Probab. Theory Relat. Fields 118, 574–594 (2000)
Digital Object Identifier (DOI) 10.1007/s004400000100

J. M. Swart

Clustering of linearly interacting diffusions and universality of their long-time limit distribution

Received: 10 May 1999 / Revised version: 18 April 2000 /
Published online: 22 November 2000 – © Springer-Verlag 2000

Abstract. Let $K \subset \mathbb{R}^d$ $(d \geq 1)$ be a compact convex set and Λ a countable Abelian group. We study a stochastic process X in K^Λ, equipped with the product topology, where each coordinate solves a SDE of the form $dX_i(t) = \sum_j a(j-i)(X_j(t)-X_i(t))dt + \sigma(X_i(t))\mathrm{dB}_i(t)$. Here $a(\cdot)$ is the kernel of a continuous-time random walk on Λ and σ is a continuous root of a diffusion matrix w on K. If $X(t)$ converges in distribution to a limit $X(\infty)$ and the symmetrized random walk with kernel $a_S(i) = a(i) + a(-i)$ is recurrent, then each component $X_i(\infty)$ is concentrated on $\{x \in K : \sigma(x) = 0\}$ and the coordinates agree, i.e., the system *clusters*. Both these statements fail if a_S is transient. Under the assumption that the class of harmonic functions of the diffusion matrix w is preserved under linear transformations of K, we show that the system clusters for all spatially ergodic initial conditions and we determine the limit distribution of the components. This distribution turns out to be universal in all recurrent kernels a_S on Abelian groups Λ.

1. Introduction

Let Λ be the set of a finite or countable Abelian group $(\Lambda, +)$ and let K be a non-empty, compact and convex subset of $\mathbb{R}^d$. We consider K^Λ-valued solutions X of the following system of stochastic differential equations:

$$dX_i^\alpha(t) = \sum_j a(j-i)(X_j^\alpha(t) - X_i^\alpha(t))dt + \sum_\beta \sigma_{\alpha\beta}(X_i(t))dB_i^\beta(t) \qquad (i \in \Lambda,\ \alpha = 1, \ldots, d,\ t \geq 0). \tag{1.1}$$

Here

$$X = (X_i)_{i\in\Lambda} = (X_i^\alpha(t))_{t\geq 0,\ i\in\Lambda}^{\alpha=1,\ldots,d} \tag{1.2}$$

and we adopt the convention that sums over Roman indices $i, j, k, \ldots$ range over Λ, while sums over Greek indices $\alpha, \beta, \gamma, \ldots$ range from 1 to d. Where both types of indices occur together, we write the Greek indices as superscipts. The $(B_i)_{i\in\Lambda}$ are independent d-dimensional Brownian motions.

J. M. Swart: TU Berlin, Sekr. MA 7-5, FB 3, Strasse des 17. Juni 136, 10623 Berlin, Germany.
e-mail: swart@math.tu-berlin.de

Mathematics Subject Classification (2000): 60K35, 82C20, 60H10, 60J60

Key words and phrases: Interacting diffusions – Clustering – Universality – Population genetics

The *interaction kernel* $a : \Lambda \to [0, \infty)$ satisfies $\sum_i a(i) < \infty$. It is the kernel of a continuous-time random walk on Λ that jumps from a point i to a point j with rate $a(j - i)$. We assume that this random walk is irreducible.

The *diffusion coefficient* σ is a continuous function from K into the real $d \times d$ matrices. The associated *diffusion matrix* $w_{\alpha\beta}(x) := \sum_\gamma \sigma_{\alpha\gamma}(x)\sigma_{\beta\gamma}(x)$ satisfies

$$\sum_{\alpha,\beta} z_\alpha w_{\alpha\beta}(x) z_\beta = 0 \quad \forall x \in K, \;\; z \in I_x^\perp, \tag{1.3}$$

where $I_x^\perp$ is the orthogonal complement of the space

$$I_x := \{y \in \mathbb{R}^d : \exists \varepsilon > 0 \text{ such that } x + \lambda y \in K \; \forall |\lambda| \le \varepsilon\}. \tag{1.4}$$

Condition (1.3) is needed to guarantee that (1.1) has a solution that stays in K^Λ.

So-called 'stepping stone models' of the form (1.1) and generalizations to non-compact or infinite-dimensional state space K have been considered by many authors. They find their origin in population biology (see Shiga [24]). For one-dimensional state space it is known that the long-time behavior of (1.1) depends on the *symmetrized random walk* on Λ that jumps from a point i to a point j with rate $a_S(j - i)$, where

$$a_S(i) := a(i) + a(-i) \qquad (i \in \Lambda). \tag{1.5}$$

If this random walk is recurrent, then for appropriate initial conditions the long-time limit $X(\infty)$ exists (in distribution) and is concentrated on $\{x \in K^\Lambda : x_i = x_j,\; w(x_i) = 0 \; \forall i, j\}$. In this case we say that the system in (1.1) clusters. On the other hand, if a_S is transient, then (1.1) does not cluster but as $t \to \infty$, $X(t)$ tends in distribution to a non-degenerate invariant measure, in which case we say that (1.1) has stable behavior. For $K = [0, 1]$ and $K = [0, \infty)$ this dichotomy between stable and clustering behavior has been established by Cox, Greven, Notohara and Shiga in [4–6, 22]. The results are also known for Wright-Fisher type diffusions in dimensions $d \ge 1$ and for infinite-dimensional Fleming-Viot diffusions (see Dawson, Greven and Vaillancourt [12]). Higher-dimensional state spaces with interaction between the components have also been considered, notably catalytic branching by Dawson, Fleischmann and Klenke in [8, 15] (for models where Λ is replaced by $\mathbb{R}^d$) and mutually catalytic branching by Dawson and Perkins in [13].

For higher-dimensional K it is often not known if solutions to (1.1) are unique. For some specific models this difficulty can be overcome, e.g. with a duality as in [13], but the general picture is not clear. We avoid this problem by proving theorems that are valid for any (weak) solution of (1.1). We prove two theorems on the long-time limit $X(\infty)$ of solutions to (1.1). The first theorem, which is based on an easy covariance calculation, shows that *if* this limit exists, then the picture with the dichotomy between clustering and stable behavior holds for all models of the form (1.1). The second theorem uses a rather restrictive assumption on the harmonic functions of the diffusion matrix w to prove in the recurrent case that the limit $X(\infty)$ indeed exists and hence the system clusters. The method also allows us to give a formula for the distribution of the $X_i(\infty)$ $(i \in \Lambda)$, which shows that this distribution does not depend on the random walk kernel a or the Abelian group Λ, as long as a_S is recurrent. Although our condition on w is rather restrictive, it seems to be satisfied by all models of the form (1.1) studied so far, with the exception of catalytic (not mutually catalytic) branching only.

2. Main results

We restrict ourselves to shift-invariant solutions of (1.1). Let K^Λ be equipped with the product topology and product-σ-field. We say that a solution X to (1.1) is shift-invariant if for each $j \in \Lambda$ the shifted process $\tilde{X}_i(t) := X_{i-j}(t)$ has the same distribution as X. Similarly we say that a probability measure μ on K^Λ is shift-invariant if it coincides with its shifts by a distance j. Since we are not assuming uniqueness for (1.1), it is a priori not clear that solutions to (1.1) with shift-invariant initial conditions are shift-invariant. We therefore show that shift-invariant solutions exist.

Proposition 2.1. *For each probability measure μ on K^Λ, there exists a weak solution $(X(t))_{t\geq 0}$ to* (1.1) *with initial condition $\mathscr{L}(X(0)) = \mu$ and sample paths in the continuous functions from $[0,\infty)$ to K^Λ. If μ is shift-invariant, then* (1.1) *has a shift-invariant solution with the same properties.*

As announced, in what follows we will not assume uniqueness of solutions to (1.1), but for the interest of the reader we mention the following result, which can be obtained by a straightforward adaptation of an argument by Shiga and Shimizu [28].

Proposition 2.2. *Assume that the function $\sigma : K \to \mathbb{R}^d \otimes \mathbb{R}^d$ is Lipschitz continuous. Then, for each K^Λ-valued initial condition $X(0)$, strong uniqueness holds for equation* (1.1).

We now formulate our first result. We use the following notation. The subset of K where the diffusion matrix w vanishes we denote by

$$\partial_w K := \{x \in K : w_{\alpha\beta}(x) = 0 \ \forall \alpha, \beta\}. \tag{2.1}$$

In typical examples, $\partial_w K$ is a subset of the (topological) boundary of K. We call $\partial_w K$ the *effective boundary* of K. We use the symbol $\Rightarrow$ for weak convergence of probability measures on K^Λ, as well as for convergence in distribution of K^Λ-valued random variables.

Theorem 1. *Let X be a shift-invariant solution to* (1.1) *and assume that there exists a K^Λ-valued random variable $X(\infty)$ such that*

$$X(t) \Rightarrow X(\infty) \qquad \text{as } t \to \infty. \tag{2.2}$$

If the random walk with kernel a_S is recurrent, then

$$\begin{array}{ll} (i) & P[X_i(\infty) \in \partial_w K \ \ \forall i \in \Lambda] = 1 \\ (ii) & P[X_i(\infty) = X_j(\infty) \ \ \forall i, j \in \Lambda] = 1. \end{array} \tag{2.3}$$

If the random walk with kernel a_S is transient, $E[X_0(0)] \notin \partial_w K$ and $\mathscr{L}(X(0))$ is spatially ergodic, then

$$\begin{array}{lll} (i) & P[X_i(\infty) \in \partial_w K] < 1 & \forall i \in \Lambda \\ (ii) & P[X_i(\infty) = X_j(\infty)] < 1 & \forall i \neq j \in \Lambda. \end{array} \tag{2.4}$$

Theorem 1 follows from a standard covariance calculation combined with compactness of the state space K. Although it seems hard to imagine a shift-invariant solution to (1.1) that does not converge as $t \to \infty$, the convergence in (2.2) is in general hard to prove.[1] For finite Λ, one may exploit the fact that $\sum_i X_i(t)$ is a bounded martingale to get (2.2), not only in the sense of weak convergence, but also in L^2-norm (see section 3.6 in Swart, [29]). For infinite Λ, convergence in L^2-norm does not generally hold.

For the statement of our second result, we need some elements of potential theory associated with the diffusion matrix w. Write $\mathscr{C}(K)$ for the continuous real functions on K and $\mathscr{C}^2(K)$ for the functions on K that can be extended to a twice continuously differentiable function on $\mathbb{R}^d$. A function $h \in \mathscr{C}^2(K)$ is called harmonic for the diffusion matrix w, in short w-*harmonic*, if

$$\sum_{\alpha,\beta} w_{\alpha\beta}(x) \frac{\partial^2}{\partial x_\alpha \partial x_\beta} h(x) = 0 \qquad (x \in K). \tag{2.5}$$

More generally,[2] a function $h \in \mathscr{C}(K)$ is called w-harmonic if it is of the form

$$h(x) = E[\phi(Z^x(\infty))] \qquad (x \in K), \tag{2.6}$$

where $\phi \in \mathscr{C}(K)$ and the $(Z^x(t))_{t \geq 0}$ are K-valued solutions of the equation

$$dZ^x_\alpha(t) = \sum_\beta \sigma_{\alpha\beta}(Z^x(t)) dB_\beta(t) \qquad (t \geq 0,\ \alpha = 1, \ldots, d), \tag{2.7}$$

with initial condition $Z^x(0) = x$. Solutions to (2.7) are bounded martingales and hence there exists a $Z^x(\infty)$ such that

$$Z^x(t) \to Z^x(\infty) \quad \text{a.s.} \quad \text{as } t \to \infty. \tag{2.8}$$

Note that (2.7) is a special case of (1.1) where Λ consists of just one point. We denote the class of continuous w-harmonic functions by $H(w)$. We introduce the following assumptions on w.[3]

[1] It is not even known if (2.3) and (2.4) hold for any weak limit point of $(X(t))_{t\geq 0}$ as $t \to \infty$. In particular, it cannot be excluded that the process X spends most time as $t \to \infty$ in a situation near (2.3) or (2.4), but makes deviations at some rare deterministic times.

[2] Indeed, if solutions to (2.7) are weakly unique, then any $\mathscr{C}^2$-function of the form (2.6) solves (2.5).

[3] Condition (A.2) implies that the Dirichlet problem for w has a continuous solution for all continuous boundary conditions on $\partial_w K$. Not all w enjoy this property. A counterexample

(A.1) Weak uniqueness holds for equation (2.7).
(A.2) For each $\phi \in \mathscr{C}(K)$, the function $x \mapsto E[\phi(Z^x(\infty))]$ is continuous on K.
(A.3) $H(w)$ is contained in the bp-closure of $H(w) \cap \mathscr{C}^2(K)$.

We now formulate our main condition on w. For any $\theta \in K$, $t \geq 0$ and $f \in \mathscr{C}(K)$ we define $T_{\theta,t} f \in \mathscr{C}(K)$ by

$$(T_{\theta,t} f)(x) := f(\theta + (x - \theta)e^{-t}). \tag{2.9}$$

Note that $(T_{\theta,t})_{t \geq 0}$ is the semigroup associated with the differential equation

$$dY_\alpha(t) = (\theta_\alpha - Y_\alpha(t))dt \qquad (t \geq 0,\ \alpha = 1, \ldots, d). \tag{2.10}$$

Definition 2.3. *We say that the diffusion matrix w has invariant harmonics if*

$$T_{\theta,t}(H(w)) \subset H(w) \text{ for all } \theta \in K,\ t \geq 0. \tag{2.11}$$

Condition (2.11) guarantees that under the evolution of equation (1.1), w-harmonic functions 'do not feel the diffusion terms', in the following sense:

Lemma 2.4. *Assume that w has invariant harmonics and satisfies* (A.1) *and* (A.3). *Let X be a solution to* (1.1) *and let Y be a solution to*

$$dY_i^\alpha(t) = \sum_j a(j - i)(Y_j^\alpha(t) - Y_i^\alpha(t))dt \qquad (i \in \Lambda,\ \alpha = 1, \ldots, d,\ t \geq 0), \tag{2.12}$$

with initial condition $\mathscr{L}(Y(0)) = \mathscr{L}(X(0))$. Then

$$E[h(X_i(t))] = E[h(Y_i(t))] \qquad \forall h \in H(w),\ i \in \Lambda,\ t \geq 0. \tag{2.13}$$

Remark. While we assume weak uniqueness for equation (2.7) here, we do not need uniqueness for equation (1.1). Uniqueness for (2.7) can often be proved under much milder conditions on w than are needed for (1.1), as will be clear from the examples given below.

We now formulate our main result.

Theorem 2. *Assume that w has invariant harmonics and satisfies* (A.1), (A.2) *and* (A.3). *Let X be a shift-invariant solution to* (1.1) *with a spatially ergodic initial condition with intensity $E[X_i(0)] = \theta$ ($i \in \Lambda$). If the random walk with kernel a_S is recurrent, then*

$$\mathscr{L}(X(t)) \Rightarrow \int_K \Gamma_\theta(dx)\delta_{\underline{x}} \qquad \text{as } t \to \infty. \tag{2.14}$$

where $\underline{x} \in K^\Lambda$ is given by $\underline{x}_i = x$ ($i \in \Lambda$), and Γ_θ is the harmonic measure with mean θ associated with w.

---

is the 'punctured ball' $K = \{x \in \mathbb{R}^2 : |x| \leq 1\}$ and $w_{\alpha\beta}(x) = \delta_{\alpha\beta} g(x)$, where $g(x) = 0$ if $|x| = 0, 1$ and $g(x) > 0$ otherwise. The bp-closure of a class of functions is the smallest set containing the class that is closed under bounded pointwise limits of sequences. The author does not know examples of w violating (A.3). In our main cases of interest, (A.3) is easily verified.

Remark. Γ_x is the distribution of the random variable $Z^x(\infty)$ in (2.8). The class of harmonic functions $H(w)$ can be expressed in terms of the harmonic measures $(\Gamma_x)_{x\in K}$ by (2.6), and hence Definition 2.3 can be reformulated as a condition on the harmonic measures of w.

The proof of Theorem 2 is an adaptation of the 'duality comparison technique' of Cox and Greven [3, 4]. They use their technique to make a comparison between systems with general diffusion matrices on $K = [0, 1]$ and systems of interacting Wright–Fisher diffusions, for which all moments are given by a duality with coalescing random walks. In our case, the role of the Wright–Fisher diffusion matrix is played by the diffusion matrix w^*, defined as

$$w^*_{\alpha\beta}(x) := \int_K \Gamma_x(dy)(y_\alpha - x_\alpha)(y_\beta - x_\beta) \qquad (x \in K,\ \alpha, \beta = 1, \ldots, d). \tag{2.15}$$

When w has invariant harmonics and satisfies (A.1), (A.2) and (A.3), then w^* is continuous in x and satisfies (1.3). By Proposition 2.1 there exist solutions of (1.1) for any continuous root σ of w^*. We do not have a duality for such systems, but we can find an expression for the time evolution of harmonic functions and second order moments, which is sufficient for our purposes. In fact, our methods yield the following.

Proposition 2.5. *Assume that w has invariant harmonics and satisfies* (A.1), (A.2) *and* (A.3) *and let w^* be as in* (2.15)*. Let X be a solution of* (1.1) *for a continuous root of λw^** ($\lambda \geq 0$)*, with initial condition $X_i(0) = \theta$* ($i \in \Lambda$)*. Then, for all $i, j \in \Lambda,\ \alpha, \beta = 1, \ldots, d$ and $t \geq 0$*

$$E[(X^\alpha_i(t) - \theta_\alpha)(X^\beta_j(t) - \theta_\beta)] = w^*_{\alpha\beta}(\theta) K^\lambda_t(i - j), \tag{2.16}$$

where $K^\lambda_t(i-j)$ denotes the probability that two delayed coalescing random walks, each with kernel a, starting in points i and j, respectively, and coalescing with rate 2λ, have coalesced before time t.

3. Examples

We give examples of diffusion matrices w satisfying the assumptions in Theorem 2. Although w^* depends on w, different w may share the same w^*. The following example describes the class of models for which w^* is the Wright–Fisher diffusion matrix.

Example 3.1. (Wright–Fisher-class models). *Assume that K is the d-dimensional simplex*

$$K_d = \{x \in \mathbb{R}^d : x_\alpha \geq 0\ \forall \alpha = 1, \ldots, d,\ \sum_\alpha x_\alpha \leq 1\}, \tag{3.1}$$

and that $x \mapsto w(x)$ is Lipschitz continuous and satisfies (*compare* (1.3))

$$\sum_{\alpha,\beta} z_\alpha w_{\alpha\beta}(x) z_\beta = 0 \quad \Leftrightarrow \quad z \in I_x^\perp \qquad (x \in K). \tag{3.2}$$

Then w has invariant harmonics and satisfies (A.1), (A.2) *and* (A.3). *The class $H(w)$ consists of all affine functions*

$$x \mapsto a + \sum_{\alpha} b_\alpha x_\alpha \qquad (a, b_1, \ldots, b_d \in \mathbb{R}), \tag{3.3}$$

and w^* is the Wright-Fisher diffusion matrix:

$$w^*_{\alpha\beta}(x) = x_\alpha(\delta_{\alpha\beta} - x_\beta) \qquad (x \in K,\ \alpha, \beta = 1, \ldots, d). \tag{3.4}$$

Note that in particular, weak uniqueness holds for equation (2.7) for all w satisfying the requirements above. It is not known if solutions to (1.1) are unique for such w.[4]

Since $H(w)$ consists only of affine functions, Lemma 2.4 is trivial in the preceding example. In the following example it is really needed.

Example 3.2. (isotropic models). *Assume that K has non-empty interior K°, and let $\partial K := K \backslash K^\circ$ denote its topological boundary. Assume that*

$$w_{\alpha\beta}(x) = \delta_{\alpha\beta} g(x) \qquad (x \in K,\ \alpha, \beta = 1, \ldots, d) \tag{3.5}$$

for some Lipschitz continuous function $g : K \to [0, \infty)$ satisfying

$$g(x) = 0 \quad \Leftrightarrow \quad x \in \partial K. \tag{3.6}$$

Then w has invariant harmonics and satisfies (A.1), (A.2) *and* (A.3). *The class $H(w)$ is given by*

$$H = \Big\{ h \in \mathscr{C}(K) : h|_{K^\circ} \in \mathscr{C}^2(K^\circ),\ \sum_{\alpha} \frac{\partial^2}{\partial x_\alpha{}^2} h(x) = 0 \text{ on } K^\circ \Big\}, \tag{3.7}$$

and w^* equals

$$w^*_{\alpha\beta}(x) = \delta_{\alpha\beta} g^*(x) \qquad (x \in K,\ \alpha, \beta = 1, \ldots, d), \tag{3.8}$$

where $g^* \in \{g \in \mathscr{C}(K) : g|_{K^\circ} \in \mathscr{C}^2(K^\circ)\}$ is the unique solution of

$$\begin{aligned} -\tfrac{1}{2} \sum_{\alpha} \frac{\partial^2}{\partial x_\alpha{}^2} g^*(x) &= 1 \qquad (x \in K^\circ) \\ g^*(x) &= 0 \qquad (x \in \partial K). \end{aligned} \tag{3.9}$$

Again, weak uniqueness is known in this case for (2.7) but not for (1.1).

The author has found a few more examples of diffusion matrices satisfying the assumptions in Theorem 2, but in all these examples the system in (1.1) can, through a linear transformation of the state space K, be reduced to a number of independent copies of Wright-Fisher-class models and isotropic models. (Note that a linear transformation of K leaves the drift term in (1.1) invariant.) The author does not know if all models satisfying the assumptions in Theorem 2 are of this form, but it seems that this could be the case.

[4] This problem has nothing to do with the fact that Λ is infinite. In fact, it exists even for simple equations of the form $dZ_\alpha(t) = c(\theta_\alpha - Z_\alpha(t))dt + \sum_\beta \sigma_{\alpha\beta}(Z(t))dB_\beta(t)$, with $\theta \in K$ and $c > 0$ in dimensions $d \geq 2$, when c and θ are such that (a part of) ∂K is accessible.

Since it is not always obvious if a model fits into one of the examples above, we give two more concrete examples and one counterexample.

A mutually catalytic Wright–Fisher model. This is one way to define an analogue of the mutually catalytic branching model of Dawson and Perkins in [13] for compact state space. Consider two populations, X^1 and X^2, taking values in $[0,1]^\Lambda$, where the diffusion rate of each population is a function of the other population, in the following way

$$\begin{aligned} dX_i^1(t) &= \sum_j a(j-i)(X_j^1(t) - X_i^1(t))dt \\ &\quad + \sqrt{X_i^1(t)(1-X_i^1(t))X_i^2(t)(1-X_i^2(t))}dB_i^1(t) \\ dX_i^2(t) &= \sum_j a(j-i)(X_j^2(t) - X_i^2(t))dt \\ &\quad + \sqrt{X_i^1(t)(1-X_i^1(t))X_i^2(t)(1-X_i^2(t))}dB_i^2(t). \end{aligned} \tag{3.10}$$

This corresponds to $w_{\alpha\beta}(x) = \delta_{\alpha\beta}x_1(1-x_1)x_2(1-x_2)$ and hence this is an isotropic model in the sense of Example 3.2. We note that the diffusion matrix w in this model is not a scalar multiple of w^*; therefore Proposition 2.5 cannot be applied to give an expression for second order moments of solutions of (3.10).

A three-type interaction model. Consider three populations X^1, X^2, X^3, subject to the conditions $X_i^\alpha(t) \geq 0$, $\sum_{\alpha=1}^3 X_i^\alpha(t) = 1$ $(i \in \Lambda,\ t \geq 0)$, solving a system of the form

$$\begin{aligned} dX_i^1(t) &= \sum_j a(j-i)(X_j^1(t) - X_i^1(t))dt \\ &\quad + \sqrt{X_i^1(t)X_i^2(t)X_i^3(t)}\left(\tfrac{2}{3}dB_i^1(t) - \tfrac{1}{3}dB_i^2(t) - \tfrac{1}{3}dB_i^3(t)\right) \\ dX_i^2(t) &= \sum_j a(j-i)(X_j^2(t) - X_i^2(t))dt \\ &\quad + \sqrt{X_i^1(t)X_i^2(t)X_i^3(t)}\left(\tfrac{2}{3}dB_i^2(t) - \tfrac{1}{3}dB_i^1(t) - \tfrac{1}{3}dB_i^3(t)\right) \\ dX_i^3(t) &= \sum_j a(j-i)(X_j^3(t) - X_i^3(t))dt \\ &\quad + \sqrt{X_i^1(t)X_i^2(t)X_i^3(t)}\left(\tfrac{2}{3}dB_i^3(t) - \tfrac{1}{3}dB_i^1(t) - \tfrac{1}{3}dB_i^2(t)\right). \end{aligned} \tag{3.11}$$

This model arises if pair resampling in the Wright-Fisher model is replaced by resampling of triples. Its diffusion matrix is $w_{\alpha\beta}(x) = (\delta_{\alpha\beta} - \frac{1}{3})x_1x_2x_3$, which can be transformed into an isotropic model in the sense of Example 3.2 through a linear transformation of the state space. A small calculation shows that w^* is a scalar multiple of w, and hence for this model Proposition 2.5 gives an explicit expression for second order moments. Higher order moments are unknown, however, and finding a duality for the model seems difficult.[5]

Catalytic Wright-Fisher diffusions. Consider two populations, X^1 and X^2, taking values in $[0,1]^\Lambda$, where the diffusion rate of the second population is proportional to the size of the first population, in the following way

$$\begin{aligned} dX_i^1(t) &= \sum_j a(j-i)(X_j^1(t) - X_i^1(t))dt + \sqrt{X_i^1(t)(1-X_i^1(t))}dB_i^1(t) \\ dX_i^2(t) &= \sum_j a(j-i)(X_j^2(t) - X_i^2(t))dt + \sqrt{X_i^1(t)X_i^2(t)(1-X_i^2(t))}dB_i^2(t). \end{aligned} \tag{3.12}$$

This corresponds to $w_{11}(x) = x_1(1-x_1)$, $w_{22}(x) = x_1x_2(1-x_2)$ and $w_{12}(x) = w_{21}(x) = 0$. With a little effort it is possible to show that formula (2.11) in Definition 2.3 is violated, i.e. w does not have invariant harmonics, in our terminology. It seems that at present nobody can treat the long-time behavior of (3.12), but Greven, Klenke and Wakolbinger [16] treat a similar case where X^1 is replaced by the voter model. They show that the model clusters[6] when a is nearest neighbor on $\Lambda = \mathbb{Z}^d$ in the recurrent dimensions $d = 1, 2$. However, the long-time limit distribution of the components depends on the dimension. Such dimension-dependent behavior in the clustering regime has been found for catalytic branching (see [8, 15]) and it can also be conjectured for (3.12).

Note that this behavior differs fundamentally from that of systems where w has invariant harmonics. Under the conditions of Theorem 2, the distribution of $X_i(\infty)$ ($i \in \Lambda$) does not depend on the choice of the Abelian group Λ, nor on the interaction a, as long as a_S is recurrent. Moreover, it coincides for all w which have the same w-harmonic functions, as occurs for the classes of diffusion matrices described in Examples 3.1 and 3.2. Thus, we may say that Theorem 2 describes a situation in which the systems exhibits 'large time-scale universality', where we borrow a term from renormalization theory. For a renormalization analysis of models of the type in (1.1) we refer to [1, 2, 9–11, 17]. We note that the 'universal object' found in this work coincides with our w^* in (2.15). In particular, the function g^* in (3.9) occurred first in [17].

---

[5] More generally, one can think of "p-type q-tuple models", with p populations and resampling of q-tuples, $2 \le q \le p$. For $q = 2$ this yields the p-type Wright-Fisher model. For $p = q$ the model can be transformed into an isotropic model; here for $p = q = 2, 3$ (but not for $p = q \ge 4$) the diffusion matrix is a scalar multiple of w^*, and Proposition 2.5 gives an explicit expression for second order moments. For $2 < q < p$ the p-type q-tuple model does not have invariant harmonics, in the sense of Definition 2.3. For details we refer to [29].

[6] Here we mean clustering in the sense of our Theorem 1, i.e. concentration on the effective boundary and the coordinates agree. Note that for the model in (3.12), $\partial_w K = \{x : x_1 = 0,\ x_2 \in [0,1]\} \cup \{x : x_1 = 1,\ x_2 \in \{0,1\}\}$.

4. Proofs

Since we are not assuming uniqueness for solutions of equation (1.1), we have to get our information about these solutions from the fact that they solve a martingale problem. We formulate this martingale problem and construct solutions to it in section 4.1. In section 4.2 we prove Theorem 1. Our main tool is an equation (formula (4.11)) for the time evolution of covariances between the components of solutions of (1.1). In section 4.3 we prove Lemma 2.4, which describes the time evolution of the expectation of w-harmonic functions of the components. For models with the diffusion matrix w^* in (2.15), the formulas for the covariances and the harmonic functions can be combined to give an expression for the covariances in closed form (Proposition 2.5). For more general models we use in section 4.4 a comparison argument (formula (4.32)) to derive an inequality for the covariances, which yields Theorem 2. In section 4.5 we prove the assertions in Examples 3.1 and 3.2.

4.1. Proof of Proposition 2.1

Just as in the finite-dimensional case, each solution of the infinite-dimensional stochastic differential equation (1.1) solves the martingale problem for an appropriate second order differential operator A, and conversely, each solution to the martingale problem for A can be represented, on an appropriate space equipped with a set of Brownian motions, as a weak solution of the stochastic differential equation (1.1) (see Shiga & Shimizu, [28]). In order to prove Proposition 2.1, it therefore suffices to construct solutions to this martingale problem with the mentioned properties.

Here, we first formulate the appropriate martingale problem, where some care is needed regarding the domain of the operator A. General theory then gives existence of solutions to the martingale problem and compactness of the space of solutions. We use this compactness to construct shift-invariant solutions.

We equip the space K^Λ with the product topology and let $\mathscr{C}(K^\Lambda)$ be the Banach space of continuous real-valued functions on K^Λ, equipped with the supremum norm $\|\cdot\|_\infty$. We write $x = (x_i^\alpha)_{i\in\Lambda}^{\alpha=1,\dots,d}$ for a point in K^Λ. Solutions to (1.1), whenever they exist, are continuous K^Λ-valued processes that solve the martingale problem (see Ethier & Kurtz, [14] for the relevant definitions) for the linear operator A on $\mathscr{C}(K^\Lambda)$ given by

$$(Af)(x) := \Big(\sum_{i,j}\sum_{\alpha} a(j-i)(x_j^\alpha - x_i^\alpha)\tfrac{\partial}{\partial x_i^\alpha} + \sum_i \sum_{\alpha,\beta} w_{\alpha\beta}(x_i)\tfrac{\partial^2}{\partial x_i^\alpha \partial x_i^\beta}\Big) f(x), \tag{4.1}$$

with domain $\mathscr{D}(A) = \mathscr{C}_{\mathrm{fin}}(K^\Lambda)$, the space of $\mathscr{C}^2$-functions depending on finitely many coordinates only.

We will occasionaly (notably in the proof of Lemma 2.4) need an extension of A to a larger domain. To this aim, we introduce the space $\mathscr{C}^2_{\mathrm{sum}}(K^\Lambda)$ of functions with continuous summable second derivatives. By definition, $f \in \mathscr{C}^2_{\mathrm{sum}}(K^\Lambda)$ if there is an open set $U \subset \mathbb{R}^d$, $U \supset K$, and an extension of f to a function with continuous first and second order partial derivatives on U^Λ, such that

$$\begin{aligned} x &\mapsto \Big(\frac{\partial}{\partial x_i^\alpha} f(x)\Big)_{i\in\Lambda}^{\alpha=1,\ldots,d} \\ x &\mapsto \Big(\frac{\partial^2}{\partial x_i^\alpha \partial x_j^\beta} f(x)\Big)_{i,j\in\Lambda}^{\alpha,\beta=1,\ldots,d} \end{aligned} \tag{4.2}$$

are continuous functions from K^Λ into the spaces $l^1(\{1,\ldots,d\} \times \Lambda)$ and $l^1(\{1,\ldots,d\}^2 \times \Lambda^2)$ of absolutely summable sequences, equipped with the l^1-norm.

One can check that for functions $f \in \mathscr{C}^2_{\rm sum}(K^\Lambda)$, the infinite sums in (4.1) converge in $\mathscr{C}(K^\Lambda)$ and the result does not depend on the summation order or on the choice of the extension of f to U^Λ. Writing A' for the extension of the operator in (4.1) to the larger domain $\mathscr{C}^2_{\rm sum}(K^\Lambda)$, one can check that A is a core for A', and hence any solution to the martingale problem for A also solves the martingale problem for A'.

We now collect the necessary facts about non-emptyness and compactness of the space of solutions to the martingale problem for A. Note that K^Λ is a compact, separable metrizable space. We write $\mathscr{D}_{K^\Lambda}[0,\infty)$ for the cadlag functions from $[0,\infty)$ to K^Λ, and $\mathscr{C}_{K^\Lambda}[0,\infty)$ for the continuous functions from $[0,\infty)$ to K^Λ. We equip $\mathscr{D}_{K^\Lambda}[0,\infty)$ with the Skorohod topology (see Ethier & Kurtz, [14], chapter 3), under wich it is a separable metrizable space. Measurable processes with sample paths in $\mathscr{D}_{K^\Lambda}[0,\infty)$ are $\mathscr{D}_{K^\Lambda}[0,\infty)$-valued random variables (measurable with respect to the Borel σ-field on $\mathscr{D}_{K^\Lambda}[0,\infty)$) and vice versa. We use the symbol $\Rightarrow$ for weak convergence of probability measures on $\mathscr{D}_{K^\Lambda}[0,\infty)$, as well as for convergence in distribution of $\mathscr{D}_{K^\Lambda}[0,\infty)$-valued random variables. By a solution to the martingale problem for an operator on $\mathscr{C}(K^\Lambda)$ we always mean a solution with sample paths in $\mathscr{D}_{K^\Lambda}[0,\infty)$.

Lemma 4.1. *For each probability measure μ on K^Λ there exists a solution to the martingale problem for the operator A in* (4.1) *with initial condition μ. Each solution to the martingale problem for A has sample paths in $\mathscr{C}_{K^\Lambda}[0,\infty)$. The space of solutions to the martingale problem for A is compact in the topology of weak convergence. If X_n, X solve the martingale problem for A, then $X_n \Rightarrow X$ implies $X_n(t) \Rightarrow X(t)$ for all $t \geq 0$.*

Proof of Lemma 4.1. By the Stone–Weierstrass theorem, $\mathscr{C}^2_{\rm fin}(K^\Lambda)$ is dense in $\mathscr{C}(K^\Lambda)$. Condition (1.3) guarantees that A satisfies the positive maximum principle. Hence the existence of solutions to the martingale problem for A follows immediately from Theorem 5.4 and Remark 5.5 in chapter 4 of Ethier & Kurtz [14]. The other assertions also follow from statements in that book. Continuity of sample paths folllows from Problem 19 in chapter 4, where one needs to use the functions

$$f_x(y) := \sum_i \gamma_i |x_i - y_i|^3 \qquad (x, y \in K^\Lambda), \tag{4.3}$$

with the γ_i chosen in such a way that $\sum_i a(j-i)\gamma_i \leq M\gamma_j$ for all $j \in \Lambda$. Note that $f_x \in \mathscr{C}^2_{\rm sum}(K^\Lambda)$ for all $x \in K^\Lambda$. Compactness of the space of solutions follows from Lemma 5.1 and Remark 5.2 from chapter 4 of [14]. Finally, weak convergence in path space of solutions X_n to the martingale problem for A implies convergence of finite-dimensional distributions by Theorem 7.8 from chapter 3 and the continuity of sample paths. □

Proof of Proposition 2.1. Existence of solutions to the martingale problem for A is guaranteed by Lemma 4.1, and hence all we need to do is to show that for shift-invariant initial conditions we can find a shift-invariant solution to the martingale problem.

We use a Cesàro-type argument. Let X be a solution to the martingale problem for A with initial condition $\mathscr{L}(X(0)) = \mu$. It is not hard to check that there exist finite sets $\Lambda_n \uparrow \Lambda$, such that

$$\lim_{n\to\infty} \frac{|\{j \in \Lambda_n : j + i \in \Lambda_n\}|}{|\Lambda_n|} = 1 \qquad \forall i \in \Lambda. \tag{4.4}$$

Define a shift-operator on $\mathscr{D}_{K^\Lambda}[0,\infty)$ in the obvious way, by putting $(\mathscr{T}_j x)_i(t) := x_{i-j}(t)$ $(i, j \in \Lambda,\ t \geq 0)$. Let (X_n) be a sequence of processes with sample paths in $\mathscr{D}_{K^\Lambda}[0,\infty)$ with law $\mathscr{L}(X_n) = |\Lambda_n|^{-1} \sum_{j\in\Lambda_n} \mathscr{L}(\mathscr{T}_j X)$. The X_n solve the martingale problem for A and by Lemma 4.1, the sequence (X_n) has a cluster point. It is easy to check that each cluster point is a shift-invariant solution to the martingale problem for A with initial condition μ. □

4.2. *Proof of Theorem 1*

We start by collecting some necessary facts about spatially ergodic measures on arbitrary Abelian groups. For $j \in \Lambda$, let the shift operator $T_j : K^\Lambda \to K^\Lambda$ be defined as $(T_j x)_i := x_{i-j}$. The σ-field of shift-invariant events is

$$\mathscr{S} := \{A \in \mathscr{B}(K^\Lambda) : T_i^{-1}(A) = A \ \ \forall i \in \Lambda\}. \tag{4.5}$$

A probability measure μ on K^Λ is spatially ergodic if for every $A \in \mathscr{S}$ either $\mu(A) = 1$ or $\mu(A) = 0$. The following is a variant on von Neumann's mean ergodic theorem (see Krengel, [19], chapter 1).

Lemma 4.2. *For $n = 1, 2, \ldots,$ let $p_n : \Lambda \to [0,\infty)$ be functions satisfying $\sum_i p_n(i) = 1$ and*

$$\lim_{n\to\infty} \sum_k |p_n(k-i) - p_n(k-j)| = 0 \qquad \forall i, j \in \Lambda. \tag{4.6}$$

Let $X = (X_i)_{i\in\Lambda}$ be a family of K-valued random variables with shift-invariant spatially ergodic law $\mathscr{L}(X)$. If $E[X_0] = x$, then

$$\lim_{n\to\infty} E\Big[\Big|x - \sum_i p_n(i) X_i\Big|^2\Big] = 0. \tag{4.7}$$

In our case, probability distributions p_n satisfying (4.6) will arise in the following way.

Lemma 4.3. *Let a be the kernel of an irreducible continuous-time random walk on Λ and let $P_t(j-i)$ be the probability that the walk is at time t in j when it starts in i. Then*

$$\lim_{t\to\infty}\sum_k |P_t(k-i)-P_t(k-j)| = 0 \qquad \forall i,j\in\Lambda. \tag{4.8}$$

This lemma can be proved by a standard coupling argument, which we leave to the reader.

Proof of Theorem 1. For any two K-valued random variables X and Y the covariance of X and Y is the quantity $\mathrm{Cov}(X,Y)=E[X\cdot Y]-E[X]\cdot E[Y]$, where $\cdot$ denotes the inner product $x\cdot y=\sum_\alpha x_\alpha y_\alpha$. By $\mathrm{tr}(w)$ we denote the trace $\mathrm{tr}(w)(x)=\sum_{\alpha=1}^d w_{\alpha\alpha}(x)$ $(x\in K)$ of the diffusion matrix w.

Let X be a shift-invariant solution to (1.1). A simple calculation involving Itô's formula shows that the intensity is conserved:

$$E[X_i(t)]=x \qquad \forall t\geq 0,\ i\in\Lambda, \tag{4.9}$$

for some $x\in K$. Moreover, there exists a function $C:[0,\infty)\times\Lambda\to\mathbb{R}$ such that

$$\mathrm{Cov}(X_i(t),X_j(t))=C_t(j-i) \qquad (t\geq 0,\ i,j\in\Lambda). \tag{4.10}$$

For each i, the function $t\mapsto C_t(i)$ is continuously differentiable and satisfies

$$\tfrac{\partial}{\partial t}C_t(i)=\sum_j a_S(j-i)(C_t(j)-C_t(i))+2\delta_{i0}E[\mathrm{tr}(w)(X_0(t))]. \tag{4.11}$$

Solutions to (4.11) can be represented in terms of the *symmetrized random walk* which jumps from a point i to a point j with rate $a_S(j-i)$. Let us denote by $P^S_t(j-i)$ the probability that this random walk is at time t in j, starting from i. A partial integration of the Markov semigroup of the random walk (see Liggett, [21], Theorem I.2.15) gives

$$C_t(i)=\sum_j P^S_t(j-i)C_0(j)+2\int_0^t P^S_s(i)E[\mathrm{tr}(w)(X_0(t-s))]ds. \tag{4.12}$$

We next show that assumption (2.2) implies that the law of $X(\infty)$ is an invariant law under the evolution in (1.1). More precisely, there exists a shift-invariant solution X^∞ to the martingale problem for the operator A in (4.1) such that $\mathscr{L}(X^\infty(t))=X(\infty)$ for all $t\geq 0$. To see this, define solutions to the martingale problem for A by $X_n(t):=X(t_n+t)$, where (t_n) is some sequence tending to infinity. By Lemma 4.1 and the fact that weak convergence of probability measures on $\mathscr{D}_{K^\Lambda}[0,\infty)$ is metrizable, we can find a subsequence $(X_{n(k)})$ that converges in distribution to some solution X^∞ to the martingale problem for A. Now

$$\mathscr{L}(X^\infty(t))=\lim_{n\to\infty}\mathscr{L}(X(t_n+t))=\mathscr{L}(X(\infty)) \qquad \forall t\geq 0, \tag{4.13}$$

where the limit denotes weak convergence of probability measures on K^Λ. It is easy to see that X^∞ is shift-invariant.

We now prove the assertions for a_S recurrent. Applying (4.12) to the process X^∞, we get

$$C_t^\infty(i) - \sum_j P_t^S(j-i)C_0^\infty(j) = 2\int_0^t P_s^S(i)E[\mathrm{tr}(w)(X_0^\infty(t-s))]ds, \quad (4.14)$$

where $C_t^\infty(i)$ refers to covariances of the process X^∞. By the compactness of the state space K, the left-hand side of (4.14) is bounded. The right-hand side is equal to

$$2E[\mathrm{tr}(w)(X_0(\infty))]\int_0^t P_s^S(i)ds. \quad (4.15)$$

By the recurrence of the random walk with kernel a_S, the integral in (4.15) diverges as t tends to infinity, and therefore (4.14) can only hold if $E[\mathrm{tr}(w)(X_0(\infty))] = 0$. The matrix w is non-negative definite and symmetric, and hence $\mathrm{tr}(w)(x) = 0$ implies $w(x) = 0$, and therefore $P[X_0(\infty) \in \partial_w K] = 1$ and by shift-invariance we arrive at (2.3) (i).

Moreover, combining what we have just proved with (4.11) we see that

$$\tfrac{\partial}{\partial t}C_t^\infty(i) = \sum_j a_S(j-i)(C_t^\infty(j) - C_t^\infty(i)), \quad (4.16)$$

which means that C^∞ is a bounded a_S-harmonic function. By the Choquet–Deny theorem (which follows easily from Lemma 4.3 –see Liggett, [21], Theorem II.1.5) it follows that C^∞ is constant. Hence $\mathrm{Cov}(X_i(\infty), X_j(\infty)) = \mathrm{Var}(X_0(\infty))$ for all $i, j \in \Lambda$ and the Cauchy–Schwarz inequality gives $P[X_i(\infty) = X_j(\infty)] = 1$, which implies (2.3) (ii).

We now prove the assertions for a_S transient. We start by noting that the spatial ergodicity of $\mathscr{L}(X(0))$ together with Lemma 4.2 and 4.3 imply that for each $i \in \Lambda$

$$\lim_{t\to\infty} \sum_j P_t^S(j-i)C_0(j) = 0. \quad (4.17)$$

Applying (4.12) to the process X, taking the limit $t \to \infty$, using the fact that a_S is transient and inserting (4.17), we get

$$\lim_{t\to\infty} C_t(i) = 2E[\mathrm{tr}(w)(X_0(\infty))]\int_0^\infty P_t^S(i)dt. \quad (4.18)$$

Now assume that $P[X_0(\infty) \in \partial_w K] = 1$. Then $E[\mathrm{tr}(w)(X_0(\infty))] = 0$ and (4.18) gives $\mathrm{Var}(X_0(\infty)) = 0$ and therefore, by (4.9), $X_0(\infty) = E[X_0(0)]$ a.s. This contradicts our assumption that $E[X_0(0)] \notin \partial_w K$ and we conclude that $P[X_0(\infty) \in \partial_w K] < 1$, and by shift-invariance (2.4) (i) holds.

To get (2.4) (ii), we note that by the fact that the random walk with kernel a_S is symmetric, $\int_0^\infty P_t^S(j-i)dt < \int_0^\infty P_t^S(0)dt$ for all $i \neq j$, and hence by (4.18), $\mathrm{Cov}(X_i(\infty), X_j(\infty)) < \mathrm{Var}(X_0(\infty))$, which implies (2.4) (ii). □

4.3. *Proof of Lemma 2.4*

Solutions to (2.12) can be expressed in terms of the random walk which jumps from a point i to a point j with rate $a(j-i)$. Let us denote by $P_t(j-i)$ the probability that this random walk is at time t in j, starting from i. Then

$$Y_i(t) = \sum_j P_t(j-i)Y_j(0) \qquad (i \in \Lambda,\ t \geq 0). \tag{4.19}$$

Let us write $(R_t)_{t\geq 0}$ for the Feller semigroup on $\mathscr{C}(K^\Lambda)$ associated with the process Y, i.e., $(R_t f)(y) := E[f(Y^y(t))]$ where Y^y is the solution of (2.12) with initial condition $Y(0) = y$.

Fix $h \in H(w) \cap \mathscr{C}^2(K)$ and $i \in \Lambda$ and define $f \in \mathscr{C}^2_{\mathrm{fin}}(K^\Lambda)$ by $f(x) := h(x_i)$. It is not hard to see that $R_t f \in \mathscr{C}^2_{\mathrm{sum}}(K^\Lambda)$ for all $t \geq 0$ (however, $R_t f \notin \mathscr{C}^2_{\mathrm{fin}}(K^\Lambda)$). Let A' be the extension of the operator A in (4.1) to the domain $\mathscr{C}^2_{\mathrm{sum}}(K^\Lambda)$. Write $A' = B + C$, where B contains the first and C the second order derivatives. General theory (see Ethier & Kurtz, [14], chapter 1) tells us that $t \mapsto R_t h$ is continuously differentiable in $\mathscr{C}(K)$ and

$$\tfrac{\partial}{\partial t} R_t h = B R_t h. \tag{4.20}$$

Using the fact that X solves the martingale problem for A' we get the following integration by parts

$$\begin{aligned} E[h(X_i(T))] - E[h(Y_i(T))] &= \\ E[(R_0 f)(X(T))] - E[(R_T f)(X(0))] &= E\int_0^T (B + C + \tfrac{\partial}{\partial t})(R_{T-t} f)(X(t))dt \\ &= E\int_0^T (C R_{T-t} f)(X(t))dt \qquad (T \geq 0). \end{aligned} \tag{4.21}$$

Here

$$(C R_{T-t} f)(x) = \sum_k P_{T-t}(k-i)^2 \sum_{\alpha,\beta} w_{\alpha\beta}(x_k)\Big(\tfrac{\partial^2}{\partial x_\alpha \partial x_\beta} h\Big) \ \Big(\sum_j P_{T-t}(j-i)x_j\Big) \qquad (x \in K^\Lambda). \tag{4.22}$$

For each k in this summation, fix $(x_j)_{j\neq k}$ and define $s \geq 0$ and $\theta \in K$ by $s = -\log P_{T-t}(k-i)$ and $\theta = (1-e^{-s})^{-1}\sum_{j\neq k} P_{T-t}(j-i)x_j$, i.e., define s and θ in such a way that $(\theta + (x_k - \theta)e^{-s}) = \sum_j P_{T-t}(j-i)x_j$. Then

$$\begin{aligned} &P_{T-t}(k-i)^2 \sum_{\alpha,\beta} w_{\alpha\beta}(x_k)\Big(\tfrac{\partial^2}{\partial x_\alpha \partial x_\beta} h\Big)\Big(\sum_j P_{T-t}(j-i)x_j\Big) \\ &\quad = \sum_{\alpha,\beta} w_{\alpha\beta}(x_k) \frac{\partial^2}{\partial x_k^\alpha \partial x_k^\beta}(T_{\theta,s} h)(x_k). \end{aligned} \tag{4.23}$$

Now use that w has invariant harmonics and that $T_{\theta,s}$ maps differentiable functions into differentiable functions, so that

$$T_{\theta,s}(H(w) \cap \mathscr{C}^2(K)) \subset H(w) \cap \mathscr{C}^2(K) \qquad \forall \theta \in K,\ s \geq 0, \tag{4.24}$$

to see that the right hand sides in (4.23), (4.22) and (4.21) are zero. This proves the lemma for $h \in H(w) \cap \mathscr{C}^2(K)$. To generalize this to arbitrary $h \in H(w)$ it suffices to note that the set of $h \in \mathscr{C}(K)$ for which (2.13) holds is bp-closed. □

4.4. Proof of Theorem 2 and Proposition 2.5

We start by collecting some elementary facts about w-harmonic functions from potential theory. If w satisfies (A.1), then the formula $(S_t f)(x) := E[f(Z^x(t))]$, with Z^x as in (2.7), defines a Feller semigroup $(S_t)_{t\geq 0}$ on $\mathscr{C}(K)$. Let G be its full generator, i.e. $Gf := \lim_{t\to 0} t^{-1}(S_t f - f)$ with domain $\mathscr{D}(G)$ the space of all $f \in \mathscr{C}(K)$ for which the limit exists in $\mathscr{C}(K)$. We add a final element to the semigroup, by $(S_\infty f)(x) := E[f(Z^x(\infty))] = \int_K \Gamma_x(dy) f(y)$.

Lemma 4.4. *Assume that w satisfies* (A.1) *and* (A.2). *Consider sets H, H', H'', H''' defined as*

$$
\begin{aligned}
H &:= \{h \in \mathscr{D}(G) : Gh = 0\} \\
H' &:= \{h \in \mathscr{C}(K) : S_t h = h \ \ \forall t \in [0,\infty]\} \\
H'' &:= \{h \in \mathscr{C}(K) : S_\infty h = h\} \\
H''' &:= \{S_\infty \phi : \phi \in \mathscr{C}(K)\}
\end{aligned}
\tag{4.25}
$$

Then $H = H' = H'' = H''' = H(w)$. For each $\phi \in \mathscr{C}(K)$ there exists a unique $h \in H(w)$ such that

$$
h(x) = \phi(x) \qquad (x \in \partial_w K) \tag{4.26}
$$

and this h is given by $h = S_\infty \phi$.

Proof of Lemma 4.4. It is easy to see that $H \subset H' \subset H'' \subset H''' = H(w)$. To see that $H''' \subset H$, note that by the martingale convergence theorem $\phi(Z^x(t)) \to \phi(Z^x(\infty))$ almost surely, so bounded convergence implies that $E[\phi(Z^x(t))] \to E[\phi(Z^x(\infty))]$ for each $x \in K$. Since $|E[\phi(Z^x(t))]| \leq \|\phi\|_\infty < \infty$, it follows that $S_t\phi \to S_\infty\phi$ as $t \to \infty$ in the sense of bounded pointwise convergence. Therefore

$$
(S_t S_\infty \phi)(x) = \lim_{s\to\infty} (S_t S_s \phi)(x) = (S_\infty \phi)(x) \qquad (x \in K). \tag{4.27}
$$

It follows that $t^{-1}(S_t - 1)S_\infty\phi = 0$ for all t, so taking the limit $t \to 0$ we see that $H''' \subset H$. Now we note that

$$
P[Z^x(\infty) \in \partial_w K] = 1 \qquad (x \in K). \tag{4.28}
$$

(Since Z^x converges, this is in fact a special case of Theorem 1.) By (4.25), $S_\infty\phi \in H$ for each $\phi \in \mathscr{C}(K)$. To see that $h := S_\infty\phi$ solves (4.26) it suffices to note that for each $x \in \partial_w K$ the process $Z^x(t) := x$ solves (2.7). To see that h is the unique w-harmonic function satisfying (4.26), suppose that $\tilde{h} \in H$ is another one. Then by (4.25) and by (4.28)

$$
\tilde{h} = S_\infty \tilde{h} = S_\infty \phi = h. \tag{4.29}
$$

□

Proof of Theorem 2. Consider the trace $\mathrm{tr}(w^*)$ of the diffusion matrix w^* in (2.15). Since $w^*(x)$ is a non-negative definite symmetric matrix, $\mathrm{tr}(w^*)$ is non-negative. Moreover, it is easy to see from (2.15) and the fact that Γ_x is concentrated on $\partial_w K$ (see (4.28)) that $\mathrm{tr}(w^*)(x) = 0$ iff $x \in \partial_w K$. Since we are assuming (A.2), $\mathrm{tr}(w^*) \in \mathscr{C}(K)$. In fact, $\mathrm{tr}(w^*) \in \mathscr{D}(G)$ (see Lemma 4.4) and

$$G\mathrm{tr}(w^*) = -2\mathrm{tr}(w). \tag{4.30}$$

To see why this is so, note that by Lemma 4.4 there exists a unique $h \in H(w)$ such that $h(x) = |x|^2$ on ∂K, and therefore $v^*(x) := h(x) - |x|^2$ is the unique $v^* \in \mathscr{D}(G)$ such that $v^* = 0$ on $\partial_w K$ and $Gv^* = -2\mathrm{tr}(w)$. Now use that a solution Z^x of (2.7) solves the martingale problem for G to write for $x \in K$

$$\begin{aligned}\mathrm{tr}(w^*)(x) = E[|Z^x(\infty) - x|^2] = 2\int_0^\infty E[\mathrm{tr}(w)(Z^x(t))]dt \\ = v^*(x) - E[v^*(Z^x(\infty))] = v^*(x).\end{aligned} \tag{4.31}$$

Like $\mathrm{tr}(w^*)$, the function $\mathrm{tr}(w)$ is continuous, non-negative and satisfies $\mathrm{tr}(w)(x) = 0 \Leftrightarrow x \in \partial_w K$. Therefore for each $\varepsilon > 0$ we can find a $\lambda > 0$ such that

$$\mathrm{tr}(w) \geq \lambda(\mathrm{tr}(w^*) - \varepsilon). \tag{4.32}$$

Inserting this into (4.11) we find that

$$\tfrac{\partial}{\partial t}C_t(i) \geq \sum_j a_S(j-i)(C_t(j) - C_t(i)) + 2\lambda\delta_{i0}\big(E[\mathrm{tr}(w^*)(X_0(t))] - \varepsilon\big) \quad (i \in \Lambda,\ t \geq 0). \tag{4.33}$$

For any function $h \in H(w)$, Lemma 2.4 and (4.19) imply that

$$E[h(X_0(t))] = E\Big[h\Big(\sum_j P_t(j)X_j(0)\Big)\Big]. \tag{4.34}$$

By Lemma 4.2, Lemma 4.3, the spatial ergodicity of $\mathscr{L}(X(0))$ and the continuity of h this implies that

$$\lim_{t\to\infty} E[h(X_0(t))] = h(\theta). \tag{4.35}$$

In particular, we may apply this to the function $x \mapsto \mathrm{tr}(w^*)(x) + |x - \theta|^2$ which is a continuous w-harmonic function by (4.30) and Lemma 4.4, to get

$$\lim_{t\to\infty} E[\mathrm{tr}(w^*)(X_0(t))] + C_t(0) = \mathrm{tr}(w^*)(\theta). \tag{4.36}$$

Combining this with (4.33) we see there exists a T such that for all $t \geq T$

$$\tfrac{\partial}{\partial t}C_t(i) \geq \sum_j a_S(j-i)(C_t(j) - C_t(i)) + 2\lambda\delta_{i0}(\mathrm{tr}(w^*)(\theta) - C_t(0) - 2\varepsilon) \qquad (i \in \Lambda,\ t \geq 0).. \tag{4.37}$$

Write $D_t(i) := \mathrm{tr}(w^*)(\theta) - C_t(i) - 2\varepsilon$ ($i \in \Lambda,\ t \geq 0$), so that

$$\frac{\partial}{\partial t} D_t(i) = \sum_j a_S(j-i)(D_t(j)-D_t(i)) - 2\lambda\delta_{i0} D_t(0) + R_t(i) \qquad (i \in \Lambda,\ t \geq T), \tag{4.38}$$

where the remainder $R_t(i)$ is non-positive and, since $C_t(i)$ is continuously differentiable, continuous in t. Consider a continuous-time random walk on Λ that jumps from a point i to a point j with rate $a_S(j-i)$ and that is killed in the origin with rate 2λ. Denote by $P_t^\lambda(j,i)$ the probability that this random walk is at time t in j, starting from i. A partial integration of (4.38) gives (see Liggett, [21], Theorem I.2.15)

$$D_{T+t}(i) \leq \sum_j P_t^\lambda(j,i) D_T(j) \qquad (t \geq 0). \tag{4.39}$$

Since we are assuming that a_S is recurrent, $\sum_j P_t^\lambda(j,i) \to 0$ as $t \to \infty$ for each $i \in \Lambda$, and using the compactness of K we see that for each $i \in \Lambda$ there exists a T' such that for all $t \geq T'$

$$C_t(i) \geq \mathrm{tr}(w^*)(\theta) - 3\varepsilon. \tag{4.40}$$

We have thus shown that $\liminf_{t\to\infty} C_t(i) \geq \mathrm{tr}(w^*)(\theta)$ for every $i \in \Lambda$. On the other hand, with the help of formula (4.36) it is easy to see that $\limsup_{t\to\infty} C_t(0) \leq \mathrm{tr}(w^*)(\theta)$. By the Cauchy–Schwarz inequality, $C_t(i) \leq C_t(0)$ for all $i \in \Lambda$ and hence

$$\lim_{t\to\infty} C_t(i) = \mathrm{tr}(w^*)(\theta) \qquad \forall i \in \Lambda. \tag{4.41}$$

We now show the convergence in distribution of $X(t)$. We start with X_0. Combining (4.41) with (4.36) one gets

$$\lim_{t\to\infty} E[\mathrm{tr}(w^*)(X_0(t))] = 0. \tag{4.42}$$

Now fix any $\phi \in \mathscr{C}(K)$. The function $(\phi - S_\infty\phi)$ is continuous on K and zero on $\partial_w K$ (see Lemma 4.4)), and from (4.42) and the fact that $\mathrm{tr}(w^*)$ is continuous, non-negative, and zero only on $\partial_w K$ it therefore follows that

$$\lim_{t\to\infty} E[\phi(X_0(t)) - (S_\infty\phi)(X_0(t))] = 0. \tag{4.43}$$

Here $S_\infty\phi \in H(w)$ (see Lemma 4.4) and hence by (4.35)

$$\lim_{t\to\infty} E[(S_\infty\phi)(X_0(t))] = (S_\infty\phi)(\theta) = \int_K \Gamma_\theta(dx)\phi(x). \tag{4.44}$$

Formulas (4.43) and (4.44) imply that $\mathscr{L}(X_0(t)) \Rightarrow \Gamma_\theta$ as $t \to \infty$. To get from this to the convergence of $X(t)$ it suffices to note that (4.41) by the Cauchy–Schwarz inequality implies that for all $i, j \in \Lambda$

$$\lim_{t\to\infty} E[|X_i(t) - X_j(t)|^2] = 0. \tag{4.45}$$

This gives weak convergence for $\mathscr{L}((X_i(t))_{i\in\Delta})$ for every finite $\Delta \subset \Lambda$, which implies (2.14). $\square$

Proof of Proposition 2.5. This proof copies the proof of Theorem 2 up to formula (4.39), with a few changes. First, replace the covariance function $C_t(i)$ by a covariance matrix function

$$C_t(j-i)_{\alpha\beta} := E[(X_i^\alpha(t)-\theta_\alpha)(X_j^\beta(t)-\theta_\beta)]. \tag{4.46}$$

Now change (4.30) into $Gw^*_{\alpha\beta} = -2w_{\alpha\beta}$ and note that (4.32) and (4.34) change and simplify to

$$\begin{aligned} w &= \lambda w^* \\ E[h(X_0(t))] &= h(\theta), \end{aligned} \tag{4.47}$$

respectively, and proceed the argument with equalities replacing inequalities, $\varepsilon = 0$ and $T = 0$ to get, instead of (4.39)

$$E[(X_i^\alpha(t)-\theta_\alpha)(X_j^\beta(t)-\theta_\beta)] = w^*_{\alpha\beta}(\theta)\Big(1-\sum_k P_t^\lambda(k, j-i)\Big), \tag{4.48}$$

which is just (2.16). □

4.5. Proof of the examples

Proof of Example 3.1. We start with weak uniqueness for (2.7). The uniqueness proof of Sato, in section 4 of [23], although stated there only for diffusion matrices of a special form, carries over to our situation. For this, the main fact one has to check is the following. For $\alpha = 1, \ldots, d+1$, let

$$\begin{aligned} F_\alpha &:= \{(x_1,\ldots,x_d) \in K_d : x_\alpha = 0\} \qquad &(\alpha = 1,\ldots,d) \\ F_\alpha &:= \{(x_1,\ldots,x_d) \in K_d : \textstyle\sum_\beta x_\beta = 1\} \qquad &(\alpha = d+1) \end{aligned} \tag{4.49}$$

be the α-face of the d-dimensional simplex K_d. Then one needs that for any solution Z^x to (2.7) with $x \in F_\alpha$

$$P[Z^x(t) \in F_\alpha \quad \forall t \geq 0] = 1, \tag{4.50}$$

but this is immediate by the martingale property of solutions to (2.7). Now one can prove strong uniqueness for the case where σ is the unique non-negative definite symmetric root of w. By (3.2), this σ is Lipschitz continuous on the interior of K_d and therefore a standard argument gives uniqueness of solutions to (2.7) up to the first hitting of a face F_α. By (4.50), the process stays in this face after hitting it. Each face is isomorphic to K_{d-1} and therefore strong uniqueness can be proved by induction. For details we refer to [23].

To see that $H(w)$ is given by (3.3), note that by (3.2), the effective boundary of K consists of the extremal points of K_d:

$$\partial_w K = \{e_1, \ldots, e_{d+1}\}, \tag{4.51}$$

where $e_1 = (1,0,\ldots,0), \ldots, e_d = (0,\ldots,0,1)$ and $e_{d+1} = (0,\ldots,0)$. By the martingale property of solutions to (2.7), the harmonic measure associated with w is given by

$$\Gamma_x = \sum_{\alpha=1}^d x_\alpha \delta_{e_\alpha} + \Big(1-\sum_{\alpha=1}^d x_\alpha\Big)\delta_{e_{d+1}}. \tag{4.52}$$

Since Γ_x is an affine function of x, every harmonic function is affine. The other assertions are now trivial. □

Proof of Example 3.2. Weak uniqueness of solutions to (2.7) is proved in the same way as in Example 3.1, where this time one needs to check that any solution Z^x starting in $x \in \partial K$ is constant with probability one.

We note that solutions to (2.7) are now time-transformed Brownian motions, and hence Γ_x is the first hitting distribution of Brownian motion started in x and stopped at ∂K. Hence the w-harmonic functions are the same as the harmonic functions for Brownian motion and we can use Proposition 4.2.7 and Theorems 4.2.12 and 4.2.19 in Karatzas & Shreve, [20], to see that (A.2) holds and that $H(w)$ is given by formula (3.7). The same references show that (3.9) has a unique solution. It follows from (3.7) that w has invariant harmonics. The other assertions in Example 3.2 are now readily checked. □

Acknowledgements. Research supported by the Dutch Association for Mathematical Physics. We thank Prof. Dr. F. den Hollander, Prof. Dr. A. van Rooij, Prof. Dr. E. Thomas and an unknown referee for valuable comments during the preparation of the manuscript.

References

[1] Baillon, J.-B., Clément, Ph., Greven, A., den, F.: Hollander, On the attracting orbit of a non-linear transformation arising from renormalization of hierarchically interacting diffusions. Part I: The compact case. Can. J. Math., **47**, 3–27 (1995)
[2] Baillon, J.-B., Clément, Ph., Greven, A., den, F.: Hollander, On the attracting orbit of a non-linear transformation arising from renormalization of hierarchically interacting diffusions. Part II: The non-compact case. J. Funct. Anal., **146**, 236–298 (1997)
[3] Cox, J.T., Fleischmann, K., Greven, A.: Comparison of interacting diffusions and applications to their ergodic theory. Probab. Theory Relat. Fields, **105**, 513–528 (1996)
[4] Cox, J.T., Greven, A.: Ergodic theorems for infinite systems of locally interacting diffusions. Ann. Probab., **22**, 833–853 (1994)
[5] Cox, J.T., Greven, A., Shiga, T.: Finite and infinite systems of interacting diffusions. Probab. Theory Relat. Fields, **103**, 165–197 (1995)
[6] Cox, J.T., Greven, A., Shiga, T.: Finite and infinite systems of interacting diffusions: cluster formation and universality properties. Math. Nachr., **192**, 105–124 (1998)
[7] Cox, J.T., Griffeath, D.: Diffusive clustering in the two dimensional voter model. Ann. Probab., **14**, 347–370 (1986)
[8] Dawson, D.A., Fleischmann, K.: Longtime behavior of a branching process controlled by branching catalysts. Stoch. Process. Appl., **71**, 241–257 (1997)
[9] Dawson, D.A., Greven, A.: Multiple time scale analysis of interacting diffusions. Probab. Theory Relat. Fields, **95**, 467–508 (1993)
[10] Dawson, D.A., Greven, A.: Hierarchical models of interacting diffusions: multiple time scale phenomena, phase transition and pattern of cluster-formation. Probab. Theory Relat. Fields, **96**, 435–473 (1993)
[11] Dawson, D.A., Greven, A.: Multiple space-time scale analysis for interacting branching models. Electronic J. of Prob., **1**, (1996), paper no. 14, pp. 1–84 (http://math.-washington.-edu/∼ejpecp/)
[12] Dawson, D.A., Greven, A., Vaillancourt, J.: Equilibria and quasi-equilibria for infinite collections of interacting Fleming–Viot processes. Trans. Am. Math. Soc., **347**, 2277–2360 (1995)

[13] Dawson, D.A., Perkins, E.A.: Long-time behavior and coexistence in a mutually catalytic branching model. Ann. Probab., **26**, 1088–1138 (1998)
[14] Ethier, S.N., Kurtz, T.G.: Markov Processes; characterization and convergence, John Wiley & Sons, New York, (1986)
[15] Fleischmann, K., Klenke, A.: Smooth density field of catalytic super–Brownian motion. Ann. Appl. Probab., **9**, 298–318 (1999)
[16] Greven, A., Klenke, A., Wakolbinger, A.: Interacting Fisher–Wright diffusions in a catalytic medium. Preprint university of Erlangen (www.mi.uni-erlangen.de/∼klenke)
[17] den Hollander, F., Swart, J.M.: Renormalization of hierarchically interacting isotropic diffusions. J. Stat. Phys., **93**, 243–291 (1998)
[18] Klenke, A.: Different clustering regimes in systems of hierarchically interacting diffusions. Ann. Probab., **24**, 660–697 (1996)
[19] Krengel, U.: Ergodic Theorems, Walter de Gruyter, Berlin, (1985)
[20] Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus, second edition, Springer, New York, (1991)
[21] Liggett, T.M.: Interacting Particle Systems, Springer, New York, (1985)
[22] Notohara, M., Shiga, T.: Convergence to genetically uniform state in stepping stone models of population genetics. J. Math. Biol., **10**, 281–294 (1980)
[23] Sato, K.-I.: Convergence to a diffusion of a multi-allelic model in population genetics. Adv. Appl. Prob., **10**, 538–562 (1978)
[24] Shiga, T.: An interacting system in population genetics. J. Math. Kyoto Univ., **20**, 213–243 (1980)
[25] Shiga, T.: Diffusion processes in population genetics. J. Math. Kyoto Univ., **21**, 133–151 (1981)
[26] Shiga, T.: A certain class of infinite-dimensional diffusion processes arising in population genetics. J. Math. Soc., Japan **39**, 17–25 (1987)
[27] Shiga, T.: Ergodic theorems and exponential decay of sample paths for certain interacting diffusion systems. Osaka J. Math., **29**, 789–807 (1992)
[28] Shiga, T., Shimizu, A.: Infinite dimensional stochastic differential equations and their applications. J. Math. Kyoto Univ., **20**, 395–416 (1980)
[29] Swart, J.M.: Large Space-Time Scale Behavior of Linearly Interacting Diffusions, doctoral thesis, University of Nijmegen, 1999 (http://helikon.ubn.kun.nl/mono/s/swart_j/largspscb.pdf)