EIGENFUNCTIONS AND THE DIRICHLET PROBLEM FOR THE CLASSICAL KIMURA DIFFUSION OPERATOR

CHARLES L. EPSTEIN* AND JON WILKENING†

Abstract. We study the classical Kimura diffusion operator defined on the n-simplex,

$$\mathrm{L}^{\mathrm{Kim}} = \sum_{1\leq i,j\leq n+1} x_i(\delta_{ij} - x_j)\partial_{x_i}\partial_{x_j},$$

which has important applications in Population Genetics. Because it is a degenerate elliptic operator acting on a singular space, special tools are required to analyze and construct solutions to elliptic and parabolic problems defined by this operator. The natural boundary value problems are the "regular" problem and the Dirichlet problem. For the regular problem, one can only specify the regularity of the solution at the boundary. For the Dirichlet problem, one can specify the boundary values, but the solution is then not smooth at the boundary. In this paper we give a computationally effective recursive method to construct the eigenfunctions of the regular operator in any dimension, and a recursive method to use them to solve the inhomogeneous equation. As noted, the Dirichlet problem does not have a regular solution. We give an explicit construction for the leading singular part along the boundary. The necessary correction term can then be found using the eigenfunctions of the regular problem.

Key words. Kimura Diffusion, Population Genetics, Degenerate Elliptic Equations, Dirichlet Problem, Eigenfunctions

AMS subject classifications. 35J25, 35J70, 33C50, 65N25

1. Introduction.

An n-simplex, Σ_n, is the subset of $\mathbb{R}^{n+1}$ given, in the *affine model*, by the relations:

$$x_1 + \cdots + x_{n+1} = 1 \text{ with } 0 \leq x_i \text{ for } 1 \leq i \leq n+1. \tag{1.1}$$

In population genetics problems, a point $(x_1, \ldots, x_{n+1}) \in \Sigma_n$ is often thought of as representing the frequencies of $n+1$ alleles or types. More generally, Σ_n can be thought of as the space of atomic probability measures with $n+1$ atoms. Mathematically, Σ_n is a *manifold with corners*; its boundary is a stratified space made up of simplices of dimensions between 0 and $n-1$.

The Kimura diffusion operator, which acts on functions of $n+1$ variables,

$$\mathrm{L}^{\mathrm{Kim}} = \sum_{1\leq i,j\leq n+1} x_i(\delta_{ij} - x_j)\partial_{x_i}\partial_{x_j} \tag{1.2}$$

appears in the infinite population limit of the $(n+1)$-allele Wright-Fisher model. It represents the limit of the random mating term, and actually appears in the infinite population limits of many Markov chain models in population genetics, see [12, 30, 16, 5]. The Kimura diffusion operator has many remarkable properties, which we employ in our analysis. The properties of this operator reflect the geometry of the simplex in much the same way as the standard Laplace operator reflects the Euclidean geometry of $\mathbb{R}^n$.

To include the effects of mutation, selection, migration, etc. the operator is modified by the addition of a vector field

$$V = \sum_{j=1}^{n+1} b_j(x)\partial_{x_j}, \tag{1.3}$$

*Department of Mathematics, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104-6395; cle@math.upenn.edu. Research supported in part by the NSF under grants DMS12-05851, and DMS-1507396; and by the ARO under grant W911NF-12-1-0552.

†Department of Mathematics, 970 Evans Hall, University of California, Berkeley, California 94720-3840; wilken@math.berkeley.edu. Research supported in part by the US Department of Energy, Office of Science, Applied Scientific Computing Research, under award number DE-AC02-05CH11231.

which is tangent to Σ_n, and inward pointing along the boundary of the simplex. In applications to population genetics, the coefficient functions $\{b_j(x)\}$ are typically polynomials. Linear terms usually suffice to model migration and mutational effects, whereas higher order terms are needed to model selection, see [30, 33, 39].

There are many statistical quantities of interest in population genetics that can be computed by solving boundary value problems of the form

$$(\mathrm{L}^{\mathrm{Kim}} + V)u = f \quad \text{in} \quad \operatorname{int}\Sigma_n, \quad \text{with} \quad u\restriction_{b\Sigma_n} = g. \tag{1.4}$$

For example, the probability of a path of the underlying process exiting through a given portion of the boundary, or the expected first arrival time at a portion of the boundary, are expressible as the solutions of such boundary value problems. Examples of this type can be found in [34], and are further discussed below.

A method for solving some of these problems, at least in principle, is given in [38], though it is not very explicit. In this note we introduce a computationally effective method for solving high-dimensional, inhomogeneous regular and Dirichlet problems for the Kimura operator itself, i.e. with $V = 0$. For the regular problem, one specifies minimal regularity requirements for the solution at the boundary. For the Dirichlet problem, the boundary values are specified, but the solution is then not smooth at the boundary. Our method also clarifies the precise regularity of the Dirichlet solution in the closed simplex, at least when f and g are sufficiently smooth. In addition we show, in somewhat greater generality, how to find the eigenfunctions of these operators, which are represented as products of functions of single variables, and how to compute the expansion coefficients.

The operator, $\mathrm{L}^{\mathrm{Kim}} + V$, and variants thereof, appear in many classical papers in population genetics, see [12, 13, 27, 26, 28, 45]. Recently, there has been a resurgence of interest in using the Kimura diffusion equation as a forward model for maximum likelihood estimators of selection coefficients, demographic models, mutation rates, effective population sizes, etc. The evolution of other observable measures of genetic variability such as the allele frequency spectrum, or site frequency spectrum, can also be shown to satisfy a variant of the Kimura diffusion equation, see [11, 19]. To use this diffusion process as a forward model, one either needs to have efficient means for solving the Kimura diffusion equation, see [19, 41, 3, 32], or one must simulate the underlying stochastic process, [11, 10]. In most previous work where the Kimura diffusion equation is solved, the underlying space is 1-dimensional. Even in one dimension, many authors employ numerical methods that rely on finite difference approximations. These are, however, not reliable for imposing the subtle boundary conditions that arise with degenerate operators like $\mathrm{L}^{\mathrm{Kim}}$, and can in fact lead to errors; see [24], and the supplement to [2]. By contrast, our approach provides a stable construction, mathematically equivalent to a Gram-Schmidt procedure, of bases of eigenfunctions. These can be used to accurately solve both elliptic and parabolic problems, as well as compute approximations to the heat kernel itself, which are of central importance in a variety of applications; see [40, 41, 42].

Our construction for the polynomial eigenfunctions of $\mathrm{L}^{\mathrm{Kim}} + V$ is applicable provided that the operator has "constant weights," see (2.15) below. The case of positive weights has been studied extensively in the literature. Kimura [29] and Karlin and McGregor [25] used hypergeometric functions to study the two-dimensional case. Littler and Fackerell [35] generalized Karlin and McGregor's approach to higher dimensions using bi-orthogonal polynomial systems. Griffiths [16] showed that the polynomial eigenfunctions of the diffusion operator corresponding to a repeated eigenvalue can be orthogonalized and grouped together to form reproducing kernel polynomials that appear in the transition function expansion of the diffusion process. One way to represent the orthogonal polynomials in the reproducing kernels is via a triangular construction of Proriol [36] and Koornwinder [31] for multivariate Jacobi polynomials. Griffiths and Spanó also discuss this construction [17], and provide probabilistic connections to multivariate versions of several families of classical orthogonal polynomials, including the Jacobi polynomials that arise here [18]. In the present work, we present a direct construction of polynomial eigenfunctions for the $V = 0$ case, so that it is not necessary to take limits of the positive weight case as the mutation rates approach zero [16]. One novelty that arises is that when $V = 0$, as functions on the n-simplex, these eigenfunctions do not belong to a single L^2-space, but each belongs to an L^2-space of some stratum of the boundary. Their coefficients in the representation of a function in terms of this spanning set are computed as inner products on these lower dimensional strata.

The constructions presented here can serve as the foundation for a perturbative method for solving Kimura-type diffusions with a more complicated vector field, also modeling selection. We will return to these and other elaborations of the theory presented here in a subsequent publication.

Acknowledgements

CLE is grateful for many useful and informative discussions with Camelia Pop, Rafe Mazzeo, and Yun Song on questions connected to the contents of this paper. We are also very grateful for the careful reading and detailed suggestions made by the referees of this paper.

2. Some Facts about the Kimura Diffusion Operator. We begin our analysis by reviewing some of the remarkable properties of $\mathrm{L}^{\mathrm{Kim}}$. A very important fact about $\mathrm{L}^{\mathrm{Kim}}$ is the result of K. Sato [37], which states that if U is a $\mathcal{C}^2$–function that vanishes on Σ_n, then $[\mathrm{L}^{\mathrm{Kim}} U]\restriction_{\Sigma_n} = 0$ as well. Thus we can start with a function $U(x_1, \dots, x_{n+1})$ defined on Σ_n, and extend it to be independent of any one of its arguments, say x_j. This uniquely defines a function on the projection of the simplex to the hyperplane $\{x_j = 0\}$, and vice versa. For definiteness we take $j = n + 1$. This gives a function

$$u(x_1, \dots, x_n) = U\big(x_1, \cdots, x_n, 1 - (x_1 + \cdots + x_n)\big) \tag{2.1}$$

defined on a *projective model* of the simplex:

$$\widetilde{\Sigma}_n = \{x \in \mathbb{R}^n : \ x_1 + \cdots + x_n \leq 1 \text{ with } x_i \geq 0 \text{ for } i = 1, \dots, n\}. \tag{2.2}$$

To compute $\mathrm{L}^{\mathrm{Kim}} U\restriction_{\Sigma_n}$, we can apply $\mathrm{L}^{\mathrm{Kim}}$ in n-variables to u :

$$\mathrm{L}^{\mathrm{Kim}} U(x_1, \dots, x_{n+1})\restriction_{\Sigma_n} = \sum_{1 \leq i,j \leq n} x_i(\delta_{ij} - x_j)\partial_{x_i}\partial_{x_j} u(x_1, \dots, x_n). \tag{2.3}$$

Here and in the sequel, when using a projective model, we let $x_{n+1} = 1-(x_1+\cdots+x_n)$. The projective model is useful for computations, whereas the affine model shows that this operator is entirely symmetric under permutations of the variables $(x_1, \dots, x_{n+1})$. In particular, it makes clear that, from the perspective of $\mathrm{L}^{\mathrm{Kim}}$, all the vertices of Σ_n are "geometrically" identical, something that is not evident in the projective model.

In this note we further investigate some remarkable properties of the operator $\mathrm{L}^{\mathrm{Kim}}$, which are hinted at in [38, 39]. We first consider a recursive construction of the polynomial eigenfunctions of $\mathrm{L}^{\mathrm{Kim}}$. After that we show how to solve the inhomogeneous Dirichlet problem for $\mathrm{L}^{\mathrm{Kim}}$, given in (1.4). It follows easily from Sato's theorem that this problem does not generally have a solution in $\mathcal{C}^2(\Sigma_n)$, even if f and g are both in $\mathcal{C}^\infty$. Our method of solution exhibits the precise form of the singularities along the various boundary strata.

2.1. Vector fields. A generalized Kimura diffusion in Σ_n, with "standard" second order part, is a second order differential operator of the form

$$\widetilde{L} = \sum_{1 \leq i,j \leq n} x_i(\delta_{ij} - x_j)\partial_{x_i}\partial_{x_j} + \sum_{j=1}^{n} \widetilde{b}_j(x)\partial_{x_j}. \tag{2.4}$$

The vector field is normally required to be inward pointing, which means that

$$\widetilde{b}_j(x)\restriction_{x_j=0} \geq 0 \text{ for } j = 1, \dots, n, \tag{2.5}$$

and

$$\sum_{j=1}^{n} \widetilde{b}_j(x)\restriction_{x_1+\dots+x_n=1} \leq 0. \tag{2.6}$$

We denote the first order terms by

$$\widetilde{V} = \sum_{j=1}^{n} \widetilde{b}_j(x)\partial_{x_j}. \tag{2.7}$$

In the affine model

$$L = \sum_{1\leq i,j\leq n+1} x_i(\delta_{ij} - x_j)\partial_{x_i}\partial_{x_j} + V, \qquad V = \sum_{j=1}^{n+1} b_j(x)\partial_{x_j}, \tag{2.8}$$

with the additional condition that

$$\sum_{j=1}^{n+1} b_j(x)\restriction_{x_1+\dots+x_{n+1}=1} = 0. \tag{2.9}$$

Considerable generalizations of this class of operators and spaces are introduced in [7]; though in the present work we concentrate on the classical cases of model operators on simplices and positive orthants, $\mathbb{R}_+^n$. In most of the Probability and Population Genetics literature a different normalization is employed, namely $\mathrm{L}^{\mathrm{Kim}} = \frac{1}{2}\left(\sum x_i(\delta_{ij} - x_j)\partial_{x_i}\partial_{x_j}\right)$. In this paper we use the normalization more common in mathematical analysis, which omits the factor of $1/2$.

The boundary of Σ_n is a stratified space. If K is a boundary face of Σ_n of codimension $n - k$, then it is again a simplex and is represented, in the affine model, by a subset of the form $x_{i_1} + \dots + x_{i_{k+1}} = 1$ with $x_{i_l} \geq 0$ for $l = 1, \dots, k+1$, which implies that $x_{j_1} = \dots = x_{j_{n-k}} = 0$. We set $\mathcal{I} = \{i_1, \dots, i_{k+1}\}$ and

$$\mathcal{J} = \{j_1, \dots, j_{n-k}\} = \{1, \dots, n+1\} \setminus \mathcal{I}. \tag{2.10}$$

We denote this boundary face by $K_{\mathcal{I}}$.

Every boundary face is a simplex, and the formula for the operator analogous to $\mathrm{L}^{\mathrm{Kim}}$ is the same for any boundary face. For example for $K_{\mathcal{I}}$, the sum in (1.2) is simply restricted to the variables $\{x_{i_1}, \dots, x_{i_{k+1}}\}$. We denote this operator by

$$\mathrm{L}_{\mathcal{I}}^{\mathrm{Kim}} = \sum_{i,j\in\mathcal{I}} x_i(\delta_{ij} - x_j)\partial_{x_i}\partial_{x_j}. \tag{2.11}$$

Let v be a $\mathcal{C}^2$-function on $K_{\mathcal{I}}$ and let $\widehat{v}$ denote any $\mathcal{C}^2$-extension of v to the ambient $\mathbb{R}^{n+1}$. K. Sato, in fact, proved, (see [37]) that

$$\mathrm{L}_{\mathcal{I}}^{\mathrm{Kim}}\, v = (\mathrm{L}^{\mathrm{Kim}}\,\widehat{v})\restriction_{K_{\mathcal{I}}}. \tag{2.12}$$

Thus, the restriction or extension of $\mathrm{L}^{\mathrm{Kim}}$ to boundary strata or higher-dimensional simplices is canonical, similar to the connection between the projective and affine models discussed at the beginning of this section. The extension property makes it more natural to label the "last" variable as x_{n+1} rather than x_0, since it may not actually be the last.

In the affine representation, a vector field V (satisfying (2.9)) is tangent to $K_{\mathcal{I}}$ provided that

$$Vx_j\restriction_{x_j=0} = 0, \qquad (j \in \mathcal{J}). \tag{2.13}$$

Condition (2.9) requires that V be tangent to Σ_n itself, i.e. $V(x_1 + \cdots + x_{n+1})\restriction_{\Sigma_n} = 0$. For vector fields, the analogues of Sato's results are obvious: if U vanishes on Σ_n then $VU = 0$, so extending a function U defined on Σ_n to be constant in any coordinate direction leads to an unambiguous value of $VU\restriction_{\Sigma_n}$; and, defining

$$V_{\mathcal{I}} = \sum_{i\in\mathcal{I}} b_i(x)\partial_{x_i}, \tag{2.14}$$

we see that if V is tangent to $K_{\mathcal{I}}$ then $V_{\mathcal{I}}v = (V\widehat{v})\restriction_{K_{\mathcal{I}}}$, where v and $\widehat{v}$ are defined as in (2.12).

2.2. Constant weights and the Hilbert space setting. In addition to $\mathrm{L}^{\mathrm{Kim}}$ there are other classes of "special" Kimura diffusion operators. Classically one singles out operators with constant "weights." (This terminology is introduced in [8].) These operators have the property that the functions $b_j(x)$ (or $\widetilde{b}_j(x)$) are linear and

$$(Vx_j)\restriction_{x_j=0} = b_j(x)\restriction_{x_j=0} = b_j, \tag{2.15}$$

a constant. In the projective model, one may readily show that such a vector field takes the form

$$\widetilde{V}_{\boldsymbol{b}} = \textstyle\sum_{j=1}^{n}(b_j - Bx_j)\partial_{x_j}, \tag{2.16}$$

where

$$\boldsymbol{b} = (b_1, \ldots, b_{n+1}) \ \text{ and } \ B = b_1 + \cdots + b_{n+1}. \tag{2.17}$$

Note that b_{n+1} enters (2.16) through B. We define

$$\mathrm{L}^{\mathrm{Kim}}_{\boldsymbol{b}} = \mathrm{L}^{\mathrm{Kim}} + V_{\boldsymbol{b}}. \tag{2.18}$$

These operators are special because they are self-adjoint with respect to the L^2 inner product on $\widetilde{\Sigma}_n$ defined by the following measure, which (up to normalization) has the Dirichlet density representing the stationary distribution of the underlying Markov process (see e.g. [16, 17, 18])

$$d\mu_{\boldsymbol{b}}(x) = w_{\boldsymbol{b}}(x)\, dx_1 \cdots dx_n, \qquad w_{\boldsymbol{b}}(x) = \prod_{j=1}^{n+1} x_j^{b_j-1}. \tag{2.19}$$

Indeed, in this case $\mathrm{L}^{\mathrm{Kim}}_{\boldsymbol{b}}$ in (2.8) may be written

$$\mathrm{L}^{\mathrm{Kim}}_{\boldsymbol{b}}\, u = \sum_{i<j}^{n+1} \frac{(\partial_{x_i} - \partial_{x_j})[w_{\boldsymbol{b}} x_i x_j (\partial_{x_i} - \partial_{x_j})u]}{w_{\boldsymbol{b}}}, \tag{2.20}$$

where the sum is over all pairs $i, j \in \{1, \dots, n+1\}$ with $i < j$, and, in the projective model

$$\langle \widetilde{\mathrm{L}_{\boldsymbol{b}}^{\mathrm{Kim}}} u, v\rangle = \langle u, \widetilde{\mathrm{L}_{\boldsymbol{b}}^{\mathrm{Kim}}} v\rangle = -\sum_{i<j}^{n+1} \int_{\widetilde{\Sigma}_n} x_i x_j [(\partial_{x_i} - \partial_{x_j})\hat{u}][(\partial_{x_i} - \partial_{x_j})\hat{v}]\, d\mu_{\boldsymbol{b}}(x). \tag{2.21}$$

Here $\hat{u}$, $\hat{v}$ are independent of x_{n+1} and agree with u, v on $\widetilde{\Sigma}_n$. In deriving (2.21), it was assumed that $w_{\boldsymbol{b}} x_i x_j [(\partial_{x_i} - \partial_{x_j})u]v = 0$ and $w_{\boldsymbol{b}} x_i x_j [(\partial_{x_i} - \partial_{x_j})v]u = 0$ on the faces $\{x_i = 0\}$ and $\{x_j = 0\}$, so care must be taken in defining the domain of $\widetilde{\mathrm{L}_{\boldsymbol{b}}^{\mathrm{Kim}}}$ when working in the Hilbert space setting with $\boldsymbol{b} = \mathbf{0}$, see [38]. Note that x_{n+1} is treated as an independent variable in these formulas when computing partial derivatives, but $x_{n+1} = 1-(x_1+\cdots+x_n)$ when evaluating integrals over $\widetilde{\Sigma}_n$. A useful variant of (2.20) is

$$\widetilde{\mathrm{L}_{\boldsymbol{b}}^{\mathrm{Kim}}} u = \sum_{i<j}^{n} \frac{(\partial_{x_i} - \partial_{x_j})[w_{\boldsymbol{b}} x_i x_j (\partial_{x_i} - \partial_{x_j})u]}{w_{\boldsymbol{b}}} + \sum_{i=1}^{n} \frac{\partial_{x_i}(w_{\boldsymbol{b}} x_i x_{n+1} \partial_{x_i} u)}{w_{\boldsymbol{b}}}, \tag{2.22}$$

where x_{n+1} is now treated as a dependent variable, i.e. $\partial x_{n+1}/\partial x_i = -1$. If all the b_j are positive, then $d\mu_{\boldsymbol{b}}(x)$ has finite total mass. In this paper we are primarily interested in the case where all of the b_j vanish, in which case the $\mu_{\boldsymbol{b}}$-volume of Σ_n is infinite.

2.3. The Dirichlet problem and alternative function spaces. In section 6 we present a method for solving the inhomogeneous Dirichlet problem in (1.4). These results have extensions to the case of Kimura diffusions with constant weights, though Dirichlet boundary conditions are only appropriate on faces $\{x_j = 0\}$ for which $b_j < 1$. For simplicity, we focus our attention here on the case when all the weights are zero, which is already of central importance in applications. We also assume that f and g in (1.4) are sufficiently smooth. A different analysis of this problem, employing blow-ups, appears in [20, 21, 22]. The blow-up approach gives a much less explicit description of the singularities that arise when the data is smooth on the simplex itself, but allows for considerably more singular data.

We use the notation and definitions of various function spaces given in the monograph [7]. The principal symbol of the Kimura diffusion operator,

$$P^{\mathrm{Kim}}(\xi) = \sum_{i,j} x_i(\delta_{ij} - x_j)\xi_i \xi_j, \tag{2.23}$$

defines the dual of the *intrinsic* metric on the simplex. This corresponding metric is singular along the boundary and incomplete, i.e., the boundary is at a finite distance from interior points. The distance function on Σ_n defined by this metric is equivalent to

$$\rho_i(x, y) = \sum_{j=1}^{n+1} |\sqrt{x_j} - \sqrt{y_j}|. \tag{2.24}$$

We also use two scales of anisotropic Hölder spaces, $\mathcal{C}_{\mathrm{WF}}^{k,\gamma}(\Sigma_n)$ and $\mathcal{C}_{\mathrm{WF}}^{k,2+\gamma}(\Sigma_n)$, $k \in \mathbb{N}_0, \gamma \in (0, 1)$, introduced in [7], with respect to which the operator $\mathrm{L}^{\mathrm{Kim}}$ has optimal mapping properties. These spaces are defined with respect to the intrinsic metric. For example $f \in \mathcal{C}_{\mathrm{WF}}^{0,\gamma}(\Sigma_n)$ if $f \in \mathcal{C}^0(\Sigma_n)$ and

$$[f]_{0,\gamma,\mathrm{WF}} = \sup_{x \neq y \in \Sigma_n} \frac{|f(x) - f(y)|}{\rho_i(x, y)^\gamma} < \infty. \tag{2.25}$$

If $\lambda < 0$ and $f \in \mathcal{C}_{\mathrm{WF}}^{k,\gamma}(\Sigma_n)$, then the elliptic equation, $(\mathrm{L}^{\mathrm{Kim}} - \lambda)u = f$, has a unique solution, $u \in \mathcal{C}_{\mathrm{WF}}^{k,2+\gamma}(\Sigma_n)$, indicating that this is indeed the correct notion of “elliptic regularity” for operators of this general type.

3. A 1-d Example. We begin by considering the Dirichlet problem in the 1d-case. These results are well known, and serve as motivation for our subsequent development. Suppose that we would like to find the solution to

$$x(1-x)\partial_x^2 u = f \text{ with } u(0) = g_0,\ u(1) = g_1. \tag{3.1}$$

We can write $u = u_0 + (1-x)g_0 + xg_1$, where u_0 vanishes at the boundary of $[0,1]$. It is apparent that u_0 cannot be $\mathcal{C}^2$ up to the boundary unless $f(0) = f(1) = 0$. In fact,

$$u_0(x) = x\log x f(0) + (1-x)\log(1-x)f(1) + \widetilde{u}_0(x), \tag{3.2}$$

where

$$x(1-x)\partial_x^2\widetilde{u}_0 = f(x) - [(1-x)f(0) + xf(1)] \overset{\text{def}}{=} f^{(1)} \text{ with } \widetilde{u}_0(0) = \widetilde{u}_0(1) = 0; \tag{3.3}$$

the right hand side, $f^{(1)}$, is as smooth as f, and vanishes at 0 and 1. As was shown in [6, 7], this equation has a unique solution, with optimal smoothness measured in the anisotropic Hölder spaces, $\mathcal{C}^{k,2+\gamma}([0,1])$. In particular, if $f \in \mathcal{C}^\infty([0,1])$ then so is $\widetilde{u}_0$.

The eigenfunctions of $x(1-x)\partial_x^2$ that vanish at the boundary are polynomials of the form $x(1-x)p_m(x)$, where p_m is the polynomial of degree $m \geq 0$ that satisfies the equation

$$\left[x(1-x)\partial_x^2 + 2(1-2x)\partial_x - 2\right]p_m = \lambda_m p_m. \tag{3.4}$$

For later reference, the left-hand side may also be written $[\mathrm{L}_{\mathbf{2}}^{\mathrm{Kim}} - 2]p_m$. By inspecting the action on $\mathcal{P}_d/\mathcal{P}_{d-1}$, where $\mathcal{P}_d$ is the space of polynomials of degree at most d, we see that $\lambda_d = -(d+1)(d+2)$. Since the eigenfunctions are orthogonal with respect to $d\mu_{\mathbf{0}}(x)$, these polynomials are orthogonal with respect to the inner product

$$\langle p, q\rangle = \int_0^1 p(x)q(x)\, x^\alpha(1-x)^\beta dx, \tag{3.5}$$

with $\alpha = \beta = 1$. Thus, they are multiples of the corresponding Jacobi polynomials [43, 14], $p_d(x) \propto P_d^{(1,1)}(2x-1)$. If we choose the normalization $\|p_d\|_{L^2([0,1];d\mu_{\mathbf{2}})} = 1$, the result would be the same as performing Gram-Schmidt on the monomials $\{1, x, x^2, \dots\}$. This yields the 3-term recurrence

$$\begin{aligned} &p_0(x) = \sqrt{1/\gamma_0}, \qquad \sqrt{b_1}\,p_1(x) = (x-a_0)p_0(x) \\ &\sqrt{b_{m+1}}\,p_{m+1}(x) = (x-a_m)p_m(x) - \sqrt{b_m}\,p_{m-1}(x), \qquad (m \geq 1), \end{aligned} \tag{3.6}$$

where $\gamma_0 = 1/6$, $a_m = 1/2$ and $b_m = \frac{m(m+2)}{4[4(m+1)^2-1]}$ in this case. Solving (3.3) for $\widetilde{u}_0$ now boils down to representing $f^{(1)}$ in the eigenbases:

$$\tilde{f}(x) = \frac{f^{(1)}(x)}{x(1-x)} = \sum_{m=0}^\infty c_m p_m(x) \quad \Rightarrow \quad \widetilde{u}_0(x) = \sum_{m=0}^\infty (c_m/\lambda_m)x(1-x)p_m(x). \tag{3.7}$$

The simplest way to obtain an approximation of $\tilde{f}(x)$ in $\mathcal{P}_{N-1}$ is to evaluate $f^{(1)}(x)$ at the zeros x_j of $p_N(x)$, and to compute the coefficients using Gauss-Lobatto quadrature:

$$c_m = \int_0^1 \tilde{f}(x)p_m(x)x(1-x)\,dx = \int_0^1 f^{(1)}(x)p_m(x)\,dx \approx \sum_{j=1}^N f^{(1)}(x_j)p_m(x_j)\omega_j. \tag{3.8}$$

Here we used the fact that $f^{(1)}(x)$ vanishes at $x_0 = 0$ and $x_{N+1} = 1$. The abscissas x_j and weights ω_j are easily found using a variant of the Golub-Welsch algorithm [14, 15]. Further details of our numerical implementation will be given elsewhere.

We can estimate the size of the coefficients $\{c_m\}$ in terms of the smoothness of $f^{(1)}$, using the facts that

$$\begin{gathered} c_m = \int_0^1 f^{(1)}(x)p_m(x)\,dx \qquad \text{and} \\ \int_0^1 \mathrm{L}^{\mathrm{Kim}}_{\mathbf{2}}\, p_m f^{(1)} dx = \int_0^1 p_m[\mathrm{L}^{\mathrm{Kim}} + 2] f^{(1)} dx. \end{gathered} \tag{3.9}$$

The second formula is a special case of (4.1) below. If $f^{(1)}$ is in $\mathcal{C}^{2l}([0,1])$, then we can iterate the integration by parts formula to obtain:

$$c_m = \frac{(-1)^l}{[m(m+3)]^l} \int_0^1 p_m(x)[\mathrm{L}^{\mathrm{Kim}} + 2]^l f^{(1)} dx. \tag{3.10}$$

As we show below, there is a constant C, independent of m so that

$$\|p_m\|_{L^\infty} \leq C\sqrt{m(m+3)}\|p_m\|_{L^2([0,1];d\mu_{\mathbf{2}})} = C\sqrt{m(m+3)}|, \tag{3.11}$$

and therefore, there is a constant $\widetilde{C}$ so that

$$|c_m| \leq \widetilde{C}\frac{\|[\mathrm{L}^{\mathrm{Kim}} + 2]^l f^{(1)}\|_{L^\infty}}{[m(m+3)]^{l-\frac{1}{2}}}. \tag{3.12}$$

In this instance it is easy to see that

$$\|[\mathrm{L}^{\mathrm{Kim}} + 2]^l f^{(1)}\|_{L^\infty} \leq C_l \|[f\|_{\mathcal{C}^{2l}}. \tag{3.13}$$

This is almost a spectral estimate, but we lose one order of decay, in part because we are estimating in the L^∞-norm, and in part because there is an implicit division of $f^{(1)}$ by $x(1-x)$ in the formula for c_m.

We can also use the L^2-norm to estimate these coefficients via

$$|c_m| \leq \widetilde{C}'\frac{\|w_{\mathbf{2}}^{-1}[\mathrm{L}^{\mathrm{Kim}} + 2]^l f^{(1)}\|_{L^2([0,1];d\mu_{\mathbf{2}})}}{[m(m+3)]^l}. \tag{3.14}$$

While the denominator is now $[m(m+3)]^l$ the numerator implicitly involves the L^2-derivative of $|[\mathrm{L}^{\mathrm{Kim}} + 2]^l f^{(1)}|^2$ at the boundary of $[0,1]$.

Remarkably, very similar approaches work to find the eigenfunctions of $\mathrm{L}^{\mathrm{Kim}}$ and solve the Dirichlet problem in any dimension. This is explained in the following two sections. In Section 4 we give a novel construction for the eigenfunctions of the neutral Kimura diffusion on the n-simplex, which highlights their vanishing properties on subsets of the boundary. In Section 6 we show how to solve the Dirichlet problem on an n-simplex, with arbitrary smooth data.

4. The Polynomial Eigenfunctions of $\mathrm{L}^{\mathrm{Kim}}$. In this section we give a hierarchical method of constructing the eigenfunctions of $\mathrm{L}^{\mathrm{Kim}}$, with considerable control over their vanishing properties on $b\Sigma_n$. As before, we let $\mathcal{P}_d$ denote polynomials of degree at most d; the variables involved will be clear from the context.

Our results are based upon a formula, which follows easily from a similar calculation in the work of Shimakura, see Section 7 of [38]. As noted above, we let $\mathrm{L}_{\boldsymbol{b}}^{\mathrm{Kim}} = \mathrm{L}^{\mathrm{Kim}} + V_{\boldsymbol{b}}$, where $V_{\boldsymbol{b}}$ has linear coefficients and assigns constant weights to the hypersurface components of $b\Sigma_n$. We then let $\mathcal{I} = \{i_1 < i_2 < \cdots < i_{k+1}\} \subset \{1, \ldots, n+1\}$. Shimakura's work implies the following formula

$$\mathrm{L}_{\boldsymbol{b}}^{\mathrm{Kim}}(w_{\mathcal{I}}\psi) = w_{\mathcal{I}}\left[\mathrm{L}_{\boldsymbol{b}'}^{\mathrm{Kim}} - \kappa_{\mathcal{I}}\right]\psi. \tag{4.1}$$

Where

$$w_{\mathcal{I}}(x) = \prod_{j=1}^{k+1} x_{i_j}^{1-b_{i_j}}; \tag{4.2}$$

$$b'_j = \begin{cases} 2 - b_j & \text{if } j \in \{i_1, \ldots, i_{k+1}\}, \\ b_j & \text{if } j \notin \{i_1, \ldots, i_{k+1}\}; \end{cases} \tag{4.3}$$

and

$$\kappa_{\mathcal{I}} = \left(\sum_{j \in \mathcal{I}} (1 - b_j)\right)\left(k + \sum_{j \notin \mathcal{I}} b_j\right). \tag{4.4}$$

If $k = n$ and $\boldsymbol{b} = \boldsymbol{0}$, then $\boldsymbol{b}' = \boldsymbol{2} = (2, \ldots, 2)$. Equation (4.1) has a wide range of applications, and is especially useful in cases where some of the $\{b_i\}$ are zero. This formula gives a very potent method to construct eigenfunctions that have a simple form and vanish on certain parts of the boundary.

Recall that the $\mathcal{C}^0$-graph closure of $\mathrm{L}^{\mathrm{Kim}}$ acting on $\mathcal{C}^3(\Sigma_n)$ is what is called the "regular operator" in [7], which is the "backward" Kolmogorov operator in applications to Population Genetics. The eigenfunctions that we construct are polynomials and hence in the domain of the regular operator. Indeed, we will show that for each natural number d, the eigenfunctions of degree less than or equal to d actually span $\mathcal{P}_d$. Hence this is the complete set of eigenfunctions for the regular operator.

As functions on the n-simplex, these eigenfunctions do not belong to a single L^2-space, but each belongs to an L^2-space of some stratum of the boundary. Their coefficients in the representation of a function in terms of this spanning set are computed as inner products on these lower dimensional strata.

4.1. Hierarchy of Polynomial Eigenfunctions. The simplest polynomial eigenfunctions of $\mathrm{L}^{\mathrm{Kim}}$ are the functions $\{x_1, \ldots, x_{n+1}\}$, which are null vectors. Each of these eigenfunctions vanishes on a codimension 1 boundary face of Σ_n. Of course, a constant function is also a null-vector for $\mathrm{L}^{\mathrm{Kim}}$, but it already belongs to the span of the others since $x_1 + \cdots + x_{n+1} = 1$. To fit with the pattern below, one can write these functions as $x_{i_1}\psi(x)$, where $\psi(x) \equiv 1$ spans the space of constant functions determined by their value at vertex i_1 (where $x_{i_1} = 1$).

Next, for each distinct pair $1 \leq i_1 < i_2 \leq n+1$, we look for eigenfunctions of the form $x_{i_1}x_{i_2}\psi(x)$. If $\psi(x)$ is only a function of x_{i_1} (or x_{i_2}) and satisfies $[\mathrm{L}_{\boldsymbol{2}}^{\mathrm{Kim}} - 2]\psi = \lambda\psi$, i.e.

$$x_{i_1}(1 - x_{i_1})\partial_{x_{i_1}}^2\psi + 2(1 - 2x_{i_1})\partial_{x_{i_1}}\psi - 2\psi = \lambda\psi, \tag{4.5}$$

then, by (4.1), $x_{i_1}x_{i_2}\psi(x_{i_1})$ is also an eigenfunction of the original operator $\mathrm{L}^{\mathrm{Kim}}$ with the same eigenvalue. Note that if $\{j_1, \ldots, j_{n-1}\} = \{1, \ldots, n+1\} \setminus \{i_1, i_2\}$, then we are solving on the edge

$$x_{j_1} = \cdots = x_{j_{n-1}} = 0. \tag{4.6}$$

Hence the eigenfunctions $x_{i_1} x_{i_2} \psi(x_{i_1})$ vanish on the boundary of this edge. Also note that if a function ψ depends only on a subset of the coordinates, then $\mathrm{L}_{\boldsymbol{b}}^{\mathrm{Kim}} \psi$ also depends only on the same subset of the coordinates. In particular, (3.4) and (4.5) agree.

With these observations, working in the different projective models, we can construct all the polynomial eigenfunctions of $\mathrm{L}^{\mathrm{Kim}}$. These take the form

$$x_{i_1} \cdots x_{i_{k+1}} \psi(x_{i_1}, \ldots, x_{i_k}),$$

for various choices of indices $\{i_1, \ldots, i_{k+1}\}$. Here ψ is an eigenfunction of the operator

$$\mathrm{L}_{\mathcal{I},\boldsymbol{b}'}^{\mathrm{Kim}} = \mathrm{L}_{\boldsymbol{b}'}^{\mathrm{Kim}} \restriction_{K_{\mathcal{I}}} \tag{4.7}$$

acting on a k-simplex, and the variables $(x_{i_1}, \ldots, x_{i_{k+1}})$ range over the face defined by the equations

$$x_{j_1} = \cdots = x_{j_{n-k}} = 0, \tag{4.8}$$

using the notation of (2.10). The weights for this operator are all equal to 2, and therefore we denote it by $\mathrm{L}_{\mathcal{I},\boldsymbol{2}}^{\mathrm{Kim}}$. A simple calculation shows that, if $|\mathcal{I}| = k+1$, then

$$\mathrm{L}_{\mathcal{I},\boldsymbol{2}}^{\mathrm{Kim}} \left(x_{i_1}^{m_1} \cdots x_{i_k}^{m_k} \right) = -\Big(|\vec{m}|^2 + (2k+1)|\vec{m}| \Big) \left(x_{i_1}^{m_1} \cdots x_{i_k}^{m_k} \right) \mod \mathcal{P}_{|\vec{m}|-1}, \tag{4.9}$$

where $|\vec{m}| = m_1 + \cdots + m_k$. (A more general formula is given in (5.9) below.) Thus, $\mathrm{L}_{\mathcal{I},\boldsymbol{2}}^{\mathrm{Kim}}$ leaves the subspaces $\mathcal{P}_d$ invariant; the dth eigenvalue is $[-d^2 - (2k+1)d]$ for $d \geq 0$; and its multiplicity is equal to $\binom{d+k-1}{k-1}$, the dimension of $\mathcal{P}_d/\mathcal{P}_{d-1}$. Moreover, applying the Gram-Schmidt procedure to the set of monomials in the variables $x_{i_1}, \ldots, x_{i_k}$, ordered by degree (but arbitrarily ordered within the set of monomials of the same degree), will lead to an orthonormal set of eigenfunctions of $\mathrm{L}_{\mathcal{I},\boldsymbol{2}}^{\mathrm{Kim}}$. As a result, the eigenfunctions $\psi_{\mathcal{I},\vec{m}}(x_{i_1}, \ldots, x_{i_k})$ of this operator are multivariate orthogonal polynomials with respect to the measure

$$d\mu_{\mathcal{I},\boldsymbol{2}} = \left(\prod_{j=1}^{k+1} x_{i_j} \right) dx_{i_1} \cdots dx_{i_k}. \tag{4.10}$$

The corresponding eigenfunctions $w_{\mathcal{I}} \psi_{\mathcal{I},\vec{m}}$ of $\mathrm{L}_{\mathcal{I}}^{\mathrm{Kim}}$ are orthogonal with respect to $d\mu_{\boldsymbol{0}}$ when $k = n$ and $\mathcal{I} = \{1, \ldots, n+1\}$, but are not normalizable (i.e. do not belong to $L^2(\Sigma_n; d\mu_{\boldsymbol{0}})$) when $k < n$. Nevertheless, they are still eigenfunctions algebraically, with eigenvalue

$$\lambda_{\mathcal{I},\vec{m}} = -|\vec{m}|^2 - (2k+1)|\vec{m}| - k(k+1), \qquad (k = |\mathcal{I}| - 1) \tag{4.11}$$

and play an essential role in solving $\mathrm{L}^{\mathrm{Kim}} u = f$ below, where we use them to adjust f to zero on the boundaries, just as was done in (3.2) and (3.3) in 1-d.

From this observation, it is not difficult to demonstrate the completeness, as a Schauder basis for $\mathcal{C}^0(\Sigma_n)$, of the eigenfunctions obtained in this way. Let $\mathcal{I} = \{i_1, \ldots, i_{k+1}\} \subset \{1, \ldots, n+1\}$ be a set of indices, and $K_{\mathcal{I}} \subset b\Sigma_n$, the corresponding boundary face. We let $\dot{\mathcal{P}}_{\mathcal{I}} \subset \mathbb{C}[x_{i_1}, \ldots, x_{i_k}]$ denote the ideal of polynomials defined on $K_{\mathcal{I}}$ that vanish on $bK_{\mathcal{I}}$. It is easy to see that

$$\dot{\mathcal{P}}_{\mathcal{I}} = x_{i_1} \cdots x_{i_{k+1}} \cdot \mathbb{C}[x_{i_1}, \ldots, x_{i_k}]. \tag{4.12}$$

Formula (4.1) shows that this ideal is invariant under $\mathrm{L}^{\mathrm{Kim}}_{\mathcal{I}}$. The operator $\mathrm{L}^{\mathrm{Kim}}_{\mathcal{I},\mathbf{2}}$ is self adjoint with respect to the measure $x_{i_1}\cdots x_{i_{k+1}}dx_{i_1}\cdots dx_{i_k}$, and it too preserves $\mathcal{P}_d$, for any d. Thus the polynomial eigenfunctions of $\mathrm{L}^{\mathrm{Kim}}_{\mathcal{I},\mathbf{2}}$ span $\mathbb{C}[x_{i_1},\dots,x_{i_k}]$, and therefore $x_{i_1}\cdots x_{i_{k+1}}$ times these eigenfunctions spans $\dot{\mathcal{P}}_{\mathcal{I}}$.

We next observe that an eigenfunction of this type must vanish on every other k-simplex in the boundary of Σ_n. This is because one of the functions $\{x_{i_j}\,:\,j=1,\dots,k+1\}$ must appear as a defining function for any other k-simplex. Using this observation, we see that any polynomial f that vanishes on every k-simplex except $K_{\mathcal{I}}$ can be expanded in these eigenfunctions. This suggests a simple recursive method for finding the regular solution of an equation of the form

$$\mathrm{L}^{\mathrm{Kim}}\,u=f. \tag{4.13}$$

As discussed below, the regular solution is required to belong to an anisotropic Hölder space in the closed simplex rather than to take on specified boundary values. Even if f is a polynomial and $g=0$, the solution to the Dirichlet problem (1.4) is generally not differentiable up to the boundary of Σ_n.

4.2. The Regular Solution of $\mathrm{L}^{\mathrm{Kim}}\,u=f$. As suggested by our construction of the eigenfunctions, the regular solution to $\mathrm{L}^{\mathrm{Kim}}\,u=f$ is found recursively by working one boundary stratum at a time. The k-skeleton of Σ_n is defined to be the union of k-simplices in $b\Sigma_n$. We denote this subset by Σ_n^k. It is connected for $k>0$, but $\Sigma_n^k\setminus\Sigma_n^{k-1}$ is a disjoint union of *open* k-simplices with boundaries contained in Σ_n^{k-1}. The regular solution takes the form

$$u=u_1+\cdots+u_n, \tag{4.14}$$

where the term u_k is found by using the eigenfunctions constructed above to solve an auxiliary inhomogeneous Dirichlet problem on $\Sigma_n^k\setminus\Sigma_n^{k-1}$.

If $u\in\mathcal{C}^2$, then $\mathrm{L}^{\mathrm{Kim}}\,u$ vanishes at each vertex. Hence, we start by assuming that f is a polynomial that vanishes on each of the vertices of Σ_n. This assumption is dropped in Section 6, where imposing inhomogeneous Dirichlet boundary conditions on $b\Sigma_n$ inevitably introduces singularities anyway. We define u_1 as the solution of the equation $\mathrm{L}^{\mathrm{Kim}}\,u_1\restriction_{\Sigma_n^1}=f\restriction_{\Sigma_n^1}$. As f vanishes on the vertices, which are the boundaries of the components of $\Sigma_n^1\setminus\Sigma_n^0$, we can use the eigenfunctions $x_{i_1}x_{i_2}\psi(x_{i_1})$ constructed above to solve this equation on each component of $\Sigma_n^1\setminus\Sigma_n^0$ independently of the others. Note that, using the eigenfunction representation, the function u_1 extends canonically to the entire n-simplex.

We next solve

$$\mathrm{L}^{\mathrm{Kim}}\,u_2\restriction_{\Sigma_n^2}=f\restriction_{\Sigma_n^2}-\mathrm{L}^{\mathrm{Kim}}\,u_1\restriction_{\Sigma_n^2}. \tag{4.15}$$

The right hand side vanishes on Σ_n^1, so we can use the eigenfunctions $w_{\mathcal{I}}\psi_{\mathcal{I},\vec{m}}$ with $\mathcal{I}=\{i_1,i_2,i_3\}$ to independently solve this equation on each connected component of $\Sigma_n^2\setminus\Sigma_n^1$. Recursively, we assume that we have found $u_1,\dots,u_{k-1}$, so that

$$f-\mathrm{L}^{\mathrm{Kim}}(u_1+\cdots+u_{k-1}) \tag{4.16}$$

vanishes on the $(k-1)$-skeleton, and then solve

$$\mathrm{L}^{\mathrm{Kim}}\,u_k\restriction_{\Sigma_n^k}=f\restriction_{\Sigma_n^k}-\mathrm{L}^{\mathrm{Kim}}(u_1+\cdots+u_{k-1})\restriction_{\Sigma_n^k}. \tag{4.17}$$

Using the eigenfunctions of the form $x_{i_1}\cdots x_{i_{k+1}}\psi(x_{i_1},\dots,x_{i_k})$, we can solve the relevant Dirichlet problems independently on each component of $\Sigma_n^k\setminus\Sigma_n^{k-1}$.

The process terminates when we reach the interior of the n-simplex, where we solve the problem

$$\mathrm{L}^{\mathrm{Kim}}\, u_n\restriction_{\Sigma_n} = f\restriction_{\Sigma_n} - \mathrm{L}^{\mathrm{Kim}}(u_1 + \cdots + u_{n-1})\restriction_{\Sigma_n} . \tag{4.18}$$

Once again the right hand side vanishes on the entire boundary of Σ_n and we can solve this equation using eigenfunctions of the form $x_1 \cdots x_{n+1}\psi(x)$. Since this can be done for any polynomial that vanishes on the vertices, it demonstrates that the eigenfunctions constructed above, including the nullspace, are in fact a complete set of eigenfunctions for the operator $\mathrm{L}^{\mathrm{Kim}}$, acting on polynomial functions defined on Σ_n. Since these functions are dense in $\mathcal{C}^2(\Sigma_n)$ it follows easily that this is also a complete set of eigenfunctions for the graph closure of $\mathrm{L}^{\mathrm{Kim}}$ acting on $\mathcal{C}^0(\Sigma_n)$. Altogether we have proved the following result:

THEOREM 4.1. *The regular operator $\mathrm{L}^{\mathrm{Kim}}$ acting on functions defined on Σ_n has a complete set of eigenfunctions of the form*

$$\mathcal{E}(\mathrm{L}^{\mathrm{Kim}}) = \{x_1, \dots, x_{n+1}\} \cup \bigcup_{k=1}^{n} \bigcup_{\mathcal{I}=\{1\leq i_1<\cdots<i_{k+1}\leq n+1\}} x_{i_1} \cdots x_{i_{k+1}} \mathcal{E}(\mathrm{L}^{\mathrm{Kim}}_{\mathcal{I},\mathbf{2}}). \tag{4.19}$$

Here $\mathcal{E}(\mathrm{L}^{\mathrm{Kim}}_{\mathcal{I},\mathbf{2}})$ denotes the eigenfunctions of the operator $\mathrm{L}^{\mathrm{Kim}}_{\mathcal{I},\mathbf{2}}$, which are polynomials in the variables $\{x_{i_1}, \dots, x_{i_k}\}$.

REMARK 4.2. *Another useful consequence of the hierarchical description of the eigenfunctions of $\mathrm{L}^{\mathrm{Kim}}$ is that it dramatically reduces the work needed to find the eigenfunctions of $\mathrm{L}^{\mathrm{Kim}}$ acting on Σ_{n+1}, once they are known for Σ_n. Indeed, up to choosing subsets of coordinates $(x_1, \dots, x_{n+1})$, all that is required is to find the “new” eigenfunctions, which are of the form*

$$x_1 \cdots x_{n+2}\psi(x_1, \dots, x_{n+1}).$$

REMARK 4.3. *At stage k of the recursive algorithm above, we are given a function f that vanishes on the $(k-1)$-skeleton of Σ_n and wish to find u_k such that $\mathrm{L}^{\mathrm{Kim}}\, u_k\restriction_{\Sigma_n^k} = f\restriction_{\Sigma_n^k}$. For simplicity of notation, we have absorbed $\mathrm{L}^{\mathrm{Kim}}(u_1 + \cdots u_{k-1})$ into f in (4.16). Since the faces of the k-skeleton decouple, we may assume $f = 0$ on all faces $K_{\mathcal{I}}$ except one of them, which we take to be $\mathcal{I} = \{1, \dots, k+1\}$. Recall from the discussion after (4.9) that the relevant eigenfunctions of $\mathrm{L}^{\mathrm{Kim}}_{\mathcal{I}}$ are of the form $(x_1 \cdots x_{k+1})\psi_{\mathcal{I},\vec{m}}(x)$, with $\{\psi_{\mathcal{I},\vec{m}}(x)\}_{\vec{m}\in\mathbb{N}_0^k}$ a set of multivariate orthogonal polynomials with respect to $d\mu_{\mathcal{I},\mathbf{2}}$ in (4.10) on $\widetilde{\Sigma}_k$. Thus, the generalization of (3.7) to k dimensions is*

$$\tilde{f}(x) = \frac{f(x)}{x_1 \cdots x_{k+1}} = \sum_{\vec{m}\in\mathbb{N}_0^k} c_{\vec{m}}\psi_{\mathcal{I},\vec{m}}(x) \quad \Rightarrow \quad u_k(x) = \sum_{\vec{m}\in\mathbb{N}_0^k} \frac{c_{\vec{m}}}{\lambda_{\vec{m}}} x_1 \cdots x_{k+1}\, \psi_{\mathcal{I},\vec{m}}(x). \tag{4.20}$$

In [7] it is shown that for k a non-negative integer, and $\gamma \in (0,1)$, the operator $\mathrm{L}^{\mathrm{Kim}}$ maps the anisotropic Hölder space $\mathcal{C}^{k,2+\gamma}_{\mathrm{WF}}(\Sigma_n)$ to $\mathcal{C}^{k,\gamma}_{\mathrm{WF}}(\Sigma_n)$. If we let $\dot{\mathcal{C}}^{*,*}_{\mathrm{WF}}$ denote the closed subspaces of functions vanishing at the vertices of Σ_n, then the inverse

$$(\mathrm{L}^{\mathrm{Kim}})^{-1} : \dot{\mathcal{C}}^{k,\gamma}_{\mathrm{WF}}(\Sigma_n) \longrightarrow \dot{\mathcal{C}}^{k,2+\gamma}_{\mathrm{WF}}(\Sigma_n) \tag{4.21}$$

is shown to be a bounded operator. Let $C_{k,\gamma}$ denote the bound on this operator. If we approximate $f \in \dot{\mathcal{C}}^{k,\gamma}_{\mathrm{WF}}(\Sigma_n)$ by a polynomial, p, then

$$\|(\mathrm{L}^{\mathrm{Kim}})^{-1}f - (\mathrm{L}^{\mathrm{Kim}})^{-1}p\|_{\mathrm{WF},k,2+\gamma} \leq C_{0,\gamma}\|f - p\|_{\mathrm{WF},k,\gamma}. \tag{4.22}$$

The classical (k,γ)-Hölder norms dominate the WF-Hölder norms, and $\mathcal{C}^{k,\gamma}_{\mathrm{WF}} \subset \mathcal{C}^k$. It is well known that the accuracy of the best degree d polynomial approximant only depends on the classical Hölder smoothness of the data, thus demonstrating the efficacy of our method for solving the equation $\mathrm{L}^{\mathrm{Kim}}\, u = f$ for any reasonably smooth data f that vanishes at the vertices of Σ_n.

One can generalize the construction of polynomial eigenfunctions to any Kimura operator, $\mathrm{L}^{\mathrm{Kim}}_{\boldsymbol{b}}$, with constant weights. If all of the weights are positive, the measure $d\mu_{\boldsymbol{b}}$ has finite mass and the polynomial eigenfunctions of the regular operator can be chosen to be orthonormal with respect to this measure. This case is, in many ways, the easiest to deal with as there is a single ambient Hilbert space containing all the eigenfunctions. It has been studied previously by solving hypergeometric equations [29, 25], by computing bi-orthogonal systems of polynomials [25, 35], or by computing orthogonal polynomials on weighted Hilbert spaces and grouping them into reproducing kernels [16, 17, 18]. The present work extends this third approach to the case of 0 weights, and, for reasons explained below, does not take the final step of forming the reproducing kernels. In the construction of Section 4.1 above, we look for eigenfunctions of the form $x_{i_1}\cdots x_{i_{k+1}}\psi(x_{i_1},\dots,x_{i_k})$ with ψ an eigenfunction of the auxiliary operator $\mathrm{L}^{\mathrm{Kim}}_{\boldsymbol{b}'}$. If all the weights are positive, the leading factor of $x_{i_1}\cdots x_{i_{k+1}}$ is dropped, k is set to n, and the eigenfunctions of $\mathrm{L}^{\mathrm{Kim}}_{\boldsymbol{b}}$ are computed directly, without recourse to an auxiliary operator. While eigenfunctions with respect to a subset of the variables are again eigenfunctions of $\mathrm{L}^{\mathrm{Kim}}_{\boldsymbol{b}}$, they are already present in the construction at level n, and do not have to be dealt with recursively along the stratification of $\partial\Sigma_n$. Computing these eigenfunctions works the same in both cases, and is mathematically equivalent to applying the Gram-Schmidt method to the monomials ordered by total degree. This is explained in the following section.

When some weights are zero and some positive, Shimakura's formula leads to a partial hierarchical structure for the eigenfunctions, similar to that described here. This is discussed, to some extent, in [38, 39]. As shown in [38] it is possible to specify homogeneous Dirichlet data on any boundary face where the weight $b_i < 1$. In general the solution is not smooth along this face, but has a leading singularity of the form $x_i^{1-b_i}$. We will return to these questions in a subsequent publication.

5. Numerical Construction of Eigenfunctions. We now give a practical method to construct the eigenfunctions described above. By a change of variables these eigenfunctions can be represented as products of functions of single variables, which are themselves eigenfunctions of differential operators. Since it is no more difficult, we consider the general case of a Kimura diffusion on Σ_n with constant weights.

In 2-d, it was observed by Proriol [36] that orthogonal polynomials on a triangle can be represented as tensor products of 1-d Jacobi polynomials. See Koornwinder [31] for further background and more complicated 2-d geometries. A rather different approach using hypergeometric functions was developed in 2-d by Kimura [29] and Karlin and McGregor [25]. This approach was extended to the simplex or n-sphere in the context of the Laplace-Beltrami equation by Kalnins, Miller and Tratnik [23]. Littler and Fackerell [35] found solutions of Kimura diffusion using expansions in bi-orthogonal polynomials. Griffiths [16] found representations for the transition function for this diffusion process using reproducing kernels expressed in terms of multivariate Jacobi polynomials; see also [17, 18]. The method we derive below is equivalent to Griffiths' approach, though we avoid computing reproducing kernels as it is more efficient and numerically stable (at the expense of symmetry) to leave the eigenfunctions separated. Wingate and Taylor [44] showed how to generalize the Proriol construction to the N-dimensional simplex in the case of a uniform weight. We adapt their method to allow more general Jacobi weights of the form

$$\langle p,q\rangle = \int_{\widetilde{\Sigma}_k} p(x)q(x)\, x_1^{\alpha_1}\cdots x_{k+1}^{\alpha_{k+1}}\, dx_1\cdots dx_k. \tag{5.1}$$

The case of interest here is $\alpha_j = b'_j - 1 = 1$, but treating the general case is no more complicated, and can be used to study Kimura diffusion with non-zero weights $\boldsymbol{b}$, as shown below.

The idea is to map the unit cube to the simplex in such a way that the desired orthogonal polynomials on the simplex pull back to have tensor product form on the cube. One way to do this, which differs somewhat from the choice made by Wingate and Taylor, is with the change of variables $x = T(X)$ given implicitly by

$$x_1 = X_1, \quad x_2 = (1-x_1)X_2, \quad \cdots \quad x_k = (1-x_1-\cdots-x_{k-1})X_k, \tag{5.2}$$

which solves to $x_j = \big(\prod_{i=1}^{j-1}(1-X_i)\big)X_j$ and $x_{k+1} = 1 - x_1 - \cdots - x_k = \prod_{i=1}^{k}(1-X_i)$. These are blow-ups similar to the changes of variables that are used in [20, 21, 22].

The Jacobian matrix $DT(X)$ is lower triangular, so its determinant is easy to compute:

$$J = |\det DT| = \textstyle\prod_{j=1}^{k}(1-X_j)^{k-j}. \tag{5.3}$$

The inner product (5.1) may now be written

$$\langle p,q\rangle = \int_{[0,1]^k} (p\circ T)(q\circ T)\Big(\prod_{j=1}^{k}\big[X_j^{\alpha_j}(1-X_j)^{(k-j)+\sum_{i=j+1}^{k+1}\alpha_i}\big]\Big)\, dX_1\cdots dX_k. \tag{5.4}$$

Next we note that if p is a polynomial (in one variable) of degree d, then

$$p(X_j)\prod_{i=1}^{j-1}(1-X_i)^d = p\left(\frac{x_j}{1-x_1-\cdots-x_{j-1}}\right)(1-x_1-\cdots-x_{j-1})^d \tag{5.5}$$

is a polynomial of degree d in the variables $x_1,\cdots,x_j$. We can therefore define $\psi_{\vec m}(x)$ via

$$\begin{aligned}\psi_{\vec m}\circ T(X) &= \prod_{j=1}^{k}\left(Q_{m_j}^{(\alpha_j,\alpha_{j\vec m})}(X_j)\prod_{i=1}^{j-1}(1-X_i)^{m_j}\right)\\ &= \prod_{j=1}^{k}\left(Q_{m_j}^{(\alpha_j,\alpha_{j\vec m})}(X_j)(1-X_j)^{\sum_{i=j+1}^{k}m_i}\right),\end{aligned} \tag{5.6}$$

where

$$\alpha_{j\vec m} = \alpha_{k+1} + \sum_{i=j+1}^{k}(\alpha_i + 2m_i + 1). \tag{5.7}$$

Here $\{Q_m^{(\alpha,\beta)}\}_{m=0}^{\infty}$ are orthogonal polynomials on $[0,1]$ in the inner product (3.5), normalized to have unit length. Substitution into (5.4) gives

$$\langle\psi_{\vec m},\psi_{\vec m'}\rangle =$$
$$\prod_{j=1}^{k}\int_0^1 Q_{m_j}^{(\alpha_j,\alpha_{j\vec m})}(X_j)Q_{m'_j}^{(\alpha_j,\alpha_{j\vec m'})}(X_j)X_j^{\alpha_j}(1-X_j)^{\alpha_{k+1}+\sum_{i=j+1}^{k}(\alpha_i+m_i+m'_i+1)}\, dX_j,$$

which, proceeding from $j = k$ to $j = 1$, may be seen to equal $\prod_{j=1}^{k} \delta_{m_j m'_j}$. We note that $\psi_{\vec{m}}(x)$ is a polynomial of degree $d = |\vec{m}| = \sum_1^k m_j$. We order the multi-indices first by degree ($\vec{m} < \vec{m}'$ if $d < d'$) and then lexicographically from right to left (if $d = d'$ then $\vec{m} < \vec{m}'$ if $m_k < m'_k$, or if $m_k = m'_k$ and $m_{k-1} < m'_{k-1}$, etc., so that $(d, 0, \dots, 0)$ is smallest). In this ordering, the $\psi_{\vec{m}}(x)$ are the same as one would get from applying Gram-Schmidt to the monomials $x^{\vec{m}}$ in the same order. Hence, all polynomials in k variables are accounted for by this construction. For example, when $k = 3$ and $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1$, the first 10 orthogonal polynomials are

$$\psi_{000} = 12\sqrt{35}, \quad \psi_{100} = 12\sqrt{105}(4x-1), \quad \psi_{010} = 12\sqrt{210}(x+3y-1),$$
$$\psi_{001} = 36\sqrt{70}(-1+x+y+2z), \quad \psi_{200} = 24\sqrt{55}(1-9x+15x^2),$$
$$\psi_{110} = 6\sqrt{2310}(-1+5x)(-1+x+3y), \quad \psi_{020} = 6\sqrt{330}(3+3x^2-21y+28y^2+3x(-2+7y)),$$
$$\psi_{101} = 18\sqrt{770}(-1+5x)(-1+x+y+2z), \quad \psi_{011} = 30\sqrt{462}(-1+x+4y)(-1+x+y+2z),$$
$$\psi_{002} = 24\sqrt{1155}(1+x^2+y^2-5z+5z^2+y(-2+5z)+x(-2+2y+5z)),$$

which span the same space as $1, x, y, z, x^2, xy, y^2, xz, yz, z^2$ and are pairwise orthogonal. In numerical computations, it is best to leave them in the form (5.6) since representation in the monomial basis is both expensive and numerically unstable when the degree is large.

The $Q_m^{(\alpha,\beta)}$ are proportional to Jacobi polynomials, $Q_m^{(\alpha,\beta)}(x) \propto P_m^{(\beta,\alpha)}(2x-1)$, but are more easily computed directly from a 3-term recurrence. Explicitly, $p_m(x) = Q_m^{(\alpha,\beta)}(x)$ satisfies the recurrence (3.6) with $\gamma_0 = \frac{\Gamma(\alpha+1)\Gamma(\beta+1)}{\Gamma(\alpha+\beta+2)}$ and

$$a_m = \frac{1}{2}\left(1 + \frac{(\alpha+\beta)(\alpha-\beta)}{(\alpha+\beta+2m+2)(\alpha+\beta+2m)}\right) \overset{m=0}{=} \frac{\alpha+1}{\alpha+\beta+2}, \qquad (m \geq 0),$$
$$b_m = \frac{m(\alpha+m)(\beta+m)(\alpha+\beta+m)}{(\alpha+\beta+2m)^2[(\alpha+\beta+2m)^2-1]} \overset{m=1}{=} \frac{(\alpha+1)(\beta+1)}{(\alpha+\beta+2)^2(\alpha+\beta+3)}, \qquad (m \geq 1).$$

These formulas can be derived from the well-known recurrence relations of the Jacobi polynomials [14, 1], modified so that $\|p_m\| = 1$. Finally, now that we have specified the orthonormal basis $\psi_{\vec{m}}(x)$, the coefficients

$$c_{\vec{m}} = \int_{\Sigma_k} \tilde{f}(x)\psi_{\vec{m}}(x)x_1^{\alpha_1}\cdots x_{k+1}^{\alpha_{k+1}}\,dx_1\cdots dx_k = \int_{\Sigma_k} f(x)\psi_{\vec{m}}(x)\,dx_1\cdots dx_k \tag{5.8}$$

in (4.20) can be computed by Gauss-Jacobi quadrature on $[0,1]^k$ via the same change of variables (5.2). Further details will be presented elsewhere.

To see that the orthogonal polynomials $\psi_{\vec{m}}$ constructed above are eigenfunctions of $L_{\mathcal{I},\boldsymbol{b}}^{\mathrm{Kim}}$, we note that the generalization of (4.9) is

$$L_{\mathcal{I},\boldsymbol{b}}^{\mathrm{Kim}}\left(x_1^{m_1}\cdots x_{k+1}^{m_{k+1}}\right) = -\Big(|\vec{m}|^2 + (B-1)|\vec{m}|\Big)\Big(x_1^{m_1}\cdots x_{k+1}^{m_{k+1}}\Big) \mod \mathcal{P}_{|\vec{m}|-1}, \tag{5.9}$$

where $B = b_1 + \cdots + b_{k+1}$ and $b_j = \alpha_j + 1$. Normally $m_{k+1} = 0$ as the simplex is parametrized by any k of the variables x_i. Equation (5.9) shows that $L_{\mathcal{I},\boldsymbol{b}}^{\mathrm{Kim}}$ is upper triangular in any monomial basis ordered by degree (with arbitrary ordering among monomials of the same degree), and the eigenvalues can be read off the diagonal. If the Gram-Schmidt procedure is applied to these monomials using the $d\mu_{\boldsymbol{b}}$ inner product, the resulting system of orthogonal polynomials will span a nested sequence of invariant subspaces for $L_{\mathcal{I},\boldsymbol{b}}^{\mathrm{Kim}}$, and hence are eigenfunctions, by self-adjointness. The above construction is equivalent to the Gram-Schmidt procedure.

Following Griffiths [16], one could combine all the eigenfunctions in each eigenspace into reproducing kernels, defined as $G_d(x,y) = \sum_{|\vec{m}|=d} \psi_{\vec{m}}(x)\psi_{\vec{m}}(y)$. In the above example, replacing (x,y,z) with (x_1,x_2,x_3), one would have e.g. $G_1(x,y) = 136080 + 181440[-x_1-x_2-x_3-y_1-y_2-y_3+x_1y_1+x_2y_2+x_3y_3+(x_1+x_2+x_3)(y_1+y_2+y_3)]$. This is theoretically appealing due to symmetry and independence of the particular ordering chosen to compute the orthogonal polynomials. However, it is more efficient computationally to work with the orthogonal polynomial representation directly. It is also numerically unstable to expand out these polynomials in terms of their coefficients to achieve symmetry in the formulas for $G_d(x,y)$ when d is large.

5.1. Coefficient Estimates. If f is a polynomial, or more generally, a smooth enough function on Σ_n, then we can use the recursive approach above to represent f in terms of the eigenfunctions of $\mathrm{L}^{\mathrm{Kim}}$, $\{w_{\mathcal{I}}\psi_{\mathcal{I},\vec{m}}\}$. Here and in the sequel we assume that all the eigenfunctions $\{\psi_{\mathcal{I},\vec{m}}\}$ are normalized by the condition:

$$\int\limits_{K_{\mathcal{I}}} |\psi_{\mathcal{I},\vec{m}}(y)|^2 d\mu_{\mathcal{I},\mathbf{2}}(y) = 1. \tag{5.10}$$

When $\mathcal{I} = \{j\}$, then $K_{\mathcal{I}}$ is a vertex; the functions on $K_{\mathcal{I}}$ are constants. We write $\psi_{\mathcal{I}}$ instead of $\psi_{\mathcal{I},\vec{m}}$, where for all j, $\psi_{\{j\}}(x) = 1$, the constant function. Note that $x_j\psi_{\{j\}}(x)$ is then a function in the null-space of $\mathrm{L}^{\mathrm{Kim}}$ that is one at e_j and vanishes at all other vertices. We let $\mu_{\mathcal{I},\mathbf{2}} = \delta_{e_j}(x)$ be the atomic measure at the vertex e_j of unit mass.

We begin by subtracting off the values of f at the vertices:

$$f^{(1)}(x) = f(x) - \sum_{j=1}^{n+1} f(e_j)x_j\psi_{\{j\}}(x). \tag{5.11}$$

The function $f^{(1)}$ now vanishes at each of the vertices of Σ_n and can therefore be expanded on the 1-skeleton in terms of the eigenfunctions of the form $\{w_{\mathcal{I}}\psi_{\mathcal{I},\vec{m}}\}$, where $\mathcal{I}$ ranges over subsets of $\{1,\ldots,n+1\}$ of cardinality 2. The coefficients are computed as integrals over the 1-dimensional strata:

$$c_{\mathcal{I},\vec{m}} = \int_0^1 \frac{f^{(1)}(x)}{x(1-x)}\psi_{\mathcal{I},\vec{m}}(x)x(1-x)dx = \int_0^1 f^{(1)}(x)\psi_{\mathcal{I},\vec{m}}(x)dx, \tag{5.12}$$

where x is a coordinate on the stratum $K_{\mathcal{I}}$. As in (3.12) and (3.14), the rate of decay of these coefficients is determined by the smoothness of $f^{(1)}$.

Subtracting off the contributions from the 1-dimensional strata we get $f^{(2)}$, which vanishes on the 1-skeleton of Σ_n. we can represent this function on the 2-skeleton in terms of eigenfunctions of the form $\{w_{\mathcal{I}}\psi_{\mathcal{I},\vec{m}}\}$, where the cardinality of $\mathcal{I}$ is 3. Proceeding in this way we get a sequence of functions $f, f^{(1)}, f^{(2)}, \ldots, f^{(k)}, \ldots, f^{(n)}$, where $f^{(k)}$ vanishes on the $k-1$-skeleton of $b\Sigma_n$.

The coefficients coming from $f^{(k)}$ are computed as integrals over the components of the k-dimensional part of the boundary of Σ_n :

$$c_{\mathcal{I},\vec{m}} = \int\limits_{K_{\mathcal{I}}} \psi_{\mathcal{I},\vec{m}}(y)f^{(k)}(y)dy. \tag{5.13}$$

Here $\mathcal{I}$ has cardinality $k+1$ and y is a linear coordinate on $K_{\mathcal{I}}$. Using Shimakura's formula, (4.1) we can show that

$$\int\limits_{K_{\mathcal{I}}} L_{\mathbf{2}}\psi_{\mathcal{I},\vec{m}}(y)f^{(k)}(y)dy = \int\limits_{K_{\mathcal{I}}} \psi_{\mathcal{I},\vec{m}}(y)[\mathrm{L}^{\mathrm{Kim}}_{\mathcal{I},\mathbf{0}} + \kappa_{\mathcal{I}}]f^{(k)}(y)dy. \tag{5.14}$$

From this we can show that if $L_{\mathbf{2}}\psi_{\mathcal{I},\vec{m}} = \lambda_{\mathcal{I},\vec{m}}\psi_{\mathcal{I},\vec{m}}$, then, for a function $f^{(k)} \in \mathcal{C}^{2l}(K_{\mathcal{I}})$ that vanishes at the boundary of $K_{\mathcal{I}}$, we have:

$$c_{\mathcal{I},\vec{m}} = \frac{1}{\lambda^l_{\mathcal{I},\vec{m}}}\int\limits_{K_{\mathcal{I}}} \psi_{\mathcal{I},\vec{m}}(y)[\mathrm{L}^{\mathrm{Kim}}_{\mathcal{I},\mathbf{0}} + \kappa_{\mathcal{I}}]^l f^{(k)}(y)dy. \tag{5.15}$$

From this relation it is clear that the rate of decay of the coefficients $\{c_{\mathcal{I},\vec{m}}\}$ is determined by the smoothness of $f^{(k)}$ on the k-skeleton, and an L^∞-estimate for the eigenfunctions $\{\psi_{\mathcal{I},\vec{m}}\}$.

To estimate $f^{(k)}$ in terms of the original data would take us too far afield, so we conclude this discussion by proving sup-norm estimates on the eigenfunctions. The eigenfunctions $\{\psi_{\mathcal{I},\vec{m}}\}$ are eigenfunctions of the operator $L_{\mathcal{I},\mathbf{2}}$, which is self adjoint with respect to the measure $d\mu_{\mathcal{I},\mathbf{2}}$, and has strictly positive weights. The kernel for the operator $e^{tL_{\mathcal{I},\mathbf{2}}}$ takes the form $p_t(y,\widetilde{y})d\mu_{\mathcal{I},\mathbf{2}}(\widetilde{y})$, with $p_t(y,\widetilde{y}) = p_t(\widetilde{y},y)$. This and the semi-group property easily imply that

$$p_{2t}(y,y) = \int_{K_{\mathcal{I}}}[p_t(y,\widetilde{y})]^2 d\mu_{\mathcal{I},\mathbf{2}}(\widetilde{y}). \tag{5.16}$$

Since the operator has positive weights, we can use the Theorem 5.2 in [8] to conclude that there is a constant C_k depending only on the dimension so that

$$p_{2t}(y,y) \leq \frac{C_k}{\mu_{\mathbf{2}}(B_{\sqrt{2t}}(y))}. \tag{5.17}$$

Here $B_r(y)$ is the ball in the intrinsic metric (see (2.24) and [8]) of radius $\sqrt{2t}$ centered at y. From the forms of the metric and the measure we can easily show that there are constants C_0, C_1, so that for small t, on strata of dimension k, we have the bounds

$$C_0 t^k \leq \mu_{\mathbf{2}}(B_{\sqrt{2t}}(y)) \leq C_1 t^{\frac{k}{2}}. \tag{5.18}$$

To prove an estimate on $\|\psi_{\mathcal{I},\vec{m}}\|_{L^\infty}$, we observe that

$$\psi_{\mathcal{I},\vec{m}}(y)e^{t\lambda_{\mathcal{I},\vec{m}}} = \int\limits_{K_{\mathcal{I}}} p_t(y,\widetilde{y})\psi_{\mathcal{I},\vec{m}}(\widetilde{y})d\mu_{\mathcal{I},\mathbf{2}}(\widetilde{y}). \tag{5.19}$$

The Cauchy-Schwarz inequality and the estimates above show that

$$|\psi_{\mathcal{I},\vec{m}}(y)| \leq \frac{C_k e^{-t\lambda_{\mathcal{I},\vec{m}}}}{t^{\frac{k}{2}}}. \tag{5.20}$$

Setting $t = -1/\lambda_{\mathcal{I},\vec{m}}$, gives the estimate

$$|\psi_{\mathcal{I},\vec{m}}(y)| \leq \widetilde{C}_k|\lambda_{\mathcal{I},\vec{m}}|^{\frac{k}{2}}. \tag{5.21}$$

Inserting this estimate into (5.15) we see that there is a constant C'_k, so that the coefficients coming from the stratum of dimension k satisfy an estimate of the form:

$$|c_{\mathcal{I},\vec{m}}| \leq C'_k\frac{\|[\mathrm{L}^{\mathrm{Kim}} + \kappa_{\mathcal{I}}]^l f^{(k)}\|_{L^\infty}}{\lambda^{l-\frac{k}{2}}_{\mathcal{I},\vec{m}}}. \tag{5.22}$$

One-half order of decay is lost with each increase in dimension. We can also give an L^2-estimate, where no such loss explicitly occurs, wherein

$$|c_{\mathcal{I},\vec{m}}| \leq C_k'' \frac{\|w_{\mathcal{I}}^{-1}[\mathrm{L}^{\mathrm{Kim}} + \kappa_{\mathcal{I}}]^l f^{(k)}\|_{L^2(K_{\mathcal{I}}, d\mu_{\mathcal{I},\mathbf{2}})}}{\lambda_{\mathcal{I},\vec{m}}^l}. \tag{5.23}$$

These estimates implicitly involve k derivatives of $|[\mathrm{L}^{\mathrm{Kim}} + \kappa_{\mathcal{I}}]^l f^{(k)}|^2$ near the boundary of $K_{\mathcal{I}}$. In both estimates there are further losses that occurs in the estimation of $[\mathrm{L}^{\mathrm{Kim}} + \kappa_{\mathcal{I}}]^l f^{(k)}$ in terms of the original data f.

6. The Dirichlet Problem. In many applications of the Kimura equation to problems in population genetics one needs to solve a problem of the form

$$\mathrm{L}^{\mathrm{Kim}} u = f \ \text{ with } \ u\restriction_S = g. \tag{6.1}$$

Here S is a subset of $b\Sigma_n$, generally assumed to be a union of faces. In the constant weight case, Shimakura showed that this problem is well-posed if the weights on the faces contained in S are all less than 1. Note that, if a weight is greater than or equal to 1, then, with probability 1, the paths of the associated stochastic process never reach the corresponding face.

In this section, for simplicity, we continue to consider the case that all weights are zero, $S = b\Sigma_n$, and that f and g have a certain degree of smoothness. With this assumption, we show that the solution to the Dirichlet problem has an asymptotic expansion along boundary with the first two terms determined by local calculations, see (6.20) and (6.57). A classical example from Population Genetics is the solution to the Dirichlet problem with $f = -1$, and $g = 0$, which is given by

$$u(x) = \sum_{j=1}^{n} \sum_{0 \leq i_1 < \cdots < i_j \leq n} \eta(x_{i_1} + \cdots + x_{i_j})(-1)^j, \tag{6.2}$$

where $\eta(\tau) = \tau \log \tau$. For a point $x \in \Sigma_n$, the value $u(x)$ is the expected time for a path starting at x to reach $b\Sigma_n$. This formula is highly suggestive of the form that the general result takes; see (6.17) and (6.42) below.

Boundary data that is not continuous on $b\Sigma_n$ does arise naturally in problems connected with exit probabilities, and cannot be directly treated by the methods in this paper. In these types of problems the boundary data is often piecewise constant assuming only the values 0 and 1, and one has recourse to explicit solutions, see [34]. In this case, one could approximate the discontinuous boundary data by smooth boundary data, and use the methods presented here along with the Feynman-Kac formula in [9] to get upper and lower bounds for the solutions of such boundary value problems. For data without some degree of regularity one should not expect the solution to have a simple, explicit asymptotic expansion along the boundary. When the solution is not continuous up to the boundary, considerable care is required in the interpretation of the partial differential equation along the boundary, and boundary condition itself, see [20, 21, 22].

To start we consider the simpler problem in a positive orthant $S_{n,0} = \mathbb{R}_+^n$, with the model operator

$$L_{\mathbf{0}} = \sum_{i=1}^{n} x_i \partial_{x_i}^2. \tag{6.3}$$

This problem is easier to solve and its solution is nearly adequate to solve the analogous problem in a simplex. We start with a simple calculus lemma.

LEMMA 6.1. *Let ψ be a $\mathcal{C}^2$ function of a single variable, then for indices $1 \leq i_1 < \cdots < i_k \leq n$, we have*

$$L_{\mathbf{0}}\psi(x_{i_1} + \cdots + x_{i_k}) = (x_{i_1} + \cdots + x_{i_k})\psi''(x_{i_1} + \cdots + x_{i_k}). \tag{6.4}$$

The proof is an elementary calculation.

We now show how to solve the problem

$$L_{\mathbf{0}}u = f \text{ in } S_{n,0} \text{ with } u\restriction_{bS_{n,0}} = g, \tag{6.5}$$

for f a compactly supported function in $\mathcal{C}^2(S_{n,0})$. We assume that g has a compactly supported extension, $\widetilde{g}$, to $S_{n,0}$, for which $L_{\mathbf{0}}\widetilde{g}$ also belongs to $\mathcal{C}^2(S_{n,0})$. It is clear that, generally, this problem cannot have a regular solution $u \in \mathcal{C}^2(S_{n,0})$. If u belongs to this space, then, for any substratum σ of $\partial S_{n,0}$, we have

$$(L_{\mathbf{0}}u)\restriction_{\sigma} = L_{\mathbf{0}}\restriction_{\sigma} u\restriction_{\sigma} . \tag{6.6}$$

Hence f and g would have to satisfy the very restrictive compatibility conditions

$$f\restriction_{\sigma} = L_{\mathbf{0}}\restriction_{\sigma} g\restriction_{\sigma} . \tag{6.7}$$

We look for a solution of the form $u = \widetilde{g} + v$, with v solving

$$L_{\mathbf{0}}v = f - L_{\mathbf{0}}\widetilde{g} \overset{\text{def}}{=} f^{(1)} \text{ and } v\restriction_{bS_{n,0}} = 0. \tag{6.8}$$

Our first goal is to give a method for solving (6.8), which gives a precise description of the singularities that arise for general smooth, compactly supported data (f, g). We begin by giving a precise definition to the meaning of the equations in (6.5): A function $u \in \mathcal{C}^2(\text{int}\, S_{n,0}) \cap \mathcal{C}^0(\overline{S}_{n,0})$ solves (6.5) if $u\restriction_{bS_{n,0}} = g$, and so that $L_{\mathbf{0}}u$, which is initially defined only in the interior of the orthant, has a continuous extension to $\overline{S}_{n,0}$ with $L_{\mathbf{0}}u = f$, throughout.

In Section 3 we showed how the analogous problem for $\mathrm{L}^{\mathrm{Kim}}$ is solved in 1-dimension. We begin this construction by explaining how to solve (6.5) when $n = 1$: For a fixed $\epsilon \ll 1$, we let $\psi \in \mathcal{C}_c^{\infty}([0, \epsilon))$ be a non-negative function, which is 1 in $[0, \frac{\epsilon}{2})$. Suppose that we want to solve

$$x\partial_x^2 u = f(x) \text{ with } u(0) = a \neq 0. \tag{6.9}$$

As noted above, if $u \in \mathcal{C}^2([0, \infty))$, then, unless $f(0) = 0$, it cannot simultaneously satisfy the differential equation as $x \to 0$, and the boundary condition. If we write

$$u(x) = \psi(x)[f(0)x \log x + a] + v(x), \tag{6.10}$$

then

$$x\partial_x^2 u = x\psi''(x)[f(0)x \log x + a] + 2x\psi'(x)[f(0)(\log x + 1)] + \psi(x)f(0) + x\partial_x^2 v. \tag{6.11}$$

The first two terms on the right hand side are smooth and compactly supported away from $x = 0$. The function v must solve the equation

$$x\partial_x^2 v = f(x) - [x\psi''(x)[f(0)x \log x + a] + 2x\psi'(x)[f(0)(\log x + 1)] + \psi(x)f(0)], \tag{6.12}$$

with $v(0) = 0$. The right hand side is smooth and vanishes at $x = 0$, hence the theory presented in [6] shows that there is a unique smooth solution to this problem.

To carry this out in higher dimensions, we need to introduce some notation. For indices $1 \leq i_1 < \cdots < i_k \leq n$, (or $n+1$) we define the vector valued function

$$X_{\hat{i}_1,\ldots,\hat{i}_k} = \sum_{\{j \notin \{i_1,\ldots,i_k\}\}} x_j e_j, \tag{6.13}$$

where e_j are the standard basis vectors for $\mathbb{R}^n$ (or $\mathbb{R}^{n+1}$ — which is needed will be clear from the context). For a function φ defined in $\mathbb{R}^n$, the composition satisfies

$$\varphi(X_{\hat{i}_1,\ldots,\hat{i}_k}) = \varphi\restriction_{x_{i_1}=\cdots=x_{i_k}=0} . \tag{6.14}$$

Our method relies on the following essentially algebraic lemma:

LEMMA 6.2. *Let $\eta(\tau) = \tau \log \tau$, and let $L_{\mathbf{0}}$ denote the operator defined in* (6.3). *For any $k \leq n$, and distinct indices $\{i_1, \ldots, , i_k\} \subset \{1, \ldots, n+1\}$, we have*

$$L_{\mathbf{0}} \eta(x_{i_1} + \cdots + x_{i_k}) = 1. \tag{6.15}$$

Proof. By the permutation symmetry of the operator $L_{\mathbf{0}}$ it suffices to consider the indices $\{1, \ldots, k\}$. Since $\partial_\tau^2 \eta(\tau) = 1/\tau$, this follows from Lemma 6.1. □

As the next step we write the function $v = v_0 + v_1$, where v_0 will be the "regular" part of the solution, vanishing on the boundary, and v_1 is the singular part, which also vanishes on the boundary and satisfies the equation

$$L_{\mathbf{0}} v_1\restriction_{bS_{n,0}} = f^{(1)}\restriction_{bS_{n,0}} . \tag{6.16}$$

As we shall see, v_1 belongs to $\mathcal{C}^{0,\gamma}_{\mathrm{WF}}(S_{n,0})$ for any $0 < \gamma < 1$. In fact we can simply write a formula for v_1 :

$$v_1(x) = \sum_{j=1}^{n} \sum_{\{1 \leq i_1 < \cdots < i_j \leq n\}} (-1)^{j-1} f^{(1)}(X_{\hat{i}_1,\ldots,\hat{i}_j}) \eta(x_{i_1} + \cdots + x_{i_j}). \tag{6.17}$$

For example, if $n = 2$, then

$$v_1(x_1, x_2) = f^{(1)}(x_1, 0)\eta(x_2) + f^{(1)}(0, x_2)\eta(x_1) - f^{(1)}(0, 0)\eta(x_1 + x_2); \tag{6.18}$$

if $n = 3$, then

(6.19)
$$\begin{aligned} v_1(x_1, x_2, x_3) = f^{(1)}(x_1, x_2, 0)\eta(x_3) + f^{(1)}(x_1, 0, x_3)\eta(x_2) + f^{(1)}(0, x_2, x_3)\eta(x_1) - \\ f^{(1)}(x_1, 0, 0)\eta(x_2 + x_3) - f^{(1)}(0, x_2, 0)\eta(x_1 + x_3) - \\ f^{(1)}(0, 0, x_3)\eta(x_1 + x_2) + f^{(1)}(0, 0, 0)\eta(x_1 + x_2 + x_3). \end{aligned}$$

REMARK 6.3. *Formula* (6.17) *should be contrasted to results like those in [4], which analyzes boundary value problems for uniformly elliptic operators in domains with corners. For this case the classical Dirichlet and Neumann boundary value problems are well posed, however, even if the data is infinitely differentiable, these problems typically do* not *have smooth solutions on domains whose boundaries have corner-type singularities. There is no analogue of the "regular" solution, which is uniquely determined in the present case. The existence of a regular solution is another indication of the remarkable relationship between* *the degeneracies of the operator* $\mathrm{L}^{\mathrm{Kim}}$ *and the singular structure of the boundary of the simplex. Singularities arise in the present case from the requirement that the solution satisfy a Dirichlet-type boundary condition, which, as we have seen, is generally impossible for a function in* $\mathcal{C}^2(\Sigma_n)$*.*

A second remarkable feature of this formula is that it provides two locally determined terms in the expansion, along $b\Sigma_n$*, of the solution* u *to the original problem in* (6.5)*. We have*

$$u(x) = \widetilde{g}(x) + v_1(x) + O(\mathrm{dist}(x, b\Sigma_n)), \tag{6.20}$$

where it should be noted that $v_1(x) = O(\mathrm{dist}(x, b\Sigma_n)\log\mathrm{dist}(x, b\Sigma_n))$*. In the non-degenerate case, determination of the second term in such an expansion requires the solution of a global problem.*

THEOREM 6.4. *Let* $f^{(1)} \in \mathcal{C}^{0,2+\gamma}_{\mathrm{WF}}(\Sigma_n)$*, then the function* v_1 *defined in* (6.17) *locally belongs to* $\mathcal{C}^{0,\gamma}_{\mathrm{WF}}(S_{n,0})$*, for any* $0 < \gamma < 1$*, as does* $L_{\mathbf{0}}v_1$*. It satisfies the following equations*

$$L_{\mathbf{0}}v_1\restriction_{bS_{n,0}} = f^{(1)}\restriction_{bS_{n,0}} \quad \textit{and} \quad v_1\restriction_{bS_{n,0}} = 0. \tag{6.21}$$

Proof. The fact that $v_1 \in \mathcal{C}^{0,\gamma}_{\mathrm{WF}}(S_{n,0})$ follows immediately from the fact that $\eta(\tau)$ is locally in $\mathcal{C}^{0,\gamma}_{\mathrm{WF}}(S_{1,0})$. We use Lemma 6.2 and the fact that $X_{\hat{i}_1,\dots,\hat{i}_j}$ and $x_{i_1} + \dots + x_{i_j}$ depend on disjoint subsets of the variables to obtain that

$$L_{\mathbf{0}}v_1 = \sum_{j=1}^{n} \sum_{\{1\le i_1<\dots<i_j\le n\}} (-1)^{j-1}\left[L_{\mathbf{0}}f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j})\eta(x_{i_1}+\dots+x_{i_j}) + f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j})\right] \tag{6.22}$$

It is clear that $L_{\mathbf{0}}v_1 \in \mathcal{C}^{0,\gamma}_{\mathrm{WF}}(S_{n,0})$ for any $0 < \gamma < 1$. The facts that $v_1\restriction_{bS_{n,0}} = 0$, and $L_{\mathbf{0}}v_1\restriction_{bS_{n,0}} = f^{(1)}\restriction_{bS_{n,0}}$ follow from a direct calculation. We first give the proof for the $n = 2$ case: Observe that

$$v_1(x_1, x_2) = f^{(1)}(x_1, 0)\eta(x_2) + f^{(1)}(0, x_2)\eta(x_1) - f^{(1)}(0, 0)\eta(x_1 + x_2). \tag{6.23}$$

We restrict to $x_2 = 0$ to obtain

$$v_1(x_1, 0) = f^{(1)}(0, 0)\eta(x_1) - f^{(1)}(0, 0)\eta(x_1) = 0. \tag{6.24}$$

The case $x_1 = 0$ is identical. Applying $L_{\mathbf{0}}$ we see that

$$\begin{aligned} L_{\mathbf{0}}v_1(x_1, x_2) &= x_1\partial_{x_1}^2 f^{(1)}(x_1, 0)\eta(x_2) + \\ &\quad f^{(1)}(x_1, 0) + x_2\partial_{x_2}^2 f^{(1)}(0, x_2)\eta(x_1) + f^{(1)}(0, x_2) - f^{(1)}(0, 0). \end{aligned} \tag{6.25}$$

Setting $x_2 = 0$ now gives

$$L_{\mathbf{0}}v_1(x_1, 0) = f^{(1)}(x_1, 0) + f^{(1)}(0, 0) - f^{(1)}(0, 0) = f^{(1)}(x_1, 0). \tag{6.26}$$

The general case is not much harder: By symmetry, and continuity it suffices to show that

$$\begin{aligned} &\sum_{j=1}^{n} \sum_{\{1\le i_1<\dots<i_j\le n\}} (-1)^{j-1} f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j})\eta(x_{i_1}+\dots+x_{i_j})\restriction_{x_n=0} = 0 \quad \text{and} \\ &\sum_{j=1}^{n} \sum_{\{1\le i_1<\dots<i_j\le n\}} (-1)^{j-1}\left[L_{\mathbf{0}}f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j})\eta(x_{i_1}+\dots+x_{i_j}) + f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j})\right]_{x_n=0} \\ &= f(X_{\widehat{n}}). \end{aligned} \tag{6.27}$$

We prove the first identity in (6.27) by observing that, if $i_j < n$, then $X_{\hat{i}_1,\dots,\hat{i}_j}\restriction_{x_n=0}= X_{\hat{i}_1,\dots,\hat{i}_j,\widehat{n}}$. We split the sum into two parts and use the fact that $\eta(x_n)\restriction_{x_n=0}=0$ to obtain:

$$\begin{aligned}(6.28)\quad &\sum_{j=1}^{n}\sum_{\{1\leq i_1<\cdots<i_j\leq n\}}(-1)^{j-1}f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j})\eta(x_{i_1}+\cdots+x_{i_j})\restriction_{x_n=0}=\\ &\quad\sum_{j=1}^{n-1}\sum_{\{1\leq i_1<\cdots<i_j\leq n-1\}}(-1)^{j-1}f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j,\widehat{n}})\eta(x_{i_1}+\cdots+x_{i_j})+\\ &\qquad\sum_{j=2}^{n}\sum_{\{1\leq i_1<\cdots<i_{j-1}\leq n-1\}}(-1)^{j-1}f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_{j-1},\widehat{n}})\eta(x_{i_1}+\cdots+x_{i_{j-1}}).\end{aligned}$$

Upon changing $j-1\to j$ in the third line, it becomes clear that the second and third lines in (6.28) differ only by a sign, and therefore sum to zero.

To prove the second identity in (6.27), we observe that the first identity already implies that

$$(6.29)\qquad \sum_{j=1}^{n}\sum_{\{1\leq i_1<\cdots<i_j\leq n\}}(-1)^{j-1}L_{\mathbf{0}}f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j})\eta(x_{i_1}+\cdots+x_{i_j})\restriction_{x_n=0}=0,$$

so we are only left to prove that

$$(6.30)\qquad \sum_{j=1}^{n}\sum_{\{1\leq i_1<\cdots<i_j\leq n\}}(-1)^{j-1}f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j})\restriction_{x_n=0}=f(X_{\widehat{n}}).$$

Applying the same decomposition as before, we see that

$$\begin{aligned}(6.31)\quad &\sum_{j=1}^{n}\sum_{\{1\leq i_1<\cdots<i_j\leq n\}}(-1)^{j-1}f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j})\restriction_{x_n=0}=\\ &\quad f^{(1)}(X_{\widehat{n}})+\sum_{j=1}^{n-1}\sum_{\{1\leq i_1<\cdots<i_j\leq n-1\}}(-1)^{j-1}f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j,\widehat{n}})+\\ &\qquad\sum_{j=2}^{n}\sum_{\{1\leq i_1<\cdots<i_{j-1}\leq n-1\}}(-1)^{j-1}f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_{j-1},\widehat{n}}),\end{aligned}$$

from which the claim is again immediate. □

This almost suffices, except for the fact that v_1 is not generally compactly supported even if the data is. This is because in the definition of v_1 we evaluate $f^{(1)}$ at subsets of the variables $(x_1,\dots,x_n)$. To repair this we multiply v_1 by a specially constructed bump function. To see that this works we need another elementary calculus lemma.

LEMMA 6.5. *Let* $\psi\in\mathcal{C}^2([0,\infty))$ *and* $h\in\mathcal{C}^2(\operatorname{int}(S_{n,0}))$, *then*

$$\begin{aligned}(6.32)\quad &L_{\mathbf{0}}[h\psi(x_1+\cdots+x_n)]=\\ &\psi(x_1+\cdots+x_n)L_{\mathbf{0}}h+2\psi'(x_1+\cdots+x_n)Rh+(x_1+\cdots+x_n)\psi''(x_1+\cdots+x_n)h.\end{aligned}$$

Here $R=x_1\partial_{x_1}+\cdots+x_n\partial_{x_n}$.

Using this lemma we can easily demonstrate the following result:

PROPOSITION 6.6. *Suppose that* $f^{(1)} \in \mathcal{C}^{0,2+\gamma}_{\mathrm{WF}}(S_{n,0})$ *is supported in the set* $\{x:\ x_1+\cdots+x_n \leq N\}$. *If* $\psi \in \mathcal{C}^2_c([0,\infty))$ *equals* 1 *in* $[0,N]$, *then*

$$\widetilde{v}_1(x) = \psi(x_1+\cdots+x_n)v_1(x) \tag{6.33}$$

is compactly supported, $\widetilde{v}_1$ *and* $L_{\mathbf{0}}\widetilde{v}_1$ *belong to* $\mathcal{C}^{0,\gamma}_{\mathrm{WF}}(S_{n,0})$. *The function* $\widetilde{v}_1$ *also satisfies the equations:*

$$L_{\mathbf{0}}\widetilde{v}_1\restriction_{bS_{n,0}} = f^{(1)}\restriction_{bS_{n,0}} \text{ and } \widetilde{v}_1\restriction_{bS_{n,0}} = 0. \tag{6.34}$$

Proof. The fact that $\widetilde{v}_1\restriction_{bS_{n,0}} = 0$ and its regularity properties are immediate. Applying Lemma 6.5 we see that

$$L_{\mathbf{0}}\widetilde{v}_1 = \psi(x_1+\cdots+x_n)L_{\mathbf{0}}v_1 + 2\psi'(x_1+\cdots+x_n)Rv_1 + (x_1+\cdots+x_n)\psi''(x_1+\cdots+x_n)v_1. \tag{6.35}$$

The fact that $\tau\partial_\tau\eta(\tau) = \tau(\log\tau+1)$ implies that Rv_1 has the same regularity as v_1, and therefore so does $L_{\mathbf{0}}\widetilde{v}_1$. Thus to prove the proposition we only need to show that Rv_1 vanishes on the interiors of the hypersurface boundary components. In a neighborhood of the interior of $x_1 = 0$ we have the representation $v_1 = x_1\log x_1 a_1(x) + x_1 a_2(x)$, where a_1 and a_2 belong locally to $\mathcal{C}^2$. Applying the vector field we see that

$$Rv_1 = x_1(\log x_1 + 1)a_1(x) + x_1\log x_1 Ra_1(x) + x_1(a_2(x) + Ra_2), \tag{6.36}$$

which obviously vanishes as $x_1 \to 0$. The other boundary faces follow by an essentially identical argument. □

We now can complete the solution of (6.8) and (6.5) as well. Write $v = \widetilde{v}_0 + \widetilde{v}_1$, where

$$L_{\mathbf{0}}\widetilde{v}_0 = f^{(1)} - L_{\mathbf{0}}\widetilde{v}_1 \text{ and } \widetilde{v}_0\restriction_{bS_{n.0}} = 0. \tag{6.37}$$

Since $f^{(1)} - L_{\mathbf{0}}\widetilde{v}_1 \in \mathcal{C}^{0,\gamma}_{\mathrm{WF}}(S_{n,0})$ is compactly supported and vanishes on the boundary, it follows from the results in [7] that (6.37) has a unique solution $\widetilde{v}_0 \in \mathcal{C}^{0,2+\gamma}_{\mathrm{WF}}(S_{n,0})$. While we call $\widetilde{v}_0$ the regular part of the solution, it will not in general be smooth. It will have complicated singularities of the form $[x_{i_1}+\cdots+x_{i_k}]^l[\log(x_{i_1}+\cdots+x_{i_k})]^m$, for $2 \leq l$, and with the maximum value of m bounded by a function of l.

With small adaptations this method can also be used to treat the case of the n-simplex, Σ_n, and the neutral Kimura diffusion operator, $\mathrm{L}^{\mathrm{Kim}}$. As noted above, the n-simplex is most symmetrically viewed in the affine plane $x_1+\cdots+x_{n+1} = 1$. This representation makes clear that every vertex is identical to every other vertex. For the construction of the solution to the Dirichlet problem:

$$\mathrm{L}^{\mathrm{Kim}} u = f \text{ in } \Sigma_n \text{ with } u\restriction_{b\Sigma_n} = g \tag{6.38}$$

it turns out to be simplest to work in the non-symmetric representation, with one vertex identified with $\mathbf{0}$. Recall that

$$\widetilde{\Sigma}_n = \{(x_1,\ldots,x_n):\ 0 \leq x_i,\ i = 1,\ldots,n\,;\ x_1+\cdots+x_n \leq 1\}. \tag{6.39}$$

It is clear that there is a choice of linear projection into $\mathbb{R}^n$ so that any given vertex of Σ_n is so identified with $\mathbf{0}$. For the moment we work in $\widetilde{\Sigma}_n$ with coordinates $(x_1,\ldots,x_n)$.

We begin by assuming that the data f, g is supported in a set of the form $\widetilde{\Sigma}_{n,2\epsilon} = \{x:\ 0 \leq x_1+\cdots+x_n \leq 1-2\epsilon\}$, for some $0 < \epsilon$. As before we let $\widetilde{g}$ denote an optimally smooth extension of g as a function with support in $\widetilde{\Sigma}_{n,\epsilon}$. We write $u = \widetilde{g} + v$, where v satisfies

$$\mathrm{L}^{\mathrm{Kim}} v = f - \mathrm{L}^{\mathrm{Kim}}\widetilde{g} = f^{(1)}, \text{ with } v\restriction_{b\widetilde{\Sigma}_n} = 0. \tag{6.40}$$

As before we write $v = v_0 + \widetilde{v}_1$, where

$$\widetilde{v}_1 = \psi(x_1 + \cdots + x_n)v_1(x), \tag{6.41}$$

with $\psi \in \mathcal{C}_c^\infty([0, 1 - \frac{\epsilon}{2}))$, satisfying $\psi(\tau) = 1$, for $\tau \in [0, 1 - \epsilon]$. The use of the bump function is much more critical here, as we do not have good control on the function v_1 on the face where $x_1 + \cdots + x_n = 1$.

The function v_1 is defined by the sum:

$$v_1(x) = \sum_{j=1}^{n} \sum_{\{1 \le i_1 < \cdots < i_j \le n\}} (-1)^{j-1} f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j})\eta(x_{i_1} + \cdots + x_{i_j}). \tag{6.42}$$

What is special about the choice of coordinates is the possibility of having the two functions, $f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j})$ and $\eta(x_{i_1} + \cdots + x_{i_j})$, depend on disjoint sets of coordinates. Elementary calculations show that

$$\mathrm{L}^{\mathrm{Kim}}\,\eta(x_{i_1} + \cdots + x_{i_j}) = 1 - (x_{i_1} + \cdots + x_{i_j}), \tag{6.43}$$

and

(6.44)

$$\begin{aligned} \mathrm{L}^{\mathrm{Kim}}[f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j})\eta(x_{i_1} + \cdots + x_{i_j})] = \eta(x_{i_1} + \cdots + x_{i_j})\,\mathrm{L}^{\mathrm{Kim}}\,f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j}) - \\ 2(x_{i_1} + \cdots + x_{i_j})\eta'(x_{i_1} + \cdots + x_{i_j})R_{\hat{i}_1,\dots,\hat{i}_j} f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j}) + \\ + (1 - (x_{i_1} + \cdots + x_{i_j}))f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j}), \end{aligned}$$

where the vector fields $R_{\hat{i}_1,\dots,\hat{i}_j}$ are defined by

$$R_{\hat{i}_1,\dots,\hat{i}_j} = \sum_{k \notin \{i_i,\dots,i_j\}} x_k \partial_{x_k}. \tag{6.45}$$

Let $b\widetilde{\Sigma}'_n = b\widetilde{\Sigma}_n \setminus \{x :\ x_1 + \cdots + x_n = 1\}$. Arguing as before we show that v_1 satisfies the equation along $b\widetilde{\Sigma}'_n$.

THEOREM 6.7. *Let* $f^{(1)} \in \mathcal{C}^{0,2+\gamma}_{\mathrm{WF}}(\Sigma_n)$. *The function* v_1 *defined in* (6.42) *belongs to* $\mathcal{C}^{0,\gamma}_{\mathrm{WF}}(\widetilde{\Sigma}_n)$, *as does* $\mathrm{L}^{\mathrm{Kim}}\,v_1$. *It satisfies the following equations*

$$\mathrm{L}^{\mathrm{Kim}}\,v_1\restriction_{b\widetilde{\Sigma}'_n} = f^{(1)}\restriction_{b\widetilde{\Sigma}'_n} \quad \textit{and} \quad v_1\restriction_{b\widetilde{\Sigma}'_n} = 0. \tag{6.46}$$

Proof. The regularity statements for v_1 and $\mathrm{L}^{\mathrm{Kim}}\,v_1$ follow easily from (6.42) and (6.44). The fact that $v_1\restriction_{b\widetilde{\Sigma}'_n} = 0$ follows from Theorem 6.4. The proof that $\mathrm{L}^{\mathrm{Kim}}\,v_1\restriction_{b\widetilde{\Sigma}'_n} = f^{(1)}\restriction_{b\widetilde{\Sigma}'_n}$ is similar to the proof of Theorem 6.4. As before continuity shows that we only need to prove this statement for the interiors of the hypersurface faces, and symmetry shows that it suffices to consider $\{x_n = 0\}$. The terms coming from the first line of (6.44) can be treated exactly as before. For the terms from the third line of (6.44), the sole difference is the coefficient $(1 - x_n)$, which equals 1 where $x_n = 0$. The first order cross terms, coming from the second line of (6.44), require some additional consideration.

We once again use the observation that if $i_j < n$, then $X_{\hat{i}_1,\dots,\hat{i}_j}\restriction_{x_n=0} = X_{\hat{i}_1,\dots,\hat{i}_j,\widehat{n}}$, as well as the facts that

$$R_{\hat{i}_1,\dots,\hat{i}_j} f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j})\restriction_{x_n=0} = R_{\hat{i}_1,\dots,\hat{i}_j,\widehat{n}} f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j,\widehat{n}}),$$

and

$$x_n \partial_{x_n} \eta(x_n) \restriction_{x_n=0} = (x_n \log x_n + x_n) \restriction_{x_n=0} = 0,$$

to split the sum into two parts, obtaining:

(6.47)

$$\sum_{j=1}^{n} \sum_{\{1\leq i_1<\cdots<i_j\leq n\}} (-1)^{j-1} 2(x_{i_1}+\cdots+x_{i_j})\eta'(x_{i_1}+\cdots+x_{i_j}) R_{\hat{i}_1,\dots,\hat{i}_j} f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j}) \restriction_{x_n=0} =$$
$$\sum_{j=1}^{n-1} \sum_{\{1\leq i_1<\cdots<i_j\leq n-1\}} (-1)^{j-1} 2(x_{i_1}+\cdots+x_{i_j})\eta'(x_{i_1}+\cdots+x_{i_j}) R_{\hat{i}_1,\dots,\hat{i}_j,\widehat{n}} f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j,\widehat{n}}) +$$
$$\sum_{j=2}^{n} \sum_{\{1\leq i_1<\cdots<i_{j-1}\leq n-1\}} (-1)^{j-1} 2(x_{i_1}+\cdots+x_{i_j})\eta'(x_{i_1}+\cdots+x_{i_j}) R_{\hat{i}_1,\dots,\hat{i}_j,\widehat{n}} f^{(1)}(X_{\hat{i}_1,\dots,\hat{i}_j,\widehat{n}}).$$

As before, after changing variables in the third line with $j \mapsto j-1$, we see that the second and third lines differ only by a sign, demonstrating that these terms sum to zero along the face given by $\{x_n = 0\}$. This completes the proof of the theorem. □

To complete the discussion of this case we need to check that $\widetilde{v}_1$ also satisfies the equations in (6.46). An elementary calculation shows that

$$\text{(6.48)}\quad \mathrm{L}^{\mathrm{Kim}}[v_1\psi(x_1+\cdots+x_n)] = \\ \psi(x_1+\cdots+x_n)\,\mathrm{L}^{\mathrm{Kim}}\, v_1 + \psi'(x_1+\cdots+x_n)\widetilde{R}v_1 + v_1\,\mathrm{L}^{\mathrm{Kim}}\,\psi,$$

where

$$\widetilde{R} = 2(1-(x_1+\cdots+x_n))\sum_{j=1}^{n} x_i\partial_{x_i}. \tag{6.49}$$

As before, $\widetilde{R}$ is tangent to $b\widetilde{\Sigma}_n$, and $\widetilde{R}v_1 \in \mathcal{C}^{0,\gamma}_{\mathrm{WF}}(\widetilde{\Sigma}_n)$. It is easy to see that

$$\widetilde{R}v_1 \restriction_{b\widetilde{\Sigma}'_n} = 0. \tag{6.50}$$

Summarizing, we have shown

PROPOSITION 6.8. *Let $f^{(1)} \in \mathcal{C}^{0,2+\gamma}_{\mathrm{WF}}(\Sigma_n)$. The function $\widetilde{v}_1$ belongs to $\mathcal{C}^{0,\gamma}_{\mathrm{WF}}(\widetilde{\Sigma}_n)$, as does $\mathrm{L}^{\mathrm{Kim}}\,\widetilde{v}_1$. If $\psi(x_1+\cdots+x_n)=1$ on $\operatorname{supp} f^{(1)} \subset \widetilde{\Sigma}_{n,\epsilon}$, for an $\epsilon>0$, then $\widetilde{v}_1$ satisfies the following equations*

$$\mathrm{L}^{\mathrm{Kim}}\,\widetilde{v}_1\restriction_{b\widetilde{\Sigma}_n} = f^{(1)}\restriction_{b\widetilde{\Sigma}_n} \quad \textit{and} \quad \widetilde{v}_1\restriction_{b\widetilde{\Sigma}_n} = 0. \tag{6.51}$$

We can now complete the solution of (6.40). We write $v = v_0 + \widetilde{v}_1$, the function v_0 must satisfy

$$\mathrm{L}^{\mathrm{Kim}}\, v_0 = f^{(1)} - \mathrm{L}^{\mathrm{Kim}}\,\widetilde{v}_1 \ \text{ with } \ v_0\restriction_{b\Sigma_n} = 0. \tag{6.52}$$

The existence of a solution $v_0 \in \mathcal{C}^{0,2+\gamma}_{\mathrm{WF}}(\widetilde{\Sigma}_n)$, which can be taken to vanish on the boundary, follows from the results in [7]. Numerically, the method of Section 5 (with $k=n$) can be used to solve (6.52) for v_0.

Finally if we have general data (f, g) so that f, and $\mathrm{L}^{\mathrm{Kim}}\,\widetilde{g} \in \mathcal{C}^{0,2+\gamma}_{\mathrm{WF}}(\Sigma_n)$, then we choose a partition of unity $\{\varphi_1, \ldots, \varphi_{n+1}\}$ so that the function φ_j equals 1 in a neighborhood of the vertex $e_j \subset \Sigma_n \subset \mathbb{R}^{n+1}$, and vanishes in a neighborhood of the opposite face, where $x_j = 0$. The data $\{\varphi_j(f, \widetilde{g}) : j = 1, \ldots, n+1\}$ satisfies the hypotheses used above, and therefore we can construct solutions $\{u_j\}$ to the equations

$$\mathrm{L}^{\mathrm{Kim}}\, u_j = \varphi_j f \text{ on } \Sigma_n, \text{ with } u_j\restriction_{b\Sigma_n} = \varphi_j g, \text{ for } j = 1, \ldots, n+1. \tag{6.53}$$

To construct the solution in a neighborhood of $e_j \in \mathbb{R}^{n+1}$ we project the simplex into the hyperplane $\{x_j = 0\}$, and use the projective representation where e_j corresponds to $\mathbf{0}$.

Now setting

$$u = u_1 + \cdots + u_{n+1}, \tag{6.54}$$

we obtain a solution to the original boundary value problem

$$\mathrm{L}^{\mathrm{Kim}}\, u = f \text{ on } \Sigma_n \text{ with } u\restriction_{b\Sigma_n} = g. \tag{6.55}$$

This proves the following general result:

THEOREM 6.9. *Let* $f \in \mathcal{C}^{0,2+\gamma}_{\mathrm{WF}}(\Sigma_n)$, *and* g *have an extension* $\widetilde{g}$ *to* Σ_n *so that* $\mathrm{L}^{\mathrm{Kim}}\,\widetilde{g} \in \mathcal{C}^{0,2+\gamma}_{\mathrm{WF}}(\Sigma_n)$. *The Dirichlet problem* (6.55) *has unique solution* u, *which takes the form* $u = \widetilde{g} + u^{(0)} + u^{(1)}$, *where* $u^{(0)} \in \mathcal{C}^{0,2+\gamma}_{\mathrm{WF}}(\Sigma_n)$, *vanishes on the boundary. The singular part,* $u^{(1)}$ *is of the form*

$$u^{(1)}(x) = \sum_{i=1}^{n+1}\sum_{j=1}^{n}\sum_{\{1\leq i_1<\cdots<i_j\leq n+1:\, i_m\neq i\}}(-1)^{j-1}F(X_{\{\hat{i}_1,\ldots,\hat{i}_j\}})\eta(x_{i_1}+\cdots+x_{i_j}) \tag{6.56}$$

Here F *is a function constructed from the pullbacks of* $\{\varphi_j f^{(1)}\}$ *to the affine model* Σ_n.

REMARK 6.10. *The solution* u *to the boundary value problem in* (6.55) *also has an explicit 2-term expansion at the boundary similar to that given in* (6.20)*:*

$$u(x) = \widetilde{g}(x) + u^{(1)}(x) + O(\mathrm{dist}(x, b\Sigma_n)). \tag{6.57}$$

Proof. Everything has been proved except the uniqueness statement. This follows from the maximum principle and the facts that u is continuous in the closed simplex, and $\mathrm{L}^{\mathrm{Kim}}$ is a strongly elliptic operator in the interior of the simplex. □

We observe that u has, in some sense, very complicated singularities, in that it includes a smooth function times $(x_{i_1} + \cdots + x_{i_j})\log(x_{i_i} + \cdots + x_{i_j})$, for each set of indices

$$1 \leq i_1 < \cdots < i_j \leq n+1, \text{ for } j \in \{1, \ldots, n\}.$$

To resolve these singularities to be classically conormal, one would, in principle, need to successively blow up all the strata of the boundary, starting with the codimension n parts and proceeding upwards to the codimension 2 part. Given the very explicit form that this singularity takes, such an approach would only obscure its simple and rather benign structure. This general approach is pursued in the papers [20, 21, 22]. These authors do not require the data (f, g) to be continuous, but allow data with complicated singularities along the boundary.

We remark that the approach proposed in Section 4 for solving the regular problem $\mathrm{L}^{\mathrm{Kim}}\, u = f$ (without boundary conditions) can be modified slightly to provide an algorithm for extending g (defined on the boundary) to $\widetilde{g}$ (in the interior) in such a way that $\mathrm{L}^{\mathrm{Kim}}\,\widetilde{g}$ is very easily computed. We write $g = g_0 + \cdots + g_{n-1}$, where $g_0 = \sum_{j=1}^{n+1} g(e_j)x_j$ agrees with g on Σ_n^0, and

$$g_k\restriction_{\Sigma_n^k} = g\restriction_{\Sigma_n^k} - \sum_{j=0}^{k-1} g_j\restriction_{\Sigma_n^k}, \qquad (1 \leq k \leq n-1).$$

The right-hand side is zero on Σ_n^{k-1}, so this equation decouples into independent homogeneous "Dirichlet" extension problems from the connected components of $\Sigma_n^k \setminus \Sigma_n^{k-1}$ to Σ_n.

Using the eigenfunction expansion to represent the right-hand side on each face of Σ_n^k gives the desired representations

$$g_k = \sum_{\mathcal{I}} c_{\mathcal{I},\vec{m}}(w_{\mathcal{I}}\psi_{\mathcal{I},\vec{m}}) \text{ and } \mathrm{L}^{\mathrm{Kim}}\, g_k = \sum_{\mathcal{I}} (\lambda_{\mathcal{I},\vec{m}} c_{\mathcal{I},\vec{m}})(w_{\mathcal{I}}\psi_{\mathcal{I},\vec{m}}), \tag{6.58}$$

with both functions canonically defined throughout Σ_n. If g is a polynomial of degree less than or equal to d on each face of $b\Sigma_n$, then $\widetilde{g} \in \mathcal{P}_d$ as well. Thus, we can replace f by $f^{(1)} = f - \mathrm{L}^{\mathrm{Kim}}\, \widetilde{g}$ in (6.40) at the outset, thereby avoiding non-homogeneous boundary conditions when working with the partition of unity.

Our approach to solving the Dirichlet problem works equally well if we add a vector field V that is everywhere tangent to $b\Sigma_n$. In a projective chart, such a vector field takes the form

$$\widetilde{V} = \sum_{j=1}^{n} b_j(x) x_j \partial_{x_j}, \tag{6.59}$$

with the additional requirement that

$$\sum_{j=1}^{n} x_j b_j(x)\restriction_{x_1+\cdots+x_n=1} = 0. \tag{6.60}$$

A simple calculation shows that

$$\widetilde{V}\eta(x_{i_1} + \cdots + x_{i_j}) = \sum_{l=1}^{n} a_l(x) x_l \Delta_{\{i_1,\ldots,i_j\}}(l)[\log(x_{i_1} + \cdots + x_{i_j}) + 1], \tag{6.61}$$

where

$$\Delta_{\{i_1,\ldots,i_j\}}(l) = \begin{cases} 1 \text{ if} & l \in \{i_1,\ldots,i_j\} \\ 0 \text{ if} & l \notin \{i_1,\ldots,i_j\}. \end{cases} \tag{6.62}$$

The function on the right hand side of (6.61) is easily seen to be continuous on the closed n-simplex. In fact these functions belong to $\mathcal{C}^{0,\gamma}_{\mathrm{WF}}(\Sigma_n)$, for any $0 < \gamma < 1$. Applying $\widetilde{V}$ to $\widetilde{v}_1$ we see that

$$\widetilde{V}\widetilde{v}_1 = v_1 \widetilde{V}\psi(x_1 + \cdots + x_n) + \psi(x_1 + \cdots + x_n)\widetilde{V} v_1, \tag{6.63}$$

from which it is clear that $\widetilde{V}\widetilde{v}_1$ is continuous on Σ_n. To show that $\widetilde{V}\widetilde{v}_1\restriction_{b\Sigma_n} = 0$, it therefore suffices to prove it in the interiors of the hypersurface boundary faces, but this is clear from (6.61) and (6.63). With these observations it follows that $(\mathrm{L}^{\mathrm{Kim}} + \widetilde{V})\widetilde{v}_1 \in \mathcal{C}^{0,\gamma}_{\mathrm{WF}}(\Sigma_n)$ and

$$(\mathrm{L}^{\mathrm{Kim}} + \widetilde{V})\widetilde{v}_1\restriction_{b\Sigma_n} = \mathrm{L}^{\mathrm{Kim}}\, \widetilde{v}_1\restriction_{b\Sigma_n} = f^{(1)}\restriction_{b\Sigma_n}. \tag{6.64}$$

Proceeding as above we easily demonstrate that, if V is tangent to the boundary $b\Sigma_n$, then the Dirichlet problem:

$$(\mathrm{L}^{\mathrm{Kim}} + V)u = f \text{ in } \Sigma_n \text{ with } u\restriction_{b\Sigma_n} = g \tag{6.65}$$

has a unique solution of the form $u = u_0 + u_1$, where u_1 is given by the formula in (6.56), and $u_0 \in \dot{\mathcal{C}}^{0,2+\gamma}_{\mathrm{WF}}(\Sigma_n)$.

REFERENCES

[1] M. Abramowitz and I. A. Stegun, eds., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, Inc., New York, 9th ed., 1972.

[2] A. Bhaskar, Y. Wang, and Y. Song, *Efficient inference of population size histories and locus-specific mutation rates from large-sample genomic variation data*, Genome Research, 25 (2015), pp. 268–279.

[3] J. P. Bollback, T. L. York, and R. Nielsen, *Estimation of 2Nes From Temporal Allele Frequency Data*, Genetics, 179 (2008), pp. 497–502.

[4] M. Costabel and M. Dauge, *Construction of corner singularities for Agmon-Douglis-Nirenberg elliptic systems*, Math. Nachr., 162 (1993), pp. 209–237.

[5] R. Der, *A Theory of Generalised Population Processes*, ProQuest LLC, Ann Arbor, MI, 2010. Thesis (Ph.D.)–University of Pennsylvania.

[6] C. L. Epstein and R. Mazzeo, *Wright-Fisher diffusion in one dimension*, SIAM J. Math. Anal., 42 (2010), pp. 568–608.

[7] ———, *Degenerate Diffusion Operators Arising in Population Biology*, vol. 185 of Annals of Mathematics Studies, Princeton University Press, Princeton, NJ, 2013.

[8] ———, *Harnack inequalities and heat-kernel estimates for degenerate diffusion operators arising in population biology*, Applied Math. Res. Express, doi:10.1093/amrx/abw002, (2016), 64 pp. arXiv:1406.1426.

[9] C. L. Epstein and C. Pop *Transition probabilities for degenerate diffusions arising in population genetics*, arXiv:1608.02119 (2016) 49pp.

[10] S. N. Evans, J. G. Schraiber, and M. Slatkin, *Bayesian inference of natural selection from allele frequency time series*, preprint, (2014), p. 35pp.

[11] S. N. Evans, Y. Shvets, and M. Slatkin, *Non-equilibrium theory of the allele frequency spectrum*, Theoretical Population Biology, 71 (2007), pp. 109–119.

[12] W. Ewens, *Mathematical Population Genetics, I, 2nd edition*, vol. 27 of Interdisciplinary Applied Mathematics, Springer Verlag, Berlin and New York, 2004.

[13] R. Fisher, *On the dominance ratio*, Proc. Roy. Soc. Edin., 42 (1922), pp. 321–341.

[14] W. Gautschi, *Orthogonal Polynomials, Computation and Approximation*, Oxford University Press, Oxford, 2004.

[15] G. Golub and J. Welsch, *Calculation of Gauss quadrature rules*, Math. Comput., 23 (1969), pp. 221–230.

[16] R. C. Griffiths, *A transition density expansion for a multi-allele diffusion model*, Adv. in Appl. Probab., 11 (1979), pp. 310–325.

[17] R. C. Griffiths and D. Spanó, *Diffusion processes and coalescent trees*, Chapter 15, 358–375. In *Probability and Mathematical Genetics, Papers in Honour of Sir John Kingman.* LMS Lecture Note Series 378. Edited by N. H. Bingham and C. M. Goldie. Cambridge University Press. (2010)

[18] ———, *Multivariate Jacobi and Laguerre polynomials, infinite-dimensional extensions, and their probabilistic connections with multivariate Hahn and Meixner polynomials*, Bernoulli., 17 (2011), pp. 1095–1125.

[19] R. Gutenkunst, R. Hernandez, S. Williamson, and C. Bustamante, *Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data*, PLoS Genetics, 5 (2009), p. e1000695.

[20] J. Hofrichter, T. D. Tran, and J. Jost, *A hierarchical extension scheme for backward solutions of the Wright-Fisher model*, preprint, (2014), p. 31pp. arXiv:1406.5146.

[21] ———, *A hierarchical extension scheme for solutions of the Wright-Fisher model*, preprint, (2014), p. 21pp. arXiv:1406.5152.

[22] ———, *The uniqueness of hierarchically extended backward solutions of the Wright-Fisher model*, (2014) to appear Communications in PDE. arXiv:1407.3067.

[23] E. G. Kalnins, W. Miller, and M. V. Tratnik, *Families of orthogonal and biorthogonal polynomials on the n-sphere*, SIAM J. Math. Anal., 22 (1991), pp. 272–294.

[24] J. A. Kamm, J. Terhorst, and Y. Song, *Efficient computation of the joint sample frequency spectra for multiple populations*, Journal of Computational and Graphical Statistics, to appear (2016), pp. 1–25.

[25] S. Karlin and J. McGregor, *On some stochastic models in genetics*, in Stochastic models in medicine and biology, J. Gurland, ed., vol. 10, Academic Press, 1963, pp. 245–271.

[26] M. Kimura, *Random genetic drift in multi-allelic locus*, Evolution, 9 (1955), pp. 419–435.

[27] ———, *Solution of a process of random genetic drift with a continuous model*, Proc. Natl. Acad. Sci., 41 (1955), pp. 144–150.

[28] ———, *Stochastic processes and distribution of gene frequencies under natural selections*, Cold Spring Harbor on Quant. Biol., 20 (1955), pp. 33–53.

[29] ———, *Random genetic drift in a tri-allelic locus; exact solution with a continuous model*, Biometrics, 12 (1956), pp. 57–66.

[30] ———, *Diffusion models in population genetics*, Journal of Applied Probability, 1 (1964), pp. 177–232.

[31] T. Koornwinder, *Two-variable analogues of the classical orthogonal polynomials*, in Theory and application of special functions, R. Askey, ed., vol. 35, Academic Press, 1975, pp. 435–495.

[32] M. Lacerda and C. Seoighe, *Population Genetics Inference for Longitudinally-Sampled Mutants Under Strong Selection*, Genetics, 198 (2014), pp. 1237–1250.

[33] W.-H. Li, *Maintenance of genetic variability under mutation and selection pressures in a finite population*, Proc. Natl. Acad. Sci. USA, 74 (1977) 2509–2513.

[34] R. A. Littler, *Loss of variability at one locus in a finite population*, Math. Biosci., 25 (1975), pp. 151-163.

[35] ———and E. D. Fackerell, *Transition densities for neutral multi-allele diffusion models*, Biometrics, 31 (1975), pp. 117–123.

[36] J. Proriol, *Sur une famille de polynomes á deux variables orthogonaux dans un triangle*, C. R. Acad. Sci. Paris, 245 (1957), pp. 2459–2461.

[37] K. Sato, *Diffusion operators in population genetics and convergence of Markov chains*, in Measure Theory Applications to Stochastic Analysis, G. Kallianpur and D. Kölzow, eds., vol. 695 of Lecture Notes in Mathematics, Springer Berlin Heidelberg, 1978, pp. 127–137.

[38] N. Shimakura, *Équations différentielles provenant de la génétique des populations*, Tôhoku Math. J., 29 (1977), pp. 287–318.

[39] ———, *Formulas for diffusion approximations of some gene frequency models*, J. Math. Kyoto Univ., 21 (1981), pp. 19–45.

[40] Y. S. Song and M. Steinrücken, *A Simple Method for Finding Explicit Analytic Transition Densities of Diffusion Processes with General Diploid Selection*, Genetics, 190 (2012), pp. 1117–1129.

[41] M. Steinrücken, A. Bhaskar, and Y. S. Song, *A novel spectral method for inferring general diploid selection from time series genetic data*, The Annals of Applied Statistics, 8 (2014), pp. 2203–2222.

[42] M. Steinrücken, Y. X. R. Wang, and Y. S. Song, *An explicit transition density expansion for a multi-allelic Wright-Fisher diffusion with general diploid selection*, Theoretical Population Biology, 83 (2013), pp. 1–14.

[43] G. Szegő, *Orthogonal Polynomials*, American Mathematical Society, Providence, 1939.

[44] B. Wingate and M. A. Taylor, *The natural function space for triangular and tetrahedral spectral elements*, Los Alamos National Laboratory Tech. Rep., LA-UR-98-1711 (1998).

[45] S. Wright, *Evolution in Mendelian populations*, Genetics, 16 (1931), pp. 97–159.