TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 369, Number 8, August 2017, Pages 5543–5579
http://dx.doi.org/10.1090/tran/6853
Article electronically published on March 1, 2017

EXISTENCE, UNIQUENESS AND THE STRONG MARKOV PROPERTY OF SOLUTIONS TO KIMURA DIFFUSIONS WITH SINGULAR DRIFT

CAMELIA A. POP

ABSTRACT. Motivated by applications to proving regularity of solutions to degenerate parabolic equations arising in population genetics, we study existence, uniqueness, and the strong Markov property of weak solutions to a class of degenerate stochastic differential equations. The stochastic differential equations considered in our article admit solutions supported in the set $[0,\infty)^n \times \mathbb{R}^m$, and they are degenerate in the sense that the diffusion matrix is not strictly elliptic, as the smallest eigenvalue converges to zero at a rate proportional to the distance to the boundary of the domain, and the drift coefficients are allowed to have power-type singularities in a neighborhood of the boundary of the domain. Under suitable regularity assumptions on the coefficients, we establish existence of solutions that satisfy the strong Markov property, and uniqueness in law in the class of Markov processes.

CONTENTS

1. INTRODUCTION

The stochastic differential equations considered in our article are motivated by their applications to population genetics, in the study of gene frequencies. One of the earliest random models for gene frequencies was proposed by Wright [36] and Fisher [21,22], and it is a discrete Markov chain. Because many quantities of interest in biology are difficult to compute in the framework of the Wright-Fisher discrete Markov chain, there has been extensive research [16,17,20,26,27,32–34] to make precise in which sense and in which scaling regime does the discrete Markov chain converge to a continuous diffusion process. In our article we are concerned with the study of continuous stochastic processes, which arise as limits and generalize the Wright-Fisher discrete Markov chain.

Received by the editors July 20, 2014 and, in revised form, August 17, 2015 and September 2, 2015.

2010 *Mathematics Subject Classification.* Primary 60J60; Secondary 35J70.

Key words and phrases. Kimura diffusions, singular drift coefficient, degenerate diffusions, degenerate elliptic operators, strong Markov property, anisotropic Hölder spaces.

A concise description of the continuous limit of the Wright-Fisher model, which we will call the Wright-Fisher process from now on, is as follows. Assume that there are $d+1$ alleles, and we denote by x_i the frequency of the i-th allele, for all $1 \leq i \leq d+1$. We see that $0 \leq x_i \leq 1$, for all $1 \leq i \leq d+1$, and $x_1 + \ldots + x_{d+1} = 1$. Fixing an allele, say the $(d+1)$-th allele, the remaining gene frequencies, $(x_1, \ldots, x_d)$, follow a diffusion supported in the d-dimensional simplex,

$$\Sigma_d := \{x \in \mathbb{R}^d : x_i \geq 0, \text{ for all } i = 1, \ldots, d, \text{ and } x_1 + \ldots + x_d \leq 1\}.$$

The Wright-Fisher model for gene frequencies is a (strong) Markov process supported in Σ_d with infinitesimal generator,

$$\widehat{L}_{\mathrm{WF}} u := \frac{1}{2} \sum_{i,j=1}^{d} x_i(\delta_{i,j} - x_j) u_{x_i x_j} + \sum_{i=1}^{d} \left(b_i - x_i \sum_{j=1}^{d+1} b_j \right) u_{x_i}, \quad \forall\, u \in C^2(\Sigma_d), \tag{1.1}$$

where b_i are nonnegative constants, for all $1 \leq i \leq d+1$; see [34, Equation (3.4)]. Above we denote by $\delta_{i,j}$ the Kronecker delta symbol.

In our article, we study a more general class of second order differential operators, which we call generalized Kimura operators, and which extend the Wright-Fisher operators, (1.1). Such operators have been introduced in the work of C. Epstein and R. Mazzeo, [13], and versions of these operators have been studied previously in the literature in connection with superprocesses; see [2, 4]. Generalized Kimura operators are defined on compact manifolds with corners, of which simplices and polyhedra are particular examples, and they have the property that in an adapted local system of coordinates the generalized Kimura operators take the form:

$$\begin{aligned} \widehat{L} u &= \sum_{i=1}^{n} \left(x_i d_{i,i}(z) u_{x_i x_i} + b_i(z) u_{x_i} \right) + \sum_{i,j=1}^{n} x_i x_j d_{i,j}(z) u_{x_i x_j} \\ &\quad + \sum_{i=1}^{n} \sum_{l=1}^{m} x_i c_{i,l}(z) u_{x_i y_l} + \sum_{k,l=1}^{m} d_{n+k,n+l}(z) u_{y_k y_l} + \sum_{l=1}^{m} e_l(z) u_{y_l}, \end{aligned} \tag{1.2}$$

where n, m are nonnegative integers, and we denote $S_{n,m} := \mathbb{R}_+^n \times \mathbb{R}^m$, and $\mathbb{R}_+ := (0, \infty)$; see [13, Proposition 2.2.3]. The main difficulties in studying second order operators of the form $\widehat{L}$, and the Markov processes associated to them, arise from the following considerations:

1. The coefficients corresponding to the second order derivatives of $\widehat{L}$ form a degenerate nonnegative definite matrix because the smallest eigenvalue converges to zero at a rate linearly proportional to the distance to the boundary, $\partial S_{n,m}$.
2. The first order terms, $b_i(z)\partial_{x_i}$, are *not* lower order terms, as it is the case for strictly elliptic operators, because the second and first order terms in the operator $\widehat{L}$ scale identically. For this reason, the sign of the coefficient $b_i(z)$ on the portion of the boundary $\partial S_{n,m} \cap \{x_i = 0\}$, for all $1 \leq i \leq n$, plays a crucial role in the analysis. We always assume that $b_i(z) \geq 0$ on $\partial S_{n,m} \cap \{x_i = 0\}$, which ensures that the Markov processes associated to the operator $\widehat{L}$ is supported in $\bar{S}_{n,m}$.
3. The domain $S_{n,m}$ is nonsmooth, it has corners and edges, which makes the analysis of the regularity of solutions to the parabolic equation defined by the operator $\widehat{L}$ nonstandard.

One of our main tasks is to prove that there is a unique Markov process associated to the operator $\widehat{L}$. For this purpose, we consider the stochastic differential equation

$$
\begin{aligned}
d\widehat{X}_i(t) &= b_i(\widehat{Z}(t))\,dt + \sqrt{\widehat{X}_i(t)}\sum_{k=1}^{n+m}\sigma_{i,k}(\widehat{Z}(t))\,d\widehat{W}_k(t),\\
d\widehat{Y}_l(t) &= e_l(\widehat{Z}(t))\,dt + \sum_{k=1}^{n+m}\sigma_{n+l,k}(\widehat{Z}(t))\,d\widehat{W}_k(t),
\end{aligned} \tag{1.3}
$$

where $1 \leq i \leq n$, $1 \leq l \leq m$, $\{\widehat{W}(t)\}_{t\geq 0}$ is an $(n+m)$-dimensional Brownian motion, and we denote $\widehat{Z}(t) := (\widehat{X}(t), \widehat{Y}(t))$, for all $t \geq 0$. Under suitable structural and regularity conditions on the coefficients of (1.3), described in Assumptions 2.1 and 2.6, we prove that the generalized Kimura equation (1.3) has a unique weak solution, and that it satisfies the strong Markov property; see Propositions 2.2 and 2.4, and Corollary 2.5.

We then extend our results to operators, $L = \widehat{L} + V$, where V is a first order vector field, which can include logarithmically divergent coefficients,

$$
\sqrt{x_i}\ln x_j\eta(z)\partial_{x_i},\quad \ln x_j\eta(z)\partial_{y_l},\quad \forall\, i,j = 1,\dots,n,\quad \forall\, l = 1,\dots,m, \tag{1.4}
$$

where $\eta : \bar{S}_{n,m} \to [0,1]$ is a smooth function with compact support, or V can include power-type coefficients,

$$
\sqrt{x_i}|x_j|^{-q}\partial_{x_i},\quad |x_j|^{-q}\partial_{y_l},\quad \forall\, i,j = 1,\dots,n,\quad \forall\, l = 1,\dots,m, \tag{1.5}
$$

when the positive power, q, is suitably chosen. Specifically, we consider stochastic differential equations of the form

$$
\begin{aligned}
dX_i(t) &= \left(b_i(Z(t)) + \sqrt{X_i(t)}\sum_{j=1}^{n} f_{i,j}(Z(t))h_{i,j}(X_j(t))\right) dt\\
&\quad + \sqrt{X_i(t)}\sum_{k=1}^{n+m}\sigma_{i,k}(Z(t))\,dW_k(t),\\
dY_l(t) &= \left(e_l(Z(t)) + \sum_{j=1}^{n} f_{n+l,j}(Z(t))h_{n+l,j}(X_j(t))\right) dt\\
&\quad + \sum_{k=1}^{n+m}\sigma_{n+l,k}(Z(t))\,dW_k(t),
\end{aligned} \tag{1.6}
$$

where $1 \leq i \leq n$, $1 \leq l \leq m$, $\{W(t)\}_{t\geq 0}$ is an $(n+m)$-dimensional Brownian motion, and we denote $Z(t) = (X(t), Y(t))$, for all $t \geq 0$. The functions $h_{i,j}(x_j)$ are required to satisfy condition (3.5), which implies that the infinitesimal generator of any Markov solution to the singular Kimura equation (1.6) has the structure $L = \widehat{L} + V$, where the vector field V can contain singular terms such as (1.4) and (1.5). In Theorems 3.1 and 3.7, we establish existence of strong Markov solutions to the generalized Kimura stochastic differential equation with singular drift (1.6), and uniqueness in law in the class of Markov processes.

Our motivation to include logarithmic singularities in the drift coefficient of generalized Kimura stochastic differential equations is two-fold. On one hand, in our proof of the Harnack inequality of nonnegative solutions to the parabolic equation defined by the Kimura operator $\widehat{L}$, [14, Theorem 1.2], we were led to the study of perturbed operators $L = \widehat{L} + V$, where the vector field V contains logarithmic singular terms such as (1.4). An essential ingredient in the proof of [14, Theorem 1.2] is that the singular Kimura stochastic differential equation (1.6) has a unique weak solution that satisfies the strong Markov property. A second motivation to study (1.6) comes from the fact that the adjoint operator of the unbounded operator $\widehat{L}$, with domain of definition included in the weighted Sobolev space $L^2(S_{n,m}; d\mu_b)$, where the weight $d\mu_b(z)$ is defined by

$$d\mu_b(z) := \prod_{i=1}^n x_i^{b_i(z)-1}\, dx_i \prod_{l=1}^m dy_l, \quad \forall\, z = (x, y) \in S_{n,m},$$

is an operator $L = \widehat{L} + V$, where V contains logarithmically divergent singularities (1.4). This observation is extensively used in our ongoing joint work with C. Epstein on the regularity and the structure of the transition probabilities of generalized Kimura processes, [15]. For more details regarding the functional analytic framework for generalized Kimura operators defined on the weighted Sobolev space $L^2(\bar{S}_{n,m}; d\mu_b)$, we refer the reader to [11], and [14, Equation (1.3) and §2]. We do not elaborate on this topic because it will not play a direct role in our article.

1.1. **Main results and outline of the article.** We begin in §2 with the analysis of the generalized Kimura stochastic differential equation, (1.3). Existence of solutions (Proposition 2.2) is an immediate consequence of classical results, and for this purpose the conditions imposed on the coefficients are more general, as outlined in Assumption 2.1. We establish uniqueness in law of solutions to the generalized Kimura stochastic differential equation in Proposition 2.4, under the more restrictive Assumption 2.6. We defer the exact technical statements of Assumptions 2.1 and 2.6 to §2.1 and §2.2.1, respectively, but we highlight here the main points. The drift coefficients $b(z)$ are only assumed to be nonnegative on the boundary of the domain $S_{n,m}$, and the coefficient functions $b(z)$, $e(z)$, and a suitable combination of the coefficients of the diffusion matrix are assumed to belong to the anisotropic Hölder spaces introduced in §2.2.1. This condition arises because our method of the proof is based on the existence, uniqueness, and regularity of solutions in anisotropic Hölder spaces to the homogeneous initial-value problem

$$\begin{aligned} u_t - \widehat{L}u &= 0 \quad \text{on } (0,T) \times S_{n,m}, \\ u(0,\cdot) &= \varphi \quad \text{on } S_{n,m}, \end{aligned} \tag{1.7}$$

where the operator $\widehat{L}$ is the generator of generalized Kimura diffusions. Regularity of solutions to parabolic equations defined by the infinitesimal generator of generalized Kimura diffusions are established in [12, 13, 29]. Our definition of the anisotropic Hölder spaces is an adaptation to our framework of the Hölder spaces introduced in [13, Chapter 5].

In §3, we prove our main results (Theorems 3.1 and 3.7) concerning the existence and uniqueness in law of weak solutions to the singular Kimura stochastic differential equation, (1.6). Our method of the proof consists in applying Girsanov's Theorem [23, Theorem 3.5.1] to the weak solutions of the generalized Kimura stochastic differential equation, (1.3), to change the probability distribution so that, under the new measure, the solutions solve the singular Kimura stochastic differential equation, (1.6). We justify the application of Girsanov's Theorem by proving that Novikov's condition [23, Corollary 3.5.13] holds, a fact that uses the Markov property of the processes we consider. Because Girsanov's Theorem is also used in the proof of uniqueness in law of weak solutions, our uniqueness result is established in the class of Markov processes. While this result is sufficient for the application to the proof of the Harnack inequality for nonnegative solutions to the parabolic Kimura equation, [14, Theorem 1.2], employing ideas used to prove [35, Theorem 12.2.4], it may be possible to prove that uniqueness in the class of Markov processes implies weak uniqueness. Notice though that [35, Theorem 12.2.4] does not apply directly to our framework because our drift coefficients are not necessarily bounded (see condition (3.5)). When the drift coefficients are bounded, that is, we consider the generalized Kimura stochastic differential equation (1.3), then we establish the weak uniqueness of solutions in Proposition 2.4.

The existence and uniqueness results for the singular Kimura equation (1.6) are established provided that Assumption 3.2 holds. We do not give the exact technical statement of Assumption 3.2 here, but we highlight the main aspects. The drift coefficient functions $b(z)$ are bounded from below on $\partial S_{n,m}$ by a positive constant b_0 (see condition (3.2)). This is a crucial ingredient in the verification of Novikov's condition in Lemmas 3.5 and 3.8. Also, the singular coefficients $h_{i,j}(x_j)$ are assumed to satisfy the growth assumption (3.5), where $q \in (0, q_0)$, and the positive constant q_0 depends on b_0, by identity (3.1).

1.2. **Comparison with previous research.** Similar processes were previously analyzed in the literature in connection with superprocesses, [2, 4], and Fleming-Viot processes arising in the dynamics of populations, [6–9]. The main applications of our results are to the study of diffusions arising in population genetics, [26, 27, 34], [18, §10.1], [24, §15.2.F], and to the study of regularity of solutions to degenerate parabolic equations, [11, 14, 15].

Our work is more closely related to [2, 4], and so we give a more detailed description of the similarities and differences between our results and the aforementioned articles. The main difference between the Kimura stochastic differential equations (1.3), and those considered in [2, 4] consist in the fact that we allow coordinates, $\{Y(t)\}_{t\geq 0}$, of the weak solutions whose dispersion coefficients are nonzero on the boundary of the domain $S_{n,m}$, and we do not require the drift coefficients to be bounded; instead we allow singularities in the drift component of the form $|x_i|^{-q}$, for $i = 1, \dots, n$, where the exponent q satisfies a suitable restriction given by inequality (3.5). In the sequel, we explain in more detail the differences between the work done in [2, 4] and our results.

In [2], the authors consider diffusions corresponding to the generator

$$\mathcal{L}u = \sum_{i=1}^{n} \left(x_i \gamma_i(x) u_{x_i x_i} + b_i(x) u_{x_i}\right),$$

where $x \in \mathbb{R}^n_+$, and $u \in C^2(\mathbb{R}^n_+)$. Under the assumption that the coefficients of the operator $\mathcal{L}$ are continuous functions on $\bar{\mathbb{R}}^n_+$, and that the drift coefficients are positive on $\partial\mathbb{R}^n_+$, it is proved in [2] that the martingale problem associated to the operator $\mathcal{L}$ has a unique solution. The method of the proof consists in proving L^2-estimates for the resolvent operators, employing a method of Krylov and Safonov to establish continuity of the resolvent operators [3, §V.7], and a localizing procedure due to Stroock and Varadhan [35, Theorem 6.6.1] to reduce the existence and uniqueness of solutions to a local problem. In §2, we recover and extend the results on existence and uniqueness in law of weak solutions obtained in [2], under the assumption that the coefficients of the operator $\mathcal{L}$ belong to the anisotropic Hölder spaces introduced in §2.2.1, and we allow the drift coefficient to be zero along the boundary of $\mathbb{R}^n_+$. Moreover, our method of the proof appears to be simpler, as we rely on existence and uniqueness of solutions in anisotropic Hölder spaces to homogeneous initial-value parabolic equations defined by the operator $\mathcal{L}$. These results were established in [10, 12, 13, 29].

In [4], the authors consider a more general class of generators of the form:

$$\mathcal{L}u = \sum_{i,j=1}^n \sqrt{x_i x_j}\gamma_{i,j}(x) u_{x_i x_j} + \sum_{i=1}^n b_i(x) u_{x_i},$$

where $x \in \mathbb{R}^n_+$, and $u \in C^2(\mathbb{R}^n_+)$. In this work, the coefficient functions $(\gamma(z))$ and $b(z)$ are assumed to belong to suitable weighted Hölder spaces, as opposed to the anisotropic Hölder spaces introduced in §2.2.1, and the drift coefficient $b(z)$ is assumed nonnegative on the boundary of the domain $\mathbb{R}^n_+$. The main difference between the weighted Hölder spaces used in [4, §2], and the anisotropic Hölder spaces used in [13, Equation (5.42)], and in our work, consists in the distance function used to define the Hölder seminorm. The weighted Hölder spaces defined in [4, §2], $C^\alpha_w(\bar{\mathbb{R}}^n_+)$, use the Euclidean distance function, while the anisotropic Hölder spaces defined in §2.2.1, $C^\alpha_{\mathrm{WF}}(\bar{S}_{n,m})$, use a distance function that is equivalent to the Riemannian metric with respect to which the second order terms of the generalized Kimura operator agree with the Laplace operator for a suitable Riemannian metric tensor. Another difference is that the weighted Hölder spaces in [4, §2] use a weight function to account for the degeneracy of the operator $\widehat{L}$, while we do not use a weight function in the definition of $C^\alpha_{\mathrm{WF}}(\bar{S}_{n,m})$, because the degeneracy of the operator is already encoded in the distance function.

The method of the proof of the results in [4] is based on establishing suitable estimates of the semigroup associated to the operator $\mathcal{L}$, and of the resolvent operators the weighted Hölder space $C^\alpha_w(\bar{\mathbb{R}}^n_+)$. These are combined with the localizing procedure of Stroock and Varadhan, [35, Theorem 6.6.1], to prove existence and uniqueness of solutions to the martingale problem associated to $\mathcal{L}$. Our results are both more general and more restrictive in certain ways, than the ones obtained in [4]. The smallness condition [4, Inequality (1.4)] on the cross-terms $\gamma_{i,j}(z)$, for $i \neq j$, of the operator $\mathcal{L}$ is less restrictive than our analogous condition (2.17) of the matrix $(a(z))$, defined in (2.11). On the other hand, we allow nondegenerate directions, $\{Y(t)\}_{t\geq 0}$, and singular, unbounded drift coefficients in our stochastic differential equation (1.6).

1.3. **Notation and conventions.** Let $N, M \in \mathbb{N}$, and U be a subset of $\mathbb{R}^N$. We let $C_{\mathrm{loc}}(U;\mathbb{R}^M)$ denote the space of functions, $u : U \to \mathbb{R}^M$, that are continuous on U, but are not necessarily bounded. If a function $u \in C_{\mathrm{loc}}(U;\mathbb{R}^M)$ is bounded, i.e.,

$$\|u\|_{C(U;\mathbb{R}^M)} := \sup_{x\in U} |u(x)| < \infty,$$

then we say that u belongs to $C(U;\mathbb{R}^M)$. When $M = 1$, we simply denote $C_{\mathrm{loc}}(U;\mathbb{R})$ by $C_{\mathrm{loc}}(U)$, and $C(U;\mathbb{R})$ by $C(U)$.

Let $U \subseteq \mathbb{R}^N$ be an open set, and $u : U \to \mathbb{R}$ be a function that admits first order derivatives on U. We denote the first order derivative of the function $u(x)$ in the i-th coordinate by u_{x_i}, and we denote the first order derivative operator in the i-th coordinate by ∂_{x_i}, for all $1 \leq i \leq N$. For a positive integer k, the space $C^k(U)$ consists of functions $u : U \to \mathbb{R}$ that admit continuous and bounded derivatives on U up to and including order k, and $C^\infty(U) := \bigcap_{k\geq 1} C^k(U)$.

Let now $U \subseteq \mathbb{R}^N$ be a closed set, such that U is the closure of $\operatorname{int}(U)$, where $\operatorname{int}(U)$ denotes the interior of U. The space $C^\infty(U)$ consists of functions $u : U \to \mathbb{R}$, such that the restriction $u \restriction_{\operatorname{int}(U)}$ has continuous derivatives of any order on $\operatorname{int}(U)$, and every partial derivative can be extended by continuity up to the boundary of the set U, and the extension is bounded. The space $C^\infty_c(U)$ consists of functions $u \in C^\infty(U)$ with the property that there is a compact set $K \subseteq U$, such that the support of the function u is contained in K.

For a Borel measurable set U, we denote by $\mathcal{B}(U)$ the collection of Borel measurable subsets of U. For all $r > 0$ and $z \in \mathbb{R}^N$, we denote by $B_r(z)$ the Euclidean ball of radius r centered at z. For $a, b \in \mathbb{R}$, we denote by $a \wedge b := \min\{a, b\}$.

2. Generalized Kimura diffusions

To establish existence, uniqueness, and the strong Markov property of weak solutions to the Kimura stochastic differential equation with singular drift (1.6), we first prove these results for the generalized Kimura diffusions, (1.3). We organize this section into three parts. In §2.1, we prove under suitable hypotheses (Assumption 2.1) that the generalized Kimura stochastic differential equation (1.3) admits weak solutions, $\{\widehat{Z}(t)\}_{t\geq 0}$, supported in $\bar{S}_{n,m}$, when the initial condition is assumed to satisfy $\widehat{Z}(0) \in \bar{S}_{n,m}$. In §2.2, we prove under more restrictive hypotheses (Assumption 2.6), that the weak solutions to the Kimura equation (1.3) are unique in law, and satisfy the strong Markov property.

2.1. Existence of weak solutions.

Existence of solutions to the generalized Kimura stochastic differential equation (1.3) can be established for a more general form of the diffusion matrix than the one implied by equations (1.3). For this reason, we consider the stochastic differential equation

$$
\begin{aligned}
d\widehat{X}_i(t) &= b_i(\widehat{Z}(t))\,dt + \sum_{k=1}^{n+m} \varsigma_{i,k}(\widehat{Z}(t))\,d\widehat{W}_k(t), \quad \forall\, t > 0, \\
d\widehat{Y}_l(t) &= e_l(\widehat{Z}(t))\,dt + \sum_{k=1}^{n+m} \varsigma_{n+l,k}(\widehat{Z}(t))\,d\widehat{W}_k(t), \quad \forall\, t > 0,
\end{aligned} \tag{2.1}
$$

where $i = 1, \ldots, n$, $l = 1, \ldots, m$, and $\{\widehat{W}(t)\}_{t\geq 0}$ is an $(n+m)$-dimensional Brownian motion.

Assumption 2.1 (Properties of the coefficients in (2.1)). *The coefficient functions of the stochastic differential equation* (2.1) *satisfy the properties:*

1. *We assume that* $b \in C_{\text{loc}}(\bar{S}_{n,m}; \mathbb{R}^n)$, $e \in C_{\text{loc}}(\bar{S}_{n,m}; \mathbb{R}^m)$, *and* $\varsigma \in C_{\text{loc}}(\bar{S}_{n,m}; \mathbb{R}^{(n+m)\times(n+m)})$.
2. *The coefficients* $b(z)$, $e(z)$ *and* $(\varsigma(z))$ *have at most linear growth in* $|z|$.
3. *We assume that*

$$
(\varsigma\varsigma^*)_{i,i}(z) = 0, \quad \forall z \in \partial S_{n,m} \cap \{x_i = 0\}, \quad \forall\, i = 1, \ldots, n, \tag{2.2}
$$

where ς^* *denotes the transpose matrix of* ς.

4. *The drift coefficients satisfy*

$$b_i(z) \geq 0, \quad \forall z \in \partial S_{n,m} \cap \{x_i = 0\}, \quad \forall\, i = 1, \ldots, n. \tag{2.3}$$

We begin with

Proposition 2.2 (Existence of weak solutions to generalized Kimura diffusions)**.** *Suppose that Assumption* 2.1 *holds. Then, for all $z \in \bar{S}_{n,m}$, there is a weak solution, $(\widehat{Z} = (\widehat{X}, \widehat{Y}), \widehat{W})$, on a filtered probability space satisfying the usual conditions, $(\Omega, \{\mathcal{F}(t)\}_{t\geq 0}, \mathcal{F}, \widehat{\mathbb{P}}^z)$, to the stochastic differential equation* (2.1)*, with initial condition $\widehat{Z}(0) = z$. Moreover, the weak solution, $\widehat{Z} = (\widehat{X}, \widehat{Y})$, is supported in $\bar{S}_{n,m}$.*

Proof. The method of the proof is similar to that of [19, Proposition 3.1 and Theorem 3.3]. We divide the proof into two steps. In Step 1, we continuously extend the coefficients of the stochastic differential equation (2.1) from $\bar{S}_{n,m}$ to $\mathbb{R}^{n+m}$, and we prove that the stochastic differential equation associated to the extended coefficients, (2.4), has a weak solution. In Step 2, we prove that any weak solution to equation (2.4) is supported in $\bar{S}_{n,m}$, when the support of the initial condition is contained in $\bar{S}_{n,m}$. Combining Steps 1 and 2, we obtain the existence of weak solutions supported in $\bar{S}_{n,m}$, to the stochastic differential equation (2.1).

Step 1 (Extension of the coefficients). By Assumption 2.1, we can extend the coefficients of the stochastic differential equation (2.1) by continuity from $\bar{S}_{n,m}$ to $\mathbb{R}^{n+m}$. We consider the function $\varphi : \mathbb{R}^{n+m} \to \bar{S}_{n,m}$ defined by

$$\varphi(z) = z', \text{ such that } z' \in \bar{S}_{n,m}, \text{ and } |z - z'| = \operatorname{dist}(z, \bar{S}_{n,m}).$$

Because $\bar{S}_{n,m}$ is a closed, convex set, the point $z' \in \bar{S}_{n,m}$ is uniquely determined for all $z \in \mathbb{R}^{n+m}$. Moreover φ is a continuous function and $\varphi(z) = z$, for all $z \in \bar{S}_{n,m}$. We define the coefficient functions $\tilde{b} := b \circ \varphi$, $\tilde{d} := d \circ \varphi$ and $\tilde{\varsigma} := \varsigma \circ \varphi$, which are continuous extensions to $\mathbb{R}^{n+m}$ of the coefficient functions b, d and σ, respectively. By Assumption 2.1, the extended coefficients are continuous functions on $\mathbb{R}^{n+m}$, and have at most linear growth in the spatial variable. We can thus apply [18, Theorem 5.3.10] to obtain that the stochastic differential equation

$$\begin{aligned} d\widetilde{X}_i(t) &= \tilde{b}_i(\widetilde{Z}(t))\, dt + \sum_{k=1}^{n+m} \tilde{\varsigma}_{i,k}(\widetilde{Z}(t))\, d\widetilde{W}_k(t), \quad \forall\, t > 0, \quad \forall\, i = 1, \ldots, n, \\ d\widetilde{Y}_l(t) &= \tilde{e}_l(\widetilde{Z}(t))\, dt + \sum_{k=1}^{n+m} \tilde{\varsigma}_{n+l;k}(\widetilde{Z}(t))\, d\widetilde{W}_k(t), \quad \forall\, t > 0, \quad \forall\, l = 1, \ldots, m, \end{aligned} \tag{2.4}$$

has a weak solution, $\{\widetilde{Z}(t) = (\widetilde{X}(t), \widetilde{Y}(t)), \widetilde{W}(t)\}_{t\geq 0}$, on a filtered probability space satisfying the usual conditions, $(\Omega, \{\mathcal{F}(t)\}_{t\geq 0}, \mathcal{F}, \widetilde{\mathbb{P}})$, for any initial condition, $\widetilde{Z}(0)$. The process $\{\widetilde{W}(t)\}_{t\geq 0}$ is an $(n+m)$-dimensional Brownian motion.

Step 2 (Support of weak solutions). Let $z \in \bar{S}_{n,m}$, and let $\{\widetilde{Z}(t) = (\widetilde{X}(t), \widetilde{Y}(t))\}_{t\geq 0}$ be a weak solution to the stochastic differential equation (2.4), with initial condition $\widetilde{Z}(0) = z$. Our goal is to show that

$$\widetilde{\mathbb{P}}^z \left(\widetilde{Z}(t) \in \bar{S}_{n,m} \right) = 1, \quad \forall\, t \geq 0, \tag{2.5}$$

where $\widetilde{\mathbb{P}}^z$ denotes the probability distribution of the process $\{\widetilde{Z}(t)\}_{t\geq 0}$, with initial condition $\widetilde{Z}(0)=z$. To prove identity (2.5), it is sufficient to show that

$$\widetilde{\mathbb{P}}^z\left(\widetilde{X}_i(t)\geq 0\right)=1,\quad \forall\, t\geq 0,\quad \forall\, i=1,\dots,n. \tag{2.6}$$

For $\varepsilon>0$, let $\eta_\varepsilon:\mathbb{R}\to[0,1]$ be a smooth function such that $\eta_\varepsilon(s)=1$ for $s\leq-\varepsilon$, $\eta_\varepsilon(s)=0$ for $s\geq 0$, and $\eta'_\varepsilon\leq 0$ on $\mathbb{R}$. We see that identity (2.6) holds, if we show that for all $\varepsilon>0$, we have that

$$\widetilde{\mathbb{P}}^z\left(\eta_\varepsilon\left(\widetilde{X}_i(t)\right)=0\right)=1,\quad \forall\, t\geq 0,\quad \forall\, i=1,\dots,n. \tag{2.7}$$

Applying Itô's rule [23, Theorem 3.3.6] to the process $\{\eta_\varepsilon(\widetilde{X}_i(t))\}_{t\geq 0}$, we obtain

$$\begin{aligned} d\eta_\varepsilon(\widetilde{X}_i(t)) &= \left(\tilde{b}_i(\widetilde{Z}(t))\eta'_\varepsilon\left(\widetilde{X}_i(t)\right)+\frac{1}{2}(\tilde{\varsigma}\tilde{\varsigma}^*)_{i,i}(\widetilde{Z}(t))\eta''_\varepsilon\left(\widetilde{X}_i(t)\right)\right)dt \\ &\quad + \eta'_\varepsilon(\widetilde{X}_i(t))\sum_{k=1}^{n+m}\widetilde{\zeta}_{i,k}(\widetilde{Z}(t))\,d\widetilde{W}_k(t), \end{aligned}$$

and integrating and taking expectation, it follows that

$$\begin{aligned} &\mathbb{E}_{\widetilde{\mathbb{P}}^z}\left[\eta_\varepsilon\left(\widetilde{X}_i(t)\right)\right] \\ &\quad = \eta_\varepsilon(z)+\mathbb{E}_{\widetilde{\mathbb{P}}^z}\left[\int_0^t\left(\tilde{b}_i(\widetilde{Z}(s))\eta'_\varepsilon\left(\widetilde{X}_i(s)\right)+\frac{1}{2}(\tilde{\varsigma}\tilde{\varsigma}^*)_{i,i}(\widetilde{Z}(s))\eta''_\varepsilon\left(\widetilde{X}_i(s)\right)\right)ds\right]. \end{aligned} \tag{2.8}$$

To justify the preceding formula, we note that because the coefficients $\widetilde{b}_i$ and $\widetilde{\zeta}_{i,k}(z)$ have at most linear growth in the spatial variable, and the derivatives η'_ε and η''_ε are bounded functions, there is a positive constant, C, such that

$$\begin{aligned} &\int_0^t\left(|\tilde{b}_i(\widetilde{Z}(s))||\eta'_\varepsilon\left(\widetilde{X}_i(s)\right)|+\frac{1}{2}|(\tilde{\varsigma}\tilde{\varsigma}^*)_{i,i}(\widetilde{Z}(s))||\eta''_\varepsilon\left(\widetilde{X}_i(s)\right)|\right)ds \\ &\qquad \leq C\int_0^t\left(1+|\widetilde{Z}(s)|^2\right)ds, \\ &\int_0^t|\eta'_\varepsilon(\widetilde{X}_i(s))|^2|\widetilde{\zeta}_{i,k}(\widetilde{Z}(s))|^2\,ds\leq C\int_0^t\left(1+|\widetilde{Z}(s)|^2\right)ds. \end{aligned}$$

Using again the linear growth of the coefficients, we can apply [23, Problem 5.3.15] to obtain that

$$\mathbb{E}_{\widetilde{\mathbb{P}}^z}\left[\int_0^t|\widetilde{Z}(s)|^2\,ds\right]<\infty.$$

The preceding three inequalities yield that the process

$$\int_0^t\eta'_\varepsilon(\widetilde{X}_i(s))\widetilde{\zeta}_{i,k}(\widetilde{Z}(s))\,d\widetilde{W}_k(s),\quad t\geq 0,$$

is a martingale, for all $k=1,\dots,n+m$, and the integral on the right-hand side of equality (2.8) is finite. This completes the justification of formula (2.8).

From condition (2.3), and the construction of the extended coefficient $\tilde{b}_i$, it follows that $\tilde{b}_i(z)$ is nonnegative on the support of the function η'_ε. Using the fact that $\eta'_\varepsilon\leq 0$, we obtain

$$\tilde{b}_i(\widetilde{Z}(s))\eta'_\varepsilon\left(\widetilde{X}_i(s)\right)\leq 0,\quad \forall\, s\in[0,t].$$

From condition (2.2), and the construction of the extended matrix coefficient $\tilde{\varsigma}$, it follows that $(\tilde{\varsigma}\tilde{\varsigma}^*)_{i,i} = 0$ on the support of η_ε''. Thus we have

$$(\tilde{\varsigma}\tilde{\varsigma}^*)_{i,i}(\widetilde{Z}(s))\eta_\varepsilon''\left(\widetilde{X}_i(s)\right) = 0, \quad \forall\, s \in [0,t].$$

Using now the fact that $\eta_\varepsilon(z) = 0$, when $z \in \bar{S}_{n,m}$, since $\eta_\varepsilon \equiv 0$ on $\mathbb{R}_+$, it follows from identity (2.8) that

$$\mathbb{E}_{\widetilde{\mathbb{P}}^z}\left[\eta_\varepsilon\left(\widetilde{X}_i(t)\right)\right] \leq 0,$$

and, because η_ε is a nonnegative function, the preceding expression holds with equality. Because $\varepsilon > 0$ was arbitrarily chosen, the preceding identity implies (2.6), for all $i = 1, \ldots, n$, and so, we conclude that (2.5) holds.

Identity (2.5) proves that, when started at points in $\bar{S}_{n,m}$, the weak solutions to the stochastic differential equation (2.4) remain in $\bar{S}_{n,m}$. Because the coefficients of the stochastic differential equations (1.3) and (2.4) agree on $\bar{S}_{n,m}$, we obtain that the weak solutions to (2.4) also solve equation (1.3). This completes the proof of Proposition 2.2. □

Remark 2.3 (Existence of weak solutions to the generalized Kimura equation). We now consider a Borel measurable matrix coefficient, $\sigma : \bar{S}_{n,m} \to \mathbb{R}^{n+m} \times \mathbb{R}^{n+m}$, such that by letting

$$\begin{aligned} \varsigma_{i,j}(z) &:= \sqrt{x_i}\sigma_{i,j}(z), \quad \forall\, i = 1, \ldots, n, \quad \forall\, j = 1, \ldots, n+m, \\ \varsigma_{i,j}(z) &:= \sigma_{i,j}(z), \quad \forall\, i = n+l, \ldots, m, \quad \forall\, j = 1, \ldots, n+m, \end{aligned} \tag{2.9}$$

the matrix $(\varsigma(z))$ verifies Assumption 2.1. Then Proposition 2.2 implies that there is a weak solution, $\{\widehat{Z}(t)\}_{t\geq 0}$, to the generalized Kimura stochastic differential equation (1.3), for any initial condition $\widehat{Z}(0)$ supported in $\bar{S}_{n,m}$, and that the solution remains supported in $\bar{S}_{n,m}$ at all subsequent times.

2.2. **Uniqueness and the strong Markov property.** In this section, we prove uniqueness in law, and the strong Markov property of weak solutions to the generalized Kimura stochastic differential equation (1.3). The main result of this section is

Proposition 2.4 (Uniqueness in law of weak solutions to (1.3)). *Suppose that the coefficients of the generalized Kimura stochastic differential equation* (1.3) *satisfy Assumption* 2.6. *Then, for all $z \in \bar{S}_{n,m}$, there is a unique weak solution, $(\widehat{Z} = (\widehat{X}, \widehat{Y}), \widehat{W})$, $(\Omega, \{\mathcal{F}(t)\}_{t\geq 0}, \mathcal{F}, \widehat{\mathbb{P}}^z)$, to the stochastic differential equation* (1.3), *satisfying the initial condition $\widehat{Z}(0) = z$.*

The exact statement of Assumption 2.6 is given in §2.2.1, where we also motivate in more detail the conditions we impose on the coefficients of the generalized Kimura stochastic differential equation (1.3).

From [23, Theorem 5.4.20], we obtain that uniqueness in law of weak solutions to the Kimura stochastic differential equation (1.3) implies that the solutions satisfy the strong Markov property. Thus we have the following corollary to Proposition 2.4.

Corollary 2.5 (The strong Markov property). *Suppose that the coefficients of the generalized Kimura stochastic differential equation* (1.3) *satisfy Assumption* 2.6. *For $z \in \bar{S}_{n,m}$, let $\{\widehat{Z}(t)\}_{t\geq 0}$ be the unique weak solution to the stochastic differential* *equation* (1.3)*, with initial condition* $\widehat{Z}(0) = z$*. Then the process* $\{\widehat{Z}(t)\}_{t\geq 0}$ *satisfies the strong Markov property.*

According to [23, Proposition 5.4.27], to prove uniqueness in law of weak solutions to the Kimura stochastic differential equation (1.3), it is sufficient to establish that for all $z \in \bar{S}_{n,m}$, any two weak solutions, $\{\widehat{Z}^i(t)\}_{t\geq 0}$, satisfying the property that $\widehat{Z}^i(0) = z$, for $i = 1, 2$, have the same one-dimensional marginal distributions. That is, for all functions, $\varphi \in C^\infty_c(\bar{S}_{n,m})$, and $T > 0$, we have that

$$\mathbb{E}_{\widehat{\mathbb{P}}^z_1}\left[\varphi(\widehat{Z}^1(T))\right] = \mathbb{E}_{\widehat{\mathbb{P}}^z_2}\left[\varphi(\widehat{Z}^2(T))\right], \tag{2.10}$$

where $\widehat{\mathbb{P}}^z_i$ denotes the probability distribution of the process $\{\widehat{Z}^i(t)\}_{t\geq 0}$, with initial condition $\widehat{Z}^i(0) = z$, for $i = 1, 2$. Property (2.10) follows if we prove that the initial-value problem (1.7) has a solution u with suitable regularity properties. We prove such a result in [29, Theorem 1.5], under the hypotheses that the generalized Kimura operator $\widehat{L}$ satisfies Assumption 2.6.

We organize this section as follows. In §2.2.1, we describe the Assumption 2.6, which we impose on the coefficients of the generalized Kimura operator $\widehat{L}$ to ensure that we can apply [29, Theorem 1.5]. In §2.2.2, we give the proof of Proposition 2.4.

2.2.1. *Assumptions on the coefficients of the Kimura operator* $\widehat{L}$. We begin by introducing the differential operator $\widehat{L}$, which will be the infinitesimal generator of the Markov solutions to the Kimura stochastic differential equation. We let

$$a(z) := \frac{1}{2}\sigma(z)\sigma^*(z), \quad \forall\, z \in \bar{S}_{n,m}, \tag{2.11}$$

and we define

$$\begin{aligned}\widehat{L}u = &\sum_{i,j=1}^n \sqrt{x_ix_j}a_{i,j}(z)u_{x_ix_j} + 2\sum_{i=1}^n\sum_{l=1}^m \sqrt{x_i}a_{i,n+l}(z)u_{x_iy_l} + \sum_{l,k=1}^m a_{n+l,n+k}(z)u_{y_ly_k}\\ &+ \sum_{i=1}^n b_i(z)u_{x_i} + \sum_{l=1}^m e_l(z)u_{y_l},\end{aligned} \tag{2.12}$$

for all $z \in S_{n,m}$, and all $u \in C^2(S_{n,m})$. Our goal is to impose conditions on the coefficients of the operator $\widehat{L}$ to ensure that it has the structure of the generalized Kimura operator defined in (1.2), and that the coefficients are regular enough to be able to apply [29, Theorem 1.5]. For this purpose, we first need to introduce a class of anisotropic Hölder spaces adapted to the degeneracy of the diffusion matrix. The following Hölder spaces are a slight modification of the Hölder spaces introduced by C. Epstein and R. Mazzeo in their study of the existence, uniqueness and regularity of solutions to the parabolic problem defined by Kimura operators, [12, 13]. Following [13, Chapter 5], we first introduce the distance function

$$\rho((t^0, z^0), (t, z)) := \rho_0(z^0, z) + \sqrt{|t^0 - t|}, \quad \forall\, (t^0, z^0), (t, z) \in [0, \infty) \times \bar{S}_{n,m}, \tag{2.13}$$

where ρ_0 is a distance function in the spatial variables z^0 and z. Because our domain $S_{n,m}$ is unbounded, as opposed to the compact manifolds with corners considered in [13], the properties of the distance function $\rho_0(z^0, z)$ depend on whether the points z^0 and z are in a neighborhood of the boundary of $S_{n,m}$, or far away from

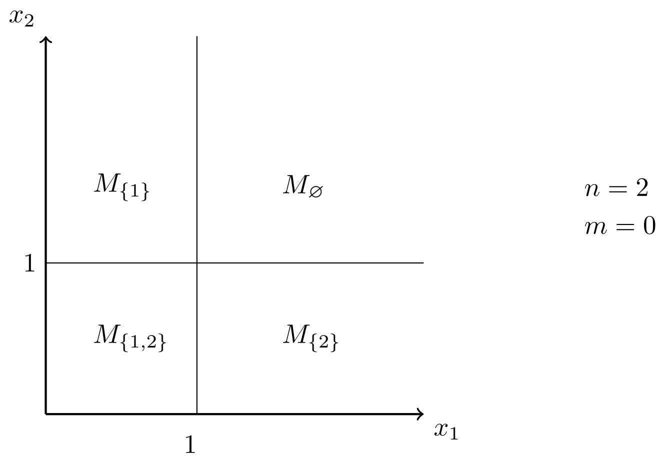

FIGURE 1. Sets M_I when $n = 2$ and $m = 0$.

the boundary of $S_{n,m}$. In a neighborhood of the boundary, for concreteness, when z^0 and z are in a neighborhood of the origin, the distance function $\rho_0(z^0, z)$ is equivalent to the Riemannian metric such that the model operator,

$$\widetilde{L}u := \sum_{i=1}^{n} x_i u_{x_i x_i} + \sum_{l=1}^{m} u_{y_l y_l},$$

agrees with the Laplacian for a suitable Riemannian metric tensor; see [13, §2.2] for a detailed explanation. In other words, in a neighborhood of the origin the distance function $\rho_0(z^0, z)$ is adapted to the degeneracy of the Kimura operator $\widehat{L}$. When z^0 and z are away from the boundary of the domain $S_{n,m}$, the operator $\widehat{L}$ is strictly elliptic, and we choose the distance function $\rho_0(z^0, z)$ such that it is equivalent to the Euclidean distance. To be more specific, we introduce the following notation: For any set of indices, $I \subseteq \{1, \dots, n\}$, we let

$$M_I := \{z = (x, y) \in S_{n,m} : x_i \in (0,1) \text{ for all } i \in I, \text{ and } x_j \in (1, \infty) \text{ for all } j \in I^c\}, \tag{2.14}$$

where we denote $I^c := \{1, \dots, n\} \backslash I$; see Figure 1.

We choose the distance function ρ_0 with the property that there is a positive constant, $c = c(n, m)$, such that for all subsets $I, J \subseteq \{1, \dots, n\}$, and all points $z^0 \in \bar{M}_I$ and $z \in \bar{M}_J$, we have that

$$\begin{aligned} &c\left(\max_{i \in I \cap J} \left|\sqrt{x_i^0} - \sqrt{x_i}\right| + \max_{j \in (I \cap J)^c} |x_j^0 - x_j| + \max_{1 \le l \le m} |y_l^0 - y_l|\right) \\ &\le \rho_0(z^0, z) \\ &\le c^{-1}\left(\max_{i \in I \cap J} \left|\sqrt{x_i^0} - \sqrt{x_i}\right| + \max_{j \in (I \cap J)^c} |x_j^0 - x_j| + \max_{1 \le l \le m} |y_l^0 - y_l|\right). \end{aligned} \tag{2.15}$$

We can construct the distance function $\rho_0(z^0, z)$ in the following way. We start with the particular case when $n = 1$ and $m = 0$, we denote $z^0 = x^0 \in [0, \infty)$ and

$z = x \in [0,\infty)$, and we set

$$\rho_0(x^0,x) := \begin{cases} \left|\sqrt{x^0}-\sqrt{x}\right|, & \text{if } x^0, x \in [0,1),\\ |x^0-x|, & \text{if } x^0,x\in[1,\infty),\\ \left|\sqrt{x^0}-1\right| + |1-x|, & \text{if } x^0\in[0,1) \text{ and } x\in[1,\infty),\\ |\sqrt{x}-1|+|1-x^0|, & \text{if } x\in[0,1) \text{ and } x^0\in[1,\infty). \end{cases}$$

It is an easy exercise to check that ρ_0, defined as above on $[0,\infty)$, is a distance function, and that the following inequalities hold:

(2.16)
$$\begin{array}{llll} \left|\sqrt{x^0}-\sqrt{x}\right| & \leq \rho_0(x^0,x) & \leq \left|\sqrt{x^0}-\sqrt{x}\right|, & \text{if } x^0,x\in[0,1),\\ |x^0-x| & \leq \rho_0(x^0,x) & \leq |x^0-x|, & \text{if } x^0,x\in[1,\infty),\\ \frac{1}{2}|x^0-x| & \leq \rho_0(x^0,x) & \leq |x^0-x|, & \text{if } x^0\in[0,1) \text{ and } x\in[1,\infty). \end{array}$$

We now turn to the more general case, when n and m are nonnegative integers, and for all subsets $I, J \subseteq \{1,\dots,n\}$, and all points $z^0 \in \bar{M}_I$ and $z\in\bar{M}_J$, we set

$$\rho_0(z^0,z) := \max_{1\leq i\leq n}\rho_0(x_i^0,x_i) + \max_{1\leq l\leq m}|y_l^0-y_l|.$$

Because $\rho_0(x^0,x)$ is a distance function for all $x^0,x\in[0,\infty)$ which satisfies properties (2.16), we conclude that $\rho_0(z^0,z)$ is a distance function for all $z^0,z\in\bar{S}_{n,m}$, and satisfies inequalities (2.15), with $c=1/2$.

Let $\alpha\in(0,1)$. Following [13, §5.2.4], we let $C^\alpha_{\mathrm{WF}}([0,T]\times\bar{S}_{n,m})$ be the Hölder space consisting of continuous functions, $u:[0,T]\times\bar{S}_{n,m}\to\mathbb{R}$, such that the following norm is finite:

$$\begin{aligned}\|u\|_{C^\alpha_{\mathrm{WF}}([0,T]\times\bar{S}_{n,m})} := &\sup_{(t,z)\in[0,T]\times\bar{S}_{n,m}}|u(t,z)|\\ &+\sup_{\substack{(t^0,z^0),(t,z)\in[0,T]\times\bar{S}_{n,m}\\(t^0,z^0)\neq(t,z)}}\frac{|u(t^0,z^0)-u(t,z)|}{\rho^\alpha((t^0,z^0),(t,z))}.\end{aligned}$$

For a set $U\subset\bar{S}_{n,m}$, we denote by $C^\alpha_{\mathrm{WF}}(U)$ the anisotropic Hölder space consisting of functions $u:U\to\mathbb{R}$, such that

$$\|u\|_{C^\alpha_{\mathrm{WF}}(U)} := \sup_{z\in U}|u(z)| + \sup_{\substack{z^0,z\in U\\ z^0\neq z}}\frac{|u(z^0)-u(z)|}{\rho_0^\alpha(z^0,z)} < \infty.$$

We use the subscript "WF" in the definition of the anisotropic Hölder spaces $C^\alpha_{\mathrm{WF}}([0,T]\times\bar{S}_{n,m})$ and $C^\alpha_{\mathrm{WF}}(U)$, to distinguish them from the classical Hölder spaces, as defined in [28, Chapter 8], and to indicate their relation to the Wright-Fisher operator, (1.1), and, more generally, to the generalized Kimura operators, (1.2).

We can now introduce the assumptions on the coefficients of the Kimura stochastic differential equation (1.6).

Assumption 2.6 (Properties of the coefficients in (1.3))**.** *The coefficient functions of the stochastic differential equation* (1.6) *satisfy the properties: Let* $\alpha\in(0,1)$, *and assume that*

1. *The coefficient functions* $b_i(z)$ *satisfy the* nonnegativity *condition* (2.3)*, for all* $i=1,\dots,n$.

2. *For all* $i, j = 1, \dots, n$ *such that* $i \neq j$, *and all* $l = 1, \dots, m$, *there are functions,* $\alpha_{i,i}, \tilde{\alpha}_{i,j}, c_{i,l} : \bar{S}_{n,m} \to \mathbb{R}$, *such that*

$$
\begin{aligned}
a_{i,j}(z) &= \delta_{i,j}\alpha_{i,i}(z) + \sqrt{x_i x_j}\tilde{\alpha}_{i,j}(z), \quad \forall\, z = (x,y) \in \bar{S}_{n,m}, \\
a_{i,n+l}(z) = a_{n+l,i}(z) &= \frac{1}{2}\sqrt{x_i} c_{i,l}(z), \quad \forall\, z = (x,y) \in \bar{S}_{n,m},
\end{aligned} \tag{2.17}
$$

where we recall that $\delta_{i,j}$ *denotes the Kronecker delta symbol.*

3. *The* strict ellipticity *condition holds: there is a positive constant,* λ, *such that for all sets of indices,* $I \subseteq \{1, \dots, n\}$, *for all* $z \in \bar{M}_I$, $\xi \in \mathbb{R}^n$ *and* $\eta \in \mathbb{R}^m$, *we have*

$$
\begin{aligned}
&\sum_{i \in I} \alpha_{i,i}(z)\xi_i^2 + \sum_{i \in I^c} x_i \alpha_{i,i}(z)\xi_i^2 + \sum_{i,j \in I} \tilde{\alpha}_{i,j}(z)\xi_i\xi_j \\
&+ \sum_{i \in I}\sum_{j \in I^c} x_j(\tilde{\alpha}_{i,j}(z) + \tilde{\alpha}_{j,i}(z))\xi_i\xi_j + \sum_{i,j \in I^c} x_i x_j \tilde{\alpha}_{i,j}(z)\xi_i\xi_j \\
&+ \sum_{i \in I}\sum_{l=1}^{m} c_{i,l}(z)\xi_i\eta_l + \sum_{i \in I^c}\sum_{l=1}^{m} x_i c_{i,l}(z)\xi_i\eta_l + \sum_{k,l=1}^{m} a_{n+k,n+l}(z)\eta_k\eta_l \\
&\geq \lambda\left(|\xi|^2 + |\eta|^2\right).
\end{aligned} \tag{2.18}
$$

4. *The coefficient functions are* Hölder continuous*: for all sets of indices,* $I \subseteq \{1, \dots, n\}$, *and for all* $i, i' \in I$, $j, j' \in I^c$, *and* $l, k = 1, \dots, m$, *we have that*

$$
\begin{aligned}
\alpha_{i,i}, \quad x_j\alpha_{j,j}, \quad \tilde{\alpha}_{i,i'}, \quad x_j\tilde{\alpha}_{i,j}, \quad x_j\tilde{\alpha}_{j,i}, \quad x_j x_{j'}\tilde{\alpha}_{j,j'} \quad &\in C^{\alpha}_{\mathrm{WF}}(\bar{M}_I), \\
a_{n+k,n+l}, \quad b_i, \quad b_j, \quad c_{i,l}, \quad x_j c_{j,l}, \quad e_l \quad &\in C^{\alpha}_{\mathrm{WF}}(\bar{M}_I).
\end{aligned} \tag{2.19}
$$

Remark 2.7 (Structure of the operator $\widehat{L}$). Condition (2.17) implies that the differential operator $\widehat{L}$ takes the form:

$$
\begin{aligned}
\widehat{L}u &= \sum_{i=1}^{n} (x_i\alpha_{i,i}(z)u_{x_ix_i} + b_i(z)u_{x_i}) + \sum_{i,j=1}^{n} x_i x_j \tilde{\alpha}_{i,j}(z)u_{x_ix_j} \\
&\quad + \sum_{i=1}^{n}\sum_{l=1}^{m} x_i c_{i,l}(z)u_{x_iy_l} + \sum_{k,l=1}^{m} a_{n+k,n+l}(z)u_{y_ky_l} + \sum_{l=1}^{m} e_l(z)u_{y_l},
\end{aligned}
$$

that is, it has the same structure as the generalized Kimura operator defined in (1.2). Assumption 2.6 ensures that the operator $\widehat{L}$ satisfies the hypotheses of [29, Theorem 1.4], which allows us to establish (2.10), and which in turn implies the conclusion of Proposition 2.4. See the proof of Proposition 2.4.

Assumption 2.6 yields some immediate boundedness conditions on the coefficients of the Kimura stochastic differential equation (1.3), which we often use in the sequel.

Lemma 2.8 (Boundedness of the coefficient functions $b(z)$ and $(\sigma(z))$)**.** *Suppose that Assumption* 2.6 *holds. Then there is a positive constant,* K, *such that for all* $i = 1, \dots, n$, *and all* $j, l = 1, \dots, n + m$, *we have that*

$$
|b_i(z)| + |\sigma_{j,l}(z)| \leq K, \quad \forall\, z \in \bar{S}_{n,m}. \tag{2.20}
$$

Proof. The boundedness of the coefficients $b(z)$ is obvious from (2.19), and the definition of the anisotropic Hölder space $C^{\alpha}_{\mathrm{WF}}(\bar{S}_{n,m})$. The boundedness of the coefficients of the matrix $(\sigma(z))$ follows from identity (2.11), and the fact that the matrix $a(z)$ is bounded, as it is implied by identity (2.17), and condition (2.19). $\square$

Lemma 2.9 (Boundedness of the matrix coefficient $(\varsigma(z))$). *Suppose that Assumption* 2.6 *holds. Then the coefficients of the matrix* $(\varsigma(z))$*, defined in* (2.9)*, are bounded.*

Proof. Let

$$D(z) := \varsigma(z)\varsigma^*(z), \quad \forall\, z \in \bar{S}_{n,m}, \tag{2.21}$$

be the diffusion matrix of the Kimura stochastic differential equation (1.3). Using (2.9), (2.11), and (2.17), it follows that, for all $i,j = 1,\dots,n$, and all $k,l = 1,\dots,m$, we have that

$$\begin{aligned} D_{i,j}(z) &= 2\left(\delta_{i,j}x_i\alpha_{i,i}(z) + x_ix_j\tilde{\alpha}_{i,j}(z)\right), \\ D_{i,n+k}(z) = D_{n+k,i}(z) &= x_ic_{i,k}(z), \\ D_{n+k,n+l}(z) &= 2a_{n+k,n+l}(z). \end{aligned} \tag{2.22}$$

By the boundedness of the coefficients implied by condition (2.19), it follows that the coefficient matrix $(D(z))$ is bounded, and so, identity (2.21) implies that the coefficient matrix $(\varsigma(z))$ is also bounded. $\square$

2.2.2. *Proof of Proposition* 2.4. In the proof of the uniqueness in law of weak solutions to the Kimura stochastic differential equation, we make use of the following version of Itô's rule, which is more general in that the function does not have to be $C^{1,2}$ on the support of the process.

Proposition 2.10 (Itô's rule for generalized Kimura diffusions). *Suppose that the coefficients* $b(z)$*,* $e(z)$*, and* $(\varsigma(z))$*, appearing in* (1.3) *and* (2.9)*, are Borel measurable functions, with at most linear growth in the spatial variable. Assume that the matrix coefficient* $(a(z))$*, defined in* (2.11)*, is bounded on compact sets in* $\bar{S}_{n,m}$*. Let* $u \in C_{\text{loc}}([0,\infty)\times\bar{S}_{n,m})$ *be such that for all* $i,j = 1,\dots,n$*, and* $l,k = 1,\dots,m$*, we have that*

$$u_t,\, u_{x_i},\, u_{y_l},\, \sqrt{x_ix_j}u_{x_ix_j},\, \sqrt{x_i}u_{x_iy_l},\, u_{y_l,y_k} \in C_{\text{loc}}([0,\infty)\times\bar{S}_{n,m}). \tag{2.23}$$

Let $z \in \bar{S}_{n,m}$*, and let* $(\widehat{Z},\widehat{W})$ *be a weak solution to the generalized Kimura stochastic differential equation* (1.3)*, with initial condition* $\widehat{Z}(0) = z$*, defined on a filtered probability space,* $(\Omega, \{\mathcal{F}(t)\}_{t\geq 0}, \mathcal{F}, \widehat{\mathbb{P}}^z)$*. Then, the following holds* $\widehat{\mathbb{P}}^z$*-a.s., for all* $t \geq 0$*:*

$$\begin{aligned} u(t,\widehat{Z}(t)) &= \int_0^t (u_s + \widehat{L}u)(s,\widehat{Z}(s))\,ds \\ &\quad + \int_0^t \sum_{j=1}^n\sum_{k=1}^{n+m} \varsigma_{j,k}(\widehat{Z}(s))u_{x_j}(s,\widehat{Z}(s))\,d\widehat{W}_k(s) \\ &\quad + \int_0^t \sum_{l=1}^m\sum_{k=1}^{n+m} \varsigma_{n+l,k}(\widehat{Z}(s))u_{y_l}(s,\widehat{Z}(s))\,d\widehat{W}_k(s), \end{aligned} \tag{2.24}$$

where we recall that the matrix coefficient $(\varsigma(z))$ *is defined in* (2.9)*.*

Remark 2.11 (Condition (2.23)). In Proposition 2.10, we assume that the function u belongs to $C_{\text{loc}}([0,\infty)\times\bar{S}_{n,m})$, and so the distributional derivatives of any order of the function u are well defined on $(0,\infty)\times S_{n,m}$. In condition (2.23), we assume that the restrictions of the distributions $u_t, u_{x_i}, u_{y_l}, \sqrt{x_ix_j}u_{x_ix_j}, \sqrt{x_i}u_{x_iy_l}, u_{y_l,y_k}$ to the set $(0,\infty)\times S_{n,m}$ are continuous functions, and that they admit continuous extensions up to the boundary of the domain $(0,\infty)\times S_{n,m}$.

In the proof of Proposition 2.10, we need the following.

Lemma 2.12 (Property of the weighted derivatives). *Suppose that* $u\in C_{\text{loc}}(\bar{S}_{n,m})$ *is such that for all* $i,j=1,\dots,n$, *and* $l=1,\dots,m$,

$$u_{x_i}, \sqrt{x_ix_j}u_{x_ix_j}, \sqrt{x_i}u_{x_iy_l}\in C_{\text{loc}}(\bar{S}_{n,m}). \tag{2.25}$$

Then the function u *has the property that*

$$\sqrt{x_ix_j}u_{x_ix_j}=0 \quad \text{on } \partial S_{n,m}\cap\{x_i=0 \text{ or } x_j=0\}, \tag{2.26}$$

$$\sqrt{x_i}u_{x_iy_l}=0 \quad \text{on } \partial S_{n,m}\cap\{x_i=0\}. \tag{2.27}$$

Before we give the proof of Lemma 2.12, we explain condition (2.25), and equalities (2.26), and (2.27).

Remark 2.13 (Conditions (2.25), (2.26), and (2.27)). Similarly to Remark 2.11, because we assume that the function u belongs to $C_{\text{loc}}(\bar{S}_{n,m})$, the distributional derivatives of any order of the function u are well defined on $S_{n,m}$. In condition (2.25), we assume that the restrictions of the distributions $u_{x_i}, \sqrt{x_ix_j}u_{x_ix_j}, \sqrt{x_i}u_{x_iy_l}$ to the set $S_{n,m}$ are continuous functions, and that they admit continuous extensions up to the boundary of the domain $S_{n,m}$. Equality (2.26) should be understood as

$$\lim_{\substack{z\to z^0\\ z\in S_{n,m}}}\sqrt{x_ix_j}u_{x_ix_j}(z)=0,$$

for all $z^0\in\partial S_{n,m}\cap\{x_i=0 \text{ or } x_j=0\}$, and equality (2.27) should be understood as

$$\lim_{\substack{z\to z^0\\ z\in S_{n,m}}}\sqrt{x_i}u_{x_iy_l}(z)=0,$$

for all $z^0\in\partial S_{n,m}\cap\{x_i=0\}$.

Proof of Lemma 2.12. We prove assertions (2.26) and (2.27) by contradiction. We consider the following cases.

Case 1 (Derivatives $x_iu_{x_ix_i}$). Assuming that the weighted derivative $x_iu_{x_ix_i}$ does not satisfy property (2.26), using the continuity of $x_iu_{x_ix_i}$ on $\bar{S}_{n,m}$ by (2.25), we can assume without loss of generality that there is a point, $z^0\in\partial S_{n,m}\cap\{x_i=0\}$, and there are positive constants, a and ε, such that $x_iu_{x_ix_i}(z)\geq a$, for all points $z\in B_\varepsilon(z^0)\cap S_{n,m}=: U_\varepsilon$. Choosing points $z^k\in U_\varepsilon$, $k=1,2$, such that all coordinates are identical to those of z^0, except for the i-th coordinate, which is x_i^k, $k=1,2$, and integrating in the x_i-variable the inequality $u_{x_ix_i}(z)\geq a/x_i$, we obtain that $u_{x_i}(z^1)-u_{x_i}(z^2)\geq a\ln(x_i^1/x_i^2)$. Letting x_i^2 converge to 0, we see that the right-hand side of the preceding inequality diverges to $+\infty$, while the left-hand side is finite by assumption (2.25). This contradiction implies that property (2.26) holds for the weighted derivative $x_iu_{x_ix_i}$.

Case 2 (Derivatives $\sqrt{x_i x_j} u_{x_i x_j}$, with $i \neq j$). We proceed similarly to Case 1. Using property (2.25), we assume without loss of generality that there is $z^0 \in \partial S_{n,m} \cap \{x_i = 0 \text{ or } x_j = 0\}$, and there are positive constants, a and ε, such that $\sqrt{x_i x_j} u_{x_i x_j}(z) \geq a$, for all points $z \in B_\varepsilon(z^0) \cap S_{n,m} =: U_\varepsilon$, where $i \neq j$, and $i, j = 1, \ldots, n$. To fix ideas we assume that the j-th coordinate of the point z^0 is 0. Choosing points $z^k \in U_\varepsilon$, $k = 1, 2$, such that all coordinates are identical to those of z^0, except for the i-th coordinate, which is x_i^k, $k = 1, 2$, and integrating in the x_i-variable the inequality $u_{x_i x_j}(z) \geq a/\sqrt{x_i x_j}$, we obtain that

$$u_{x_j}(z^1) - u_{x_j}(z^2) \geq \frac{2a}{\sqrt{x_j}} \left(\sqrt{x_i^1} - \sqrt{x_i^2} \right). \tag{2.28}$$

We can choose the points $z^1, z^2 \in U_\varepsilon$, such that there is a positive constant, c, with the property that

$$\sqrt{x_i^1} - \sqrt{x_i^2} \geq c. \tag{2.29}$$

Letting x_j converge to 0 in inequality (2.28), while the lower bound (2.29) is satisfied, we see that the right-hand side of (2.28) diverges to $+\infty$, while the left-hand side is finite by assumption (2.25). This contradiction implies that property (2.26) holds for the weighted derivative $\sqrt{x_i x_j} u_{x_i x_j}$, where $i \neq j$.

Case 3 (Derivatives $\sqrt{x_i} u_{x_i y_l}$). The proof of property (2.27) can be done using the argument of Case 2 to prove (2.26) for the weighted derivative $\sqrt{x_i x_j} u_{x_i x_j}$, where $i \neq j$, with the observation that instead of integrating in the x_i-variable, we integrate in the y_l-variable.

Combining Cases 1, 2 and 3, we obtain the conclusion of the lemma. □

Proof of Proposition 2.10. The method of the proof of (2.24) is very similar to that of [19, Proposition 3.5], but we include it for completeness. We choose $\varepsilon \geq 0$, and let

$$z^\varepsilon := (x_1 + \varepsilon, \ldots, x_n + \varepsilon, y), \quad \forall\, z = (x, y) \in \bar{S}_{n,m},$$
$$\widehat{Z}^\varepsilon(t) := \left(\widehat{X}_1(t) + \varepsilon, \ldots, \widehat{X}_n(t) + \varepsilon, \widehat{Y}(t) \right), \quad \forall\, t \geq 0.$$

The proof follows by applying the standard Itô's formula, [23, Theorem 3.3.6], to the processes $\{\widehat{Z}^\varepsilon(t)\}_{t \geq 0}$, and taking limit as ε tends to zero. This will require the use of condition (2.23). For all $N \in \mathbb{N}$, let τ_N be the first time the process $\{\widehat{Z}(t)\}_{t \geq 0}$ exits the Euclidean ball of radius N centered at the origin. Because the coefficients $b(z)$, $e(z)$, and $(\varsigma(z))$ appearing in (1.3) and (2.9) have at most linear growth in the spatial variable, we obtain by [23, Problem 5.3.15], that for all $M \geq 1$ and $t \geq 0$, there is a positive constant, $C = C(M, t)$, such that

$$\mathbb{E}_{\widehat{\mathbb{P}}^z} \left[\max_{s \in [0,t]} |\widehat{Z}(s)|^{2M} \right] \leq C \left(1 + |z|^{2M} \right). \tag{2.30}$$

Then it follows by (2.30) that the nondecreasing sequence of stopping times $\{\tau_N\}_{N \geq 1}$ satisfies

$$\lim_{N \to \infty} \tau_N = +\infty \quad \widehat{\mathbb{P}}^z\text{-a.s.} \tag{2.31}$$

If this were not the case, then there would be a deterministic time $t > 0$ such that

$$\lim_{N \to \infty} \widehat{\mathbb{P}}^z \left(\tau_N \leq t \right) > 0. \tag{2.32}$$

But, $\widehat{\mathbb{P}}^z(\tau_N \le t) = \widehat{\mathbb{P}}^z\left(\sup_{s\in[0,t]} |\widehat{Z}(s)| \ge N\right)$, and we have

$$\widehat{\mathbb{P}}^z\left(\sup_{s\in[0,t]} |\widehat{Z}(s)| \ge N\right) \le \frac{1}{N^2}\mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\max_{s\in[0,t]} |\widehat{Z}(s)|^2\right] \le \frac{C(1+|z|^2)}{N^2} \quad \text{(by (2.30))}.$$

Since the preceding expression converges to zero, as N goes to ∞, we obtain a contradiction in (2.32), and so (2.31) holds. By (2.31), it suffices to prove (2.24) for the stopped process, that is,

$$\begin{aligned} u(t\wedge\tau_N, \widehat{Z}(t\wedge\tau_N)) &= \int_0^{t\wedge\tau_N} (u_s + \widehat{L}u)(s, \widehat{Z}(s))\,ds \\ &\quad + \int_0^{t\wedge\tau_N} \sum_{j=1}^n \sum_{k=1}^{n+m} \varsigma_{j,k}(\widehat{Z}(s)) u_{x_j}(s, \widehat{Z}(s))\, d\widehat{W}_k(s) \\ &\quad + \int_0^{t\wedge\tau_N} \sum_{l=1}^m \sum_{k=1}^{n+m} \varsigma_{n+l,k}(\widehat{Z}(s)) u_{y_l}(s, \widehat{Z}(s))\, d\widehat{W}_k(s). \end{aligned} \tag{2.33}$$

Because the process $\{\widehat{Z}^\varepsilon(t)\}_{t\ge0}$ is supported in $[\varepsilon,\infty)^n \times \mathbb{R}^m \subset S_{n,m}$, and the function u belongs to $C^{1,2}([0,\infty)\times S_{n,m})$, we can apply the standard Itô's formula, [23, Theorem 3.3.6], to obtain

$$\begin{aligned} &u(t\wedge\tau_N, \widehat{Z}^\varepsilon(t\wedge\tau_N)) = u(0, \widehat{Z}^\varepsilon(0)) \\ &\quad + \int_0^{t\wedge\tau_N} \sum_{k=1}^{n+m} \sum_{j=1}^n \varsigma_{j,k}(\widehat{Z}(s)) u_{x_j}(s, \widehat{Z}^\varepsilon(s))\, d\widehat{W}_k(s) \\ &\quad + \int_0^{t\wedge\tau_N} \sum_{l=1}^m \sum_{k=1}^{n+m} \varsigma_{n+l,k}(\widehat{Z}(s)) u_{y_l}(s, \widehat{Z}^\varepsilon(s))\, d\widehat{W}_k(s) \\ &\quad + \int_0^{t\wedge\tau_N} \Bigg(u_s(s, \widehat{Z}^\varepsilon(s)) + \sum_{i=1}^n b_i(\widehat{Z}(s)) u_{x_i}(s, \widehat{Z}^\varepsilon(s)) + \sum_{l=1}^m e_l(\widehat{Z}(s)) u_{y_l}(s, \widehat{Z}^\varepsilon(s)) \\ &\qquad + \frac{1}{2}\sum_{i,j=1}^n \sqrt{\widehat{X}_i(s)\widehat{X}_j(s)}\, a_{i,j}(\widehat{Z}(s)) u_{x_i x_j}(s, \widehat{Z}^\varepsilon(s)) \\ &\qquad + \sum_{i=1}^n \sum_{l=1}^m \sqrt{\widehat{X}_i(s)}\, a_{i,n+l}(\widehat{Z}(s)) u_{x_i y_l}(s, \widehat{Z}^\varepsilon(s)) \\ &\qquad + \frac{1}{2}\sum_{l,k=1}^m a_{n+l,n+k}(\widehat{Z}(s)) u_{y_l y_k}(s, \widehat{Z}^\varepsilon(s)) \Bigg)\, ds. \end{aligned} \tag{2.34}$$

Our goal is to show that, by taking the limit as $\varepsilon \downarrow 0$, the left-hand and the right-hand side in (2.34) converge in probability to the corresponding expressions in (2.33).

Because the function u belongs to $C_{\mathrm{loc}}([0,\infty)\times \bar{S}_{n,m})$, we have for all $s\in[0,t]$,

$$u(s\wedge\tau_n, \widehat{Z}^\varepsilon(s\wedge\tau_N)) \to u(s\wedge\tau_N, \widehat{Z}(s\wedge\tau_N)) \quad \widehat{\mathbb{P}}^z\text{-a.s. when } \varepsilon\downarrow 0. \tag{2.35}$$

The terms in (2.34) containing the pure Itô integrals can be evaluated in the following way. We describe the details for the terms involving the derivatives in the x_i-variable, as the terms containing the derivatives in the y_l-variable can be treated similarly. We define, for all $i = 1 \dots, n$, $k = 1, \dots, n+m$, and $\varepsilon \geq 0$,

$$H^{\varepsilon}_{i,k}(s) := \varsigma_{i,k}(\widehat{Z}(s)) u_{x_i}(s, \widehat{Z}^{\varepsilon}(s)) \mathbf{1}_{\{|\widehat{Z}(s)| \leq N\}}.$$

Because we assume that the derivative u_{x_i} is continuous on $[0,\infty) \times \bar{S}_{n,m}$, and the coefficient $(\varsigma(z))$ has at most linear growth in the spatial variable, we see that the sequence $\{H^{\varepsilon}_{i,k}\}_{\varepsilon \geq 0}$ is uniformly bounded on $[0,t]$, and converges $\widehat{\mathbb{P}}^z$-a.s. to $H^0_{i,k}$, for all $s \in [0,t]$. Then [31, Theorem IV.2.32] implies that

$$\int_0^t H^{\varepsilon}_{i,k}(s)\, d\widehat{W}_k(s) \longrightarrow \int_0^t H^{0}_{i,k}(s)\, d\widehat{W}_k(s), \qquad \text{as } \varepsilon \downarrow 0,$$

where the convergence takes place in probability. Using the fact that

$$\int_0^{t \wedge \tau_N} \varsigma_{i,k}(s, \widehat{Z}(s)) u_{x_i}(s, \widehat{Z}^{\varepsilon}(s))\, d\widehat{W}_k(s) = \int_0^t H^{\varepsilon}_{i,k}(s)\ d\widehat{W}_k(s), \quad \forall\, \varepsilon \geq 0,$$

we see that

(2.36)

$$\int_0^{t \wedge \tau_N} \varsigma_{i,k}(s, \widehat{Z}(s)) u_{x_i}(s, \widehat{Z}^{\varepsilon}(s))\, d\widehat{W}_k(s) \longrightarrow \int_0^{t \wedge \tau_N} \varsigma_{i,k}(s, \widehat{Z}(s)) u_{x_i}(s, \widehat{Z}(s))\, d\widehat{W}_k(s),$$

where the convergence takes place in probability, as ε tends to zero. Analogously to $H^{\varepsilon}_{i,k}(s)$, we define $G^{\varepsilon}(s)$ to be equal to the integrand of the ds-term on the right-hand side of (2.34), multiplied by the indicator of the set $\{|\widehat{Z}(s)| \leq N\}$, for all $s \in [0,t]$, and all $\varepsilon \geq 0$. Using the fact that the derivatives u_s, u_{x_i}, u_{y_l}, and $u_{y_l y_k}$ are continuous on $[0,\infty) \times \bar{S}_{n,m}$, and the fact that the process $\{\widehat{Z}(t)\}_{t \geq 0}$ has continuous paths, we have the $\widehat{\mathbb{P}}^z$-a.s. convergence, as $\varepsilon \downarrow 0$,

$$\begin{aligned}
&u_s(s, \widehat{Z}^{\varepsilon}(s)) + \sum_{i=1}^n b_i(\widehat{Z}(s)) u_{x_i}(s, \widehat{Z}^{\varepsilon}(s)) \\
&\quad + \sum_{l=1}^m e_l(\widehat{Z}(s)) u_{y_l}(s, \widehat{Z}^{\varepsilon}(s)) + \frac{1}{2} \sum_{l,k=1}^m a_{n+l,n+k}(\widehat{Z}(s)) u_{y_l y_k}(s, \widehat{Z}^{\varepsilon}(s)) \\
&\qquad \longrightarrow u_s(s, \widehat{Z}(s)) + \sum_{i=1}^n b_i(\widehat{Z}(s)) u_{x_i}(s, \widehat{Z}(s)) \\
&\qquad\quad + \sum_{l=1}^m e_l(\widehat{Z}(s)) u_{y_l}(s, \widehat{Z}(s)) + \frac{1}{2} \sum_{l,k=1}^m a_{n+l,n+k}(\widehat{Z}(s)) u_{y_l y_k}(s, \widehat{Z}(s)),
\end{aligned}$$

for all $s \in [0,t]$. Recall from Lemma 2.12 that the weighted derivatives $\sqrt{x_i x_j} u_{x_i x_j}$ and $\sqrt{x_i} u_{x_i y_l}$ satisfy properties (2.26) and (2.27). Using these properties together with the continuity of $\sqrt{x_i x_j} u_{x_i x_j}$ and $\sqrt{x_i} u_{x_i y_l}$ on $[0,\infty) \times \bar{S}_{n,m}$, and the fact that

$\{\widehat{Z}(t)\}_{t\geq 0}$ has continuous paths, by writing

$$\begin{aligned}
&\sqrt{\widehat{X}_i(s)\widehat{X}_j(s)}a_{i,j}(\widehat{Z}(s))u_{x_ix_j}(s,\widehat{Z}^\varepsilon(s))\\
&\quad= \frac{\sqrt{\widehat{X}_i(s)\widehat{X}_j(s)}}{\sqrt{\widehat{X}^\varepsilon_i(s)\widehat{X}^\varepsilon_j(s)}}a_{i,j}(\widehat{Z}(s))\sqrt{\widehat{X}^\varepsilon_i(s)\widehat{X}^\varepsilon_j(s)}u_{x_ix_j}(s,\widehat{Z}^\varepsilon(s)),\\
&\sqrt{\widehat{X}_i(s)}a_{i,n+l}(\widehat{Z}(s))u_{x_iy_l}(s,\widehat{Z}^\varepsilon(s))\\
&\quad= \frac{\sqrt{\widehat{X}_i(s)}}{\sqrt{\widehat{X}^\varepsilon_i(s)}}a_{i,n+l}(\widehat{Z}(s))\sqrt{\widehat{X}^\varepsilon_i(s)}u_{x_iy_l}(s,\widehat{Z}^\varepsilon(s)),
\end{aligned}$$

we see that, as $\varepsilon\downarrow 0$,

$$\begin{aligned}
\sqrt{\widehat{X}_i(s)\widehat{X}_j(s)}a_{i,j}(\widehat{Z}(s))u_{x_ix_j}(s,\widehat{Z}^\varepsilon(s)) &\longrightarrow \sqrt{\widehat{X}_i(s)\widehat{X}_j(s)}a_{i,j}(\widehat{Z}(s))u_{x_ix_j}(s,\widehat{Z}(s)),\\
\sqrt{\widehat{X}_i(s)}a_{i,n+l}(\widehat{Z}(s))u_{x_iy_l}(s,\widehat{Z}^\varepsilon(s)) &\longrightarrow \sqrt{\widehat{X}_i(s)}a_{i,n+l}(\widehat{Z}(s))u_{x_iy_l}(s,\widehat{Z}(s)),
\end{aligned}$$

where the convergence takes place $\widehat{\mathbb{P}}^z$-a.s., for all $s\in[0,t]$. Combining this with the fact that the coefficients $b(z)$, $e(z)$, and $(a(z))$ are bounded on compact sets in $\bar{S}_{n,m}$, we see that the sequence $\{G^\varepsilon\}_{\varepsilon\geq 0}$ is uniformly bounded on $[0,t]$, and converges $\widehat{\mathbb{P}}^z$-a.s. to G^0, for all $s\in[0,t]$. Then [31, Theorem IV.2.32] implies that

$$\int_0^t G^\varepsilon(s)\,ds \longrightarrow \int_0^t G^0(s)\,ds,\quad \text{in probability as } \varepsilon\downarrow 0.$$

The preceding convergence property is equivalent to the fact that the ds-term on the right-hand side of (2.34) converges in probability to the ds-term on the right-hand side of (2.33). By combining the latter convergence in probability with (2.35) and (2.36), we find that the right-hand side of (2.34) converges in probability to the right-hand side in (2.33), as ε tends to zero. This concludes the proof. □

We recall from [29, §2] that the anisotropic Hölder space $C^{2+\alpha}_{\mathrm{WF}}([0,T]\times\bar{S}_{n,m})$ consists of functions, $u:[0,T]\times\bar{S}_{n,m}\to\mathbb{R}$, that belong to $C([0,T]\times\bar{S}_{n,m})$, such that the distributional derivatives $u_t, u_{x_i}, u_{y_l}, u_{x_ix_j}, u_{x_iy_l}, u_{y_ly_k}$ are continuous functions on $(0,T)\times S_{n,m}$, for all $1\leq i,j\leq n$ and $1\leq l,k\leq m$. Moreover, the following functions have continuous extensions up to the boundary of the domain $(0,T)\times S_{n,m}$, and they belong to the corresponding anisotropic Hölder space: For all subsets $I\subseteq\{1,\dots,n\}$, we assume that

(2.37)
$$\begin{aligned}
u, u_t, u_{x_i}, u_{y_l}, u_{y_ly_k} &\in C^\alpha_{\mathrm{WF}}([0,T]\times\bar{S}_{n,m}),\quad \forall\, i,j=1,\dots,n,\quad \forall\, l,k=1,\dots,m,\\
\sqrt{x_ix_j}u_{x_ix_j}, \sqrt{x_i}u_{x_iy_l} &\in C^\alpha_{\mathrm{WF}}([0,T]\times\bar{M}_I),\quad \forall\, i,j\in I,\quad \forall\, l,k=1,\dots,m,\\
\sqrt{x_i}u_{x_ix_j}, u_{x_jx_k} &\in C^\alpha_{\mathrm{WF}}([0,T]\times\bar{M}_I),\quad \forall\, i\in I,\quad \forall\, j,k\in I^c,
\end{aligned}$$

where we recall the definition of the sets M_I in (2.14). We can now give the

Proof of Proposition 2.4. The method of the proof is similar to that of [19, Theorem 1.3 and Proposition 3.6]. As stated in §2.2, it suffices to show that for all $\varphi\in C^\infty_c(\bar{S}_{n,m})$, $z\in\bar{S}_{n,m}$, and $T>0$, if $(\widehat{Z}^i,\widehat{W}^i)$, for $i=1,2$, are two weak solutions to the stochastic differential equation (1.3) with initial condition $\widehat{Z}^i(0)=z$, then identity (2.10) holds. Assumption 2.6 guarantees that we can apply [29, Theorem

1.4] to the operator $\widehat{L}$, to conclude that the homogeneous initial-value problem (1.7) has a unique solution, $u \in C^{2+\alpha}_{\mathrm{WF}}([0,T] \times \bar{S}_{n,m})$. By (2.37), the function $(t,z) \mapsto u(T-t,z)$ satisfies the hypotheses of Proposition 2.10, and property (2.24) yields

(2.38)

$$\begin{aligned} du(T-t,\widehat{Z}^i(t)) &= -(u_t + \widehat{L}u)(T-t,\widehat{Z}^i(t))\,dt \\ &\quad + \sum_{j=1}^{n}\sum_{k=1}^{n+m} \varsigma_{j,k}(\widehat{Z}^i(t)) u_{x_j}(T-t,\widehat{Z}^i(t))\,d\widehat{W}_k(t) \\ &\quad + \sum_{l=1}^{m}\sum_{k=1}^{n+m} \varsigma_{n+l,k}(\widehat{Z}^i(t)) u_{y_l}(T-t,\widehat{Z}^i(t))\,d\widehat{W}_k(t), \quad \forall\, i=1,2, \end{aligned}$$

where we recall that the coefficient matrix $(\varsigma(z))$ is defined by (2.9). The $d\widehat{W}_k(t)$-terms in the preceding identity define martingales because the coefficient matrix $(\varsigma(z))$ is bounded, by Lemma 2.9, and the derivatives u_{x_j} and u_{y_l} are bounded functions on $[0,T] \times \bar{S}_{n,m}$, by (2.37). Combining the preceding observation with the fact that u is a solution to the initial-value problem (1.7), we obtain from identity (2.38) that

$$u(0,z) = \mathbb{E}_{\mathbb{P}^z_i}\left[\varphi(Z^i(T))\right], \quad \forall\, z \in \bar{S}_{n,m}, \quad \forall\, i=1,2.$$

In particular, identity (2.10) holds, which implies by [23, Proposition 5.4.27], that uniqueness in law holds for solutions to the Kimura stochastic differential equation (1.3). This completes the proof. ☐

3. Kimura diffusions with singular drift

In this section we prove existence, uniqueness in law, and the strong Markov property of weak solutions to Kimura stochastic differential equations with *singular drift*, (1.6). Our strategy of the proof is to apply Girsanov's Theorem [23, Theorem 3.5.1] to build a new probability measure so that solutions to the generalized Kimura stochastic differential equation (1.3) become solutions to the equation with singular drift (1.6), under the new probability measure. The weak solutions obtained by this method satisfy the strong Markov property. Girsanov's Theorem also allows us to prove that uniqueness in law of weak solutions to equation (1.6) holds, in the class of Markov processes.

3.1. Existence of weak solutions.

In this section we prove the existence of weak solutions to the singular Kimura equation (1.6). The solutions that we build in Theorem 3.1 satisfy the strong Markov property.

Theorem 3.1 (Existence of weak solutions to the Kimura equation with singular drift (1.6))**.** *Suppose that the coefficients of the Kimura stochastic differential equation with singular drift* (1.6) *satisfy Assumption* 3.2*. Then, for all $z \in \bar{S}_{n,m}$, there is a weak solution $(Z=(X,Y),W)$, $(\Omega, \{\mathcal{F}_t\}_{t\geq 0}, \mathcal{F}, \mathbb{P}^z)$, to equation* (1.6)*, with initial condition $Z(0)=z$. Moreover, the solution satisfies the strong Markov property.*

To prove existence of weak solutions to the Kimura stochastic differential equation with singular drift (1.6), we assume that the coefficients satisfy the following conditions.

Assumption 3.2 (Properties of the coefficients in (1.6))**.** *Let $q \in (0, q_0)$, where q_0 is given by*

$$q_0 := \min\left\{\frac{1}{4}, \frac{b_0}{(n+m)K^2}\right\}, \tag{3.1}$$

and K is the positive constant appearing in Lemma 2.8*, and b_0 satisfies condition* (3.2) *below. We assume that:*

1. *The functions $b(z)$, $e(z)$, and $(\sigma(z))$ satisfy Assumption* 2.6*.*
2. *The drift coefficients satisfy the* positivity condition*: there is a positive constant, b_0, such that for all $i = 1, \ldots, n$, we have that*

$$b_i(z) \geq b_0 > 0, \quad \forall\, z \in \partial S_{n,m} \cap \{x_i = 0\}. \tag{3.2}$$

3. *The coefficients $f_{i,j} : \bar{S}_{n,m} \to \mathbb{R}$, and $h_{i,j} : [0, \infty) \to \mathbb{R}$ are Borel measurable functions, for all $i, k = 1, \ldots, n+m$, and all $j = 1, \ldots, n$, and there is a positive constant, K_0, such that*

$$|f_{i,j}(z)| \leq K_0, \quad \forall\, z \in \bar{S}_{n,m}, \tag{3.3}$$

$$|(\sigma^{-1})_{k,i}(z) f_{i,j}(z)| \leq K_0, \quad \forall\, z \in \bar{S}_{n,m}, \tag{3.4}$$

$$|h_{i,j}(s)| \leq K_0 s^{-q}, \quad \forall\, s \in (0, \infty). \tag{3.5}$$

Condition (3.4) uses the fact that the matrix coefficient $(\sigma(z))$ is invertible on $\bar{S}_{n,m}$. We next prove that this property is a consequence of conditions (2.17) and (2.18).

Lemma 3.3 (Invertibility of the matrix coefficient $(\sigma(z))$)**.** *Suppose that the matrix $(a(z))$, defined in* (2.11)*, satisfies properties* (2.17) *and* (2.18)*. Then $(\sigma(z))$ is invertible on $\bar{S}_{n,m}$, and $(\sigma^{-1}(z))$ has bounded coefficients on compact sets in $\bar{S}_{n,m}$.*

Proof. Let $r \in (0, 1)$, and for $z = (x, y) \in \bar{S}_{n,m}$, let I denote the set of indices $i \in \{1, \ldots, n\}$ such that $x_i \in [0, r]$, and I^c denoted the set of indices $j \in \{1, \ldots, n\}$ such that $x_j > r$. Let $R > 0$ be such that $x_j \leq R$, for all $j \in I^c$. Using identities (2.17), for all $\zeta \in \mathbb{R}^n$ and $\eta \in \mathbb{R}^m$, we have that

$$\begin{aligned}
&\sum_{i,j=1}^n a_{i,j}(z)\zeta_i\zeta_j + \sum_{i=1}^n\sum_{l=1}^m \left(a_{i,n+l}(z) + a_{n+l,i}(z)\right)\zeta_i\eta_l + \sum_{l,k=1}^m a_{n+l,n+k}(z)\eta_l\eta_k \\
&\quad = \sum_{i=1}^n \alpha_{i,i}(z)\zeta_i^2 + \sum_{i,j=1}^n \sqrt{x_i x_j}\tilde{\alpha}_{i,j}(z)\zeta_i\zeta_j \\
&\qquad + \sum_{i=1}^n\sum_{l=1}^m \sqrt{x_i} c_{i,l}(z)\zeta_i\eta_l + \sum_{k,l=1}^m a_{n+k,n+l}(z)\eta_k\eta_l \\
&\quad = \sum_{i\in I} x_i\alpha_{i,i}(z)\zeta_i^2 + \sum_{i\in I^c} \alpha_{i,i}(z)\zeta_i^2 + \sum_{i,j=1}^n \sqrt{x_i x_j}\tilde{\alpha}_{i,j}(z)\zeta_i\zeta_j \\
&\qquad + \sum_{i=1}^n\sum_{l=1}^m \sqrt{x_i} c_{i,l}(z)\zeta_i\eta_l + \sum_{k,l=1}^m a_{n+k,n+l}(z)\eta_k\eta_l + \sum_{i\in I}(1 - x_i)\alpha_{i,i}(z)\zeta_i^2.
\end{aligned}$$

The last inequality, and the strict ellipticity condition (2.18), applied with the vector $\xi \in \mathbb{R}^n$ chosen such that $\xi_i = \sqrt{x_i}\zeta_i$, for all $i \in I$, and $\xi_j = \frac{1}{\sqrt{x_j}}\zeta_j$, for all

$j \in I^c$, yield

$$\begin{aligned}
\sum_{i,j=1}^n a_{i,j}(z)\zeta_i\zeta_j &+ \sum_{i=1}^n\sum_{l=1}^m \left(a_{i,n+l}(z)+a_{n+l,i}(z)\right)\zeta_i\eta_l + \sum_{l,k=1}^m a_{n+l,n+k}(z)\eta_l\eta_k \\
&\geq \lambda\left(\sum_{i\in I} x_i|\zeta_i|^2 + \sum_{j\in I^c}\frac{1}{x_j}|\zeta_j|^2 + |\eta|^2\right) + (1-r)\lambda\sum_{i\in I}|\zeta_i|^2 \\
&\geq \lambda\min\left\{\frac{1}{R}, 1-r\right\}\left(|\zeta|^2+|\eta|^2\right), \qquad (3.6)
\end{aligned}$$

where we used the fact that $r \in (0,1)$, $1-x_i \geq 1-r$, for all $i \in I$, and $x_j \leq R$, for all $j \in I^c$. Thus, indeed the matrix $(a(z))$ is strictly positive definite, and so, the matrix $(\sigma(z))$ is invertible, for all $z \in \bar{S}_{n,m}$. Inequality (3.6), and identity (2.11) gives us that the coefficients of $(\sigma^{-1}(z))$ are bounded on compact sets in $\bar{S}_{n,m}$. This completes the proof. □

Remark 3.4 (Condition (3.4)). In general, the boundedness condition (3.4) is not a consequence of Assumption 2.6. Our main motivation to study the singular Kimura stochastic differential equation (1.6) is for its application to the proof of the Harnack inequality for nonnegative solutions to the generalized Kimura parabolic equation (1.7); see [14, Theorems 1.2 and 1.3]. For such applications, it is sufficient to assume that the coefficients $f_{i,j} : S_{n,m} \to \mathbb{R}$ are bounded, Borel measurable functions, for all $i = 1,\dots,n+m$, and all $j = 1,\dots,n$, and that they have compact support in $\bar{S}_{n,m}$. Under such assumptions, Lemma 3.3 gives us that condition (3.4) is a consequence of (2.17), (2.18), and the upper bound (3.3).

The proof of Theorem 3.1 is based on an application of Girsanov's Theorem. We change the probability distributions of the weak solutions of the generalized Kimura equation (1.3) obtained in Proposition 2.2, so that we add a singular drift as in equation (1.6). In order to justify the application of Girsanov's Theorem, we prove that Novikov's condition, [23, Corollary 3.5.13], for generalized Kimura diffusions holds.

Lemma 3.5 (Novikov's condition for generalized Kimura diffusions)**.** *Suppose that the coefficients of the Kimura stochastic differential equation* (1.3) *satisfy Assumption* 2.6*, and condition* (3.2)*. Let $q \in (0,q_0)$, where the positive constant q_0 is given by* (3.1)*. Then, for all $\Lambda > 0$ and $T > 0$, we have*

$$\sup_{z\in\bar{S}_{n,m}} \mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\exp\left(\Lambda\int_0^T\sum_{i=1}^n |\widehat{X}_i(t)|^{-2q}\,dt\right)\right] < \infty, \qquad (3.7)$$

where $\{\widehat{Z}(t) = (\widehat{X}(t),\widehat{Y}(t))\}_{t\geq 0}$ is the unique weak solution to the Kimura stochastic equation (1.3)*, with initial condition $\widehat{Z}(0) = z$.*

An elementary method to guarantee that condition (3.7) holds is to prove that the hypotheses of Khas'minskii's Lemma [5, 25, 30] are satisfied. A statement of Khas'minskii's Lemma when the underlying process is Brownian motion can be found in [1, Theorem 1.2], which states that if $\{W(t)\}_{t\geq 0}$ is a d-dimensional Brownian motion, and $g : \mathbb{R}^d \to [0,\infty)$ is a (nonnegative) Borel measurable function such that there is a constant, $\delta \in (0,1)$, with the property that

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}_x \left[\int_0^t g(W(s))\,ds\right] = \delta,$$

then we have that

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}_x \left[\exp\left(\int_0^t g(W(s))\,ds\right)\right] \leq \frac{1}{1-\delta}.$$

Lemma 3.6 (Verification of the hypotheses of Khas'minskii's Lemma)**.** *Suppose that the coefficients of the Kimura stochastic differential equation* (1.3) *satisfy Assumption* 2.6*, and condition* (3.2)*. Let* Λ *be a positive constant and* $q \in (0, q_0)$*, where the positive constant* q_0 *is given by* (3.1)*. Then for all* $\delta \in (0,1)$*, there is a positive constant,* $T = T(b_0, \|b\|_{C^\alpha_{\mathrm{WF}}(\bar S_{n,m})}, \delta, K, \Lambda, m, n, q)$*, such that*

$$\sup_{z\in\bar S_{n,m}} \mathbb{E}_{\widehat{\mathbb{P}}^z} \left[\int_0^T \Lambda \sum_{i=1}^n |\widehat X_i(t)|^{-2q}\,dt\right] < \delta, \tag{3.8}$$

where $\{\widehat Z(t) = (\widehat X(t), \widehat Y(t))\}_{t\geq 0}$ *is the unique weak solution to the Kimura stochastic equation* (1.3)*, with initial condition* $\widehat Z(0) = z$*.*

Proof. Without loss of generality, we may assume that $\Lambda = 1$. Using condition (3.2) and the uniform continuity of the coefficient $b_i(z)$ implied by (2.19), we obtain that for all $\rho \in (0,1)$, there is a positive constant, r, such that

$$b_i(z) \geq \frac{b_0}{1+\rho} \quad \text{on } \{z = (x,y) \in S_{n,m} :\ x_i \in [0,r]\}, \quad \forall\, i = 1,\dots,n. \tag{3.9}$$

Let $\varphi : [0,\infty) \to [0,1]$ be a smooth cut-off function, such that $\varphi(s) = 1$ for $s \leq r/2$, and $\varphi(s) = 0$ for $s \geq r$, and such that there is a positive constant, c, with the property that

$$|\varphi'(s)| \leq cr^{-1}, \quad \text{and} \quad |\varphi''(s)| \leq cr^{-2}, \quad \forall\, s \geq 0. \tag{3.10}$$

For all $\varepsilon \in (0,1)$, we let $\widehat X^\varepsilon_i(t) := \widehat X_i(t) + \varepsilon$, and $x^\varepsilon_i := x_i + \varepsilon$. By Itô's rule [23, Theorem 3.3.6] applied to the process $\varphi(X^\varepsilon_i(t))(X^\varepsilon_i(t))^{1-2q}$, we obtain

$$\begin{aligned}
d\varphi(\widehat X^\varepsilon_i(t))(\widehat X^\varepsilon_i(t))^{1-2q}
&= (1-2q)\varphi(\widehat X^\varepsilon_i(t))(\widehat X^\varepsilon_i(t))^{-2q}\left(b_i(\widehat Z(t)) - q|\sigma_i(\widehat Z(t))|^2 \frac{\widehat X_i(t)}{\widehat X^\varepsilon_i(t)}\right) dt \\
&\quad + b_i(\widehat Z(t))\varphi'(\widehat X^\varepsilon_i(t))(\widehat X^\varepsilon_i(t))^{1-2q}\,dt \\
&\quad + \frac{|\sigma_i(\widehat Z(t))|^2}{2}\widehat X_i(t)(\widehat X^\varepsilon_i(t))^{-2q}\left(\widehat X^\varepsilon_i(t)\varphi''(\widehat X^\varepsilon_i(t)) + 2(1-2q)\varphi'(\widehat X^\varepsilon_i(t))\right) dt \\
&\quad + (\widehat X^\varepsilon_i(t))^{-2q}\left(\widehat X^\varepsilon_i(t)\varphi'(\widehat X^\varepsilon_i(t))\right. \\
&\quad \left. + (1-2q)\varphi(\widehat X^\varepsilon_i(t))\right)\sqrt{\widehat X_i(t)}\sum_{k=1}^{n+m}\sigma_{i,k}(\widehat Z(t))\,d\widehat W_k(t),
\end{aligned}$$

where we denote by $\sigma_i(z)$ the i-th row of the matrix $(\sigma(z))$. From Lemma 2.9 and identity (2.9), we see that the coefficients $(\sqrt{x_i}\sigma_i(z))$ are bounded, and so, the

$d\widehat{W}_k(t)$-terms in the preceding equality define martingales. We obtain

$$
\begin{aligned}
&\mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\varphi(\widehat{X}_i^\varepsilon(T))(\widehat{X}_i^\varepsilon(T))^{1-2q}\right] = \varphi(x_i^\varepsilon)(x^\varepsilon)^{1-2q}\\
&\quad + (1-2q)\mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\int_0^T \varphi(\widehat{X}_i^\varepsilon(t))(\widehat{X}_i^\varepsilon(t))^{-2q}\left(b_i(\widehat{Z}(t)) - q|\sigma_i(\widehat{Z}(t))|^2\frac{\widehat{X}_i(t)}{\widehat{X}_i^\varepsilon(t)}\right)dt\right]\\
&\quad + \mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\int_0^T b_i(\widehat{Z}(t))\varphi'(\widehat{X}_i^\varepsilon(t))(\widehat{X}_i^\varepsilon(t))^{1-2q}\,dt\right]\\
&\quad + \mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\int_0^T \frac{|\sigma_i(\widehat{Z}(t))|^2}{2}\widehat{X}_i(t)(\widehat{X}_i^\varepsilon(t))^{-2q}\left(\widehat{X}_i^\varepsilon(t)\varphi''(\widehat{X}_i^\varepsilon(t))\right.\right.\\
&\qquad\qquad\qquad\qquad\qquad\left.\left. + 2(1-2q)\varphi'(\widehat{X}_i^\varepsilon(t))\right)dt\right],
\end{aligned}
$$

for all $z = (x, y) \in \bar{S}_{n,m}$. The preceding identity together with the boundedness of the coefficients $b(z)$ and $(\sigma(z))$ (see inequality (2.20)), and the choice of the cut-off function φ, and (3.10), give us that there is a positive constant, $C = C(K, m, n)$, such that

$$
\begin{aligned}
&(1-2q)\mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\int_0^T \varphi(\widehat{X}_i^\varepsilon(t))(\widehat{X}_i^\varepsilon(t))^{-2q}\left(b_i(\widehat{Z}(t)) - q|\sigma_i(\widehat{Z}(t))|^2\frac{\widehat{X}_i(t)}{\widehat{X}_i^\varepsilon(t)}\right)dt\right]\\
&\qquad \le Cr^{-2q}T.
\end{aligned}
$$

Using inequalities (3.9) and (2.20), we see that

(3.11)
$$
\begin{aligned}
&(1-2q)\mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\int_0^T \varphi(\widehat{X}_i^\varepsilon(t))(\widehat{X}_i^\varepsilon(t))^{-2q}\left(b_i(\widehat{Z}(t)) - q|\sigma_i(\widehat{Z}(t))|^2\frac{\widehat{X}_i(t)}{\widehat{X}_i^\varepsilon(t)}\right)dt\right]\\
&\qquad \ge (1-2q)\left(\frac{b_0}{1+\rho} - q(n+m)K^2\right)\mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\int_0^T (\widehat{X}_i^\varepsilon(t))^{-2q}\mathbf{1}_{\{\widehat{X}_i(t)\in[0,r/2]\}}\,dt\right].
\end{aligned}
$$

Combining the preceding two inequalities, and letting ε tend to 0, we obtain that there is a positive constant, $C = C(K, m, n)$, such that

$$
\mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\int_0^T (\widehat{X}_i(t))^{-2q}\,dt\right] \le \frac{Cr^{-2q}T}{(1-2q)\left(\frac{b_0}{1+\rho} - q(n+m)K^2\right)}. \tag{3.12}
$$

Note that by choosing $q \in (0, q_0)$, where the positive constant q_0 is given by (3.1), we can find a positive constant, $\rho_0 = \rho_0(K, m, n)$, such that

$$
\frac{b_0}{1+\rho_0} - q(n+m)K^2 > 0.
$$

Then we choose $r_0 = r_0(\|b\|_{C^\alpha_{\mathrm{WF}}(\bar{S}_{n,m})}, K, m, n)$, such that inequality (3.9) holds with ρ replaced by ρ_0, for all $r \in (0, r_0)$. For all $\delta \in (0, 1)$, let

$$
r = r(\|b\|_{C^\alpha_{\mathrm{WF}}(\bar{S}_{n,m})}, \delta, K, m, n, q)
$$

and

$$
T = T(b_0, \|b\|_{C^\alpha_{\mathrm{WF}}(\bar{S}_{n,m})}, \delta, K, m, n, q)
$$

be chosen small enough such that using inequality (3.12), we obtain that estimate (3.8) holds. This completes the proof. □

Using Lemma 3.6, we can now give the proof of

Proof of Lemma 3.5. By Corollary 2.5, the solutions to generalized Kimura stochastic differential equations (1.3) satisfy the Markov property. Thus, the proof of [1, Theorem 1.2] applies equally well to generalized Kimura diffusions in place of standard Brownian motion, and using Lemma 3.6, we obtain that for all $\delta > 0$, there is $T_\delta > 0$, such that

$$\sup_{z\in\bar{S}_{n,m}} \mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\exp\left(\int_0^{T_\delta} \varphi(\widehat{X}(t))\,dt\right)\right] < \frac{1}{1-\delta}, \tag{3.13}$$

where we denote for brevity, $\varphi(x) := \Lambda\sum_{i=1}^n |x_i|^{-2q}$, for all $x\in\mathbb{R}^n_+$. Let $T>0$ and set $k := \lceil T/T_\delta\rceil$. We consider the sequence $T_i := T-(k-i)T_\delta$, for all $i=1,\ldots,k$, and $T_0 = 0$. We have, for all $z\in\bar{S}_{n,m}$,

$$\begin{aligned}
\mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\exp\left(\int_0^T \varphi(\widehat{X}(t))\,dt\right)\right] &= \mathbb{E}_{\widehat{\mathbb{P}}^z}\left[e^{\int_0^{T_{k-1}} \varphi(\widehat{X}(t))\,dt} e^{\int_{T_{k-1}}^{T_k} \varphi(\widehat{X}(t))\,dt}\right]\\
&= \mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\mathbb{E}_{\widehat{\mathbb{P}}^z}\left[e^{\int_0^{T_{k-1}} \varphi(\widehat{X}(t))\,dt} e^{\int_{T_{k-1}}^{T_k} \varphi(\widehat{X}(t))\,dt}\Big|\mathcal{F}_{T_{k-1}}\right]\right]\\
&= \mathbb{E}_{\widehat{\mathbb{P}}^z}\left[e^{\int_0^{T_{k-1}} \varphi(\widehat{X}(t))\,dt}\mathbb{E}_{\widehat{\mathbb{P}}^{\widehat{Z}(T_{k-1})}}\left[e^{\int_0^{T_\delta} \varphi(\widehat{X}(t))\,dt}\right]\right].
\end{aligned}$$

Inequality (3.13) gives us

$$\mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\exp\left(\int_0^T \varphi(\widehat{X}(t))\,dt\right)\right] \le \frac{1}{1-\delta}\mathbb{E}_{\widehat{\mathbb{P}}^z}\left[e^{\int_0^{T_{k-1}} \varphi(\widehat{X}(t))\,dt}\right],$$

and iterating the preceding argument k times, we obtain

$$\mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\exp\left(\int_0^T \varphi(\widehat{X}(t))\,dt\right)\right] \le \frac{1}{(1-\delta)^k}, \quad \forall z\in\bar{S}_{n,m},$$

from where inequality (3.7) now follows. □

Lemma 3.5 allows us to establish the

Proof of Theorem 3.1. We divide the proof into two steps. In Step 1, we prove existence of weak solutions to the Kimura stochastic differential equation with singular drift (1.6), via Girsanov's Theorem, and in Step 2, we show that the solution constructed in Step 1 satisfies the strong Markov property.

Step 1 (Existence of weak solutions to equation (1.6)). Let $z\in\bar{S}_{n,m}$. Because the coefficient functions b, e, and σ satisfy Assumptions 2.1 and 2.6, Propositions 2.2 and 2.4 show that there is a unique weak solution, $(\widehat{Z} = (\widehat{X},\widehat{Y}),\widehat{W})$, $(\Omega,\{\mathcal{F}_t\}_{t\ge0},\mathcal{F},\widehat{\mathbb{P}}^z)$, to the Kimura stochastic differential equation (1.3), with initial condition $\widehat{Z}(0)=z$. Let $\theta:\bar{S}_{n,m}\to\mathbb{R}^{n+m}$ be the Borel measurable vector field defined by

$$\theta := \sigma^{-1}\xi, \tag{3.14}$$

where the function $\xi : \bar{S}_{n,m} \to \mathbb{R}^{n+m}$ is defined by

$$\xi_i(z) := \sum_{j=1}^n f_{i,j}(z) h_{i,j}(x_j), \quad \forall\, z \in \bar{S}_{n,m}, \quad \forall\, i = 1, \ldots, n+m. \tag{3.15}$$

Definitions (3.14) and (3.15) give us that

$$\theta_k(z) = \sum_{i=1}^{n+m} \sum_{j=1}^{n} (\sigma^{-1})_{k,i}(z) f_{i,j}(z) h_{i,j}(x_j), \quad \forall\, z \in \bar{S}_{n,m}, \quad \forall\, k = 1, \ldots, n+m,$$

and using conditions (3.4), and (3.5), it follows that there is a positive constant, Λ, such that

$$|\theta(z)| \le \Lambda \sum_{i=1}^n |x_i|^{-q}, \quad \forall\, z \in \bar{S}_{n,m}. \tag{3.16}$$

Let T be a positive constant. Lemma 3.5 and inequality (3.16) show that

$$\mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\exp\left(\frac{1}{2}\int_0^T |\theta(t)|^2\, dt\right)\right] < \infty,$$

and so, [23, Corollary 3.5.13] implies that the process $\{\widehat{M}(t)\}_{0\le t\le T}$ defined by

$$\widehat{M}(t) := \exp\left(\int_0^t \sum_{i=1}^{n+m} \theta_i(\widehat{Z}(s))\, d\widehat{W}_i(s) - \frac{1}{2}\int_0^t |\theta(\widehat{Z}(s))|^2\, ds\right), \quad \forall\, t \in [0,T],$$

is a $\widehat{\mathbb{P}}^z$-martingale. We can apply Girsanov's Theorem [23, Theorem 3.5.1] to construct a new probability measure, $\mathbb{P}^z$, by letting

$$\frac{d\mathbb{P}^z}{d\widehat{\mathbb{P}}^z} = \widehat{M}(T), \tag{3.17}$$

such that the process

$$W(t) := \widehat{W}(t) - \int_0^t \theta(\widehat{Z}(s))\, ds, \quad \forall\, t \in [0,T], \tag{3.18}$$

is a $\mathbb{P}^z$-Brownian motion. Using (3.14), we see that by letting $Z(t) := \widehat{Z}(t)$, for all $t \in [0,T]$, we obtain that the process $\{Z(t), W(t)\}_{0\le t\le T}$, $(\Omega, \{\mathcal{F}_t\}_{0\le t\le T}, \mathcal{F}, \mathbb{P}^z)$ is a weak solution to the Kimura stochastic differential equation with singular drift (1.6), with initial condition $Z(0) = z$. To see this, using identity (3.18) in (1.3), we have that

$$\begin{aligned} d\widehat{X}_i(t) &= \left(b_i(\widehat{Z}(t)) + \sqrt{\widehat{X}_i(t)} \sum_{k=1}^{n+m} \sigma_{i,k}(\widehat{Z}(t))\theta_k(\widehat{Z}(t))\right) dt \\ &\quad + \sqrt{\widehat{X}_i(t)} \sum_{k=1}^{n+m} \sigma_{i,k}(\widehat{Z}(t))\, dW_k(t), \\ d\widehat{Y}_l(t) &= \left(e_l(\widehat{Z}(t)) + \sum_{k=1}^{n+m} \sigma_{n+l,k}(\widehat{Z}(t))\theta_k(\widehat{Z}(t))\right) dt + \sum_{k=1}^{n+m} \sigma_{n+l,k}(\widehat{Z}(t))\, dW_k(t), \end{aligned}$$

while identities (3.14) and (3.15) give us that

$$\sum_{k=1}^{n+m} \sigma_{i,k}(z)\theta_k(z) = \sum_{j=1}^{n} f_{i,j}(z) h_{i,j}(x_j), \quad \forall\, i = 1, \ldots, n+m.$$

The preceding three identities imply that the process $\{Z(t) := \widehat{Z}(t), W(t)\}_{\{0\leq t\leq T\}}$ is a weak solution to (1.6). This completes the proof of Step 1.

Step 2 (The strong Markov property). Let $z \in \bar{S}_{n,m}$, and let (Z,W), $(\Omega, \{\mathcal{F}_t\}_{t\geq 0}, \mathcal{F}, \mathbb{P}^z)$ be the weak solution to the Kimura stochastic differential equation with singular drift (1.6), with initial condition $Z(0)=z$, constructed in Step 1. We now show that the process $\{Z(t)\}_{t\geq 0}$ satisfies the strong Markov property, that is, for all stopping times, τ, for all $t\geq 0$, and $B\in\mathcal{B}(\bar{S}_{n,m})$, we have that

$$\mathbb{P}^z\left(Z(\tau+t)\in B|\mathcal{F}_\tau\right) = \mathbb{P}^z\left(Z(\tau+t)\in B|Z(\tau)\right) \quad \mathbb{P}^z\text{-a.s. on } \{\tau<\infty\}. \tag{3.19}$$

It is sufficient to prove identity (3.19) for all bounded stopping times in order to conclude that the strong Markov property (3.19) holds for arbitrary stopping times. Let $T>0$ and let τ be a stopping time such that $\tau\leq T$, $\mathbb{P}^z$-a.s. Notice that using (3.17), for all $\mathcal{F}_\tau$-measurable and bounded random variables, Y, we have that

$$\begin{aligned}\mathbb{E}_{\widehat{\mathbb{P}}^z}[Y] &= \mathbb{E}_{\mathbb{P}^z}\left[\widehat{M}(T)^{-1}Y\right] = \mathbb{E}_{\mathbb{P}^z}\left[\mathbb{E}_{\mathbb{P}^z}\left[\widehat{M}(T)^{-1}Y|\mathcal{F}_\tau\right]\right]\\ &= \mathbb{E}_{\mathbb{P}^z}\left[Y\mathbb{E}_{\mathbb{P}^z}\left[\widehat{M}(T)^{-1}|\mathcal{F}_\tau\right]\right] = \mathbb{E}_{\mathbb{P}^z}\left[\widehat{M}(\tau)^{-1}Y\right],\end{aligned} \tag{3.20}$$

where in the last equality we applied the Optional Sampling Theorem [23, Theorem 1.3.22] to the $\mathbb{P}^z$-martingale $\{\widehat{M}(t)^{-1}\}_{0\leq t\leq T}$. We use identity (3.20) to prove that for all $t\geq 0$ and all $\mathcal{F}_{\tau+t}$-measurable and bounded random variables, Z, we have that

$$\mathbb{E}_{\mathbb{P}^z}[Z|\mathcal{F}_\tau] = \frac{1}{\widehat{M}(\tau)}\mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\widehat{M}(\tau+t)Z\middle|\mathcal{F}_\tau\right]. \tag{3.21}$$

The preceding identity gives us the analogue of [23, Lemma 3.5.3] for general stopping times, as opposed to deterministic stopping times. To see the validity of identity (3.21), it is sufficient to show that, for all sets $A\in\mathcal{F}_\tau$, we have

$$\int Z\mathbf{1}_A\,d\mathbb{P}^z = \int \frac{1}{\widehat{M}(\tau)}\mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\widehat{M}(\tau+t)Z\middle|\mathcal{F}_\tau\right]\mathbf{1}_A\,d\mathbb{P}^z. \tag{3.22}$$

Applying identity (3.20) on the right-hand side of the preceding identity, with the choice $Y := \mathbb{E}_{\widehat{\mathbb{P}}^z}[\widehat{M}(\tau+t)Z|\mathcal{F}_\tau]$, we see that

$$\int \frac{1}{\widehat{M}(\tau)}\mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\widehat{M}(\tau+t)Z\middle|\mathcal{F}_\tau\right]\mathbf{1}_A\,d\mathbb{P}^z = \int \mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\widehat{M}(\tau+t)Z\middle|\mathcal{F}_\tau\right]\mathbf{1}_A\,d\widehat{\mathbb{P}}^z,$$

and using the tower property of conditional expectation on the right-hand side, it follows that

$$\int \frac{1}{\widehat{M}(\tau)}\mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\widehat{M}(\tau+t)Z\middle|\mathcal{F}_\tau\right]\mathbf{1}_A\,d\mathbb{P}^z = \int \widehat{M}(\tau+t)Z\mathbf{1}_A\,d\widehat{\mathbb{P}}^z.$$

Another application of identity (3.20) with $Y := \widehat{M}(\tau+t)Z\mathbf{1}_A$, and τ replaced by $\tau+t$, gives us that (3.22) holds, which implies identity (3.21).

We now prove that the strong Markov property (3.19) holds. We have

$$\begin{aligned}\mathbb{P}^z\left(Z(t+\tau)\in B|\mathcal{F}_\tau\right) &= \mathbb{E}_{\mathbb{P}^z}\left[\mathbf{1}_{\{Z(\tau+t)\in B\}}\middle|\mathcal{F}_\tau\right]\\ &= \mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\frac{\widehat{M}(\tau+t)}{\widehat{M}(\tau)}\mathbf{1}_{\{\widehat{Z}(\tau+t)\in B\}}\middle|\mathcal{F}_\tau\right] \quad \text{(by (3.21))}\\ &= \mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\frac{\widehat{M}(\tau+t)}{\widehat{M}(\tau)}\mathbf{1}_{\{\widehat{Z}(\tau+t)\in B\}}\middle|\widehat{Z}(\tau)\right] \quad \text{(by Corollary 2.5).}\end{aligned}\tag{3.23}$$

The last equality holds, if for all measurable sets $A\in\mathcal{F}_\tau$, the following take place:

$$\int \mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\frac{\widehat{M}(\tau+t)}{\widehat{M}(\tau)}\mathbf{1}_{\{\widehat{Z}(\tau+t)\in B\}}\middle|\widehat{Z}(\tau)\right]\mathbf{1}_{\{Z(\tau)\in A\}}\,d\mathbb{P}^z = \int \mathbf{1}_{\{Z(\tau+t)\in B\}}\mathbf{1}_{\{Z(\tau)\in A\}}\,d\mathbb{P}^z.$$

We have that

$$\begin{aligned}&\int \mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\frac{\widehat{M}(\tau+t)}{\widehat{M}(\tau)}\mathbf{1}_{\{\widehat{Z}(\tau+t)\in B\}}\middle|\widehat{Z}(\tau)\right]\mathbf{1}_{\{Z(\tau)\in A\}}\,d\mathbb{P}^z\\ &\quad= \int \mathbb{E}_{\widehat{\mathbb{P}}^z}\left[\frac{\widehat{M}(\tau+t)}{\widehat{M}(\tau)}\mathbf{1}_{\{\widehat{Z}(\tau+t)\in B\}}\middle|\widehat{Z}(\tau)\right]\mathbf{1}_{\{Z(\tau)\in A\}}\widehat{M}(\tau)\,d\widehat{\mathbb{P}}^z \quad \text{(by (3.20))}\\ &\quad= \int \widehat{M}(\tau+t)\mathbf{1}_{\{\widehat{Z}(\tau+t)\in B\}}\mathbf{1}_{\{\widehat{Z}(\tau)\in A\}}\,d\widehat{\mathbb{P}}^z\\ &\quad= \int \mathbf{1}_{\{Z(\tau+t)\in B\}}\mathbf{1}_{\{Z(\tau)\in A\}}\,d\mathbb{P}^z \quad \text{(by (3.17)).}\end{aligned}$$

Since the preceding identity is true for all measurable sets, $A\in\mathcal{F}_\tau$, it follows that (3.23), which in turn implies that (3.19) holds. Thus, the process $\{Z(t)\}_{t\geq 0}$ satisfies the strong Markov property.

This completes the proof. □

3.2. **Uniqueness and the strong Markov property.**

Using the uniqueness in law of solutions to the generalized Kimura stochastic differential equation (1.3), we can establish uniqueness in law of solutions to the Kimura stochastic differential equation with singular drift (1.6).

Theorem 3.7 (Uniqueness in law of solutions to Kimura equation with singular drift)**.** *Suppose that the coefficients of the Kimura stochastic differential equation with singular drift* (1.6) *satisfy Assumption* 3.2. *Then, for all $z\in\bar{S}_{n,m}$, there is a unique weak solution to the stochastic differential equation* (1.6) *that satisfies the strong Markov property, with initial condition $Z(0)=z$.*

We remark that Theorem 3.7 establishes uniqueness of solutions only in the class of Markov processes. The reason for this restriction is due to our method of the proof which consists in applying Girsanov's Theorem to remove the singular drift in equation (1.6) and reduce our problem to a generalized Kimura equation (1.3), for which we know that uniqueness in law of solutions holds by Proposition 2.4. In applying Girsanov's Theorem, we need to establish the fact that the process $\{M(t)\}_{t\geq 0}$ defined in (3.31) is a martingale. As we will see in the proofs of Lemmas 3.8 and 3.9, this requires us to assume that the solution to the singular Kimura equation (1.6) satisfies the Markov property. This is the reason why our method of the proof yields uniqueness of solutions only in the class of Markov processes. This result suffices for the analytical applications of the Kimura stochastic differential equation with singular drift, (1.6), to the proof of the Harnack inequality and stochastic representation formula in [14, Theorems 1.2 and 1.3]. It is very likely that adapting the ideas to prove [35, Theorem 12.2.4], the existence and uniqueness of Markov solutions to (1.6) already implies the existence and uniqueness of weak solutions. The impediment to applying [35, Theorem 12.2.4] to our framework is due to the fact that the coefficients of the singular Kimura equation (1.6) are unbounded, and so greater care is needed in establishing certain crucial compactness arguments in the proof of [35, Theorem 12.2.4].

We begin with the analogue of Lemma 3.5 for Kimura diffusions with singular drift.

Lemma 3.8 (Novikov's condition for Kimura diffusions with singular drift)**.** *Suppose that the coefficients of the Kimura stochastic differential equation with singular drift* (1.6) *satisfy Assumption* 3.2*. Let $q \in (0, q_0)$, where the positive constant q_0 is given by* (3.1)*. Then for all $\Lambda > 0$ and $T > 0$, we have*

$$\sup_{z\in\bar{S}_{n,m}} \mathbb{E}_{\mathbb{P}^z}\left[\exp\left(\Lambda\int_0^T \sum_{i=1}^n |X_i(t)|^{-2q}\,dt\right)\right] < \infty, \tag{3.24}$$

where $\{Z(t) = (X(t), Y(t))\}_{t\geq 0}$ is a solution to the singular Kimura stochastic differential equation (1.6) *that satisfies the Markov property, with initial condition $Z(0) = z$.*

We prove Lemma 3.8 with the aid of the analogue of Lemma 3.6 for the Kimura stochastic differential equation with singular drift.

Lemma 3.9 (Verification of the hypotheses of Khas'minskii's Lemma for singular Kimura diffusions)**.** *Suppose that the coefficients of the Kimura stochastic differential equation with singular drift* (1.6) *satisfy Assumption* 3.2*. Let $q \in (0, q_0)$, where the positive constant q_0 is given by* (3.1)*. Then for all positive constants, $\delta \in (0,1)$ and Λ, there is a positive constant, $T = T(b_0, \|b\|_{C^\alpha_{\mathrm{WF}}(\bar{S}_{n,m})}, \delta, K_0, K, \Lambda, m, n, q)$, such that*

$$\sup_{z\in\bar{S}_{n,m}} \mathbb{E}_{\mathbb{P}^z}\left[\Lambda\int_0^T \sum_{i=1}^n |X_i(t)|^{-2q}\,dt\right] < \delta, \tag{3.25}$$

where $\{Z(t) = (X(t), Y(t))\}_{t\geq 0}$ is a solution to the singular Kimura stochastic differential equation (1.6)*, with initial condition $Z(0) = z$.*

Proof. The proof of Lemma 3.9 is similar to that of Lemma 3.6, but we have to pay closer attention to the singular component of the drift coefficient in the stochastic differential equation (1.6). We let the positive constants ρ, r, and the cut-off function φ be as in the proof of Lemma 3.6. Without loss of generality, we may assume that $\Lambda = 1$. We consider the auxiliary function

$$\psi(x) = \sum_{i=1}^n x_i^{1-2q}\varphi(x_i), \quad \forall\, x \in \mathbb{R}^n_+.$$

For all $\varepsilon \in (0,1)$, we recall that we denote $X^\varepsilon_i(t) := X_i(t)+\varepsilon$, and $x^\varepsilon_i := x_i+\varepsilon$. Applying Itô's rule to the process $\{\psi(X^\varepsilon(t))\}_{t\geq 0}$, we obtain, for all $T>0$ and $\varepsilon>0$,

(3.26)

$$\begin{aligned}\psi(X^\varepsilon(T)) &= \psi(x^\varepsilon) + \sum_{i=1}^n \int_0^T b_i(Z(t))\varphi'(X^\varepsilon_i(t))(X^\varepsilon_i(t))^{1-2q}\,dt\\ &+\sum_{i=1}^n \int_0^T \frac{|\sigma_i(Z(t))|^2}{2} X_i(t)(X^\varepsilon_i(t))^{-2q}\left(X^\varepsilon_i(t)\varphi''(X^\varepsilon_i(t))+2(1-2q)\varphi'(X^\varepsilon_i(t))\right)dt\\ &+(1-2q)\sum_{i=1}^n \int_0^T \varphi(X^\varepsilon_i(t))(X^\varepsilon_i(t))^{-2q}\left(b_i(Z(t))-q|\sigma_i(Z(t))|^2\frac{X_i(t)}{X^\varepsilon_i(t)}\right)dt\\ &+(1-2q)\sum_{i=1}^n \int_0^T \varphi(X^\varepsilon_i(t))(X^\varepsilon_i(t))^{-2q}\sqrt{X_i(t)}\sum_{j=1}^n f_{i,j}(Z(t))h_{i,j}(X_j(t))\,dt\\ &+\sum_{i=1}^n \int_0^T (X^\varepsilon_i(t))^{1-2q}\varphi'(X^\varepsilon_i(t))\sqrt{X_i(t)}\sum_{j=1}^n f_{i,j}(Z(t))h_{i,j}(X_j(t))\,dt\\ &+\sum_{i=1}^n \int_0^T (X^\varepsilon_i(t))^{-2q}\left(X^\varepsilon_i(t)\varphi'(X^\varepsilon_i(t))+(1-2q)\varphi(X^\varepsilon_i(t))\right)\\ &\qquad\times\sqrt{X_i(t)}\sum_{k=1}^{n+m}\sigma_{i,k}(Z(t))\,dW_k(t),\end{aligned}$$

where we recall that $\sigma_i(z)$ denotes the i-th row of the matrix $(\sigma(z))$. By Lemma 2.8, the coefficient functions $b(z)$ and $(\sigma(z))$ are bounded, and using the properties of the cut-off function φ, there is a positive constant, $C=C(K,m,n)$, such that

(3.27)

$$\begin{aligned}&\psi(x^\varepsilon) + \sum_{i=1}^n \int_0^T b_i(Z(t))\varphi'(X^\varepsilon_i(t))(X^\varepsilon_i(t))^{1-2q}\,dt\\ &+\sum_{i=1}^n \int_0^T \frac{|\sigma_i(Z(t))|^2}{2} X_i(t)(X^\varepsilon_i(t))^{-2q}\left(X^\varepsilon_i(t)\varphi''(X^\varepsilon_i(t))+2(1-2q)\varphi'(X^\varepsilon_i(t))\right)dt\\ &\leq nr^{1-2q}+Cr^{-2q}T,\quad \forall\,\varepsilon>0.\end{aligned}$$

Inequality (3.11) applied with $\widehat{X}^\varepsilon_i(t)$ replaced by $X^\varepsilon_i(t)$, together with (3.26), (3.27), and the fact that the cut-off function φ has support in $[0,r]$, yields

$$\begin{aligned}C_0\sum_{i=1}^n\int_0^T |X^\varepsilon_i(t)|^{-q}\,\mathbf{1}_{\{X_i(t)\in[0,r/2]\}}\,dt &\leq 2nr^{1-2q}+Cr^{-2q}T\\ &+r^{-2q}\sum_{i=1}^n\int_0^T\sqrt{X_i(t)}\mathbf{1}_{\{X_i(t)\in[0,r]\}}\left|\sum_{j=1}^n f_{i,j}(Z(t))h_{i,j}(X_j(t))\right|dt\\ &-\sum_{i=1}^n\int_0^T\sqrt{X_i(t)}(X^\varepsilon_i(t))^{-2q}\left(X^\varepsilon_i(t)\varphi'(X^\varepsilon_i(t))+(1-2q)\varphi(X^\varepsilon_i(t))\right)\\ &\qquad\times\sum_{k=1}^{n+m}\sigma_{i,k}(Z(t))\,dW_k(t),\end{aligned}\tag{3.28}$$

for all $\varepsilon > 0$, where we denote for brevity

$$C_0 := (1-2q)\left(\frac{b_0}{1+\rho} - q(n+m)K^2\right).$$

Note that the dt-integral term on the right-hand side of (3.28) is finite, from our assumption that the process $\{Z(t)\}_{t\geq 0}$ is a weak solution to equation (1.6), which implies that

$$\int_0^T \sqrt{X_i(t)}\left|\sum_{j=1}^n f_{i,j}(Z(t))h_{i,j}(X_j(t))\right| dt < \infty, \quad \mathbb{P}^z\text{-a.s.}$$

Moreover using the fact that $q < 1/4$ from (3.1), we see that $\sqrt{x_i}(x_i^\varepsilon)^{-2q}$ is bounded as $\varepsilon \downarrow 0$. From Lemma 2.8, it follows that the matrix coefficient $(\sigma(z))$ is bounded, and so, the integrand of the $dW(t)$-term on the right-hand side of inequality (3.28) is uniformly bounded, for all $\varepsilon \in (0,1)$. We can now apply [31, Theorem IV.2.32] to conclude that the $dW(t)$-integral on the right-hand side of (3.28) converges in probability, as ε tends to 0, to the corresponding expression with $\varepsilon = 0$. Inequality (3.28) becomes, as $\varepsilon \downarrow 0$,

$$\begin{aligned}
C_0\sum_{i=1}^n \int_0^T |X_i(t)|^{-2q}\, \mathbf{1}_{\{X_i(t)\in[0,r/2]\}}\, dt &\leq 2nr^{1-2q} + Cr^{-2q}T \\
&+ r^{-2q}\sum_{i=1}^n \int_0^T \sqrt{X_i(t)}\mathbf{1}_{\{X_i(t)\in[0,r]\}}\left|\sum_{j=1}^n f_{i,j}(Z(t))h_{i,j}(X_j(t))\right| dt \\
&- \sum_{i=1}^n \int_0^T \sqrt{X_i(t)}(X_i(t))^{-2q}\left(X_i(t)\varphi'(X_i(t)) + (1-2q)\varphi(X_i(t))\right) \\
&\times \sum_{k=1}^{n+m} \sigma_{i,k}(Z(t))\, dW_k(t).
\end{aligned}$$

This implies that the integral on the left-hand side of the preceding inequality is finite. We may now use the upper bounds (2.20), (3.3), and (3.5), to conclude that there is a positive constant, $C = C(K_0, K, m, n)$, such that

$$\begin{aligned}
\sum_{i=1}^n \int_0^T |X_i(t)|^{-2q}\, \mathbf{1}_{\{X_i(t)\in[0,r/2]\}}\, dt &\leq \frac{2nr^{1-2q} + Cr^{-2q}T}{C_0} \\
&+ \frac{Cr^{1/2-2q}}{C_0}\sum_{i=1}^n \int_0^T |X_i(t)|^{-q}\mathbf{1}_{\{X_i(t)\in[0,r/2]\}}\, dt \\
&- \frac{1}{C_0}\sum_{i=1}^n \int_0^T \sqrt{X_i(t)}(X_i(t))^{-2q}\left(X_i(t)\varphi'(X_i(t)) + (1-2q)\varphi(X_i(t))\right) \\
&\times \sum_{k=1}^{n+m} \sigma_{i,k}(Z(t))\, dW_k(t).
\end{aligned}$$

Because we choose $q < 1/4$ by (3.1), we may choose a positive constant, $r_1 = r_1(C_0, C, q)$, small enough so that $Cr_1^{1/2-2q}/C_0 \leq 1/2$. The preceding inequality gives us that

$$\begin{aligned}\sum_{i=1}^n \int_0^T |X_i(t)|^{-2q} \mathbf{1}_{\{X_i(t)\in[0,r/2]\}}\, dt &\leq \frac{4nr^{1-2q} + 2Cr^{-2q}T}{C_0} \\ &\quad - \frac{2}{C_0} \sum_{i=1}^n \int_0^T \sqrt{X_i(t)}(X_i(t))^{-2q} \left(X_i(t)\varphi'(X_i(t)) + (1-2q)\varphi(X_i(t))\right) \\ &\quad\quad \times \sum_{k=1}^{n+m} \sigma_{i,k}(Z(t))\, dW_k(t),\end{aligned}$$

for all $r \in (0, r_1)$. Because the $dW(t)$-term in the preceding equality defines a martingale, we obtain, for all $T > 0$ and $r \in (0, r_1)$,

$$\mathbb{E}_{\mathbb{P}^z}\left[\int_0^T \sum_{i=1}^n |X_i(t)|^{-2q} \mathbf{1}_{\{X_i(t)\in[0,r/2]\}}\, dt\right] \leq \frac{4nr^{1-2q} + 2Cr^{-2q}T}{C_0}.$$

Removing the indicator function in the preceding inequality, we obtain

$$\mathbb{E}_{\mathbb{P}^z}\left[\int_0^T \sum_{i=1}^n |X_i(t)|^{-2q}\, dt\right] \leq \frac{4nr^{1-2q} + 2Cr^{-2q}T}{C_0} + nr^{-2q}T.$$

For all $\delta \in (0,1)$, we may now choose positive constants

$$r = r(b_0, \|b\|_{C^\alpha_{\mathrm{WF}}(\bar{S}_{n,m})}, \delta, K_0, K, m, n, q)$$

and

$$T = T(b_0, \|b\|_{C^\alpha_{\mathrm{WF}}(\bar{S}_{n,m})}, \delta, K_0, K, m, n, q)$$

small enough so that inequality (3.25) holds. This completes the proof. □

We can now give the

Proof of Lemma 3.8. The proof of Lemma 3.8 is identical to that of Lemma 3.5, only in place of Lemma 3.6 we use Lemma 3.9, and so, we omit the detailed proof. □

To prove Theorem 3.7, in addition to Lemma 3.8, we need the following result which proves uniqueness in law of the joint probability distributions of any weak solution $\{(\widehat{Z}(t), \widehat{W}(t))\}_{t\geq 0}$, to the generalized Kimura stochastic differential equation, (1.3).

Lemma 3.10 (Uniqueness of the joint law of weak solutions $(\widehat{Z}, \widehat{W})$ to the generalized Kimura equation)**.** *Suppose that Assumption* 2.6 *holds. Let $z \in \bar{S}_{n,m}$, and let $(\widehat{Z}^i, \widehat{W}^i)$, $(\Omega^i, \{\mathcal{F}^i_t\}_{t\geq 0}, \mathcal{F}^i, \widehat{\mathbb{P}}^z_i)$, for $i = 1, 2$, be two weak solutions to the generalized Kimura equation* (1.3)*, with initial condition $\widehat{Z}^1(0) = \widehat{Z}^2(0) = z$. Then the joint probability laws of the processes $(\widehat{Z}^i, \widehat{W}^i)$, for $i = 1, 2$, agree.*

Proof. From Proposition 2.4, it follows that the probability laws of the processes $\{\widehat{Z}^i(t)\}_{t\geq 0}$, for $i=1,2$, agree. For $i=1,2$, we consider the $(n+m)$-dimensional processes defined by

$$\begin{aligned}\widehat{N}^i_j(t) &:= \widehat{X}^i_j(t) - \widehat{X}^i_j(0) - \int_0^t \widehat{b}_j(\widehat{Z}^i(s))\,ds, \quad \forall\, j=1,\ldots,n,\\ \widehat{N}^i_{n+l}(t) &:= \widehat{Y}^i_l(t) - \widehat{Y}^i_l(0) - \int_0^t e_l(\widehat{Z}^i(s))\,ds, \quad \forall\, l=1,\ldots,m.\end{aligned} \tag{3.29}$$

Our goal is to prove that the following identity holds:

$$\widehat{W}^i(t) = \int_0^t \varsigma^{-1}(\widehat{Z}^i(s))\mathbf{1}_{\{\widehat{Z}^i(s)\in S_{n,m}\}}d\widehat{N}^i(s), \quad \forall\, t\geq 0, \tag{3.30}$$

$\widehat{\mathbb{P}}^z_i$-a.s, for $i=1,2$. Notice that on the right-hand side of the preceding identity we used the invertibility of the matrix function $(\varsigma(z))$, for all $z\in S_{n,m}$. This follows from identity (2.9), and the fact that the matrix $(\sigma(z))$ is invertible, by Lemma 3.3. Properties (3.30) and (3.29), together with the fact that the probability laws of the processes $\{\widehat{Z}^i(t)\}_{t\geq 0}$, for $i=1,2$, agree, imply that the joint probability laws of the processes $(\widehat{Z}^i, \widehat{W}^i)$, for $i=1,2$, also agree.

We now proceed to the proof of identity (3.30). From definition (2.9) of the matrix function $(\varsigma(z))$, the choice of the processes $\{\widehat{N}^i(t)\}_{t\geq 0}$, for $i=1,2$, and from equation (1.3), we see that identity (3.30) is equivalent to

$$\widehat{W}^i(t) = \int_0^t \mathbf{1}_{\{\widehat{Z}^i(s)\in S_{n,m}\}}d\widehat{W}^i(s) \quad \widehat{\mathbb{P}}^z_i\text{-a.s.}, \quad \forall\, t\geq 0, \quad i=1,2.$$

Thus identity (3.30) holds if and only if

$$\int_0^t \mathbf{1}_{\{\widehat{Z}^i(s)\in \partial S_{n,m}\}}d\widehat{W}^i(s) = 0 \quad \widehat{\mathbb{P}}^z_i\text{-a.s.}, \quad \forall\, t\geq 0, \quad i=1,2.$$

The preceding equality is equivalent to proving that

$$\mathbb{E}_{\widehat{\mathbb{P}}^z_i}\left[\int_0^t \mathbf{1}_{\{\widehat{Z}^i(s)\in \partial S_{n,m}\}}\,ds\right] = 0, \quad \forall\, t\geq 0, \quad i=1,2,$$

but this clearly holds from the fact that the quantity defined in (3.7) is finite. This completes the proof. □

We can now give the proof of Theorem 3.7 with the aid of Lemmas 3.8 and 3.10.

Proof of Theorem 3.7. Let (Z^i, W^i), $(\Omega^i, \{\mathcal{F}^i_t\}_{0\leq t\leq T}, \mathcal{F}^i, \mathbb{P}^z_i)$, be two weak solutions to the Kimura stochastic differential equation with singular drift (1.6), satisfying the initial condition $Z^i(0)=z$, for $i=1,2$, where we assume that $z\in \bar{S}_{n,m}$. Assume that the two weak solutions satisfy the Markov property. Our goal is to show that the laws of the processes $\{Z^i(t)\}_{t\in[0,T]}$, are the same, for all $T>0$, for $i=1,2$. Let $\theta: S_{n,m}\to\mathbb{R}^{n+m}$ be the vector field defined in (3.14), and recall that the function $\theta(z)$ satisfies inequality (3.16). Lemma 3.8 together with inequality (3.16) shows that condition (3.24) holds, and so, [23, Corollary 3.5.13] yields that the processes $\{M^i(t)\}_{0\leq t\leq T}$ defined by

$$M^i(t) := \exp\left(-\int_0^t \sum_{k=1}^{n+m}\theta_k(Z^i(s))\,dW^i_k(s) - \frac{1}{2}\int_0^t |\theta(Z^i(s))|^2\,ds\right), \quad \forall\, t\in[0,T], \tag{3.31}$$

are $\mathbb{P}_i^z$-martingale, for $i = 1, 2$. We can apply Girsanov's Theorem [23, Theorem 3.5.1] to construct new probability measures, $\widehat{\mathbb{P}}_i^z$, by letting

$$\frac{d\widehat{\mathbb{P}}_i^z}{d\mathbb{P}_i^z} = M^i(T), \quad i = 1, 2. \tag{3.32}$$

Then the process

$$\widehat{W}^i(t) := W^i(t) + \int_0^t \theta(Z^i(s))\, dt, \quad \forall\, t \in [0, T], \tag{3.33}$$

is an $(n+m)$-dimensional Brownian motion with respect to the probability measure $\widehat{\mathbb{P}}_i^z$, for $i = 1, 2$. Using definition (3.14) of the function $\theta(z)$, we see that by letting $\widehat{Z}^i(t) := Z^i(t)$, for all $t \in [0, T]$, we obtain that the processes $\{\widehat{Z}^i(t), \widehat{W}^i(t)\}_{0 \le t \le T}$, $(\Omega, \{\mathcal{F}_t\}_{0 \le t \le T}, \mathcal{F}, \widehat{\mathbb{P}}_i^z)$ are weak solutions to the generalized Kimura stochastic differential equation (1.3), with initial condition $\widehat{Z}^i(0) = z$, for $i = 1, 2$. From Lemma 3.10, it follows that the joint law of the processes $\{\widehat{Z}^i(t), \widehat{W}^i(t)\}_{0 \le t \le T}$, for $i = 1, 2$, agree. From definitions (3.31) and (3.33), we have that

$$M^i(t) := \exp\left(-\int_0^t \theta(\widehat{Z}^i(s)) \cdot d\widehat{W}^i(s) + \frac{1}{2}\int_0^t |\theta(\widehat{Z}^i(s))|^2\, ds\right), \quad i = 1, 2,$$

and so, the laws of the processes $\{M^1(t)\}_{0 \le t \le T}$ and $\{M^2(t)\}_{0 \le t \le T}$ also agree. Thus, it follows from (3.32), that the probability laws of the processes $\{Z^i(t)\}_{t \in [0,T]}$ are the same, for all $T > 0$, for $i = 1, 2$. This completes the proof. □

Acknowledgments

We are very grateful to the referee for the careful reading of our manuscript and for providing valuable comments that led to the improvement of our article. We would also like to thank Charles Epstein for suggesting this problem and for many very helpful discussions on degenerate equations.

References

[1] M. Aizenman and B. Simon, *Brownian motion and Harnack inequality for Schrödinger operators*, Comm. Pure Appl. Math. **35** (1982), no. 2, 209–273, DOI 10.1002/cpa.3160350206. MR644024

[2] S. R. Athreya, M. T. Barlow, R. F. Bass, and E. A. Perkins, *Degenerate stochastic differential equations and super-Markov chains*, Probab. Theory Related Fields **123** (2002), no. 4, 484–520, DOI 10.1007/s004400100191. MR1921011

[3] Richard F. Bass, *Diffusions and elliptic operators*, Probability and its Applications (New York), Springer-Verlag, New York, 1998. MR1483890

[4] Richard F. Bass and Edwin A. Perkins, *Degenerate stochastic differential equations with Hölder continuous coefficients and super-Markov chains*, Trans. Amer. Math. Soc. **355** (2003), no. 1, 373–405 (electronic), DOI 10.1090/S0002-9947-02-03120-3. MR1928092

[5] Anne-Marie Berthier and Bernard Gaveau, *Critère de convergence des fonctionnelles de Kac et application en mécanique quantique et en géométrie* (French), J. Funct. Anal. **29** (1978), no. 3, 416–424, DOI 10.1016/0022-1236(78)90039-3. MR512253

[6] Sandra Cerrai and Philippe Clément, *On a class of degenerate elliptic operators arising from Fleming-Viot processes*, J. Evol. Equ. **1** (2001), no. 3, 243–276, DOI 10.1007/PL00001370. Dedicated to Ralph S. Phillips. MR1861222

[7] Sandra Cerrai and Philippe Clément, *Schauder estimates for a class of second order elliptic operators on a cube*, Bull. Sci. Math. **127** (2003), no. 8, 669–688, DOI 10.1016/S0007-4497(03)00058-7. MR2014753

[8] Sandra Cerrai and Philippe Clément, *Well-posedness of the martingale problem for some degenerate diffusion processes occurring in dynamics of populations*, Bull. Sci. Math. **128** (2004), no. 5, 355–389, DOI 10.1016/j.bulsci.2004.03.004. MR2066345

[9] Sandra Cerrai and Philippe Clément, *Schauder estimates for a degenerate second order elliptic operator on a cube*, J. Differential Equations **242** (2007), no. 2, 287–321, DOI 10.1016/j.jde.2007.08.002. MR2363317

[10] C. L. Epstein and R. Mazzeo, *C^0-estimates for diagonal degenerate diffusion operators arising in population biology*, pp. 19, preprint.

[11] Charles L. Epstein and Rafe Mazzeo, *Harnack inequalities and heat kernel estimates for degenerate diffusion operators arising in population biology*, Appl. Math. Res. Express. AMRX **2** (2016), 217–280, DOI 10.1093/amrx/abw002. MR3551776

[12] Charles L. Epstein and Rafe Mazzeo, *Wright-Fisher diffusion in one dimension*, SIAM J. Math. Anal. **42** (2010), no. 2, 568–608, DOI 10.1137/090766152. MR2607922

[13] Charles L. Epstein and Rafe Mazzeo, *Degenerate diffusion operators arising in population biology*, Annals of Mathematics Studies, vol. 185, Princeton University Press, Princeton, NJ, 2013. MR3202406

[14] C. L. Epstein and C. A. Pop, *Harnack inequalities for degenerate diffusions*, pp. 57, submitted, arXiv:1406.4759.

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

[16] S. N. Ethier, *Limit theorems for absorption times of genetic models*, Ann. Probab. **7** (1979), no. 4, 622–638. MR537210

[17] S. N. Ethier, *An error estimate for the diffusion approximation in population genetics*, Ph.D. thesis, Wisconsin Univ., 1985.

[18] Stewart N. Ethier and Thomas G. Kurtz, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1986. Characterization and convergence. MR838085

[19] Paul M. N. Feehan and Camelia A. Pop, *On the martingale problem for degenerate-parabolic partial differential operators with unbounded coefficients and a mimicking theorem for Itô processes*, Trans. Amer. Math. Soc. **367** (2015), no. 11, 7565–7593, DOI 10.1090/tran/6243. MR3391893

[20] William Feller, *Diffusion processes in genetics*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950, University of California Press, Berkeley and Los Angeles, 1951, pp. 227–246. MR0046022

[21] R. A. Fisher, *On the dominance ratio*, Proc. Roy. Soc. Edin. **42** (1922), 321–431.

[22] R. A. Fisher, *The distributions of gene ratios of rare mutations*, Proc. Roy. Soc. Edin. **50** (1930), 205–220.

[23] Ioannis Karatzas and Steven E. Shreve, *Brownian motion and stochastic calculus*, 2nd ed., Graduate Texts in Mathematics, vol. 113, Springer-Verlag, New York, 1991. MR1121940

[24] S. Karlin and Taylor, *A second course on stochastic processes*, Academic, New York, 1981.

[25] R. Z. Has′minskiĭ, *On positive solutions of the equation $\mathcal{U} + Vu = 0$*, Theor. Probability Appl. **4** (1959), 309–318. MR0123373

[26] Motoo Kimura, *Some problems of stochastic processes in genetics*, Ann. Math. Statist. **28** (1957), 882–901. MR0094267

[27] Motoo Kimura, *Diffusion models in population genetics*, J. Appl. Probability **1** (1964), 177–232. MR0172727

[28] N. V. Krylov, *Lectures on elliptic and parabolic equations in Hölder spaces*, Graduate Studies in Mathematics, vol. 12, American Mathematical Society, Providence, RI, 1996. MR1406091

[29] Camelia A. Pop, *C^0-estimates and smoothness of solutions to the parabolic equation defined by Kimura operators*, J. Funct. Anal. **272** (2017), no. 1, 47–82, DOI 10.1016/j.jfa.2016.10.014. MR3567501

[30] N. I. Portenko, *Diffusion processes with an unbounded drift coefficient* (Russian, with English summary), Teor. Verojatnost. i Primenen. **20** (1975), 29–39. MR0375483

[31] Philip E. Protter, *Stochastic integration and differential equations*, Stochastic Modelling and Applied Probability, vol. 21, Springer-Verlag, Berlin, 2005. Second edition. Version 2.1; Corrected third printing. MR2273672

[32] Ken-iti Sato, *Diffusion processes and a class of Markov chains related to population genetics*, Osaka J. Math. **13** (1976), no. 3, 631–659. MR0445627

[33] Ken-iti Sato, *Diffusion operators in population genetics and convergence of Markov chains*, Measure theory applications to stochastic analysis (Proc. Conf., Res. Inst. Math., Oberwolfach, 1977), Lecture Notes in Math., vol. 695, Springer, Berlin, 1978, pp. 127–137. MR527079

[34] Norio Shimakura, *Formulas for diffusion approximations of some gene frequency models*, J. Math. Kyoto Univ. **21** (1981), no. 1, 19–45. MR606311

[35] Daniel W. Stroock and S. R. Srinivasa Varadhan, *Multidimensional diffusion processes*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 233, Springer-Verlag, Berlin-New York, 1979. MR532498

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

School of Mathematics, University of Minnesota, 127 Vincent Hall, 206 Church Street SE, Minneapolis, Minnesota 55455

E-mail address: capop@umn.edu