

Max-Planck-Institut  
für Mathematik  
in den Naturwissenschaften  
Leipzig

A hierarchical extension scheme for solutions of  
the WrightFisher model

by

*Julian Hofrichter, Tat Dat Tran, and Jürgen Jost*

Preprint no.: 61

2014





# A hierarchical extension scheme for solutions of the Wright–Fisher model

Julian Hofrichter\*, Tat Dat Tran<sup>†</sup>, Jürgen Jost<sup>‡</sup>

June 19, 2014

We develop a global and hierarchical scheme for the forward Kolmogorov (Fokker-Planck) equation of the diffusion approximation of the Wright-Fisher model of population genetics. That model describes the random genetic drift of several alleles at the same locus in a population. The key of our scheme is to connect the solutions before and after the loss of an allele. Whereas in an approach via stochastic processes or partial differential equations, such a loss of an allele leads to a boundary singularity, from a biological or geometric perspective, this is a natural process that can be analyzed in detail. Our method depends on evolution equations for the moments of the process and a careful analysis of the boundary flux.

**Keywords:** Wright-Fisher model; forward Kolmogorov equation; random genetic drift; hierarchical solution

## 1 Introduction

The Wright-Fisher model [15, 38], the basic model of population genetics, models random genetic drift in a finite population of fixed size. It describes the evolution of the probabilities between non-overlapping generations of several alleles. These probabilities are obtained from random sampling in the parental generation. The basic model covers a single locus only. Extensions to several loci are possible, as is the inclusion of mutation, selection, or a spatial population structure, and this has driven research in mathematical population genetics ([14, 4]), starting with the pioneering work of Kimura [25, 26, 27]. In fact, the important insight was that one should not work with the original model of a finite population evolving in discrete time steps, but rather consider the diffusion approximation for an infinite population in continuous time. The latter is well described

---

\*hofricht@mis.mpg.de

<sup>†</sup>trandat@mis.mpg.de

<sup>‡</sup>jost@mis.mpg.de, phone +49-341-9959550

by two partial differential equations, the forward and backward Kolmogorov equations. The forward equation is a partial differential equation of parabolic type, and the backward equation is its adjoint w.r.t. a suitable product (and with a boundary contribution whose analysis will be the main new tool of the present paper). The forward equations describes how an initial probability distribution for the alleles in a population evolves in time. Since alleles can disappear from the population by random genetic drift and will then be lost forever (unless the possibility of mutations is included, which, however, we do not consider in this paper), one needs to dynamically connect strata corresponding to different allele numbers. This is the main source of difficulties addressed in this paper. The backward equation, in contrast, describes the evolution of the probabilities of the distributions of ancestral states giving rise to the present allele distribution. Solutions of an inhomogeneous backward equation also yield formulas for expected loss of allele times. In the backward equation, the transition between strata is simpler, because the contributions from the various strata essentially add up. That is, the degeneracy at the boundary is easier for the backward than for the forward equations, and there already exists a substantial literature corresponding expansions for the solution of the backward equations, see for instance [29, 31, 16, 17, 32, 33, 10], as well as the more recent investigations of the polynomial expansions in [18, 19, 20]. A powerful tool in this line of research has been Kingman's coalescent [28], that is, the method of tracing lines of descent back into the past and analyzing their merging patterns (for a quick introduction to that theory, see also [23]). We shall describe some of these achievements in more detail below. Here, as already mentioned, we analyze the more difficult boundary transitions for the forward equation. For this purpose, we shall also utilize the duality between the forward and the backward equations in an essential manner.

Since the original work of Fisher, Wright and Kimura, the Wright-Fisher model as well as several extensions or generalizations of it have been investigated in considerable detail. One research line incorporated the model into the general theory of stochastic processes or partial differential equations and derived existence and regularity results from that general perspective, see for instance [9, 11, 12, 24] for stochastic processes or [7, 8] for partial differential equations. A lot of research was also concerned with explicit results and formulas, for instance for the expected time of loss of an allele, and therefore, while relying on the abstract theory, had to work out the specific and explicit structure of the model. We propose a third line of research on the Wright-Fisher model, based on geometric constructions, more specifically on information geometry, that is, the geometry of probability distributions, see [1, 2].

In fact, the structure of the model is surprisingly rich, in particular when viewed from a geometric perspective. We are carrying out a systematic research project (see [34, 21, 35, 36, 37]) to develop this structure and thereby obtain an understanding of the Wright-Fisher model that is both deeper and more explicit, in the sense that precise formulae for the quantities of interest can be derived. The present paper, which is based on [21], is a key component of this project.

There are many solution schemes for the Kolmogorov forward equation (3.1) in the literature. As early as 1956, Kimura presented a local solution scheme for the 3-allelic case ( $n = 2$ ) in [27]. Another approach by separation of variables was presented by

Baxter, Blythe and McKane in [3], this time for an arbitrary number of alleles. For the Kolmogorov backward equation, the situation is even better. The starting point of much of the literature was the observation of Wright [39] that when one includes mutation, the degeneracy at the boundary is removed. And when the probability of a mutation of allele  $i$  into allele  $j$  depends only on the target  $j$ , then the backward process possesses a unique stationary distribution, at least as long as those mutation rates are positive. This then lead to explicit representation formulas in [29, 31, 16, 17, 32, 33, 10]. Some of these formulas, in particular those of [32, 10] also pertain to the limit of vanishing mutation rates. In [32], a superposition of the contributions from the various strata was achieved whereas [10] could write down an explicit formula in terms of a Dirichlet distribution. As already mentioned, for the Kolmogorov forward equation, the situation is more subtle because such a superposition for the contributions from the various strata no longer holds. The difference is the following. The backward equation produces the probability distribution of ancestral states giving rise to a current distribution. That latter distribution may involve states with different numbers of alleles present. Their ancestral distributions, however, do not interfere, regardless of the numbers of alleles they involve. Thus, some superposition principle holds, and the Kolmogorov backward equation nicely extends to the boundary. In contrast, the Kolmogorov forward equation yields the future distribution evolving from a current one. Here, however, the probability of some boundary state does not only depend on the flow within the corresponding boundary stratum, but also on the distribution in the interior, because at any time, there is some probability that an interior state loses some allele and turns into a boundary state. Thus, there is a continuous flux into the boundary strata from the interior. Therefore, the extension of the flow from the interior to the boundary strata is different from the intrinsic flows in those strata, and no superposition principle holds.

Let us now describe in more specific terms what we achieve in this paper. The key is the degeneracy at the boundary of the Kolmogorov equations. While from an analytical perspective, this presents a profound difficulty for obtaining boundary regularity of the solutions of the equations, from a biological or geometric perspective, this is very natural because it corresponds to the loss by random drift of some alleles from the population in finite time. And from a stochastic perspective, this has to happen almost surely. Now, however, even after an allele gets lost, the population keeps evolving by random genetic drift according to the Wright-Fisher scheme, simply with fewer alleles than before, until only one allele is left and the evolution comes to a halt (in this basic model). Therefore, it is biologically essential and geometrically natural to connect the processes before and after the loss of an allele. When the original process, starting with, say,  $n + 1$  alleles, takes place on an  $n$ -dimensional probability simplex, after the loss of an allele, we have a process on an  $(n - 1)$ -dimensional simplex. The latter then should be identified as a facet of the former, that is, the loss of an allele simply means that the process moves from the interior into the boundary of the original simplex. Of course, this will then be repeated when further alleles get lost, and the process moves to lower and lower dimensional boundary strata until it gets stuck in a corner. In this paper, we therefore construct a global solution that incorporates and connects these successive loss of allele events. In technical terms, we develop a hierarchical scheme that relies on equations for the

moments of the process, the interplay between the forward and the backward Kolmogorov equations, representations of the solutions in terms of Gegenbauer polynomials and on a careful analysis of the boundary flux. We should also point out that while the existing literature draws upon concepts and results from the theory of stochastic processes in an essential way, our approach is analytical and geometric and therefore offers an alternative to the existing ones. The processes for fewer alleles that occur in our hierarchical scheme combine the intrinsic dynamics with a reduced allele number with the contributions through loss-of-allele events from larger allele sets. This avoids any singularities.

In geometric terms, we carefully investigate the boundary flux from a simplex into its various boundary faces. In analytical terms, we reduce the Kolmogorov type PDE to a family of ODEs, the moment evolution equations. Earlier research in that direction is due to Dawson-Hochberg [5] and Dynkin [6], within a more general and therefore less explicit scheme. In our approach, the moment equations are global and guarantee the consistency of the process beyond loss-of-allele events and across the various boundary faces, see Theorem 4.12. And in [35, 36, 37], we have already constructed a global solution, which (implicitly) made use of the hierarchical scheme presented here.

Extensions to several loci and the effects of genetic recombination, processes with mutations, presence of selective forces, etc. will be studied elsewhere.

## 2 Preliminaries and notation

In the Wright–Fisher-model, we have a population of finite size  $N$  carrying initially  $n + 1$  different alleles, at a single locus in the basic version that we are concerned with here. For each member of a new generation – the model works with discrete time steps –, randomly (with replacement) a mother is drawn from the previous generation who then donates her allele to that offspring. Eventually, all but one allele will get lost by random drift from the population. For the mathematical analysis, one goes to the diffusion approximation of the process, that is, rescales time as  $t = \frac{1}{N}$  and lets  $N \rightarrow \infty$ . The evolution of the probability distribution of the alleles in the population is then described by a Fokker-Planck equation, also called the forward Kolmogorov equation. The state space here is the  $n$ -dimensional probability simplex.

Since we plan to develop a hierarchical scheme for the solution of the Kolmogorov equations on the various boundary strata of the standard simplex, we need to develop some notation which allows for the recursive application of our scheme on different boundary strata. We also need suitable hierarchical products. That is the purpose of this section.

We consider relative frequencies  $x^0, x^1, \dots, x^n$  of alleles  $0, 1, \dots, n$ . Thus we have the normalization  $\sum_{j=0}^n x^j = 1$ , and we therefore have  $x^0 = 1 - \sum_{i=1}^n x^i$ . We shall therefore work with the (open)  $n$ -dimensional standard orthogonal simplex

$$\Delta_n := \left\{ (x^1, \dots, x^n) \in \mathbb{R}^n \mid x^i > 0 \text{ for } i = 1, \dots, n \text{ and } \sum_{i=1}^n x^i < 1 \right\}, \quad (2.1)$$

or equivalently,

$$\Delta_n = \left\{ (x^1, \dots, x^n) \in \mathbb{R}^n \mid x^j > 0 \text{ for } j = 0, 1, \dots, n \text{ and } \sum_{j=0}^n x^j = 1 \right\}. \quad (2.2)$$

Its topological closure is

$$\overline{\Delta}_n = \left\{ (x^1, \dots, x^n) \in \mathbb{R}^n \mid x^i \geq 0 \text{ for } i = 1, \dots, n \text{ and } \sum_{i=1}^n x^i \leq 1 \right\}. \quad (2.3)$$

In order to include time  $t \in [0, \infty)$ , we shall also use the notation

$$(\Delta_n)_\infty := \Delta_n \times (0, \infty).$$

The boundary  $\partial\Delta_n = \overline{\Delta}_n \setminus \Delta_n$  consists of various subsimplices of descending dimensions called *faces*, starting from the  $(n-1)$ -dimensional *facets* down to the *vertices* (which represent 0-dimensional faces). Each subsimplex of dimension  $k \leq n-1$  is isomorphic to the  $k$ -dimensional standard orthogonal simplex  $\Delta_k$ . For an index set  $I_k = \{i_0, i_1, \dots, i_k\} \subset \{0, \dots, n\}$  with  $i_j \neq i_l$  for  $j \neq l$  we put

$$\Delta_k^{(I_k)} := \left\{ (x^1, \dots, x^n) \in \overline{\Delta}_n \mid x^i > 0 \text{ for } i \in I_k; x^i = 0 \text{ for } i \in I_n \setminus I_k \right\}. \quad (2.4)$$

We note that  $\Delta_n = \Delta_n^{(I_n)}$ .

For a given  $k \leq n-1$ , there are of course  $\binom{n+1}{k+1}$  different such subsets  $I_k$  of  $I_n$ , each of which corresponds to a certain boundary face  $\Delta_k^{(I_k)}$ . We therefore introduce the  $k$ -dimensional boundary  $\partial_k\Delta_n$  of  $\Delta_n$  by putting

$$\partial_k\Delta_n^{(I_n)} := \bigcup_{I_k \subset I_n} \Delta_k^{(I_k)} \subset \partial\Delta_n^{(I_n)} \quad \text{for } 0 \leq k \leq n-1. \quad (2.5)$$

With this notation, we have  $\partial_n\Delta_n = \Delta_n$ , although this is not a boundary component. The concept of the  $k$ -dimensional boundary also applies to simplices which are themselves boundary instances of some  $\Delta_l^{(I_l)}$ ,  $I_l \subset I_n$  for  $0 \leq k < l \leq n$ , thus

$$\partial_k\Delta_l^{(I_l)} = \bigcup_{I_k \subset I_l} \Delta_k^{(I_k)} \subset \partial\Delta_l^{(I_l)}. \quad (2.6)$$

In the Wright–Fisher model,  $\Delta_n$  corresponds to the state of all  $n+1$  alleles being present, whereas  $\partial_k\Delta_n$  represents the state of exactly (any)  $k+1$  alleles being present in the population. The individual  $\Delta_k^{\{i_0, \dots, i_k\}}$  comprising  $\partial_k\Delta_n$  correspond to the state of exactly the alleles  $i_0, \dots, i_k$  being present in the population. Likewise,  $\partial_{k-1}\Delta_k^{\{i_0, \dots, i_k\}}$  corresponds to the state of exactly one further allele out of  $i_0, \dots, i_k$  being eliminated from the population.

In order to define integral products on  $\Delta_n$  and its faces, we need appropriate spaces of square integrable functions. Thus,

$$L^2\left(\bigcup_{k=0}^n \partial_k \Delta_n\right) := \left\{ f: \bar{\Delta}_n \longrightarrow \mathbb{R} \mid f|_{\partial_k \Delta_n} \text{ is } \mathfrak{L}_k\text{-measurable and} \right. \\ \left. \int_{\partial_k \Delta_n} |f(x)|^2 \mathfrak{L}_k(dx) < \infty \text{ for all } k = 0, \dots, n \right\}. \quad (2.7)$$

Here,  $\mathfrak{L}_k$  stands for  $k$ -dimensional Lebesgue measure, but when integrating over some  $\Delta_k^{(I_k)}$  with  $0 \notin I_k$ , the measure needs to be replaced with the one induced on  $\Delta_k^{(I_k)}$  by the Lebesgue measure of the containing  $\mathbb{R}^{k+1}$  – this measure, however, will still be denoted by  $\mathfrak{L}_k$  as it is clear from the domain of integration  $\Delta_k^{(I_k)}$  with either  $0 \in I_k$  or  $0 \notin I_k$  which version is actually used. In particular, for the top-dimensional simplex, we simply have

$$L^2(\Delta_n) := \left\{ f: \Delta_n \longrightarrow \mathbb{R} \mid f \text{ is } \mathfrak{L}_n\text{-measurable and } \int_{\Delta_n} |f(x)|^2 \mathfrak{L}_n(dx) < \infty \right\}. \quad (2.8)$$

Furthermore, we also define for  $k \in \mathbb{N} \cup \{\infty\}$

$$C_0^k(\bar{\Delta}_n) := \{f \in C^k(\bar{\Delta}_n) \mid f|_{\partial \Delta_n} = 0\}, \quad (2.9)$$

$$C_0^k(\Delta_n) := \{f \in C^k(\Delta_n) \mid \exists \bar{f} \in C_0^k(\bar{\Delta}_n) \text{ with } \bar{f}|_{\Delta_n} = f\} \quad (2.10)$$

as well as

$$C_c^k(\bar{\Delta}_n) := \{f \in C^k(\bar{\Delta}_n) \mid \text{supp}(f) \subsetneq \Delta_n\}. \quad (2.11)$$

We can now introduce a *product* of functions  $u, v \in L^2(\Delta_n)$  by

$$(u, v)_n := \int_{\Delta_n} u(x)v(x) \mathfrak{L}_n(dx). \quad (2.12)$$

Importantly, the integration here is carried out only over the interior  $\Delta_n$ ; the index – if no confusion is to be expected – may be omitted. As  $n$  was arbitrary, the product may also be recursively applied on some  $\Delta_k^{(I_k)} \subset \partial \Delta_n$ .

Utilising the various products  $(\cdot, \cdot)_k$  on  $\partial_k \Delta_n^{(I_n)}$  for  $k = 0, \dots, n$ , we can now define a *hierarchical product* on the closure of the simplex  $\bar{\Delta}_n$ . For functions  $u, v: \bar{\Delta}_n \longrightarrow \mathbb{R}$  with  $u, v|_{\partial_k \Delta_n^{(I_n)}} \in L^2(\partial_k \Delta_n^{(I_n)})$  for all  $k = 0, \dots, n$ , we put

$$[u, v]_n := \sum_{k=0}^n (u, v)_k \quad (2.13)$$

with  $(u, v)_k$  denoting the integral over the full  $k$ -dimensional boundary  $\partial_k \Delta_n$  of  $\Delta_n$  (cf. equation (2.5)), thus

$$[u, v]_n = \sum_{k=0}^n (u, v)_k = \sum_{k=0}^n \int_{\partial_k \Delta_n} u(x)v(x) \mathfrak{L}_k(dx) = \sum_{k=0}^n \sum_{I_k \subset I_n} \int_{\Delta_k^{(I_k)}} u(x)v(x) \mathfrak{L}_k(dx) \quad (2.14)$$



with  $\mathbb{A}_k$  again denoting either the Lebesgue measure of  $\mathbb{R}^k$  or – if the domain of integration is some  $\Delta_k^{(I_k)}$  with  $0 \notin I_k$  – the measure induced on  $\Delta_k^{(I_k)}$  by the Lebesgue measure of the containing  $\mathbb{R}^{k+1}$ .

### 3 The Kolmogorov operators

The *Kolmogorov forward equation* for the diffusion approximation of the  $n$ -allelic 1-locus Wright–Fisher model reads

$$\begin{cases} \frac{\partial}{\partial t} u(x, t) = L_n u(x, t) & \text{in } (\Delta_n)_\infty = \Delta_n \times (0, \infty) \\ u(x, 0) = f(x) & \text{in } \Delta_n, f \in L^2(\Delta_n) \end{cases} \quad (3.1)$$

for  $u(\cdot, t) \in C^2(\Delta_n)$  for each fixed  $t \in (0, \infty)$  and  $u(x, \cdot) \in C^1((0, \infty))$  for each fixed  $x \in \Delta_n$  and with

$$L_n u(x, t) := \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2}{\partial x^i \partial x^j} (x^i (\delta_j^i - x^j) u(x, t)) \quad (3.2)$$

being the *forward operator*. Analogously, we have the *backward operator*

$$L_n^* u(x, t) := \frac{1}{2} \sum_{i,j=1}^n (x^i (\delta_j^i - x^j)) \frac{\partial^2}{\partial x^i \partial x^j} u(x, t), \quad (3.3)$$

appearing in the corresponding Kolmogorov backward equation. The definitions of the operators given in equations (3.2) and (3.3) also apply to the closure  $\overline{\Delta}_n$ , and in fact, we shall also consider extensions of the solution and the differential equation to the boundary.

For relations between the two operators, we immediately have the following two lemmas:

**3.1 Lemma.**  $L_n$  and  $L_n^*$  are (formal) adjoints with respect to the product  $(\cdot, \cdot)_n$  in the sense that

$$(L_n u, \varphi)_n = (u, L_n^* \varphi)_n \quad \text{for } u \in C^2(\overline{\Delta}_n), \varphi \in C_0^2(\overline{\Delta}_n). \quad (3.4)$$

*Proof.* The assertion directly follows from proposition 4.4 below.  $\square$

**3.2 Lemma.** For an eigenfunction  $\varphi \in C^2(\overline{\Delta}_n)$  of  $L_n$  and  $\omega_n := \prod_{k=1}^n x^k (1 - \sum_{l=1}^n x^l)$ , we have:  $\omega_n \varphi \in C_0^2(\overline{\Delta}_n)$  is an eigenfunction of  $L_n^*$  corresponding to the same eigenvalue and conversely.

*Proof.* Looking for a function  $\omega_n$  with  $L_n^*(\omega_n u) = \omega_n L_n(u)$  (and hence for  $L_n$ -eigenfunctions  $\varphi$  consequently  $L_n^*(\omega_n \varphi) = \omega_n L_n(\varphi) = -\lambda \omega_n \varphi$ ), we have on the one hand

$$L_n u = -\frac{n(n+1)}{2} u + \sum_i (1 - (n+1)x^i) \frac{\partial}{\partial x^i} u + \frac{1}{2} \sum_{i,j} x^i (\delta_j^i - x^j) \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} u \quad (3.5)$$

and on the other hand

$$\begin{aligned} L_n^*(\omega_n u) &= \frac{1}{2} \sum_{i,j} x^i (\delta_j^i - x^j) \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} (\omega_n u) \\ &= \frac{1}{2} \sum_{i,j} x^i (\delta_j^i - x^j) \left( \left( \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} \omega_n \right) u + 2 \frac{\partial}{\partial x^j} \omega_n \frac{\partial}{\partial x^i} u + \omega_n \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} u \right). \end{aligned} \quad (3.6)$$

Thus, it suffices that such a function  $\omega_n$  satisfies

$$\begin{cases} \sum_{i,j} x^i (\delta_j^i - x^j) \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} \omega_n = -n(n+1)\omega_n \\ \sum_j x^i (\delta_j^i - x^j) \frac{\partial}{\partial x^j} \omega_n = (1 - (n+1)x^i)\omega_n \quad \text{for all } i, \end{cases} \quad (3.7)$$

which is the case for  $\omega_n = \prod_{k=1}^n x^k (1 - \sum_{l=1}^n x^l)$  as may easily be verified by direct computation.  $\square$

For our hierarchical scheme, it will be important that the operator  $L_n^*$ , if restricted to subsimplices  $\Delta_k^{(I_k)} \cong \Delta_k$  in  $\partial\Delta_n^{(I_n)}$  of any dimension  $k$ , is the adjoint of the differential operator  $L_k$  corresponding to the evolution of a  $(k+1)$ -allelic process in  $\Delta_k$ :

**3.3 Lemma.** *For  $0 \leq k < n$  and  $I_k \subset \{0, \dots, n\}$ ,  $|I_k| = k$ , we have*

$$L_n^*|_{\Delta_k^{(I_k)}} = L_k^*. \quad (3.8)$$

*Proof.* For  $I_k \subset \{1, \dots, n\}$ ,  $|I_k| = k$ , we directly have:

$$\begin{aligned} L_n^*|_{\Delta_k^{(I_k)}} &= \frac{1}{2} \sum_{i,j=1}^n (x^i (\delta_j^i - x^j)) \frac{\partial^2}{\partial x^i \partial x^j} \Big|_{\Delta_k^{(I_k)}} \\ &= \frac{1}{2} \sum_{i,j \in I_k} (x^i (\delta_j^i - x^j)) \frac{\partial^2}{\partial x^i \partial x^j} \equiv L_k^*. \end{aligned} \quad (3.9)$$

By symmetry, this then also holds for  $I_k$  with  $0 \in I_k$ , hence for arbitrary  $I_k$ .  $\square$

We may therefore omit the index  $k$  in  $L_k^*$  whenever convenient, in particular when considering domains where (parts of) the boundary are included. For the operator  $L_n$ , in contrast, we do not have such a restriction property.

## 4 Solution schemes for the Kolmogorov forward equation

The starting point of our solution scheme will be the solution constructed in [36]. That solution depends on

**4.1 Proposition.** *For  $n \in \mathbb{N}_+$  and each multi-index  $\alpha = (\alpha^1, \dots, \alpha^n)$  with  $|\alpha| = \alpha^1 + \dots + \alpha^n = m \geq 0$ ,*

$$C_{m,\alpha}(x) := x^\alpha + \sum_{|\beta|=0}^{m-1} a_{m,\beta} x^\beta, \quad x \in \Delta_n \quad (4.1)$$

with  $a_{m,\beta}$  inductively defined by  $a_{m,\beta} := \delta_\beta^\alpha$  for  $|\beta| = m$  and

$$a_{m,\beta} := -\frac{\sum_{i=1}^n (\beta^i + 2)(\beta^i + 1)a_{m,(\beta^1, \dots, \beta^i+1, \dots, \beta^n)}}{(m - |\beta|)(m + |\beta| + 2n + 1)} \quad \text{for all } 0 \leq |\beta| \leq m - 1, \quad (4.2)$$

is an eigenfunction of  $L_n$  in  $\bar{\Delta}_n$  corresponding to the eigenvalue  $\lambda_m^{(n)} = \frac{(m+n)(m+n+1)}{2}$ .

Knowing the eigenfunctions, it is straightforward to reconstruct the local solution of [27], [3] (for details cf. [34, 36]).

**4.2 Proposition.** *For  $n \in \mathbb{N}$  and any initial condition  $f \in L^2(\Delta_n)$ , the Kolmogorov forward equation corresponding to the diffusion approximation of the  $n$ -dimensional Wright–Fisher model (3.1) always allows a unique solution  $u: (\Delta_n)_\infty \rightarrow \mathbb{R}$  with  $u(\cdot, t) \in C^\infty(\Delta_n)$  for each fixed  $t \in (0, \infty)$  and  $u(x, \cdot) \in C^\infty((0, \infty))$  for each fixed  $x \in \Delta_n$ . Furthermore, this solution (and all its spatial derivatives) may be extended continuously to the boundary  $\partial\Delta_n$ .*

The regularity, which follows from the regularity of the generalised Gegenbauer polynomials, of course agrees with standard PDE theory (cf. e. g. [22]).

#### 4.1 Moments and the weak formulation of the Kolmogorov forward equation

The solution of equation (3.1) in  $\Delta_n$  lacks conservation properties: As the smallest eigenvalue of  $L_n$  is  $\lambda_0^{(n)} = \frac{n(n+1)}{2} > 0$ , a solution vanishes everywhere in  $\Delta_n$  for  $t \rightarrow \infty$ , which in particular implies that the total mass and all other moments are generally not preserved. However, again these properties are an important aspect of the validity of the model, and what disappears in the interior of the simplex should accumulate in its boundary. After all, the process should continue after the loss of one or several alleles. We shall therefore introduce a suitable extended solution on the entire  $\bar{\Delta}_n$ . This solution will be derived from the conservation of the moments of the process. In fact, the moments of the  $n$ -dimensional process as obtained as limits of those from the underlying discrete model satisfy the *moment evolution equations*

$$\frac{\partial}{\partial t} \bar{\mu}_\alpha(t) = -\frac{|\alpha|(|\alpha| - 1)}{2} \bar{\mu}_\alpha(t) + \sum_{i=1}^n \frac{\alpha_i(\alpha_i - 1)}{2} \bar{\mu}_{\alpha - e_i}(t) \quad (4.3)$$

for  $\alpha = (\alpha_1, \dots, \alpha_n)$ ,  $|\alpha| \geq 1$ , whereas  $\frac{\partial}{\partial t} \bar{\mu}_0(t) = 0$  (with  $e_i$  denoting the multi-index  $(0, \dots, 0, 1, 0, \dots, 0)$  with 1 appearing at the  $i$ -th position). These moments can be defined as

$$\bar{\mu}_\alpha(t) := [U, x^\alpha]_n \equiv \sum_{k=0}^n \int_{\partial_k \Delta_n} U(x, t) x^\alpha \mathfrak{A}_k(dx), \quad t \geq 0, \alpha = (\alpha_1, \dots, \alpha_n), \quad (4.4)$$

utilising the hierarchical product as introduced in equation (2.13). This now involves an integration over  $\bar{\Delta}_n$ , that is, including the boundary  $\partial\Delta_n$  of the state space, which

corresponds to configurations of the model where some allele frequencies may be zero. Correspondingly, the capitalised  $U: (\overline{\Delta}_n)_\infty \rightarrow \mathbb{R}$  is introduced as an extended solution which is assumed to be the probability density function of the diffusion approximation of the  $n$ -dimensional Wright–Fisher process on the entire  $\overline{\Delta}_n$  (thus in particular  $U|_{\Delta_n}$  is a solution of the Kolmogorov forward equation (3.1) in  $\Delta_n$ ).

We shall now discuss the consistency between the moments evolution equation (4.3) and the Kolmogorov backward operator  $L^*$  in  $\overline{\Delta}_n$  as defined in equation (3.3); however, the following considerations also hold for a generic product  $[\cdot, \cdot]$ : Since  $L^*$  has polynomial coefficients, it maps polynomials to polynomials, and we have

$$\begin{aligned} L^* x^\alpha &= \frac{1}{2} \sum_{i,j=1}^n (x^i \delta_j^i - x^j) \frac{\partial^2}{\partial x^i \partial x^j} x^\alpha \\ &= \frac{1}{2} \sum_{i=1}^n \alpha_i (\alpha_i - 1) (x^{\alpha - e_i} - x^\alpha) - \frac{1}{2} \sum_{i \neq j} \alpha_i \alpha_j x^\alpha \\ &= \frac{1}{2} \sum_{i=1}^n \alpha_i (\alpha_i - 1) x^{\alpha - e_i} - \frac{1}{2} |\alpha| (|\alpha| - 1) x^\alpha \quad \text{for } x \in \overline{\Delta}_n, \end{aligned} \quad (4.5)$$

which yields, using the notation of equation (4.4),

$$[U(t), L_n^* x^\alpha]_n = \frac{1}{2} \sum_{i=1}^n \alpha_i (\alpha_i - 1) \bar{\mu}_{\alpha - e_i}(t) - \frac{1}{2} |\alpha| (|\alpha| - 1) \bar{\mu}^\alpha(t) \quad (4.6)$$

with the right-hand side being equal to that of equation (4.3). Thus, if the moments equation is fulfilled for some probability density function  $U$ , we may equivalently write

$$\frac{\partial}{\partial t} \bar{\mu}_\alpha(t) = \left[ \frac{\partial}{\partial t} U(t), x^\alpha \right]_n = [U(t), L_n^* x^\alpha]_n \quad \text{for } t \in (0, \infty). \quad (4.7)$$

Since the  $x^\alpha$  generate the space of all polynomials and since the polynomials are dense in  $C^\infty$ , we therefore also have such relations for arbitrary test functions  $\varphi$ ,

$$\left[ \frac{\partial}{\partial t} U(t), \varphi \right]_n = [U(t), L_n^* \varphi]_n \quad \text{for } \varphi \in C^\infty(\overline{\Delta}_n) \text{ and all } t \in (0, \infty). \quad (4.8)$$

This is our *weak formulation* of the Kolmogorov forward equation (3.1). We may also write the initial condition<sup>1</sup> weakly as

$$[U(\cdot, 0), \varphi]_n = [f, \varphi]_n \quad \text{for all } \varphi \in C^\infty(\overline{\Delta}_n), \quad (4.9)$$

---

<sup>1</sup>As the integration is over  $\overline{\Delta}_n$ ,  $f$  may now also be formulated as an extended initial condition on the entire  $\overline{\Delta}_n$ . Then,  $f|_{\partial \Delta_n} \neq 0$  would correspond to the process (partially) already starting on certain boundary instances. However, these parts of the process exactly evolve as a proper process of corresponding dimension, and hence do not yield any further insight into the nature of the process. For this reason, we will usually assume  $f|_{\partial \Delta_n} \equiv 0$  or that  $f$  is extended that way if it is only given on  $\Delta_n$ .

which requires no explicit regularity towards the boundary (yet, we will need that its restriction to interior instances is continuously extendable to the corresponding boundary). Only, an integrability condition applies, which is  $U(\cdot, t), \frac{\partial}{\partial t}U(\cdot, t), f \in L^2(\bigcup_{k=0}^n \partial_k \Delta_n)$  for  $t \geq 0$ .

Summarising our findings, we have:

**4.3 Lemma.** *A function  $U: (\overline{\Delta}_n)_\infty \rightarrow \mathbb{R}$ ,  $U(\cdot, t), \frac{\partial}{\partial t}U(\cdot, t) \in L^2(\bigcup_{k=0}^n \partial_k \Delta_n)$  for  $t \geq 0$  with corresponding moments  $\bar{\mu}_\alpha(t) = [U(t), x^\alpha]_n$ ,  $\alpha = (\alpha_1, \dots, \alpha_n)$ ,  $t \geq 0$  fulfilling the moments evolution equation (4.3) also solves the weak formulation of the Kolmogorov forward equation (4.8) and conversely.*

## 4.2 The boundary flux and a hierarchical extension of solutions

In order to construct suitable boundary values as required for an extended solution  $U: (\overline{\Delta}_n)_\infty \rightarrow \mathbb{R}$ , we shall now introduce the concept of boundary flux. The investigation of the boundary flux as the basis for a hierarchical solution scheme is the main new contribution of this paper.

The flux  $G_u: (\Delta_n)_\infty \rightarrow \mathbb{R}^n$  of a solution  $u: (\Delta_n)_\infty \rightarrow \mathbb{R}^n$  of equation (3.1) is given in terms of its components

$$G_u^i(x, t) := -\frac{1}{2} \sum_{j=1}^n \frac{\partial}{\partial x^j} (x^i (\delta_j^i - x^j) u(x, t)) = -\frac{1}{2} \sum_{j=1}^n \frac{\partial}{\partial x^j} (a^{ij} u(x, t)), \quad i = 1, \dots, n. \quad (4.10)$$

In particular, we have

$$\operatorname{div} G_u = \sum_{i=1}^n \frac{\partial}{\partial x^i} G_u^i = -L_n u = -u_t. \quad (4.11)$$

Again, this concept directly extends to boundary instances of  $\overline{\Delta}_n$  if  $u$  is extendable to the boundary such that the extension is of class  $C^2$  with respect to the spatial variables (which is the case for a solution as in proposition 4.2).

With the flux at hand, we may now state the generalised form of lemma 3.1, which gives the adjointness relation for the Kolmogorov operators  $L_n$  and  $L_n^*$  with non-vanishing boundary terms:

**4.4 Proposition.** *For  $n \in \mathbb{N}_+$  and  $u, \varphi \in C^2(\overline{\Delta}_n)$ , we have*

$$(L_n u, \varphi)_n = - \int_{\partial_{n-1} \Delta_n} \varphi G_u \cdot \nu \, d\lambda_{n-1} + (u, L_n^* \varphi)_n \quad (4.12)$$

with  $G_u$  being the flux of  $u$  and  $\nu$  the outward unit surface normal to  $\partial \Delta_n$ .

*Proof.* We use the integration by parts formula

$$\int_{\Omega} \frac{\partial u}{\partial x^i} \varphi \, d\Omega = \int_{\partial \Omega} \varphi u \nu^i \, d\partial \Omega - \int_{\Omega} u \frac{\partial \varphi}{\partial x^i} \, d\Omega, \quad (4.13)$$

holding for a domain  $\Omega$  with piecewise continuous boundary  $\partial\Omega$ ,  $u, \varphi \in C^1(\overline{\Omega})$  and  $\nu^i$  being the  $i$ -th component of the outward unit surface normal to  $\partial\Omega$ . This yields

$$\begin{aligned} (L_n u, \varphi)_n &= - \int_{\Delta_n} \sum_i \frac{\partial}{\partial x^i} G_u^i \varphi d\mathbb{A}_n \\ &= - \int_{\partial\Delta_n} \sum_i G_u^i \nu^i \varphi d\mathbb{A}_{n-1} + \int_{\Delta_n} \sum_i G_u^i \frac{\partial}{\partial x^i} \varphi d\mathbb{A}_n. \end{aligned} \quad (4.14)$$

$\bigcup_{k=0}^{n-2} \partial_k \Delta_n$  clearly is a null set with respect to  $\mathbb{A}_{n-1}$ , and we may hence replace the domain of integration at the first summand by  $\partial_{n-1} \Delta_n$ . For the second term, we apply the integration by parts formula again (also using the modified domain of integration):

$$\begin{aligned} \int_{\Delta_n} \sum_i G_u^i \frac{\partial}{\partial x^i} \varphi d\mathbb{A}_n &= - \int_{\partial_{n-1} \Delta_n} \frac{1}{2} \sum_{i,j} x^i (\delta_j^i - x^j) u \nu^j \frac{\partial}{\partial x^i} \varphi d\mathbb{A}_{n-1} \\ &\quad + \int_{\Delta_n} \frac{1}{2} \sum_{i,j} a^{ij} u \frac{\partial^2}{\partial x^i \partial x^j} \varphi d\mathbb{A}_n. \end{aligned} \quad (4.15)$$

For the boundary integral over  $\partial_{n-1} \Delta_n = \bigcup_{l=0}^n \Delta_{n-1}^{(I_n \setminus \{l\})}$ , we have  $\nu^j = -\delta_l^j$  on  $\Delta_{n-1}^{(I_n \setminus \{l\})}$ ,  $l = 1, \dots, n$  and  $\nu^j = \frac{1}{\sqrt{n}}$  on  $\Delta_{n-1}^{(I_n \setminus \{0\})}$ , which yields

$$\sum_j x^i (\delta_j^i - x^j) u \nu^j = -x^i (\delta_l^i - x^l) u = 0 \quad \text{on } \Delta_{n-1}^{(I_n \setminus \{l\})} = \{x^l = 0\} \quad (4.16)$$

and

$$\begin{aligned} \sum_j x^i (\delta_j^i - x^j) u \nu^j &= \frac{1}{\sqrt{n}} \sum_j x^i (\delta_j^i - x^j) u \\ &= \frac{1}{\sqrt{n}} x^i \left(1 - \sum_j x^j\right) u = 0 \quad \text{on } \Delta_{n-1}^{(I_n \setminus \{0\})} = \left\{1 - \sum_j x^j = 0\right\}. \end{aligned} \quad (4.17)$$

Thus, the second integral over  $\partial_{n-1} \Delta_n$  vanishes. Altogether, we have

$$\begin{aligned} (L_n u, \varphi)_n &= - \int_{\partial_{n-1} \Delta_n} \sum_i G_u^i \nu^i \varphi d\mathbb{A}_{n-1} + \int_{\Delta_n} u \frac{1}{2} \sum_{i,j} a^{ij} \frac{\partial^2}{\partial x^i \partial x^j} \varphi d\mathbb{A}_n \\ &= - \int_{\partial_{n-1} \Delta_n} G_u \cdot \nu \varphi d\mathbb{A}_{n-1} + (u, L_n^* \varphi)_n. \quad \square \end{aligned} \quad (4.18)$$

If  $\varphi: \overline{\Delta_n} \rightarrow \mathbb{R}$  is a polynomial of degree less than 2, we have  $L^* \varphi = 0$ . Integrating the flux  $G_u$  on  $\partial_{n-1} \Delta_n$  over time as boundary values for a solution  $u$  of equation (3.1) (resp.

for its continuous extension to  $\partial\Delta_n$ ), proposition 4.4 already yields the behaviour for the 0th and the 1st moment which is prescribed by the moments evolution equation (4.4). Thus, the total mass and the expectation value of the process are preserved.

This concept of a solution in  $\Delta_n$  plus accumulated flux on the boundary  $\partial_{n-1}\Delta_n$  is, however, not yet sufficient. In general it does not yield the desired evolution laws for moments of degree 2 or higher, nor does  $\partial_{n-1}\Delta_n$  account for the full boundary  $\partial\Delta_n$ . To resolve this, instead of accumulating the incoming flux on  $\partial_{n-1}\Delta_n$  for  $n \geq 2$  statically, we assume that it rather evolves as if it was an  $(n-1)$ -dimensional Wright–Fisher process, i. e. as a subsolution on  $\partial_{n-1}\Delta_n$ . Carrying forward the construction of boundary data to the boundary instance of subsequent lower dimension by assessing the respective boundary flux of the subsolutions on each  $\partial_{n-2}\Delta_{n-1}$  and continuing successively leads us to

**4.5 Definition.** For  $\Delta_n^{(I_n)}$  with  $I_n = \{0, 1, \dots, n\}$  and a solution  $u: (\Delta_n^{(I_n)})_\infty \rightarrow \mathbb{R}$  of the Kolmogorov forward equation (3.1) for given  $f: \Delta_n^{(I_n)} \rightarrow \mathbb{R}$  as in proposition 4.2, a *hierarchical extension*

$$U: (\overline{\Delta_n^{(I_n)}})_\infty \rightarrow \mathbb{R} \quad \text{with} \quad U(x, t) := \sum_{k=0}^n U_k(x, t) \chi_{\partial_k \Delta_n^{(I_n)}}(x) \quad (4.19)$$

is given by

$$U_k: (\partial_k \Delta_n^{(I_n)})_\infty \rightarrow \mathbb{R} \quad \text{with} \quad U_k(x, t) := \begin{cases} u(x, t) & \text{for } x \in \Delta_n^{(I_n)} \equiv \partial_n \Delta_n^{(I_n)} \\ U_{k, I_k}(x, t) & \text{for } x \in \Delta_k^{(I_k)} \subset \partial_k \Delta_n^{(I_n)}, I_k \subset I_n \\ 0 & \text{else} \end{cases} \quad (4.20)$$

for all  $0 \leq k \leq n$  and

$$U_{k, I_k}: (\Delta_k^{(I_k)})_\infty \rightarrow \mathbb{R} \quad \text{with} \quad U_{k, I_k}(x, t) := \int_0^t u_{k, I_k}^\tau(x, t - \tau) d\tau \quad (4.21)$$

for  $0 \leq k \leq n-1$  and for all subsets  $I_k \subset I_n$ . Here,  $u_{k, I_k}^\tau(x, t): (\Delta_k^{(I_k)})_\infty \rightarrow \mathbb{R}$  is a solution of

$$\begin{cases} L_k u(x, t) = \frac{\partial}{\partial t} u(x, t) & (x, t) \in (\Delta_k^{(I_k)})_\infty \\ u(x, 0) = \sum_{I_{k+1} \supset I_k} G_{U_{k+1, I_{k+1}}}^\perp(x, \tau) & x \in \Delta_k^{(I_k)} \end{cases} \quad (4.22)$$

for all  $\tau > 0$  as in proposition 4.2 and  $G_{U_{k+1, I_{k+1}}}^\perp$  is the normal component of the flux of the continuous extension of  $U_{k+1, I_{k+1}}$  to  $\overline{\Delta_{k+1}^{(I_{k+1})}}$ .

**4.6 Remark.** For a given solution  $u$  of equation (3.1), the induced boundary functions  $U_k$  on  $\partial_k \Delta_n^{(I_n)}$  for  $0 \leq k \leq n-1$  in general do not fulfil the equation  $\frac{\partial}{\partial t} U_k = L_k U_k$  in some  $\Delta_k^{(I_k)} \subset \partial_k \Delta_n^{(I_n)}$  and consequently are not solutions of the corresponding  $k$ -dimensional problem (3.1) in  $\Delta_k^{(I_k)}$ .

### 4.3 An application of the hierarchical conception

For the hierarchically extended solution and the product  $[\cdot, \cdot]_n$ , we may now continue the line of reasoning of lemma 3.1 and proposition 4.4.

**4.7 Proposition.** *A hierarchical extension  $U: \overline{(\Delta_n^{(I_n)})}_\infty \rightarrow \mathbb{R}$  (cf. definition 4.5) of a solution  $u$  of the Kolmogorov forward equation (3.1) in  $\Delta_n$  satisfies*

$$\left[ \frac{\partial}{\partial t} U(t), \varphi \right]_n = [U(t), L^* \varphi]_n \quad (4.23)$$

for  $\varphi \in C^\infty(\overline{(\Delta_n^{(I_n)})})$  and for all  $t \in (0, \infty)$ .

*Proof.* By proposition 4.4 we have for  $U_n \equiv u$  and arbitrary  $\varphi \in C^\infty(\overline{(\Delta_n^{(I_n)})})$

$$\left( \frac{\partial}{\partial t} U_n, \varphi \right)_n = (L_n U_n, \varphi)_n = - \int_{\partial_{n-1} \Delta_n^{(I_n)}} \varphi G_{U_n}^\perp d\lambda_{n-1} + (U_n, L_n^* \varphi)_n \quad (4.24)$$

with  $G_{U_n}^\perp = G_{U_n} \cdot \nu$  denoting the (normal) flux of the continuous extension of  $U_n$  to  $\partial_{n-1} \Delta_n^{(I_n)}$ . The boundary integral may be expressed in terms of the evolution of the boundary function  $U_{n-1}$  that lives on  $\partial_{n-1} \Delta_n^{(I_n)}$ . As this implies a hierarchical dependence on the particular subprocesses, we directly start our consideration for arbitrary  $k \in \{1, \dots, n\}$ . Then we have by proposition 4.4 and for all  $I_k \subset I_n$

$$\int_{\Delta_k^{(I_k)}} (L_k U_{k, I_k}) \varphi d\lambda_k = - \int_{\partial_{k-1} \Delta_k^{(I_k)}} \varphi G_{U_{k, I_k}}^\perp d\lambda_{k-1} + \int_{\Delta_k^{(I_k)}} U_{k, I_k} L_k^* \varphi d\lambda_k \quad (4.25)$$

with  $G_{U_{k, I_k}}$  again denoting the flux of the continuous extension of  $U_{k, I_k}$  to  $\partial_{k-1} \Delta_k^{(I_k)}$  (not to be confused with the proper boundary function  $U_{k-1}$  on  $\partial_{k-1} \Delta_n^{(I_n)}$ ). This may be summarised for the whole  $k$ -dimensional boundary  $\partial_k \Delta_n^{(I_n)}$  of  $\Delta_n^{(I_n)}$  by summing over all  $\Delta_k^{(I_k)} \subset \partial_k \Delta_n^{(I_n)}$  resp. all subsets  $I_k \subset I_n$ . This yields (because of  $\bigcup_{I_k \subset I_n} \Delta_k^{(I_k)} = \partial_k \Delta_n^{(I_n)}$  and the definition of  $U_k$ )

$$\int_{\partial_k \Delta_n^{(I_n)}} (L_k U_k) \varphi d\lambda_k = \sum_{I_k \subset I_n} \int_{\partial_{k-1} \Delta_k^{(I_k)}} \varphi G_{U_{k, I_k}}^\perp d\lambda_{k-1} + \int_{\partial_k \Delta_n^{(I_n)}} U_k L_k^* \varphi d\lambda_k. \quad (4.26)$$

Transforming the boundary term using  $\bigcup_{I_k \subset I_n} \partial_{k-1} \Delta_k^{(I_k)} = \bigcup_{I_{k-1} \subset I_n} \Delta_{k-1}^{(I_{k-1})}$  and employing the product notation, we get

$$(L_k U_k, \varphi)_k = \sum_{I_{k-1} \subset I_n} \int_{\Delta_{k-1}^{(I_{k-1})}} \varphi \sum_{I_k \supset I_{k-1}} G_{U_{k, I_k}}^\perp d\lambda_{k-1} + (U_k, L_k^* \varphi)_k. \quad (4.27)$$



Now, the sum of fluxes appearing here may be expressed in terms of the evolution of the associated boundary function  $U_{k-1, I_{k-1}}$  on  $\Delta_{k-1}^{(I_{k-1})}$  for every  $I_{k-1} \subset I_n$ . By the chain rule, we have on  $\Delta_{k-1}^{(I_{k-1})}$

$$\begin{aligned} \frac{\partial}{\partial t} U_{k-1, I_{k-1}}(x, t) &= \frac{\partial}{\partial t} \int_0^t u_{k-1, I_{k-1}}^\tau(x, t - \tau) d\tau \\ &= u_{k-1, I_{k-1}}^\tau(x, t - \tau) \Big|_{\tau=t} + \int_0^t \frac{\partial}{\partial t} u_{k-1, I_{k-1}}^\tau(x, t - \tau) d\tau \\ &= u_{k-1, I_{k-1}}^t(x, 0) + \int_0^t L_{k-1} u_{k-1, I_{k-1}}^\tau(x, t - \tau) \end{aligned} \quad (4.28)$$

by the solution property of  $u_{k-1, I_{k-1}}^\tau$ . Interchanging  $L_{k-1}$  with the  $\tau$ -integration and substituting  $u_{k-1, I_{k-1}}^t(x, 0)$  by the initial values as prescribed altogether yields

$$- \sum_{I_k \supset I_{k-1}} G_{U_k, I_k}^\perp(x, t) = - \frac{\partial}{\partial t} U_{k-1, I_{k-1}}(x, t) + L_{k-1} U_{k-1, I_{k-1}}(x, t). \quad (4.29)$$

Multiplying this with  $\varphi$ , integrating over  $\Delta_{k-1}^{(I_{k-1})}$  and summing over all  $I_{k-1} \subset I_n$  results in

$$\begin{aligned} &- \sum_{I_{k-1} \subset I_n} \int_{\Delta_{k-1}^{(I_{k-1})}} \varphi \sum_{I_k \supset I_{k-1}} G_{U_k, I_k}^\perp d\mathfrak{A}_{k-1} \\ &= - \sum_{I_{k-1} \subset I_n} \int_{\Delta_{k-1}^{(I_{k-1})}} \varphi \frac{\partial}{\partial t} U_{k-1, I_{k-1}} d\mathfrak{A}_{k-1} + \sum_{I_{k-1} \subset I_n} \int_{\Delta_{k-1}^{(I_{k-1})}} \varphi L_{k-1} U_{k-1, I_{k-1}} d\mathfrak{A}_{k-1} \\ &= - \left( \frac{\partial}{\partial t} U_{k-1}, \varphi \right)_{k-1} + (L_{k-1} U_{k-1}, \varphi)_{k-1} \end{aligned} \quad (4.30)$$

because of  $\bigcup_{I_{k-1} \subset I_n} \Delta_{k-1}^{(I_{k-1})} = \partial_{k-1} \Delta_n^{(I_n)}$ . Combining this with equation (4.27), we get

$$(L_k U_k, \varphi)_k = - \left( \frac{\partial}{\partial t} U_{k-1}, \varphi \right)_{k-1} + (L_{k-1} U_{k-1}, \varphi)_{k-1} + (U_k, L_k^* \varphi)_k, \quad (4.31)$$

which – by assumption – holds for all  $k \in \{1, \dots, n\}$ . Hence, this formula may be iterated



*Proof.* From proposition 4.1 we already see that the linear combinations of the eigenfunctions of  $L_n$  are dense in  $C^\infty(\overline{\Delta}_n)$ . Dividing a function  $f \in C_c^\infty(\overline{\Delta}_n)$  by  $\omega_n$  (cf. lemma 3.2) again yields a function in  $C_c^\infty(\overline{\Delta}_n) \subset C_0^\infty(\overline{\Delta}_n)$  as  $\omega_n$  is in  $C_0^\infty(\overline{\Delta}_n)$  itself and positive in the interior  $\Delta_n$ .  $\square$

*Proof of proposition 4.9.* Assume that  $U': (\overline{\Delta}_n)_\infty \rightarrow \mathbb{R}$  is another solution of equation (4.8) for a given initial condition  $f$ . We need to show that  $U$  and  $U'$  agree on all  $\partial_k \Delta_n \subset \overline{\Delta}_n$  for  $k = n, \dots, 0$ . We start with  $\partial_n \Delta_n \equiv \Delta_n$ . For an eigenfunction  $\varphi \in C^\infty(\overline{\Delta}_n)$  of  $L_n$  (corresponding to the eigenvalue  $\lambda$ ), we obtain by lemma 3.2 that  $\psi := \omega_n \varphi$  is an eigenfunction of  $L_n^*$  corresponding to the eigenvalue  $\lambda$  and, by the properties of  $\omega_n$ , that  $\psi \in C_0^\infty(\overline{\Delta}_n)$ . For such a  $\psi$ , the weak Kolmogorov forward equation (4.8) then reduces to

$$\left( \frac{\partial}{\partial t} U, \psi \right)_n = (U, L_n^* \psi)_n \equiv -\lambda(U, \psi)_n \quad (4.33)$$

and

$$\left( \frac{\partial}{\partial t} U', \psi \right)_n = (U', L_n^* \psi)_n \equiv -\lambda(U', \psi)_n \quad (4.34)$$

respectively. Consequently, by  $t$ -integration we have

$$(U(t), \psi)_n = e^{-\lambda t} (U(0), \psi)_n, \quad (4.35)$$

$$(U'(t), \psi)_n = e^{-\lambda t} (U'(0), \psi)_n, \quad (4.36)$$

from which we obtain via  $U(0) = U'(0) = f$

$$(U(t), \psi)_n = (U'(t), \psi)_n \quad \text{for } t \geq 0 \quad (4.37)$$

and for all eigenfunctions  $\psi$ . Since the linear span of these functions is dense in  $C_c^\infty(\overline{\Delta}_n)$  (cf. lemma 4.11),  $U$  and  $U'$  agree in  $\Delta_n$ , indeed.

Now, we proceed inductively. Assume that we have already shown that  $U$  and  $U'$  agree on all  $\partial_k \Delta_n \subset \overline{\Delta}_n$  with  $k > m$ . Then for an eigenfunction  $\varphi: \overline{\Delta}_m \rightarrow \mathbb{R}$  of  $L_m$  (corresponding to the eigenvalue  $\lambda$ ),  $\psi := \omega_m \varphi$  again is an eigenfunction of  $L_m^*$  corresponding to the eigenvalue  $\lambda$  and  $\psi \in C_0^\infty(\overline{\Delta}_m)$ . From any such  $\psi: \overline{\Delta}_m \rightarrow \mathbb{R}$ , a function  $\overline{\psi}: \overline{\Delta}_n^{(I_n)} \rightarrow \mathbb{R}$  may be composed, e. g. by copying  $\psi$  to  $\Delta_m^{(I_m)} \subset \partial_m \Delta_n$  for all  $I_m \subset I_n$  and employing convex combinations of the boundary values to spread to all higher dimensional (boundary) instances subsequently while putting  $\overline{\psi} := 0$  on all lower dimensional boundary instances. Of course,  $\overline{\psi}$  is generally not an eigenfunction of  $L^*$  in  $\overline{\Delta}_n$ , but we still have  $(L^* \overline{\psi})|_{\Delta_m^{(I_m)}} = L_m^* \psi = -\lambda \psi$  for all  $\Delta_m^{(I_m)} \subset \partial \Delta_n$ .

For such a  $\overline{\psi}$ , the weak Kolmogorov forward equation (4.8) is converted into

$$\left( \frac{\partial}{\partial t} U, \overline{\psi} \right)_m = -(U, L_n^* \overline{\psi})_m + \sum_{k=m+1}^n \left( \left( \frac{\partial}{\partial t} U, \overline{\psi} \right)_k - (U, L_k^* \overline{\psi})_k \right) \quad (4.38)$$

and

$$\left(\frac{\partial}{\partial t}U', \bar{\psi}\right)_m = -(U', L_n^*\bar{\psi})_m + \sum_{k=m+1}^n \left(\left(\frac{\partial}{\partial t}U', \bar{\psi}\right)_k - (U', L_k^*\bar{\psi})_k\right) \quad (4.39)$$

with the sums on the right agreeing as  $U = U'$  on all  $\partial_k\Delta_n$  with  $k > m$ , hence

$$\left(\frac{\partial}{\partial t}(U - U'), \bar{\psi}\right)_m = (U' - U, L_n^*\bar{\psi})_m \equiv \lambda(U' - U, \bar{\psi})_m, \quad (4.40)$$

which yields – analogously to our considerations above –  $U = U'$  in  $\partial_m\Delta_n$  on account of the completeness of the  $\bar{\psi}$ 's and the initial condition.  $\square$

Thus, with the additional assumptions that the moments of the process coincide with the limits of the moments of the underlying discrete processes, we altogether have:

**4.12 Theorem.** *For  $n \in \mathbb{N}$  and a given initial condition  $f \in L^2(\Delta_n)$ , the Kolmogorov forward equation corresponding to the diffusion approximation of the  $n$ -dimensional Wright–Fisher model (3.1) always allows a unique extended solution  $U: (\bar{\Delta}_n)_\infty \rightarrow \mathbb{R}$  in the sense that  $U|_{\Delta_n}$  is a solution of equation (3.1) and that its moments  $\bar{\mu}_\alpha(t) := [U(t), x^\alpha]_n$ ,  $t \geq 0$  (cf. equation (4.4)) satisfy the  $n$ -dimensional moments evolution equation (4.3).*

### Acknowledgements

The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n° 267087. J.H. and T.D.T. have also been supported by scholarships from the IMPRS “Mathematics in the Sciences” during earlier stages of this work.

### References

- [1] S. Amari, H. Nagaoka, *Methods of information geometry*, Translations of mathematical monographs; v. **191**, American Mathematical Society, 2000.
- [2] N. Ay, J. Jost, H. V. Lê, L. Schwachhöfer, *Information geometry*, Monograph, to appear.
- [3] G. J. Baxter, R. A. Blythe, A. J. McKane. *Exact Solution of the Multi-Allelic Diffusion Model*, *Math. Biosci.*, **209** (1) (2007) 124–170.
- [4] R. Bürger, *The mathematical theory of selection, recombination, and mutation*, Wiley series in Mathematical and Computational Biology, 2000.
- [5] D. Dawson, K. Hochberg, *Wandering random measures in the Fleming-Viot model*, *Annals Prob.* **10** (1982), 554–580.

- [6] E. Dynkin, *Three classes of infinite-dimensional diffusions*, J. Funct. Anal. **86** (1989), 75–110.
- [7] C. L. Epstein, R. Mazzeo, *Wright-Fisher diffusion in one dimension*, SIAM J. Math. Anal. **42** (2010), 568–608.
- [8] C. L. Epstein, R. Mazzeo, *Degenerate diffusion operators arising in population biology*, Princeton Univ. Press, 2013
- [9] S. N. Ethier, *A class of degenerate diffusion processes occurring in population genetics*, Comm. Pure Appl. Math. **29** (1976), 483–493.
- [10] S.N. Ethier, R.C. Griffiths, *The Transition Function of a Fleming-Viot Process*, Ann. Probab. **21** (1993) 1571–1590.
- [11] S. N. Ethier, T. Kurtz, *Markov processes: characterization and convergence*, Wiley series in Probability and Mathematical Statistics, 2005.
- [12] S. N. Ethier, T. Nagylaki, *Diffusion Approximations of Markov Chains with Two Time Scales and Applications to Population Genetics*, Adv. Appl. Probab. **12** (1980), 14–49.
- [13] S. N. Ethier, T. Nagylaki, *Diffusion approximations of the two-locus Wright-Fisher model*, J. Math. Biol. **27** (1989), 17–28.
- [14] W. J. Ewens, *Mathematical Population Genetics I. Theoretical Introduction*, Springer-Verlag New York Inc., Interdisciplinary Applied Mathematics, 2nd ed., 2004.
- [15] R. A. Fisher, *On the dominance ratio*, Proc. Roy. Soc. Edinb., **42** (1922), 321–341.
- [16] R. C. Griffiths, *A Transition Density Expansion for a Multi-Allele Diffusion Model*, Adv. Appl. Probab. **11** (1979) 310–325.
- [17] R. C. Griffiths, *Lines of descent in the diffusion approximation of neutral Wright-Fisher models*, Theor. Pop. Biol. **17** (1980) 37–50.
- [18] R. C. Griffiths, D. Spanó, *Diffusion processes and coalescent trees*, Chapter 15, pp. 358-375. In: Probability and Mathematical Genetics, Papers in Honour of Sir John Kingman. LMS Lecture Note Series 378. ed Bingham, N. H. and Goldie, C. M., Cambridge University Press, 2010.
- [19] R. C. Griffiths, D. Spanó, *Multivariate Jacobi and Laguerre polynomials, infinite-dimensional extensions, and their probabilistic connections with multivariate Hahn and Meixner polynomials*, Bernoulli **17** (2011) 1095–1125.
- [20] R. C. Griffiths, D. Spanó, *Orthogonal polynomial kernels and canonical correlations for Dirichlet measures*, Bernoulli **19** (2013) 548–498.
- [21] J. Hofrichter, *On the diffusion approximation of Wright-Fisher models with several alleles and loci and its geometry*, PhD thesis, University of Leipzig, 2014.

- [22] J. Jost, *Partial Differential Equations*, Graduate Texts in Mathematics, v. **214** Springer, Berlin, Heidelberg, 3rd edition, 2013.
- [23] J. Jost, *Mathematical methods in biology and neurobiology*, Universitext, Springer, London, Heidelberg, 2014
- [24] S. Karlin, H. M. Taylor, *A second course in stochastic processes*, Academic Press, 1981
- [25] M. Kimura, *Solution of a Process of Random Genetic Drift with a Continuous Model*, PNAS–USA, Vol. **41**, No. 3, (1955), 144–150.
- [26] M. Kimura, *Random genetic drift in multi-allele locus*, *Evolution*, **9** (1955), 419–435.
- [27] M. Kimura, *Random genetic drift in a tri-allelic locus; exact solution with a continuous model*, *Biometrics*, **12** (1956), 57–66.
- [28] J. F. C. Kingman, *The coalescent*, *Stoch. Proc. Appl.* **13** (1982) 235–248.
- [29] R. A. Littler, E. D. Fackerell, *Transition Densities for Neutral Multi-Allele Diffusion Models*, *Biometrics* Vol **31** No 1 (1975) 117–123.
- [30] K. Sato, *Diffusion operators in population genetics and convergence of Markov chains*, In *Measure Theory Application to Stochastic Analysis*, LNM **695** (1978) 127–137.
- [31] N. Shimakura, *Equations différentielles provenant de la génétique des populations*, *Tohoku Math. J.*, **29** (1977) 287–318.
- [32] N. Shimakura, *Formulas for diffusion approximations of some gene frequency models*, *J. Math. Kyoto Univ.* **21** (1981) 19–45.
- [33] S. Tavaré, *Line-of-descent and genealogical processes, and their applications in population genetics models*, *Theor. Popn. Biol.* **26** (1984) 119–164.
- [34] T. D. Tran. *Information Geometry and the Wright-Fisher Model of Mathematical Population Genetics*. PhD thesis, University of Leipzig, 2012.
- [35] T. D. Tran, J. Hofrichter, J. Jost, *An introduction to the mathematical structure of the Wright-Fisher model of population genetics*, *Theory Biosc.* **132** (2013), 73–82.
- [36] T. D. Tran, J. Hofrichter, J. Jost, *A general solution of the Wright-Fisher model of random genetic drift*, submitted.
- [37] T. D. Tran, J. Hofrichter, J. Jost, *The evolution of moment generating functions for the Wright-Fisher model of population genetics*, submitted.
- [38] S. Wright, *Evolution in Mendelian populations*, *Genetics*, **16** (1931), 97–159.
- [39] S. Wright, *Adaptation and selection*, In: G.Jepson, E.Mayr, G.Simpson (eds.), *Genetics, Paleontology, and Evolution*, pp.365–389, Princeton Univ.Press, 1949.