Partitioning the Impacts of Spatial-Temporal Variation in Demography and Dispersal on Metapopulation Growth Rates

Sebastian J. Schreiber
Department of Evolution and Ecology
University of California, Davis USA 95616

Abstract. Spatial-temporal variation in environmental conditions is ubiquitous in nature. This variation simultaneously impacts survival, reproduction, and movement of individuals and, thereby, the rate at which metapopulations grow. Using the tools of stochastic demography, the metapopulation growth rate is decomposed five components corresponding to temporal, spatial, and spatial-temporal variation in fitness, and spatial and spatial-temporal covariation in dispersal and fitness. While temporal variation in fitness always reduces the metapopulation growth rate, all other sources of variation can either increase or reduce the metapopulation growth rate. Increases occur either by reducing the impacts of temporal variation or by generating a positive fitness-density covariance where individuals tend to concentrate in higher-quality patches. For example, positive auto-correlations in spatial-temporal variability in fitness generate this positive fitness-density covariance for less dispersive populations, but decrease this covariance for highly dispersive populations e.g. migratory species. Negative autocorrelations in spatial-temporal variability have the opposite effects. Positive covariances between movement and future fitness, on short or long timescales, increase growth rates. These positive covariances can arise is unexpected ways. For example, the win-stay,lose-shift dispersal strategy in negatively autocorrelated environments can generate positive spatial covariances that exceed negative spatial-temporal covariances. This decomposition of the metapopulation growth rate provides a way to quantify the relative importance of fundamental sources of variation on metapopulation persistence.

Introduction

All populations experience spatial and temporal variation in environmental conditions. Within populations, this variation generates individual variation in survival, growth, and fecundity across space and time. Faced with these demographic consequences, individuals may disperse between habitat patches and, thereby, alter their demographic fates. As environmental conditions vary across time, so may patterns of dispersal. This covariation in demography and dispersal in space and time can impact the rate at which populations grow or decline. Many studies have identified important ways that this covariation impacts population growth [Hastings, 1983, Levin et al., 1984, Wiener and Tuljapurkar, 1994, Jansen and Yoshimura, 1998, Schmidt, 2004, Roy et al., 2005, Schreiber and Lloyd-Smith, 2009, Schreiber, 2010, Evans et al., 2013, Williams and Hastings, 2013, Snyder et al., 2014, Crone, 2016]. However, as these studies typically focused on only one or two sources of variation, there remain gaps in our understanding of these impacts on population growth and the relative importance of these different sources of variation for population growth rates.

Temporal and spatial variation in demography, in and of themselves, can have opposing effects on population growth. For unstructured populations (e.g. no spatial structure), temporal variation in survival or fecundity reduces population growth rates [Lewontin and Cohen, 1969, Tuljapurkar, 1990, Boyce et al., 2006]. This reduction stems from population growth being a multiplicative process and reductions in survival or fecundity having a larger impact on population growth than comparable increases in these vital rates. For sedentary populations, spatial variability in survival or fecundity always increases population growth rates [Karlin, 1976, Hastings, 1983, Schreiber and Lloyd-Smith, 2009, Snyder and Chesson, 2003, Cantrell et al., 2017]. This positive effect stems from populations growing most rapidly where fitness is highest and individuals rarely dispersing elsewhere. As the population concentrates in locations with high fitness, the covariance between population density and fitness becomes positive i.e. a positive fitness-density covariance [Chesson, 2000]. Crone [2016] illustrated these theoretical conclusions using integral projection models for the prairie forb, *Pulsatilla patens*. For these relatively sedentary populations, Crone [2016] found that "environmental stochasticity reduced population growth rates relative to the average, whereas spatial heterogeneity increased population growth rates." For highly mobile populations, theory predicts that the positive effect of spatial variability on population growth is reduced and possibly reversed [Karlin, 1976, Hastings, 1983, Kirkland et al., 2006, Schreiber and Lloyd-Smith, 2009, Altenberg, 2012a,b]. For example, following rapid environmental change, individuals may mistakenly prefer patches with lower fitness. This generates a negative fitness-density covariance and reduces the population growth rate [Battin, 2004, Schreiber and Lloyd-Smith, 2009, Hale and Swearer, 2016]. However, the exact conditions under which spatial heterogeneity, in general, reduces population growth rates are not known.

The simultaneous impacts of spatial and temporal variation in demography on population growth are more nuanced. For spatially structured populations, serially uncorrelated temporal variation in demographic rates still reduce population growth rates [Tuljapurkar, 1990, Boyce et al., 2006]. However, if this temporal variation is spatially asynchronous, dispersal can act as a spatial bet-hedging strategy and neutralize the negative impact of temporal variation on population growth [Kuno, 1981, Metz et al., 1983, Wiener and Tuljapurkar, 1994, Jansen and Yoshimura, 1998, Evans et al., 2013, Jaggi et al., 2023, Kortessis et al., 2023]. Notably, populations coupled by dispersal can have positive growth rates despite all local growth rates being negative [Metz et al., 1983, Jansen and Yoshimura, 1998, Evans et al., 2013, Jaggi et al., 2023]. In contrast to serially uncorrelated temporal variation, positively auto-correlated fluctuations in survival or reproduction can increase growth rates of moderately dispersing populations [Roy et al., 2005, Schreiber, 2010, Kortessis et al., 2020]. Intuitively, positive auto-correlations imply that higher quality patches at one moment tend to remain higher quality. Population densities rapidly build up in these patches and dispersal seeds other patches to capitalize on shifting environmental conditions. In the long-term, this results in population densities accumulating in higher quality patches and higher population growth rates [Roy et al., 2005, Kortessis et al., 2023]. Consistent with these theoretical predictions, experiments with *Paramecium aurelia* found that temporally autocorrelated fluctuations in temperature and moderate dispersal rates led to positive population growth rates despite negative growth rates on average [Matthews and Gonzalez, 2007]. These studies,

however, didn't consider the effects of spatial heterogeneity or negatively auto-correlated fluctuations on population growth.

Variation in environmental conditions can drive variation in rates of active or passive dispersal [Schmidt, 2004, Haugen et al., 2006, Ellner and Schreiber, 2012, Williams and Hastings, 2013, Snyder et al., 2014, Catalano et al., 2021, Mason et al., 2022, Peniston et al., 2023]. Active dispersers may use environmental cues to determine whether or not to leave their current patch [Schmidt, 2004, Clobert, 2012, Carroll et al., 2018]. For example, little penguins (*Eudyptula minor*) exhibit behavior consistent with the win-stay,lose-shift rule. Individuals were less likely to leave sites where they had greater foraging success and more likely to leave sites where foraging was less successful [Carroll et al., 2018]. Simulating mathematical models, Schmidt [2004] found that the win-stay,lose-shift rule combined with positively auto-correlated fluctuations in site quality increased population growth rates. Intuitively, with positive auto-correlations, site quality tends to persist over time. Hence, by remaining in higher quality sites and leaving lower quality sites, individuals tends to concentrate in higher quality sites and increase the population growth rate. Movement of passive dispersers, on the other hand, varies with the speed and direction of environmental currents (e.g. air, wind), their dispersal vectors, and their physiological state [O'Connor et al., 2007, Clobert, 2012, Catalano et al., 2021, Mason et al., 2022, Peniston et al., 2023]. For example, pelagic larvae transported by spatially and temporally varying ocean currents [Catalano et al., 2021] and the time spent riding these currents is temperature dependent [O'Connor et al., 2007]. Theory predicts that this temporal variation in habitat connectivity can reduce or inflate population growth rates [Williams and Hastings, 2013, Snyder et al., 2014]. However, as these theories consider specific scenarios (i.e. two patches with no demographic variation, or serially uncorrelated fluctuations), they do no provide general principles of how variable dispersal impacts population growth rates.

Collectively, these earlier studies highlight the complex ways that spatial and temporal variation in demography and dispersal can impact population growth. Despite significant progress, there remain several gaps in understanding the impacts of certain forms of variation. More importantly, the relative importance of these different sources of variation on population growth is not well understood. To address these issues, I analyze a class of stochastic, matrix models with any number of patches and accounting for many forms of spatial-temporal variation in demography and dispersal. The analysis provides a method to decompose the population growth rate into 6 contributions due to the average fitness, temporal, spatial, and spatial-temporal variation in fitness, and spatial and spatial-temporal covariation in fitness and dispersal. Using this decomposition one can evaluate the relative importance of different sources of variation for metapopulation persistence.

Model and Methods

A Metapopulation Model. The model represents a metapopulation living in k patches coupled by dispersal. These patches may correspond to geographically, distinct patches of one or more habitat types, or discretizations of a continuous landscape. The population density in the i-th patch is $n_{i,t}$ at time t. The population census times $t = 0, 1, 2, 3, \ldots$ are discrete. Between successive censuses, I

assume that survival and reproduction are followed by dispersal. This ordering, however, is not critical for the main results. Survival and reproduction in patch i changes the local population density by a multiplicative factor $\lambda_{i,t}$ at time t. Even though the model doesn't assume that individuals only live for one time step, I call this multiplicative factor $\lambda_{i,t}$ the fitness of an individual in patch i at time t. After survival and reproduction, a fraction $d_{ij,t}$ of individuals from patch j disperse to patch i. Summing over all the contributions of individuals to each patch determines the metapopulation dynamics:

$$n_{i,t+1} = \sum_{j=1}^{k} d_{ij,t}\lambda_{j,t}n_{j,t} \text{ for } i = 1, 2, \ldots, k. \tag{1}$$

For mathematical convenience, I assume all dispersing individuals arrive in some patch i.e. $\sum_{j=1}^{k} d_{ji} = 1$. To account for mortality due to dispersal, one can reduce the fitness in each patch by the fraction dying due to dispersal i.e. if a fraction $\delta_{ij,t}$ die dispersing from patch j to i then $\lambda_{j,t}$ becomes $\lambda_{j,t}\sum_{i=1}^{k} d_{ij,t}(1-\delta_{ij,t})$. While this model does not include density-dependence, one can view it as a linearization of a density-dependent model about the extinction equilibrium. This linearization determines how metapopulations grow at low population densities and, consequently, whether the metapopulation persists or not [Benaïm and Schreiber, 2009, Roth and Schreiber, 2014, Benaïm and Schreiber, 2019].

One can rewrite the model (1) in a simpler form using vector-matrix notation. The vector of population densities at time t is

$$\mathbf{n}_t = \begin{pmatrix} n_{1,t} \\ \vdots \\ n_{k,t} \end{pmatrix} \quad .$$

Survival and reproduction at time t corresponds to multiplying by the demography matrix

$$\Lambda_t = \begin{pmatrix} \lambda_{1,t} & 0 & \ldots & 0 \\ 0 & \lambda_{2,t} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_{k,t} \end{pmatrix}.$$

Dispersal at time t corresponds to multiplying by the dispersal matrix

$$D_t = \begin{pmatrix} d_{11,t} & d_{12,t} & \ldots & d_{1k,t} \\ d_{21,t} & d_{22,t} & \ldots & d_{2k,t} \\ \vdots & \vdots & \ddots & \vdots \\ d_{k1,t} & d_{k2,t} & \ldots & d_{kk,t} \end{pmatrix}.$$

With this vector-matrix notation, the model (1) becomes

$$\mathbf{n}_{t+1} = D_t\Lambda_t\mathbf{n}_t. \tag{2}$$

Fluctuations in environmental conditions drive the fluctuations in demography Λ_t and dispersal D_t. The environmental state θ_t at time t lies in the set Θ of possible environmental states. This set Θ may consist of a finite number of environmental states $\{1, 2, \ldots, m\}$ e.g. wet and cool versus hot and dry. Alternatively, Θ may consist of continuum of states $\Theta \subset (-\infty, \infty)^m$ (e.g. temperature driven vital rates), or some combination of continuous and discrete states. Consistent with classical stochastic demography [Ruelle, 1979, Tuljapurkar, 1990, Boyce et al., 2006], I assume that the sequence of environmental states $\theta_1, \theta_2, \theta_3, \ldots$ is (asymptotically) stationary and ergodic. For example, for a discrete environmental state space, the environmental dynamics may be given by an irreducible, aperiodic Markov chain on $\Theta = \{1, 2, \ldots, m\}$. Alternatively, for a continuous environmental state space, the dynamics may be a multivariate auto-regressive process. To couple population dynamics to the environmental dynamics, the entries of the demographic and dispersal matrices are viewed as functions of the environment. Namely, in environment θ, demography matrix is the diagonal matrix $\Lambda(\theta)$ with diagonal entries $\lambda_i(\theta)$ and the dispersal matrix is $\mathbf{D}(\theta)$ with entries $d_{ij}(\theta)$. With this coupling to the environmental dynamics, the demographic and dispersal matrices at time t in (2) are

$$\Lambda_t = \Lambda(\theta_t) \text{ and } D_t = \mathbf{D}(\theta_t). \tag{3}$$

Under suitable technical assumptions, Ruelle [1979] provided that there is quantity r characterizing the long-term growth rate of the metapopulation. Specifically, if

$$\text{the total density } |\mathbf{n}_t| = \sum_{i=1}^{k} n_{i,t}$$

at time t, then

$$\text{the metapopulation growth rate } r = \lim_{t\to\infty} \frac{1}{t} \ln |\mathbf{n}_t| \text{ whenever } |\mathbf{n}_0| > 0.$$

Roughly, the total metapopulation population density, $|\mathbf{n}_t|$, grows exponentially at rate r; increasing when $r > 0$ and decreasing when $r < 0$. The sign of the metapopulation growth rate r has implications for models accounting for density-dependence or demographic stochasticity. For models with negative density-dependence, positive values of r imply the metapopulation persists, while negative values imply the metapopulation approaches extinction exponentially quickly [Benaïm and Schreiber, 2009, Benaïm and Schreiber, 2019]. For density-independent models with demographic stochasticity (i.e. a multi-type branching processes in a random environment [Athreya and Ney, 2004]), $r > 0$ implies that there is a positive probability of metapopulation establishment, while $r \leq 0$ implies establishment is unlikely.

Decomposing the sources of variation. In order to identify the relative roles of spatial and temporal variation on the metapopulation growth rate r, I decompose the demography and dispersal matrices into components corresponding to temporal variation, spatial variation, or spatial-temporal variation.

Averaging the fitnesses $\boldsymbol{\lambda}_t = \begin{pmatrix} \lambda_{1,t} \\ \lambda_{2,t} \\ \vdots \\ \lambda_{k,t} \end{pmatrix}$ across space at time t yields the scalar

$$\text{the spatial average of fitness at time } t \ \langle \boldsymbol{\lambda}_t \rangle = \frac{1}{k} \sum_{i=1}^{k} \lambda_{i,t}.$$

Averaging the fitnesses across time yields the vector

$$\text{the temporal averages of fitnesses } \overline{\boldsymbol{\lambda}} = \begin{pmatrix} \mathbb{E}[\lambda_{1,t}] \\ \mathbb{E}[\lambda_{2,t}] \\ \vdots \\ \mathbb{E}[\lambda_{k,t}] \end{pmatrix}$$

where $\overline{\lambda_i} = \mathbb{E}[\lambda_{i,t}]$ denotes the expected fitness in patch i. By stationarity of the environmental process, these expected fitnesses correspond to the long-term temporally averaged fitnesses and, therefore, do not depend on time t i.e. $\overline{\lambda_i} = \mathbb{E}[\lambda_{i,1}] = \mathbb{E}[\lambda_{i,2}] = \ldots$. Averaging simultaneously across space and time yields

$$\text{the spatial-temporal average of fitness } \overline{\langle \boldsymbol{\lambda} \rangle} = \frac{1}{k} \sum_{i=1}^{k} \mathbb{E}[\lambda_{i,1}].$$

Using these averages, the fitnesses $\boldsymbol{\lambda}_t$ can be decomposed into four components. The first component, $\overline{\langle \boldsymbol{\lambda} \rangle}$, is the average fitness experienced by the population across space and time. Temporal, spatial, and spatial-variability in the fitnesses are quantified as deviations from this spatial-temporal average. Subtracting the spatial-temporal average from spatial average in year t, $\langle \boldsymbol{\lambda}_t \rangle - \overline{\langle \boldsymbol{\lambda} \rangle}$, yields the temporal deviation in year t. Positive temporal deviations correspond to years where individuals, on average across space, have higher fitness. Subtracting the spatial-temporal average from the vector of temporal averages, $\overline{\boldsymbol{\lambda}} - \overline{\langle \boldsymbol{\lambda} \rangle}$, yields the vector of spatial deviations. Positive components of this vector correspond to patches that, on average across time, have higher fitness. The remaining deviations in fitness are the vector of spatial-temporal deviations

$$\boldsymbol{\lambda}_t - \overline{\langle \boldsymbol{\lambda} \rangle} - \underbrace{(\langle \boldsymbol{\lambda}_t \rangle - \overline{\langle \boldsymbol{\lambda} \rangle})}_{\text{temporal deviations}} - \underbrace{(\overline{\boldsymbol{\lambda}} - \overline{\langle \boldsymbol{\lambda} \rangle})}_{\text{spatial deviations}} = \underbrace{\boldsymbol{\lambda}_t - \langle \boldsymbol{\lambda}_t \rangle - \overline{\boldsymbol{\lambda}} + \overline{\langle \boldsymbol{\lambda} \rangle}}_{\text{spatial-temporal deviations}} . \tag{4}$$

Positive components of this vector correspond to years of higher fitness for a given patch. Rearranging (4) yields the following decomposition of the demography:

$$\boldsymbol{\lambda}_t = \overline{\langle \boldsymbol{\lambda} \rangle} + \underbrace{\langle \boldsymbol{\lambda}_t \rangle - \overline{\langle \boldsymbol{\lambda} \rangle}}_{\text{temporal deviations}} + \underbrace{\overline{\boldsymbol{\lambda}} - \overline{\langle \boldsymbol{\lambda} \rangle}}_{\text{spatial deviations}} + \underbrace{\boldsymbol{\lambda}_t - \langle \boldsymbol{\lambda}_t \rangle - \overline{\boldsymbol{\lambda}} + \overline{\langle \boldsymbol{\lambda} \rangle}}_{\text{spatial-temporal deviations}} \tag{5}$$

This decomposition of the demography is similar to decompositions used by Johnson and Hastings [2023] and Kortessis et al. [2023].

To quantify variation in dispersal, I assume there is a symmetric dispersal matrix D_{sym} (i.e. $d_{ij} = d_{ji}$ for all $i \neq j$) around which the spatial variation in dispersal occurs. For two patch models, this symmetric matrix is

$$D_{\text{sym}} = \begin{pmatrix} 1-d & d \\ d & 1-d \end{pmatrix} \tag{6}$$

where d is the probability an individual disperses to the other patch. For models with more patches, this symmetric matrix could represent a random walk or diffusion with any pattern of connectivity between the patches [Chung, 1997, Spielman, 2019]. There are two advantages of using a symmetric matrix D_{sym} as a frame of reference. First, in the absence of spatial variation in the fitnesses, symmetric dispersal ensures that, in the long-term, individuals are uniformly distributed across the patches i.e. there is no spatial heterogeneity in population densities. Second, symmetric matrices have an orthogonal basis of eigenvectors providing a natural way to decompose spatial variation using a generalization of discrete Fourier analysis [Stein and Shakarchi, 2011].

Variation of the dispersal matrix D_t around the reference dispersal matrix D_{sym} takes two forms: spatial variation and spatial-temporal variation. The spatial variation corresponds to the difference between the time-averaged dispersal matrix $\overline{D}$ and the reference matrix D_{sym}. A positive entry in this difference corresponds to increased movement, on average over time, between a pair of patches. Unlike demography, there is no natural way to define pure temporal variation in D_t. This stems from the spatial average of the entries of D_t not varying over time i.e. $\frac{1}{k^2}\sum_{i,j} d_{ij,t} = \frac{1}{k}$ for all t. Intuitively, movement into one patch is balanced by a corresponding movement out of another patch. Hence, beyond the spatial deviations, there are only the spatial-temporal deviations yielding the decomposition

$$D_t = D_{\text{sym}} + \underbrace{\overline{D} - D_{\text{sym}}}_{\substack{\text{spatial}\\ \text{deviations}}} + \underbrace{D_t - \overline{D}}_{\substack{\text{spatial-temporal}\\ \text{deviations}}} \tag{7}$$

Positive entries in the spatial-temporal deviations $D_t - \overline{D}$ correspond to years with more movement from one patch to another than in a typical year.

Analytical and numerical approximations of the metapopulation growth rate r. Finding explicit analytic expressions for the metapopulation growth rate r is generally intractable. However, using methods of stochastic demography [Ruelle, 1979, Tuljapurkar, 1990], I derive an analytical approximation of r when the temporal and spatial variations in demography and dispersal are small. To validate the accuracy of these analytical approximations, I numerically compute r using the power method [Caswell, 2001].

To derive the small variance approximation in a mathematically rigorous manner, I use the work of Ruelle [1979] who proved that r is analytic function of the environment-demography mapping $\theta \to \mathbf{D}(\theta)\Lambda(\theta) =: \mathbf{A}(\theta)$. Ruelle [1979]'s proof of this fact provides an approach to explicitly compute all

derivatives of r as a function of the environment-demography mapping. In particular, this approach provides an alternative, more direct method to compute the first order and second order approximations of Tuljapurkar [1990]. Furthermore, Ruelle [1979]'s approach also allows one to derive second-order mixed derivatives of r that are used to quantify the interactive effect of covariation in dispersal and demography on r. Using Ruelle [1979]'a approach, Appendix A derives these first-order and second-order derivatives for general stochastic, matrix models. Using these derivative formulas, Appendix B derives a second order Taylor expansion for r around the unperturbed model $\bar{A} = D_{\text{sym}}\overline{\langle \boldsymbol{\lambda} \rangle}$. The Taylor expansion consists of sum of three terms: the zero-th order term $\ln \overline{\langle \boldsymbol{\lambda} \rangle}$, a second-order correction term due to demographic variation, and a second-order mixed correction term due co-variation between demography and dispersal.

To approximate r numerically, I use the power method [see, e.g., Caswell, 2001] where one renormalizes the matrix dynamics in (2) at each time step:

$$\mathbf{v}_{t+1} = \frac{A_t \mathbf{v}_t}{|A_t \mathbf{v}_t|} \text{ with } \mathbf{v}_0 = \begin{pmatrix} 1/k \\ 1/k \\ \vdots \\ 1/k \end{pmatrix}. \tag{8}$$

The work of Ruelle [1979] implies that

$$r = \lim_{t \to \infty} \frac{1}{t} \sum_{s=1}^{t} \ln |A_s \mathbf{v}_s| \text{ with probability one.}$$

Hence,

$$r \approx \frac{1}{T} \sum_{s=1}^{T} \ln |A_s \mathbf{v}_s| \text{ for sufficiently large } T.$$

I use $T = 100,000$ time steps for these numerical approximations.

The role of the fitness-density covariance. The work of Ruelle [1979] implies that $\mathbf{v}_t$ converges to a stationary sequence $\widehat{\mathbf{v}}_t$ representing the spatial-temporal distribution of individuals in the population in the long-term. Notably, the metapopulation growth rate equals

$$r = \overline{\ln |A_t \widehat{\mathbf{v}}_t|}.$$

This equation provides a link to the fitness-density covariance of Chesson [2000]. Specifically, if

$$\text{cov}_{\text{space}}(\mathbf{u}, \mathbf{v}) = \langle \mathbf{u} - \langle \mathbf{u} \rangle, \mathbf{v} - \langle \mathbf{v} \rangle \rangle$$

denotes the spatial covariance of two arbitrary vectors $\mathbf{u}, \mathbf{v}$, then

$$\begin{aligned}|A_t\hat{\mathbf{v}}_t| =& \sum_{i=1}^{k} \lambda_{i,t} v_{i,t} \\ =& k\langle \boldsymbol{\lambda}_t\rangle\langle \mathbf{v}_t\rangle + k\,\mathrm{cov}_{\mathrm{space}}(\boldsymbol{\lambda}_{\mathrm{t}}, \mathbf{v}_{\mathrm{t}}) \\ =& \langle \boldsymbol{\lambda}_t\rangle + k\,\mathrm{cov}_{\mathrm{space}}(\boldsymbol{\lambda}_{\mathrm{t}}, \mathbf{v}_{\mathrm{t}})\end{aligned}$$

where the first line follows from the dispersal matrices being column stochastic, the second line follows from the definition of the spatial covariance, and the final line follows from $\langle \mathbf{v}_t\rangle = 1/k$ i.e. $|\mathbf{v}_t| = 1$. The expression, $k\,\mathrm{cov}(\boldsymbol{\lambda}_{\mathrm{t}}, \mathbf{v}_{\mathrm{t}})$, corresponds to the spatial covariance between the fitnesses at time t and the (normalized) population densities at time t. Chesson [2000] calls this expression, the fitness-density covariance. Putting these pieces together, one arrives at the following representation of the metapopulation growth rate r

$$r = \overline{\ln\left(\langle \boldsymbol{\lambda}_t\rangle + k\,\mathrm{cov}_{\mathrm{space}}(\boldsymbol{\lambda}_{\mathrm{t}}, \mathbf{v}_{\mathrm{t}})\right)}. \tag{9}$$

Roy et al. [2005] derived a similar formula to (9).

Results

The main result is an analytical formula for how spatial and temporal variation in demography and dispersal contribute to the metapopulation growth rate r. This formula consists of two main components: the contribution of demographic variation and an interactive effect demographic and dispersal variation. To highlight the key features of these two components, I present them first for the case of $k = 2$ patches and then extend to the case of any number of patches. I conclude by deriving a relationship between the fitness-density covariance of Chesson [2000] and the metapopulation growth rate decomposition.

Contributions of demographic variation to r for two patches. In case of $k = 2$ patches without dispersal variation, the dispersal matrix $D_t = D_{\mathrm{sym}}$ is characterized by the probability d that an individual disperses to the other patch, see equation (6). In the absence of births and deaths, individuals dispersing according to this dispersal matrix will ultimately by equally distributed between the two patches i.e. the stable spatial distribution is $\begin{pmatrix}\frac{1}{2}\\ \frac{1}{2}\end{pmatrix}$. Deviations away from this spatially homogeneous distribution decay exponentially like $(1-2d)^t$ as $t \to \infty$. This decay occurs more quickly when the fraction dispersing is closer to 50%. Moreover, this decay occurs monotonically when no more than 50% of the population is dispersing ($d < 0.5$). In this case, I call the population *philopatric* (home-loving). When the fraction dispersing is greater than 50% ($d > 0.5$), I call the population *miseopatric* (home-hating). The spatial distribution of miseopatric populations exhibits dampened oscillations when approaching a spatially homogeneous state; the patch with the greatest number of individuals alternating between the two patches from time step to time step.

In the presence of spatially and temporally varying fitnesses, the metapopulation experiences a temporal variance $\sigma^2_{\text{time}} = \overline{(\langle \boldsymbol{\lambda}_t \rangle - \overline{\langle \boldsymbol{\lambda} \rangle})^2}$ in its spatially averaged fitnesses with coefficient of variation $\text{CV}_{\text{time}} = \sigma_{\text{time}}/\overline{\langle \boldsymbol{\lambda} \rangle}$, and a spatial variance $\sigma^2_{\text{space}} = (\overline{\lambda_1} - \overline{\langle \boldsymbol{\lambda} \rangle})^2 = (\overline{\lambda_2} - \overline{\langle \boldsymbol{\lambda} \rangle})^2$ in its temporally averaged fitnesses with coefficient of variation $\text{CV}_{\text{space}} = \sigma_{\text{space}}/\overline{\langle \boldsymbol{\lambda} \rangle}$. In the case of $k = 2$ patches, the vector of spatial-temporal deviations (final term in equation (5)) always have opposite sign; the spatial-temporal deviations are

$$\begin{pmatrix} \alpha_t \\ -\alpha_t \end{pmatrix} \text{ where } \alpha_t = \lambda_{1,t} - \overline{\lambda_1} - \langle \boldsymbol{\lambda}_t - \overline{\boldsymbol{\lambda}} \rangle.$$

α_t measures the direction and magnitude of the spatial asynchrony between the two patches. When α_t is zero, fitness in both patches are equally higher or equally lower than their respective temporal averages. When α_t is positive, fitness in patch 1 is higher than its average fitness while fitness in patch 2 is correspondingly lower than its average fitness. Let $\sigma^2_{\text{st}} = \overline{\alpha^2}$ be the temporal variation in the spatial asynchrony and $\text{CV}_{\text{st}} = \sigma_{\text{st}}/\overline{\langle \boldsymbol{\lambda} \rangle}$ the associated coefficient of variation. The ratio of variances, $\frac{\sigma^2_{\text{st}}}{\sigma^2_{\text{st}} + \sigma^2_{\text{time}}}$, is a metric of spatial asynchrony: zero corresponds to perfect spatial synchrony while one corresponds to perfect spatial asynchrony. Let $\rho_{\text{st}} = \overline{\alpha_t \alpha_{t+1}}/\sigma^2_{\text{st}}$ be the temporal autocorrelation in the α_t. When the autocorrelation ρ_{st} is positive, time steps in which the fitness in patch 1 is greater than its average tend to be followed by time steps where the fitness in patch 1 continues to be higher than its average. When the autocorrelation ρ_{st} is negative, time steps in which the fitness in patch 1 is higher than its average tend to be followed by time steps where the fitness in patch 2 is higher than its average.

The analytical approximation of the metapopulation growth rate r depends on the average fitness and the temporal, spatial, and spatial-temporal deviations from this average i.e. each of the terms of the decomposition of fitness in equation (5). Specifically,

$$r \approx \ln \overline{\langle \boldsymbol{\lambda} \rangle} \underbrace{-\frac{1}{2}\text{CV}^2_{\text{time}}}_{\substack{\text{temporal}\\ \text{deviations}}} \underbrace{+\frac{1-2d}{2d}\text{CV}^2_{\text{space}}}_{\substack{\text{spatial}\\ \text{deviations}}} \underbrace{+\frac{\rho_{\text{st}}(1-2d)}{1-(1-2d)\rho_{\text{st}}}\text{CV}^2_{\text{st}}}_{\substack{\text{spatial-temporal}\\ \text{deviations}}} \tag{10}$$

The first term $\ln \overline{\langle \boldsymbol{\lambda} \rangle}$ corresponds to the metapopulation growth rate in the absence of spatial and temporal variation. Consistent with classical stochastic demography, the second term, $-\frac{1}{2}\text{CV}^2_{\text{time}}$, implies that temporal variation in demography reduces the metapopulation growth rate. In contrast, spatial variation (the third term) increases the metapopulation growth rate for philopatric populations ($d < 0.5$) and decreases the metapopulation growth rate for miseopatric populations ($d > 0.5$). The impact of spatially asynchronous fluctuations on the metapopulation growth rate (the fourth term) depend simultaneously on the temporal auto-correlation ρ_{st} and whether the populations are philopatric or miseopatric. For philopatric populations, positive autocorrelations increase the metapopulation growth rate, while negative autocorrelations reduce the metapopulation growth rate. For miseopatric populations, these impacts are reversed.

To illustrate these theoretical implications, I first consider numerical scenarios inspired by the experiments of Matthews and Gonzalez [2007] with metapopulations of *Paramecium aurelia*. By varying

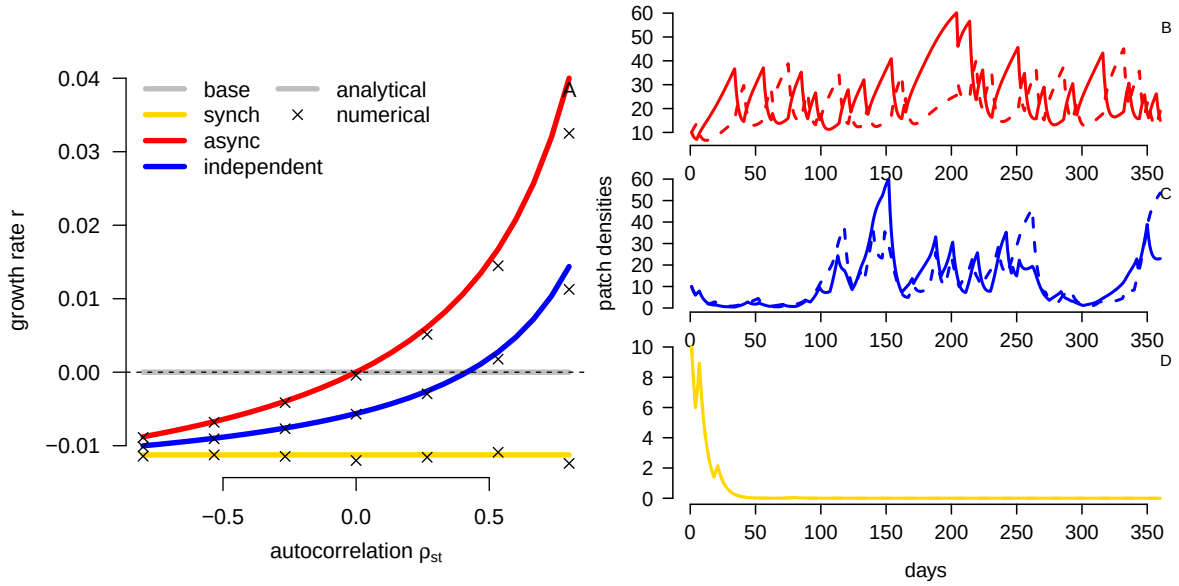

FIGURE 1. Impacts of temporal and spatial-temporal variation in fitness on two patch metapopulations. In (A), the metapopulation growth rate r as a function of the autocorrelation in the spatial-temporal variation. Solid curves correspond to the analytical approximation (13) and crosses correspond to numerical approximations. Spatially synchronized fluctuations in yellow, spatially asynchronous fluctuations in green, and spatially independent fluctuations in yellow. The gray line corresponds to $\ln \overline{\langle \boldsymbol{\lambda} \rangle}$. In (B)-(D), population dynamics with negative density-dependence added when $\rho_{\text{st}} = 0.8$. Different colors as described for (A). Parameters: $\lambda_{i,t} = 0.85$ and 1.15 with equal likelihood, and $d = 0.1$. For the spatially synchronous (respectively, asynchronous) models there are two environmental states corresponding to $\lambda_{i,t} = 1.15$ in both patches or $\lambda_{i,t} = 0.85$ in both patches (respectively, 1.15 in one patch, and 0.85 in the other patch). For the spatially independent model, there are four environmental states corresponding to $\lambda_{i,t} = 1.15$ in both patches, $\lambda_{i,t}$ =0.85 in both patches, and $\lambda_{i,t}$ =1.15 in one patch, $\lambda_{i,t}$ =0.85 in the other. For B-D, density-dependence is added by multiplying $\lambda_{i,t}$ by $\exp(-0.001 n_{i,t})$.

temperature between two values, Matthews and Gonzalez [2007] varied fitness within flasks between low ($\lambda_{i,t} < 1$) and high ($\lambda_{i,t} > 1$) values. This temperature variation was chosen so that the average fitness was less than one ($\ln \overline{\langle \boldsymbol{\lambda} \rangle} < 0$). Each day (a time step), 10% of the individuals were moved between pairs of flasks ($d = 0.1$). Like their experiments, my numerical experiments varied temporal and spatial correlations in fitness (Fig. 1). In these scenarios, the analytical approximation of the metapopulation growth rate r closely matched the numerically estimated values of r (solid curves versus

crosses in Fig 1A). For spatially synchronized metapopulations ($\sigma^2_{\text{st}} = 0$), temporal variation in fitness reduced the metapopulation growth rate (yellow versus gray curves in Fig. 1A). When the patches are completely spatially asynchronous (i.e. when one patch experiences high fitness, the other experiences low fitness), the metapopulation only experiences spatial-temporal fluctuations in fitness ($\sigma^2_{\text{time}} = 0$); there is no reduction due to temporal variation. As the metapopulation is philopatric ($d = 0.1$), positive autocorrelations in this spatial-temporal variation increase the metapopulation growth rate, while negative autocorrelations decrease the metapopulation growth rate (green versus gray curves in Fig. 1A). Sufficiently positive autocorrelations lead to a positive metapopulation growth rate despite both patches being sinks (green curve above dashed line in Fig. 1A). When the patches are spatially independent, the joint effects of temporal variation and spatial-temporal variation result in intermediate metapopulation growth rates (orange versus yellow and green curves in Fig. 1A). This stems from spatial independence simultaneously reducing the temporal variation and the spatial-temporal variation terms of the metapopulation growth rate decomposition (10). To illustrate these implications for populations experiencing density-dependence, I ran simulations replacing the density-independent fitnesses $\lambda_{i,t}$ with density-dependent fitnesses $\lambda_{i,t}e^{-0.001n_{i,t}}$. Consistent with mathematical theory [Benaïm and Schreiber, 2019], the metapopulation growth rate r for the density-independent model predicts the long-term fate of metapopulations experiencing negative density-dependence (Fig. 1B–D); the metapopulation persisting when $r > 0$ (Fig. 1B,C), and going extinct when $r < 0$ (Fig. 1D).

To illustrate the joint effects of spatial variation and dispersal rates on the metapopulation growth rate, I consider a scenario where patch 1 has, on average, higher but more variable fitness, while patch 2 experiences no temporal variation (Fig. 2A). While dispersal rates and temporal autocorrelations have no effect on the temporal variation component of the metapopulation growth rate r (not shown), the spatial and spatial-temporal components of r depend on both factors (Fig. 2B). For all levels of temporal autocorrelation, the spatial component of r decreases with the dispersal rate d and becomes negative when the population is miseopatric (gray dashed curve in Fig. 2B). In contrast, the effect of dispersal on the spatial-temporal component of r depends on the sign of the temporal autocorrelation: a positive effect for negative temporal autocorrelation (green curve in Fig. 2B) and a negative effect for positive temporal autocorrelation (orange curve in Fig. 2B). For negative temporal autocorrelations, the opposing effects of dispersal on two components of r result in a non-monotonic effect of dispersal on r (orange curve in Fig. 2A). Indeed, r is locally maximized at low and high dispersal rates and is minimized at intermediate dispersal rates.

Interactive effects of dispersal and fitness variation for two patches. In the case of $k = 2$ patches, spatial and temporal variation in dispersal is determined by the changes $\delta_{i,t}$ in the fraction d dispersing from patch $i = 1, 2$. Namely,

$$D_t = D_{\text{sym}} + \begin{pmatrix} -\delta_{1,t} & \delta_{2,t} \\ \delta_{1,t} & -\delta_{2,t} \end{pmatrix} \tag{11}$$

The spatial component in the dispersal matrix decomposition (7) is determined by the temporal averages $\overline{\delta_i}$ and spatial-temporal component is determined by the deviations $\delta_{i,t} - \overline{\delta_i}$.

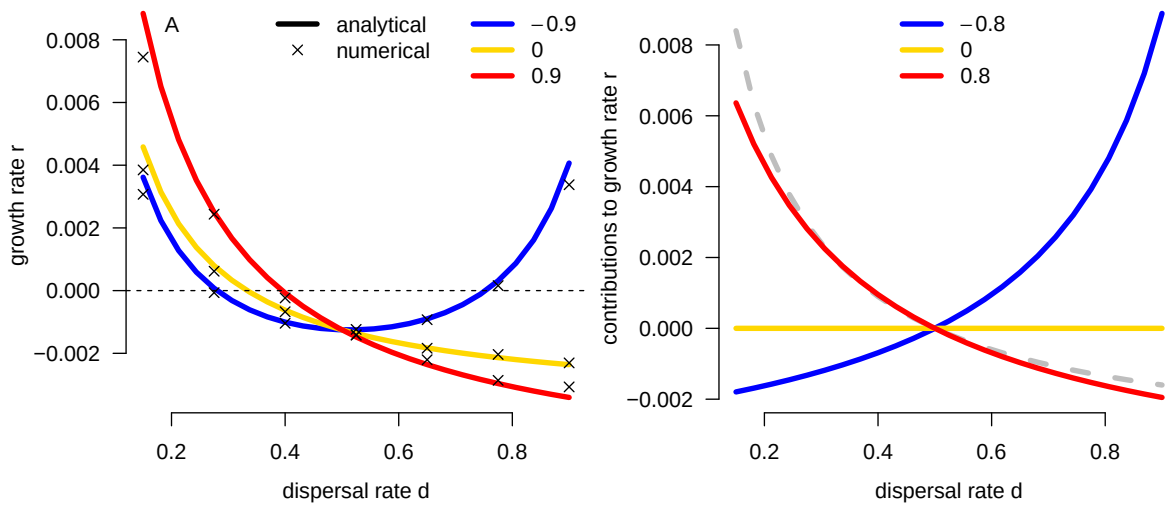

FIGURE 2. Impacts of spatial variation, temporal auto-correlation, and dispersal on the metapopulation growth rate r. Fitness is variable and, on average, higher in patch 1 and constant, but lower in patch 2. In (A), r as function of the dispersal rate d. Solid curves correspond to the analytical approximation (10) and crosses correspond to the numerical approximations. Different colors correspond to different values of spatial-temporal auto-correlation. In (B), the contributions of spatial variation (dashed gray) and spatial-temporal variation (yellow,orange,green) on r. Parameters: $\lambda_{i,t}$ in patch $i = 1$ are $0.95, 1.15$ with equal probability, and always 0.95 in patch $i = 2$.

While dispersal variation, in and of itself, has no impact on the metapopulation growth rate, it does have an impact when it covaries with demographic variation. The spatial component of this impact stems from the spatial covariance between dispersal and demography,

$$\mathrm{cov}_{D\Lambda\text{-space}} = \mathrm{cov}_{\mathrm{space}}(\overline{\boldsymbol{\lambda}}, \overline{\delta}) = (\overline{\lambda_1} - \overline{\langle \boldsymbol{\lambda} \rangle})(\overline{\delta_2} - \overline{\delta_1}).$$

Positive values of this covariance occur when individuals, on average across time, exhibit biased movement to the patch with the higher average fitness. The impact of spatial-temporal variation on the metapopulation growth rate depends on the temporal covariance between the difference in dispersal rates between the two patches at one time step and the corresponding difference in fitness between the two patches in the next time step:

$$\mathrm{cov}_{D\Lambda\text{-st}} = \mathrm{cov}_{\mathrm{time}}[\delta_{2,\mathrm{t}} - \delta_{1,\mathrm{t}}, \lambda_{1,\mathrm{t+1}} - \lambda_{2,\mathrm{t+1}}]$$

where $\mathrm{cov}_{\mathrm{time}}(\mathrm{x}, \mathrm{y}) = \overline{(\mathrm{x} - \overline{\mathrm{x}})(\mathrm{y} - \overline{\mathrm{y}})}$ is the covariance for two stationary time series x_t, y_t. Let $\rho_{D\Lambda\text{-st}}$ be the correlation associated with the covariance $\mathrm{cov}_{D\Lambda\text{-st}}$. Positive correlations occur when individuals

are more likely to disperse to the patch with the higher fitness in the next time step. The effects of these spatial and spatial-temporal covariances on the metapopulation growth rate r are

$$\underbrace{\frac{1}{2d}\frac{\mathrm{cov}_{D\Lambda\text{-space}}}{\overline{\langle\boldsymbol{\lambda}\rangle}^2}}_{\substack{\text{spatial}\\ \text{covariance}}} + \underbrace{\frac{1-2d}{1-(1-2d)\rho_{D\Lambda\text{-st}}}\frac{\mathrm{cov}_{D\Lambda\text{-st}}}{\overline{\langle\boldsymbol{\lambda}\rangle}^2}}_{\substack{\text{spatial-temporal}\\ \text{covariance}}}. \tag{12}$$

Each contribution is positive only if the corresponding covariance is positive i.e. individuals tend to move to patches of higher quality. The magnitude of these contributions decreases with the base dispersal rate d.

To illustrate some of the implications of the covariation of dispersal and demography on the metapopulation growth rate, Figure 3 considers a population whose individuals make use of the win-stay,lose-shift (WSLS) dispersal strategy [Schmidt, 2004]. Individuals using this strategy remain in patches that supported the highest fitness, else disperse to the other patch. Dispersing individuals (a fraction d of the population) utilize the WSLS strategy with probability p. In the numerical experiments, fitness in the patches fluctuates between lower and higher values. Patch 1, on average, supports the higher fitness, but has the lowest fitness one-quarter of the time. Greater use of the WSLS strategy (higher p) always increases the metapopulation growth rate (Fig 3A), but this increase is weakest when there are negative temporal autocorrelations (orange curve in Fig. 3A). At first, it may seem surprising that the WSLS strategy enhances metapopulation growth in a strongly negatively autocorrelated environment. Indeed, individuals using this strategy are more likely to end up in a patch with lower fitness in the next time step leading to a negative spatial-temporal covariance between dispersal and fitness (orange curve in Fig. 3B). However, the spatial component always has a positive effect on the metapopulation growth rate (dashed gray curve in Fig. 3B). This positive effect stems from the WSLS strategy ensuring individuals spend three-quarters of their time, on average, in the patch with the higher average fitness.

Decomposition for any number of patches. In the case of $k = 2$ patches, there is only one degree of freedom for spatial variation around the temporal averages of fitness, $\overline{\boldsymbol{\lambda}}$, corresponding to multiples of the vector $\begin{pmatrix}1\\-1\end{pmatrix}$. With $k \geq 3$ patches, there are $k-1 \geq 2$ degrees of freedom. A mathematically convenient way to characterize these $k-1$ degrees of freedom is to use a form of discrete Fourier analysis that decomposes spatial deviations from a spatially homogeneous state into a weighted sum of $k-1$ spatial modes of variation (SMV) i.e. an orthogonal basis of subdominant eigenvectors $\mathbf{u}(1), \mathbf{u}(2), \ldots, \mathbf{u}(k-1)$, of the dispersal matrix D_{sym} normalized such that $\sum_{i=1}^{k} u_i(j)^2 = k$. In the absence of variation in fitness or dispersal, spatial variation in the population densities along SMV $\mathbf{u}(i)$ decays at a fixed rate μ_i due to dispersal according to D_{sym} where μ_i is the eigenvalue associated with $\mathbf{u}(i)$. While the magnitude of μ_i is always less than one, its sign can be positive or negative. I call spatial modes of variation for which $\mu_i > 0$ *philopatric spatial modes of variation*; when the initial spatial distribution of the population has all of its variation in a philopatric spatial mode,

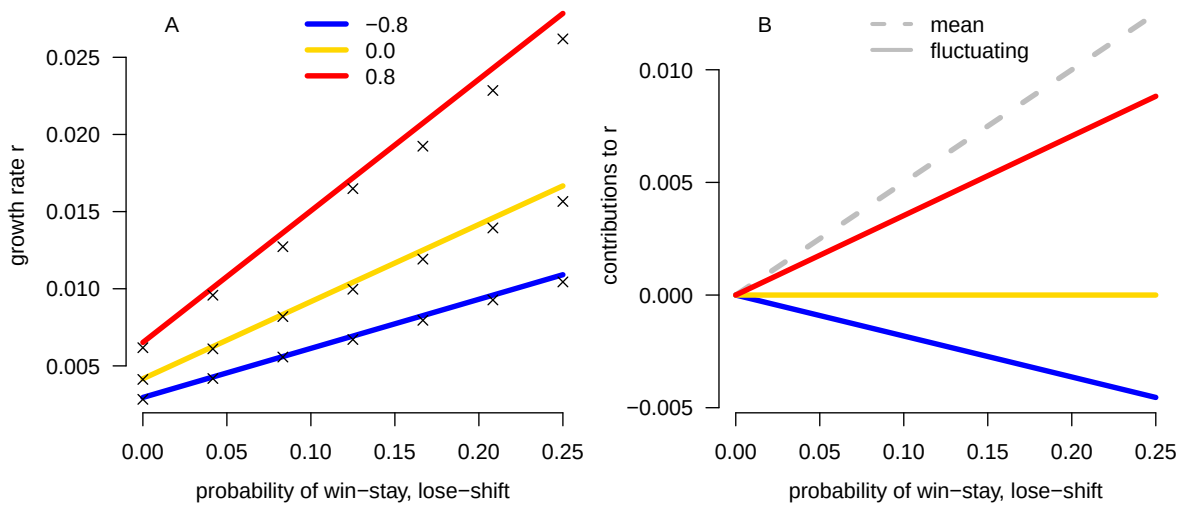

FIGURE 3. Impacts of fluctuating dispersal and demography on population growth. The fitness in patch 1 is, on average in time, higher than in patch 2, but is lower one-quarter of the time. In each time step, individuals disperse with probability d from patches with lower fitness and with probability pd from patches with higher fitness. In (A), the population growth rate plotted as function of the probability p of a dispersing individual playing the win-stay,lose-shift strategy. Solid curves correspond to the analytical approximation (13) and crosses correspond numerical approximations. Different colors correspond to different values of temporal auto-correlation. In (B), the contribution of the mean (dashed gray) and fluctuating (solid curves) biases in dispersal on the population growth rate. Parameters: Parameters: $\lambda_{i,t}$ in patch $i = 1$ are $1.2, 1.0$ with equal probability, and $1.0, 0.8$ in patch $i = 2$ with equal probability. Fitness are spatially independent and $d = 0.3$.

patches with higher densities will continue to have higher densities in future time steps. When the initial spatial distribution has all of its spatial variation in a SMV with negative values of μ_i, patches with the highest densities will have the lower densities in the next time step. These are *miseopatric spatial modes of variation*. By the Gershgorin circle theorem [Horn and Johnson, 1990], miseopatric SMV only occur if there is at least one patch where at least 50% of the individuals leave i.e. there is a diagonal entry of D_{sym} whose value is greater than 0.5. For example, In the case of $k = 2$ patches, there is only one SMV $\mathbf{u}(1) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ corresponding to complete spatial asynchrony. The eigenvalue associated with $\mathbf{u}(1)$ is $\mu_1 = 1 - 2d$ where d is the fraction of dispersing individuals.

As in the case of $k = 2$ patches, each source of variation in fitness (see (5)) and dispersal (see (7)) has a distinctive impact on the metapopulation growth rate r. To identify these impacts, I decompose

the spatial sources of variation along the SMVs $\mathbf{u}(1), \ldots, \mathbf{u}(k)$. For the spatial variation in fitness, this decomposition takes the form

$$\overline{\boldsymbol{\lambda}} - \overline{\langle \boldsymbol{\lambda} \rangle} = \sum_{i=1}^{k-1} \overline{\alpha}_i \mathbf{u}(i) \text{ where } \overline{\alpha}_i = \frac{1}{k} \sum_{j=1}^{k} (\overline{\lambda}_j - \overline{\langle \boldsymbol{\lambda} \rangle}) u_j(i).$$

For the spatial-temporal variation, the decomposition is

$$\boldsymbol{\lambda}_t - \langle \boldsymbol{\lambda}_t \rangle - \overline{\boldsymbol{\lambda}} + \overline{\langle \boldsymbol{\lambda} \rangle} = \sum_{i=1}^{k-1} \alpha_{i,t} \mathbf{u}(i) \text{ where } \alpha_{i,t} = \frac{1}{k} \sum_{j=1}^{k} \left(\lambda_{i,t} - \langle \boldsymbol{\lambda}_t \rangle - \overline{\lambda_i} + \overline{\langle \boldsymbol{\lambda} \rangle} \right) u_j(i).$$

For dispersal variation, only net changes in movement for each patch impact the metapopulation growth rate

$$\mathbf{m}_t = \text{ row sums of } D_t - D_{\text{sym}}.$$

As the column sum of $\mathbf{m}_t$ always equals zero, one can express this vector as a unique linear combination of the spatial modes of variation:

$$\mathbf{m}_t = \sum_{i=1}^{k-1} \overline{\beta}_i \mathbf{u}(i) + \sum_{i=1}^{k-1} \beta_{i,t} \mathbf{u}(i) \text{ where } \overline{\beta}_i = \frac{1}{k} \sum_{j=1}^{k} \overline{m}_j u_j(i) \text{ and } \beta_{i,t} = \frac{1}{k} \sum_{j=1}^{k} (m_{j,t} - \overline{m}_j) u_j(i).$$

As in the two patch case, temporal variation $\sigma^2_{\text{time}} = \overline{\left(\langle \boldsymbol{\lambda}_t \rangle - \overline{\langle \boldsymbol{\lambda} \rangle} \right)^2}$ in demography reduces the metapopulation growth rate. The effect of variation along a spatial mode depends on its variance $\sigma^2_{\text{space},i} = \overline{\alpha}_i^2$ and the eigenvalue μ_i associated with this mode. In particular, variation along a spatial mode enhances the metapopulation growth rate only when movement along this spatial mode is philopatric i.e. $\mu_i > 0$. The effect of spatial-temporal variation in demography depends on its variation $\sigma^2_{\text{st},i} = \text{Var}[\alpha^2_{\text{i,t}}]$ along a spatial mode and the associated temporal autocorrelation $\rho_{\text{st},i} = \text{Corr}[\alpha_{\text{i,t}}, \alpha_{\text{i,t+1}}]$. For philopatric spatial modes of variation ($\mu_i > 0$), positive autocorrelations enhance metapopulation growth rates, while negative autocorrelations reduce r. For miseopatric spatial modes of variation, the direction of these impacts is reversed. The impact of spatial covariation in demography and dispersal on r depends on the dispersal-demography covariance $\text{cov}_{D\Lambda\text{-space},i}$; positive covariances along a spatial mode increase r. Similarly, positive spatial-temporal covariances $\text{cov}_{D\Lambda\text{-st},i}$ or autocorrelations $\rho_{D\Lambda\text{-st},\text{i}}$ also increase the metapopulation growth rate. Putting all of these contributions

together yield the following decomposition of the metapopulation growth rate:

$$
\begin{aligned}
r \approx \ln \overline{\langle \boldsymbol{\lambda} \rangle} & \underbrace{-\frac{1}{2}\mathrm{CV}^2_{\text{time}}}_{\substack{\text{temporal}\\ \text{deviations}}} + \underbrace{\sum_{i=1}^{k-1} \frac{\mu_i}{1-\mu_i}\mathrm{CV}^2_{\text{space},i}}_{\substack{\text{spatial}\\ \text{deviations}}} + \underbrace{\sum_{i=1}^{k-1} \frac{\rho_{\text{st},i}\mu_i}{1-\rho_{\text{st},i}\mu_i}\mathrm{CV}^2_{\text{st},i}}_{\substack{\text{spatial-temporal}\\ \text{deviations}}} \\
& + \underbrace{\sum_{i=1}^{k-1} \frac{1}{1-\mu_i}\frac{\mathrm{cov}_{D\Lambda\text{-space},i}}{\overline{\langle \boldsymbol{\lambda} \rangle}^2}}_{\substack{\text{spatial}\\ \text{covariance}}} + \underbrace{\sum_{i=1}^{k-1} \frac{\mu_i}{1-\rho_{D\Lambda,i}\mu_i}\frac{\mathrm{cov}_{D\Lambda\text{-st},i}}{\overline{\langle \boldsymbol{\lambda} \rangle}^2}}_{\substack{\text{spatial-temporal}\\ \text{covariance}}}
\end{aligned}
\tag{13}
$$

To illustrate the role of multiple spatial modes on population growth, Figures 4 and 5 consider a one-dimensional gradient of $k = 15$ patches with periodic boundary conditions. Individuals disperse preferentially to patches with higher fitness. The reference dispersal matrix D_{sym} corresponds to dispersal in the absence of preferences; half of the individuals in a patch move to the left, and the other half to the right. The SMVs of D_{sym} correspond to different spatial frequencies

$$
\mathbf{u}_i = \sqrt{2k}\left(1 \quad \cos\left(\tfrac{2\pi i}{k} - \tfrac{\pi}{4}\right) \quad \cos\left(\tfrac{4\pi i}{k} - \tfrac{\pi}{4}\right) \quad \ldots \quad \cos\left(\tfrac{2(k-1)\pi i}{k} - \tfrac{\pi}{4}\right)\right) \text{ for } 1 \le i \le k-1
$$

with the associated eigenvalue μ_i equal to $\cos\left(\frac{2\pi i}{k}\right)$ [Berlin and Kac, 1952]. Importantly, movement along modes corresponding to higher frequencies (i.e. $\frac{k}{8\pi} < i < \frac{3k}{8\pi}$) is miseopatric (i.e. $\mu_i < 0$) and otherwise philopatric. It follows that lower versus higher frequencies of spatial variation have opposing impacts on population growth rate (Fig. 4 versus Fig. 5).

When spatial variation always occurs at lower frequencies (Fig. 4), this variation increases r; $r = 0.0034 > 0$ from both the analytical formula and numerical approximation despite $\ln \overline{\langle \boldsymbol{\lambda} \rangle} < 0$ (sum of colored bars versus gray bar in Fig. 4B). While temporal variation reduces the metapopulation growth rate slightly, spatial and spatial-temporal variation in demography and dispersal more than compensate (colored bars in Fig. 4B). These positive effects stem from the spatial modes corresponding to the low frequencies (thickest curves in Fig. 4C).

When spatial variation occurs at higher frequencies (Fig. 5A), it reduces the metapopulation growth rate ($r = -0.025$ from analytical formula, $r = -0.018$ from numerical approximation) (Fig. 5B). This negative impact stems from the dominant spatial modes of variation having higher frequencies (thick orange and green curves in Fig. 5C) and, consequently, being miseopatric. Hence, spatial and spatial-temporal variation in demography have negative impacts on metapopulation growth rates (orange and yellow bars in Fig. 5B). These negative impacts exceed the slightly positive impacts of spatial-temporal covariation in dispersal and demography (green bar in Fig. 5B).

Relationship to the fitness-density covariance. As shown by equation (9), the metapopulation growth rate r is determined by the temporal variation of the spatial averages $\langle \boldsymbol{\lambda}_t \rangle$ and the temporal variation in the fitness-density covariances $\mathrm{cov}_{\text{space}}(\boldsymbol{\lambda}_{\text{t}}, \hat{\mathbf{v}}_{\text{t}})$. When spatial and temporal variation is

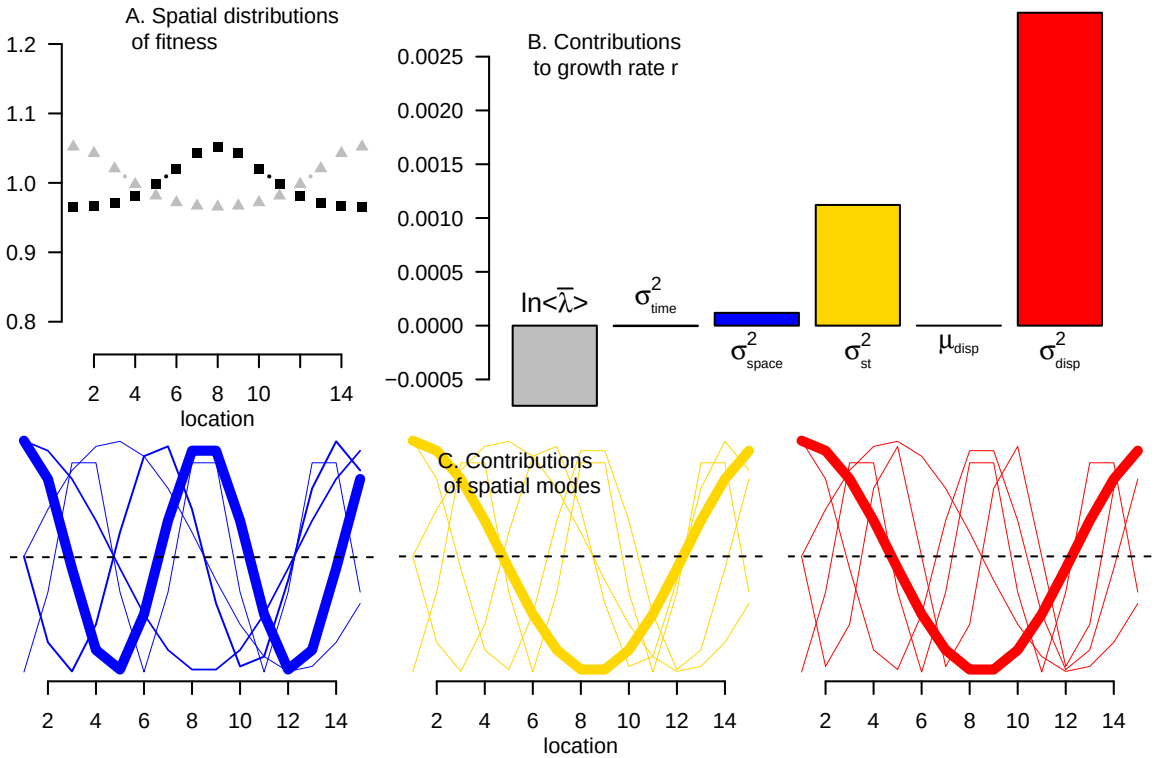

FIGURE 4. The impact of low frequency spatial variation in fitness on the metapopulation growth rate. A population living along a spatial gradient of $k = 15$ patches that varies over time between higher fitness in the center (black curve in A) and higher fitness at the edges (gray curve in A). Individuals disperse to neighboring patches with a preference for patches supporting a higher fitness. Analytical decomposition of the population growth rate in B. In C, all spatial modes of variation (SMV) plotted with the thickness of the curves proportional to the magnitude of their contribution to the metapopulation growth rate.

small, the second-order Taylor approximation of $\ln(\overline{\langle\boldsymbol{\lambda}\rangle} + x) \approx \ln\overline{\langle\boldsymbol{\lambda}\rangle} + \frac{x}{\overline{\langle\boldsymbol{\lambda}\rangle}} - \frac{x^2}{2\overline{\langle\boldsymbol{\lambda}\rangle}^2}$ implies that the metapopulation growth rate satisfies

$$r \approx \ln\overline{\langle\boldsymbol{\lambda}\rangle} + \frac{k}{\overline{\langle\boldsymbol{\lambda}\rangle}}\overline{\mathrm{cov}(\boldsymbol{\lambda}_\mathrm{t}, \mathbf{v}_\mathrm{t})} - \frac{1}{2}\mathrm{CV}^2_{\mathrm{time}}. \tag{14}$$

Hence, all but the first two terms of the metapopulation growth rate decomposition (13) correspond to the time-averaged fitness-density covariance. Positive values of the remaining terms (due to spatial or

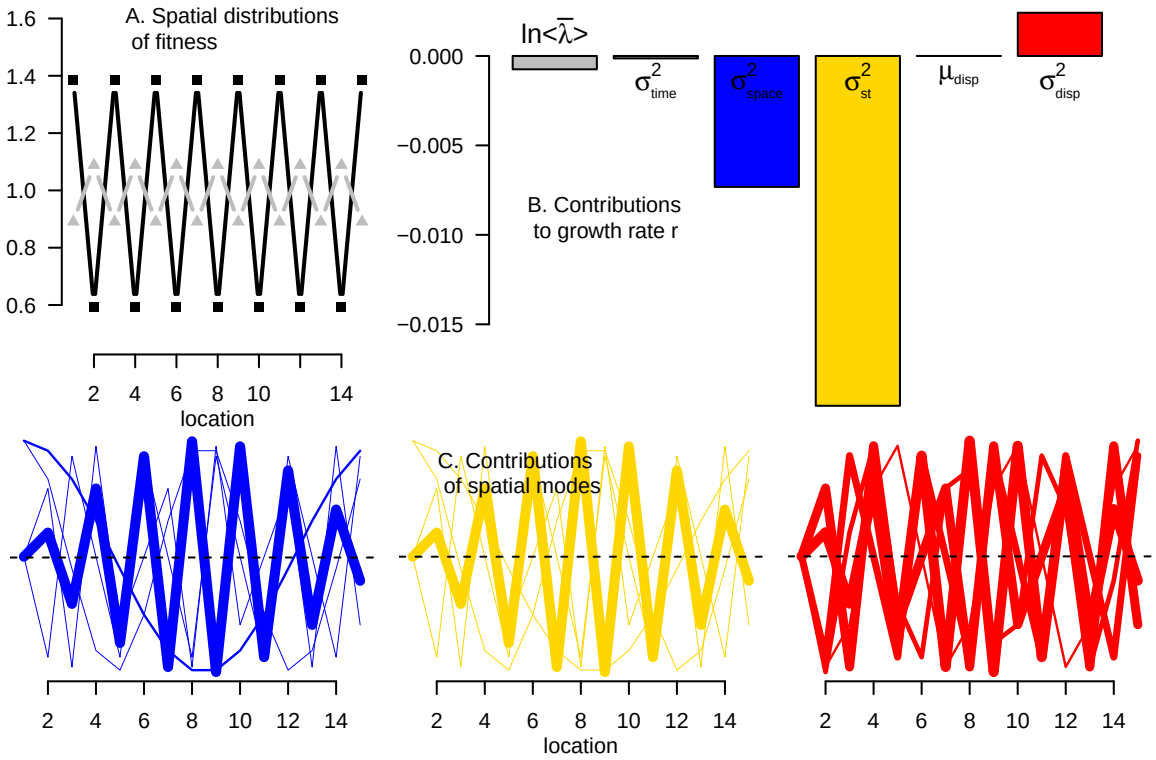

FIGURE 5. The impact of high frequency spatial variation in fitness on the metapopulation growth rate. A population living along a spatial gradient of $k = 15$ patches that varies over time between higher fitness occurring at odd number patches (black curve in A) and even numbered patches (gray curve in A). Individuals disperse to neighboring patches with a preference for patches supporting higher fitness. Analytical decomposition of the population growth rate in B. In C, all spatial modes of variation (SMV) plotted with the thickness of the curves proportional to the magnitude of their contribution to the metapopulation growth rate.

spatial-temporal variation in fitness, or spatial or spatial-temporal covariation in fitness and dispersal) contribute to a positive fitness-density covariance.

DISCUSSION

Spatial and temporal variations in biotic and abiotic conditions are a ubiquitous feature of the natural world. This variation, via its effects on fitness and dispersal, modulates how quickly metapopulations grow or decline in abundance. Here, I presented a framework to explicitly quantify the contributions of each of these sources of variation to the metapopulation growth rate. Aside from temporal variation in

the spatially averaged fitness, all other sources of variation determine how population densities covary with patch quality – the fitness-density covariance of Chesson [2000]. When this covariance is positive, individuals tend concentrate in patches supporting higher fitness resulting in a higher metapopulation growth rate. This framework recovers classical results about the opposing effects of spatial and temporal variation on population growth, spatial bet-hedging by reducing temporal variation, and inflationary effects due to temporally auto-correlated variation. Furthermore, this framework highlights new mechanisms yielding inflationary effects and positive fitness-density covariances, identifies which sources of variation have the largest impact on population growth, and provides new opportunities and challenges for future theoretical and empirical work.

Temporal versus spatial variation

A well-known dichotomy in demography is that temporal variation in fitness reduces population growth rates [Lewontin and Cohen, 1969, Boyce et al., 2006], while spatial variation can increase metapopulation growth rates [Karlin, 1976, Hastings, 1983, Snyder and Chesson, 2003, Schreiber and Lloyd-Smith, 2009, Crone, 2016]. For spatially and temporally variable environments, this dichotomy remains if temporal variation corresponds to the temporal variation in the spatially averaged fitnesses, and spatial variation corresponds to the spatial variation of the temporally averaged fitnesses. The decomposition of the metapopulation growth rate in equation (13) implies that temporal variation in spatial averages always reduces the metapopulation growth rate, and identities when spatial variation in temporal averages increase the metapopulation growth rate. The positive effect of spatial variation only occurs if the population is philopatric i.e. no more than 50% of individuals leave a patch in a time step; otherwise spatial variation has a negative impact. For these philopatric populations, the positive effect stems from individuals accumulating more quickly in the patches with higher fitness resulting in a positive fitness-density covariance [Chesson, 2000, Snyder, 2007]. For miseopatric populations where at least half of the individuals are leaving some patches, spatial variation reduces metapopulation growth rates. The simplest illustration of this negative impact occurs in a population dispersing between two patches with unequal fitnesses $\lambda_1 \neq \lambda_2$. If individuals mostly disperse to the other patch each time step, then the population increases by a factor of $\lambda_1\lambda_2$ over two time steps. Hence, the metapopulation growth rate equals the logarithm of the geometric mean $\sqrt{\lambda_1\lambda_2}$. As the geometric mean is strictly smaller than the spatially averaged fitness $\frac{\lambda_1+\lambda_2}{2}$, spatial variation in fitness reduces this miseopatric population's growth rate. More generally, the metapopulation growth rate decomposition in equation (13) implies that the positive impact of spatial variation decreases continually with increasing dispersal rates i.e. the $\frac{1-2d}{2d}$ term in (10). For models without temporal variation, this prediction follows from Karlin's reduction principle [Karlin, 1976, Hastings, 1983, Kirkland et al., 2006, Altenberg, 2012a].

The decomposition of the metapopulation growth rate clarifies under what conditions spatial variation has a larger impact on metapopulation growth rates than an equivalent amount of temporal variation i.e. $\sigma_{\text{time}} = \sigma_{\text{space}}$. This occurs whenever dispersal rates are sufficiently low (e.g. $d < \frac{1}{3}$ for two patch systems), otherwise temporal variation has a larger impact. This theoretical prediction is consistent with a demographic analysis of the relatively sedentary, perennial prairie forb *Pulsatilla patens* [Crone, 2016]. Using data from a 10-year demographic monitoring study, Crone [2016] analyzed

two versions of an integral projection model: a spatially explicit model accounting for between-site variation in the vital rates and a stochastic model accounting for between-year variation in the vital rates. Despite spatial variation in vital rates being smaller than temporal variation in the vital rates, Crone [2016] found spatial variation had a larger effect (+0.006) on the population growth rate than temporal variation (−0.002).

Spatial buffering and inflationary effects

Spatial-temporal variation, here, corresponds to the residual variation in fitness beyond spatial variation and temporal variation. This definition of spatial-variation coincides with recent work by Johnson and Hastings [2023] and Kortessis et al. [2023]. Spatial-temporal variation always corresponds to complete spatial asynchrony in fitness i.e. all spatial modes of variation have entries of opposite sign. By reducing the amount of temporal variation (in the spatially averaged fitnesses), this spatial asynchrony buffers population growth against the adverse effects of temporal variation – a classical result in metapopulation theory [Hanski, 1999]. This buffering effect is most easily understood for a two patch system with each patch exhibiting the same distribution of fitnesses over time. If the two patches are perfectly spatially synchronized, then fitness only varies over time, not space; there is a reduction of the metapopulation growth rate. If the two patches are always completely, spatially asynchronous, then fitness only varies over space, not time. When the spatial-temporal variation is temporally uncorrelated, this spatial asynchrony buffers the metapopulation growth rate by shifting all of the temporal variation to spatial-temporal variation. This well-known buffering effect can select for dispersal [Kuno, 1981, Metz et al., 1983, Levin et al., 1984], lead to evolutionarily stable sink populations [Holt, 1997, Jansen and Yoshimura, 1998, Schreiber, 2012], and rescue metapopulations from extinction [Jansen and Yoshimura, 1998, Schreiber, 2012, Evans et al., 2013]. This buffering effect, however, does not generate a fitness-density covariance.

Beyond its buffering effect, spatial-temporal variation in fitness can inflate metapopulation growth rates. Consistent with earlier work [Roy et al., 2005, Schreiber, 2010], positive autocorrelations in spatial-temporal variation increase the metapopulation growth rates of philopatric populations. Positive autocorrelation, here, corresponds to one configuration of spatial asynchrony (e.g. fitness in patch 1 being greater than fitness in patch 2) being maintained for several time steps before shifting to a different configuration of asynchrony (e.g. fitness in patch 2 being greater than fitness in patch 1). This positive effect however, is always less than, the positive effect due to a comparable amount of spatial variation (e.g. $\frac{1-2d}{d} > \frac{\rho_{\mathrm{st}}(1-2d)}{1-\rho_{\mathrm{st}}(1-2d)}$ for $d < 1/2$ and $\rho_{\mathrm{st}} > 0$), vanishes in well-mixed populations [Schreiber, 2010] and becomes negative when populations are miseopatric.

An inflationary effect also occurs for miseopatric populations when spatial-temporal variation is negatively autocorrelated. Intuitively, if patches supporting high fitness tend to have low fitness in the next time step, moving to other patches results in the metapopulation continuing to concentrate in patches with higher fitness – a positive fitness-density covariance [Chesson, 2000]. This inflationary effect likely occurs in migratory bird species dispersing seasonally between breeding grounds and wintering areas [Dingle and Drake, 2007]. Similar to the scenario considered in Figure 2, fitnesses in

breeding grounds (patch 1) often vary between higher values in the summer and lower values in the winter [Lundberg, 1988, Shaw and Levin, 2011, Dingle, 2014]. In contrast, fitness in wintering areas (patch 2) are higher than the breeding grounds in the winter, but lower in the summer. If a time step in the metapopulation model is a season, then the seasonal fluctuations in fitnesses generates a negative spatial-temporal autocorrelation. If most individuals disperse each season ($d \approx 1$), then seasonal migration generates an inflationary effect. In Holartic regions with less severe winter conditions [Lundberg, 1988], the breeding ground may support higher fitness, on average, than the wintering areas. This spatial variation can result in populations of sedentary individuals ($d \approx 0$) and populations migratory individuals ($d \approx 1$) having locally optimal metapopulation growth rates, but populations of individuals partially dispersing ($0 < d < 1$) having low metapopulation growth rates (Fig. 2A). Hence, this combination of spatial variation and negatively auto-correlated spatial-temporal variation may contribute to the evolution of partial migration [Lundberg, 1988, Taylor and Norris, 2007, Shaw and Levin, 2011, Lundberg, 2013, De Leenheer et al., 2017] where one subpopulation always remains in the breeding grounds ($d = 0$) and the rest of the population migrates ($d = 1$).

The metapopulation growth decomposition in equation (13) also highlights how spatial-temporal variation in dispersal generates inflationary effects. This occurs when spatial-temporal variation in dispersal positively covaries with spatial-temporal variation in fitness i.e. the final term in (13). Intuitively, a tendency of individuals to track patches that will support higher fitness increases the metapopulation growth rate by creating a positive fitness-density covariance. Northern pike (*Esox lucius*) in Lake Windermere exhibit this type of positive covariance [Haugen et al., 2006]. Over a 40 year period, year-to-year movement between the southern and northern basins were positively correlated to the fitnesses in the two basins. One behavioral strategy that produces this positive fitness-dispersal covariance is the win-stay, lose-shift (WSLS) strategy. Individuals playing this strategy only leave patches after experiencing low fitness. The metapopulation growth rate decomposition (13) implies that this strategy generates a positive covariance only when the spatial-temporal variation in fitness is positively autocorrelated, a finding consistent with earlier simulation studies [Schmidt, 2004]. However, in negatively autocorrelated environments, the WSLS can enhance metapopulation growth rates provided there is pure spatial variation in fitness. WSLS can lead to individuals, on average, spending more time in these higher quality patches and, thereby, still produce a positive fitness-density covariance.

Challenges and opportunities: Theoretical and empirical

The decomposition of the metapopulation growth rate offers a multitude of challenges and opportunities for future theoretical and empirical work. From the theoretical side, I considered environmental stochasticity, but not demographic stochasticity. In the face of demographic stochasticity and negative density-dependence, extinction in finite time is inevitable [Schreiber, 2017]. Hence, the question is not whether the population goes extinct, but how soon does it go extinct? Recently, Prodhomme and Strickler [2021] showed that a positive metapopulation growth rate r ensures that the time to extinction increases like a power of the total habitat size A i.e. A^θ for some $\theta > 0$. In contrast, a negative metapopulation growth rate r implies that the time to extinction only increases logarithmically with the habitat size i.e. $\log A$. Hence, in large habitats, a positive metapopulation growth

rate r implies long persistence times, but not otherwise. Despite this fundamental dichotomy, more positive metapopulation growth rates need not, in general, imply longer persistence times or higher establishment probabilities [Schreiber and Lloyd-Smith, 2009, Pande et al., 2020a, Ellner et al., 2020, Pande et al., 2020b]. This issue stems from the metapopulation growth rate r only describing the mean rate of increase and not variation around this mean trend. For non-spatial models, Pande et al. [2022] introduced new metrics that account for this variation and do a much better job of predicting establishment success. To what extent these or related metrics can be used to identify the relative contributions of spatial and temporal variation to persistence times remains to be seen.

The results presented here offer many empirically testable hypotheses. Three of these hypotheses concern two patch systems for which patch-specific fitnesses and dispersal rates can be manipulated e.g. [Matthews and Gonzalez, 2007]. When one patch consistently has consistently higher fitness than the other patch, our theory predicts that the metapopulation growth rate at low dispersal (i.e. $d < 0.5$) will be higher than the growth rate for a well-mixed population, but the growth rate at high dispersal ($d > 0.5$) will be lower than the metapopulation growth rate for a well-mixed population. For the well-mixed population, one could either manipulate dispersal to achieve individuals experiencing average fitness conditions (i.e. $d = 0.5$) or manipulate the fitness in the patches to correspond to the average fitness of the spatially variable treatments. The second hypothesis concerns negative autocorrelation in space and time; one patch alternates between high and low fitness, while the other patch alternates between low and high fitness. Theory predicts that metapopulation growth rate at low dispersal rates will be higher than the growth rate for a well-mixed population, but the metapopulation growth rate at high dispersal rates will be higher than the growth rate for a well-mixed population; this corresponds to direct opposite of the first hypothesis. The third hypothesis concerns combining the features of the first two hypothesis; the system exhibits spatial variation and negatively-autocorrelated spatial-temporal variation. There are several ways to combine these features. For example, one patch could have, on average, the higher fitness ($\overline{\lambda_1} > \overline{\lambda_2}$) and the patches exhibit negatively autocorrelated fluctuations around these averages ($\lambda_{1,t} = \overline{\lambda_1} \pm \Delta\lambda$, $\lambda_{2,t} = \overline{\lambda_2} \mp \Delta\lambda$). Alternatively, one could create a variable source patch and constant sink patch as described in Figure 2. Theory predicts that the metapopulation growth rates for low and high dispersal will be higher than the growth rate of a well-mixed population; this hypothesis correspond to one part of each of the first two hypotheses. Beyond these scenarios with constant dispersal, it would be exciting to test these hypothesis for different forms of variable dispersal e.g. win-stay,lose-shift dispersal versus lose-stay,win-shift dispersal.

The decomposition of the metapopulation growth rate is conceptually similar (but simpler) to the low-density growth rate decompositions in Chesson's quantitative coexistence theory [Chesson, 1994, 2000, 2018]. Applications of Chesson's theory to empirical systems have allowed ecologist to answer "why do these species coexist?" by evaluating the relative importance of alternative coexistence mechanisms [Adler et al., 2010, Ellner et al., 2016, 2018, Spaak and Schreiber, 2023]. Similarly, applying the metapopulation growth rate decomposition (13) to empirical systems, one could answer "why is this metapopulation growing (or declining)?" by evaluating the relative importance of different sources of variation in demography and dispersal. For example, is a metapopulation growing mostly due to uniformly high quality patches or individuals actively moving to higher quality patches? Is a

metapopulation declining primarily due to temporal variation in fitness or maladaptive movement to lower quality habitats?

To answer these questions for a specific system, one needs to parameterize a spatially explicit model similar to (2). If the model involves only 2 patches, the symmetric dispersal matrix D_{sym} is given by (6) with $d = \overline{\frac{d_{12}+d_{21}}{2}}$. If the model involves more than two patches, there is no canonical choice of D_{sym}, but a natural option, when possible, is using $\overline{\frac{d_{ij}+d_{ji}}{2}}$ for the off-diagonal elements. Using the model, there are two approaches to doing the decomposition of the metapopulation growth rate. The computationally less expensive approach is akin to using of the second order approximations formulas of Tuljapurkar [1990] to estimate the population growth rate [Morris and Doak, 2002, Crone et al., 2013, Crone, 2016]. In our setting, one would simulate (or directly use empirical) time series of the fitnesses $\lambda_{i,t}$ and the dispersal rates $d_{ij,t}$, decompose these time series with respect to the spatial modes variation (as described in the methods), compute the appropriate variances and covariances, and directly evaluate the decomposition of the metapopulation growth rate (13).While there is precedent for this approach in stochastic demography [Morris and Doak, 2002, Boyce et al., 2006, Crone et al., 2013, Crone, 2016], it may produce spurious results if some sources of variation are not sufficiently small.

To deal with larger sources variation in dispersal and demography, one can follow the approach of Ellner et al. [2020] for Chesson's quantitative coexistence theory. This involves computationally more intensive calculations using a multi-step, hierarchical procedure and is based on the fitness-density covariance representation of r in equation (9). First, one partitions the metapopulation growth rate r into three terms: growth $\ln \overline{\langle \boldsymbol{\lambda} \rangle}$ due to average fitness, a reduction Δr_{time} due to temporal fluctuations, and the remaining growth Δr_{FD} due to fitness-density covariance:

$$
\begin{aligned}
r &= \ln \overline{\langle \boldsymbol{\lambda} \rangle} + \Delta r_{\text{time}} + \Delta r_{\text{FD}} \quad \text{where} \\
\Delta r_{\text{time}} &= \overline{\ln \langle \boldsymbol{\lambda}_t \rangle} - \ln \overline{\langle \boldsymbol{\lambda} \rangle} < 0 \text{ and } \Delta r_{\text{FD}} = r - \Delta r_{\text{time}} - \ln \overline{\langle \boldsymbol{\lambda} \rangle}
\end{aligned} \tag{15}
$$

where r is approximated numerically as described in Models and Methods. The next step would partition the growth Δr_{FD} due to the fitness-density covariance into contributions due spatial and spatial-temporal variation in demography and dispersal. For example, if there is only variation in demography, Δr_{FD} decomposes into growth $\Delta r_{\text{FD,space}}$ due to spatial variation and growth $\Delta r_{\text{FD,space-time}}$ due to spatial-temporal variation:

$$
\begin{aligned}
\Delta r_{\text{FD}} &= \Delta r_{\text{FD,space}} + \Delta r_{\text{FD,space-time}} \quad \text{where} \\
\Delta r_{\text{FD,space}} &= \ln \lambda_{\text{space}} - \ln \overline{\langle \boldsymbol{\lambda} \rangle} - \Delta r_{\text{time}} \text{ and } \Delta r_{\text{FD,space-time}} = \Delta r_{\text{FD}} - \Delta r_{\text{FD, space}}
\end{aligned} \tag{16}
$$

where λ_{space} is the dominant eigenvalue of the temporally averaged matrix $\overline{D\Lambda}$.

Applying one or both of these approaches to empirically based models could provide insights into what sources of variation have the greatest impacts the growth rates of real populations. Moreover, identifying the most important sources of variation for population growth may be useful for conservation and management efforts.

References

P.B. Adler, S.P. Ellner, and J.M. Levine. Coexistence of perennial plants: an embarrassment of niches. *Ecology letters*, 13:1019–1029, 2010.

L Altenberg. Resolvent positive linear operators exhibit the reduction phenomenon. *Proceedings of the National Academy of Sciences*, 109:3705–3710, 2012a.

L. Altenberg. The evolution of dispersal in random environments and the principle of partial control. *Ecological Monographs*, 82:297–333, 2012b.

K. B. Athreya and P. E. Ney. *Branching processes.* Dover Publications Inc., Mineola, NY, 2004.

J. Battin. When good animals love bad habitats: ecological traps and the conservation of animal populations. *Conservation Biology*, 18:1482–1491, 2004.

M. Benaïm and S. J. Schreiber. Persistence of structured populations in random environments. *Theoretical Population Biology*, 76:19–34, 2009.

M. Benaïm and S.J. Schreiber. Persistence and extinction for stochastic ecological models with internal and external variables. *Journal of Mathematical Biology*, 79:393–431, 2019.

T.H. Berlin and M. Kac. The spherical model of a ferromagnet. *Physical Review*, 86:821, 1952.

M. S. Boyce, C. V. Haridas, C. T. Lee, and the NCEAS Stochastic Demography Working Group. Demography in an increasingly variable world. *Trends in Ecology and Evolution*, 21:141–148, 2006.

R.S. Cantrell, C. Cosner, Y. Lou, and S.J. Schreiber. Evolution of natal dispersal in spatially heterogenous environments. *Mathematical Biosciences*, 283:136–144, 2017.

G. Carroll, R. Harcourt, B.J. Pitcher, D. Slip, and I. Jonsen. Recent prey capture experience and dynamic habitat quality mediate short-term foraging site fidelity in a seabird. *Proceedings of the Royal Society B: Biological Sciences*, 285:20180788, 2018.

H. Caswell. *Matrix Population Models.* Sinauer, Sunderland, Massachuesetts, 2001.

K.A. Catalano, A.G. Dedrick, M.R. Stuart, J.B. Puritz, H.R. Montes Jr, and M.L. Pinsky. Quantifying dispersal variability among nearshore marine populations. *Molecular Ecology*, 30:2366–2377, 2021.

P. Chesson. Multispecies competition in variable environments. *Theoretical Population Biology*, 45: 227–276, 1994.

P. Chesson. Updates on mechanisms of maintenance of species diversity. *Journal of Ecology*, 106: 1773–1794, 2018.

P.L. Chesson. General theory of competitive coexistence in spatially-varying environments. *Theoretical Population Biology*, 58:211–237, 2000.

F.R.K. Chung. *Spectral graph theory*, volume 92. American Mathematical Soc., 1997.

J. Clobert. *Dispersal ecology and evolution*. Oxford University Press, 2012.

E.E. Crone. Contrasting effects of spatial heterogeneity and environmental stochasticity on population dynamics of a perennial wildflower. *Journal of Ecology*, 104:281–291, 2016.

E.E. Crone, M.M. Ellis, W.F. Morris, A. Stanley, T. Bell, P. Bierzychudek, J. Ehrlan, T.N. Kaye, T.M. Knight, P. Lesica, G. Oostermeijer, P.F. Quintana-Ascencio, T. Ticktin, T. Valverde, J.L. Williams, D.F. Doak, R. Ganesan, A.S. Mcearchern, K.and Thorpe, and E.S. Menges. Ability of matrix models to explain the past and predict the future of plant populations. *Conservation Biology*, 27:968–978, 2013. doi: https://doi.org/10.1111/cobi.12049. URL `https://conbio.onlinelibrary.wiley.com/doi/abs/10.1111/cobi.12049`.

P. De Leenheer, A. Mohapatra, H.A. Ohms, D.A. Lytle, and J.M. Cushing. The puzzle of partial migration: adaptive dynamics and evolutionary game theory perspectives. *Journal of Theoretical Biology*, 412:172–185, 2017.

H. Dingle. *Migration: the biology of life on the move*. Oxford University Press, USA, 2014.

H. Dingle and V.A. Drake. What is migration? *Bioscience*, 57:113–121, 2007.

S. P. Ellner, R. E. Snyder, P. B. Adler, and G. Hooker. An expanded modern coexistence theory for empirical applications. *Ecology Letters*, 22:3–18, 2018. doi: 10.1111/ele.13159.

S.P. Ellner and S.J. Schreiber. Temporally variable dispersal and demography can accelerate the spread of invading species. *Theoretical Population Biology*, 82:283–298, 2012.

S.P. Ellner, R.E. Snyder, and P.B. Adler. How to quantify the temporal storage effect using simulations instead of math. *Ecology Letters*, 19:1333–1342, 2016.

S.P. Ellner, R.E. Snyder, P..B. Adler, G. Hooker, and S.J. Schreiber. Technical Comment on Pande et al.(2020): Why invasion analysis is important for understanding coexistence. *Ecology Letters*, 23: 1721–1724, 2020.

S. N. Evans, P. Ralph, S. J. Schreiber, and A. Sen. Stochastic growth rates in spatio-temporal heterogeneous environments. *Journal of Mathematical Biology*, 66:423–476, 2013.

R. Hale and S.E. Swearer. Ecological traps: current evidence and future directions. *Proceedings of the Royal Society B: Biological Sciences*, 283(1824):20152647, 2016.

I. Hanski. *Metapopulation Ecology*. Oxford Series in Ecology and Evolution. Oxford University Press, 1999.

A. Hastings. Can spatial variation alone lead to selection for dispersal? *Theoretical Population Biology*, 24:244–251, 1983.

T.O. Haugen, I.J. Winfield, L.A. Vollestad, J.M. Fletcher, J.B. James, and N.C. Stenseth. The ideal free pike: 50 years of fitness-maximizing dispersal in Windermere. *Proceedings of the Royal Society of London B: Biological Sciences*, 273:2917–2924, 2006.

R. D. Holt. On the evolutionary stability of sink populations. *Evolutionary Ecology*, 11:723–731, 1997.

R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1990. Corrected reprint of the 1985 original.

H. Jaggi, D. Steinsaltz, and S. Tuljapurkar. Migration in uncertain times. *bioRxiv*, 2023.

V. A. A. Jansen and J. Yoshimura. Populations can persist in an environment consisting of sink habitats only. *Proceeding of the National Academy of Sciences USA*, 95:3696–3698, 1998.

E.C. Johnson and A. Hastings. Coexistence in spatiotemporally fluctuating environments. *Theoretical Ecology*, pages 1–34, 2023.

S. Karlin. *Population genetics and ecology*, chapter Population subdivision and selection migration interaction, pages 616–657. Academic Press, 1976.

S. Kirkland, C.K. Li, and S. J. Schreiber. On the evolution of dispersal in patchy landscapes. *SIAM Journal on Applied Mathematics*, 66(4):1366–1382, 2006.

N. Kortessis, M.W. Simon, M. Barfield, G.E. Glass, B.H. Singer, and R.D. Holt. The interplay of movement and spatiotemporal variation in transmission degrades pandemic control. *Proceedings of the National Academy of Sciences*, 117:30104–30106, 2020.

N. Kortessis, G. Glass, A. Gonzalez, N.W. Ruktanonchai, M.W. Simon, B. Singer, and R.D. Holt. Neglected consequences of spatio-temporal heterogeneity and dispersal: Metapopulations, the inflationary effect, and real-world consequences for public health. 2023.

E. Kuno. Dispersal and the persistence of populations in unstable habitats: a theoretical note. *Oecologia*, 49:123–126, 1981.

S. A. Levin, D. Cohen, and A. Hastings. Dispersal strategies in patchy environments. *Theoretical Population Biology*, 26(2):165 – 191, 1984.

R. C. Lewontin and D. Cohen. On population growth in a randomly varying environment. *Proceedings of the National Academy of Sciences USA*, 62:1056—1060, 1969.

P. Lundberg. The evolution of partial migration in birds. *Trends in Ecology & Evolution*, 3:172–175, 1988.

P. Lundberg. On the evolutionary stability of partial migration. *Journal of Theoretical Biology*, 321: 36–39, 2013.

D.S. Mason, C. Baruzzi, and M.A. Lashley. Passive directed dispersal of plants by animals. *Biological Reviews*, 97:1908–1929, 2022.

D. P. Matthews and A. Gonzalez. The inflationary effects of environmental fluctuations ensure the persistence of sink metapopulations. *Ecology*, 88:2848–2856, 2007.

J. A. J. Metz, T. J. de Jong, and P. G. L. Klinkhamer. What are the advantages of dispersing; a paper by Kuno extended. *Oecologia*, 57:166–169, 1983.

William F Morris and Daniel F Doak. *Quantitative conservation biology: theory and practice of population viability analysis*. Sinauer Associates, 2002.

M.I. O'Connor, J.F. Bruno, S.D. Gaines, B.S. Halpern, S.E. Lester, B.P. Kinlan, and J. M. Weiss. Temperature control of larval dispersal and the implications for marine ecology, evolution, and conservation. *Proceedings of the National Academy of Sciences*, 104:1266–1271, 2007.

J. Pande, T. Fung, R. Chisholm, and N.M. Shnerb. Mean growth rate when rare is not a reliable metric for persistence of species. *Ecology letters*, 23:274–282, 2020a.

J Pande, T Fung, R Chisholm, and N.M. Shnerb. Invasion growth rate and its relevance to persistence: a response to Technical Comment by Ellner et al. *Ecology Letters*, 23:1725–1726, 2020b.

J. Pande, Y. Tsubery, and N.M. Shnerb. Quantifying invasibility. *Ecology Letters*, 25:1783–1794, 2022.

J.H. Peniston, G.A. Backus, M.L. Baskett, R.J. Fletcher, and R.D. Holt. Ecological and evolutionary consequences of temporal variation in dispersal. *Ecography*, page e06699, 2023.

A. Prodhomme and É. Strickler. Large population asymptotics for a multitype stochastic SIS epidemic model in randomly switched environment, 2021.

G. Roth and S.J. Schreiber. Persistence in fluctuating environments for interacting structured populations. *Journal of Mathematical Biology*, 69:1267–1317, 2014.

M. Roy, R.D. Holt, and M. Barfield. Temporal autocorrelation can enhance the persistence and abundance of metapopulations comprised of coupled sinks. *American Naturalist*, 166:246–261, 2005.

D. Ruelle. Analyticity properties of the characteristic exponents of random matrix products. *Advances in Mathematics*, 32(1):68–80, 1979. ISSN 0001-8708.

K.A. Schmidt. Site fidelity in temporally correlated environments enhances population persistence. *Ecology Letters*, 7:176—184, 2004.

S. J. Schreiber and J. O. Lloyd-Smith. Invasion dynamics in spatially heterogenous environments. *American Naturalist*, 174:490–505, 2009.

S.J. Schreiber. Interactive effects of temporal correlations, spatial heterogeneity, and dispersal on population persistence. *Proceedings of the Royal Society: Biological Sciences*, 277:1907–1914, 2010.

S.J. Schreiber. The evolution of patch selection in stochastic environments. *The American Naturalist*, 180:17–34, 2012.

S.J. Schreiber. Coexistence in the face of uncertainty. In *Recent Progress and Modern Challenges in Applied Mathematics, Modeling and Computational Science*, pages 349–384. Springer, 2017.

A.K. Shaw and S.A. Levin. To breed or not to breed: a model of partial migration. *Oikos*, 120: 1871–1879, 2011.

R.E. Snyder. Spatiotemporal population distributions and their implications for species coexistence in a variable environment. *Theoretical population biology*, 72(1):7–20, 2007.

R.E. Snyder and P.L. Chesson. Local dispersal can facilitate coexistence in the presence of permanent spatial heterogeneity. *Ecology Letters*, 6(4):301–309, 2003.

R.E. Snyder, C.B. Paris, and A.C. Vaz. How much do marine connectivity fluctuations matter? *The American Naturalist*, 184:523–530, 2014.

J.W. Spaak and S.J. Schreiber. Building modern coexistence theory from the ground up: the role of community assembly. *bioRxiv*, pages 2023–01, 2023.

D.A. Spielman. Spectral and algebraic graph theory, 2019. URL `http://cs-www.cs.yale.edu/homes/spielman/sagt/`.

E.M. Stein and R. Shakarchi. *Fourier analysis: an introduction*, volume 1. Princeton University Press, 2011.

C.M. Taylor and D.R. Norris. Predicting conditions for migration: effects of density dependence and habitat quality. *Biology letters*, 3:280–284, 2007.

S. Tuljapurkar. *Population Dynamics in Variable Environments*. Springer-Verlag, New York, 1990.

P. Wiener and S. Tuljapurkar. Migration in variable environments: Exploring life-history evolution using structured population models. *Journal of Theoretical Biology*, 166:75–90, 1994.

P.D. Williams and A. Hastings. Stochastic dispersal and population persistence in marine organisms. *The American Naturalist*, 182:271–282, 2013.

Appendix A. General Formulas for First and Second Order Derivatives of Lyapunov Exponents

This appendix derives formulas for first order and second derivatives of the top Lyapunov exponent of a random product of a stationary sequence of positive matrices. I follow the approach of Ruelle [1979] who proved under what conditions these derivatives exist and derive them directly using multivariate calculus. As Ruelle's approach is presented in a skew-product formulation that can obfuscate the heart of the calculations, I use more standard probabilistic notation. Moreover, the derivations below correct an error in Ruelle's formula for the second order derivatives, and provide explicit formulas for mixed second order derivatives not found in classical stochastic demography literature [Tuljapurkar, 1990].

To describe the general class of random matrix products, let the set Θ of all environmental states be a compact metric space and let $\theta_1, \theta_2, \theta_3, \ldots$ be an ergodic stationary sequence taking values in Θ. θ_t is the state of the environment at time t. Let $\mathbf{A}(\theta)$ be a non-negative, primitive matrix associated with environmental state θ. Assume that the mapping $\theta \mapsto \mathbf{A}(\theta)$ is continuous. $\mathbf{A}$ and θ_t defines a stationary sequence of matrices $A_1 = \mathbf{A}(\theta_1), A_2 = \mathbf{A}(\theta_2), \ldots$ and a corresponding matrix model

$$\mathbf{n}_{t+1} = A_t \mathbf{n}_t$$

Associated with this linear model is the Lyapunov exponent

$$r(\mathbf{A}) = \lim_{t\to\infty} \frac{1}{t} \mathbb{E}[\ln \|A_t \ldots A_1\|]$$

for which

$$\lim_{t\to\infty} \frac{1}{t} \ln \frac{\sum_i n_i(t)}{\sum_i n_i(0)} = r(\mathbf{A})$$

with probability one whenever the total initial population size is positive i.e. $\sum_i n_i(0) > 0$. Furthermore, for any choice of positive row and column vectors, $\mathbf{v}$ and $\mathbf{w}$, Ruelle [1979, Proposition 3.2] implies that

$$r(\mathbf{A}) = \lim_{t\to\infty} f_t(\mathbf{A}) \text{ where } f_t(\mathbf{A}) = \frac{1}{t} \ln \mathbf{v} A_t \ldots A_1 \mathbf{w}$$

Under suitable conditions, Ruelle [1979, Theorem 3.1] proved that $r(\mathbf{A})$ is an analytic function of $\mathbf{A}$. In particular, it is continuous and continuously differentiable with respect to the environment-demography mapping $\mathbf{A}$. Moreover, the first derivative at $\mathbf{A}$ in the $\mathbf{B}$ direction is

$$\frac{\partial r}{\partial \mathbf{B}}(\mathbf{A}) = \lim_{\varepsilon\to 0} \frac{r(\mathbf{A} + \varepsilon \mathbf{B}) - r(\mathbf{A})}{\varepsilon}$$

where $\theta \mapsto \mathbf{B}(\theta)$ is a continuous mapping into the space of $k \times k$ matrices. Ruelle proved that

$$\frac{\partial r}{\partial \mathbf{B}}(\mathbf{A}) = \lim_{t\to\infty} \frac{\partial f_t}{\partial \mathbf{B}}(\mathbf{A})$$

The second order derivative in the $\mathbf{B}$ and $\mathbf{C}$ directions is

$$\frac{\partial^2 r}{\partial \mathbf{B}\partial \mathbf{C}}(\mathbf{A}) = \lim_{\varepsilon \to 0} \frac{\frac{\partial r}{\partial \mathbf{B}}(\mathbf{A} + \varepsilon \mathbf{C}) - \frac{\partial r}{\partial \mathbf{B}}(\mathbf{A})}{\varepsilon}$$

where $\theta \mapsto \mathbf{B}(\theta), \theta \mapsto \mathbf{C}(\theta)$ are continuous mappings into the space of $k \times k$ matrices. Ruelle proved that

$$\frac{\partial^2 r}{\partial \mathbf{B}\partial \mathbf{C}}(\mathbf{A}) = \lim_{t \to \infty} \frac{\partial^2 f_t}{\partial \mathbf{B}\partial \mathbf{C}}(\mathbf{A})$$

To compute $\frac{\partial r}{\partial \mathbf{B}}(\mathbf{A})$, the chain rule and the product rule imply

$$\frac{\partial f_t}{\partial \mathbf{B}}(\mathbf{A}) = \frac{1}{t} \sum_{s=1}^{t} \mathbb{E}\left[\frac{\mathbf{v} A_t \dots A_{s+1} B_s A_{s-1} \dots A_1 \mathbf{w}}{\mathbf{v} A_t \dots A_1 \mathbf{w}}\right]$$

In the case of interest when $\mathbf{A}(\theta) = A$ for all $\theta \in \Theta$, $\mathbf{v}A = \lambda \mathbf{v}$, $A\mathbf{w} = \lambda \mathbf{w}$ and $\mathbf{v}\mathbf{w} = 1$, then

$$\frac{\partial r}{\partial \mathbf{B}}(\mathbf{A}) = \lim_{t \to \infty} \frac{1}{t\lambda} \sum_{s=1}^{t} \mathbb{E}\left[\mathbf{v} B_s \mathbf{w}\right] = \frac{1}{\lambda} \mathbb{E}[\mathbf{v} B_1 \mathbf{w}]$$

where the second equality follows from stationarity of B_t. Thus,

$$\frac{\partial r}{\partial \mathbf{B}}(\mathbf{A}) = \frac{1}{\lambda} \mathbb{E}[\mathbf{v} B_1 \mathbf{w}] \tag{A2}$$

To compute a second order derivative in the directions $\mathbf{B}$ and $\mathbf{C}$, the chain rule and product rule imply

$$\frac{\partial^2 f_t}{\partial \mathbf{B}\partial \mathbf{C}}(\mathbf{A}) =$$
$$\frac{1}{t} \sum_{s=2}^{t} \sum_{u=1}^{s-1} \mathbb{E}\left[\frac{\mathbf{v} A_t \dots A_{s+1} B_s A_{s-1} \dots A_{u+1} C_u A_{u-1} \dots A_1 \mathbf{w} + \mathbf{v} A_t \dots A_{s+1} C_s A_{s-1} \dots A_{u+1} B_u A_{u-1} \dots A_1 \mathbf{w}}{\mathbf{v} A_t \dots A_1 \mathbf{w}}\right]$$
$$- \frac{1}{t} \mathbb{E}\left[\frac{(\sum_{s=1}^{t} \mathbf{v} A_t \dots A_{s+1} B_s A_{s-1} \dots A_1 \mathbf{w})(\sum_{s=1}^{t} \mathbf{v} A_t \dots A_{s+1} C_s A_{s-1} \dots A_1 \mathbf{w})}{(\mathbf{v} A_t \dots A_1 \mathbf{w})^2}\right]$$

In the case of interest when $\mathbf{A}(\theta) = A$ for all $\theta \in \Theta$ and $\mathbf{v}A = \lambda \mathbf{v}$ and $A\mathbf{w} = \lambda \mathbf{w}$ and $\mathbf{v}\mathbf{w} = 1$, then

$$
\begin{aligned}
&\frac{\partial^2 f_t}{\partial \mathbf{B}\partial \mathbf{C}}(\mathbf{A}) = \\
&\frac{1}{t}\sum_{s=2}^{t}\sum_{u=1}^{s-1}\mathbb{E}\left[\frac{\mathbf{v}\lambda^{t-s}B_sA^{s-1-u}C_u\lambda^{u-1}\mathbf{w}+\mathbf{v}\lambda^{t-s}C_sA^{s-1-u}B_u\lambda^{u-1}\mathbf{w}}{\lambda^t}\right] \\
&-\frac{1}{t}\mathbb{E}\left[\frac{(\sum_{s=1}^{t}\mathbf{v}\lambda^{t-s}B_s\lambda^{s-1}\mathbf{w})(\sum_{s=1}^{t}\mathbf{v}\lambda^{t-s}C_s\lambda^{s-1}\mathbf{w})}{\lambda^{2t}}\right] \\
&=\underbrace{\frac{1}{t}\sum_{s=2}^{t}\sum_{u=1}^{s-1}\frac{1}{\lambda^{s-u+1}}\mathbb{E}\left[\mathbf{v}B_sA^{s-1-u}C_u\mathbf{w}+\mathbf{v}C_sA^{s-1-u}B_u\mathbf{w}\right]}_{=:\clubsuit} \\
&-\underbrace{\frac{1}{t\lambda^2}\mathbb{E}\left[\sum_{s=1}^{t}\mathbf{v}B_s\mathbf{w}\cdot\sum_{s=1}^{t}\mathbf{v}C_s\mathbf{w}\right]}_{=:\spadesuit}
\end{aligned}
$$

By stationarity,

$$
\spadesuit = \frac{1}{\lambda^2}\mathbb{E}\left[\mathbf{v}B_1\mathbf{w}\cdot\mathbf{v}C_1\mathbf{w}\right]+\sum_{s=2}^{t}\frac{t-s+1}{t\lambda^2}\mathbb{E}[\mathbf{v}B_s\mathbf{w}\cdot\mathbf{v}C_1\mathbf{w}+\mathbf{v}C_s\mathbf{w}\cdot\mathbf{v}B_1\mathbf{w}]
$$

By ergodicity,

$$
\lim_{t\to\infty}\spadesuit = \frac{1}{\lambda^2}\mathbb{E}\left[\mathbf{v}B_1\mathbf{w}\cdot\mathbf{v}C_1\mathbf{w}\right]+\frac{1}{\lambda^2}\sum_{s=2}^{\infty}\mathbb{E}[\mathbf{v}B_s\mathbf{w}\cdot\mathbf{v}C_1\mathbf{w}+\mathbf{v}C_s\mathbf{w}\cdot\mathbf{v}B_1\mathbf{w}] \tag{A3}
$$

Now for $\clubsuit$. Reordering the sum yields

$$
\clubsuit = \frac{1}{t}\sum_{u=1}^{t-1}\sum_{s=u+1}^{t}\frac{1}{\lambda^{s-u+1}}\mathbb{E}\left[\mathbf{v}B_sA^{s-1-u}C_u\mathbf{w}+\mathbf{v}C_sA^{s-1-u}B_u\mathbf{w}\right]
$$

Re-indexing with $i = s-u+1$ yields

$$
\clubsuit = \frac{1}{t}\sum_{u=1}^{t-1}\sum_{i=2}^{t-u+1}\frac{1}{\lambda^i}\mathbb{E}\left[\mathbf{v}B_{u+i-1}A^{i-2}C_u\mathbf{w}+\mathbf{v}C_{u+i-1}A^{i-2}B_u\mathbf{w}\right]
$$

By stationarity,

$$
\clubsuit = \frac{1}{t}\sum_{i=2}^{t}\frac{t+i-1}{\lambda^i}\mathbb{E}\left[\mathbf{v}B_iA^{i-2}C_1\mathbf{w}+\mathbf{v}C_iA^{i-2}B_1\mathbf{w}\right]
$$

By ergodicity

$$\lim_{t\to\infty} \clubsuit = \sum_{i=2}^{\infty} \frac{1}{\lambda^i} \mathbb{E}\left[\mathbf{v}B_i A^{i-2} C_1 \mathbf{w} + \mathbf{v}C_i A^{i-2} B_1 \mathbf{w}\right] \tag{A4}$$

Hence, (A3) and (A4) imply
(A5)

$$\frac{\partial^2 r}{\partial \mathbf{B}\partial \mathbf{C}}(\mathbf{A}) = \lim_{t\to\infty} \frac{\partial^2 f_t}{\partial \mathbf{B}\partial \mathbf{C}}(\mathbf{A}) =$$
$$\sum_{t=2}^{\infty} \frac{1}{\lambda^t} \mathbf{v}\mathbb{E}\left[B_t A^{t-2} C_1 + C_t A^{t-2} B_1\right] \mathbf{w} - \frac{1}{\lambda^2}\mathbb{E}\left[\mathbf{v}B_1\mathbf{w}\cdot\mathbf{v}C_1\mathbf{w}\right] - \frac{1}{\lambda^2}\sum_{t=2}^{\infty}\mathbb{E}[\mathbf{v}B_t\mathbf{w}\cdot\mathbf{v}C_1\mathbf{w} + \mathbf{v}C_t\mathbf{w}\cdot\mathbf{v}B_1\mathbf{w}]$$

In the special case $\mathbf{C} = \mathbf{B}$, one has

$$\frac{\partial^2 r}{\partial \mathbf{B}\partial \mathbf{C}}((\mathbf{A}) = \sum_{t=2}^{\infty} \frac{2}{\lambda^t} \mathbf{v}\mathbb{E}\left[B_t A^{t-2} B_1\right] \mathbf{w} - \frac{1}{\lambda^2}\mathbb{E}\left[(\mathbf{v}B_1\mathbf{w})^2\right] - \frac{2}{\lambda^2}\sum_{t=2}^{\infty}\mathbb{E}[\mathbf{v}B_t\mathbf{w}\cdot\mathbf{v}B_1\mathbf{w}] \tag{A6}$$

Appendix B. The Metapopulation Growth Rate Approximation

In this Appendix, I derive the metapopulation growth rate decomposition (13) using the first and second order derivative formulas from Appendix A. Recall, that the metapopulation model is

$$\mathbf{n}_{t+1} = D_t \Lambda_t \mathbf{n}_t \tag{B2}$$

where $D_t = \mathbf{D}(\theta_t)$ is the dispersal matrix, $\Lambda_t = \Lambda(\theta_t)$ is the diagonal matrix of fitnesses $\boldsymbol{\lambda}_t = \boldsymbol{\lambda}(\theta_t)$, and θ_t is an ergodic, stationary sequence in the environmental state space Θ.

As described in the methods, the fitness vector can be written as $\boldsymbol{\lambda}(\theta) = \overline{\langle\boldsymbol{\lambda}\rangle} + \bar{\delta} + \delta(\theta)$ where $\bar{\delta} = \mathbb{E}[\boldsymbol{\lambda}_t] - \overline{\langle\boldsymbol{\lambda}\rangle}$ is the vector of spatial deviations, and $\delta_t = \delta(\theta_t)$ is the sum of the temporal and spatial-temporal deviations. Specifically, $\langle\delta_t\rangle$ and $\delta_t - \langle\delta_t\rangle$ are the temporal and spatial-temporal deviations at time t. The dispersal matrix $\mathbf{D}(\theta)$ is given by $\mathbf{D}(\theta) = D_{\text{sym}} + \bar{\Delta} + \Delta(\theta)$ where D_{sym} is a symmetric, column stochastic matrix, $\bar{\Delta} = \mathbb{E}[\mathbf{D}(\theta_t)] - D_{\text{sym}}$ is the spatial deviation, and $\Delta(\theta) = \mathbf{D}(\theta) - D_{\text{sym}} - \bar{\Delta}$ is the spatial-temporal deviation.

To use the derivative formulas in Appendix A, define the unperturbed matrix model as the constant matrix $\mathbf{A}(\theta) = D_{\text{sym}}\overline{\langle\boldsymbol{\lambda}\rangle}$ for all environmental states $\theta \in \Theta$. There are two types of perturbations to this base matrix model. First, the fitness perturbation is $\mathbf{B}(\theta) = D_{\text{sym}}\text{diag}(\bar{\delta} + \delta(\theta))$. Second, the dispersal perturbation is $\mathbf{C}(\theta) = \bar{\Delta} + \Delta(\theta)$ Using these perturbations, one has

$$\mathbf{D}(\theta)\Lambda(\theta) = (D_{\text{sym}} + \bar{\Delta} + \Delta(\theta))\text{diag}(\overline{\langle\boldsymbol{\lambda}\rangle} + \bar{\delta} + \delta(\theta)) = \mathrm{D}_{\text{sym}}\overline{\langle\boldsymbol{\lambda}\rangle} + \mathbf{B}(\theta) + \overline{\langle\boldsymbol{\lambda}\rangle}\mathbf{C}(\theta) + \mathbf{C}(\theta)\mathbf{B}(\theta) \tag{B3}$$

where $\text{diag}(\mathbf{v})$ denotes a diagonal matrix whose diagonal entries are given by the vector $\mathbf{v}$. To understand the effects of these perturbation, I compute all the relevant first and second order derivatives i.e. $\frac{\partial r}{\partial \mathbf{B}}(\mathbf{A})$, $\frac{\partial r}{\partial \mathbf{C}}(\mathbf{A})$, $\frac{\partial^2 r}{\partial \mathbf{B}^2}(\mathbf{A})$, $\frac{\partial^2 r}{\partial \mathbf{C}^2}(\mathbf{A})$ and $\frac{\partial^2 r}{\partial \mathbf{B}\partial \mathbf{C}}(\mathbf{A})$.

I begin by computing the first and second order derivatives of r with respect to $\mathbf{B}$. Let $\mathbf{1}$ be $k \times 1$ column vector of ones, and let $\mathbf{1}^T$ denote its transpose i.e. the $1 \times k$ row vector of ones. Let $\mathbf{0}$ be the $k \times 1$ vector of zeros. Recall, that $\langle \mathbf{v} \rangle = \frac{1}{k}\sum_i v_i$ corresponds to the spatial average of a vector $\mathbf{v}$. Using equation (A2) from Appendix A with $B_t = D_{\text{sym}}\text{diag}(\bar{\delta} + \delta_{\text{t}})$, $\mathbf{v} = \mathbf{1}^T$ and $\mathbf{w} = \mathbf{1}/k$ yields

$$\begin{aligned}\frac{\partial r}{\partial \mathbf{B}}(\mathbf{A}) =& \mathbb{E}[\mathbf{1}^T D_{\text{sym}}(\bar{\delta} + \delta_1)/k] \\ =& \mathbb{E}[\mathbf{1}^T(\bar{\delta} + \delta_1)/k] \\ =& \mathbb{E}[\mathbf{1}^T\bar{\delta}] + \mathbb{E}[\langle \delta_1 \rangle] = 0\end{aligned}$$

where the second line follows from $\mathbf{1}^T D_{\text{sym}} = \mathbf{1}^T$, and the last line follows from the average spatial deviation, $\langle \bar{\delta} \rangle$, being zero, and the vector of temporal deviations having zero expectation i.e. $\mathbb{E}[\delta_1] = \mathbf{0}$. To find the second order derivative $\frac{\partial^2 r}{\partial \mathbf{B}^2}(\mathbf{A})$, I expand the three terms in (A6) from Appendix A. For these expansions,

$$\mathbf{v}B_t\mathbf{w} = \mathbf{1}^T(\bar{\delta} + \delta_t)/k = \langle \delta_t \rangle$$

as $\mathbf{1}^T\bar{\delta} = 0$. Therefore,

$$\mathbb{E}\left[(\mathbf{v}B_1\mathbf{w})^2\right] = \mathbb{E}[\langle \delta_1 \rangle^2]$$

and

$$\mathbb{E}[\mathbf{v}B_t\mathbf{w} \cdot \mathbf{v}B_1\mathbf{w}] = \mathbb{E}[\langle \delta_t \rangle \langle \delta_1 \rangle]$$

Furthermore, as $\mathbb{E}[\delta_t] = \mathbf{0}$,

$$\mathbf{v}\mathbb{E}\left[B_t A^{t-2} B_1\right]\mathbf{w} = \frac{\overline{\langle \boldsymbol{\lambda} \rangle}^{t-2}}{k}\mathbb{E}[(\bar{\delta} + \delta_t)^T D_{\text{sym}}^{t-1}(\bar{\delta} + \delta_1)] = \frac{\overline{\langle \boldsymbol{\lambda} \rangle}^{t-2}}{k}\mathbb{E}[\bar{\delta}^T D_{\text{sym}}^{t-1}\bar{\delta} + \delta_t^T D_{\text{sym}}^{t-1}\delta_1)].$$

Substituting these expressions in (A6) from Appendix A yields

$$\text{(B4)} \qquad \frac{\partial^2 r}{\partial \mathbf{B}^2}(\mathbf{A}) = \frac{2}{k\overline{\langle \boldsymbol{\lambda} \rangle}^2}\sum_{t=2}^{\infty}\mathbb{E}[\bar{\delta}^T D_{\text{sym}}^{t-1}\bar{\delta} + \delta_t^T D_{\text{sym}}^{t-1}\delta_1)] - \frac{1}{\overline{\langle \boldsymbol{\lambda} \rangle}^2}\mathbb{E}[\langle \delta_1 \rangle^2] - \frac{2}{\overline{\langle \boldsymbol{\lambda} \rangle}^2}\sum_{t=2}^{\infty}\mathbb{E}[\langle \delta_t \rangle \langle \delta_1 \rangle]$$

As D_{sym} is symmetric, one can choose a basis of eigenvectors $\mathbf{u}(1), \ldots, \mathbf{u}(k)$ with associated eigenvalues $\mu_1, \ldots, \mu_k$ such that $\mathbf{u}(i)^T\mathbf{u}(i) = k$ for all i, $\mathbf{u}(j)^T\mathbf{u}(i) = 0$ for $i \neq j$, $\mathbf{u}(1) = \mathbf{1}$ and $\mu_1 = 1$. Let $\bar{\alpha}_i$ and $\alpha_{i,t}$ be such that

$$\bar{\delta} = \sum_{i=1}^{k}\bar{\alpha}_i\mathbf{u}(i) \text{ and } \delta_t = \sum_{i=1}^{k}\alpha_{i,t}\mathbf{u}(i)$$

where $\bar{\alpha}_1 = 0$. Then (B4) becomes

$$\text{(B5)} \qquad \frac{\partial^2 r}{\partial \mathbf{B}^2}(\mathbf{A}) = \frac{2}{\overline{\langle \boldsymbol{\lambda} \rangle}^2}\sum_{t=2}^{\infty}\sum_{i=1}^{k}\mu_i^{t-1}(\bar{\alpha}_i^2 + \mathbb{E}[\alpha_{i,t}\alpha_{i,1}]) - \frac{1}{\overline{\langle \boldsymbol{\lambda} \rangle}^2}\mathbb{E}[\langle \delta_1 \rangle^2] - \frac{2}{\overline{\langle \boldsymbol{\lambda} \rangle}^2}\sum_{t=2}^{\infty}\mathbb{E}[\langle \delta_t \rangle \langle \delta_1 \rangle]$$

To simplify this expression further, notice that

$$\langle \delta_t \rangle = \frac{1}{k}\mathbf{1}^T \sum_{i=1}^{k} \alpha_{i,t}\mathbf{u}(i) = \frac{1}{k}\alpha_{1,t}\mathbf{1}^T\mathbf{u}(1) = \alpha_{1,t}$$

and, consequently,

$$\frac{2}{\overline{\langle \boldsymbol{\lambda} \rangle}^2} \sum_{t=2}^{\infty} \mathbb{E}[\alpha_{i,t}\alpha_{i,1}] = \frac{2}{\overline{\langle \boldsymbol{\lambda} \rangle}^2} \sum_{t=2}^{\infty} \mathbb{E}[\langle \delta_t \rangle \langle \delta_1 \rangle].$$

Assume $\mathbb{E}[\alpha_{i,t}\alpha_{i,1}] = \rho_i^{t-1}\sigma_i^2$ for $t \geq 1$ i.e. there is an exponential decay of correlations. Then using the fact that $\bar{\alpha}_1 = 0$, equation (B5) becomes

$$\frac{\partial^2 r}{\partial \mathbf{B}^2}(\mathbf{A}) = \frac{2}{\overline{\langle \boldsymbol{\lambda} \rangle}^2} \sum_{t=2}^{\infty}\sum_{i=2}^{k} \mu_i^{t-1}(\bar{\alpha}_i^2 + \rho_i^{t-1}\sigma_i^2) - \frac{1}{\overline{\langle \boldsymbol{\lambda} \rangle}^2}\mathbb{E}[\langle \delta_1 \rangle^2]$$

and simplifies to

$$\frac{\partial^2 r}{\partial \mathbf{B}^2}(\mathbf{A}) = \frac{2}{\overline{\langle \boldsymbol{\lambda} \rangle}^2} \sum_{i=2}^{k} \left(\frac{\mu_i}{1-\mu_i}\bar{\alpha}_i^2 + \frac{\mu_i\rho_i}{1-\mu_i\rho_i}\sigma_i^2 \right) - \frac{1}{\overline{\langle \boldsymbol{\lambda} \rangle}^2}\mathbb{E}[\langle \delta_1 \rangle^2] \tag{B6}$$

Next, I compute the derivatives of r with respect to $\mathbf{C}$. As $\mathbf{1}^T\bar{\Delta} = \mathbf{0}^T$, $\mathbf{1}^T\Delta_t = \mathbf{0}^T$, and $\mathbb{E}[\Delta_t] = \mathbf{0}$, one gets $\frac{\partial r}{\partial \mathbf{C}}(\mathbf{A}) = \frac{\partial^2 r}{\partial \mathbf{C}^2}(\mathbf{A}) = 0$. Therefore, the only contribution of the perturbation $\mathbf{C}$ is through the mixed derivative $\frac{\partial^2 r}{\partial \mathbf{B}\partial \mathbf{C}}(\mathbf{A})$. To compute this mixed derivative, I compute the terms in equation (A5) of Appendix A. To this end, notice that

$$\mathbf{v}C_t\mathbf{w} = \mathbf{1}^T(\bar{\Delta} + \Delta_t)\mathbf{1}/k = 0 \text{ and } \mathbf{v}B_t\mathbf{w} = \mathbf{1}^T D_{\text{sym}}\text{diag}(\bar{\delta} + \delta_\text{t})\mathbf{1}/\text{k} = \langle \delta_\text{t} \rangle$$

Furthermore,

$$\begin{aligned}
\mathbb{E}[\mathbf{v}B_sA^{s-2}C_1\mathbf{w}] =& \mathbb{E}[\mathbf{1}^T D_{\text{sym}}\text{diag}(\bar{\delta} + \delta_\text{s})\overline{\langle \boldsymbol{\lambda} \rangle}^{\text{s}-2}\text{D}_{\text{sym}}^{\text{s}-2}(\bar{\Delta} + \Delta_1)\mathbf{1}/\text{k}] \\
=& \overline{\langle \boldsymbol{\lambda} \rangle}^{s-2}\mathbb{E}[(\bar{\delta} + \delta_s)^T D_{\text{sym}}^{s-2}(\bar{\Delta} + \Delta_1)\mathbf{1}/k] \\
=& \overline{\langle \boldsymbol{\lambda} \rangle}^{s-2}\frac{1}{k}\sum_{i=1}^{k}\mu_i^{s-2}\mathbb{E}[(\bar{\alpha}_i + \alpha_{i,s})\mathbf{u}(i)^T(\bar{\Delta} + \Delta_1)\mathbf{1}] \\
=& \overline{\langle \boldsymbol{\lambda} \rangle}^{s-2}\frac{1}{k}\sum_{i=2}^{k}\mu_i^{s-2}(\bar{\alpha}_i\mathbf{u}(i)^T\bar{\Delta}\mathbf{1} + \mathbb{E}[\alpha_{i,s}\mathbf{u}(i)^T\Delta_1\mathbf{1}])
\end{aligned}$$

where the last equality follows from $\bar{\alpha}_1 = 0$ and $\mathbf{u}(1)^T\Delta_t = 0$. On the other hand,

$$\mathbb{E}[\mathbf{v}C_sA^{s-2}B_1\mathbf{w}] = \mathbb{E}[\mathbf{1}^T(\bar{\Delta} + \Delta_t)\overline{\langle \boldsymbol{\lambda} \rangle}^{s-2}D_{\text{sym}}^{s-2}D_{\text{sym}}\text{diag}(\bar{\delta} + \delta_1)\mathbf{1}/\text{k}] = 0$$

Substituting these expressions in equation (A5) yields

$$\frac{\partial^2 r}{\partial \mathbf{B}\partial \mathbf{C}}(\mathbf{A}) = \frac{1}{k\overline{\langle \boldsymbol{\lambda} \rangle}^2} \sum_{t=2}^{\infty} \sum_{i=2}^{k} \mu_i^{t-2} (\bar{\alpha}_i \mathbf{u}(i)^T \bar{\Delta}\mathbf{1} + \mathbb{E}[\alpha_{i,t} \mathbf{u}(i)^T \Delta_1 \mathbf{1}])$$

Choose $\bar{\beta}_i$ and $\beta_{i,t}$ such that

$$\bar{\Delta}\mathbf{1} = \sum_i \bar{\beta}_i \mathbf{u}(i) \text{ and } \Delta_t \mathbf{1} = \sum_i \beta_{i,t} \mathbf{u}(i).$$

Then,

$$\frac{\partial^2 r}{\partial \mathbf{B}\partial \mathbf{C}}(\mathbf{A}) = \frac{1}{\overline{\langle \boldsymbol{\lambda} \rangle}^2} \sum_{t=2}^{\infty} \sum_{i=2}^{k} \mu_i^{t-2} (\bar{\alpha}_i \bar{\beta}_i + \mathbb{E}[\alpha_{i,t} \beta_{i,1}])$$

Assuming $\mathbb{E}[\alpha_{i,t}\beta_{i,1}] = p_i^{t-1} v_i$ yields

$$\frac{\partial^2 r}{\partial \mathbf{B}\partial \mathbf{C}}(\mathbf{A}) = \frac{1}{\overline{\langle \boldsymbol{\lambda} \rangle}^2} \sum_{i=2}^{k} \frac{\bar{\alpha}_i \bar{\beta}_i}{1-\mu_i} + \frac{p_i v_i}{1-\mu_i p_i}. \tag{B7}$$

Using these derivatives and equation (B3) yields the desired second order approximation

$$\begin{aligned} r \approx \ln \overline{\langle \boldsymbol{\lambda} \rangle} - \frac{\mathbb{E}[\langle \delta_1 \rangle^2]}{2\overline{\langle \boldsymbol{\lambda} \rangle}^2} + \sum_{i=2}^{k} \frac{\mu_i}{1-\mu_i} \frac{\bar{\alpha}_i^2}{\overline{\langle \boldsymbol{\lambda} \rangle}^2} + \sum_{i=2}^{k} \frac{\mu_i \rho_i}{1-\mu_i \rho_i} \frac{\sigma_i^2}{\overline{\langle \boldsymbol{\lambda} \rangle}^2} \\ + \sum_{i=2}^{k} \frac{1}{1-\mu_i} \frac{\bar{\alpha}_i \bar{\beta}_i}{\overline{\langle \boldsymbol{\lambda} \rangle}^2} + \sum_{i=2}^{k} \frac{1}{1-\mu_i p_i} \frac{p_i v_i}{\overline{\langle \boldsymbol{\lambda} \rangle}^2} \end{aligned} \tag{B8}$$

Equation (B8) gives equation (13) by re-indexing $i \mapsto i-1$ and defining $\mathrm{CV}^2_{\text{time}} = \frac{\mathbb{E}[\langle \delta_1 \rangle^2]}{\overline{\langle \boldsymbol{\lambda} \rangle}^2}$, $\mathrm{CV}^2_{\text{space},i} = \frac{\bar{\alpha}_i^2}{\overline{\langle \boldsymbol{\lambda} \rangle}^2}$, $\mathrm{cov}_{D\Lambda\text{-space},i} = \bar{\alpha}_i \bar{\beta}_i$, $\mathrm{CV}^2_{\text{st},i} = \frac{\rho_i \sigma_i^2}{\overline{\langle \boldsymbol{\lambda} \rangle}^2}$, $\rho_{\text{st},i} = \rho_i$, $\rho_{D\Lambda\text{-st},\mathrm{i}} = p_i$, and $\mathrm{cov}_{D\Lambda\text{-space},i} = p_i v_i$.