal individual and by $z_{\mathbf{k},t}$ the average phenotype of individuals living in a patch "$\mathbf{k}, t$", i.e., in a patch $\mathbf{k} \in \mathcal{G}$ at t demographic time points (or "generations") prior to the focal generation, with the important caveat that $z_{\mathbf{0},0}$ denotes the average phenotype among the neighbors of the focal

individual in the focal generation (thus excluding the focal from the average). Let us further denote by $\mathbf{z}_{\mathbf{0},t} = (z_{\mathbf{0},t}, ..., z_{\mathbf{k},t}, ..., z_{\mathbf{D}_{-1},t})$ the vector collecting all such realized phenotypes in $\mathcal{G}$ in lexicographic order and terminating with the one at position $\mathbf{D}_{-1} = (D_1 - 1, D_2 - 1, ..., D_d - 1)$. Let us also denote by $n_{\mathbf{k},t}$ the environmental variable in patch $\mathbf{k}$ at $t \geq 0$ generations prior to the focal generation. Finally, denote by $\mathbf{n}_{\mathbf{0},t} = (n_{\mathbf{0},t}, ..., n_{\mathbf{k},t}, ..., n_{\mathbf{D}_{-1},t})$ the collection of all environmental variables across all patches.

In terms of realized trait values and environmental state variables, we consider the class of models where the fitness of the focal individual is determined by the function $w : \mathbb{R} \times \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}_+$ defined such that

$$w(z_\bullet, \mathbf{z}_{\mathbf{0},0}, \mathbf{n}_{\mathbf{0},0}) \tag{1}$$

is the expected number of successful offspring (including the surviving self) produced over one demographic time by a focal individual with trait $z_\bullet$ when the collection of average traits among other individuals at the different spatial positions is $\mathbf{z}_{\mathbf{0},0}$ and the collection of environmental variables is $\mathbf{n}_{\mathbf{0},0}$. These environmental variables are obtained as solutions to the system of equations

$$n_{\mathbf{k},t} = F(\mathbf{z}_{\mathbf{k},t+1}, \mathbf{n}_{\mathbf{k},t+1}) \qquad \text{for } \mathbf{k} \in \mathcal{G}, \tag{2}$$

where $F : \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}$ is a transition map for the dynamics of the environmental variable, and where $\mathbf{z}_{\mathbf{k},t}$ is obtained, for all t, by a circular permutation of the elements of $\mathbf{z}_{\mathbf{0},t}$ having $z_{\mathbf{k},t}$ as first element [e.g., for a one dimensional lattice where $d = 1$, we have $\mathbf{z}_{\mathbf{0},t} = (z_{0,t}, z_{1,t}, ..., z_{D-1,t})$, $\mathbf{z}_{\mathbf{1},t} = (z_{1,t}, z_{2,t}, ..., z_{0,t})$, $\mathbf{z}_{\mathbf{2},t} = (z_{2,t}, z_{3,t}, ..., z_{0,t}, z_{1,t})$, etc.]. Likewise, the vector $\mathbf{n}_{\mathbf{k},0}$ is the value at $t = 0$ of the vector $\mathbf{n}_{\mathbf{k},t}$ obtained by a circular permutation of the elements of $\mathbf{n}_{\mathbf{0},t}$ having $n_{\mathbf{k},t}$ as first element [e.g., for a one dimensional lattice, we have $\mathbf{n}_{\mathbf{1},t} = (n_{1,t}, n_{2,t}, ..., n_{0,t})$, $\mathbf{n}_{\mathbf{2},t} = (n_{2,t}, n_{3,t}, ..., n_{0,t}, n_{1,t})$, etc.]. Hence, for each patch $\mathbf{k}$, the function F (which is the same in all patches and all times) maps the current average phenotypes and environmental variables in the whole population to the new value of the environment in that patch. Note that the time indices in eq. (2) run backwards in time owing to the fact that individual fitness depends necessarily on the realized past states. It follows from the recursive structure of eq. (2) that the vector of environmental variables $\mathbf{n}_{\mathbf{0},0}$ depends potentially on the whole history of traits $\mathbf{z}_{\mathrm{H}} = (\mathbf{z}_{\mathbf{0},1}, \mathbf{z}_{\mathbf{0},2}, ...)$ in the population prior to the focal generation; namely on the vector collecting all $z_{\mathbf{k},t}$ for all $\mathbf{k} \in \mathcal{G}$ and all t prior to the focal generation. Thus the fitness of a focal individual (1) depends potentially on all other previous living individuals across space and time. To bring this dependence upfront, we write the fitness of the focal individual as $w(z_\bullet, \mathbf{z}_{\mathbf{0},0}, \mathbf{n}_{\mathbf{0},0}(\mathbf{z}_{\mathrm{H}}))$.

We make the additional assumption that in a monomorphic population where all individuals have one and the same (resident) phenotype z, the deterministic ecological dynamics described by the map F has a unique hyperbolically stable equilibrium point, identical in each patch, and satisfying

$$\hat{n} = F(\mathbf{z}, \hat{n}), \tag{3}$$

where $\mathbf{z} = (z, ..., z)$ and $\hat{\boldsymbol{n}} = (\hat{n}, ..., \hat{n})$ are vectors of dimension D whose entries are all equal to trait value z and environmental state variable $\hat{n}$ respectively. In the study of multi-patch ecological systems, this is sometimes called the spatially homogeneous or flat solution (Jansen and Lloyd, 2000, p. 235). Note that demographic consistency implies that in a resident population at the environmental equilibrium fitness is equal to one, i.e., $w(z, \mathbf{z}, \hat{\boldsymbol{n}}(z)) = 1$ holds for all feasible z.

2.3 Evolutionary dynamics

We assume that the quantitative trait evolves through rare mutations of weak phenotypic effects such that evolutionary dynamics are mutation limited and proceed as a trait substitution sequence on the state space $\mathcal{Z} \subseteq \mathbb{R}$ in a finite population without interactions between mutation and selection (i.e., the process of "long term evolution" for finite populations described in Van Cleve, 2015). The direction of evolution under such a trait substitution sequence can be ascertained by computing a selection gradient, which consists in the first-order effects of selection on the fixation probability of a single mutant copy. The local attractors of the evolutionary dynamics (i.e., convergent stable trait values) are points where the selection gradient vanishes (Rousset and Billiard, 2000; Rousset, 2003; Van Cleve, 2015). Our main objective in this paper is to derive the selection gradient for this process under the scenario described above in a way that is both computationally accessible and easy to interpret biologically. With these aims in sight, we arrive at different equivalent expressions to characterize selection. Most technical details about this derivation can be found in the appendices and in the accompanying boxes. Our main findings are summarized below.

3 Results

3.1 Recipient-centered perspective: direct and indirect fitness effects

Let $\mathcal{Z} \subset \mathbb{R}$ denote the state space of the evolutionary dynamics. Then, we show in Appendix A that a resident trait value z^* in the interior of $\mathcal{Z}$ is a local evolutionary trap, i.e., z^* is locally convergence stable, if and only if both

$$s(z^*) = 0 \quad \text{and} \quad \left.\frac{\mathrm{d}s(z)}{\mathrm{d}z}\right|_{z=z^*} < 0 \tag{4}$$

hold, where

$$s(z) = s_{\mathrm{w}}(z) + s_{\mathrm{e}}(z), \tag{5}$$

with

$$s_{\mathrm{w}}(z) = \frac{\partial w(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0}(\mathbf{z}_{\mathrm{H}}))}{\partial z_\bullet} + \sum_{\mathbf{k} \in \mathcal{G}} \frac{\partial w(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0}(\mathbf{z}_{\mathrm{H}}))}{\partial z_{\mathbf{k},0}} R_{\mathbf{k},0} \tag{6a}$$

8

$$s_{\text{e}}(z) = \sum_{t=1}^{\infty} \sum_{\mathbf{k} \in \mathcal{G}} \frac{\partial w(z_{\bullet}, \mathbf{z_{0,0}}, \mathbf{n_{0,0}}(\mathbf{z_{\text{H}}}))}{\partial z_{\mathbf{k},t}} R_{\mathbf{k},t}. \tag{6b}$$

Both $s_{\text{w}}(z)$ and $s_{\text{e}}(z)$ depend on marginal fitness effects, i.e., on derivatives of focal fitness, which here and hereafter are evaluated in a monomorphic population where all individuals express the resident trait value z, and where the environmental state variable in all patches is at equilibrium $\hat{n}$ (eq. 3). The quantity $R_{\mathbf{k},t}$ weighing fitness effects in eq. (6) is the relatedness between the focal individual and another randomly sampled individual from patch $\mathbf{k}, t$. It is defined as,

$$R_{\mathbf{k},t} = \lim_{\mu \to 0} \frac{Q_{\mathbf{k},t} - \overline{Q}_t}{1 - \overline{Q}_0}, \tag{7}$$

where μ is the mutation rate at the evolving locus; $Q_{\mathbf{k},t}$ is the stationary probability that an allele sampled in the focal individual is identical by descent with a homologous allele sampled in another individual chosen at random from patch $\mathbf{k}, t$ under neutrality (i.e., in a population monomorphic for z); and $\overline{Q}_t = \sum_{\mathbf{k} \in \mathcal{G}} Q_{\mathbf{k},t}/D$ is the average probability of identity between two homologous alleles sampled in two individuals living t generations apart. The probability of identity by descent $Q_{\mathbf{k},t}$, and thus $R_{\mathbf{k},t}$, may depend on the resident phenotype z but we leave this dependence implicit for readability.

Relatedness $R_{\mathbf{k},t}$ can be interpreted as the likelihood that an individual sampled in patch $\mathbf{k}, t$ carries an allele identical by descent to the focal individual, compared to a randomly sampled individual at generation t (prior to the focal generation). In a panmictic or randomly mixing population (where $m_{\mathbf{k}} = 1/D$ for all $\mathbf{k} \in \mathcal{G}$), relatedness between any two individuals is zero. However, as soon as soon as dispersal is limited (where $m_0 > 1/D$), relatedness between individuals of the same patch becomes positive, i.e., $R_{\mathbf{0},t} > 0$. Since the average relatedness is zero (i.e., $\sum_{\mathbf{k} \in \mathcal{G}} R_{\mathbf{k},t}/D = 0$ holds) this then entails that some individuals in the population are negatively related to the focal. To illustrate the concept of relatedness, consider a Wright-Fisher process (i.e., there is no adult survival and individuals are semelparous), which is the reference model under which probabilities of identity by descent are evaluated under isolation by distance (e.g., Malécot, 1975; Sawyer, 1976; Rousset, 2004). The relatedness coefficients can then be obtained for $t = 1, 2, 3, \dots$ as

$$R_{\mathbf{k},t} = \frac{1}{ND + G} \sum_{\mathbf{h} \in \mathcal{G} \setminus \mathbf{0}} \frac{\mathcal{M}(\mathbf{h})^t}{1 - \mathcal{M}(\mathbf{h})^2} \overline{\chi}_{\mathbf{k}}(\mathbf{h}), \tag{8}$$

where $\mathcal{G} \setminus \mathbf{0}$ means the set $\mathcal{G}$ with element $\mathbf{0}$ removed, $G = \sum_{\mathbf{h} \in \mathcal{G} \setminus \mathbf{0}} \mathcal{M}(\mathbf{h})^2 / \left(1 - \mathcal{M}(\mathbf{h})^2\right)$, and

$$\mathcal{M}(\mathbf{h}) = \sum_{\mathbf{k} \in \mathcal{G}} m_{\mathbf{k}} \chi_{\mathbf{k}}(\mathbf{h}) \tag{9}$$

is the Fourier transform of the dispersal distribution (see Box 1 for the definition of the character functions $\chi_{\mathbf{k}}(\mathbf{h})$ and more details). The relatedness coefficient between two individuals sampled at $\mathbf{k}$ steps apart in the same generation ($t = 0$) is given by substituting $t = 2$ into eq. 8; that is, $R_{\mathbf{k},0} = R_{\mathbf{k},2}$ (for

a derivation of eq. (8) see Lehmann, 2010). Fig. 1 shows the typical patterns of relatedness for populations structured as one-dimensional and two-dimensional lattices, and illustrates that relatedness can readily take negative values.

In eq. (6a), the partial derivative $\partial w / \partial z_{\bullet}$ represents the effect of a change in trait expression by the focal individual on its own fitness. By contrast, $\partial w / \partial z_{\mathbf{k},t}$ (with $t = 0$ in eq. 6a and $t \geq 1$ in eq. 6b) is the effect of the whole set of individuals living in patch $\mathbf{k}, t$ on the fitness of the focal individual. The first term $s_{\mathrm{w}}(z)$ in eq. (5) thus captures the net effect of selection owing to all intra-temporal effects on fitness (i.e., effects within a demographic period), while the second term $s_{\mathrm{e}}(z)$ captures selection owing to all inter-temporal effects (i.e., effects between demographic periods). As such, eq. (5) gives the sum of the effects of all individuals that are or were alive on the fitness of a focal individual alive in a focal generation, which is thus regarded as the recipient of all such phenotypic effects.

3.2 Actor-centered perspective: inclusive fitness and extended phenotypic effects

Owing to the symmetry between individuals in a resident population, relatedness eq. (8) can also be interpreted as a measure of the extent to which an individual sampled in patch $\mathbf{k}$ at time t in the future is more (or less) likely to carry an allele identical by descent to that in the focal individual than is a randomly sampled individual at time t. Owing to space-time homogeneity, all components of the selection gradient can in fact be interpreted in an actor-centered way where fitness effects are grouped by an actor changing trait expression (i.e., the inclusive fitness way of grouping fitness effects, e.g., Hamilton, 1970; Frank, 1998; Rousset, 2004; Lehmann, 2008). For instance, $\partial w / \partial z_{\mathbf{k},t}$ can be read as the effect of a single individual in the focal patch on the fitness of the whole set of individuals in a patch at spatial "distance" $\mathbf{k}$ at t steps in the future, which readily leads to an actor-centered perspective of the fitness component $s_{\mathrm{w}}(z)$. Yet, because fitness w is mediated by the environmental variable and its value depends on past phenotypic expression, the actor-centered perspective is not explicit in the fitness component $s_{\mathrm{e}}(z)$. To make this explicit and further make $s_{\mathrm{e}}(z)$ more easily computable, we next derive an actor centered representation of $s_{\mathrm{e}}(z)$ by formalizing the extended phenotypic effect of an individual on its environment.

From now on, we let t index time forward so that $\mathbf{n}_{\mathbf{k},t}$ is the value of the environmental state variable in patch $\mathbf{k}$ at t time steps in the future of the focal generation, and likewise let $\mathbf{z}_{\mathbf{k},t}$ denote the collection of population phenotypes at time t in the future, as viewed from the perspective of individuals in patch $\mathbf{k}$. This allows us to rewrite eq. (2) by looking forward in time as

$$
n_{\mathbf{k},t+1} = \begin{cases} F(\mathbf{z}_{\mathbf{k},0}^{\mathrm{R}}, \mathbf{n}_{\mathbf{k},0}) & \text{for } t = 0 \\ F(\mathbf{z}_{\mathbf{k},t}, \mathbf{n}_{\mathbf{k},t}) & \text{for } t \geq 1 \end{cases} , \tag{10}
$$

where $\mathbf{z}_{\mathbf{k},0}^R$ is equal to $\mathbf{z}_{\mathbf{k},0}$ except that the component $z_{\mathbf{0},0}$ in this vector is replaced with $z_{\mathbf{0},0}^R = \frac{1}{N} z_\bullet + \frac{N-1}{N} z_{\mathbf{0},0}$, i.e., the average phenotype in the focal patch including the focal individual. The map (10) brings upfront that the expression of the phenotype of the focal in the focal patch has cascading downstream effects on the value the environmental variable can take in any patch of the population, possibly at any time in the future. To quantify such effect, we define by

$$e_{\mathbf{k},t} = \frac{\partial n_{\mathbf{k},t}}{\partial z_\bullet} \tag{11}$$

the extended phenotypic effect of the focal individual on the environmental variable in patch $\mathbf{k}$ at t generations in the future (where, as usual, all derivatives are evaluated at the resident values z and $\hat{n}$). The extended phenotypic effect (11) can be computed from knowledge of only $F(\mathbf{z}_{\mathbf{0},0}^R, \mathbf{n}_{\mathbf{0},0})$, which is usually intuitive to write down given specific modeling assumptions (see example in Section (3.4)). To that end, let

$$\psi_{\mathbf{k}} = \frac{\partial F(\mathbf{z}_{\mathbf{k},0}^R, \mathbf{n}_{\mathbf{k},0})}{\partial z_\bullet} = \begin{cases} \dfrac{1}{N} \dfrac{\partial F(\mathbf{z}_{\mathbf{0},0}^R, \mathbf{n}_{\mathbf{0},0})}{\partial z_{\mathbf{0},0}^R} & \text{for } \mathbf{k} = \mathbf{0} \\[4mm] \dfrac{1}{N} \dfrac{\partial F(\mathbf{z}_{\mathbf{0},0}^R, \mathbf{n}_{\mathbf{0},0})}{\partial z_{\mathbf{k},0}} & \text{otherwise} \end{cases} \tag{12}$$

be the change in the abundance of the ecological variable in the patch at $\mathbf{k}$ steps apart from the focal patch and resulting from a change in the phenotype of the focal individual, where the second equality follows from spatial homogeneity and is useful to analyze concrete models (see exampl in Section 3.4). Further, let

$$c_{\mathbf{k}} = \frac{\partial F(\mathbf{z}_{\mathbf{k},0}^R, \mathbf{n}_{\mathbf{k},0})}{\partial n_{\mathbf{0},0}} = \frac{\partial F(\mathbf{z}_{\mathbf{0},0}^R, \mathbf{n}_{\mathbf{0},0})}{\partial n_{\mathbf{k},0}} \tag{13}$$

be the effect of the variation of the abundance of the ecological variable in the focal patch on the change of the dynamics of the variable in patch $\mathbf{k}$, where the second equality also follows from spatial homogeneity and is useful to analyze concrete models. Now let

$$\Psi(\mathbf{h}) = \sum_{\mathbf{k} \in \mathcal{G}} \psi_{\mathbf{k}} \chi_{\mathbf{k}}(\mathbf{h}) \tag{14}$$

and

$$\mathcal{C}(\mathbf{h}) = \sum_{\mathbf{k} \in \mathcal{G}} c_{\mathbf{k}} \chi_{\mathbf{k}}(\mathbf{h}) \tag{15}$$

be, respectively, the Fourier transform of the coefficients defined in eq. (12) and eq. (13). Then, we show in Appendix B.1 that the extended phenotypic effect can be be conveniently computed as the inverse Fourier transform:

$$e_{\mathbf{k},t} = \frac{1}{D} \sum_{\mathbf{h} \in \mathcal{G}} \mathcal{E}_t(\mathbf{h}) \overline{\chi}_{\mathbf{k}}(\mathbf{h}), \tag{16}$$

where

$$\mathcal{E}_t(\mathbf{h}) = \mathcal{C}(\mathbf{h})^{t-1} \Psi(\mathbf{h}). \tag{17}$$

This brings upfront that $e_{\mathbf{k},t}$ is conceptually equivalent to a press perturbation, i.e., a unit perturbation in the dynamics of some environmental state variable that propagates to the future (e.g., Yodzis, 1989; Novak et al., 2016). Here, this press perturbation has its origin in a single individual affecting the dynamics of the state variable in possibly different locations in the habitat (captured by the coefficients $\psi(\mathbf{h})$); this change brings forth cascading effects over $t-1$ demographic time steps to finally affect the environment t generations downstream. A concrete example of such press perturbations is given in Section 3.4.

In terms of the extended phenotypic effects just defined, we show in Appendix B.2 that the selection component $s_{\mathrm{e}}(z)$ can be written as

$$s_{\mathrm{e}}(z) = \sum_{t=1}^{\infty} \sum_{\mathbf{k} \in \mathcal{G}} \sum_{\mathbf{j} \in \mathcal{G}} N R_{\mathbf{k},t} \frac{\partial w(z_{\bullet}, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial n_{\mathbf{j},0}} e_{\mathbf{j}-\mathbf{k},t}, \tag{18}$$

which can be read as an inclusive fitness effect of how the focal individual changes the fitness of all downstream relatives by changing the value of the environmental variable at each location of the habitat to which change the relatives are exposed. Beyond providing a biologically relevant interpretation of the nature of natural selection on environmentally mediated social interactions, eqs. (16)–(18) also make the selection component computable.

Eqs. (18)-(17) generalizes to a multi-dimensional environmental dynamics, i.e., when there are n_{e} biotic or abiotic state variables that can be affected by the evolving trait (see Box 2). Importantly, the focal species abundance N could itself be taken as one of these state variables. In this case its dynamic map (one of the components of eq. 48) needs to be written consistently with the life-cycle assumptions and its value in all expressions of the selection gradient needs to be evaluated at the resident environmental equilibrium (eq. 49), so the focal species abundance will depend endogenously on the evolving trait, i.e., $N(z)$ (see Mullon et al., 2018 for an application to the island model of dispersal, an approach that carries over *mutatis mutandis* to isolation by distance by applying the results of Box 2). Before turning to an application of result (18), we further simplify the selection gradient under some common assumptions about social interactions commonly endorsed in the literature.

3.3 Payoff-mediated fitness: scaled relatedness or the payoff value of others

We now assume that individual fitness is determined by some payoff function, which summarizes the social interactions between individuals and is a common assumption in evolutionary game theory (e.g., McNamara and Leimar, 2020a). To that end, we introduce the payoff function $\pi : \mathbb{R} \times \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}_+$ defined such that $\pi(z_{\bullet}, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})$ is the payoff to an individual with phenotype $z_{\bullet}$ when the collection of average phenotypes among all other actors on its payoff is $\mathbf{z}_{0,0}$ and the collection of environmental state variable is $\mathbf{n}_{0,0}$. We then assume that fitness of a focal individual can be written as a function of the

payoff to each actor on its fitness (namely, itself, an average patch neighbour, and average individual from each other patch) as

$$w(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0}) = \tilde{w}(\boldsymbol{\pi}(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})), \tag{19}$$

where

$$\boldsymbol{\pi}(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0}) = \{\underbrace{\pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}_{\pi_\bullet}, \underbrace{\pi(\mathbf{z}_{0,0}, \mathbf{z}_{0,0}^{\mathrm{n}}, \mathbf{n}_{0,0})}_{\pi_0}, \dots, \underbrace{\pi(z_{\mathbf{j},0}, \mathbf{z}_{\mathbf{j},0}, \mathbf{n}_{\mathbf{j},0})}_{\pi_{\mathbf{j}}}\} \tag{20}$$

is the vector collecting the payoffs to each type of actor. In the payoff π_0 to an average patch neighbour to the focal, $\mathbf{z}_{0,0}^{\mathrm{n}}$ is equivalent to $\mathbf{z}_{0,0}$, except that the first entry is given by $z_{0,0}^{\mathrm{n}} = \frac{1}{N-1}z_\bullet + \left(\frac{N-2}{N-1}\right)z_{0,0}$, i.e., the average trait among the neighbors of a neighbor of the focal individual. This arises because fitness depends on the payoff to patch neighbours of a focal individual, and these neighbors have their payoff possibly affected by the focal individual.

To illustrate the construction above, consider a life cycle where individuals are iteroparous (i.e., individuals can survive over demographic time steps) and where the fitness of the focal individual expressed in payoff terms is given by

$$\tilde{w}(\boldsymbol{\pi}) = s(\pi_\bullet) + \sum_{\mathbf{i} \in \mathcal{G}} m_{\mathbf{i}} \left[1 - s^{\mathrm{R}}(\pi_{\mathbf{i}})\right] \frac{f(\pi_\bullet)}{\sum_{\mathbf{j} \in \mathcal{G}} m_{\mathbf{i}-\mathbf{j}} f^{\mathrm{R}}(\pi_{\mathbf{j}})}. \tag{21}$$

Here, the payoff to an individual can affect both its fecundity and survival, $f : \mathbb{R}_+ \to \mathbb{R}_+$ and $s : \mathbb{R}_+ \to \mathbb{R}_+$ are fecundity and survival functions, respectively, and quantities with R as superscript are defined such that $s^{\mathrm{R}}(\pi_0) = \frac{1}{N}s(\pi_\bullet) + \frac{N-1}{N}s(\pi_0)$ is the average survival in the focal patch, otherwise $s^{\mathrm{R}}(\pi_{\mathbf{i}}) = s(\pi_{\mathbf{i}})$ for $\mathbf{i} \neq \mathbf{0}$; and $f^{\mathrm{R}}(\pi_{\mathbf{i}}) = \frac{1}{N}f(\pi_\bullet) + \frac{N-1}{N}f(\pi_0)$ is the average fecundity in the focal patch, otherwise $f^{\mathrm{R}}(\pi_{\mathbf{i}}) = f(\pi_{\mathbf{i}})$ for $\mathbf{i} \neq \mathbf{0}$. When survival is zero, eq. 21 is the fitness for a Wright-Fisher process; with no dependence on environmental state, it boils down to the fitness model considered in Rousset (2000, eq. 3) except that traits do not affect fecundity here. This captures standard "fecundity effects" which implements a form of 'death–birth' updating protocol in the language of evolutionary graph theory (e.g., Ohtsuki et al., 2006; Ohtsuki and Nowak, 2006). By contrast, in the absence of traits effect on fecundity, eq. 21 provides a model with only "survival effects" which implements a form of 'birth–death' updating protocol.

In Appendix C, we prove that with fitness written in the form of eq. (19), the selection gradient (5) can be expressed as

$$s(z) \propto \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial z_\bullet} + \sum_{\mathbf{k} \in \mathcal{G}} \kappa_{\mathbf{k},0} \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial z_{\mathbf{k},0}} + \sum_{t=1}^{\infty} \sum_{\mathbf{k} \in \mathcal{G}} N\kappa_{\mathbf{k},t} \sum_{\mathbf{j} \in \mathcal{G}} \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial n_{\mathbf{j},0}} e_{\mathbf{k}-\mathbf{j},t}. \tag{22}$$

13

Here,

$$
\kappa_{\mathbf{k},t} = \begin{cases} \dfrac{R_{\mathbf{0},0} - \frac{1}{N-1}\lambda_{\mathbf{0}}(1 - R_{\mathbf{0},0}) - \sum_{\mathbf{j}\in\mathcal{G}} \lambda_{\mathbf{j}} R_{\mathbf{j},0}}{1 - \sum_{\mathbf{j}\in\mathcal{G}} \lambda_{\mathbf{j}} R_{\mathbf{j},0}} & \text{for } (\mathbf{k}, t) = (\mathbf{0}, 0) \\[4mm] \dfrac{R_{\mathbf{k},t} - \sum_{\mathbf{j}\in\mathcal{G}} \lambda_{\mathbf{j}} R_{\mathbf{j}-\mathbf{k},t}}{1 - \sum_{\mathbf{j}\in\mathcal{G}} \lambda_{\mathbf{j}} R_{\mathbf{j},0}} & \text{otherwise,} \end{cases}
\tag{23}
$$

where

$$
\lambda_{\mathbf{j}} = -\frac{\partial \tilde{w}}{\partial \pi_{\mathbf{j}}} \bigg/ \frac{\partial \tilde{w}}{\partial \pi_{\bullet}}.
\tag{24}
$$

is a coefficient of fitness interdependence measuring the marginal effect of the material payoff of a neighbor at distance $\mathbf{j}$ on own fitness, relative to the marginal effect of own material payoff on own fitness. A positive $\lambda_{\mathbf{j}}$ can be then be interpreted as competition for local resources: an increase in the material payoffs to neighbors reduces an individual's fitness. The coefficient $\kappa_{\mathbf{k},t}$ (23) is a measure of scaled relatedness (Lehmann and Rousset, 2010; Van Cleve, 2015) and can be interpreted as a marginal substitution rate; namely, as the number of units of own payoff the focal individual is willing to exchange for one unit of payoff accruing to the whole set of individuals living in group $\mathbf{k}, t$, the value of which depends on the nature of competition between individuals. The scaled relatedness coefficient $\kappa_{\mathbf{k},t}$ can thus be seen as the "genetic value" of others *sensu* evolutionary biology in units of payoff. Eq. (22) can be read as an inclusive fitness effect at the payoff level, where the actor values individuals in each patch according to their scaled relatedness to the actor (sum over $\mathbf{k}$ in eq. (22)) and how the focal changes the payoff to each such type of recipient, which for downstream relatives is mediated by how the focal changes the value of the environmental variable at each location of the habitat to which change the relatives are exposed (sum over $\mathbf{j}$ in the third summand in eq. (22)).

While eq. (23) may at first hand appear complicated to evaluate, manageable expressions of $\kappa_{\mathbf{k},t}$ can be obtained under the fitness model of eq. (21), which covers both the Wright-Fisher process and overlapping generation processes with arbitrary dispersal distributions. These expressions for $\kappa_{\mathbf{k},t}$ are detailed in Box 3. For instance, for a Wright-Fisher process ($s = 0$) we have

$$
\kappa_{\mathbf{k},t} = \begin{cases} -\dfrac{1}{ND - 1} & \text{if } t = 0 \\[4mm] \dfrac{D p_{\mathbf{k},t} - 1}{DN - 1} & \text{otherwise,} \end{cases}
\tag{25}
$$

where

$$
p_{\mathbf{k},t} = \frac{1}{D} \sum_{\mathbf{h}\in\mathcal{G}} \mathcal{M}(\mathbf{h})^t \overline{\chi}_{\mathbf{k}}(\mathbf{h})
\tag{26}
$$

is the probability that a random line of genes descending from an individual residing in the focal patch will be in patch $\mathbf{k}$ at $t > 0$ steps in the future (recall Box 1 and note that $\sum_{\mathbf{k}\in\mathcal{G}} p_{\mathbf{k},t} = 1$ and $\sum_{\mathbf{k}\in\mathcal{G}} \kappa_{\mathbf{k},t} = 0$). The collection of these probabilities thus yield the distribution of a standard random walk with step distribution given by $m_{\mathbf{i}}$. As was true for relatedness, eq. (25) entails that for some

$t > 0$ some individuals in the population are positive and other negatively (scaled) related to the focal individual. The exact nature depends on the dispersal distribution (see Fig 1 for an illustration of the scaled relatedness).

This example shows that our formalization allows for recovering a number of previous results from the social evolution literature. For instance, consider that there are no ecologically meditated interactions, i.e., $\partial \pi(z_\bullet, z_{0,0}, n_{0,0})/\partial n_{0,0} = 0$, that $-C(z) = \partial \pi(z_\bullet, z_{0,0}, n_{0,0})/\partial z_\bullet < 0$ is a net payoff cost to self and that $-B(z) = \partial \pi(z_\bullet, z_{0,0}, n_{0,0})/\partial z_{0,0}$ is a payoff "benefit" to group neigbours as would occur for instance when individuals play a prisoner's dilemma game or more generally under the assumptions of models for the evolution of helping behavior. Then, eq. (22) entails that the cost-to-benefit ratio that needs to be overcome for helping to be favored is $\kappa_{0,0} > B(z)/C(z)$ with $\kappa_{0,0} = -1/(ND-1)$, which illustrates two notions. First, helping neighbours can never spread under a Wright-Fisher process regardless of the pattern of population structure since the cost to benefit ratio cannot be satisfied as long as $B(z) > 0$. For $D \to \infty$, this result was derived for a lattice structured population by Taylor (1992b) [and generalized to any abelian group structure by Taylor et al., 2011] and for finite D for a circular one dimensional habitat by Rousset (2004) [and generalized to any abelian group structure by Lehmann and Rousset, 2012]. For this latter case, we see that harming neighbours, i.e., $B(z) < 0$, can be favored by selection. Second, the scaled relatedness coefficient $\kappa_{\mathbf{k},t}$ can be more generally thought of as the cost-to-benefit ratio that needs to be overcome when interactions occur between individuals sampled at "distance" $\mathbf{k}$ at t times steps in the future, a perspective that allows to recover several results about the evolution of helping and harming behavior in lattice structured populations for different type of reproductive processes and assumptions about the effect of behavior on survival and reproduction (see Lehmann and Rousset, 2012 for more details). Finally, in the presence of ecologically mediated interactions, $\partial \pi(z_\bullet, z_{0,0}, n_{0,0})/\partial n_{0,0} > 0$, under a Wright-Fisher process. Setting $s_{\mathbf{k},t} = N \left(\partial \pi(z_\bullet, z_{0,0}, n_{0,0})/\partial n_{0,0} \right) e_{\mathbf{k},t}$, we recover eq. (A.21) in Lehmann (2008).

Our formalization of selection extends on this previous work in three novel directions. First, in the absence of environmentally mediated interactions, the model allows us to study the evolution of social interaction at any arbitrary, while the bulk of standard evolutionary game theory modeling in this area consider same patch or nearest-neighbour interactions. Second, for environmentally mediated interactions, we have an explicit representation of the extended phenotypic effect of an individual for all spatial positions and future times (recall eqs. 16–17), and, third, the expression for the selection gradient, eq. (22), holds for a broad class of life cycles (and are not limited to a specific one as usually in the literature). This together allows for a more general and less involved analysis of selection on environmentally mediated social interactions. For instance, when the payoff depends only on the phenotypes of neighbors and on the local state variable; namely, the payoff to the focal individual is

given by $\pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}}) = \pi(z_\bullet, z_{0,0}, n_{0,0})$. In this case, the selection gradient eq. (A-31) reduces to

$$s(z) \propto \frac{\partial \pi(z_\bullet, z_{0,0}, n_{0,0})}{\partial z_\bullet} + \frac{\partial \pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial z_{0,0}} \kappa_{0,0} + \frac{\partial \pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial n_{\mathbf{j},0}} NK, \tag{27}$$

where

$$K = \sum_{t=1}^{\infty} \sum_{\mathbf{k} \in \mathcal{G}} e_{\mathbf{k},t} \kappa_{\mathbf{k},t} \tag{28}$$

has a readily computable representation given by eq. (61).

We next illustrate these results with an example.

3.4 Inter-temporal helping and harming through a lasting commons

To gain more specific insights into how isolation by distance influences how selection shapes environmentally mediated interactions, we consider a scenario where the environmental variable n is some lasting commons (e.g., a common-pool resource, or a toxic compound) that can move in space, and whose production depends on the evolving trait, that is individually costly to express. The commons may be a "good" when $\hat{n} > 0$, or a "bad" when $\hat{n} < 0$; the trait leads to the production of the former when $z > 0$ and to the latter when $z < 0$. The trait can thus be broadly thought of as environmentally mediated helping (increasing survival and reproduction to recipients) when $z > 0$ and harming (decreasing survival and reproduction to recipients) when $z < 0$.

More specifically, we assume that fitness takes the form of eq. (21) with payoff given by

$$\pi(z_\bullet, z_{0,0}, n_{0,0}) = \exp\left(B n_{0,0}^{\alpha_B} - C z_\bullet^{\alpha_C}\right), \tag{29}$$

where $B > 0$ and $C > 0$ are parameters that respectively tune the effects on focal payoff of the commons in the focal patch $n_{0,0}$ and of the modifying trait $z_\bullet$ of the focal individual. These effects also depend on the shape parameters, $\alpha_B > 0$ and $\alpha_C > 0$, which we assume are respectively odd and even. We also assume that costs increase more steeply than benefits i.e., $\alpha_C > \alpha_B$. The local commons thus respectively increases and decreases payoff when $n_{0,0} > 0$ and $n_{0,0} < 0$, but any trait expression, i.e., any z away from 0, is individually costly and reduces payoff. Meanwhile, how the trait modifies the commons is determined by the function F (eq. 2), which we assume here is given by

$$F(\mathbf{z_{0,0}^R}, \mathbf{n_{0,0}}) = d_\mathbf{0}\left((1-\epsilon)n_{0,0} + P(z_{0,0}^R)\right) + \sum_{\mathbf{j} \in \mathcal{G} \backslash \mathbf{0}} d_\mathbf{j}\left((1-\epsilon)n_{\mathbf{j},0} + P(z_{\mathbf{j},0})\right). \tag{30}$$

In words, the commons changes from demographic time point to the next due to three processes. First, the commons is modified or "produced" in each patch according to a function $P : \mathbb{R} \to \mathbb{R}$ of the focal patch trait average, i.e., $z_{0,0}^R = \frac{1}{N} z_\bullet + \frac{N-1}{N} z_{0,0}$ in the focal patch, and $z_{\mathbf{j},0}$ otherwise (i.e., for $\mathbf{j} \neq \mathbf{0}$). We

assume that the function P is such that $P(0) = 0$ and that it is monotonically increasing with z, i.e., $P'(z) > 0$ holds for all $z \in \mathbb{R}$. Second, each unit of commons in a patch "diffuses" or moves with probability $d_{\mathbf{j}}$ to the patch at distance $\mathbf{j}$ from the source patch. The probability distribution defined by $d_{\mathbf{j}}$ can be thought of as the environmental equivalent of the dispersal probability distribution $m_{\mathbf{j}}$, and we let

$$\mathcal{D}(\mathbf{h}) = \sum_{\mathbf{i} \in \mathcal{G}} d_{\mathbf{i}} \chi_{\mathbf{i}}(\mathbf{h}) \tag{31}$$

denote the characteristic function of this distribution for future use (i.e., $\mathcal{D}(\mathbf{h})$ is to $d_{\mathbf{j}}$ as $\mathcal{M}(\mathbf{h})$ is to $m_{\mathbf{j}}$, eq. 9). Third, and finally, a unit of commons decays with rate $0 < \epsilon \leq 1$ from one time step to the next. Overall, eq. (30) entails that in a monomorphic population for z, the dynamics of the commons stabilises to

$$\hat{n} = \frac{P(z)}{\epsilon}, \tag{32}$$

which is positive when $z > 0$ and negative when $z < 0$, and whose absolute value increases as the rate of decay ϵ decreases, as expected.

We can then perform the analysis of selection described in Sections 3.2 and 3.3 for which we need to compute $e_{\mathbf{k},t}$, since $\kappa_{\mathbf{k},t}$ is given by the expressions in Box 2. Substituting eq. (30) into eq. (12) and eq. (13) we get $\psi_{\mathbf{k}} = d_{\mathbf{k}} P'(z)/N$, $c_{\mathbf{k}} = (1 - \epsilon)d_{\mathbf{k}}$, $\Psi(\mathbf{h}) = \mathcal{D}(\mathbf{h})P'(z)/N$, and $\mathcal{C}(\mathbf{h}) = \mathcal{D}(\mathbf{h})(1 - \epsilon)$, which yields $\mathcal{E}_t(\mathbf{h}) = \mathcal{D}(\mathbf{h})^t(1 - \epsilon)^{t-1}P'(z)/N$. Thereby the extended phenotypic effect $e_{\mathbf{k},t}$ to patch $\mathbf{k}, t$ (from eq. 16) can be expressed as

$$e_{\mathbf{k},t} = \frac{P'(z)}{N}(1 - \epsilon)^{t-1} q_{\mathbf{k},t}, \tag{33}$$

where

$$q_{\mathbf{k},t} = \frac{1}{D} \sum_{\mathbf{h} \in \mathcal{G}} \mathcal{D}(\mathbf{h})^t \overline{\chi}_{\mathbf{k}}(\mathbf{h}). \tag{34}$$

Equation eq. (33) can be interpreted as follows. A focal individual produces $P'(z)/N$ additional units of commons by changing trait value and each such unit decays through time according to $(1 - \epsilon)^{t-1}$ and influences patch $\mathbf{k}, t$ according to $q_{\mathbf{k},t}$, which can be interpreted as the probability that a non-decaying unit of the commons modified in the focal patch is located in patch $\mathbf{k}$ at t generations in the future. Extended phenotypic effects thus depend on the way the commons moves in space, as captured by the distribution $d_{\mathbf{i}}$ (see Fig. 3 for examples of $e_{\mathbf{k},t}$ in a 1D model).

How selection depends on extended phenotypic effects is found by substituting eq. (29) and eq. (33) into eq. (22). From this, we obtain

$$s(z) \propto BP'(z)\alpha_{\mathrm{B}} \left(\frac{P(z)}{\epsilon}\right)^{\alpha_{\mathrm{B}}-1} \Omega - C\alpha_C z^{\alpha_C - 1}, \tag{35}$$

where

$$\Omega = \sum_{t=1}^{\infty} \sum_{\mathbf{k} \in \mathcal{G}} \kappa_{\mathbf{k},t} (1 - \epsilon)^{t-1} q_{\mathbf{k},t} \qquad (36)$$

can be thought of as the expected genetic value of all the individuals in the future that are affected by a unit of the commons in the focal patch. As indicated by eq. (35), the greater this expectation, the greater the z favoured by selection. When at $z = 0$, $\alpha_B = 1$ and $P'(0) > 0$, the selection gradient reduces to $s(0) \propto \Omega$. This shows that in a population where the trait is initially absent so that individuals have no effect on the commons, selection favours environmental modifications leading to a common good ($z > 0$) when $\Omega > 0$ or to a common bad ($z < 0$) when $\Omega < 0$. Put differently, selection favours environmentally mediated inter-temporal helping when the recipient of such help has positive genetic value on average, and conversely, inter-temporal harming when it has negative genetic value.

Further information can be obtained if we additionally assume that P is linear, i.e., $P(z) = P_0 z$. In that case, we can rearrange eq. (35) to obtain that the singular trait value z^* is such that

$$(z^*)^{\alpha_C - \alpha_B} = \frac{B \alpha_B}{C \alpha_C} \cdot \frac{P_0^{\alpha_B}}{\epsilon^{\alpha_B - 1}} \cdot \Omega. \qquad (37)$$

This value can be shown to be convergence stable owing to our assumption that $\alpha_C > \alpha_B$ (note that since α_C is even and α_B is odd, $\alpha_C - \alpha_B$ is also odd). Eq. (37) shows that the absolute value of z^* increases with the benefit-to-cost ratio $B\alpha_B/(C\alpha_C)$ and with the environmental effect of the trait P_0. However, whether z^* is positive or negative, so whether helping or harming evolves, ultimately depends on the sign of Ω, i.e., whether the expected genetic value Ω of a modification to the commons is positive or negative.

How species dispersal and commons movement interact to determine Ω can be seen most easily by assuming that payoff influences fecundity under a Wright-Fisher process (i.e., where $f' > 0$ and $s' = 0$ and $s = 0$). In that case, helping corresponds to altruism and harming to spite (*sensu* Rousset, 2004, chapter 7) as an individual can never obtain direct benefits from its own trait expression. We show in Appendix F.1.1 that under Wright-Fisher life cycle, Ω can be expressed as

$$\Omega = \frac{D}{ND - 1} \sum_{t=1}^{\infty} (1 - \epsilon)^{t-1} \text{cov}_t^{\mathrm{T}}(p_t, q_t), \qquad (38)$$

where $\text{cov}_t^{\mathrm{T}}(p_t, q_t)$ is the covariance between the two distributions $p_t = (p_{\mathbf{k},t})_{\mathbf{k} \in \mathcal{G}}$ (eq. 26) and $q_t = (q_{\mathbf{k},t})_{\mathbf{k} \in \mathcal{G}}$ (eq. 34) summed over all spatial positions. This covariance is positive when there is a positive association between gene lineages and the commons these lineages modify. In other words, Ω is positive and altruism is favoured when an environmental modification owing to the expression of a gene is most likely to be experienced by future carriers of that same gene and its identical-by-descent copies. Conversely, Ω tends to be negative and spite is favoured when this environmental modification is less likely to be experienced by future carriers. Eq. (38) offers intuition on the biological condi-

tions leading to positive or negative Ω values, but it is more readily computed from the characteristic functions as

$$\Omega = \frac{1}{ND-1} \sum_{\mathbf{j} \in \mathcal{G} \backslash \mathbf{0}} \frac{\mathcal{D}(\mathbf{j})\mathcal{M}(\mathbf{j})}{1 - (1-\epsilon)\mathcal{D}(\mathbf{j})\mathcal{M}(\mathbf{j})}, \tag{39}$$

which is obtained from eq. (61) by noting that $\Omega = KN/P'(z)$. Figure 4AB gives the sign of Ω under a binomial model for the distance of both the dispersal of the focal species and the movement of the commons (using the characteristic functions detailed in Appendix F.3 to evaluate eq. (39)). These figures show that such model of dispersal allows for both positive and negative values of Ω. In particular, spite tends to be favoured by: (1) high levels of dispersal in the evolving species; (2) high levels of movement of the commons; (3) high environmental decay ϵ; and (4) large differences in the dispersal distance of the species and of the commons (e.g., when individuals disperse short distances while the commons move far away from their original patch). This is because these conditions tend to lead to a negative association between gene lineages and the commons these lineages modify. Conversely, altruism tends to be favoured when dispersal and movement are weak, environmental decay is low, and similar dispersal dispersal distributions of the species and of the commons (Figure 4AB, white region). In fact, under weak dispersal of the both the focal species and the commons (so that $m_0 = 1 - m$ and $d_0 = 1 - d$ with m and d close to zero), we show in Appendix F.1.2 that, regardless of the dispersal distribution,

$$\Omega = \left(\frac{D-1}{ND-1}\right)\left(\frac{1}{\epsilon} - \frac{m+d}{\epsilon^2}\right) \tag{40}$$

holds, which is always positive when m and d are small (smaller than ϵ).

Let us now briefly consider survival effects in the sense that payoff influences survival ($s' > 0$ and $f' = 0$). The expression of Ω for this case expressed in terms of characteristic dispersal functions can be found in Appendix F.2. Using this, we show in Fig 5 that selection is not fundamentally different than when payoff influences fecundity. One further insight that the analysis of survival effects provides is that harming tends to be favoured when baseline survival s is low, especially when environmental decay is low (Fig 5C). This is because otherwise, an individual may harm itself in the future.

4   Discussion

The continual feedback between organisms and their environment makes the evolution of genetic traits that influence this feedback challenging to track mathematically, and consequently to understand. This is especially true in spatially structured populations, where local sampling effects lead to an explosion of the evolutionary and ecological state space on which eco-evolutionary dynamics occur. Here, we contributed to solving this problem by characterising the selection gradient on traits that impact dynamics of environmental state variables which then feed back on survival and reproduction, when the population is subdivided among patches arranged such that fitness and environmental effects are

homogeneous in space. We described the selection gradient in two complementary ways. Each makes the evolutionary analysis of environmentally mediated social interactions readily tractable, computationally convenient, and easily connected to fundamental biological processes.

First, we expressed the selection gradient in terms of an inclusive fitness effect; namely a sum of relatedness-weighted fitness differentials capturing how an individual expressing a change in trait value influences (i) its own fitness and that of current relatives living at the different spatial positions of the population; and (ii) the fitness of all downstream relatives living at the different spatial positions through modifications of the local values of the environmental state variables these individuals experience (eq. 5 with eq. 6a and eq. 18 or eq. 50 for the multi-dimensional environmental case). These modifications are captured by a formalization of the extended phenotypic effect (*sensu* Dawkins, 1982) of an individual on its environment. We showed that this extended phenotypic effect takes the form of press perturbations of ecology (Yodzis, 1989; Novak et al., 2016), i.e., the intensity by which a unit perturbation in the dynamics of the environmental states variable induced by a change in an individual's trait expression propagates to change the values of this variable to all spatial positions and all future times in the population (eq. 50). Exactly how such effects will be distributed in space and time depends on the specificity of environmental dynamics, e.g., how it diffuses or disperses, how it decays, or how it is added.

Second, we showed that when interactions between individuals are mediated through a payoff function, which in turn determines survival and reproduction, then the selection gradient can be expressed in terms of an inclusive fitness payoff of how an individual expressing a change in trait value influences (i) its own payoff and that of current relatives; and (ii) the payoff of all downstream relatives and where the weights in the inclusive fitness are given by extended phenotypic effect and scaled-relatedness coefficients (eqs. (22)–(24)). A scaled relatedness coefficient to a set of individuals inhabitating a certain spatial position of the habitat in the current or downstream generation can be interpreted as a marginal rate of substitution; namely, as of the number of units of own payoff a focal individual is willing to exchange at an evolutionary equilibrium for one unit of payoff accruing to the whole set of individuals living in the patch under consideration. The value of a scaled relatedness depends on the type of competition that individuals experience within and between patches and thus on the details of the life cycle, whose consequences can thus be analyzed, e.g., the intensity of kin competition can be measured by scaled relatedness. For our core life-cycle assumptions (Section (2)) where individual can survive over demographic time points, we were able to fully resolve the scaled relatedness coefficient and obtain a readily computable representation of the selection gradient that can be applied to investigate the evolution of environmentally mediated interactions under various scenarios. Our formalization was particularly influenced and tuned to investigate the evolution of traits affecting the evolution of environmentally mediated social dilemmas *sensu* Estrela et al. (2019).

As an application of these results, we derived the selection gradient on a trait affecting the production

of a commons that can move arbitrarily in space, and whose production comes at a fecundity and/or survival cost to individuals. We evaluated the convergence stable level of investment of the resource for both fecundity and the survival effects. The analysis under fecundity effects reveals that where there is a negative covariance in the spatial pattern of dispersal of individuals of the evolving species and the movement of the commons, then inter-generational indiscriminate harming that is spiteful can readily evolve by natural selection. By "spite", we mean a trait whose expression decreases the fitness of the individual actor and results in a decrease in the survival or fecundity of the primary recipients of the extended phenotypic effect of the trait. This is a strong form of spite (*sensu* Rousset, 2004, see chapter 7), which needs to be contrasted to the examples of weak spite reported in Hamilton (1971) and Schaffer (1988), where the behavior directly benefit the actor. As illustrated by Fig. 4, there is scope for the evolution of spite at a distance under isolation by distance in finite populations under both fecundity and survival effects. Our explicit model shows that this can indeed occur readily when the commons is such as to fall predominantly on negatively related individuals. Otherwise, inter-generational altruism is more likely to evolve.

Two points about our finding about inter-generational spite are worth mentioning. First, previous model of the evolution of strong spite considered the evolution of discriminate spite, where there is either implicitly or explicitly discrimination between individuals (e.g., Gardner and West, 2004; Lehmann et al., 2009; West and Gardner, 2010; Smead and Forber, 2016; Bruner and Smead, 2022). By contrast, our model displays the evolution of indiscriminate spite. Of course, spite must preferentially benefit the spiteful lineage through decrease of competition, but at a proximal level there is no recognition and discrimination by individual to preferentially target the trait effects (in contrast to previous work). Second, we observed the evolution of spite in populations whose sizes are non-neglibible (e.g., of local size 50, Fig. 4). More generally, the analysis of our example shows that whether natural selection favors anyone of the particular type of environmentally mediated social behavior, which have been classified in terms of of "enrich", "detox","pollute", and "deplete" scenarios (Estrela et al., 2019), depends on the spatial distribution of both evolving species and the common-pool resource.

Overall, our analysis demonstrates that the evolution of environmentally mediated social interactions is tractable under different life-cycle assumptions and biologically relevant class of isolation-by-distance models of population structure when the underlying dynamics of states variables are locally deterministic. An interesting avenue for future work would be to investigate whether these results also allow to approximate multi-species co-evolution under stochastic local dynamics for isolation by distance, as such approximations have actually been shown to hold for the island model of dispersal (Mullon and Lehmann, 2018). Co-evolution of different traits such as dispersal and allocation efforts could also be considered under our framework. More broadly, we hope that our model will serve as a source of inspiration to further explore the evolution of environmentally mediated social behavior.

Box 1. Fourier analysis on finite abelian groups. We assume that the set of patches $\mathcal{G}$ is endowed with an abelian group structure, which entails that it can be written as a direct product of finite circles, i.e.,

$$\mathcal{G} = \mathbb{Z}_{D_1} \times \ldots \times \mathbb{Z}_{D_d}, \tag{41}$$

where $\mathbb{Z}_{D_i} = \{0, \ldots, D_i - 1\}$ is the (additive) group of integers modulo D_i and d can be interpreted as the number of "dimensions" of the population. The group $\mathcal{G}$ defined by the right-hand side of eq. (41) consists of all vectors $\mathbf{x} = (x_1, \ldots, x_d)$ where $x_i \in \mathbb{Z}_{D_i}$, and the addition of two vectors is component-wise. We define the discrete Fourier transform $\mathcal{F}(\mathbf{x})$ of function f at $\mathbf{x}$ by

$$\mathcal{F}(\mathbf{x}) = \sum_{\mathbf{y} \in \mathcal{G}} f_{\mathbf{y}} \chi_{\mathbf{y}}(\mathbf{x}), \tag{42}$$

where the "character" function

$$\chi_{\mathbf{y}}(\mathbf{x}) = \prod_{i=1}^{d} \exp\left(\frac{2\pi \iota x_i y_i}{D_i}\right) = \exp\left(2\pi \iota \sum_{i=1}^{d} \frac{x_i y_i}{D_i}\right) \tag{43}$$

is defined for all $\mathbf{x} = (x_1, \ldots, x_d) \in \mathcal{G}$ and $\mathbf{y} = (y_1, \ldots, y_d) \in \mathcal{G}$ and where $\iota = \sqrt{-1}$ (recall that the total number of patches is $D = \prod_{i=1}^{d} D_i$). Here, we followed the convention of population genetics (e.g., Malécot, 1975; Sawyer, 1976; Rousset, 2004) and probability theory in defining the Fourier transform in terms of the character $\chi_{\mathbf{y}}(\mathbf{x})$ (instead of defining it in terms of its conjugate given by eq. 45, which is more standard in mathematics and engineering), whereby the Fourier transform defines a characteristic function when f is a probability distribution. For instance, the Fourier transform $\mathcal{M}(\mathbf{h}) = \sum_{\mathbf{k} \in \mathcal{G}} m_{\mathbf{k}} \chi_{\mathbf{k}}(\mathbf{h})$ is the characteristic function of the dispersal distribution. Discrete Fourier transforms are useful in the context of our analysis because knowing the Fourier transform of a function allows us to obtain the original function by using the inversion property

$$f_{\mathbf{x}} = \mathcal{L}_{\mathbf{x}}(\mathcal{F}) = \frac{1}{D} \sum_{\mathbf{y} \in \mathcal{G}} \mathcal{F}(\mathbf{y}) \overline{\chi}_{\mathbf{x}}(\mathbf{y}), \tag{44}$$

where $\mathcal{L}_{\mathbf{x}}(\mathcal{F})$ is the inverse transform of $\mathcal{F}$ at $\mathbf{x}$, which is defined in terms of the (complex) conjugate of $\chi_{\mathbf{y}}(\mathbf{x})$:

$$\overline{\chi}_{\mathbf{x}}(\mathbf{y}) = \prod_{i=1}^{d} \exp\left(-\frac{2\pi \iota x_i y_i}{D_i}\right) = \exp\left(-2\pi \iota \sum_{i=1}^{d} \frac{x_i y_i}{D_i}\right) \tag{45}$$

(e.g., Terras). Another useful property that we use in our analysis is the orthogonality relation of the characters:

$$\sum_{\mathbf{k} \in \mathcal{G}} \overline{\chi}_{\mathbf{k}}(\mathbf{i}) \overline{\chi}_{\mathbf{k}}(\mathbf{j}) = \sum_{\mathbf{k} \in \mathcal{G}} \overline{\chi}_{\mathbf{k}}(\mathbf{i} + \mathbf{j}) = \begin{cases} |\mathcal{G}| = D & \text{if } \mathbf{j} + \mathbf{i} = \mathbf{0} \\ 0 & \text{otherwise,} \end{cases} \tag{46}$$

(Terras, 1999, p. 169).

Box 2. Multi-dimensional environment. We here generalize $s(z) = s_{\mathrm{w}}(z) + s_{\mathrm{e}}(z)$ to the case where there are $n_{\mathrm{e}} > 1$ environmental state variables and denote by $\mathbf{n}_{\mathbf{k},t} = (n_{1,\mathbf{k},t}, n_{2,\mathbf{k},t}, ... n_{n_{\mathrm{e}},\mathbf{k},t})$ the vector of such variables in patch $\mathbf{k}, t$ with $n_{i,\mathbf{k},t} \in \mathbb{R}$ denoting the value of the ith environment. Then, the fitness of the focal individual is given by

$$w(z_\bullet, \mathbf{z}_{\mathbf{0},0}, \vec{\boldsymbol{n}}_{\mathbf{0},0}) \qquad (\text{formally } w : \mathbb{R} \times \mathbb{R}^D \times \mathbb{R}^{D n_{\mathrm{e}}} \to \mathbb{R}_+), \tag{47}$$

where $\vec{\boldsymbol{n}}_{\mathbf{0},0}$ is the values in $t = 0$ of $\vec{\boldsymbol{n}}_{\mathbf{0},t} = (\mathbf{n}_{\mathbf{0},t}, \mathbf{n}_{\mathbf{1},t}, ..., \mathbf{n}_{\mathbf{k},t}, ..., \mathbf{n}_{\mathbf{D}_{-1},t})$ whose elements are solutions of

$$n_{i,\mathbf{k},t+1} = F_i(\mathbf{z}_{\mathbf{k},t}, \vec{\boldsymbol{n}}_{\mathbf{k},t}) \qquad \text{for } i = 1, 2, .., n_{\mathrm{e}} \text{ and all } \mathbf{k} \in \mathcal{G}, \tag{48}$$

where F_i is the transition map for environmental variable i. As for the one dimensional case, we assume that in a monomorphic population z, there is a hyperbollically stable fixed point to the environmental dynamics

$$\hat{n}_i = F_i(\mathbf{z}, \hat{\vec{\boldsymbol{n}}}) \qquad \text{for } i = 1, 2, ..., n_{\mathrm{e}}, \tag{49}$$

where $\hat{\vec{\boldsymbol{n}}} = (\hat{\boldsymbol{n}}, ..., \hat{\boldsymbol{n}})$ is a vector of dimension D whose entries are all given by $\hat{\boldsymbol{n}} = (\hat{n}_1, ..., \hat{n}_{n_{\mathrm{e}}})$. With fitness model eq. (47), the selection component $s_{\mathrm{w}}(z)$ is formally unchanged and given by eq. (6a) by substituting $w(z_\bullet, \mathbf{z}_{\mathbf{0},0}, \vec{\boldsymbol{n}}_{\mathbf{0},0}(\mathbf{z}_{\mathrm{H}}))$ for $w(z_\bullet, \mathbf{z}_{\mathbf{0},0}, \mathbf{n}_{\mathbf{0},0}(\mathbf{z}_{\mathrm{H}}))$. For selection component $s_{\mathrm{e}}(z)$, we carry out *mutadis mutandis* exactly the same calculations for the multi-dimensional case as we carried out in Section 3.2 and Appendix B.1 for the one-dimensional case to obtain

$$s_{\mathrm{e}}(z) = \sum_{t=1}^{\infty} \sum_{i=1}^{n_{\mathrm{e}}} \sum_{\mathbf{j} \in \mathcal{G}} \sum_{\mathbf{k} \in \mathcal{G}} N \frac{\partial w(z_\bullet, \mathbf{z}_{\mathbf{0},0}, \vec{\boldsymbol{n}}_{\mathbf{0},0})}{\partial n_{i,\mathbf{j},0}} e_{i,\mathbf{j}-\mathbf{k},t} R_{\mathbf{k},t}, \tag{50}$$

where $e_{i,\mathbf{k},t}$ is the extended phenotypic effect on environmental variable i in patch $\mathbf{k}, t$. This is obtained as the inverse transform

$$e_{i,\mathbf{k},t} = \frac{1}{D} \sum_{\mathbf{h} \in \mathcal{G}} \mathcal{E}_{i,t}(\mathbf{h}) \overline{\chi}_{\mathbf{k}}(\mathbf{h}), \tag{51}$$

where $\mathcal{E}_{i,t}(\mathbf{h})$ is the ith element of the vector $\vec{\mathcal{E}}_t(\mathbf{h}) = (\mathcal{E}_{1,t}(\mathbf{h}), \mathcal{E}_{2,t}(\mathbf{h}), ..., \mathcal{E}_{n_{\mathrm{e}},t}(\mathbf{h}))$ which is obtained from

$$\vec{\mathcal{E}}_t(\mathbf{h}) = C(\mathbf{h})^{t-1} \vec{\Psi}(\mathbf{h}). \tag{52}$$

Here, the community matrix $C(\mathbf{h})$ has its ij-th element given by $\mathcal{C}_{ij}(\boldsymbol{\theta}) = \sum_{\mathbf{k} \in \mathcal{G}} c_{i,\mathbf{0} \leftarrow j,\mathbf{k}} \chi_{\mathbf{k}}(\mathbf{h})$, where

$$c_{i,\mathbf{0} \leftarrow j,\mathbf{k}} = \frac{\partial F_i(\mathbf{z}_{\mathbf{0},0}^{\mathrm{R}}, \vec{\boldsymbol{n}}_{\mathbf{0},0})}{\partial n_{j,\mathbf{k},t}}, \tag{53}$$

which can be interpreted as the effect of the variation of the abundance of environment j in the focal patch on the abundance of environment i located in patch $\mathbf{k}$. The vector $\vec{\Psi}(\mathbf{h})$ has its i-th element given by $\Psi_i(\mathbf{h}) = \sum_{\mathbf{k} \in \mathcal{G}} \psi_{\mathbf{k},i} \chi_{\mathbf{k}}(\mathbf{h})$, which is the Fourier transform of

$$\psi_{i,\mathbf{k}} = \begin{cases} \frac{1}{N} \frac{\partial F_i(\mathbf{z}_{\mathbf{0},0}^{\mathrm{R}}, \vec{\boldsymbol{n}}_{\mathbf{0},0})}{\partial z_{\mathbf{0},0}^{\mathrm{R}}} & \text{for } \mathbf{k} = \mathbf{0} \\ \frac{1}{N} \frac{\partial F_i(\mathbf{z}_{\mathbf{0},0}^{\mathrm{R}}, \vec{\boldsymbol{n}}_{\mathbf{0},0})}{\partial z_{\mathbf{k},0}} & \text{otherwise.} \end{cases} \tag{54}$$

For an infinite island model of dispersal, $R_{\mathbf{k},t} = 0$, $e_{i,\mathbf{k},t} = 0$, and $c_{i,\mathbf{0} \leftarrow j,\mathbf{k}} = 0$ for all $\mathbf{k} \in \mathcal{G}$ except $\mathbf{k} = \mathbf{0}$, whereby eq. (50) reduces to eqs. 15-16 of Mullon and Lehmann (2018).

Box 3. Scaled relatedness coefficients. For the fitness model defined by eq (21), we show in Appendix E that

$$
\kappa_{\mathbf{k},t} = \begin{cases} \dfrac{\mathcal{L}_{\mathbf{k}}(F)-(1+s)[s'f+2f'(1-s)]/(2D)}{N[s'f+f'(1-s)]+\mathcal{L}_0(F)-(1+s)[s'f+2f'(1-s)]/(2D)} & \text{if } t=0 \\[1em] \dfrac{\mathcal{L}_{\mathbf{k}}(G_t)-(1+s)[s'f+2f'(1-s)]/(2D)}{N[s'f+f'(1-s)]+\mathcal{L}_0(F)-(1+s)[s'f+2f'(1-s)]/(2D)} & \text{otherwise,} \end{cases}
\tag{55}
$$

where $f' = \partial f\,(\pi_\bullet)\,/\partial\pi_\bullet$, $s' = \partial s\,(\pi_\bullet)\,/\partial\pi_\bullet$, all functions s, s', f, and f', are as usual, evaluated at the resident trait value z, and $\mathcal{L}_{\mathbf{k}}(D)$ is the inverse Fourier transform of a the function D at $\mathbf{k}$ (eq. (44) of Box 1). Here the functions F and G_t are defined in $\mathbf{h}$ as

$$
\begin{aligned}
F(\mathbf{h}) &= -\frac{(1-s)\,[s'f - f'2s]\,\mathcal{M}(\mathbf{h})}{1+s+(1-s)\mathcal{M}(\mathbf{h})}, \\
G_t(\mathbf{h}) &= \frac{(1+s)\,[s'f + f'(1-s)(1+\mathcal{M}(\mathbf{h}))]\,[s+(1-s)\mathcal{M}(\mathbf{h})]^t}{1+s+(1-s)\mathcal{M}(\mathbf{h})}.
\end{aligned}
\tag{56}
$$

For fecundity effects only ($s'=0$), eq. (55) reduces to

$$
\kappa_{\mathbf{k},t} = \begin{cases} \dfrac{\mathcal{L}_{\mathbf{k}}(F)-(1+s)/D}{N+\mathcal{L}_0(F)-(1+s)/D} & \text{if } t=0 \\[1em] \dfrac{\mathcal{L}_{\mathbf{k}}(G_t)-(1+s)/D}{N+\mathcal{L}_0(F)-(1+s)/D} & \text{otherwise,} \end{cases}
\tag{57}
$$

where

$$
\begin{aligned}
F(\mathbf{h}) &= \frac{2s\mathcal{M}(\mathbf{h})}{1+s+(1-s)\mathcal{M}(\mathbf{h})}, \\
G_t(\mathbf{h}) &= \frac{(1+s)(1+\mathcal{M}(\mathbf{h}))\,[s+(1-s)\mathcal{M}(\mathbf{h})]^t}{1+s+(1-s)\mathcal{M}(\mathbf{h})}.
\end{aligned}
\tag{58}
$$

Note that these eq. (56)–(58) apply to the case where $s>0$ but the evolving trait does not affect survival. For a Wrigh-Fisher process, we further have $s=0$, whereby $F(\mathbf{h})=0$ and $G_t(\mathbf{h})=\mathcal{M}(\mathbf{h})^t$, which yields eq. (25). By contrast, for only survival effects, ($f'=0$), eq. (55) reduces to

$$
\kappa_{\mathbf{k},t} = \begin{cases} \dfrac{\mathcal{L}_{\mathbf{k}}(F)-(1+s)/(2D)}{N+\mathcal{L}_0(F)-(1+s)/(2D)} & \text{if } t=0 \\[1em] \dfrac{\mathcal{L}_{\mathbf{k}}(G_t)-(1+s)/(2D)}{N+\mathcal{L}_0(F)-(1+s)/(2D)} & \text{otherwise,} \end{cases}
\tag{59}
$$

where

$$
\begin{aligned}
F(\mathbf{h}) &= -\frac{(1-s)\mathcal{M}(\mathbf{h})}{1+s+(1-s)\mathcal{M}(\mathbf{h})}, \\
G_t(\mathbf{h}) &= \frac{(1+s)\,[s+(1-s)\mathcal{M}(\mathbf{h})]^t}{1+s+(1-s)\mathcal{M}(\mathbf{h})}.
\end{aligned}
\tag{60}
$$

Using eq (16) and eq. (55) we show in Appendix that

$$
K = \frac{1}{H}\sum_{\mathbf{j}\in\mathcal{G}\setminus\mathbf{0}} \frac{(1+s)\,[s'f + f'(1-s)(1+\mathcal{M}(\mathbf{j}))]\,[s+(1-s)\mathcal{M}(\mathbf{j})-\mathcal{C}(-\mathbf{j})\mathcal{M}(\mathbf{j})]\,\Psi(-\mathbf{j})}{[1+s+(1-s)\mathcal{M}(\mathbf{j})]\,[1-\mathcal{C}(-\mathbf{j})]\,[1-\mathcal{C}(-\mathbf{j})\mathcal{M}(\mathbf{j})]},
\tag{61}
$$

where $\mathcal{C}(-\mathbf{j}) = \mathcal{C}(\mathbf{j})$ and $\Psi(-\mathbf{j}) = \Psi(\mathbf{j})$ under symmetry of the commons migration and phenotypic effects (which is a natural assumption for lattice structured populations), and

$$
H = (DN-1)\,[s'f + f'(1-s)] - \sum_{\mathbf{j}\in\mathcal{G}\setminus\mathbf{0}} \frac{(1-s)\,[s'f-2f's]\,\mathcal{M}(\mathbf{j})}{1+s+(1-s)\mathcal{M}(\mathbf{j})}.
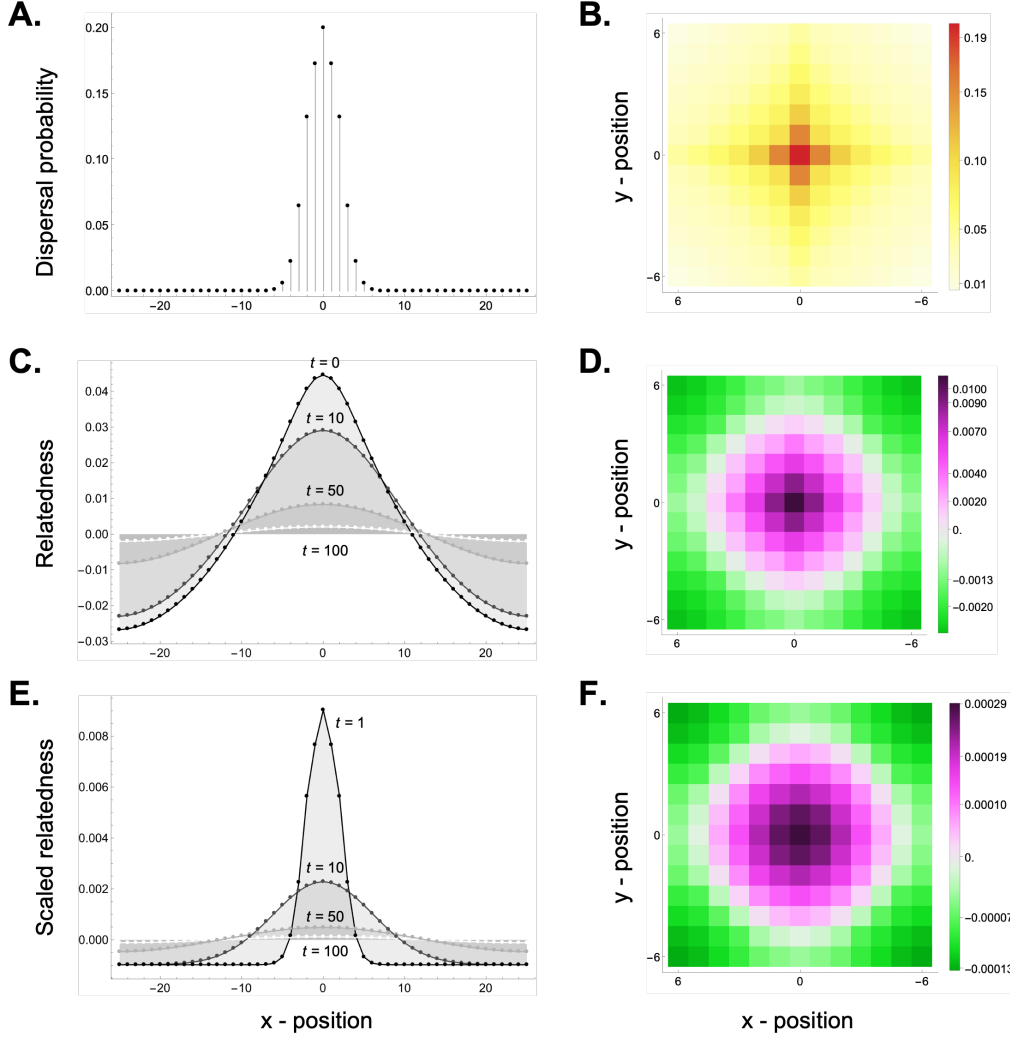\tag{62}
$$

Figures

Figure 1: **Dispersal distribution, relatedness and scaled relatedness with short range dispersal.** Panels **A-B.** Dispersal distribution $m_{\mathbf{k}}$ in a lattice-structured population in a one-dimensional habitat (henceforth 1D, panel A with $D_1 = 51$) and a two-dimensional habitat (2D, panel B with $D_1 = D_2 = 13$). An offspring leaves its natal patch with probability $1 - m_{\mathbf{0}} = m = 0.8$ and disperses to a patch at a Manhattan distance that follows a truncated binomial distribution (eq. A-71 in Appendix F.3) with mean $\lambda_{\mathrm{m}} = 1.5$, leading to short-range dispersal in this example (see Appendix F.3 for details). The distance dispersed along each dimension of the habitat is uniformly distributed across all dimensions and directions (see Appendix F.3 for details). Panels **C-D.** Relatedness $R_{\mathbf{k},t}$ for the 1D and 2D population settings shown in panels A-B (using eq. 8 with patch size $N = 20$ and no adult survival $s = 0$). Panel C displays the inter-temporal relatedness coefficients $R_{\mathbf{k},0}$, $R_{\mathbf{k},10}$, $R_{\mathbf{k},50}$ and $R_{\mathbf{k},100}$ (from dark to light gray) highlighting how relatedness decays in time and space, becoming negative away from the focal deme when dispersal is local (i.e., peaked close to 0). Panel D displays the intra-temporal relatedness $R_{\mathbf{k},0}$. Panels **E-F.** Scaled relatedness $\kappa_{\mathbf{k},t}$ for the 1D and 2D population settings shown shown in A-B for a Wright-Fisher model with fecundity effects (using eq. 25 with patch size $N = 20$). Panel E displays $\kappa_{\mathbf{k},1}$, $\kappa_{\mathbf{k},10}$, $\kappa_{\mathbf{k},50}$ and $\kappa_{\mathbf{k},100}$ (from dark to light gray) while panel F displays $\kappa_{\mathbf{k},10}$.
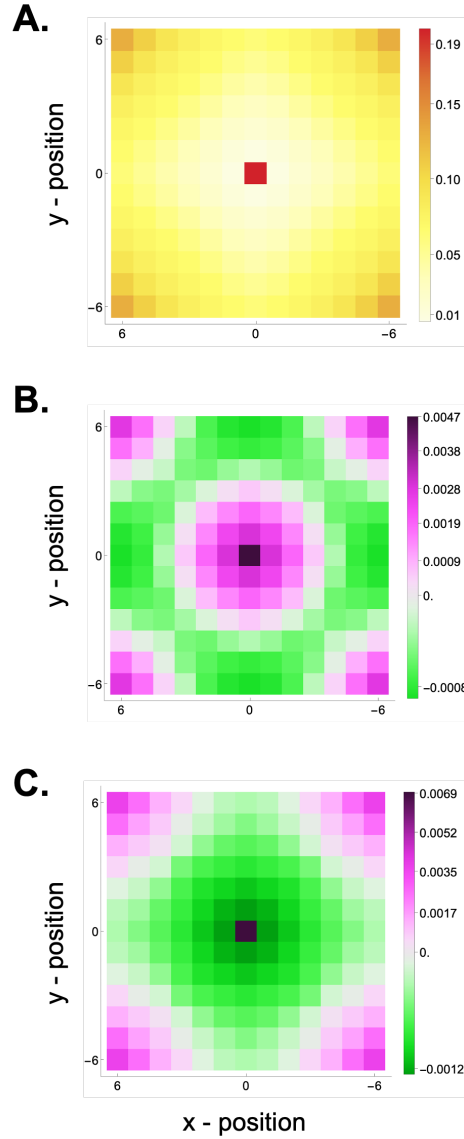
Figure 2: **Dispersal distribution, relatedness and scaled relatedness in 2D with long range dispersal.** Panel **A.** Dispersal distribution $m_{\mathbf{k}}$ in 2D (with $D_1 = D_2 = 13$). An offspring leaves its natal patch with probability $1 - m_0 = m = 0.8$, and if so, the Manhattan distance dispersed follows a Binomial distribution with mean $\lambda_{\mathbf{m}} = 11$ (Appendix Appendix F.3 for details). This leads to long range dispersal whereby dispersing offspring immigrate into patches far away from their natal patch. Panel **B.** Relatedness $R_{\mathbf{k},0}$ from the dispersal distributions shown in A (using eq. 8 with patch size $N = 20$ and no adult survival $s = 0$). In contrast to the case of short range dispersal (Fig. 1D), relatedness here is negative between individuals living at intermediate, rather than long, distance. Panel **C.** Scaled relatedness $\kappa_{\mathbf{k},1}$ from the dispersal distributions shown in panel A for a Wright-Fisher model (using eq. 25 with patch size $N = 20$). inter-generational scaled relatedness here is positive between individuals living in the same patch or separated by long distances, and negative otherwise.
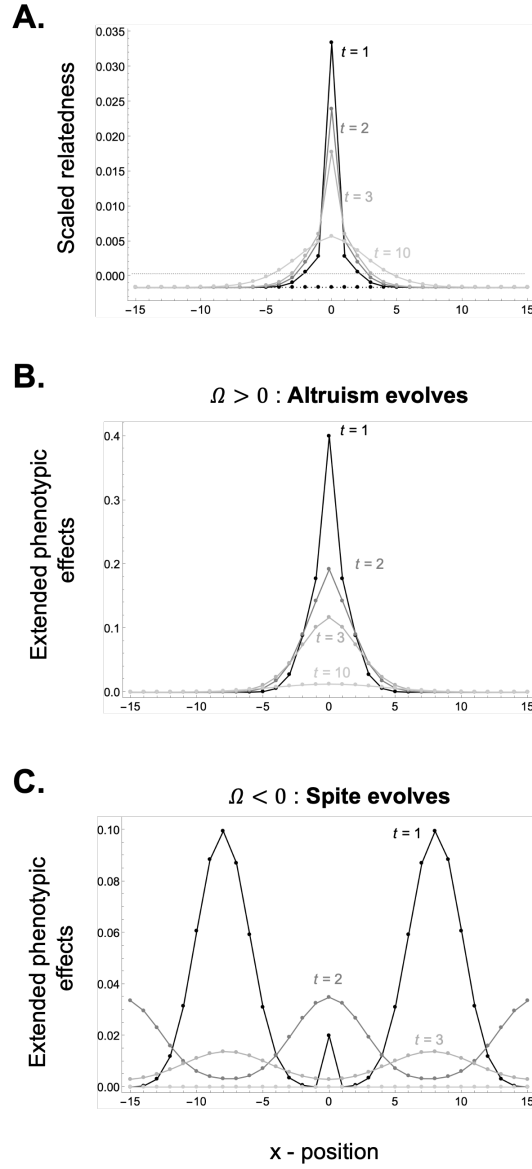
Figure 3: **How scaled relatedness and extended phenotypic effects determine whether altruism or spite evolves.** Whether selection favours $z > 0$ (altruism) or $z < 0$ (spite) is determined by Ω (eq. 36), which consists in the convolution between scaled relatedness $\kappa_{\mathbf{k},t}$ and extended phenotypic effects $e_{\mathbf{k},t}$. In this 1D example of the evolution of z (with $D_1 = 31$) displayed in panels **A**-**C** dispersal is short range (with $m = 0.3$ and $\lambda_{\mathrm{m}} = 1$). Scaled relatedness $\kappa_{\mathbf{k},t}$ shown in panel **A** decays in time and space away from the focal deme (using eq. 25 with patch size $N = 20$). When the resource moves locally, extended phenotypic effects $e_{\mathbf{k},t}$ also decay in time and space away from the focal deme as shown in panel **B** (using eq. 33 with movement parameters $d = 0.6$ and $\lambda_{\mathrm{d}} = 1$; production function $P(z) = Nz$, i.e., each unit of z contribute to one unit of resource; and decay rate $\epsilon = 0.2$). This overlap between scaled relatedness and extended phenotypic effects leads to $\Omega = 0.033 > 0$, indicating that selection favours the evolution of altruism here. This is because individuals that have the greatest genetic value (greatest $\kappa_{\mathbf{k},t}$) are most likely to be affected by resource modifications (which are quantified by $e_{\mathbf{k},t}$). In contrast, when the resource moves greater distances, extended phenotypic effects $e_{\mathbf{k},t}$ are greatest further away from the focal deme, as shown in panel **C** (using eq. 33 with movement parameters $d = 0.98$ and $\lambda_{\mathrm{d}} = 8$; production function $P(z) = Nz$; and decay rate $\epsilon = 0.5$). As a result, these effects disproportionately affect individuals with negative genetic value (i.e., where $\kappa_{\mathbf{k},t} < 0$ see A), thus favouring the evolution of spite (as indicated here by $\Omega = -0.00027 < 0$).
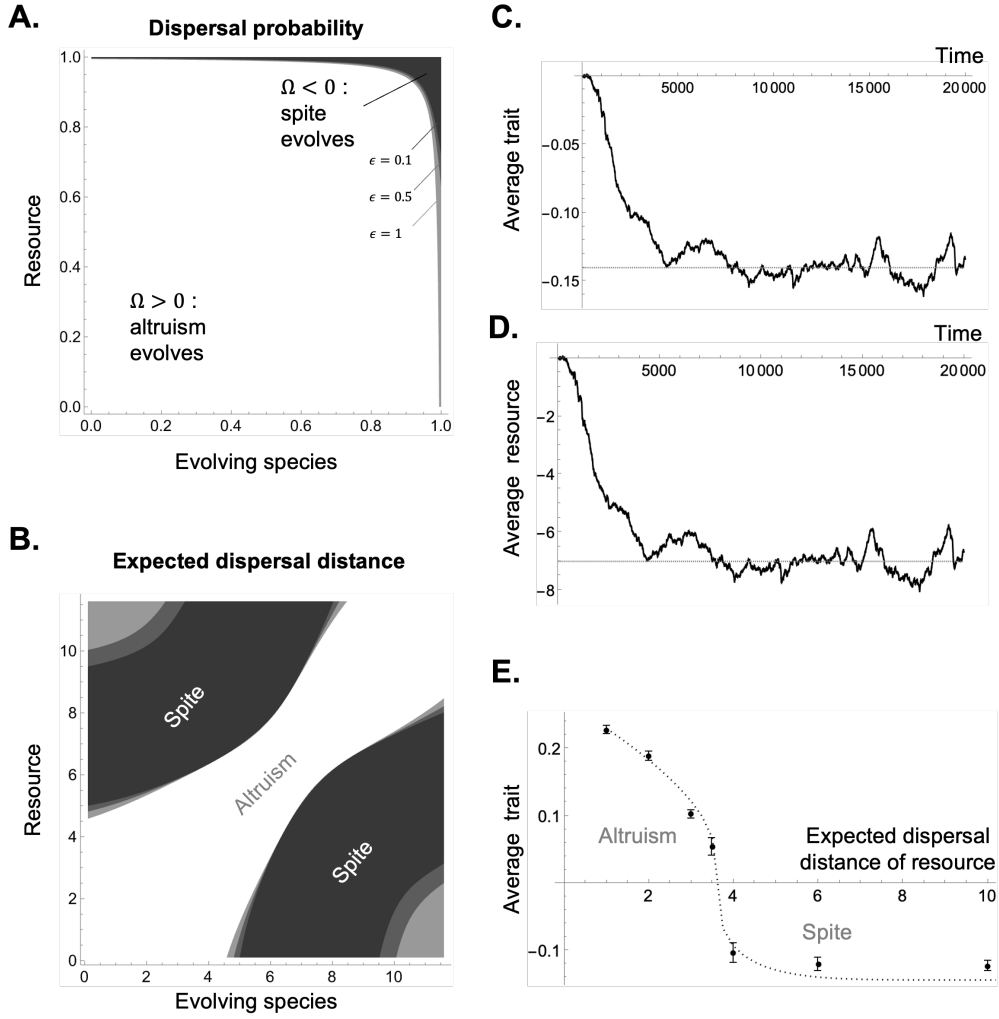
27

Figure 4: **Selection favours altruism or spite depending on dispersal of the evolving species and the resource.** Panels **A-B**. Regions of dispersal parameters leading to the evolution of altruism, $\Omega > 0$ (in white), or of spite, $\Omega < 0$ (in gray) for an example in 2D (with $D_1 = D_2 = 13$ and $N = 50$) under a Wright-Fisher life-cycle with fecundity effects (with Ω computed from eq. 39). **A**: Combination of dispersal probability of the evolving species m (x-axis) and of the resource d (y-axis) for different levels of environmental decay ϵ in different shades of gray ($\epsilon = 0.1, 0.5, 1$) with expected dispersal distance fixed ($\lambda_m = 1$ and $\lambda_d = 8$). This shows that spite is favoured by high levels of dispersal and high levels of environmental decay. **B**: Combination of expected dispersal distance of the evolving species λ_m (x-axis) and of the resource λ_d (y-axis) for different levels of environmental decay ϵ in different shades of gray (A for legend) with dispersal probability fixed ($m = 0.98$ and $d = 1$). This shows that spite is favoured by dispersal asymmetry between the evolving species and the resource. **C-D** Evolution of spite in individual based simulations under a Wright-Fisher life-cycle with fecundity effects (with $D_1 = D_2 = 13$, $N = 50$, $m = 0.3$, $\lambda_m = 1$, $d = 1$, $\lambda_d = 8$, $B = 2$ $\alpha_B = 1$, $C = 1$, $\alpha_C = 4$, $P(z) = Nz$; for mutation: the trait mutates during reproduction with probability 10^{-4}, in which case a normally distributed deviation with mean 0 and standard deviation 10^{-2} is added to the parental trait value). C shows the average trait z in the population and D shows the average resource level or environmental variable n (with simulations in full and analytical prediction in dashed – from eq. 37 for z and 32 for n). **E** Observed *vs.* predicted equilibrium trait value in individual based simulations running for 20'000 generations under different expected dispersal distance of the resource λ_d leading to altruism ($z > 0$) and spite ($z < 0$). Other parameters: same as in C-D. Prediction is computed from eq. (37) and shown as a dashed line. Observed values of the trait average in the population are shown as black dots for the average from generation 5000 to 15000, with error bars for standard deviation over the same 15000 generations. Simulations were initialised at the predicted convergence stable trait value.
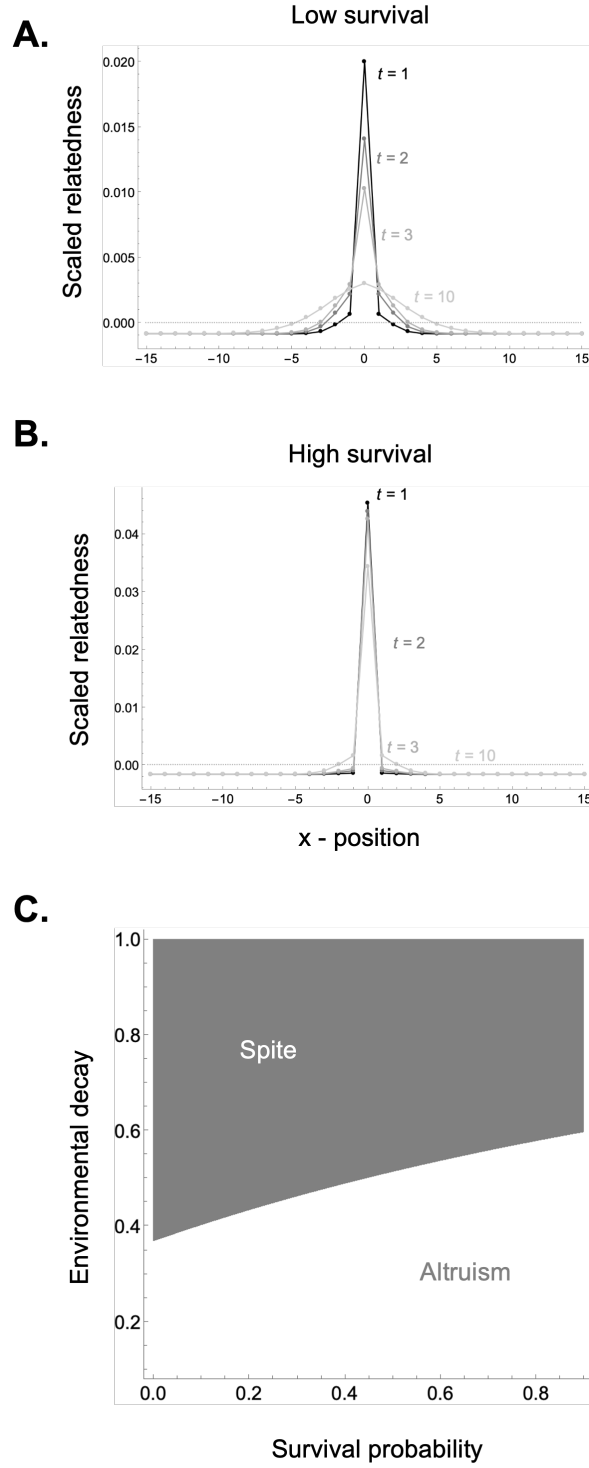
28

Figure 5: **Scaled relatedness and selection under survival effects. A-B:** Scaled relatedness $\kappa_{\mathbf{k},t}$ in 1D under survival effects (from eqs. (59) and (60), with $s = 0$ in A and $s = 0.9$ in B; other parameters: same as in Fig. 3A). These shows that genetic value decays away from the focal deme especially quickly when baseline survival is high (compare A and B). Otherwise, these profiles of scaled relatedness are similar to to those in Fig. 3A, which suggests that selection act similarly when the trait affects survival or fecundity. **C.** Parameter region where selection favours the evolution of altruism $\Omega > 0$ or spite $\Omega < 0$ under survival effects with adult survival probability s on the x-axis and environmental decay ϵ on the y-axis (Ω computed from eq. 36 using eq. A-69; other parameters: same as in Fig. 3C.

Appendix A   Convergence stability from fixation probability

We here show eq. (4) from fixation probability considerations. Suppose then we introduce a single mutant with trait value $z + \delta$ into a population monomorphic for resident trait z and let $\Pi(z + \delta, z)$ denote the fixation probability of this single mutant, and denote by

$$\phi(z) = \left. \frac{\mathrm{d}\Pi(z + \delta, z)}{\mathrm{d}\delta} \right|_{\delta = 0} \tag{A-1}$$

the derivative of the fixation probability with respect to mutant deviation. Then, a measure of the convergence stability of trait value z^* under a strait substition sequence is that

$$\phi(z^*) = 0 \quad \text{and} \quad \left. \frac{\mathrm{d}\phi(z)}{\mathrm{d}z} \right|_{z = z^*} < 0 \tag{A-2}$$

(Rousset and Billiard, 2000; Rousset, 2004; Van Cleve, 2015). Under our modeling assumptions, the perturbation of the fixation probability can be computed as

$$\phi(z) = \lim_{\mu \to 0} \left(\frac{1 - \overline{Q}_0}{1 - Q_{0,0}} \right) \times \left(\frac{\partial w(z_\bullet, \mathbf{z}_{\mathbf{0},0}, \mathbf{n}_{\mathbf{0},0}(\mathbf{z}_{\mathrm{H}}))}{\partial z_\bullet} + \sum_{t=0}^{\infty} \sum_{\mathbf{k} \in \mathcal{G}} \frac{\partial w(z_\bullet, \mathbf{z}_{\mathbf{0},0}, \mathbf{n}_{\mathbf{0},0}(\mathbf{z}_{\mathrm{H}}))}{\partial z_{\mathbf{k},t}} R_{\mathbf{k},t} \right), \tag{A-3}$$

(Lehmann, 2010, eq. 1 together with Lehmann, 2007, eq. A11), which can be expressed as

$$\phi(z) = \underbrace{\lim_{\mu \to 0} \left(\frac{1 - \overline{Q}_0}{1 - Q_{0,0}} \right)}_{> 0} s(z), \tag{A-4}$$

where $s(z)$ is given by eqs. (5)–(6b). Because the limit is always positive as long as $N > 1$ (Rousset, 2004), the condition for convergence stability (A-2) is equivalently given by eq. (4).

The condition for convergence stability (A-2) connects to the stationary probability density function $p(z)$ that (mean) trait value z is observed in the population under our trait substitution sequence process assumptions with rare, random mutations of weak phenotypic effects (recall section 2.3). This probability density function is given by

$$p(z) = K \exp \left[2 \int_l^z N D \phi(y) \, \mathrm{d}y \right], \tag{A-5}$$

(Lehmann, 2012; Van Cleve, 2015) and has a local maximum at z^* if condition (A-2) is satisfied (e.g., (Van Cleve, 2015) for details).

The density function (A-5) can be useful to evaluate the expected phenotypic variance in the population, but requires to have the full representation of $\phi(z)$, which requires to evaluate $\lim_{\mu \to 0} \left(\frac{1 - \overline{Q}_0}{1 - Q_{0,0}} \right)$,

which is process specific. For instance, for the Wright-Fisher process, we have

$$\lim_{\mu \to 0} \left(\frac{1 - \overline{Q}_0}{1 - Q_{0,0}} \right) = \left(\frac{ND + G}{N} \right) \tag{A-6}$$

(Lehmann, 2010, eq. A10). Then, for the example of section 3.4, eq. (35) gives the term in parenthesis ineq. (A-31), and adding the constant of proportionality given by eq. (A-54) for a Wright-Fisher process yields that the perturbation of the fixation probability is

$$\phi(z) = \left(\frac{ND - 1}{D} \right) \left(\frac{ND + G}{N} \right) \left(BP'(z)\alpha_B \left(\frac{P(z)}{\epsilon} \right)^{\alpha_B - 1} \Omega - C\alpha_C z^{\alpha_C - 1} \right). \tag{A-7}$$

Appendix B   Extended phenotypic effects

Appendix B.1   Extended phenotypic effects as press perturbations

We here show eqs. (16)–(17) by deriving a recurrence equation for the extended phenotypic effect $e_{\mathbf{k},t} = \partial n_{\mathbf{k},t}/\partial z_\bullet$ defined by eq. (11). For this, take the derivative on both side of eq. (10) with respect to $z_\bullet$, which yields

$$\frac{\partial n_{\mathbf{k},t+1}}{\partial z_\bullet} = \delta_{t,0} \frac{\partial F(\mathbf{z}_{\mathbf{k},0}^{R}, \mathbf{n}_{\mathbf{k},0})}{\partial z_\bullet} + \sum_{\mathbf{i} \in \mathcal{G}} \frac{\partial F(\mathbf{z}_{\mathbf{k},t}^{R}, \mathbf{n}_{\mathbf{k},t})}{\partial n_{\mathbf{i},t}} \frac{\partial n_{\mathbf{i},t}}{\partial z_\bullet}, \tag{A-8}$$

where we used the fact that $\partial F(\mathbf{z}_{\mathbf{k},t}, \mathbf{n}_{\mathbf{k},t})/\partial n_{\mathbf{i},t} = \partial F(\mathbf{z}_{\mathbf{k},t}^{R}, \mathbf{n}_{\mathbf{k},t})/\partial n_{\mathbf{i},t} =$, since all derivatives are evaluated at z and $\hat{n}$. This further entails that the derivatives of the transition map F are independent of time, which allows us to write

$$e_{\mathbf{k},t+1} = \delta_{t,0}\psi_{\mathbf{k}} + \sum_{\mathbf{i} \in \mathcal{G}} c_{\mathbf{k}-\mathbf{i}} e_{\mathbf{i},t}, \tag{A-9}$$

with

$$\psi_{\mathbf{k}} = \frac{\partial F(\mathbf{z}_{\mathbf{k},0}^{R}, \mathbf{n}_{\mathbf{k},0})}{\partial z_\bullet} = \begin{cases} \frac{1}{N} \frac{\partial F(\mathbf{z}_{0,0}^{R}, \mathbf{n}_{0,0})}{\partial z_{0,0}^{R}} & \text{for } \mathbf{k} = \mathbf{0} \\ \frac{1}{N} \frac{\partial F(\mathbf{z}_{0,0}^{R}, \mathbf{n}_{0,0})}{\partial z_{\mathbf{k},0}} & \text{otherwise,} \end{cases} \tag{A-10}$$

and

$$c_{\mathbf{k}-\mathbf{i}} = \frac{\partial F(\mathbf{z}_{\mathbf{k},t}^{R}, \mathbf{n}_{\mathbf{k},t})}{\partial n_{\mathbf{i},t}} = \frac{\partial F(\mathbf{z}_{\mathbf{k},0}^{R}, \mathbf{n}_{\mathbf{k},0})}{\partial n_{\mathbf{i},0}} = \frac{\partial F(\mathbf{z}_{0,0}^{R}, \mathbf{n}_{0,0})}{\partial n_{\mathbf{k}-\mathbf{i},t}}, \tag{A-11}$$

where the last equalities in both equation follow from spatial homogeneity and are usefull in concrete competition since only $F(\mathbf{z}_{0,0}^{R}, \mathbf{n}_{0,0})$ needs to be specified.

We now use Fourier analysis (see Box 1) to solve eq. (A-9) by using the Fourier transforms $\mathcal{E}_t(\mathbf{h}) = \sum_{\mathbf{k} \in \mathcal{G}} e_{\mathbf{k},t} \chi_{\mathbf{k}}(\mathbf{h})$, $\mathcal{C}(\mathbf{h}) = \sum_{\mathbf{k} \in \mathcal{G}} c_{\mathbf{k}} \chi_{\mathbf{k}}(\mathbf{h})$ and $\Psi(\mathbf{h}) = \sum_{\mathbf{k} \in \mathcal{G}} \psi_{\mathbf{k}} \chi_{\mathbf{k}}(\mathbf{h})$. Then, from (A-9), and noting that

$\chi_{\mathbf{k}}(\mathbf{h}) = \chi_{\mathbf{k}-\mathbf{i}}(\mathbf{h})\chi_{\mathbf{i}}(\mathbf{h})$, we have

$$\underbrace{\sum_{\mathbf{k}\in\mathcal{G}} e_{i,\mathbf{k},t+1}\chi_{\mathbf{k}}(\mathbf{h})}_{\mathcal{E}_{t+1}(\mathbf{h})} = \delta_{t,0}\underbrace{\sum_{\mathbf{k}\in\mathcal{G}} \psi_{\mathbf{k}}\chi_{\mathbf{k}}(\mathbf{h})}_{\Psi(\mathbf{h})} + \underbrace{\sum_{i\in\mathcal{G}} e_{\mathbf{i},t}\chi_{\mathbf{i}}(\mathbf{h})}_{\mathcal{E}_t(\mathbf{h})} \underbrace{\sum_{\mathbf{k}\in\mathcal{G}} c_{\mathbf{k}-\mathbf{i}}\chi_{\mathbf{k}-\mathbf{i}}(\mathbf{h})}_{\mathcal{C}(\mathbf{h})}, \tag{A-12}$$

where the second underbrace holds because the sum is over all (homogeneous) space and thus is independent of $\mathbf{i}$. Therefore

$$\mathcal{E}_{t+1}(\mathbf{h}) = \delta_{t,0}\Psi(\mathbf{h}) + \mathcal{C}(\mathbf{h})\mathcal{E}_t(\mathbf{h}), \tag{A-13}$$

whose solution given the initial condition $\mathcal{E}_0(\mathbf{h}) = 0$ (there are no extended phenotypic effects in the focal generation) is eq. (17).

Appendix B.2   Actor-centered representation of inter-temporal effects

We here derive eq. (18). For this apply the chain rule to eq. (1) whereby we have for $t \geq 1$ that

$$\frac{\partial w(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0}(\mathbf{z}_{\mathrm{H}}))}{\partial z_{\mathbf{k},t}} = \sum_{\mathbf{j}\in\mathcal{G}} \left(N\frac{\partial w(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial n_{\mathbf{j},0}}\right) e_{\mathbf{j}-\mathbf{k},t}, \tag{A-14}$$

where

$$e_{\mathbf{j}-\mathbf{k},t} = \frac{1}{N}\frac{\partial n_{\mathbf{j},0}(\mathbf{z}_{\mathrm{H}})}{\partial z_{\mathbf{k},t}} = \frac{1}{N}\frac{\partial n_{\mathbf{j}-\mathbf{k},t}(\mathbf{z}_{\mathrm{H}})}{\partial z_{0,t}}. \tag{A-15}$$

The first equality defines $e_{\mathbf{j}-\mathbf{k},t}$ and the second follows from spatial homogenity. Here, $e_{\mathbf{j}-\mathbf{k},t}$ is the extended phenotypic effect of a *single* individual residing in the focal patch at t time steps in the past on the value that the environmental variable takes in patch $\mathbf{j}-\mathbf{k}$ in the present. Because eq. (2) entails time homogeneity since the map F does not depend on time, $e_{\mathbf{j}-\mathbf{k},t}$ is also the effect of a focal individual residing in the focal patch on the value that the environmental variable takes in patch $\mathbf{j}-\mathbf{k}$ at t time steps in the future. Hence, we have

$$e_{\mathbf{j}-\mathbf{k},t} = \frac{\partial n_{\mathbf{j}-\mathbf{k},t}}{\partial z_\bullet}, \tag{A-16}$$

where $n_{\mathbf{k},t}$ now stands for the value of the environmental variable in patch $\mathbf{k}$ at t steps in the future. Then, substituting eq. (A-16) into eq. (A-14), which then substituted into eq. (6b), we get eq. (18).

Appendix C Selection gradient in terms of scaled relatedness

We here derive eq. (22), which is premised on the fitness of the focal individual taking the form $w(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0}) = \tilde{w}(\boldsymbol{\pi}(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0}))$ with payoff vector

$$\boldsymbol{\pi}(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0}) = \{\underbrace{\pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}_{\pi_\bullet}, \underbrace{\pi(z_{0,0}, \mathbf{z}_{0,0}^{\mathrm{n}}, \mathbf{n}_{0,0})}_{\pi_0}, \dots, \underbrace{\pi(z_{\mathbf{j},0}, \mathbf{z}_{\mathbf{j},0}, \mathbf{n}_{\mathbf{j},0})}_{\pi_{\mathbf{j}}}\}, \tag{A-17}$$

where in the underbraces we labeled each payoff according to the type of actor bearing that payoff, each of which thus depends on the whole distribution of average phenotypes of actors on that payoff and environmental state variable in the focal generation. In the payoff π_0 to an average patch neighbour to the focal, $\mathbf{z}_{0,0}^{\mathrm{n}}$ is equivalent to $\mathbf{z}_{0,0}$ except for the first entry given by $z_{0,0}^{\mathrm{n}}$.

In order to simplify the operation of taking derivatives, we first express the derivatives of the payoff $\pi_{\mathbf{j}}$ appearing in eq. (A-17) with respect to its various arguments in terms of the derivatives of the payoff of the focal individual. Recalling that $z_{0,0}^{\mathrm{n}} = \frac{1}{N-1} z_\bullet + \left(\frac{N-2}{N-1}\right) z_{0,0}$, we then have for phenotypic derivatives that

$$\frac{\partial \pi_\bullet}{\partial z_\bullet} = \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial z_\bullet} \quad \text{and} \quad \frac{\partial \pi_0}{\partial z_\bullet} = \frac{1}{N-1} \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial z_{0,0}}, \tag{A-18}$$

$$\frac{\partial \pi_{\mathbf{j}}}{\partial z_{\mathbf{j},0}} = \begin{cases} \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial z_\bullet} + \left(\frac{N-2}{N-1}\right) \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial z_{0,0}} & \text{for } \mathbf{j} = \mathbf{0} \\ \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial z_\bullet} + \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial z_{0,0}} & \text{for } \mathbf{j} \neq \mathbf{0}, \end{cases} \tag{A-19}$$

and for $\mathbf{k} \neq \mathbf{j}$

$$\frac{\partial \pi_{\mathbf{j}}}{\partial z_{\mathbf{k},0}} = \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial z_{\mathbf{k},0}}. \tag{A-20}$$

For derivatives of payoffs with respect to state variables, we have

$$\frac{\partial \pi_\bullet}{\partial n_{0,0}} = \frac{\partial \pi_{\mathbf{j}}}{\partial n_{\mathbf{j},0}} = \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial n_{0,0}} \qquad \text{for all } \mathbf{j} \in \mathcal{G}, \tag{A-21}$$

where the second equality is a consequence of spatial homogeneity, and for $\mathbf{j} \neq \mathbf{0}$

$$\frac{\partial \pi_{\mathbf{k}}}{\partial n_{\mathbf{k}-\mathbf{j},0}} = \frac{\partial \pi_\bullet}{\partial n_{\mathbf{j},0}} = \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial n_{\mathbf{j},0}}, \tag{A-22}$$

where the second equality is again a consequence of spatial homogeneity. In terms of these notations, and using the chain rule, we can write the derivative of fitness $w(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0}) = \tilde{w}(\boldsymbol{\pi}(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0}))$ with respect to its different arguments that appear in the selection gradient (eqs. (6a)–(6b)) in terms of the focal's payoff derivatives as follows. First, the derivatives with repect to the focal's phenotype is

$$\frac{\partial w(z_\bullet, \mathbf{z}_0, \mathbf{n}_{0,0})}{\partial z_\bullet} = \frac{\partial \tilde{w}}{\partial \pi_\bullet} \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial z_\bullet} + \frac{\partial \tilde{w}}{\partial \pi_0} \frac{1}{N-1} \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial z_{0,0}}, \tag{A-23}$$

that with respect to average patch neighbor phenotype is

$$\frac{\partial w(z_\bullet, \mathbf{z_0}, \mathbf{n_{0,0}})}{\partial z_{\mathbf{0},0}} = \frac{\partial \tilde{w}}{\partial \pi_\bullet} \frac{\partial \pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial z_{\mathbf{0},0}} + \frac{\partial \tilde{w}}{\partial \pi_0} \left(\frac{\partial \pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial z_\bullet} - \left(\frac{1}{N-1} \right) \frac{\partial \pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial z_{\mathbf{0},0}} \right)$$
$$+ \sum_{\mathbf{k} \in \mathcal{G}} \frac{\partial \tilde{w}}{\partial \pi_\mathbf{k}} \frac{\partial \pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial z_{\mathbf{k},0}}, \quad \text{(A-24)}$$

and that with respect to the average phenotype in any patch $\mathbf{j} \neq \mathbf{0}$ is

$$\frac{\partial w(z_\bullet, \mathbf{z_0}, \mathbf{n_{0,0}})}{\partial z_{\mathbf{j},0}} = \frac{\partial \tilde{w}}{\partial \pi_\bullet} \frac{\partial \pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial z_{\mathbf{j},0}} + \frac{\partial \tilde{w}}{\partial \pi_\mathbf{j}} \frac{\partial \pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial z_\bullet} + \sum_{\mathbf{k} \in \mathcal{G}} \frac{\partial \tilde{w}}{\partial \pi_\mathbf{k}} \frac{\partial \pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial z_{\mathbf{j-k},0}}. \quad \text{(A-25)}$$

Second, the derivative with respect to the state variable in patch $\mathbf{j}$ is

$$\frac{\partial w(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial n_{\mathbf{j},0}} = \frac{\partial \tilde{w}}{\partial \pi_\bullet} \frac{\partial \pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial n_{\mathbf{j},0}} + \sum_{\mathbf{k} \in \mathcal{G}} \frac{\partial \tilde{w}}{\partial \pi_\mathbf{k}} \frac{\partial \pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial n_{\mathbf{j-k},0}}. \quad \text{(A-26)}$$

Now denote by

$$\lambda_\mathbf{j} = -\frac{\partial \tilde{w}}{\partial \pi_\mathbf{j}} \bigg/ \frac{\partial \tilde{w}}{\partial \pi_\bullet} \quad \text{(A-27)}$$

the coefficient of fitness interdependence between individuals in the focal patch and that in patch $\mathbf{j}$. Substituting eqs. (A-23)–(A-25) along eq. (A-27) into eq. (6a) yields

$$s_\text{w}(z) = \frac{\partial \tilde{w}}{\partial \pi_\bullet} \left[\frac{\partial \pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial z_\bullet} \left(1 - \sum_{\mathbf{j} \in \mathcal{G}} \lambda_\mathbf{j} R_{\mathbf{j},0} \right) - \frac{1}{N-1} \frac{\partial \pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial z_\mathbf{0}} \lambda_0 (1 - R_{\mathbf{0},0}) \right.$$
$$\left. + \sum_{\mathbf{k} \in \mathcal{G}} \frac{\partial \pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial z_{\mathbf{k},0}} \left(R_{\mathbf{k},t} - \sum_{\mathbf{j} \in \mathcal{G}} \lambda_\mathbf{j} R_{\mathbf{j-k},t} \right) \right], \quad \text{(A-28)}$$

where for rearrangement we used the identity

$$\sum_{\mathbf{j} \in \mathcal{G}} \sum_{\mathbf{k} \in \mathcal{G}} \lambda_\mathbf{k} \frac{\partial \pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial z_{\mathbf{j-k},0}} R_{\mathbf{j},0} = \sum_{\mathbf{j} \in \mathcal{G}} \sum_{\mathbf{k} \in \mathcal{G}} \lambda_\mathbf{k} \frac{\partial \pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial z_{\mathbf{j},0}} R_{\mathbf{j-k},\mathbf{0}}. \quad \text{(A-29)}$$

Now substituting eq. (A-26) along eq. (A-27) into eq. (18) produces

$$s_\text{e}(z) = N \frac{\partial \tilde{w}}{\partial \pi_\bullet} \sum_{t=1}^\infty \sum_{\mathbf{k} \in \mathcal{G}} \sum_{\mathbf{j} \in \mathcal{G}} N \left[\frac{\partial \pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial n_{\mathbf{j},0}} - \sum_{\mathbf{i} \in \mathcal{G}} \lambda_\mathbf{i} \frac{\partial \pi(z_\bullet, \mathbf{z_{0,0}}, \mathbf{n_{0,0}})}{\partial n_{\mathbf{j-i},0}} \right] e_{\mathbf{j-k},t} R_{\mathbf{k},t}, \quad \text{(A-30)}$$

Owing to spatial homogeneity, we can further write

$$
\begin{aligned}
s_{\mathrm{e}}(z) &= N\frac{\partial \tilde{w}}{\partial \pi_\bullet} \sum_{t=1}^{\infty} \left[\sum_{\mathbf{j}\in\mathcal{G}} \sum_{\mathbf{k}\in\mathcal{G}} \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial n_{\mathbf{j},0}} e_{\mathbf{j}-\mathbf{k},t} R_{\mathbf{k},t} - \sum_{\mathbf{j}\in\mathcal{G}} \sum_{\mathbf{i}\in\mathcal{G}} \sum_{\mathbf{k}\in\mathcal{G}} \lambda_{\mathbf{i}} \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial n_{\mathbf{j}-\mathbf{i},0}} e_{\mathbf{j}-\mathbf{k},t} R_{\mathbf{k},t} \right] \\
&= N\frac{\partial \tilde{w}}{\partial \pi_\bullet} \sum_{t=1}^{\infty} \left[\sum_{\mathbf{j}\in\mathcal{G}} \sum_{\mathbf{k}\in\mathcal{G}} \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial n_{\mathbf{j},0}} e_{\mathbf{k},t} R_{\mathbf{j}-\mathbf{k},t} - \sum_{\mathbf{j}\in\mathcal{G}} \sum_{\mathbf{i}\in\mathcal{G}} \sum_{\mathbf{k}\in\mathcal{G}} \lambda_{\mathbf{i}} \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial n_{\mathbf{j},0}} e_{\mathbf{j}-\mathbf{i}-\mathbf{k},t} R_{\mathbf{k},t} \right] \\
&= N\frac{\partial \tilde{w}}{\partial \pi_\bullet} \sum_{t=1}^{\infty} \left[\sum_{\mathbf{j}\in\mathcal{G}} \sum_{\mathbf{k}\in\mathcal{G}} \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial n_{\mathbf{j},0}} e_{\mathbf{k},t} R_{\mathbf{j}-\mathbf{k},t} - \sum_{\mathbf{j}\in\mathcal{G}} \sum_{\mathbf{i}\in\mathcal{G}} \sum_{\mathbf{k}\in\mathcal{G}} \lambda_{\mathbf{i}} \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial n_{\mathbf{j},0}} e_{\mathbf{k},t} R_{\mathbf{j}-\mathbf{i}-\mathbf{k},t} \right] \\
&= N\frac{\partial \tilde{w}}{\partial \pi_\bullet} \sum_{t=1}^{\infty} \sum_{\mathbf{j}\in\mathcal{G}} \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial n_{\mathbf{j},0}} \sum_{\mathbf{k}\in\mathcal{G}} e_{\mathbf{k},t} \left[R_{\mathbf{j}-\mathbf{k},t} - \sum_{\mathbf{i}\in\mathcal{G}} \lambda_{\mathbf{i}} R_{\mathbf{j}-\mathbf{k}-\mathbf{i},t} \right],
\end{aligned}
$$

where in the second equality we interchanged the dummy variables of the $n_{\mathbf{j}}$'s and the $e_{\mathbf{j}}$'s in the second summand, and in the third equality we interchanged the dummy variables of the $e_{\mathbf{j}}$'s and the $R_{\mathbf{j}}$'s. Now adding up intra-temporal and inter-temporal effects, we obtain

$$
\begin{aligned}
s(z) &= L \left(\frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial z_\bullet} + \sum_{\mathbf{k}\in\mathcal{G}} \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial z_{\mathbf{k},0}} \kappa_{\mathbf{k},0} + N \sum_{t=1}^{\infty} \sum_{\mathbf{j}\in\mathcal{G}} \sum_{\mathbf{k}\in\mathcal{G}} \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial n_{\mathbf{j},0}} e_{\mathbf{k},t} \kappa_{\mathbf{k}-\mathbf{j},t} \right) \\
&\qquad\text{(A-31)} \\
&= L \left(\frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial z_\bullet} + \sum_{\mathbf{k}\in\mathcal{G}} \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial z_{\mathbf{k},0}} \kappa_{\mathbf{k},0} + N \sum_{t=1}^{\infty} \sum_{\mathbf{k}\in\mathcal{G}} \sum_{\mathbf{j}\in\mathcal{G}} \kappa_{\mathbf{k},t} \frac{\partial \pi(z_\bullet, \mathbf{z}_{0,0}, \mathbf{n}_{0,0})}{\partial n_{\mathbf{j},0}} e_{\mathbf{k}-\mathbf{j},t} \right)
\end{aligned}
$$

where the second line follows exchanging dummy variable of the $\kappa_{\mathbf{j}}$'s and the $e_{\mathbf{j}}$'s,

$$
L = \frac{\partial \tilde{w}}{\partial \pi_\bullet} \left(1 - \sum_{\mathbf{j}\in\mathcal{G}} \lambda_{\mathbf{j}} R_{\mathbf{j},0} \right), \tag{A-32}
$$

for $t = 0$ we have

$$
\kappa_{\mathbf{k},0} = \begin{cases} \dfrac{R_{0,0} - \frac{1}{N-1}\lambda_{0}(1-R_{0,0}) - \sum_{\mathbf{j}\in\mathcal{G}} \lambda_{\mathbf{j}} R_{\mathbf{j},0}}{1 - \sum_{\mathbf{j}\in\mathcal{G}} \lambda_{\mathbf{j}} R_{\mathbf{j},0}} & \text{for } \mathbf{k} = 0 \\[2ex] \dfrac{R_{\mathbf{k},0} - \sum_{\mathbf{j}\in\mathcal{G}} \lambda_{\mathbf{j}} R_{\mathbf{j}-\mathbf{k},0}}{1 - \sum_{\mathbf{j}\in\mathcal{G}} \lambda_{\mathbf{j}} R_{\mathbf{j},0}} & \text{otherwise,} \end{cases} \tag{A-33}
$$

while for $t > 0$ we have

$$
\kappa_{\mathbf{k},t} = \frac{R_{\mathbf{k},t} - \sum_{\mathbf{i}\in\mathcal{G}} \lambda_{\mathbf{i}} R_{\mathbf{i}-\mathbf{k},t}}{1 - \sum_{\mathbf{i}\in\mathcal{G}} \lambda_{\mathbf{i}} R_{\mathbf{i},0}}. \tag{A-34}
$$

For the infinite island model of dispersal $R_{\mathbf{j},0} = 0$ for all $\mathbf{j} \neq \mathbf{0}$ in which case $\kappa_{0,0}$ reduces to eq. 22 of Alger et al. (2020) as it should be, and this provides a consistency check of the derivation.

Appendix D   Explicit expressions for scaled relatedness

We here derive eq. (55) and a representation for L in eq. (A-31), which is premised on individual fitness being given by eq. (21); that is

$$\tilde{w}(\boldsymbol{\pi}) = s\left(\pi_\bullet\right) + \sum_{\mathbf{i}\in\mathcal{G}} m_{\mathbf{i}} \left[1 - s^{\mathrm{R}}\left(\pi_{\mathbf{i}}\right)\right] \frac{f(\pi_\bullet)}{\sum_{\mathbf{j}\in\mathcal{G}} m_{\mathbf{i}-\mathbf{j}} f^{\mathrm{R}}(\pi_{\mathbf{j}})}, \tag{A-35}$$

where $s^{\mathrm{R}}(\pi_{\mathbf{0}}) = \frac{1}{N} s(\pi_\bullet) + \frac{N-1}{N} s(\pi_{\mathbf{0}})$, otherwise $s^{\mathrm{R}}(\pi_{\mathbf{i}}) = s(\pi_{\mathbf{i}})$ for $\mathbf{i} \neq 0$; and $f^{\mathrm{R}}(\pi_{\mathbf{i}}) = \frac{1}{N} f(\pi_\bullet) + \frac{N-1}{N} f(\pi_{\mathbf{0}})$, otherwise $f^{\mathrm{R}}(\pi_{\mathbf{i}}) = f(\pi_{\mathbf{i}})$ for $\mathbf{i} \neq 0$. Direct calculations, then produce

$$\frac{\partial \tilde{w}(\boldsymbol{\pi})}{\partial \pi_{\mathbf{j}}} = \begin{cases} s'\left(1 - \frac{1}{N} m_{\mathbf{0}}\right) + \frac{f'}{f}(1-s)\left(1 - \frac{1}{N}\sum_{\mathbf{i}\in\mathcal{G}} m_{\mathbf{i}}^2\right) & \text{for } \mathbf{j} = \bullet \\ -s' \frac{N-1}{N} m_{\mathbf{0}} - \frac{f'}{f}(1-s)\frac{N-1}{N}\sum_{\mathbf{i}\in\mathcal{G}} m_{\mathbf{i}}^2, & \text{for } \mathbf{j} = 0 \\ -s' m_{\mathbf{j}} - \frac{f'}{f}(1-s)\sum_{\mathbf{i}\in\mathcal{G}} m_{\mathbf{i}} m_{\mathbf{i}-\mathbf{j}} & \text{otherwise,} \end{cases} \tag{A-36}$$

where $f' = \partial f\left(\pi_\bullet\right)/\partial \pi_\bullet$ and $s' = \partial s\left(\pi_\bullet\right)/\partial \pi_\bullet$ and all functions are, as usual, evaluated at the resident trait value z and equilibrium $\hat{n}$. Thereby, the coefficient of fitness interdependence (A-37) can be written

$$\lambda_{\mathbf{j}} = -\frac{\partial \tilde{w}(\boldsymbol{\pi})/\partial \pi_{\mathbf{j}}}{\partial \tilde{w}(\boldsymbol{\pi})/\partial \pi_\bullet} = \begin{cases} \frac{s' f \frac{N-1}{N} m_{\mathbf{0}} + f'(1-s)\frac{N-1}{N} P_{\mathbf{0}}}{s' f\left(1 - \frac{1}{N} m_{\mathbf{0}}\right) + f'(1-s)(1-(1/N)P_{\mathbf{0}})} & \text{for } \mathbf{j} = 0 \\ \frac{s' f m_{\mathbf{j}} + f'(1-s)P_{\mathbf{j}}}{s' f\left(1 - \frac{1}{N} m_{\mathbf{0}}\right) + f'(1-s)(1-(1/N)P_{\mathbf{0}})} & \text{otherwise,} \end{cases} \tag{A-37}$$

where $P_{\mathbf{j}} = \sum_{\mathbf{i}\in\mathcal{G}} m_{\mathbf{i}} m_{\mathbf{i}-\mathbf{j}}$.

Now substituting eq. (A-37) into the first line of eq. (A-33), we get

$$\kappa_{0,0} = \frac{R_{0,0} - \frac{1}{N-1}\lambda_{\mathbf{0}}(1 - R_{0,0}) - \sum_{\mathbf{j}\in\mathcal{G}} \lambda_{\mathbf{j}} R_{\mathbf{j},0}}{1 - \sum_{\mathbf{j}\in\mathcal{G}} \lambda_{\mathbf{j}} R_{\mathbf{j},0}} \tag{A-38}$$

$$= \frac{R_{0,0} - \frac{1}{N-1}(1 - R_{0,0})\lambda_{\mathbf{0}} - R_{0,0}\lambda_{\mathbf{0}} - \sum_{\mathbf{j}\in\mathcal{G}\setminus\mathbf{0}} \lambda_{\mathbf{j}} R_{\mathbf{j},0}}{1 - \lambda_{\mathbf{0}} R_{0,0} - \sum_{\mathbf{j}\in\mathcal{G}\setminus\mathbf{0}} \lambda_{\mathbf{j}} R_{\mathbf{j},0}}$$

$$= \frac{R_{0,0} - \left[\frac{1}{N-1}(1 - R_{0,0}) + R_{0,0}\right] \frac{s' f m_{\mathbf{0}} + f'(1-s)\frac{N-1}{N} P_{\mathbf{0}}}{s' f\left(1 - \frac{1}{N} m_{\mathbf{0}}\right) + f'(1-s)(1-(1/N)P_{\mathbf{0}})} - \sum_{\mathbf{j}\in\mathcal{G}\setminus\mathbf{0}} \frac{s' f m_{\mathbf{j}} + f'(1-s)P_{\mathbf{j}}}{s' f\left(1 - \frac{1}{N} m_{\mathbf{0}}\right) + f'(1-s)(1-(1/N)P_{\mathbf{0}})} R_{\mathbf{j},0}}{1 - \frac{s' f m_{\mathbf{j}} + f'(1-s)\frac{N-1}{N} P_{\mathbf{0}}}{s' f\left(1 - \frac{1}{N} m_{\mathbf{0}}\right) + f'(1-s)(1-(1/N)P_{\mathbf{0}})} R_{0,0} - \sum_{\mathbf{j}\in\mathcal{G}\setminus\mathbf{0}} \frac{s' f m_{\mathbf{j}} + f'(1-s)P_{\mathbf{j}}}{s' f\left(1 - \frac{1}{N} m_{\mathbf{0}}\right) + f'(1-s)(1-(1/N)P_{\mathbf{0}})} R_{\mathbf{j},0}}$$

$$= \frac{R_{0,0}\left(s' f + f'(1-s)\right) - s' f\left(\sum_{\mathbf{j}\in\mathcal{G}} m_{\mathbf{j}} R_{\mathbf{j},0} + m_{\mathbf{0}}\frac{(1-R_{0,0})}{N}\right) - f'(1-s)\left(\sum_{\mathbf{j}\in\mathcal{G}} P_{\mathbf{j}} R_{\mathbf{j},0} + P_{\mathbf{0}}\frac{(1-R_{0,0})}{N}\right)}{s' f + f'(1-s) - s' f\left(\sum_{\mathbf{j}\in\mathcal{G}} m_{\mathbf{j}} R_{\mathbf{j},0} + m_{\mathbf{0}}\frac{(1-R_{0,0})}{N}\right) - f'(1-s)\left(\sum_{\mathbf{j}\in\mathcal{G}} P_{\mathbf{j}} R_{\mathbf{j},0} + P_{\mathbf{0}}\frac{(1-R_{0,0})}{N}\right)}$$

$$= \frac{s' f\left[R_{0,0} - \left(\sum_{\mathbf{j}\in\mathcal{G}} m_{\mathbf{j}} R_{\mathbf{j},0} + m_{\mathbf{0}}\frac{(1-R_{0,0})}{N}\right)\right] + f'(1-s)\left[R_{0,0} - \left(\sum_{\mathbf{j}\in\mathcal{G}} P_{\mathbf{j}} R_{\mathbf{j},0} + P_{\mathbf{0}}\frac{(1-R_{0,0})}{N}\right)\right]}{s' f\left[1 - \left(\sum_{\mathbf{j}\in\mathcal{G}} m_{\mathbf{j}} R_{\mathbf{j},0} + m_{\mathbf{0}}\frac{(1-R_{0,0})}{N}\right)\right] + f'(1-s)\left[1 - \left(\sum_{\mathbf{j}\in\mathcal{G}} P_{\mathbf{j}} R_{\mathbf{j},0} + P_{\mathbf{0}}\frac{(1-R_{0,0})}{N}\right)\right]},$$

while substituting eq. (A-37) into the second line of of eq. (A-33) and eq. (23), we get for all $(\mathbf{k}, t) \neq$

$(\mathbf{0}, 0)$ that

$$
\begin{aligned}
\kappa_{\mathbf{k},t} &= \frac{R_{\mathbf{k},t} - \sum_{\mathbf{j}\in\mathcal{G}} \lambda_{\mathbf{j}} R_{\mathbf{j}-\mathbf{k},t}}{1 - \sum_{\mathbf{j}\in\mathcal{G}} \lambda_{\mathbf{j}} R_{\mathbf{j},0}} \\[2mm]
&= \frac{R_{\mathbf{k},t} - \lambda_{\mathbf{0}} R_{\mathbf{k},t} - \sum_{\mathbf{j}\in\mathcal{G}\backslash\mathbf{0}} \lambda_{\mathbf{j}} R_{\mathbf{j}-\mathbf{k},t}}{1 - \lambda_{\mathbf{0}} R_{\mathbf{0},0} - \sum_{\mathbf{j}\in\mathcal{G}\backslash\mathbf{0}} \lambda_{\mathbf{j}} R_{\mathbf{j},0}} \\[2mm]
&= \frac{R_{\mathbf{k},t} - \frac{s'fm_{\mathbf{0}}+f'(1-s)\frac{N-1}{N}P_{\mathbf{0}}}{s'f\left(1-\frac{1}{N}m_{\mathbf{0}}\right)+f'(1-s)(1-(1/N)P_{\mathbf{0}})} R_{\mathbf{k},t} - \sum_{\mathbf{j}\in\mathcal{G}\backslash\mathbf{0}} \frac{s'fm_{\mathbf{j}}+f'(1-s)P_{\mathbf{j}}}{s'f\left(1-\frac{1}{N}m_{\mathbf{0}}\right)+f'(1-s)(1-(1/N)P_{\mathbf{0}})} R_{\mathbf{j}-\mathbf{k},t}}{1 - \frac{s'fm_{\mathbf{j}}+f'(1-s)\frac{N-1}{N}P_{\mathbf{0}}}{s'f\left(1-\frac{1}{N}m_{\mathbf{0}}\right)+f'(1-s)(1-(1/N)P_{\mathbf{0}})} R_{\mathbf{0},0} - \sum_{\mathbf{j}\in\mathcal{G}\backslash\mathbf{0}} \frac{s'fm_{\mathbf{j}}+f'(1-s)P_{\mathbf{j}}}{s'f\left(1-\frac{1}{N}m_{\mathbf{0}}\right)+f'(1-s)(1-(1/N)P_{\mathbf{0}})} R_{\mathbf{j},0}} \\[2mm]
&= \frac{R_{\mathbf{k},t}\left(s'f + f'(1-s)\right) - s'f\sum_{\mathbf{j}\in\mathcal{G}} m_{\mathbf{j}} R_{\mathbf{j}-\mathbf{k},t} - f'(1-s)\sum_{\mathbf{j}\in\mathcal{G}} P_{\mathbf{j}} R_{\mathbf{j}-\mathbf{k},t}}{s'f + f'(1-s) - s'f\left(\sum_{\mathbf{j}\in\mathcal{G}} m_{\mathbf{j}} R_{\mathbf{j},0} + m_{\mathbf{0}}\frac{(1-R_{\mathbf{0},0})}{N}\right) - f'(1-s)\left(\sum_{\mathbf{j}\in\mathcal{G}} P_{\mathbf{j}} R_{\mathbf{j},0} + P_{\mathbf{0}}\frac{(1-R_{\mathbf{0},0})}{N}\right)} \\[2mm]
&= \frac{s'f\left[R_{\mathbf{k},t} - \sum_{\mathbf{j}\in\mathcal{G}} m_{\mathbf{j}} R_{\mathbf{j}-\mathbf{k},t}\right] + f'(1-s)\left[R_{\mathbf{k},t} - \sum_{\mathbf{j}\in\mathcal{G}} P_{\mathbf{j}} R_{\mathbf{j}-\mathbf{k},t}\right]}{s'f\left[1 - \left(\sum_{\mathbf{j}\in\mathcal{G}} m_{\mathbf{j}} R_{\mathbf{j},0} + m_{\mathbf{0}}\frac{(1-R_{\mathbf{0},0})}{N}\right)\right] + f'(1-s)\left[1 - \left(\sum_{\mathbf{j}\in\mathcal{G}} P_{\mathbf{j}} R_{\mathbf{j},0} + P_{\mathbf{0}}\frac{(1-R_{\mathbf{0},0})}{N}\right)\right]}.
\end{aligned}
$$
(A-39)

Now on substituting eq. (7) into the last two equations and canceling common terms, we obtain

$$
\kappa_{\mathbf{0},0} = \frac{s'f\lim_{\mu\to0}\left[Q_{\mathbf{0},0} - \left(\sum_{\mathbf{j}\in\mathcal{G}} m_{\mathbf{j}} Q_{\mathbf{j},0} + m_{\mathbf{0}}\frac{(1-Q_{\mathbf{0},0})}{N}\right)\right] + f'(1-s)\lim_{\mu\to0}\left[Q_{\mathbf{0},0} - \left(\sum_{\mathbf{j}\in\mathcal{G}} P_{\mathbf{j}} Q_{\mathbf{j},0} + P_{\mathbf{0}}\frac{(1-Q_{\mathbf{0},0})}{N}\right)\right]}{s'f\lim_{\mu\to0}\left[1 - \left(\sum_{\mathbf{j}\in\mathcal{G}} m_{\mathbf{j}} Q_{\mathbf{j},0} + m_{\mathbf{0}}\frac{(1-Q_{\mathbf{0},0})}{N}\right)\right] + f'(1-s)\lim_{\mu\to0}\left[1 - \left(\sum_{\mathbf{j}\in\mathcal{G}} P_{\mathbf{j}} Q_{\mathbf{j},0} + P_{\mathbf{0}}\frac{(1-Q_{\mathbf{0},0})}{N}\right)\right]}
$$
(A-40)

and for all $(\mathbf{k},t) \neq (\mathbf{0},0)$ that

$$
\kappa_{\mathbf{k},t} = \frac{s'f\lim_{\mu\to0}\left[Q_{\mathbf{k},t} - \sum_{\mathbf{j}\in\mathcal{G}} m_{\mathbf{j}} Q_{\mathbf{j}-\mathbf{k},t}\right] + f'(1-s)\lim_{\mu\to0}\left[Q_{\mathbf{k},t} - \sum_{\mathbf{j}\in\mathcal{G}} P_{\mathbf{j}} Q_{\mathbf{j}-\mathbf{k},t}\right]}{s'f\lim_{\mu\to0}\left[1 - \left(\sum_{\mathbf{j}\in\mathcal{G}} m_{\mathbf{j}} Q_{\mathbf{j},0} + m_{\mathbf{0}}\frac{(1-Q_{\mathbf{0},0})}{N}\right)\right] + f'(1-s)\lim_{\mu\to0}\left[1 - \left(\sum_{\mathbf{j}\in\mathcal{G}} P_{\mathbf{j}} Q_{\mathbf{j},0} + P_{\mathbf{0}}\frac{(1-Q_{\mathbf{0},0})}{N}\right)\right]}.
$$
(A-41)

In force of eqs. (A.42), (A.48), and (A.51)[1] in Lehmann and Rousset (2012), we have

$$
\lim_{\mu\to0} \frac{1}{(1-Q_{\mathbf{0},0})}\left[1 - \left(\sum_{\mathbf{j}\in\mathcal{G}} m_{\mathbf{j}} Q_{\mathbf{j},0} + m_{\mathbf{0}}\frac{(1-Q_{\mathbf{0},0})}{N}\right)\right] = \frac{1}{N}\left[N + \mathcal{L}_{\mathbf{0}}\left(F^{\mathrm{s}}\right) - \frac{1+s}{2D}\right]
$$
(A-42)

$$
\lim_{\mu\to0} \frac{1}{(1-Q_{\mathbf{0},0})}\left[Q_{\mathbf{0},0} - \left(\sum_{\mathbf{j}\in\mathcal{G}} m_{\mathbf{j}} Q_{\mathbf{j},0} + m_{\mathbf{0}}\frac{(1-Q_{\mathbf{0},0})}{N}\right)\right] = \frac{1}{N}\left[1 + \mathcal{L}_{\mathbf{0}}\left(F^{\mathrm{s}}\right) - \frac{1+s}{2D}\right]
$$

$$
\lim_{\mu\to0} \frac{1}{(1-Q_{\mathbf{0},0})}\left[Q_{\mathbf{k},t} - \sum_{\mathbf{j}\in\mathcal{G}} m_{\mathbf{j}} Q_{\mathbf{j}-\mathbf{k},t}\right] = \frac{1}{N}\left[\mathcal{L}_{\mathbf{k}}\left(G_t^{\mathrm{s}}\right) - \frac{1+s}{2D}\right],
$$

where the functions F^{s} and G_t^{s} are defined in $\mathbf{h}$ as

$$
\begin{aligned}
F^{\mathrm{s}}(\mathbf{h}) &= -\frac{(1-s)\mathcal{M}(\mathbf{h})}{1+s+(1-s)\mathcal{M}(\mathbf{h})}, \\[2mm]
G_t^{\mathrm{s}}(\mathbf{h}) &= \frac{(1+s)\left[s+(1-s)\mathcal{M}(\mathbf{h})\right]^t}{1+s+(1-s)\mathcal{M}(\mathbf{h})}.
\end{aligned}
$$
(A-43)

---

[1]This eq. (A.51) in Lehmann and Rousset (2012) applies for for all $\mathbf{k}\in\mathcal{G}$, the condition " if $\mathbf{k}\neq\mathbf{0}$" therein is a typo.

In force of eqs. (A.32), (A.38, and (A.41)[2] in Lehmann and Rousset (2012), we have

$$\lim_{\mu \to 0} \frac{1}{(1 - Q_{0,0})} \left[1 - \left(\sum_{j \in \mathcal{G}} P_j Q_{j,0} + P_0 \frac{(1 - Q_{0,0})}{N} \right) \right] = \frac{1}{N} \left[N + \mathcal{L}_0 \left(F^f \right) - \frac{1 + s}{D} \right] \quad \text{(A-44)}$$

$$\lim_{\mu \to 0} \frac{1}{(1 - Q_{0,0})} \left[Q_{0,0} - \left(\sum_{j \in \mathcal{G}} P_j Q_{j,0} + P_0 \frac{(1 - Q_{0,0})}{N} \right) \right] = \frac{1}{N} \left[\mathcal{L}_0 \left(F^f \right) - \frac{1 + s}{D} \right]$$

$$\lim_{\mu \to 0} \frac{1}{(1 - Q_{0,0})} \left[Q_{k,t} - \sum_{j \in \mathcal{G}} P_j Q_{j-k,t} \right] = \frac{1}{N} \left[\mathcal{L}_k \left(G^f \right) - \frac{1 + s}{D} \right],$$

where the functions F^f and G_t^f are defined in $\mathbf{h}$ as

$$F^f(\mathbf{h}) = \frac{2s \mathcal{M}(\mathbf{h})}{1 + s + (1 - s) \mathcal{M}(\mathbf{h})},$$

$$G_t^f(\mathbf{h}) = \frac{(1 + s)(1 + \mathcal{M}(\mathbf{h})) \left[s + (1 - s) \mathcal{M}(\mathbf{h}) \right]^t}{1 + s + (1 - s) \mathcal{M}(\mathbf{h})}. \quad \text{(A-45)}$$

On substituting eq. (A-42) and eq. (A-44) into eqs. (A-41)–(A-40) produces for $t = 0$,

$$\kappa_{k,0} = \frac{s' f \left[1 + \mathcal{L}_k (F^s) - \frac{1+s}{2D} \right] + f'(1 - s) \left[\mathcal{L}_k \left(F^f \right) - \frac{1+s}{D} \right]}{s' f \left[N + \mathcal{L}_0 (F^s) - \frac{1+s}{2D} \right] + f'(1 - s) \left[N + \mathcal{L}_0 \left(F^f \right) - \frac{1+s}{D} \right]} \quad \text{(A-46)}$$

and for $t > 0$,

$$\kappa_{k,t} = \frac{s' f \left[\mathcal{L}_k (G_t^s) - \frac{1+s}{2D} \right] + f'(1 - s) \left[\mathcal{L}_k \left(G^f \right) - \frac{1+s}{D} \right]}{s' f \left[N + \mathcal{L}_0 (F^s) - \frac{1+s}{2D} \right] + f'(1 - s) \left[N + \mathcal{L}_0 \left(F^f \right) - \frac{1+s}{D} \right]}. \quad \text{(A-47)}$$

Finally substituting eqs. (A-43)–(A-45) and rearranging yields eq. (55) of Box 2:

$$\kappa_{k,t} = \begin{cases} \frac{\mathcal{L}_k(F) - (1+s)[s'f + 2f'(1-s)]/(2D)}{N[s'f + f'(1-s)] + \mathcal{L}_0(F) - (1+s)[s'f + 2f'(1-s)]/(2D)} & \text{if } t = 0 \\ \frac{\mathcal{L}_k(G_t) - (1+s)[s'f + 2f'(1-s)]/(2D)}{N[s'f + f'(1-s)] + \mathcal{L}_0(F) - (1+s)[s'f + 2f'(1-s)]/(2D)} & \text{otherwise,} \end{cases} \quad \text{(A-48)}$$

where the functions F and G_t are defined in $\mathbf{h}$ as

$$F(\mathbf{h}) = -\frac{(1 - s) \left[s' f - f' 2s \right] \mathcal{M}(\mathbf{h})}{1 + s + (1 - s) \mathcal{M}(\mathbf{h})},$$

$$G_t(\mathbf{h}) = \frac{(1 + s) \left[s' f + f'(1 - s)(1 + \mathcal{M}(\mathbf{h})) \right] \left[s + (1 - s) \mathcal{M}(\mathbf{h}) \right]^t}{1 + s + (1 - s) \mathcal{M}(\mathbf{h})}. \quad \text{(A-49)}$$

Finally, we evaluate $L = \partial \tilde{w} / \partial \pi_\bullet \left(1 - \sum_{j \in \mathcal{G}} \lambda_j R_{j,0} \right)$ in eq. (A-31), which is needed if one aims to

---

[2]This eq. (A.41) in Lehmann and Rousset (2012) applies for for all $\mathbf{k} \in \mathcal{G}$, the condition " if $\mathbf{k} \neq \mathbf{0}$" therein is a typo.

evaluated the trait stationary density function (A-5). Using eqs. (A-36)–(A-37) gives

$$L = \frac{1}{f}\left(s'f\left(1 - \frac{1}{N}m_0\right) + f'(1-s)\left(1 - \frac{1}{N}P_0\right)\right)\left(1 - \frac{s'fm_j + f'(1-s)\frac{N-1}{N}P_0}{s'f\left(1 - \frac{1}{N}m_0\right) + f'(1-s)\left(1 - (1/N)P_0\right)}R_{0,0}\right.$$
$$\left. - \sum_{j \in \mathcal{G}\backslash 0} \frac{s'fm_j + f'(1-s)P_j}{s'f\left(1 - \frac{1}{N}m_0\right) + f'(1-s)\left(1 - (1/N)P_0\right)}R_{j,0}\right), \quad (A\text{-}50)$$

and on simplification yields

$$L = \frac{1}{f}\left(s'f\left[1 - \left(\sum_{j \in \mathcal{G}} m_j R_{j,0} + m_0 \frac{(1 - R_{0,0})}{N}\right)\right] + f'(1-s)\left[1 - \left(\sum_{j \in \mathcal{G}} P_j R_{j,0} + P_0 \frac{(1 - R_{0,0})}{N}\right)\right]\right).$$
$$(A\text{-}51)$$

Substituting eq. (7), we get

$$L = \frac{1}{f}\left(s'f \lim_{\mu \to 0} \frac{1}{(1 - Q_{0,0})}\left[1 - \left(\sum_{j \in \mathcal{G}} m_j Q_{j,0} + m_0 \frac{(1 - Q_{0,0})}{N}\right)\right]\right.$$
$$\left. + f'(1-s)\lim_{\mu \to 0} \frac{1}{(1 - Q_{0,0})}\left[1 - \left(\sum_{j \in \mathcal{G}} P_j Q_{j,0} + P_0 \frac{(1 - Q_{0,0})}{N}\right)\right]\right), \quad (A\text{-}52)$$

where the term in the outmost parenthesis is the same as the denominator of eq. (A-40) and thus in force of eq. (A-48) gives

$$L = \frac{1}{f}\left(N\left[s'f + f'(1-s)\right] + \mathcal{L}_0(F) - \frac{(1+s)\left[s'f + 2f'(1-s)\right]}{2D}\right) \quad (A\text{-}53)$$

For a Wright-Fisher process where $s' = 0$ and $s = 0$, we have

$$L = \frac{f'}{f}\left(N - \frac{1}{D}\right) = \frac{f'(ND - 1)}{fD}. \quad (A\text{-}54)$$

Appendix E Explicit coefficient for the selection gradient

We here derive eq. (61). Inserting the second line of eq. (55) into eq. (28) yields

$$K = \sum_{t=1}^{\infty} \sum_{\mathbf{k} \in \mathcal{G}} e_{\mathbf{k},t} \left[\frac{\mathcal{L}_{\mathbf{k}}(G_t) - (1+s) \left[s'f + 2f'(1-s) \right] / (2D)}{N \left[s'f + f'(1-s) \right] + \mathcal{L}_{\mathbf{0}}(F) - (1+s) \left[s'f + 2f'(1-s) \right] / (2D)} \right] \tag{A-55}$$

$$= \sum_{t=1}^{\infty} \sum_{\mathbf{k} \in \mathcal{G}} e_{\mathbf{k},t} \left[\frac{\frac{1}{D} \sum_{\mathbf{j} \in \mathcal{G}} G_t(\mathbf{j}) \overline{\chi}_{\mathbf{k}}(\mathbf{j}) - \frac{(1+s)[s'f+2f'(1-s)]}{2D}}{N \left[s'f + f'(1-s) \right] + \frac{1}{D} \sum_{\mathbf{j} \in \mathcal{G}} F(\mathbf{j}) - \frac{(1+s)[s'f+2f'(1-s)]}{2D}} \right]$$

$$= \sum_{t=1}^{\infty} \sum_{\mathbf{k} \in \mathcal{G}} e_{\mathbf{k},t} \left[\frac{\frac{1}{D} \sum_{\mathbf{j} \in \mathcal{G} \backslash \mathbf{0}} G_t(\mathbf{j}) \overline{\chi}_{\mathbf{k}}(\mathbf{j})}{N \left[s'f + f'(1-s) \right] + \frac{1}{D} \sum_{\mathbf{j} \in \mathcal{G} \backslash \mathbf{0}} F(\mathbf{j}) - \frac{[s'f+f'(1-s)]}{D}} \right]$$

$$= \sum_{t=1}^{\infty} \sum_{\mathbf{k} \in \mathcal{G}} e_{\mathbf{k},t} \left[\frac{\sum_{\mathbf{j} \in \mathcal{G} \backslash \mathbf{0}} G_t(\mathbf{j}) \overline{\chi}_{\mathbf{k}}(\mathbf{j})}{(ND - 1) \left[s'f + f'(1-s) \right] + \sum_{\mathbf{j} \in \mathcal{G} \backslash \mathbf{0}} F(\mathbf{j})} \right],$$

where the peultimate equality follows from eq. (56) and noting that

$$G_t(\mathbf{0}) = \frac{(1+s) \left[s'f + f'(1-s)(1+\mathcal{M}(\mathbf{0})) \right] \left[s + (1-s)\mathcal{M}(\mathbf{0}) \right]^t}{1 + s + (1-s)\mathcal{M}(\mathbf{0})} = \frac{(1+s) \left[s'f + 2f'(1-s) \right]}{2} \tag{A-56}$$

and

$$F(\mathbf{0}) = -\frac{(1-s) \left[s'f - f'2s \right] \mathcal{M}(\mathbf{h})}{1 + s + (1-s)\mathcal{M}(\mathbf{h})} = -\frac{(1-s) \left[s'f - 2f's \right]}{2}. \tag{A-57}$$

Now setting $H = (ND - 1) \left[s'f + f'(1-s) \right] + \sum_{\mathbf{j} \in \mathcal{G} \backslash \mathbf{0}} F(\mathbf{j})$ and using eq. (16), we get

$$K = \frac{1}{H} \sum_{t=1}^{\infty} \sum_{\mathbf{k} \in \mathcal{G}} \left[\frac{1}{D} \sum_{\mathbf{i} \in \mathcal{G}} \mathcal{C}(\mathbf{i})^{t-1} \Psi(\mathbf{i}) \overline{\chi}_{\mathbf{k}}(\mathbf{i}) \right] \left[\sum_{\mathbf{j} \in \mathcal{G} \backslash \mathbf{0}} G_t(\mathbf{j}) \overline{\chi}_{\mathbf{k}}(\mathbf{j}) \right] \tag{A-58}$$

$$= \frac{1}{H} \sum_{t=1}^{\infty} \sum_{\mathbf{i} \in \mathcal{G}} \sum_{\mathbf{j} \in \mathcal{G} \backslash \mathbf{0}} \frac{1}{D} \mathcal{C}(\mathbf{i})^{t-1} \Psi(\mathbf{i}) G_t(\mathbf{j}) \sum_{\mathbf{k} \in \mathcal{G}} \overline{\chi}_{\mathbf{k}}(\mathbf{i}) \overline{\chi}_{\mathbf{k}}(\mathbf{j})$$

$$= \frac{1}{H} \sum_{t=1}^{\infty} \sum_{\mathbf{j} \in \mathcal{G} \backslash \mathbf{0}} \mathcal{C}(-\mathbf{j})^{t-1} \Psi(-\mathbf{j}) G_t(\mathbf{j}),$$

where the penultimate equality follows from using eq. (46) Inserting eq. (56) we get

$$K = \frac{1}{H} \sum_{t=1}^{\infty} \sum_{\mathbf{j} \in \mathcal{G} \backslash \mathbf{0}} \mathcal{C}(-\mathbf{j})^{t-1} \Psi(-\mathbf{j}) \frac{(1+s) \left[s'f + f'(1-s)(1+\mathcal{M}(\mathbf{j})) \right] \left[s + (1-s)\mathcal{M}(\mathbf{j}) \right]^t}{1 + s + (1-s)\mathcal{M}(\mathbf{j})} \tag{A-59}$$

$$= \frac{1}{H} \sum_{\mathbf{j} \in \mathcal{G} \backslash \mathbf{0}} \frac{(1+s) \left[s'f + f'(1-s)(1+\mathcal{M}(\mathbf{j})) \right] \left[s + (1-s)\mathcal{M}(\mathbf{j}) - \mathcal{C}(-\mathbf{j})\mathcal{M}(\mathbf{j}) \right] \Psi(-\mathbf{j})}{\left[1 + s + (1-s)\mathcal{M}(\mathbf{j}) \right] \left[1 - \mathcal{C}(-\mathbf{j}) \right] \left[1 - \mathcal{C}(-\mathbf{j})\mathcal{M}(\mathbf{j}) \right]},$$

where the second equality follows from taking the geometric series which then yields eq. (61). If the moduli of $\mathcal{M}(\mathbf{j})$ and $\mathcal{C}(\mathbf{j})$ are smaller than one (i.e., $|\mathcal{M}(\mathbf{j})| < 1$ and $|\mathcal{C}(\mathbf{j})| < 1$) for all $\mathbf{j} \in \mathcal{G} \backslash \mathbf{0}$, then the geometric series converges. Indeed, the complex numbers $\mathcal{M}(\mathbf{j})$ and $\mathcal{C}(\mathbf{j})$ are within the unit circle for the following two reasons. First, the property of characteristic functions of probability distribution

entails that $\mathcal{M}(\mathbf{0}) = 1$, and $|\mathcal{M}(\mathbf{j})| < 1$ for $\mathbf{j} \in \mathcal{G} \setminus \mathbf{0}$ (Grimmett and Stirzaker, 2001, p. 182). Second, $|\mathcal{C}(\mathbf{j})| < 1$ is ensured by our assumption that the dynamical system eq. (2) has an hyperbolically stable equilibrium point, which in turn requires that all the eigenvalues of the Jacobian matrix of eq. (2) have modulus smaller than one (e.g., Galor, 2007, p. 103). This is because the coefficients $\mathcal{C}(\mathbf{j})$ are in fact the eigenvalues of the Jacobian matrix. To see this, first note that from eq. (2) the Jacobian of this discrete-time dynamical system around the equilibrium $\hat{n}$ defined by eq. (3) is given by

$$
\mathbf{J} = \begin{pmatrix}
\dfrac{\partial F(\mathbf{z_{0,t}}, \mathbf{n_{0,t}})}{\partial n_{\mathbf{0}}} & \dfrac{\partial F(\mathbf{z_{0,t}}, \mathbf{n_{0,t}})}{\partial n_{\mathbf{1}}} & \dfrac{\partial F(\mathbf{z_{0,t}}, \mathbf{n_{0,t}})}{\partial n_{\mathbf{2}}} & \cdots & \dfrac{\partial F(\mathbf{z_{0,t}}, \mathbf{n_{0,t}})}{\partial n_{\mathbf{D_{-1}}}} \\[2.5ex]
\dfrac{\partial F(\mathbf{z_{1,t}}, \mathbf{n_{1,t}})}{\partial n_{\mathbf{0}}} & \dfrac{\partial F(\mathbf{z_{1,t}}, \mathbf{n_{1,t}})}{\partial n_{\mathbf{1}}} & \dfrac{\partial F(\mathbf{z_{1,t}}, \mathbf{n_{1,t}})}{\partial n_{\mathbf{2}}} & \cdots & \dfrac{\partial F(\mathbf{z_{1,t}}, \mathbf{n_{1,t}})}{\partial n_{\mathbf{D_{-1}}}} \\[2.5ex]
\dfrac{\partial F(\mathbf{z_{2,t}}, \mathbf{n_{2,t}})}{\partial n_{\mathbf{0}}} & \dfrac{\partial F(\mathbf{z_{2,t}}, \mathbf{n_{2,t}})}{\partial n_{\mathbf{1}}} & \dfrac{\partial F(\mathbf{z_{2,t}}, \mathbf{n_{2,t}})}{\partial n_{\mathbf{2}}} & \cdots & \dfrac{\partial F(\mathbf{z_{2,t}}, \mathbf{n_{2,t}})}{\partial n_{\mathbf{D_{-1}}}} \\[2.5ex]
\vdots & \vdots & \vdots & \ddots & \vdots \\[2.5ex]
\dfrac{\partial F(\mathbf{z_{D_{-1},t}}, \mathbf{n_{D_{-1},t}})}{\partial n_{\mathbf{0}}} & \dfrac{\partial F(\mathbf{z_{D_{-1},t}}, \mathbf{n_{D_{-1},t}})}{\partial n_{\mathbf{1}}} & \dfrac{\partial F(\mathbf{z_{D_{-1},t}}, \mathbf{n_{D_{-1},t}})}{\partial n_{\mathbf{2}}} & \cdots & \dfrac{\partial F(\mathbf{z_{D_{-1},t}}, \mathbf{n_{D_{-1},t}})}{\partial n_{\mathbf{D_{-1}}}}
\end{pmatrix},
$$
$$\text{(A-60)}$$

where all derivatives are evaluated at z and $\hat{n}$. Now, recalling the notations defined in eq. (A-11), the entries of this matrix are of the form

$$
c_{\mathbf{k-i}} = \frac{\partial F(\mathbf{z_{k,t}}, \mathbf{n_{k,t}})}{\partial n_{\mathbf{i},t}} = \frac{\partial F(\mathbf{z_{k-i,t}}, \mathbf{n_{k-i,t}})}{\partial n_{\mathbf{0},t}} = \frac{\partial F(\mathbf{z_{0,t}}, \mathbf{n_{0,t}})}{\partial n_{\mathbf{k-i},t}}, \tag{A-61}
$$

which is the same as eq. (A-11) since all phenotypes vectors, here and there, are set to $(z, ..., z)$ when computing the derivative. From the first equality in the previous equation, the Jacobian (A-62) can be written as

$$
\mathbf{J} = \begin{pmatrix}
c_{\mathbf{0}} & c_{\mathbf{-1}} & c_{\mathbf{-2}} & \cdots \\[2ex]
c_{\mathbf{1}} & c_{\mathbf{0}} & c_{\mathbf{-1}} & \cdots \\[2ex]
\vdots & \vdots & \vdots & \cdots \\[2ex]
c_{\mathbf{D_{-1}}} & c_{\mathbf{D_{-2}}} & c_{\mathbf{D_{-3}}} & \ddots
\end{pmatrix}, \tag{A-62}
$$

where we defined $\mathbf{D_{-2}} = \mathbf{D_{-1}} - \mathbf{1}$, $\mathbf{D_{-3}} = \mathbf{D_{-1}} - \mathbf{2}$, etc. Written in this form, it is clear that the Jacobian (A-62) is a $\mathcal{G}$-group circulant matrix (e.g., Diaconis, 1988, p. 50), with eigenvalues given by the Fourier transform of $c_{\mathbf{j}}$ (Diaconis, 1988, Theorem 8). Hence, the $\mathbf{k}$-th eigenvalue of $\mathbf{J}$ is $\mathcal{C}(\mathbf{k}) = \sum_{\mathbf{j} \in \mathcal{G}} c_{\mathbf{j}} \chi_{\mathbf{j}}(\mathbf{k})$.

Appendix F Public good diffusion example

Appendix F.1 Fecundity effects

Appendix F.1.1 Covariance between dispersal and resource diffusion

We here derive eq. (38). On inserting eq. (25) into eq. (36) yields

$$\Omega = \sum_{t=1}^{\infty} \sum_{\mathbf{k}\in\mathcal{G}} (1-\epsilon)^{t-1} q_{\mathbf{k},t} \left(\frac{D p_{\mathbf{k},t} - 1}{DN - 1} \right) \tag{A-63}$$

$$= \sum_{t=1}^{\infty} \sum_{\mathbf{k}\in\mathcal{G}} (1-\epsilon)^{t-1} \left(\frac{D q_{\mathbf{k},t} p_{\mathbf{k},t} - 1}{DN - 1} \right)$$

$$= \frac{D}{ND - 1} \sum_{t=1}^{\infty} (1-\epsilon)^{t-1} \left(\sum_{\mathbf{k}\in\mathcal{G}} p_{\mathbf{k},t} q_{\mathbf{k},t} - \frac{1}{D} \right).$$

This expression can be written in terms of the population covariance of $p_{\mathbf{k},t}$ and $q_{\mathbf{k},t}$ in the following way. Recall that the population covariance of two vectors x and y of length n is

$$\mathrm{cov}(x, y) = \frac{1}{n} \sum_{j=1}^{n} x_j y_j - \bar{x}\bar{y} \tag{A-64}$$

where $\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j$ and $\bar{y} = \frac{1}{n}\sum_{j=1}^{n} y_j$. Using this definition, we can write

$$\mathrm{cov}(p_t, q_t) = \frac{1}{D} \sum_{\mathbf{k}\in\mathcal{G}} p_{\mathbf{k},t} q_{\mathbf{k},t} - \left(\frac{1}{D} \sum_{\mathbf{k}\in\mathcal{G}} p_{\mathbf{k},t} \right) \left(\frac{1}{D} \sum_{\mathbf{k}\in\mathcal{G}} q_{\mathbf{k},t} \right) \tag{A-65}$$

$$= \frac{1}{D} \sum_{\mathbf{k}\in\mathcal{G}} p_{\mathbf{k},t} q_{\mathbf{k},t} - \frac{1}{D^2}$$

$$= \frac{1}{D} \left(\sum_{\mathbf{k}\in\mathcal{G}} p_{\mathbf{k},t} q_{\mathbf{k},t} - \frac{1}{D} \right)$$

where the second line follows from the fact that both $p_{\mathbf{k},t}$ and $q_{\mathbf{k},t}$ are probability mass functions and hence satisfy $\sum_{\mathbf{k}\in\mathcal{G}} p_{\mathbf{k},t} = \sum_{\mathbf{k}\in\mathcal{G}} q_{\mathbf{k},t} = 1$. Now setting $\mathrm{cov}_t^{\mathrm{T}}(p_t, q_t) = D\mathrm{cov}(p_t, q_t)$, we obtain

$$\Omega = \frac{D}{ND - 1} \sum_{t=1}^{\infty} (1-\epsilon)^{t-1} \mathrm{cov}_t^{\mathrm{T}}(p_t, q_t). \tag{A-66}$$

Appendix F.1.2 Weak dispersal

We here derive eq. (40) by using the common approach to evaluate a weak migration approximation (Rousset, 2004, chapter 3). For this, set $m_0 = (1 - m)$ and $d_0 = (1 - d)$, where m and d are the net

dispersal probabilitie of the focal species and the environmental variable. Further write $m_{\mathbf{i}} = mg_{\mathbf{i}}^{\mathbf{m}}$ and $d_{\mathbf{i}} = dg_{\mathbf{i}}^{\mathbf{d}}$, whereby the characteristic functions of the dispersal distributions can be expressed as $\mathcal{M}(\mathbf{j}) = 1 - mx^{\mathbf{m}}(\mathbf{j})$ and $\mathcal{D}(\mathbf{j}) = 1 - dx^{\mathbf{d}}(\mathbf{j})$, where $x^{\mathbf{m}}(\mathbf{j}) = 1 - \sum_{\mathbf{i}\neq\mathbf{0}} g_{\mathbf{i}}^{\mathbf{m}}\chi_{\mathbf{i}}(\mathbf{j})$ and $x^{\mathbf{d}}(\mathbf{j}) = 1 - \sum_{\mathbf{i}\neq\mathbf{0}} g_{\mathbf{i}}^{\mathbf{d}}\chi_{\mathbf{i}}(\mathbf{j})$. Substituting these expressions into the last line of eq. (38) , Taylor expanding around $m = 0$ and $d = 0$, inserting back $x^{\mathbf{m}}(\mathbf{j}) = (1 - \mathcal{M}(\mathbf{j}) - 1)/m$ and $x^{\mathbf{d}}(\mathbf{j}) = (1 - \mathcal{D}(\mathbf{j}) - 1)/d$, we obtain that

$$\Omega = \frac{1}{ND - 1} \sum_{\mathbf{j}\in\mathcal{G}\backslash\mathbf{0}} \frac{\mathcal{D}(\mathbf{j}) + \mathcal{M}(\mathbf{j}) + \epsilon - 2}{\epsilon^2} + \text{h.o.t.,} \tag{A-67}$$

where "h.o.t." refers to higher order, e.g., m^2, md, d^2 etc. Neglecting these and using $\mathcal{M}(\mathbf{j}) = \sum_{\mathbf{k}\in\mathcal{G}} m_{\mathbf{k}}\chi_{\mathbf{k}}(\mathbf{j})$, and $\mathcal{D}(\mathbf{j}) = \sum_{\mathbf{k}\in\mathcal{G}} d_{\mathbf{k}}\chi_{\mathbf{k}}(\mathbf{j})$ produces

$$\begin{aligned}
\Omega &= \frac{1}{ND - 1}\left[\sum_{\mathbf{j}\in\mathcal{G}\backslash\mathbf{0}}\sum_{\mathbf{k}\in\mathcal{G}}\left(\frac{m_{\mathbf{k}} + d_{\mathbf{k}}}{\epsilon^2}\right)\chi_{\mathbf{k}}(\mathbf{j}) + \sum_{\mathbf{j}\in\mathcal{G}\backslash\mathbf{0}}\left(\frac{\epsilon - 2}{\epsilon^2}\right)\right] \tag{A-68} \\
&= \frac{1}{ND - 1}\left[\sum_{\mathbf{k}\in\mathcal{G}}\left(\frac{m_{\mathbf{k}} + d_{\mathbf{k}}}{\epsilon^2}\right)\left(\sum_{\mathbf{j}\in\mathcal{G}}\chi_{\mathbf{k}}(\mathbf{j}) - 1\right) + \sum_{\mathbf{j}\in\mathcal{G}\backslash\mathbf{0}}\left(\frac{\epsilon - 2}{\epsilon^2}\right)\right] \\
&= \frac{1}{ND - 1}\left[\left(\frac{m_{\mathbf{0}} + d_{\mathbf{0}}}{\epsilon^2}\right)(D - 1) + (D - 1)\left(\frac{\epsilon - 2}{\epsilon^2}\right)\right] \\
&= \left(\frac{D - 1}{ND - 1}\right)\left(\frac{\epsilon - m - d}{\epsilon^2}\right),
\end{aligned}$$

where the penultimate equality follows from the fact that $\sum_{\mathbf{k}\in\mathcal{G}}\chi_{\mathbf{k}}(\mathbf{j}) = D$ if $\mathbf{j} = \mathbf{0}$, zero otherwise (recall eq. (46)) and $\sum_{\mathbf{j}\in\mathcal{G}\backslash\mathbf{0}} = D - 1$.

Appendix F.2   Survival effects

Consider now only survival effects ($f' = 0$). Then inserting eq. (25) into eq. (36) yields

$$\Omega = \frac{1}{(DN - 1) - \sum_{\mathbf{j}\in\mathcal{G}\backslash\mathbf{0}}\frac{(1-s)\mathcal{M}(\mathbf{j})}{1+s+(1-s)\mathcal{M}(\mathbf{j})}} \sum_{\mathbf{j}\in\mathcal{G}\backslash\mathbf{0}} \frac{(1+s)\left[s + (1-s)\mathcal{M}(\mathbf{j}) - (1-\epsilon)\mathcal{M}(\mathbf{j})\mathcal{D}(\mathbf{j})\right]\mathcal{D}(\mathbf{j})}{\left[1 + s + (1-s)\mathcal{M}(\mathbf{j})\right]\left[1 - (1-\epsilon)\mathcal{D}(\mathbf{j})\right]\left[1 - (1-\epsilon)\mathcal{M}(\mathbf{j})\mathcal{D}(\mathbf{j})\right]}, \tag{A-69}$$

which remains a somewhat complicated expression. When $s \to 1$, we get

$$\Omega = \frac{1}{ND - 1}\sum_{\mathbf{j}\in\mathcal{G}\backslash\mathbf{0}}\frac{\mathcal{D}(\mathbf{j})}{1 - (1 - \epsilon)\mathcal{D}(\mathbf{j})}. \tag{A-70}$$

This can be thought as a special case where investment into the common-pool resource occurs in a population of immortal individuals that therefore become mortal through endogeneously induced deaths.

Appendix F.3   Characteristic function of dispersal and commons movement

Here, we specify the probability distributions and characteristic functions for the dispersal of the focal species and the "dispersal" (movement) of the commons, which we used to generate the figures of the main text.

Appendix F.3.1   One-dimensional habitat

Let us first consider a one-dimensional habitat consisting of a circular lattice, so that the set of patches is $\mathcal{G} = \mathbb{Z}_D$ where $\mathbb{Z}_D = \{0, \dots, D-1\}$ is the set of integers modulo D. We assume that D is odd. We assume an individual disperses with probability m and with probability $1-m$ stays in its natal patch. Given an individual disperses, it disperses with equal probability "clockwise" or "counterclockwise" a number $j \in \{1, 2, \dots, (D-1)/2\}$ of steps, which we assume follows a truncated binomial distribution with probability mass function,

$$p_j(N_{\mathrm{s}}, q) = \frac{1}{1 - (1-q)^{N_{\mathrm{s}}}} \binom{N_{\mathrm{s}}}{j} q^j (1-q)^{N_{\mathrm{s}}-j}, \tag{A-71}$$

where $N_{\mathrm{s}} = (D-1)/2$ is the number of trials, $q = 2\lambda_{\mathrm{m}}/(D-1)$ is the probability of success, such that $\lambda_{\mathrm{m}} = N_{\mathrm{s}}q$ is approximately the mean number of steps an individual disperses conditional on dispersal (λ_{m} is actually the mean of the non-truncated distribution). From these assumptions, we have $m_0 = (1-m)$ and $m_j = mp_j$. The characteristic function of the dispersal distribution then is

$$\begin{aligned}
\mathcal{M}(k) =& \frac{1}{2} \sum_{j=0}^{\frac{D-1}{2}} m_j \chi_j(k) + \frac{1}{2} \sum_{j=0}^{\frac{D-1}{2}} m_j \overline{\chi}_j(k) \\
=& (1-m) + m \sum_{j=1}^{\frac{D-1}{2}} p_j((D-1)/2, 2\lambda_{\mathrm{m}}/(D-1)) \left(\frac{\overline{\chi}_j(k) + \chi_j(k)}{2} \right) \\
=& (1-m) + m \sum_{j=1}^{\frac{D-1}{2}} p_j((D-1)/2, 2\lambda_{\mathrm{m}}/(D-1)) \cos(2\pi jk/D),
\end{aligned} \tag{A-72}$$

which is determined by the parameters D, m, and λ_{m}.

For the distribution of movement in space of the commons, we make the same assumptions as above but let the probability of movement be d (instead of m) and the mean number of steps be λ_{d} (instead of λ_{m}) so that the characteristic function is

$$\mathcal{D}(k) = (1-d) + d \sum_{j=1}^{\frac{D-1}{2}} p_j((D-1)/2, 2\lambda_{\mathrm{d}}/(D-1)) \cos(2\pi jk/D). \tag{A-73}$$

Appendix F.3.2   Two-dimensional habitat

For the two dimensional case, we assume that patches are arranged on a torus with the same number of patches in each dimension so that $\mathcal{G} = \{(k_1, k_2) : 0 \leq k_j < D^{1/2}\}$ for k_1 and k_2 modulo $D^{1/2}$. The dispersal distribution of the focal species $m_{\mathbf{k}}$ for $\mathbf{k} = (k_1, k_2) \in \mathcal{G}$, is constructed similarly as above. First, an individual disperses with probability m and with probability $1 - m$ stays in its natal patch. Second, conditional on dispersal, we sample the number of steps $j \in \{1, 2, ..., D^{1/2} - 1\}$ an individual disperses on the lattice (maximum $D^{1/2} - 1$) from a truncated Binomial distribution $p_j(N_s, q)$ (eq. A-71) with parameters $N_s = D^{1/2} - 1$ and $q = \lambda_m / (D^{1/2} - 1)$. Third, we determine how this total number of steps j is divided between j_1 steps in dimension 1 and j_2 steps in dimension 2 (so with $j = j_1 + j_2$) assuming that dispersal in either dimension has the same distribution. We do so by sampling j_1 from a discrete uniform distribution $\text{unif}(j_{\min}, j_{\max})$, where the minimum and maximum are given by

$$
\begin{aligned}
j_{\min} &= \max\left(0, j - \frac{D^{1/2} - 1}{2}\right) \\
j_{\max} &= \min\left(j, \frac{D^{1/2} - 1}{2}\right)
\end{aligned}
\tag{A-74}
$$

and setting $j_2 = j - j_1$. Finally, given the number of steps in each dimension j_1 and j_2, these are then equally likely to occur in either direction away from the focal patch. Details on this computation and on how to obtain the generating functions used in the figures can be found in the accompanying Mathematica notebook.

For the distribution of movement in space of the commons under this two dimensional habitat we make exactly the same assumptions as the commons, except that the probability of movement out of the focal patch is d (instead of m) and the mean number of steps of movement is λ_d (instead of λ_m).

References

Alger, I., J. W. Weibull, and L. Lehmann. 2020. Evolution of preferences in structured populations: Genes, guns, and culture. *Journal of Economic Theory* 185:1–45.

Bruner, J. P. and R. Smead. 2022. Tag-based spite with correlated interactions. *Journal of Theoretical Biology* 540:111052.

Cadet, C., R. Ferrière, J. A. J. Metz, and M. van Baalen. 2003. The evolution of dispersal under demographic stochasticity. *American Naturalist* 162:427–41.

Comins, H. N., W. D. Hamilton, and R. M. May. 1980. Evolutionarily stable dispersal strategies. *Journal ot Theoretical Biology* 82:205–230.

Dawkins, R. 1982. The Extended Phenotype. Oxford University Press, Oxford.

Diaconis, P. 1988. Group representations in probability and statistics. Institute of Mathematical Statistics Lecture Notes.

Eshel, I. 1983. Evolutionary and continuous stability. *Journal ot Theoretical Biology* 103:99–111.

Estrela, S., E. Libby, J. Van Cleve, F. Débarre, M. Deforet, W. R. Harcombe, J. Peña, S. P. Brown, and M. E. Hochberg. 2019. Environmentally Mediated Social Dilemmas. *Trends Ecol Evol* 34:6–18.

Frank, S. A. 1998. Foundations of Social Evolution. Princeton University Press, Princeton, NJ.

Galor, O. 2007. Discrete Dynamical Systems. Springer, Berlin.

Gandon, S. and F. Rousset. 1999. Evolution of stepping-stone dispersal rates. *Proceedings of the Royal Society of London Series B-Biological Sciences* 221:2507–2513.

Gardner, A. and S. A. West. 2004. Spite and the scale of competition. *Journal of Evolutionary Biology* 17:1195–1203.

Grafen, A. and M. Archetti. 2008. Natural selection of altruism in inelastic viscous homogeneous populations. *Journal of Theoretical Biology* 252:694–710.

Grimmett, G. and D. Stirzaker. 2001. Probability and Random Processes. Oxford University Press, Oxford.

Hamilton, W. D. 1970. Selfish and spiteful behavior in an evolutionary model. *Nature* 228:1218–1220.

Hamilton, W. D. 1971. Selection of selfish and altruistic behaviour in some extreme models. In Eisenberg, J. and W. Dillon (eds.), *Man and Beast: Comparative Social Behavior*, pp. 59–91. Smithsonian Institutions Press, Washington, DC.

Irwin, A. J. and P. D. Taylor. 2001. Evolution of altruism in stepping-stone populations with overlapping generations. *Theoretical Population Biology* 60:315–325.

Jansen, V. A. A. and A. L. Lloyd. 2000. Local stability analysis of spatially homogeneous solutions of multi-patch systems. *J. Math. Biol.* 41:232–252.

Le Galliard, J., R. Ferrière, and U. Dieckmann. 2003. The adaptive dynamics of altruism in spatially heterogeneous populations. *Evolution* 57:1–17.

Le Galliard, J., R. Ferrière, and U. Dieckmann. 2005. Adaptive Evolution of Social Traits: Origin, Trajectories. *American Naturalist* 165:206–224.

Lehmann, L. 2007. The evolution of trans-generational altruism: kin selection meets niche construction. *Journal of Evolutionary Biology* 20:181–189.

Lehmann, L. 2008. The adaptive dynamics of niche constructing traits in spatially subdivided populations: evolving posthumous extended phenotypes. *Evolution* 62:549–566.

Lehmann, L. 2010. Space-time relatedness and Hamilton's rule for long-lasting behaviors in viscous populations. *American Naturalist* 175:136–143.

Lehmann, L. 2012. The stationary distribution of a continuously varying strategy in a class-structured population under mutation-selection-drift balance. *Journal of Evolutionary Biology* 25:770–787.

Lehmann, L., M. W. Feldman, and F. Rousset. 2009. On the evolution of harming and recognition in finite panmictic and infinite structured populations. *Evolution* 63:2896–2913.

Lehmann, L., L. Keller, and D. Sumpter. 2007. The evolution of helping and harming on graphs: the return of the inclusive fitness effect. *Journal of Evolutionary Biology* 20:2284–229.

Lehmann, L., N. Perrin, and F. Rousset. 2006. Population demography and the evolution of helping behaviors. *Evolution* 60:1137–1151.

Lehmann, L. and F. Rousset. 2010. How life-history and demography promote or inhibit the evolution of helping behaviors. *Philosophical Transactions of the Royal Society B* 365:2599–2617.

Lehmann, L. and F. Rousset. 2012. The evolution of social discounting in hierarchically clustered populations. *Molecular Ecology* 21:447–471.

Lehmann, L. and F. Rousset. 2014. The genetical theory of social behaviour. *Philosophical Transactions of the Royal Society B* 369:1–18.

Lion, S. and M. van Baalen. 2007. From infanticide to parental care: Why spatial structure can help adults be good parents. *American Naturalist* 170:E26–E46.

Malécot, G. 1975. Heterozygosity and relationship in regularly subdivided populations. *Theoretical Population Biology* 8:212–241.

McNamara, J. M. and O. Leimar. 2020a. Game Theory in Biology. Oxford University Press.

McNamara, J. M. and O. Leimar. 2020b. Game theory in biology: concepts and frontiers. Oxford University Press.

Mullon, C., L. Keller, and L. Lehmann. 2018. Social polymorphism is favoured by the co-evolution of dispersal with social behaviour. *Nat Ecol Evol* 2:132–140.

Mullon, C. and L. Lehmann. 2018. Eco-evolutionary dynamics in metacommunities: ecological inheritance, helping within species, and harming between species. *American Naturalist* 192:664–685.

Nakamaru, M. and Y. Iwasa. 2005. The evolution of altruism by costly punishment in lattice-structured populations: score-dependent viability versus score-dependent fertility. *Evolutionary Ecology Research* 7:853–870.

Nakamaru, M., H. Matsuda, and Y. Iwasa. 1997. The evolution of cooperation in a lattice-structured population. *Journal of Theoretical Biology* 184:65–81.

Novak, M., J. D. Yeakel, A. E. Noble, D. F. Doak, M. Emmerson, J. A. Estes, U. Jacob, M. T. Tinker, and J. T. Wootton. 2016. Characterizing species interactions to understand press perturbations: what is the community matrix? *Annual Review of Ecology, Evolution, and Systematics* 47:409–432.

Ohtsuki, H. 2010. Evolutionary games in Wright's island model: kin selection meets evolutionary game theory. *Evolution* 64:3344–53.

Ohtsuki, H., C. Hauert, E. Lieberman, and M. A. Nowak. 2006. A simple rule for the evolution of cooperation on graphs and social networks. *Nature* 441:502–505.

Ohtsuki, H. and M. A. Nowak. 2006. Evolutionary games on cycles. *Proceedings of the Royal Society of London Series B-Biological Sciences* 273:2249–2256.

Ohtsuki, H., C. Rueffler, J. Y. Wakano, K. Parvinen, and L. Lehmann. 2020. The components of directional and disruptive selection in heterogeneous group-structured populations. *Journal of Theoretical Biology* 507:1–16.

Parvinen, K., U. Dieckmann, M. Gyllenberg, and J. A. J. Metz. 2003. Evolution of dispersal in metapopulations with local density dependence and demographic stochasticity. *Journal of Evolutionary Biology* 16:143–53.

Peña, J., B. Wu, J. Arranz, and A. Traulsen. 2016. Evolutionary Games of Multiplayer Cooperation on Graphs. *PLOS Computational Biology* 12:1–15.

Rogers, A. R. 1990. Group selection by selective emigration: the effects of migration and kin structure. *American Naturalist* 135:398–413.

Rousset, F. 2000. Genetic differentiation between individuals. *Journal of Evolutionary Biology* 13:58–62.

Rousset, F. 2003. A minimal derivation of convergence stability measures. *Journal of Theoretical Biology* 221:665–668.

Rousset, F. 2004. Genetic Structure and Selection in Subdivided Populations. Princeton University Press, Princeton, NJ.

Rousset, F. and S. Billiard. 2000. A theoretical basis for measures of kin selection in subdivided populations: finite populations and localized dispersal. *Journal of Evolutionary Biology* 13:814–825.

Rousset, F. and S. Gandon. 2002. Evolution of the distribution of dispersal distance under distance-dependent cost of dispersal. *Journal of Evolutionary Biology* 15:515–523.

Rousset, F. and O. Ronce. 2004. Inclusive fitness for traits affecting metapopulation demography. *Theoretical Population Biology* 65:127–141.

Sawyer, S. 1976. Results for the stepping stone model for migration in population genetics. *Annals of Probability* 4:699–728.

Schaffer, M. E. 1988. Evolutionarily stable strategies for a finite population and a variable contest size. *Journal of Theoretical Biology* 132:469–478.

Smead, R. and P. Forber. 2016. The coevolution of recognition and social behavior. *Scientifc Reports* 6:25813.

Taylor, P. D. 1992a. Altruism in viscous populations - an inclusive fitness model. *Evolutionary Ecology* 6:352–356.

Taylor, P. D. 1992b. Inclusive fitness in a homogeneous environment. *Proceedings of the Royal Society of London Series B-Biological Sciences* 240:299–302.

Taylor, P. D. 1994. Sex ratio in a stepping-stone population with sex-specific dispersal. *Theoretical Population Biology* 45:203–218.

Taylor, P. D., T. Day, and G. Wild. 2007. Evolution of cooperation in a finite homogeneous graph. *Nature* 447:469–472.

Taylor, P. D., T. Lilicrap, and C. D. 2011. Inclusive fitness analysis on mathematical groups. *Evolution* pp. 849–59.

van Baalen, M. and A. Rand. 1998. The unit of selection in viscous populations and the evolution of altruism. *Journal of Theoretical Biology* 193:631–648.

Van Cleve, J. 2015. Social evolution and genetic interactions in the short and long term. *Theoretical Population Biology* 103:2–26.

West, S. A. and A. Gardner. 2010. Altruism, spite, and greenbeards. *Science* 327:1341–1344.

Wild, G., A. Gardner, and S. A. West. 2009. Adaptation and the evolution of parasite virulence in a connected world. *Nature* 459:983–986.

Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16:97–159.

Yodzis, P. 1989. Introduction to Theoretical Ecology. Harper and Row, New York.