



Contents lists available at ScienceDirect

Physica D

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

Well posedness and maximum entropy approximation for the dynamics of quantitative traits

Katarína Boďová^a, Jan Haskovec^{b,*}, Peter Markowich^b

^a Institute of Science and Technology Austria (IST Austria), Klosterneuburg A-3400, Austria

^b Computer, Electrical and Mathematical Sciences & Engineering, King Abdullah University of Science and Technology, 23955 Thuwal, Saudi Arabia

HIGHLIGHTS

- We study a Fokker–Planck equation with degenerate diffusion coefficients.
- The equation describes the dynamics of quantitative traits.
- We construct mass and positivity preserving weak solutions of the equation.
- Existence of a positive spectral gap implies exponential convergence to equilibrium.
- We study analytically and numerically the “Dynamic Maximum Entropy” approximation.

ARTICLE INFO

Article history:

Received 27 April 2017

Received in revised form 25 October 2017

Accepted 29 October 2017

Available online xxx

Dedicated to E. Titi's 60. Birthday

Keywords:

Quantitative traits

Fokker–Planck equation

Boundary condition

Quasi-stationary approximation

Dynamic maximum entropy

ABSTRACT

We study the Fokker–Planck equation derived in the large system limit of the Markovian process describing the dynamics of quantitative traits. The Fokker–Planck equation is posed on a bounded domain and its transport and diffusion coefficients vanish on the domain's boundary. We first argue that, despite this degeneracy, the standard no-flux boundary condition is valid. We derive the weak formulation of the problem and prove the existence and uniqueness of its solutions by constructing the corresponding contraction semigroup on a suitable function space. Then, we prove that for the parameter regime with high enough mutation rate the problem exhibits a positive spectral gap, which implies exponential convergence to equilibrium.

Next, we provide a simple derivation of the so-called Dynamic Maximum Entropy (DynMaxEnt) method for approximation of observables (moments) of the Fokker–Planck solution, which can be interpreted as a nonlinear Galerkin approximation. The limited applicability of the DynMaxEnt method inspires us to introduce its modified version that is valid for the whole range of admissible parameters. Finally, we present several numerical experiments to demonstrate the performance of both the original and modified DynMaxEnt methods. We observe that in the parameter regimes where both methods are valid, the modified one exhibits slightly better approximation properties compared to the original one.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Quantitative genetics studies evolution of biologically important traits which are continuous (e.g., size, weight, lifespan, color) and depend on many loci; some of the loci contribute to the trait with a large effect but majority contribute with very small effect. Selection, mutation and stochasticity affect the evolution

of quantitative traits and lead to complex dynamics that is often studied using mathematical models. However, most of these models are discrete and stochastic and their complexity obscures rigorous analysis of the effects of various parameters, that are key for biological interpretation. Therefore, the discrete processes are often replaced by continuous approximation, leading to the so-called *diffusion theory*, which accurately approximates multiple classes of population genetics models [1–4].

The diffusion theory describes the dynamics of allele frequencies $\mathbf{x} = (x_1, \dots, x_L)$, where L is the number of loci that contribute to the trait. Each locus allows two possible alleles (type 0 and 1) and x_i is the proportion of alleles 1 at the i th locus in a population

* Corresponding author.

E-mail addresses: kbodova@ist.ac.at (K. Boďová), jan.haskovec@kaust.edu.sa (J. Haskovec), peter.markowich@kaust.edu.sa (P. Markowich).

<https://doi.org/10.1016/j.physd.2017.10.015>

0167-2789/© 2017 Elsevier B.V. All rights reserved.

of N diploid individuals. The joint probability density of allele frequencies $u = u(t, \mathbf{x})$ then satisfies the linear Fokker–Planck equation (i.e., the deterministic forward Kolmogorov equation) of a form

$$\frac{\partial u}{\partial t} = -\frac{1}{2} \sum_{i=1}^L \frac{\partial}{\partial x_i} \left(\xi_i \frac{\partial(\boldsymbol{\alpha} \cdot \mathbf{A})}{\partial x_i} u \right) + \frac{1}{4N} \sum_{i=1}^L \frac{\partial^2}{\partial x_i^2} (\xi_i u), \quad (1.1)$$

on $\Omega_{\mathbf{x}} := (0, 1)^L$, where we denoted $\xi_i := \xi(x_i) = x_i(1 - x_i)$ for $i = 1, \dots, L$. The diffusion term captures the stochasticity of the allele frequencies arising from random sampling. Here we assume that linkage disequilibria are negligible, otherwise this term would be of cross-diffusion type, reflecting correlations between loci [5]. This biologically corresponds to a sufficiently fast recombination such that the population can be represented by allele frequencies. The Eq. (1.1) describes a broad range of evolutionary scenarios. These depend on the form of the drift term, which captures deterministic effects on allele frequencies that are described by a vector of coefficients $\boldsymbol{\alpha}$ and a vector of complementary quantities \mathbf{A} . We assume the mathematically simplest form of directional selection, extended to dominance, which adds selection on heterozygous individuals. Moreover, we assume symmetrical mutation (easily generalizable to asymmetrical mutation), which together, using the notation of [5], corresponds to the choice

$$\mathbf{A} = (\xi'_1, \dots, \xi'_L, \xi_1, \dots, \xi_L, \ln \xi_1, \dots, \ln \xi_L)$$

and

$$\boldsymbol{\alpha} \cdot \mathbf{A} = -\beta \sum_{i=1}^L \gamma_i \xi'_i + 2h \sum_{i=1}^L \eta_i \xi_i + 2\mu \sum_{i=1}^L \ln \xi_i, \quad (1.2)$$

where the nondimensional parameters $\beta, h, \gamma_i, \eta_i \in \mathbb{R}$ represent the effects of loci on the traits, $\mu > 0$ is the mutation rate, and $\xi'_i := \xi'(x_i) = 1 - 2x_i$. For notational simplicity and without loss of generality, we set $\beta = h = 1$ in the sequel, so that

$$\boldsymbol{\alpha} = (-\gamma_1, \dots, -\gamma_L, 2\eta_1, \dots, 2\eta_L, 2\mu, \dots, 2\mu) \in \mathbb{R}^{3L}.$$

The main difficulty for analysis of the Fokker–Planck equation (1.1) is the degeneracy of the diffusion coefficients $\xi_i = x_i(1 - x_i)$ at the boundary of $\Omega_{\mathbf{x}}$, which arises due to finite population size. Consequently, the task of prescribing boundary conditions that lead to a well-posed problem is far from obvious; see also [6,7] for related issues in population genetics problems. As noted above, we aim at interpreting the solution u as a time-dependent probability density, which calls for a no-flux boundary condition. In Section 2 we argue that the standard no-flux boundary condition is indeed appropriate for (1.1). In Section 3 we derive the weak formulation of (1.1) subject to the no-flux boundary condition and prove the existence and uniqueness of its solutions by constructing the corresponding contraction semigroup. Then, in Section 4 we prove that for the parameter regime with high enough mutation rate the problem exhibits a positive spectral gap, which implies exponential convergence to equilibrium.

The main challenge of quantitative genetics is to predict the dynamics of the complex traits, particularly when the genetic forces are changing in time. If all evolutionary forces $\boldsymbol{\alpha}$ (with a possible time-dependence) are known one may simply solve the Fokker–Planck equation (1.1) and use the allele frequency distribution $u(t, \mathbf{x})$ to obtain the trait mean and other observables of interest by computing expectations of functions of allele frequencies subject to $u(t, \mathbf{x})$. However, Eq. (1.1) is too complex both for analytical solutions and for effective numerical methods in the multi-locus case. Therefore, we present a method that approximates the macroscopic observables of the Fokker–Planck solution without needing to solve the PDE (for more biological details see [5]). Section 5 is devoted to the study of this approximation, the so-called Dynamic Maximum Entropy (DynMaxEnt). We

first show in Section 5.1 that a related constrained entropy maximization is equivalent to a moment-matching problem, which we solve in a simple case. This provides an interpretation of the observables and evolutionary forces as the variational constraints and the corresponding Lagrange multipliers. Then, in Section 5.2 we provide a simple and straightforward derivation of the DynMaxEnt method by adopting a quasi-stationary approximation, by assuming that at every time the allele frequency distribution can be approximated by an equilibrium distribution with some evolutionary parameters $\boldsymbol{\alpha}(t)$, that change in time to match the observables. This results in a nonlinear low-dimensional system of ordinary differential equations. It should be noted that the approximation is exact if the system is in equilibrium. It can be interpreted as a nonlinear Galerkin approximation of the Fokker–Planck equation (1.1).

However, this “original” DynMaxEnt method cannot be applied in the biologically realistic regime of small mutations, i.e., when $4N\mu \leq 1$. This inspires us to introduce a modified version, which is valid for the whole range of admissible parameters, Section 5.3. Finally, in Section 6 we present several numerical experiments to demonstrate the performance of both the original and modified DynMaxEnt methods. We observe that in the parameter regimes where both methods are valid, the modified one exhibits slightly better approximation properties compared to the original one.

The surprisingly good approximation properties of the DynMaxEnt method, as documented by the numerical results in [5] and Section 6 of this paper, suggest that the infinitely-dimensional dynamics of the Fokker–Planck equation (1.1) can be well approximated by suitable finitely-dimensional dynamical systems. This is reminiscent of the recent series of works of E. Titi and collaborators [8–12] where a data assimilation (downscaling) approach to fluid flow problems is developed, inspired by ideas applied for designing finite-parameters feedback control for dissipative systems. The goal of a data assimilation algorithm is to obtain (numerical) approximation of a solution of an infinitely-dimensional dynamical system corresponding to given measurements of a finite number of observables. In particular, in [8], it has been shown that solutions of the two-dimensional Navier–Stokes equations can be well reconstructed from a relatively low number of low Fourier modes or local averages over finite volume elements. In [9], continuous data assimilation (CPA) algorithm was proposed and analyzed for a two-dimensional Bénard convection problem, where the observables were incorporated as a feedback (nudging) term in the evolution equation of the horizontal velocity. In [10] CPA was applied for downscaling a coarse resolution configuration of the 2D Bénard convection equations into a finer grid, while in [11] the CPA method is studied for a three-dimensional Brinkman–Forchheimer-extended Darcy model of porous media, and in [12] for the three-dimensional Navier–Stokes– α model. Finally, in [13] numerical performance of the CPA algorithm in the context of the two-dimensional incompressible Navier–Stokes equations was studied. It was shown that the numerical method is computationally efficient and performs far better than the analytical estimates suggest. This is similar to our numerical observations showing very good approximation properties of the DynMaxEnt method applied to the Fokker–Planck equation (1.1).

2. Boundary conditions for the stationary problem

As pointed out in Section 1, the degeneracy of the diffusion coefficients $\xi_i = x_i(1 - x_i)$ at the boundary $\partial\Omega_{\mathbf{x}}$ renders the question which type of boundary conditions shall be prescribed (if any at all) to produce a well-posed problem. However, the essential property that we want to impose on the solution is total mass conservation (together with nonnegativity preservation), so that

Download English Version:

<https://daneshyari.com/en/article/8256189>

Download Persian Version:

<https://daneshyari.com/article/8256189>

[Daneshyari.com](https://daneshyari.com)