Mathematical Statistics
Stockholm University

Spatial Autocorrelation for Subdivided Populations with Invariant Migration Schemes

Ola Hössjer

Research Report 2012:11

ISSN 1650-0377

Postal address:

Mathematical Statistics
Dept. of Mathematics
Stockholm University
SE-106 91 Stockholm
Sweden

Internet:

http://www.math.su.se/matstat

Mathematical Statistics
Stockholm University
Research Report **2012:11**,
http://www.math.su.se/matstat

Spatial Autocorrelation for Subdivided Populations with Invariant Migration Schemes

Ola Hössjer*

October 7, 2012

Abstract

For populations with geographic substructure and selectively neutral genetic data, the short term dynamics is a balance between migration and genetic drift. Before fixation of any allele, the system enters into a quasi equilibrium (QE) state. Hössjer and Ryman (2012) developed a general QE methodology for computing approximations of spatial autocorrelations between subpopulations, subpopulation differentiation (fixation indexes) and variance effective population sizes. In this paper we treat a class of models with translationally invariant migration and use Fourier transforms for computing these quantities. We show how the QE approach is related to other methods based on conditional kinship coefficients between subpopulations under mutation-migration-drift equilibrium. We also verify that QE autocorrelations are closely related to the expected value of Moran's autocorrelation function and treat limits of continuous spatial location (isolation by distance) and infinite gitters of subpopulations. The theory is illustrated with several examples including island models, circular and torus stepping stone models, von Mises models, hierarchical island models and Gaussian models. It is well known that the fixation index contains information about the effective number of migrants. The spatial autocorrelations is complementary and typically reveal the type of migration (local or global).

Key words: Fixation index, Fourier transforms, kinship coefficients, quasi equilibrium, spatial autocorrelations, translationally invariant migration.

Mathematics subject classification: 92D25, 60F99.

*Department of Mathematics, Div. of Mathematical Statistics, Stockholm University.

1 Introduction

It is often the case that genetic variables of a population exhibit geographic variation. This can either be modeled by dividing the population into a discrete set of subpopulations, or treating spatial location as a continuous variable. In either case, models for the subsequent microevolution of the population involve four sources of variation; genetic drift due to random fertilization, migration between geographic sites, mutation and selection, see for instance Nei (1977) or Durrett (2008).

Genetic data involves markers at a number of loci along the genome, sampled for individuals at a number of geographic sites at different points in time. In this paper we are interested the relatively short time scales encountered in consservational biology (see for instance Palstra and Ruzzante, 2008, and Hare et al., 2011). We focus on neutral genetic markers and disregard the occurrence of new mutations. The dynamics of the system is then a balance between genetic drift and migration. In absence of new mutations, one of the alleles of each marker will eventually become fixed. However, before this happens the population converges to a state of quasi equilibrium between genetic drift and migration, as formalized by Hössjer et al. (2012) for the island model and more generally in Hössjer and Ryman (2012).

In order to assess the nature and magnitude of genetic drift and migration, it is customary to compute a few summary statistics from data, from which a genetic model is fitted. This includes the effective population size (Wright, 1931, 1938) that quantifies the amount of genetic drift, the fixation index (Wright, 1951) that quantifies the amount of spatial genetic variation, and spatial autocorrelation functions (Sokal and Oden, 1978, Slatkin and Arter, 1991, Sokal et al., 1997, Rousset, 2001, Hardy and Vekemans, 2002).

Hössjer and Ryman (2012) derived formulas for approximations of so called standardized genetic drift and spatial covariance matrices. This was used in order to approximate the variance effective size, the fixation index and spatial autocorrelations of the population at quasi equilibrium. In this paper, we treat populations with spatial symmetry and translationally invariant migration and use Fourier analysis in order to derive fast algorithms for computing these quantities, with special attention to spatial autocorrelation. This is of interest, for instance, in plant genetics, since the fine-scale genetic structure indicates the amount of inbreeding and selfing in the population (Vekemans and Hardy, 2004, Zhao et al., 2009).

In more detail, the paper is organized as follows: The models of migration and reproduction are defined in Section 2 and the standardized genetic drift and spatial covariance matrices in Section 3. In particular, we show that

the entries of the latter can be interpreted as conditional kinship coefficients between subpopulations. The quasi equilibrium approach is introduced in Section 4, and in Section 5 we treat the special case of translationally invariant migration. In particular, we derive explicit formulas for quasi equilibrium approximations of the fixation index, variance effective size and spatial autocorrelation. We also show that the latter is closely related to the expected value of Moran's autocorrelation function. In Section 6 we consider the traditional approach based on kinship coefficients between subpopulations under migration-mutation-drift equilibrium. We extend this approach by allowing the local effective size of each subpopulation to be different from the local census size, and we also assume that genes are drawn *with* replacement when computing kinship coefficients. When the mutation probability tends to zero, it turns out that this methodology is similar but not identical to the quasi equilibrium approach. In Section 7 we consider models with infinitely many subpopulations; either continuous spatial location (isolation by distance) or infinite gitters. In Section 8 we illustrate the theory for several models and in Section 9 we present algorithms and numerical results. A summary and discussion is provided in Section 10 and proofs are collected in the appendix.

2 Migration and Reproduction Model

Consider a population of N diploid individuals divided into a set $\mathcal{G}$ of s subpopulations. For each generation $t = 1, 2, \ldots$, these subpopulations have sizes Nu_j, $j \in \mathcal{G}$. We will focus on a polymorphic biallelic region of DNA, which, for simplicity, we refer to as a gene, although it is typically a Single Nucleotide Polymorphism (SNP) or some other genetic marker. Hence there are $2Nu_j$ genes in subpopulation j, and P_{tj} is the fraction of these having one of the two alleles. The overall frequency of this allele is

$$P_t^{\boldsymbol{w}} = \sum_{j \in \mathcal{G}} w_j P_{tj},$$

when subpopulations are assigned non-negative weights $\boldsymbol{w} = (w_j;\, j \in \mathcal{G})$, with $\sum_{j \in \mathcal{G}} w_j = 1$. In particular, all genes in the population are assigned the same weight if $\boldsymbol{w}$ equals $\boldsymbol{u} = (u_j;\, j \in \mathcal{G})$, so that subpopulations are weighted proportionally to their sizes. The vector

$$\boldsymbol{P}_t = (P_{tj};\, j \in \mathcal{G})^T \tag{1}$$

summarizes how the allele frequency is spatially distributed over subpopulations, with T the Hermitian or conjugate transpose operator.

We assume that the migration rate from subpopulation k to j is M_{kj}, loosely defined as the expected number of offspring genes in subpopulation j that each subpopulation k gene passes on to the next generation. The constant subpopulation sizes require that

$$u_j = \sum_{k \in \mathcal{G}} u_k M_{kj} \tag{2}$$

holds for all $j \in \mathcal{G}$. This can also be formulated as $\boldsymbol{u}$ being a left eigenvector of the migration matrix $\boldsymbol{M} = (M_{kj})_{k,j \in \mathcal{G}}$ with eigenvalue 1. In order to study the dynamics of $\boldsymbol{P}_t$, it is more relevant to consider the backward migration matrix $\boldsymbol{B} = (B_{jk})$ rather than $\boldsymbol{M}$, where

$$B_{jk} = \frac{u_k M_{kj}}{u_j} \tag{3}$$

is the fraction of genes of subpopulation j that originate from subpopulation k in the previous generation. It follows from (2)-(3) that $\boldsymbol{B}$ has row sums equal to one, and hence it is the transition matrix of a Markov chain, assumed to be irreducible and aperiodic with a unique stationary distribution $\boldsymbol{\gamma} = (\gamma_j;\, j \in \mathcal{G})$.

It is shown in Hössjer and Ryman (2012) that

$$\boldsymbol{P}_{t+1} = \boldsymbol{B}\boldsymbol{P}_t + \boldsymbol{\varepsilon}_{t+1}$$

evolves as vector valued time inhomogeneous autoregressive process for a large class of reproduction models and migration schemes, with error term $\boldsymbol{\varepsilon}_{t+1}$ satisfying $E(\boldsymbol{\varepsilon}_{t+1}|\boldsymbol{P}_t) = 0$ for a selectively neutral allele. The exact distributional properties of $\boldsymbol{\varepsilon}_t$ depends on the chosen reproduction model.

We will assume that N_{ej} is the variance effective size of subpopulation j and use one of the reproduction models considered by Hössjer et al. (2012) and Hössjer and Ryman (2012), with fertilization preceding migration (FM). It can be summarized as follows:

FM1 Gamete formation: In each subpopulation k, the reproduction cycle from generation t to $t+1$ starts by randomly selecting $2N_{ek}$ breeding genes without replacement from the set of all $2Nu_k$ genes. The $2N_{ek}$ breeding genes of subpopulation k generate an infinite gamete pool, with equal contribution from each gene.

FM2 Fertilization: For each pair k, j of subpopulations, draw $2Nu_kM_{kj}$ gametes binomially from gamete pool k.

FM3 Migration: The $2Nu_kM_{kj}$ genes of the previous step migrate from k to j. This is repeated for all pairs k, j.

3 Measures of Spatial and Temporal Allele Frequency Change

The allele frequency vector can be divided into two orthogonal components,

$$\boldsymbol{P}_t = P_t\mathbf{1} + \boldsymbol{P}_t^0, \tag{4}$$

where

$$P_t = P_t^{\boldsymbol{\gamma}} = \sum_{j\in\mathcal{G}} \gamma_j P_{tj}$$

is the global allele frequency of generation t when subpopulations are weighted as

$$\boldsymbol{w} = \boldsymbol{\gamma}, \tag{5}$$

and $\mathbf{1} = (1,\ldots,1)^T$ is an $s \times 1$ column vector of ones. Hence the first term $P_t\mathbf{1}$ of (4) gives the overall genetic drift whereas the second term $\boldsymbol{P}_t^0$ describes spatio-temporal fluctuations of the allele frequencies around the overall mean.

The dynamics of the time series $\{\boldsymbol{P}_t\}_{t\geq 1}$ is well characterized by the standardized genetic drift covariance matrix

$$\boldsymbol{\Sigma}_t = \frac{\mathrm{Cov}(\boldsymbol{\varepsilon}_{t+1}|P_t)}{P_t(1-P_t)}$$

and the standardized spatial covariance matrix

$$\boldsymbol{V}_t = \frac{E(\boldsymbol{P}_t^0(\boldsymbol{P}_t^0)^T|P_t)}{P_t(1-P_t)}. \tag{6}$$

It turns out that $\boldsymbol{V}_t = (V_{tjk})$ can be characterized in terms of kinship coefficients (Malécot, 1948). Define, for each pair j,k of subpopulations, the kinship coefficient $f_{t,jk}$ of generation t as the probability that two randomly drawn (with replacement if $j=k$) genes are identical by descent (IBD). In the same way, f_t, the apriori kinship, is taken to be the probability that two genes in generation t are IBD when drawn with replacement from the whole population. Assuming an infinite alleles model, it follows that IBD is equivalent to having the same allele, i.e. being identical by state. Hence

$$\begin{aligned} f_{t,jk} &= P_{tj}P_{tk} + (1-P_{tj})(1-P_{tk}), \\ f_t &= P_t^2 + (1-P_t)^2. \end{aligned} \tag{7}$$

Yet another definition

$$\tilde{f}_{t,jk} = \frac{(P_{tj}-P_t)(P_{tk}-P_t)}{P_t(1-P_t)} \tag{8}$$

of the kinship coefficient between subpopulations j and k is given by Barbujani (1987). The following result shows that (7) and (8) are both closely related to V_{tjk}:

Proposition 1 *The j,k entry of the standardized spatial covariance matrix equals*

$$V_{tjk} = E(\tilde{f}_{t,jk}|P_t) \tag{9}$$

and

$$V_{tjk} = \frac{E(f_{t,jk}|P_t) - f_t + (1-2P_t)E(P_{tj} + P_{tk} - 2P_t|P_t)}{1 - f_t} \tag{10}$$

respectively. In particular, if the components of $\boldsymbol{P}_t$ in (1) are exchangeable, it follows that $E(\boldsymbol{P}_t^0|P_t) = 0$ and (10) simplifies to

$$V_{tjk} = \frac{E(f_{tjk}|P_t) - f_t}{1 - f_t}. \tag{11}$$

The right hand side of (11) is sometimes referred to as a conditional kinship coefficient (Morton, 1973, Hardy and Vekemans, 1999, Rousset, 2002), and $E(f_{tjk}|P_t)$ as the average kinship coefficient (7) between subpopulations j and k. Even without the exchangeability condition on $\boldsymbol{P}_t$, we expect $E(\boldsymbol{P}_t^0|P_t)$ to be close to 0, so that (10) holds approximately.

In this paper, we will mainly focus on the spatial autocorrelation function

$$\rho_{tjk}^{\boldsymbol{w}} = \mathrm{Corr}(P_{tj}, P_{tk}|P_t^{\boldsymbol{w}}) \overset{\boldsymbol{w}=\boldsymbol{\gamma}}{=} \mathrm{Corr}(P_{tj}, P_{tk}|P_t) \overset{E(\boldsymbol{P}_t^0|P_t)=0}{=} \frac{V_{tjk}}{\sqrt{V_{tjj}}\sqrt{V_{tkk}}} \tag{12}$$

between all pairs j,k of subpopulations in generation t, when local allele frequencies are weighted according to $\boldsymbol{w}$. It is shown by Hössjer and Ryman (2012) that (5) weights genes proportionally to their reproductive values. We will refer to it as the canonical weighting scheme, and it has been advocated e.g. by Felsenstein (1971) and Waples and Yokota (2007). Whenever this scheme is used, we will drop superscript $\boldsymbol{w}$ for any quantity that involves $\boldsymbol{w}$.

Nei (1973) considered a distance

$$D_{tjk} = -\log\left(\frac{f_{t,jk}}{\sqrt{f_{t,jj}}\sqrt{f_{t,jj}}}\right)$$

between subpopulations j and k. It can be seen that the multilocus generalization of this distance is somewhat related to

$$D_{tjk}^* = -\log\left(\frac{E(f_{t,jk}|P_t)}{\sqrt{E(f_{t,jj}|P_t)}\sqrt{E(f_{t,jj}|P_t)}}\right). \tag{13}$$

The following result shows that D^*_{tjk} is well approximated by $-\log(\rho_{tjk})$ whenever the apriori kinship coefficient f_t is small:

Proposition 2 *Given that (11) holds, the generalized version (13) of Nei's distance between subpopulations j and k at time t, can be written as*

$$D^*_{tjk} = -\log\left(\frac{V_{tjk} + f_t(1-V_{tjk})}{\sqrt{V_{tjj} + f_t(1-V_{tjj})}\sqrt{V_{tkk} + f_t(1-V_{tkk})}}\right) \overset{f_t \approx 0}{\approx} -\log(\rho_{tjk}). \tag{14}$$

Formula (14) can be verified by combining the definition of D^*_{tjk} in (13) with (11). The $f_t \approx 0$ approximation follows from the definition of ρ_{tjk} on the right hand side of (12).

The two other quantities we will study are the fixation index

$$F^{\boldsymbol{w}}_{ST,t} = \frac{\sum_{j\in\mathcal{G}} w_j (P_{tj} - P^{\boldsymbol{w}}_t)^2}{P^{\boldsymbol{w}}_t(1-P^{\boldsymbol{w}}_t)} \overset{\boldsymbol{w}=\boldsymbol{\gamma}}{=} \frac{\sum_{j\in\mathcal{G}} \gamma_i (P_{tj} - P_t)^2}{P_t(1-P_t)} \tag{15}$$

the variance effective size

$$N^{\boldsymbol{w}}_{eV,t} = \frac{P^{\boldsymbol{w}}_t(1-P^{\boldsymbol{w}}_t)}{2\mathrm{Var}(P^{\boldsymbol{w}}_{t+1} - P^{\boldsymbol{w}}_t | P^{\boldsymbol{w}}_t)} \overset{\boldsymbol{w}=\boldsymbol{\gamma}}{=} \frac{P_t(1-P_t)}{2\mathrm{Var}(P_{t+1} - P_t | P_t)} \tag{16}$$

of the population.

4 Quasi Equilibrium

Eventually, as $t \to \infty$, one of the two alleles will become fixed in all subpopulation, unless new mutations take place. However, for large populations the time for this to happen is large. Before fixation the system converges to a quasi equilibrium mode at a rate that depends on the amount of migration between subpopulations. Hössjer and Ryman (2012) formulated this mathematically, and proved in particular that conditionally on non-fixation of both alleles, $\boldsymbol{\Sigma}_t$ and $\boldsymbol{V}_t$ are well approximated, under quasi equilibrium, by covariance matrices $\boldsymbol{\Sigma} = (\Sigma_{jk})$ and $\boldsymbol{V} = (V_{jk})$ for large populations and small amounts of allele frequency fluctuations between subpopulations. They also showed that $\boldsymbol{\Sigma}$ and $\boldsymbol{V}$ can be computed from a recursive set of equations, whose first part is given by

$$\boldsymbol{V} = \boldsymbol{B}\boldsymbol{V}\boldsymbol{B}^T + \tilde{\boldsymbol{\Sigma}}, \tag{17}$$

where

$$\tilde{\boldsymbol{\Sigma}} = \mathrm{Cov}(\boldsymbol{\varepsilon}_t - (\boldsymbol{\gamma}\boldsymbol{\varepsilon}_t)\mathbf{1}) = (\boldsymbol{I} - \mathbf{1}\boldsymbol{\gamma})\boldsymbol{\Sigma}(\boldsymbol{I} - \mathbf{1}\boldsymbol{\gamma})^T, \tag{18}$$

and $\boldsymbol{I}$ is the identity matrix of order s. The form of the second part of these equations depends on the particular reproduction scheme. For FM1-FM3 they are given by

$$\Sigma_{jk} = \sum_{l\in\mathcal{G}} \left(\frac{1}{2N_{el}} - \frac{1}{2Nu_l}\right) B_{jl}B_{kl}(1 - V_{ll}) + \frac{1_{\{j=k\}}}{2Nu_j}\left(1 - \sum_{l\in\mathcal{G}} B_{jl}V_{ll}\right), \tag{19}$$

for all $j, k \in \mathcal{G}$.

By the symmetry of $\boldsymbol{V}$ and $\boldsymbol{\Sigma}$, each matrix contains $s(s+1)/2$ unknown entries on or below the diagonal, and therefore (17)-(19) defines a linear system of equations for the $s(s+1)$ unknowns $\{V_{jk}, \Sigma_{jk};\, j \geq k\}$, given that the subpopulations in $\mathcal{G}$ are ordered in some way.

Quasi equilibrium values of the spatial autocorrelation function, fixation index and effective population size are defined as limits of the expected values of (12), (15) and (16) as $t \to \infty$, conditionally on that no allele gets fixed, see Nei et al. (1977), Hössjer et al. (2011) and Hössjer and Ryman (2012). Approximate expressions $\rho_{jk}^{\mathrm{appr},\boldsymbol{w}}$, $F_{ST}^{\mathrm{appr},\boldsymbol{w}}$ and $N_{eV}^{\mathrm{appr},\boldsymbol{w}}$ for these quasi equilibrium limits are computed from $\boldsymbol{\Sigma}$ and $\boldsymbol{V}$ for a selectively neutral allele. They simplify considerably for the canonical weighting scheme (5) and equal

$$\rho_{jk}^{\mathrm{appr}} \overset{E(\boldsymbol{P}_t^0|P_t)=0}{=} \frac{V_{jk}}{\sqrt{V_{jj}V_{kk}}}, \tag{20}$$

$$F_{ST}^{\mathrm{appr}} = \sum_{j\in\mathcal{G}} \gamma_j V_{jj} \tag{21}$$

and

$$N_{eV}^{\mathrm{appr}} = \frac{1}{2\boldsymbol{\gamma}\boldsymbol{\Sigma}\boldsymbol{\gamma}^T} \tag{22}$$

respectively.

5 Spatially Invariant Migration

We assume that the subpopulations are located at geographic sites $\mathcal{X} = \{x_j;\, j \in \mathcal{G}\}$ that form a group under $+$, with $-$ the inverse operation. This induces the corresponding (inverse) group operation $+$ $(-)$ on $\mathcal{G}$ through $x_{j+k} = x_j + x_k$ $(x_{j-k} = x_j - x_k)$. Migration is invariant with respect $+$, i.e. $M_{jk} = M_{j'k'}$ for all pairs j, k and j', k' such that $(j', k') = (j + l, k + l)$ for

some $l \in \mathcal{G}$. Hence there exists a vector $\boldsymbol{m} = (m_j;\, j \in \mathcal{G})$ with $\sum_{j\in\mathcal{G}} m_j = 1$ such that

$$M_{kj} = m_{j-k} \tag{23}$$

for all j, k. From this it follows that (5) holds with

$$\boldsymbol{u} = \boldsymbol{\gamma} = \mathbf{1}^T/s, \tag{24}$$

and plugging (24) into (3) we find that $\boldsymbol{B}$ and $\boldsymbol{M}$ are both doubly stochastic matrices, that is, they both have row and column sums equal to one, with

$$\boldsymbol{B} = \boldsymbol{M}^T. \tag{25}$$

For this reason, the migration rates M_{kj} will also be referred to as migration probabilities. We will assume that the

$$s = \prod_{l=1}^{d} s_l \tag{26}$$

subpopulations are located as equispaced elements $x_j = \left(e^{2\pi i j_1/s_1}, \ldots, e^{2\pi i j_d/s_d}\right)$ of the d-dimensional torus. That is, $\mathcal{X}$ is a commutative (Abelian) group, isomorphic to the direct product

$$\mathcal{G} = \mathbb{Z}_{s_1} \oplus \ldots \oplus \mathbb{Z}_{s_d}$$

of the d cyclic groups $\mathbb{Z}_{s_l}$ of order s_l, so that each element of $\mathcal{G}$ can be written as multiindex $j = (j_1, \ldots, j_d)$, with $0 \le j_l \le s_l - 1$.

We will also assume that the reproduction model is spatially symmetric, with a constant local effective size

$$N_e = N_{ej} \tag{27}$$

of all subpopulations j. It follows from (24) that the local census size

$$N_c = N u_j = N/s \tag{28}$$

is also the same for all subpopulations. Equations (27)-(28), the translationally invariant migration (23) and (25) imply that the $\boldsymbol{\Sigma}$ is also translationally invariant. Indeed, (19) simplifies to

$$\begin{aligned}
\Sigma_{jk} &= \left(\tfrac{1}{2N_e} - \tfrac{1}{2N_c}\right) \textstyle\sum_{l\in\mathcal{G}} M_{lj} M_{lk} (1 - V_{ll}) + \tfrac{1_{\{j=k\}}}{2N_c} \left(1 - \textstyle\sum_{l\in\mathcal{G}} M_{lj} V_{ll}\right) \\
&= \left(\tfrac{1}{2N_e} - \tfrac{1}{2N_c}\right) \textstyle\sum_{l\in\mathcal{G}} m_{j-l} m_{k-l} (1 - V_{ll}) + \tfrac{1_{\{j=k\}}}{2N_c} \left(1 - \textstyle\sum_{l\in\mathcal{G}} m_{j-l} V_{ll}\right) \\
&= \left(\tfrac{1}{2N_e} - \tfrac{1}{2N_c}\right) \textstyle\sum_{l\in\mathcal{G}} m^{-}_{l-j} m_{k-l} (1 - v_0) + \tfrac{1_{\{j=k\}}}{2N_c} \left(1 - \textstyle\sum_{l\in\mathcal{G}} m_{j-l} v_0\right) \\
&= \sigma_{k-j},
\end{aligned} \tag{29}$$

with $m_j^- = m_{-j}$,

$$\sigma_j = (1 - v_{\mathbf{0}}) \left\{ \left(\frac{1}{2N_e} - \frac{1}{2N_c} \right) (\boldsymbol{m}^- * \boldsymbol{m})_j + \frac{1_{\{j=\mathbf{0}\}}}{2N_c} \right\} \tag{30}$$

and $\boldsymbol{m}^- * \boldsymbol{m}$ the convolution operator between the vectors $\boldsymbol{m}$ and $\boldsymbol{m}^- = (m_j^-)$. We interpret the elements of the vector $\boldsymbol{\sigma} = (\sigma_j;\, j \in \mathcal{G})$ as the (approximate) covariance function of the spatially stationary random field $\boldsymbol{\varepsilon}_{t+1}/\sqrt{P_t(1-P_t)}$, when quasi equilibrium has been attained.

In the third step of (29) we used (17), (18) and (24) to deduce that $\boldsymbol{V}$ is also translationally invariant, i.e. there exists a vector $\boldsymbol{v} = (v_j;\, j \in \mathcal{G})$ such that

$$V_{jk} = v_{k-j} \tag{31}$$

for all j, k. We interpret $\boldsymbol{v}$ as the (approximate) covariance function of the spatially invariant random field $\boldsymbol{P}_t^0/\sqrt{P_t(1-P_t)}$ under quasi equilibrium.

The system of equations in (17) and (19) for computing $\boldsymbol{\Sigma}$ and $\boldsymbol{V}$ simplifies considerably under spatial invariance (23), (29) and (31). To this end, we introduce the scalar product

$$(\boldsymbol{q}, \boldsymbol{q}') = \sum_{j \in \mathcal{G}} q_j \bar{q}'_j \tag{32}$$

for column vectors $\boldsymbol{q} = (q_j;\, j \in \mathcal{G})^T$ and $\boldsymbol{q}' = (q'_j;\, j \in \mathcal{G})^T$, with $\bar{q}'_j$ the complex conjugate of q'_j. For any $r = (r_1, \ldots, r_d) \in \mathcal{G}$ we define the unit column vector $\boldsymbol{q}_r = (q_{rj};\, j \in \mathcal{G})^T$ with components

$$q_{rj} = \exp\left(2\pi i \sum_{l=1}^{d} r_l j_l / s_l \right) / \sqrt{s}.$$

As shown in the appendix, the translational migration invariance (23) implies that $\{\boldsymbol{q}_r\}_{r \in \mathcal{G}}$ forms an orthonormal system of right eigenvectors of $\boldsymbol{B}$ with

$$\boldsymbol{B}\boldsymbol{q}_r = \hat{m}_r \boldsymbol{q}_r, \tag{33}$$

for all $r \in \mathcal{G}$, where

$$\hat{m}_r = \sum_{j \in \mathcal{G}} m_j \exp(-2\pi i \sum_{l=1}^{d} r_l j_l / s_l) \tag{34}$$

are the Fourier coefficients of $\boldsymbol{m}$. Notice that all $\hat{m}_r$ are real when $\boldsymbol{M}$ is symmetric and $m_{-j} = m_j$, although this need not generally be the case.

Put $\mathbf{0} = (0, \ldots, 0)$, and define, for each $j \in \mathcal{G}$,

$$S_j = \frac{1}{s} \sum_{r \in \mathcal{G} \setminus \mathbf{0}} \frac{\left(\frac{1}{2N_e} - \frac{1}{2N_c}\right) |\hat{m}_r|^2 + \frac{1}{2N_c}}{1 - |\hat{m}_r|^2} \cos\left(2\pi \sum_{l=1}^{d} r_l j_l / s_l\right), \tag{35}$$

so that in particular

$$S_{\mathbf{0}} = \frac{1}{s} \sum_{r \in \mathcal{G} \setminus \mathbf{0}} \frac{\left(\frac{1}{2N_e} - \frac{1}{2N_c}\right) |\hat{m}_r|^2 + \frac{1}{2N_c}}{1 - |\hat{m}_r|^2},$$

and

$$T_j = \frac{1}{s} \sum_{r \in \mathcal{G}} \left\{ \left(\frac{1}{2N_e} - \frac{1}{2N_c}\right) |\hat{m}_r|^2 + \frac{1}{2N_c} \right\} \cos\left(2\pi \sum_{l=1}^{d} r_l j_l / s_l\right). \tag{36}$$

We then have the following result that is proved in the appendix:

Theorem 1 *Consider a genetic model with reproduction scheme FM1-FM3 and spatially invariant migration (23). Then the entries of the standardized genetic drift and spatial covariance matrices* $\mathbf{\Sigma}$ *and* $\boldsymbol{V}$ *in (29) and (31) can be retrieved from*

$$\sigma_j = \frac{T_j}{1 + S_{\mathbf{0}}} \tag{37}$$

and

$$v_j = \frac{S_j}{1 + S_{\mathbf{0}}} \tag{38}$$

respectively, with S_j *and* T_j *as defined in (35) and (36). Moreover, the approximations (21), (22) and (20) of the quasi equilibrium limits of the fixation index, variance effective size and spatial autocorrelation function simplify to*

$$F_{ST}^{appr} = v_{\mathbf{0}} = \frac{S_{\mathbf{0}}}{1 + S_{\mathbf{0}}}, \tag{39}$$

$$N_{eV}^{appr} = \frac{sN_e}{1 - F_{ST}^{appr}} \tag{40}$$

and

$$\rho_j^{appr} = \rho_{k,k+j}^{appr} = \frac{v_j}{v_{\mathbf{0}}} = \frac{S_j}{S_{\mathbf{0}}} \tag{41}$$

respectively, for the canonical weights (5).

We notice in particular that (40) agrees with the well known expression for the fixation index derived by Wright (1951) for the infinite island model, and more generally by Wang and Caballero (1999, eqn. 15).

For spatially invariant migration schemes, it turns out that ρ_j^{appr} in (41) is related to a version of Moran's autocorrelation function

$$I_{tj} = \frac{s \sum_{i,k \in \mathcal{G}} w_{jik}(P_{ti} - P_t)(P_{tk} - P_t)}{\sum_{i,k \in \mathcal{G}} w_{jik} \sum_{i \in \mathcal{G}} (P_{ti} - P_t)^2}. \tag{42}$$

between subpopulations at distance $j \in \mathcal{G}$ in generation t (Moran, 1950, Sokal and Oden, 1978). The weights w_{jik} are binary entries of a connectivity matrix $\boldsymbol{W}_j = (w_{jik})$, such that $w_{jik} = 1$ if $k - i = \pm j$ and 0 otherwise. We denote the ratio of the expected values of the numerator and denominator in (42) as

$$E(I_{tj}|P_t)^* = \frac{s \sum_{i,k \in \mathcal{G}} w_{jik} E\left((P_{ti} - P_t)(P_{tk} - P_t)|P_t\right)}{\sum_{i,k \in \mathcal{G}} w_{jik} \sum_{i \in \mathcal{G}} E\left((P_{ti} - P_t)^2|P_t\right)}. \tag{43}$$

In general this quantity differs in from $E(I_{tj}|P_t)$, although the difference is small when the number of terms $2s$ in the numerator and denominator of (42) is large. This is even more true for multilocus generalizations of I_{tj}, when contributions from several subpopulations *and* loci are added over separately in the numerator and denominator.

The connection between $E(I_{tj}|P_t)^*$ and ρ_j^{appr} can summarized as follows:

Proposition 3 *The approximation (43) of $E(I_{tj}|P_t)$ satisfies*

$$E(I_{tj}|P_t)^* = \frac{s \sum_{i,k \in \mathcal{G}} w_{jik} V_{tik}}{\sum_{i,k \in \mathcal{G}} w_{jik} \sum_{i \in \mathcal{G}} V_{tii}}. \tag{44}$$

In particular, the quasi equilibrium approximation of (44), obtained by replacing V_{tik} by V_{ik}, equals

$$I_j^{appr} = \rho_j^{appr} \tag{45}$$

for spatially invariant migration schemes.

It is well known that autocorrelations need to be averaged over several alleles or loci in order to remove noise from single locus biallelic autocorrelations, see for instance Slatkin and Arter (1991). This indicates, in view of (45), that $\{\rho_j^{\text{appr}};\ j \in \mathcal{G}\}$ should be useful summary statistics for inferential purposes. The same conclusion can be drawn from the simulation study of Sokal et al. (1997).

6 Equilibrium Between Mutation, Migration and Drift

When each gene has a mutation probability $\mu > 0$ per generation, it is possible to attain equilibrium between mutation, migration and drift. This approach was introduced by Malécot (1950, 1951), and has subsequently been studied by several authors, see for instance Sawyer (1976) and Durrett (2008). We extend this work and consider a scenario where the local census and effective population sizes are separate entities. The reproduction model is slightly different from FM1-FM3, in that migration precedes fertilization (MF), as follows:

MF1 Gamete formation: Same as FM1, but with a probability μ_1 that each gamete mutates.

MF2 Migration: The infinitely sized gamete pools mix, so that after migration, gamete pool j contains exact proportions $\{B_{jk}\}_{k\in\mathcal{G}}$ of the contents of all pre-migration gamete pools $k \in \mathcal{G}$.

MF3 Fertilization: In each subpopulation j, $2Nu_j$ genes are drawn binomially from its infinitely sized post-migration gamete pool j. The mutation probability during fertilization for each gamete is μ_2.

This reproduction scheme was introduced by Hössjer et al. (2012) and generalizes the stochastic migration scheme of Sved and Latter (1977) for the island model, when $N_e = N_c$. Assuming that the mutation events in MF1 or MF3 are independent, the probability is

$$\mu = \mu_1 + \mu_2 - \mu_1\mu_2$$

for a gene to mutate at least once during the whole reproduction cycle.

It is shown in the appendix that MF1-MF3 facilitates computation of the kinship coefficient or IBD probability $f_{jk}^{\mathrm{eq}}(\mu)$ at equilibrium between two genes sampled randomly from subpopulations j and k. We use μ as argument, since it turns out that f_{jk}^{eq} depends on μ_1 and μ_2 only through μ. Latter and Sved (1981) studied the island model and required that when $j = k$, the two genes are drawn *with* replacement from subpopulation j. We follow this approach, since it offers additional flexibility and generality. We will also assume that translational invariance holds, as in Section 5, so that

$$f_j^{\mathrm{eq}}(\mu) = f_{k,k+j}^{\mathrm{eq}}(\mu)$$

independently of $k \in \mathcal{G}$. The average equilibrium kinship coefficient for two randomly drawn genes from the whole population is

$$\begin{aligned} f^{\mathrm{eq}}(\mu) &= \tfrac{1}{s^2} \textstyle\sum_{j,k \in \mathcal{G}} f_{jk}^{\mathrm{eq}}(\mu) \\ &= \tfrac{1}{s} \textstyle\sum_{j \in \mathcal{G}} f_{j}^{\mathrm{eq}}(\mu), \end{aligned}$$

and the conditional kinship coefficient at equilibrium is

$$v_j^{\mathrm{eq}}(\mu) = \frac{f_j^{\mathrm{eq}}(\mu) - f^{\mathrm{eq}}(\mu)}{1 - f^{\mathrm{eq}}(\mu)}. \tag{46}$$

Hardy and Vekemans (1999) noticed the connection between $v_j^{\mathrm{eq}}(\mu)$ and measures of spatial autocorrelation. In view of (11), it is also of interest to compare this quantity for small mutation probabilities with the quasi equilibrium approximation v_j of the standardized spatial covariance (38). The following result shows that the two quantities are indeed closely related but not identical:

Theorem 2 *Assume a reproduction scheme MR1-MR3 and spatially invariant migration, as described in Section 5. Let $\mu > 0$ be the mutation probability and $v_j^{eq}(\mu)$ the conditional kinship coefficient (46) under mutation-migration-drift equilibrium for genes in subpopulations at distance $j \in \mathcal{G}$. Then*

$$v_j^{eq}(\mu) = \frac{\tilde{S}_j(\mu)}{1 + \tilde{S}_0(\mu)}, \tag{47}$$

where

$$\tilde{S}_j(\mu) = \frac{1}{s\left(1 - \frac{1}{2N_c}\right)} \sum_{r \in \mathcal{G} \setminus \mathbf{0}} \frac{(1-\mu)^2 \left(\frac{1}{2N_e} - \frac{1}{2N_c}\right) |\hat{m}_r|^2 + \frac{1}{2N_c}}{1 - (1-\mu)^2 |\hat{m}_r|^2} \cos\left(2\pi \sum_{l=1}^{d} \frac{r_l j_l}{s_l}\right). \tag{48}$$

Hence, in the limit of small migration probabilities,

$$\lim_{\mu \to 0} v_j^{eq}(\mu) = \frac{\tilde{S}_j}{1 + \tilde{S}_0}, \tag{49}$$

where

$$\begin{aligned} \tilde{S}_j &= \lim_{\mu \to 0} \tilde{S}_j(\mu) \\ &= \tfrac{1}{s\left(1 - \frac{1}{2N_c}\right)} \textstyle\sum_{r \in \mathcal{G} \setminus \mathbf{0}} \frac{\left(\frac{1}{2N_e} - \frac{1}{2N_c}\right) |\hat{m}_r|^2 + \frac{1}{2N_c}}{1 - |\hat{m}_r|^2} \cos\left(2\pi \textstyle\sum_{l=1}^{d} \frac{r_l j_l}{s_l}\right). \end{aligned} \tag{50}$$

We notice that $\tilde{S}_j$ only differs from S_j by a term $(1-1/(2N_c))^{-1}$. Therefore the limit in (49) is very close to v_j unless N_c is very small, and the two quantities coincide when $N_c = \infty$. It follows from the proof of Theorem 2 in the appendix that $N_c = \infty$ is equivalent to drawing genes from the same subpopulation without rather than with replacement.

We will refer to

$$F_{ST}^{\text{appr}} = \lim_{\mu \to 0} v_{\mathbf{0}}(\mu) = \frac{\tilde{S}_{\mathbf{0}}}{1 + \tilde{S}_{\mathbf{0}}}, \tag{51}$$

and

$$\rho_j^{\text{appr}} = \lim_{\mu \to 0} \frac{f_j^{\text{eq}}(\mu) - f^{\text{eq}}(\mu)}{f_0^{\text{eq}}(\mu) - f^{\text{eq}}(\mu)} = \frac{\tilde{S}_j}{\tilde{S}_{\mathbf{0}}} \tag{52}$$

as approximations of the fixation index and spatial autocorrelation function at equilibrium, based on the mutation-migration-drift approach.

7 Infinite Number of Subpopulations

In this section will let the number of subpopulations grow by requiring that

$$\min(s_1, \ldots, s_d) \to \infty, \tag{53}$$

in (26), so that $s \to \infty$. We will do this in two different ways:

7.1 Isolation by distance models

Wright (1943, 1946) introduced isolation by distance models, where discrete subpopulation membership is replaced by a continuous spatial location. Many authors have since then studied such models, including for instance Malécot (1948) and Rohlf and Schnell (1971).

We model isolation by distance by keeping the total effective population $N_{e,\text{tot}} = sN_e$ fixed, and divide the population into an increasingly large number s of small subpopulations. In order to avoid a conditional genetic drift covariance matrix (29)-(30) that converges to continuous white noise, we put $N = \infty$. For each subpopulation k, this corresponds to having an allele frequency of genes that migrate from k to subpopulation j that equals the allele frequency of gamete pool k, independently of j. From Theorem 2 we deduce that the quasi equilibrium approach is equivalent to mutation-migration-drift equilibrium for translationally invariant migration, when the mutation probability μ tends to 0 and $N = \infty$. Therefore, the formulas below are also analogues of continuous location result of Malécot (1950), Maruyama (1972), Nagylaki (1976), Sawyer (1977) and Barton et al. (2002) as $\mu \to 0$.

We rescale subpopulation membership j as $y = (y_1, \ldots, y_d)$, with $y_l = j_l/s_l$, and denote the corresponding locations $x(y) = (e^{2\pi i y_1}, \ldots, e^{2\pi i y_d})$. Hence, in the limit

$$\mathcal{G}_\infty = [0,1]^d$$

and $\mathcal{X} = \mathbb{S}^d$, the d-dimensional torus, with $\mathbb{S}$ the unit sphere. The migration probabilities of the vector $\boldsymbol{m} = (m_j;\ j \in \mathcal{G})$ become a migration density $m = \{m(y);\ y \in \mathcal{G}_\infty\}$ according to

$$sm_j \to m(y) \text{ as } s \to \infty. \tag{54}$$

This implies

$$\begin{array}{rcl} \hat{m}_r & = & \lim_{s\to\infty} \sum_{j\in\mathcal{G}} m_j \exp(-2\pi i \sum_{l=1}^d r_l j_l/s_l) \\ & = & \int_{[0,1]^d} m(y) \exp(-2\pi i \sum_{l=1}^d r_l y_l) dy \end{array} \tag{55}$$

for all $r \in \mathbb{Z}^d$. Analogously we get from (35) and (36) that

$$S_j \to S(y) = \frac{1}{2N_{e,\text{tot}}} \sum_{r\in\mathbb{Z}^d\setminus\mathbf{0}} \frac{|\hat{m}_r|^2}{1-|\hat{m}_r|^2} \cos\left(2\pi \sum_{l=1}^d r_l y_l\right) \tag{56}$$

and

$$T_j \to T(y) = \frac{1}{2N_{e,\text{tot}}} \sum_{r\in\mathbb{Z}^d} |\hat{m}_r|^2 \cos\left(2\pi \sum_{l=1}^d r_l y_l\right), \tag{57}$$

when $j_l, s_l \to \infty$ in such a way that $y_l = j_l/s_l$ are fixed, and $N = \infty$.

When the number of components of $\boldsymbol{\varepsilon}_{t+1}/\sqrt{P_t(1-P_t)}$ and $\boldsymbol{P}_t^0/\sqrt{P_t(1-P_t)}$ grow, they tend to spatially stationary random fields with a continuous index set and covariance functions $\{\Sigma(z, z+y) = \sigma(y);\ y, z \in \mathcal{G}_\infty\}$ and $\{V(z, z+y) = v(y);\ y, z \in \mathcal{G}_\infty\}$ respectively, with

$$\sigma(y) = \frac{T(y)}{1+S(0)} \tag{58}$$

and

$$v(y) = \frac{S(y)}{1+S(0)}. \tag{59}$$

The variance effective size formula (40) holds as before, the fixation index (39) is rewritten as

$$F_{ST}^{\text{appr}} = v(0) = \frac{S(0)}{1+S(0)}, \tag{60}$$

and the spatial autocorrelation function of $\boldsymbol{P}_t^0/\sqrt{P_t(1-P_t)}$ becomes

$$\rho^{\text{appr}}(y) = \rho^{\text{appr}}(z, z+y) = \frac{v(y)}{v(0)} = \frac{S(y)}{S(0)} \tag{61}$$

in the limit of large s.

7.2 Fixed Local Population Size

We now keep the local effective and census subpopulation sizes N_e and N_c (cf. (28)) fixed, whereas the total subpopulation size N grows, as a consequence of (53). In the limit we get an infinite d-dimensional gitter $\mathcal{G}_\infty = \mathbb{Z}^d$ of subpopulation indexes. We also keep the migration rates m_j between subpopulations k and $k+j$ fixed. It is therefore reasonable to assume that the pairwise distances between k and $k+j$ are kept fixed when $s \to \infty$. This is achieved in the limit by having geographical locations along a d-dimensional gitter, i.e. $x_j = j$ and $\mathcal{X}_\infty = \mathbb{Z}^d$.

Given r and s, define $x = (x_1, \ldots, x_d)$ through $x_l = r_l/s_l$, and let the components of r and s grow in such a way that those of x are kept fixed. Then

$$\begin{array}{rcl} \hat{m}_r & = & \sum_{j \in \mathcal{G}} m_j \exp(-2\pi i \sum_{l=1}^d r_l j_l / s_l) \\ & \to & \sum_{j \in \mathcal{G}_\infty} m_j \exp(-2\pi i \sum_{l=1}^d x_l j_l) \\ & = & \hat{m}(x) \end{array}$$

as all $s_l \to \infty$ according to (53), with $\hat{m} : [0,1]^d \to \mathbb{C}$ the Fourier transform of $\boldsymbol{m} = \{m_j;\, j \in \mathcal{G}_\infty\}$. We can also view (35) and (36) as Riemann sums that we would like to converge to integrals

$$S_j = \int_{[0,1]^d} \frac{\left(\frac{1}{2N_e} - \frac{1}{2N_c}\right) |\hat{m}(x)|^2 + \frac{1}{2N_c}}{1 - |\hat{m}(x)|^2} \cos\left(2\pi \sum_{l=1}^d x_l j_l\right) dx, \qquad (62)$$

and

$$T_j = \int_{[0,1]^d} \left(\left(\frac{1}{2N_e} - \frac{1}{2N_c}\right) |\hat{m}(x)|^2 + \frac{1}{2N_c}\right) \cos\left(2\pi \sum_{l=1}^d x_l j_l\right) dx \qquad (63)$$

respectively. Once convergence is established, it follows that (37), (38), (39) and (41) hold for the limiting infinite population model, with S_j and T_j as defined in (62) and (63).

In order to check convergence, let $J = (J_1, \ldots, J_d) \in \mathbb{Z}^d$ denote a random vector with distribution $\boldsymbol{m}$. Since

$$\hat{m}(x) = 1 - 2\pi i x E(J)^T - \frac{(2\pi)^2}{2} x \mathrm{Cov}(J) x^T + o(|x|^2),$$

as the Euclidean norm $|x|$ of x tends to zero, it follows that

$$|\hat{m}(x)|^2 = 1 + O(|x|^2).$$

Hence, given any $j \in \mathbb{Z}^d$, the integral in (62) has a singularity at $\mathbf{0}$ in dimensions $d = 1, 2$, whereas the integral converges for $d \geq 3$. If we formally

compute the fixation index and autocorrelation function at quasi equilibrium from (39) and (41), the conclusion is that $F_{ST}^{\text{appr}} = \rho_j^{\text{appr}} = 1$ for $d = 1, 2$ and $j \in \mathbb{Z}^d$, whereas F_{ST}^{appr} and ρ_j^{appr} are both less then one for $d \geq 3$.

This indicates that the spatial allele frequency fluctuations on $\mathbb{Z}^d$ are locally isolated for $d = 1, 2$, when there is less space for small distance migration to reach out to remotely distant subpopulations. We can motive this in terms of the mutation-migration-drift equilibrium approach of Section 6: Let τ be the coalescence time between two randomly chosen genes. Slatkin (1991) has shown that

$$\begin{array}{rcl} f_j^{\text{eq}}(\mu) & = & E_j\left((1-\mu)^{2\tau}\right) \approx 1 - 2\mu E_j(\tau), \\ f^{\text{eq}}(\mu) & = & E\left((1-\mu)^{2\tau}\right) \approx 1 - 2\mu E(\tau), \end{array} \tag{64}$$

where E_j and E denote expectation given that the two genes are picked from subpopulations $\mathbf{0}$ and j or randomly from the whole population respectively. The approximations in (64) hold when the mutation probability μ is small. Cox and Durrett (2002) have derived formulas for $E_j(\tau)$ and $E(\tau)$ for the two-dimensional stepping stone model and reproductive scheme MR1-MR3 when $N_e = N_c$ and $s_1 = s_2 = \sqrt{s} \to \infty$. By inserting these expressions for $E_j(\tau)$ and $E(\tau)$ into (64) and (46), it follows that F_{ST}^{appr} and ρ_j^{appr} in (51) and (52) both equal 1 in the limit when $\mu \to 0$ *and* $s \to \infty$ in such a way that $\mu s \log(s) \to 0$. Somewhat related results for the one-dimensional stepping stone model can be found in Durrett and Restrepo (2008).

8 Examples

Example 1 (Island model.) The island model (Wright 1943; Maruyama 1970; Latter 1973) is the simplest possible way of describing a subdivided population. We assume a total migration probability $0 < m \leq 1$ from any subpopulation, with equal probability $m/(s-1)$ of migrating to any other subpopulation. Since migration probabilities are not dependent on the distance between subpopulations, the island model can be incorporated into any dimension $d = 1, 2, \ldots$, with

$$\begin{array}{rcl} m_j & = & (1-m)1_{\{j=0\}} + \frac{m}{s-1}1_{\{j\neq 0\}} \\ & = & (1-m')1_{\{j=0\}} + \frac{m'}{s}, \end{array}$$

where $m' = ms/(s-1)$. This yields

$$\hat{m}_r = (1-m') + m'1_{\{r=0\}}. \tag{65}$$

With the quasi equilibrium approach, we insert (65) into (35), (39) and (41) in order to get

$$F_{ST}^{\text{appr}} = \frac{1}{\frac{s}{s-1}\left(1-(1-m')^2\right)2\tilde{N}+1} \tag{66}$$

for the fixation index and

$$\rho_j^{\text{appr}} = \begin{cases} 1, & j = 0, \\ -1/(s-1), & j \neq 0, \end{cases} \tag{67}$$

for the spatial autocorrelation function. Here $\tilde{N}$ is a local population size, defined as an harmonic mean

$$\frac{1}{\tilde{N}} = \frac{(1-m')^2}{N_e} + \frac{1-(1-m')^2}{N_c}$$

of N_e and N_c. Formula (66) (and refinements thereof) has also been derived by other methods in Hössjer et al. (2011) and Hössjer and Ryman (2012).

The mutation-migration-drift approach yields a spatial autocorrelation function (52) identical to (67) and an expression

$$\begin{aligned} F_{ST}^{\text{appr}} &= \frac{1}{\frac{s}{s-1}(1-(1-m')^2)2\tilde{N}\frac{2N_c-1}{2N_c}+1} \\ &= \begin{cases} \frac{(1-m')^2}{\frac{s}{s-1}(1-(1-m')^2)2N_e+(1-m')^2}, & N_c = \infty, \\ \frac{1}{\frac{s}{s-1}(1-(1-m')^2)(2N_c-1)+1}, & N_c = N_e, \end{cases} \end{aligned} \tag{68}$$

for the fixation index. The lower part of (68) can also be interpreted as having $N_c = N_e$ and drawing genes from the same subpopulation *without* replacement, see for instance Takahata (1983) and Takahata and Nei (1984). The main advantage of drawing genes from the same subpopulation *with* replacement is that F_{ST}^{appr} does not tend to zero as m' tends to 1. See Hössjer et al. (2012) for a more detailed discussion on this topic within a quasi equilibrium framework. □

Example 2 (Stepping stone model.) Kimura (1953) proposed a class of so called stepping stone models, where migration occurs to subpopulations in a local neighbourhood of the present one. For a mathematical treatment of these models, see for instance Kimura and Weiss (1964), Weiss and Kimura (1965) and Durrett (2008).

In one dimension ($d = 1$), the circular stepping stone model is defined as follows: Assume $0 < m \leq 1$ and $0 \leq p \leq 1$. Migration is only possible to the

closest neighbours, with a probability mp of migrating one step to the right and probability $m(1-p)$ of migrating one step to the left, so that

$$m_j = \begin{cases} 1-m, & j=0, \\ mp, & j=1, \\ m(1-p), & j=-1, \\ 0, & \text{otherwise} \end{cases}$$

and

$$\hat{m}_r = (1-m) + m\cos(2\pi r/s) + im(1-2p)\sin(2\pi r/s).$$

We will refer to $p = 0.5$ as the symmetric one-dimensional stepping stone model.

In $d = 2$ dimensions, migration from $(0,0)$ in one generation is possible to a neighbourhood

$$\mathcal{N} = \{(-1,0),(0,-1),(0,1),(1,0)\}$$

of four subpopulations, by moving one step, horizontally or vertically. Define $p_j \geq 0$ for $j \in \mathcal{N}$ so that $\sum_{j\in\mathcal{N}} p_j = 1$. Then put

$$m_j = \begin{cases} 1-m, & j=(0,0), \\ mp_j, & j \in \mathcal{N}, \\ 0, & \text{otherwise}, \end{cases}$$

where $m > 0$ is the total migration probability, as before. Taking the Fourier transform we get

$$\begin{aligned} \hat{m}_r &= 1-m \\ &+ m(p_{(-1,0)} + p_{(1,0)})\cos\left(\tfrac{2\pi r_1}{s_1}\right) + m(p_{(0,-1)} + p_{(0,1)})\cos\left(\tfrac{2\pi r_2}{s_2}\right) \\ &+ im(p_{(-1,0)} - p_{(1,0)})\sin\left(\tfrac{2\pi r_1}{s_1}\right) + im(p_{(0,-1)} - p_{(0,1)})\sin\left(\tfrac{2\pi r_2}{s_2}\right). \end{aligned}$$

We refer to $p_j \equiv 0.25$, $j \in \mathcal{N}$, as the symmetric two-dimensional torus stepping stone model. □

Example 3 (von Mises models.) Starting with the one-dimensional ($d = 1$) case, we let $0 < m \leq 1$ refer to the total migration probability and put

$$p_j = \frac{\exp\left(\kappa\cos(2\pi j/s)\right)}{\sum_{k=1}^{s-1}\exp\left(\kappa\cos(2\pi k/s)\right)} \tag{69}$$

for $j = 1, \ldots, s-1$, where $\kappa \geq 0$, and

$$m_j = \begin{cases} 1-m, & j = 0, \\ mp_j, & j \neq 0. \end{cases} \tag{70}$$

This migration distribution is symmetric, with κ an inverse variance parameter quantifying the magnitude of jumps. We refer to it as the one-dimensional discrete von Mises model. The two extreme choices are the island model ($\kappa = 0$) and the symmetric one-dimensional circular stepping stone model ($\kappa = \infty$).

Let $s \to \infty$, as described in Section 7.1. Both m and κ are kept fixed in (69)-(70), and in order for (54) to hold we must therefore put $m = 1$. In the limit we get a migration density

$$m(y) = \frac{1}{I_0(\kappa)} \exp\left(\kappa \cos(2\pi y)\right), \quad 0 \leq y \leq 1, \tag{71}$$

with $I_0(\kappa) = \int_0^1 \exp\left(\kappa \cos(2\pi y)\right) dy$ a modified Bessel function of the first kind of order 0 (cf. Section 9.6 of Abramowitz and Stegun, 1972). We refer to (71) as the migration density of the continuous one-dimensional von Mises model. Because of the symmetry of $m(\cdot)$ we deduce from (55) that

$$\hat{m}_r = \frac{I_r(\kappa)}{I_0(\kappa)}, \quad r \in \mathbb{Z},$$

where $I_r(\kappa) = \int_0^1 \exp(\kappa \cos(2\pi y)) \cos(2\pi r y) dy$ is a modified Bessel function of the first kind of order r.

In $d = 2$ dimensions, we let κ_1 and κ_2 be non-negative dispersion parameters and define

$$p_j = \frac{\exp\left(\kappa_1 \cos(2\pi j_1/s_1) + \kappa_2 \cos(2\pi j_2/s_2)\right)}{\sum_{k \in \mathcal{G} \setminus (0,0)} \exp\left(\kappa_1 \cos(2\pi k_1/s_1) + \kappa_2 \cos(2\pi k_2/s_2)\right)}, \quad j \neq (0,0).$$

We let $0 < m \leq 1$ refer to the total migration probability and put

$$m_j = \begin{cases} 1-m, & j = (0,0), \\ mp_j, & j \neq (0,0). \end{cases}$$

The island model corresponds to $\kappa_1 = \kappa_2 = 0$ and the symmetric torus stepping stone model to $\kappa_1 = \kappa_2 = \infty$. The continuous analogue, obtained

by putting $m = 1$ and letting $s \to \infty$, as described in Section 7.1, has a migration density with two independent components (71), i.e.

$$m(y) = \frac{1}{I_0(\kappa_1)I_0(\kappa_2)} \exp\left(\kappa_1 \cos(2\pi y_1) + \kappa_2 \cos(2\pi y_2)\right), \quad y \in [0,1]^2,$$

and the Fourier coefficients

$$\hat{m}_r = \frac{I_{r_1}(\kappa_1)I_{r_2}(\kappa_2)}{I_0(\kappa_1)I_0(\kappa_2)}, \quad r \in \mathbb{Z}^2,$$

are obtained from the modified Bessel function of the first kind, as before. □

Example 4 (Hierarchical island model.) Assume $d = 2$ and that all subpopulations $j = (j_1, j_2)$ can be divided into s_1 groups of equal size s_2, so that j_1 is the group number and j_2 the subpopulation number within group j_1. The total migration probability $0 < m \leq 1$ is either within a group, with total probability $m_w = pm$, where $0 \leq p \leq 1$, divided equally $m_w/(s_2 - 1)$ between subpopulations within the group, or with total probability $m_b = (1-p)m$ between groups, with each subpopulations outside the group having the same probability $m_b/((s_1 - 1)s_2)$ of being reached. This model has been treated by Carmelli and Cavalli-Sforza (1976), Sawyer and Felsenstein (1983) and Slatkin and Voelm (1991), and the latter authors referred to it as the hierarchical island model.

It will be convenient to introduce $m'_w = m_w s_2/(s_2 - 1)$ and $m'_b = m_b s_1/(s_1 - 1)$. Then we can write

$$\begin{aligned} m_j &= (1-m)1_{\{j=(0,0)\}} + \tfrac{m_w}{s_2-1}1_{\{j_1=0, j_2\neq 0\}} + \tfrac{m_b}{(s_1-1)s_2}1_{\{j_1\neq 0\}} \\ &= \left(1 - m - \tfrac{m'_w}{s_2}\right)1_{\{j=(0,0)\}} + \left(\tfrac{m'_w}{s_2} - \tfrac{m'_b}{s_1 s_2}\right)1_{\{j_1=0\}} + \tfrac{m'_b}{s_1 s_2}, \end{aligned}$$

so that

$$\hat{m}_r = \left(1 - m - \frac{m'_w}{s_2}\right) + \left(m'_w - \frac{m'_b}{s_1}\right)1_{\{r_2=0\}} + m'_b 1_{\{r=(0,0)\}}.$$

□

Example 5 (Spherically symmetric Gaussian models.) This is a continuous model with migration density

$$m(y) = \sum_{r\in\mathbb{Z}^d} \frac{1}{(2\pi)^{d/2}(\sigma/\sqrt{2})^d} \exp\left(-\frac{\sum_{l=1}^d (y_l + r_l)^2}{\sigma^2}\right), \quad y \in [0,1]^d,$$

so that jumps are taken independently in each direction of $\mathbb{R}^d$ according to Gaussian distributions with mean 0 and variance $\sigma^2/2$, and then wrapped onto the d-dimensional torus. This model has been studied for $d = 1, 2$ by Maruyama (1972). If two genes are located at the same point, their offspring in the next generation have diverged according to a dispersal density

$$d(y) = \sum_{r\in\mathbb{Z}^d} \frac{1}{(2\pi)^{d/2}\sigma^d} \exp\left(-\frac{\sum_{l=1}^d (y_l + r_l)^2}{2\sigma^2}\right), \quad y \in [0,1]^d,$$

with variance σ^2 in each direction before projection onto $[0,1]^d$. In this case

$$\hat{m}_r = \exp\left(-\pi^2\sigma^2 \sum_{l=1}^d y_l^2\right).$$

□

9 Numerical Algorithm and Results

The numerical algorithm for computing F_{ST}^{appr}, N_{eV}^{appr} and $\boldsymbol{\rho}^{\text{appr}} = (\rho_j^{\text{appr}})$ can be summarized as follows:

1. INPUT PARAMETERS: $\boldsymbol{s} = (s_1, \ldots, s_d)$, $\boldsymbol{m} = (m_j)$, N and N_e.
2. COMPUTE $\hat{\boldsymbol{m}} = (\hat{m}_r)$ as the d-dimensional discrete Fourier transform of $\boldsymbol{m}$.
3. COMPUTE $\hat{\boldsymbol{S}} = (\hat{S}_r)$, with
$$\hat{S}_r = 1_{\{r\neq\mathbf{0}\}} \frac{((2N_e)^{-1} - (2N_c)^{-1})\,|\hat{m}_r|^2 + (2N_c)^{-1}}{1 - |\hat{m}_r|^2}.$$
4. COMPUTE $\boldsymbol{S} = (S_j)$ as the inverse d-dimensional discrete Fourier transform of $\hat{\boldsymbol{S}}$.
5. COMPUTE $\boldsymbol{v} = (v_j)$ from (38).

6. COMPUTE OUTPUT PARAMETERS F_{ST}^{appr}, N_{eV}^{appr} and $\boldsymbol{\rho}^{\text{appr}}$ from (39), (40) and (41).

The algorithm for the mutation-migration-drift equilibrium approach is completely analogous, replacing S_j by $\tilde{S}_j$.

Since the properties of F_{ST}^{appr} and N_{eV}^{appr} have been numerically studied in Hössjer et al. (2012) and Hössjer and Ryman (2012), we mainly focus on autocorrelations here. In Figure 1 we plot the autocorrelation function for the circular stepping stone model with $s = 25$ subpopulations, for various choices of m and p. In view of Proposition 3, we can use the terminology of Slatkin and Arter (1991) and refer to these plots as averaged correlograms. It is seen that ρ_j^{appr} for each fixed j depends very little on m and p. An exception is migration probabilities m close to 1, when ρ_j^{appr} exhibits an oscillatory behaviour as a function of j, with period 2. This is more pronounced the closer to symmetric ($p = 0.5$) the migration scheme is.

Averaged correlograms are shown in Figure 2 for the one-dimensional von Mises-model with $s = 25$. It is seen that $j \to \rho_j^{\text{appr}}$ depends a lot on the concentration parameter κ, but very little on m, N_e and N_c.

There are only two values of the autocorrelation function for the hierarchical island model; the autocorrelation ρ_w^{appr} between subpopulations within the same group, and the autocorrelation ρ_b^{appr} for subpopulations belonging to different groups. Figure 3 displays both of these quantities as well as the fixation index F_{ST}^{appr} for a grid of 5×5 subpopulations. On the x-axis of each plot is p, the probability of migrating to a subpopulation within the same group. The within autocorrelations increase with p, whereas the between autocorrelations decrease with p. The migration probability m affects the fixation index a lot, but hardly at all the autocorrelations.

In Figure 4 we plot the fixation index and autocorrelations for different lags j for the torus stepping stone model when $s_1 = s_2 = 5$. As for the one-dimensional stepping stone model, the autocorrelations depend very little on m and $(p_j;\, j \in \mathcal{N})$ except for large values of m, whereas the fixation index varies quite a lot with m.

Figure 5 depicts convergence to an isolation by distance model, as described in Section 7.1. More specifically, the plots illustrate how the one-dimensional von Mises models with $\kappa = 5$ converge to its continuous analogue in that the autocorrelations ρ_j^{appr} approach $\rho^{\text{appr}}(y)$ quite rapidly as $s \to \infty$ and $j/s \to y$,

Figure 6, on the other hand, illustrates the theory of Section 7.2 when $d = 1$, for the symmetric circular stepping stone model. The autocorrelations ρ_j^{appr} converge to 1 when $s \to \infty$ while j is kept fixed.

10 Discussion

In this paper we developed a general methodology for computing approximations of autocorrelations, fixation indexes and variance effective population sizes under quasi equilibrium. We considered a class of population genetic models with geographic substructure for which migration is translationally invariant and developed a fast algorithm based on Fourier transforms. This class of models includes, for instance, circular and torus stepping stone models, von Mises models, hierarchical island models and Gaussian models. We also considered limits of continuous spatial location and infinite gitters of subpopulations. We established connections between our framework and conditional kinship coefficients on one hand and Moran's autocorrelation functions on the other. Finally, we proved new results for kinship coefficients between subpopulations under mutation-migration-drift equilibrium, and showed how they relate to the approximate quasi equilibrium formulas.

Spatial autocorrelation functions (or transformations thereof) can naturally be estimated from data. Several authors have discussed the degree of information contained in the autocorrelations, see for instance Slatkin and Arter (1991), Rousset (1997, 2000) and Sokal et al. (1989, 1997). It follows from the results of Barbujani (1997) and Hardy and Vekemans (1999), that the fixation index and spatial correlation functions contain complementary information.

Theorem 1 is consistent with these findings, showing that both the fixation index and autocorrelations are needed in order to compute all $\boldsymbol{S} = (S_j;\, j \in \mathcal{G})$ uniquely. This suggests that both should be used for improving estimates of genetic model parameters. On the other hand, the conditional kinship coefficients contain no *additional* information for the class of models studied in this paper, since

$$v_j = F_{ST}^{\text{appr}} \rho_j^{\text{appr}}.$$

Our numerical results indicate that the autocorrelations are quite insensitive to variations of the overall migration probability

$$m = \sum_{j \in \mathcal{G} \setminus \mathbf{0}} m_j.$$

On the other hand, the averaged correlograms are useful in determining whether migration is local or extends over larger regions. Our numerical results for the one-dimensional von Mises model (cf. Figure 2) reveal that the local behaviour of ρ_j^{appr} around the origin contains substantially more information about the concentration parameter κ than does the genetic patch size, i.e. the Euclidean distance $|j|$ when ρ_j^{appr} is close to zero.

Several extensions of our work are possible. For instance, it is of interest to study the quasi equilibrium behavior for different settings, for instance multilocus autocorrelation functions defined between individuals rather than subpopulations (Smousse and Peakall, 1999), or local autocorrelations functions (Sokal et al., 1998).

Acknowledgment

The author wants to thank Nils Ryman for fruitful discussions and for providing valuable references. Ola Hössjer's research was financially supported by the Swedish Research Council, contract nr. 621-2008-4946, and the Gustafsson Foundation for Research in Natural Sciences and Medicine.

Appendix

Proof of Proposition 1. Equation (9) follows immediately from the definition of V_{tjk}. In order to prove (10), we deduce from (7) that

$$\begin{aligned} E(f_{t,jk}|P_t) &= E(P_{tj}P_{tk}|P_t) + E\left((1-P_{tj})(1-P_{tk})|P_t\right) \\ &= P_t^2 + (2P_t-1)E(P_{tj}-P_t|P_t) + \mathrm{Cov}(P_{tj},P_{tk}) \\ &+ (1-P_t)^2 + (2P_t-1)E(P_{tk}-P_t|P_t) + \mathrm{Cov}(1-P_{tj},1-P_{tk}) \\ &= f_t + (2P_t-1)E(P_{tj}+P_{tk}-2P_t|P_t) + 2\mathrm{Cov}(P_{tj},P_{tk}) \\ &= f_t + (2P_t-1)E(P_{tj}+P_{tk}-2P_t|P_t) + 2P_t(1-P_t)V_{tjk} \end{aligned}$$

and hence

$$\begin{aligned} V_{tjk} &= \frac{E(f_{t,jk}|P_t)-f_t+(1-2P_t)E(P_{tj}+P_{tk}-2P_t|P_t)}{2P_t(1-P_t)} \\ &= \frac{E(f_{t,jk}|P_t)-f_t+(1-2P_t)E(P_{tj}+P_{tk}-2P_t|P_t)}{1-f_t}. \end{aligned}$$

When $\{P_{tj};\, j\in\mathcal{G}\}$ are exchangeable, $E(P_{tj}^0|P_t) = E(P_{tj}-P_t|P_t)$ is independent of j. Since

$$\sum_{j\in\mathcal{G}} \gamma_j E(P_{tj}^0|P_t) = E\left(\sum_{j\in\mathcal{G}} \gamma_j P_{tj}^0|P_t\right) = E(0|P_t) = 0,$$

$E(P_{tj}^0|P_t) = 0$ must hold for all j, and so (11) follows from (10). □

Rewriting (17) as an infinite series. We can rewrite (17) as

$$\boldsymbol{V} = \sum_{l=0}^{\infty} \boldsymbol{B}^l \tilde{\boldsymbol{\Sigma}} (\boldsymbol{B}^T)^l = \boldsymbol{Q}\left(\sum_{l=0}^{\infty} \boldsymbol{\Lambda}_B^l \boldsymbol{Q}^{-1}\tilde{\boldsymbol{\Sigma}}\boldsymbol{Q}^{-T}(\boldsymbol{\Lambda}_B^T)^l\right)\boldsymbol{Q}^T, \qquad (72)$$

where

$$\boldsymbol{B} = \boldsymbol{Q}\boldsymbol{\Lambda}_B\boldsymbol{Q}^{-1}. \tag{73}$$

If $\boldsymbol{B}$ is symmetric, we can choose $\boldsymbol{\Lambda}_B = \mathrm{diag}(\lambda_j;\, j \in \mathcal{G})$ to be diagonal in (73), with the elements of $\mathcal{G}$ ordered so that the real-valued eigenvalues of $\boldsymbol{B}$ appear along the diagonal in decreasing order

$$1 = \lambda_0 > |\lambda_1| \geq \ldots \geq |\lambda_{s-1}| \geq 0 \tag{74}$$

with respect to their moduli. The first inequality in (74) is strict since $\boldsymbol{B}$ corresponds to an irreducible Markov chain. In general, when $\boldsymbol{B}$ is not be symmetric, we can use the Perron Frobenius Theorem (see for instance Cox and Miller, 1965) and interpret (73) as the Jordan canonical form of the non-negative matrix $\boldsymbol{B}$, with $\boldsymbol{\Lambda}_B$ an upper triangular matrix of order s with the possibly complex valued eigenvalues along the diagonal satisfying (74), with $|\lambda_j|$ the modulus of the complex number λ_j. In any case, the leftmost column of $\boldsymbol{Q}$ is proportional to a right eigenvector $\mathbf{1}$ of $\boldsymbol{B}$, and the first row of $\boldsymbol{Q}^{-1}$ is proportional to a left eigenvector $\boldsymbol{\gamma}$ of $\boldsymbol{B}$, both with eigenvalue $\lambda_0 = 1$. □

Eigenvectors and eigenvalues of $\boldsymbol{B}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{V}$. We first prove (33). Indeed, for any multiindex $k = (k_1, \ldots, k_d) \in \mathcal{G}$,

$$\begin{aligned}
(\boldsymbol{B}\boldsymbol{q}_r)_k &= \textstyle\sum_{j\in\mathcal{G}} B_{kj} q_{rj} \\
&= \textstyle\sum_{j\in\mathcal{G}} M_{jk} q_{rj} \\
&= \textstyle\sum_{j\in\mathcal{G}} m_{k-j} \exp\left(2\pi i \sum_{l=1}^d r_l j_l / s_l\right) / \sqrt{s} \\
&= q_{rk} \textstyle\sum_{j\in\mathcal{G}} m_{k-j} \exp\left(2\pi i \sum_{l=1}^d r_l (j_l - k_l) / s_l\right) \\
&= \hat{m}_r q_{rk}.
\end{aligned} \tag{75}$$

Therefore (73) holds with $\boldsymbol{Q}$ having columns $\boldsymbol{q}_r$ and $\boldsymbol{\Lambda}_B$ diagonal with eigenvalues $\lambda_r = \hat{m}_r$ for all $r \in \mathcal{G}$. It can be shown, analogously to (33), that

$$\boldsymbol{\Sigma}\boldsymbol{q}_r = \hat{\sigma}_r \boldsymbol{q}_r \tag{76}$$

and

$$\boldsymbol{V}\boldsymbol{q}_r = \hat{v}_r \boldsymbol{q}_r, \tag{77}$$

for all $r \in \mathcal{G}$, where $\hat{\sigma}_r$ and $\hat{v}_r$ are the Fourier coefficients of $\boldsymbol{\sigma}$ and $\boldsymbol{v}$, defined analogously to (34). Since $\sigma_{-j} = \sigma_j$ and $\lambda_{-j} = \lambda_j$, it follows that $\hat{\sigma}_r$ and $\hat{v}_r$ are both real. □

Proof of Theorem 1. In order to verify (38), we notice, since $\{\boldsymbol{q}_r\}_{r\in\mathcal{G}}$ is an orthonormal system of vectors with respect to the scalar product (32), that

$$\boldsymbol{v} = \sum_{r\in\mathcal{G}} (\boldsymbol{v}, \boldsymbol{q}_r)\boldsymbol{q}_r = \sum_{r\in\mathcal{G}} \hat{v}_r \boldsymbol{q}_r / \sqrt{s}, \tag{78}$$

and hence, taking the j:th coordinate of (78),

$$\begin{aligned} v_j &= \textstyle\sum_{r\in\mathcal{G}} \hat{v}_r q_{rj}/\sqrt{s} \\ &= \textstyle\sum_{r\in\mathcal{G}} \hat{v}_r \exp\left(2\pi i \sum_{l=1}^d r_l j_l/s_l\right)/s \\ &= \textstyle\sum_{r\in\mathcal{G}} \hat{v}_r \cos\left(2\pi \sum_{l=1}^d r_l j_l/s_l\right)/s, \end{aligned} \tag{79}$$

where in the last step we used that $\boldsymbol{v}$ is symmetric and hence all $\hat{v}_r$ real-valued. We may also deduce (79) directly, without using (78), since it is the inverse discrete Fourier transform of $\{\hat{v}_r\}$.

The next step is to compute $\hat{v}_r$ more explicitly as functions of $\hat{m}_r$ and $\hat{\sigma}_r$. We notice that (76) implies $\boldsymbol{\Sigma} = \boldsymbol{Q}\boldsymbol{\Lambda}_\Sigma\boldsymbol{Q}^{-1}$, with $\boldsymbol{\Lambda}_\Sigma = \mathrm{diag}(\hat{\sigma}_r)$. Since (18) simplifies to $\tilde{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma} - \hat{\sigma}_0\boldsymbol{q}_0\boldsymbol{q}_0^T$, we can write $\tilde{\boldsymbol{\Sigma}} = \boldsymbol{Q}\boldsymbol{\Lambda}_{\tilde{\Sigma}}\boldsymbol{Q}^{-1}$, with $\boldsymbol{\Lambda}_{\tilde{\Sigma}} = \mathrm{diag}(\hat{\sigma}_r 1_{\{r\neq\mathbf{0}\}})$. Analogously, it follows from (77) that $\boldsymbol{V} = \boldsymbol{Q}\boldsymbol{\Lambda}_V\boldsymbol{Q}^{-1}$, with $\boldsymbol{\Lambda}_V = \mathrm{diag}(\hat{v}_r)$. But $\boldsymbol{Q}^{-1} = \boldsymbol{Q}^T$, since $\{\boldsymbol{q}_r\}$ forms an orthonormal system of eigenvectors, so that equation (72) implies

$$\boldsymbol{V} = \boldsymbol{Q}\left(\sum_{l=0}^{\infty} \boldsymbol{\Lambda}_B^l \boldsymbol{\Lambda}_{\tilde{\Sigma}} (\boldsymbol{\Lambda}_B^T)^l\right)\boldsymbol{Q}^{-1} =: \boldsymbol{Q}\boldsymbol{\Lambda}_V\boldsymbol{Q}^{-1},$$

where $\boldsymbol{\Lambda}_V = \mathrm{diag}(\hat{v}_r)$ has elements

$$\hat{v}_r = 1_{\{r\neq\mathbf{0}\}} \sum_{l=0}^{\infty} \hat{m}_r^l \hat{\sigma}_r \bar{\hat{m}}_r^{\,l} = 1_{\{r\neq\mathbf{0}\}} \frac{\hat{\sigma}_r}{1-|\hat{m}_r|^2}. \tag{80}$$

On the other hand, taking the Fourier transform of (30) and using $\widehat{(\boldsymbol{m}^- * \boldsymbol{m})}_r = |\hat{m}_r|^2$, we find that

$$\hat{\sigma}_r = (1-v_0)\left\{\left(\frac{1}{2N_e} - \frac{1}{2N_c}\right)|\hat{m}_r|^2 + \frac{1}{2N_c}\right\} \tag{81}$$

for all $r \in \mathcal{G}$. Inserting (80) and (81) into (79), we get

$$v_j = (1-v_{\mathbf{0}})S_j. \tag{82}$$

In particular, inserting $j = \mathbf{0}$ in (82), it follows that $v_{\mathbf{0}} = S_{\mathbf{0}}/(1+S_{\mathbf{0}})$. Resubstituting this result back into (82), we obtain (38).

In order to prove (37), we write

$$\sigma_j = \sum_{r\in\mathcal{G}} \hat{\sigma}_r \cos\left(2\pi \sum_{l=1}^d r_l j_l/s_l\right)/s \tag{83}$$

as the inverse discrete Fourier transform of $\{\hat{\sigma}_r\}$, analogously to (79), and then insert (81) into (83) and use (38) with $j = 0$.

Equation (39) follows directly from (21), (24) and (31), equation (41) from (20) and (31). Finally, for equation (40) we make use of (5), (22), (24), (29) and (81) to deduce that

$$\begin{aligned} N_{eV}^{\text{appr}} &= 1/(2\boldsymbol{\gamma}\boldsymbol{\Sigma}\boldsymbol{\gamma}^T) \\ &= s^2/(2\mathbf{1}\boldsymbol{\Sigma}\mathbf{1}^T) \\ &= s/(2\textstyle\sum_{j\in\mathcal{G}}\sigma_j) \\ &= s/(2\hat{\sigma}_0) \\ &= s/(2(1-F_{ST}^{\text{appr}})\left\{\left(\tfrac{1}{2N_e}-\tfrac{1}{2N_c}\right)|\hat{m}_0|^2+\tfrac{1}{2N_c}\right\}) \\ &= sN_e/(1-F_{ST}^{\text{appr}}), \end{aligned}$$

where in the last step we used that $\hat{m}_0 = \sum_{j\in\mathcal{G}} m_j = 1$. □

Proof of Proposition 3. Formula (44) follows from the definition of V_{tik} in (6), and by rewriting (43) as

$$E(I_{tj}|P_t)^* = \frac{s\sum_{i,k\in\mathcal{G}} w_{jik}\frac{E((P_{ti}-P_t)(P_{tk}-P_t)|P_t)}{P_t(1-P_t)}}{\sum_{i,k\in\mathcal{G}} w_{jik}\sum_{i\in\mathcal{G}}\frac{E((P_{ti}-P_t)^2|P_t)}{P_t(1-P_t)}}.$$

We define I_j^{appr} by replacing V_{tik} and V_{tii} with V_{ik} and V_{ii} in the numerator and denominator of (44) respectively. There are $2s$ terms in the numerator of I_j^{appr} with $w_{jik}=1$, and for all of them $V_{ik}=v_j$ since $k-i=\pm j$ and $v_j = v_{-j}$. (Unless $j=-j$, then the number of terms is s.) Moreover, there are s terms in the denominator of I_j^{appr} with $V_{ii}=v_0$. This yields

$$I_j^{\text{appr}} = \frac{s\cdot 2s\cdot v_j}{2s\cdot s\cdot v_0} = \frac{v_j}{v_0} = \rho_j^{\text{appr}},$$

thus proving (45). □

Proof of Theorem 2. For simplicity of notation, we drop superscript eq and argument μ, and write $f_{jk}^{\text{eq}}(\mu)=f_{jk}$, $f_j^{\text{eq}}(\mu)=f_j$ and $f^{\text{eq}}(\mu)=f$. Let $\gamma_1=(1-\mu_1)^2$, $\gamma_2=(1-\mu_2)^2$ and $\gamma=\gamma_1\gamma_2=(1-\mu)^2$ be the probabilities that none of two genes mutates during gamete formation, fertilization and the whole reproduction cycle respectively.

Consider two genes draw at random from subpopulations $\mathbf{0}$ and j, with replacement if $j=\mathbf{0}$. Following Malécot (1950), we can set up equilibrium equations

$$\begin{aligned} f_j &= f_{0j} \\ &= 1_{\{j\neq\mathbf{0}\}}\gamma_2\left(\textstyle\sum_{k\neq l}B_{0k}B_{jl}\gamma_1 f_{l-k}+\sum_k B_{0k}B_{jk}\left(\frac{\gamma_1}{2N_e}+\gamma_1(1-\frac{1}{2N_e})\frac{f_0-\frac{1}{2N_c}}{1-\frac{1}{2N_c}}\right)\right) \\ &+ 1_{\{j=\mathbf{0}\}}\left\{\tfrac{1}{2N_c}+(1-\tfrac{1}{2N_c})\gamma_2\left(\textstyle\sum_{k\neq l}B_{0k}B_{0l}\gamma_1 f_{l-k}\right.\right. \\ &\qquad\left.\left.+\sum_k B_{0k}B_{0k}\left(\frac{\gamma_1}{2N_e}+\gamma_1(1-\frac{1}{2N_e})\frac{f_0-\frac{1}{2N_c}}{1-\frac{1}{2N_c}}\right)\right)\right\} \end{aligned} \tag{84}$$

In order to motivate (84), we start with the $j \neq \mathbf{0}$ term. The probability is γ_2 that none of the two gametes in post-migration pools $\mathbf{0}$ and j mutates during fertilization. Given this, the probability is $B_{\mathbf{0}k}B_{jl}\gamma_1$ that the parental genes of the two gametes from post-migration pools $\mathbf{0}$ and j originate from subpopulations k and l, and that none of them mutated during gamete formation. If, so the probability for them to be IBD is f_{l-k}, by the equilibrium assumption. If $k = l$, the probability is $\gamma_1/(2N_e)$ that the two parental gametes (which are different since the gamete pools are infinite) originate from the same parental gene and that none of them mutated. The probability is $\gamma_1(1 - 1/(2N_e))$ that they originate from different parental genes and that none of them of them mutates. Given this, the probability is

$$\frac{f_0 - \frac{1}{2N_c}}{1 - \frac{1}{2N_c}}, \tag{85}$$

for the two parental genes to be IBD. The probability in (85) is the same as the IBD probability of two genes that are drawn *without* replacement from the same subpopulation.

When $j = \mathbf{0}$, both genes are drawn from the same subpopulation, the probability is $1/(2N_c)$ for them to be the same, and the probability is $1 - 1/(2N_c)$ for them to be different. In the latter case we argue as above when $j = \mathbf{0}$.

A little algebra shows that

$$\frac{1}{2N_e} + \left(1 - \frac{1}{2N_e}\right)\frac{f_0 - \frac{1}{2N_c}}{1 - \frac{1}{2N_c}} = f_0 + \frac{\frac{1}{2N_e} - \frac{1}{2N_c}}{1 - \frac{1}{2N_c}}(1 - f_0) =: f_0 + \alpha(1 - f_0).$$

Hence we can rewrite (84) as

$$\begin{aligned} f_j &= \gamma\left(\textstyle\sum_{k,l} B_{0k}B_{jl}f_{l-k} + \alpha(1 - f_0)\sum_k B_{0k}B_{jk}\right) + \frac{A1_{\{j=0\}}}{2N_c} \\ &= \gamma\left((\boldsymbol{m}^- * \boldsymbol{m} * \boldsymbol{f})_j + \alpha(1 - f_0)(\boldsymbol{m}^- * \boldsymbol{m})_j\right) + \frac{A1_{\{j=0\}}}{2N_c}, \end{aligned} \tag{86}$$

with $\boldsymbol{m}^-$ and $\boldsymbol{m}$ as defined in Section 5, $\boldsymbol{f} = (f_j; j \in \mathcal{G})$, and

$$A = 1 - \gamma\left((\boldsymbol{m}^- * \boldsymbol{m} * \boldsymbol{f})_0 + \alpha(1 - f_0)(\boldsymbol{m}^- * \boldsymbol{m})_0\right).$$

Putting $j = 0$ in (86) and solving the resulting equation for A, we find that

$$f_0 = (1 - A) + \frac{A}{2N_c} \Leftrightarrow A = \frac{1 - f_0}{1 - \frac{1}{2N_c}}. \tag{87}$$

Taking the Fourier transform of both sides of (84), solving for $\hat{f}_r$ and using the definitions of α and A, we then obtain

$$\begin{aligned} \hat{f}_r &= \tfrac{\gamma|\hat{m}_r|^2\alpha(1-f_0)+\frac{A}{2N_c}}{1-\gamma|\hat{m}_r|^2} \\ &= \tfrac{1-f_0}{1-\frac{1}{2N_c}} \cdot \tfrac{\gamma|\hat{m}_r|^2\left(\frac{1}{2N_e}-\frac{1}{2N_c}\right)+\frac{1}{2N_c}}{1-\gamma|\hat{m}_r|^2} \end{aligned} \tag{88}$$

for all $r \in \mathcal{G}$. We then use (87) and take the inverse Fourier transform of (88) in order to get an explicit expression

$$f_j = \frac{1-f_0}{s(1-\frac{1}{2N_c})} \sum_{r\in\mathcal{G}} \frac{\gamma|\hat{m}_r|^2\left(\frac{1}{2N_e}-\frac{1}{2N_c}\right)+\frac{1}{2N_c}}{1-\gamma|\hat{m}_r|^2} \cos\left(2\pi \sum_{l=1}^{d} \frac{r_l j_l}{s_l}\right) \tag{89}$$

for the equilibrium kinship coefficient. By averaging (89) over j and using $\hat{m}_r = 1$, we find that

$$\begin{aligned} f &= \tfrac{1}{s}\textstyle\sum_{j\in\mathcal{G}} f_j \\ &= \tfrac{1-f_0}{s(1-\frac{1}{2N_c})} \cdot \tfrac{\gamma\left(\frac{1}{2N_e}-\frac{1}{2N_c}\right)+\frac{1}{2N_c}}{1-\gamma}. \end{aligned} \tag{90}$$

When $j = 0$ is inserted into (89), we can first solve for f_0 and then compute

$$\frac{1}{1-f_0} = 1 + \frac{1}{s(1-\frac{1}{2N_c})} \sum_{r\in\mathcal{G}} \frac{\gamma|\hat{m}_r|^2\left(\frac{1}{2N_e}-\frac{1}{2N_c}\right)+\frac{1}{2N_c}}{1-\gamma|\hat{m}_r|^2} \cos\left(2\pi \sum_{l=1}^{d} \frac{r_l j_l}{s_l}\right). \tag{91}$$

Combining (89), (90) and (91) we arrive at

$$\begin{aligned} v_j^{\mathrm{eq}}(\mu) &= \tfrac{f_j - f}{1-f} \\ &= \frac{\frac{1}{s(1-\frac{1}{2N_c})}\sum_{r\in\mathcal{G}\setminus 0} \frac{\gamma|\hat{m}_r|^2\left(\frac{1}{2N_e}-\frac{1}{2N_c}\right)+\frac{1}{2N_c}}{1-\gamma|\hat{m}_r|^2} \cos\left(2\pi\sum_{l=1}^{d}\frac{r_l j_l}{s_l}\right)}{\frac{1}{1-f_0} - \frac{1}{s(1-\frac{1}{2N_c})}\cdot\frac{\gamma\left(\frac{1}{2N_e}-\frac{1}{2N_c}\right)+\frac{1}{2N_c}}{1-\gamma}} \\ &= \tfrac{\tilde{S}_j(\mu)}{1+\tilde{S}_0(\mu)}, \end{aligned} \tag{92}$$

where

$$\tilde{S}_j(\mu) = \frac{1}{s(1-\frac{1}{2N_c})} \sum_{r\in\mathcal{G}\setminus\mathbf{0}} \frac{\gamma|\hat{m}_r|^2\left(\frac{1}{2N_e}-\frac{1}{2N_c}\right)+\frac{1}{2N_c}}{1-\gamma|\hat{m}_r|^2} \cos\left(2\pi \sum_{l=1}^{d} \frac{r_l j_l}{s_l}\right).$$

But (92) is identical to (47). Formula (49) then follows easily by letting $\mu \to 0$. □

References

Abramowitz, M. and Stegun, I.A. (1972). *Handbook of Mathematical Functions. With Formulas, Graphs, and Mathematical Tables.* Tenth printing, U.S. Government Printing Office Washington, D.C.

Barbujani, G. (1987). Autocorrelation of gene frequencies under isolation by distance. *Genetics* **117**, 777-782.

Barton, N.H., Depaulis, F. and Etheridge, A.M. (2002). Neutral evolution in spatially continuous populations. *Theor. Pop. Biol.* **61**, 31-48.

Carmelli, D. and Cavalli-Sforza, L.L. (1976). Some models of population structure and evolution. *Theor. Pop. Biol.* **9**, 329-359.

Cox, D.R. and Miller, H.D. (1965). *The theory of stochastic processes.* Methuen & Co Ltd, London.

Cox, J.T. and Durrett, R. (2002). The stepping stone model: New formulas expose old myths. *Ann. Appl. Prob.* **12**(4), 1348-1377.

Durrett, R. (2008). *Probability models for DNA sequence evolution*, 2nd ed., Springer, New York.

Durrett, R. and Restrepo. M. (2008). One-dimensional stepping stone modes, sardine genetics and Brownian local time. *Ann. Appl. Prob.* **18**(1), 334-358.

Felsenstein, J. (1971). Inbreeding and variance effective numbers in populations with overlapping generations. *Genetics* **68**, 581-597.

Hare, M.P., Nunney, L., Schwartz, M.K., Ruzzante, D.E., Burford, M., Waples, R., Ruegg, K. and Palstra, F. (2011). Understanding and estimating effective population size for practical applications in marine species management. *Conservational Biology* **25**(3), 438-449.

Hardy, O.J. and Vekemans, X. (1999). Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics models. *Heredity* **83**, 145-154.

Hardy, O.J. and Vekemans, X. (2002). SPAGeDI: a versatile computer program to analyse spatial genetic structure at the individual or population model. *Molecular Ecology Notes* **2**, 618-620.

Hössjer, O., Jorde, P.E. and Ryman, N. (2012). Quasi Equilibrium Approximations of the Fixation Index of the Island Model under Neutrality. Report 2012:5, Mathematical Statistics, Stockholm University. Under revision for *Theoretical Population Biology*.

Hössjer, O. and Ryman, N. (2012). Quasi Equilibrium, Variance Effective Population Size and Fixation Index for Models with Spatial Structure. Re-

port 2012:4, Mathematical Statistics, Stockholm University. Under revision for *Journal of Mathematical Biology*.

Kimura, M. (1953). 'Stepping stone' model of population. *Ann. Rep. Natl. Inst. Genet. Japan* **3**, 62-63.

Kimura, M. and Weiss, G.H. (1964). The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* **61**, 763-771.

Latter, B.D.H. (1973). The island model of population differentiation: a general solution. *Genetics* **73**, 147-157.

Latter, B.D.H. and Sved, J.A. (1981). Migration and mutation in stochastic models of gene frequency change. II. Stochastic migration with a finite number of islands. *J. Math. Biology* **13**, 95-104.

Malécot, G. (1948). *Les Mathématiques de l'Hérédité*. Masson et Cie, Paris.

Malécot, G. (1950). Quelques schémas probabilistes sur la variabilité des populations naturelles. *Annales de l'Université de Lyon A* **13**, 37-60.

Malécot, G. (1951). Un treatment stochastique des problemès linairés (mutation, linkage, migration) en géneétique de populations. *Annales de l'Université de Lyon A* **14**, 79-117.

Maruyama, T. (1970). Effective number of alleles in subdivided populations. *Theor. Pop. Biol.* **1**, 273-306.

Maruyama, T. (1972). Rate of decrease of genetic variability in a two-dimensional continuous population of finite size. *Genetics* **70**, 639-651.

Moran, P.A.P. (1950). Notes on continuous stochastic phenomena. *Biometrika* **37**(1), 17-23.

Morton, N.E. (1973). Kinship and population structure. In: Morton, N.E. (ed.) *Genetic Structure of Populations*, pp. 66-69. University of Hawaii Press, Honolulu.

Nagylaki, T. (1976). The decay of genetic variability in geographically structured populations. II. *Theor. Pop. Biol.* **10**, 70-82.

Nei, M. (1973). Genetic distance between populations. *American Naturalist* **106**, 283-292.

Nei, M. (1977). Molecular evolution and population genetics. North-Holland Publishing Company, Amsterdam.

Nei, M., Chakravarti, A. and Tateng, Y. (1977). Mean and variance of F_{ST} in a finite number of incompletely isolated populations. *Theor. Pop. Biol.* **11**, 291-306.

Palstra, F.P. and Ruzzante, D.E. (2008). Genetic estimates of contemporary effective population size: what can they tell us about the importance of

genetic stochasticity for wild populations persistence? *Molecular Ecology* **17**, 3428-3447.

Rohlf, F.J. and Schnell, G.D. (1971). An investigation of the isolation-by-distance model. *American Naturalist* **105**, 295-324.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance.

Rousset, F. (2000). Genetic differentiation between individuals. *J. Evol. Biol.* **13**, 58-62.

Rousset, F. (2001). Inferences from spatial population genetics. In *Handbook of Statistical Genetics*, eds. Balding, B.J., Bishop, M. and Cannings, C., 239-269.

Rousset, F. (2002). Inbreeding and relatedness coefficients: what do they measure? *Heredity* **88**, 371-380.

Sawyer, S. (1976). Results for the stepping stone model for migration in population genetics. *Ann. Prob.* **4**, 699-728.

Sawyer, S. (1977). Asymptotic properties of the equilibrium probability of identity in a geographically structured population. *Adv. Appl. Prob.* **9**, 268-282.

Sawyer, S. and Felsenstein, J. (1983). Isolation by distance in a hierarchically clustered population. *J. Appl. Prob.* **20**, 1-10.

Slatkin, M. (1991). Inbreeding coefficients and coalescent times. *Genet. Res. Cambridge* **58**, 167-175.

Slatkin, M. and Arter, H.E. (1991). Spatial autocorrelation methods in population genetics. *American Naturalist* **138**(2), 499-517.

Slatkin, M. and Voelm, L. (1991). F_{ST} in a hierarchical model. *Genetics* **127**, 627-629.

Smouse, P.E. and Peakall, R. (1998). Spatial autocorrelation analysis of individual allele and multilocus genetic structure. *Heredity* **82**, 561-573.

Sokal, R.R., Jacques, M.J. and Wooten, M.C. (1989). Spatial autocorrelation analysis of migration and selection. *Genetics* **121**, 845-855.

Sokal, R.R. and Oden, N.L. (1978). Spatial autocorrelation in biology. *Biol. J. Linnean Society* **10**, 199-228.

Sokal, R.R., Oden, N.L. and Thomson, B.A. (1997). A simulation study of microevolutionary inferences by spatial autocorrelation analysis. *Biol. J. Linnean Society* **60**, 73-93.

Sokal, R.R., Oden, N.L. and Thomson, B.A. (1998). Local spatial autocorrelation in biological variables. *Biol. J. Linnean Society* **65**, 41-62.

Sved, J.A. and Latter, B.D.H. (1977). Migration and mutation in stochastic models of gene frequency change. *J. Math. Biology* **5**, 61-73.

Takahata, N. 1983. Gene identity and genetic differentiation of populations in the finite island model. *Genetics* **104**, 497-512.

Takahata, N. and Nei, M. 1984. F_{ST} and G_{ST} statistics in the finite island model. *Genetics* **107**, 501-504.

Vekemans, X. and Hardy, O.J. (2004). New insights from fine-scale genetic structure analyses in plant populations. *Molecular Ecology* **13**, 921-935.

Wang, J. and Caballero, A. (1999). Developments in predicting the effective size of subdivided populations. *Heredity* **82**, 212-226.

Waples, R.S. and Yokota, M. (2007). Temporal estimates of effective population size in species with overlapping generations. *Genetics* **175**, 219-233

Weiss, G.H. and Kimura, M. (1965). A mathematical analysis of the stepping stone model of genetic correlation. *J. Appl. Prob.* **2**, 129-149.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics* **16**, 97-159.

Wright, S. (1938). Size of population and breeding structure in relation to evolution. *Science* **87**, 430-431.

Wright, S. (1943). Isolation by distance. *Genetics* **28**, 114-138.

Wright, S. (1946). Isolation by distance under diverse systems of mating. *Genetics* **31**, 39-59.

Wright, S. (1951). The general structure of populations. *Annals of Eugenics* **15**, 323-354.

Zhao, R., Xia, H. and Lu, B-R. (2009). Fine-scale genetic structure enhances biparental inbreeding by promoting mating events between more related individuals in wild soybean populations. *Am. J. of Botany* **96**(6), 1138-1147.

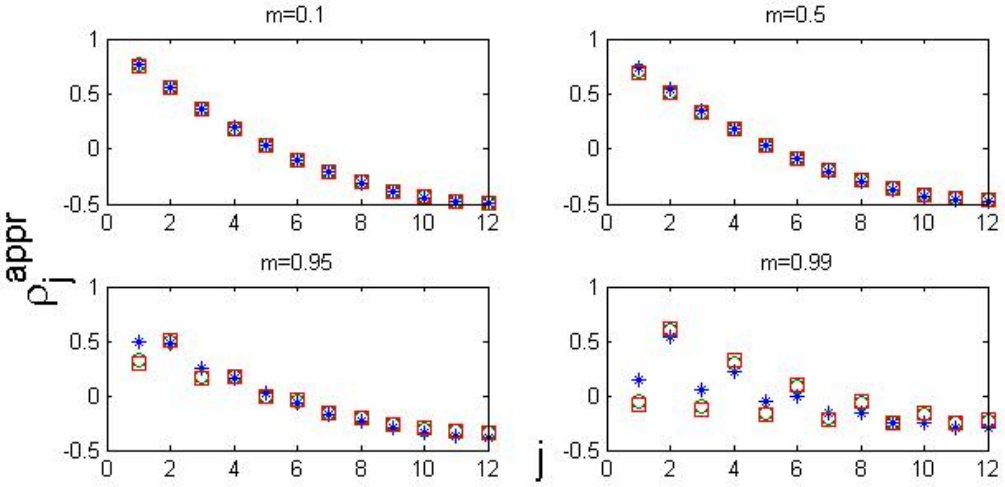

Figure 1: Plots of autocorrelation ρ_j^{appr} versus j for the one-dimensional (circular) stepping stone model and four different migration rates m, with $s = 25$, $N_e = N_c = 50$, $p = 0.1$ (asterisks), $p = 0.3$ (circles) and $p = 0.5$ (squares).

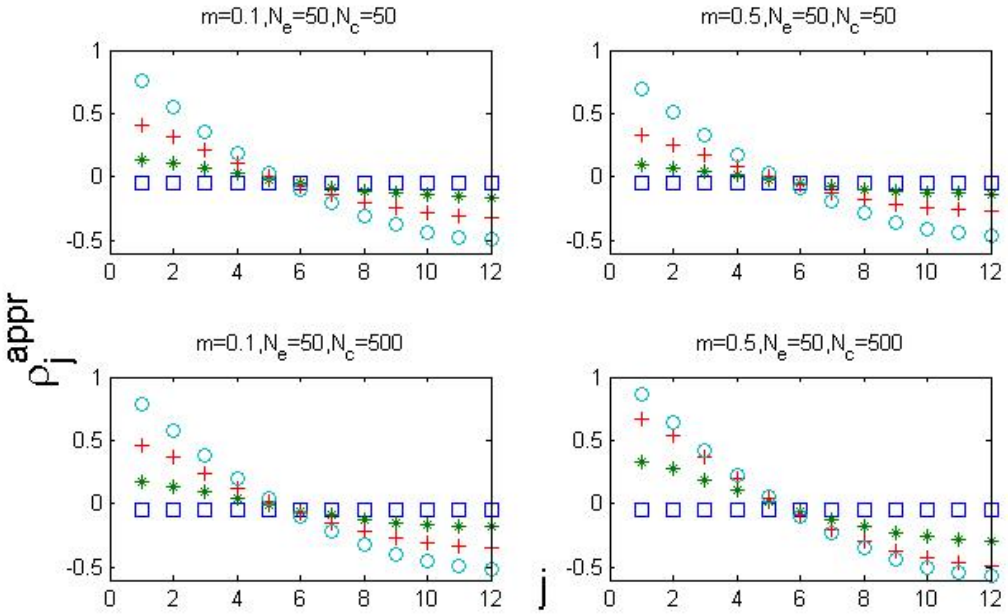

Figure 2: Plots of autocorrelation ρ_j^{appr} versus j for the one-dimensional von Mises model with $s = 25$ and four different combinations of the migration rate m, the local effective population size N_e and the local census population size N_c. The concentration parameter κ equals 0 (squares = island model), 2 (asterisks), 5 (plus signs) and 50 (circles, close to circular stepping stone).

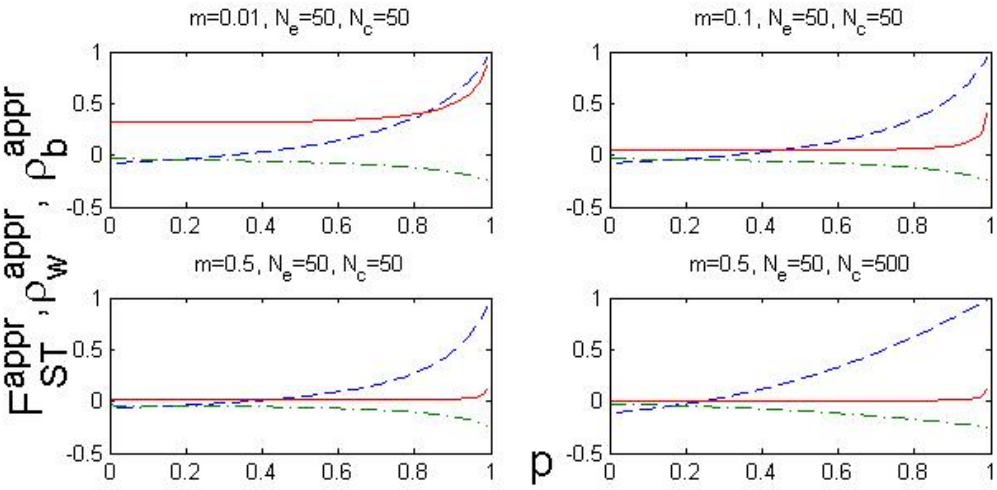

Figure 3: Plots of fixation index F_{ST}^{appr} (solid), autocorrelation ρ_w^{appr} within the same (dashed) and ρ_b^{appr} between different (dash-dotted) groups of subpopulations, for the hierarchical island model, with $s_1 = s_2 = 5$ and various combinations of m, N_c and N_e. The probability p of migrating to the same group varies between 0 and 1. The island model corresponds to $\rho_w^{\text{appr}} = \rho_b^{\text{appr}}$ and $p = (s_1 - 1)/s = 4/24 = 1/6$.

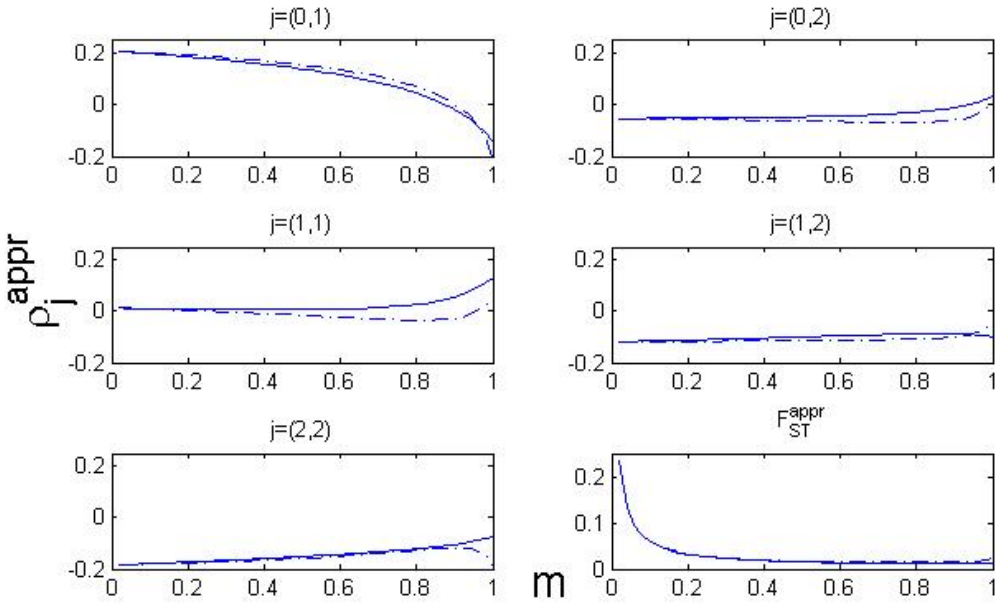

Figure 4: Plots of autocorrelation ρ_j^{appr} and fixation index F_{ST}^{appr} (lower right) as function of the migration rate m for a two-dimensional (torus) stepping stone model with $s_1 = s_2 = 5$ and $N_e = N_c = 50$. The migration probabilities are either symmetric ($p_j = 0.25$ for all $j \in \{(-1,0),(0,1),(0,-1),(0,1)\}$, solid), or non-symmetric ($p_j = 0.45$ for $j \in \{(0,1),(1,0)\}$, $p_j = 0.05$ for $j \in \{(0,-1),(-1,0)\}$, dash-dotted).

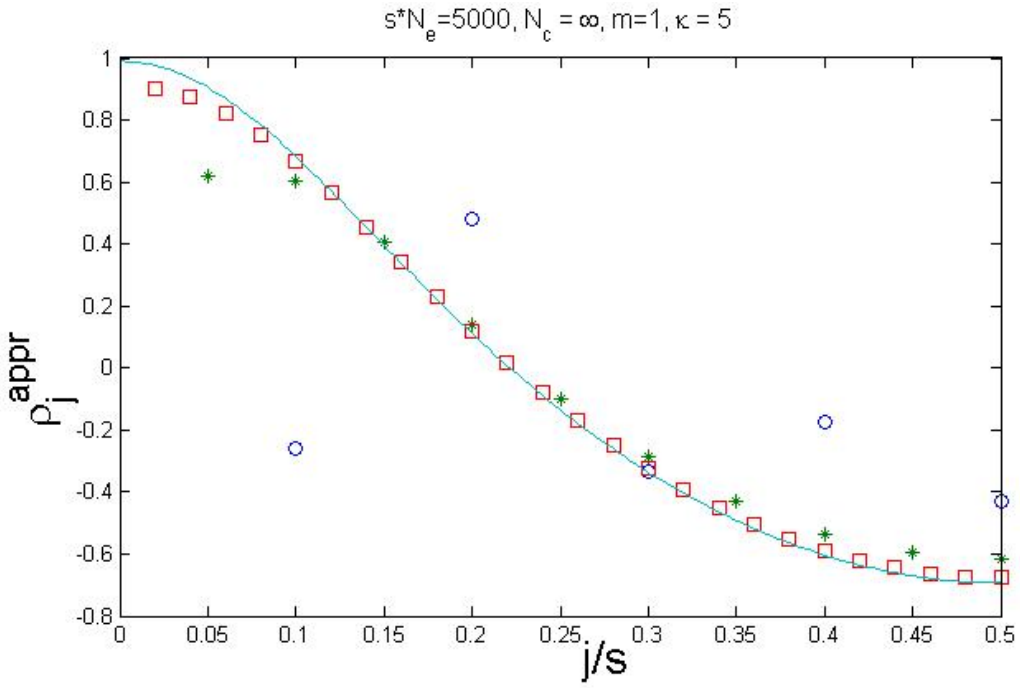

Figure 5: Plots of autocorrelation ρ_j^{appr} versus j/s for the one-dimensional von Mises model when $N_{e,\text{tot}} = sN_e = 5000$, $N_c = \infty$, $m = 1$ and $\kappa = 5$. The four plots correspond to $s = 10$ (circles), $s = 20$ (asterisks), $s = 50$ (squares) and $s = 300$ (solid).

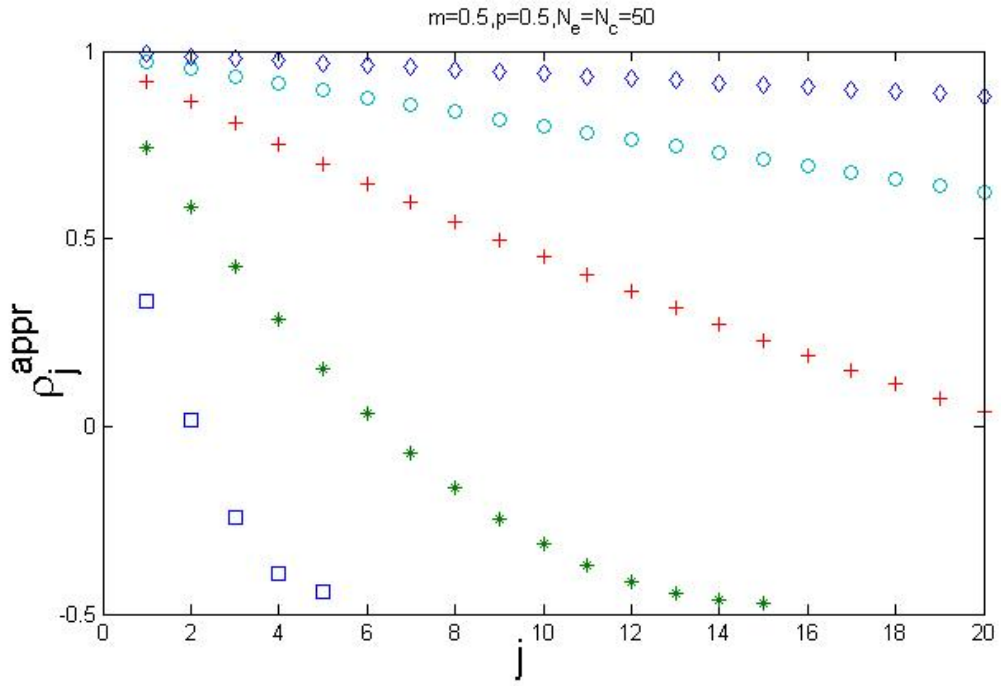

Figure 6: Plots of autocorrelation ρ_j^{appr} as function of j for the one-dimensional (circular) stepping stone model with $m = p = 0.5$ and $N_e = N_c = 50$. The four curves correspond to $s = 10$ (squares), $s = 30$ (asterisks), $s = 100$ (plus signs), $s = 300$ (circles) and $s = 1000$ (diamonds).