Stationary distributions and convergence for Walsh diffusions

TOMOYUKI ICHIBA[1] and ANDREY SARANTSEV[2]

[1]*Department of Statistics and Applied Probability, University of California, Santa Barbara, South Hall 5607A, Santa Barbara, CA, 93106, USA. E-mail: ichiba@pstat.ucsb.edu*
[2]*Department of Mathematics and Statistics, University of Nevada, Reno, 1664 N Virginia St, Reno, NV, 89557, USA. E-mail: asarantsev@unr.edu*

A Walsh diffusion on Euclidean space moves along each ray from the origin, as a solution to a stochastic differential equation with certain drift and diffusion coefficients, as long as it stays away from the origin. As it hits the origin, it instantaneously chooses a new direction according to a given probability law, called the spinning measure. A special example is a real-valued diffusion with skew reflections at the origin. This process continuously (in the weak sense) depends on the spinning measure. We determine a stationary measure for such process, explore long-term convergence to this distribution and establish an explicit rate of exponential convergence.

Keywords: ergodic process; invariant measure; Lyapunov function; reflected diffusion; stationary distribution; stochastic differential equation; Walsh Brownian motion; Walsh diffusion

1. Introduction

1.1. Informal description of Walsh Brownian motions and Walsh diffusions

Fix a positive integer $d \geq 1$, a dimension of $\mathbb{R}^d$ with Euclidean norm

$$\|x\| := \left(x_1^2 + \cdots + x_d^2\right)^{1/2}.$$

Take a Borel probability measure μ on the unit sphere $\mathbb{S} := \{z \in \mathbb{R}^d \mid \|z\| = 1\}$. The origin in $\mathbb{R}^d$ will be denoted by $\mathbf{0}$, to distinguish it from the zero on the real line. With rays $\mathcal{R}_\theta := \{s\theta \in \mathbb{R}^d \mid s \geq 0\}$, $\theta \in \mathbb{S}$, we see $\mathbb{R}^d = \bigcup_{\theta \in \mathbb{S}} \mathcal{R}_\theta$. Take a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbf{P})$ with the filtration satisfying the usual conditions. A *Walsh Brownian motion* in $\mathbb{R}^d$ with *spinning measure* μ is an adapted, continuous stochastic process $X = (X(t), t \geq 0)$ which is informally described as follows.

Let us take a one-dimensional reflected Brownian motion $S = (S(t), t \geq 0)$ with values in $\mathbb{R}_+ := [0, \infty)$, starting from the origin and reflected at the origin. Its sample path can be split into *excursions* in a measurable way. For *every* excursion, choose an $\mathbb{S}$-valued random variable $\boldsymbol{\theta}$ distributed in μ, independent of these variables for other excursions and of the underlying reflected Brownian motion S. Define the d-dimensional stochastic process $X(t) = \boldsymbol{\theta} S(t)$ for *each* t in the open interval which the corresponding excursion straddles, and $X(t) = \mathbf{0}$ for all other t (where $S(t) = 0$). We call $X = (X(t), t \geq 0)$ a Walsh Brownian motion in $\mathbb{R}^d$ with spinning measure μ.

A particular case is the *skew Brownian motion*, when the measure μ is supported on two opposite polar points (North Pole and South Pole) on $\mathbb{S}$. In this case, the process is essentially one-dimensional. In fact, the skew Brownian motion is usually defined as an $\mathbb{R}$-valued process, where, instead of North and South Pole, we choose the positive and negative half-line. For example, if the probabilities attached to the North and South Poles are equal to $1/2$, then we recreate a standard Brownian motion. Usually, Walsh Brownian motions in $\mathbb{R}^2$ are considered in the literature, however, it was originally pointed out by Walsh in Walsh [34] that the construction and the corresponding theory for the general d-dimensional case is basically the same as the case of $d = 2$.

We can also construct *Walsh diffusions*, given the measure μ on $\mathbb{S}$ and some coefficients $g : \mathbb{R}^d \to \mathbb{R}$ and $\sigma : \mathbb{R}^d \to \mathbb{R}_+$. Similarly to a Walsh Brownian motion, Walsh diffusions move along the rays $\mathcal{R}_\theta$, $\theta \in \mathbb{S}$, that is, as long as this process is on the ray $\mathcal{R}_\theta$ for a certain $\theta \in \mathbb{S}$, it behaves as a solution of a stochastic differential equation (SDE) on $(0, \infty)$ with drift coefficient $g(\cdot, \theta)$ and diffusion coefficient $\sigma^2(\cdot, \theta)$. If necessary, we stress their dependence on both θ and $r = \|x\|$ by writing them as $g(r, \theta)$ and $\sigma(r, \theta)$. Note that the meaning of spinning measure μ for this Walsh diffusion is slightly different from what we describe for the Walsh Brownian motion in the above. More precisely, for every subset $A \subseteq \mathbb{S}$, $\mu(A)$ is the share of local time accumulated at the origin corresponding to the excursions of the Walsh diffusion on rays $\mathcal{R}_\theta$, $\theta \in A$ (see (2.5) below).

Our primary purpose of study in this paper is the long-term convergence of the Walsh diffusion to the stationary measure under appropriate conditions. Along the course, we examine in detail the local behavior of the Walsh diffusions in the neighborhood of the origin, and examine the Feller continuity and some other properties of Walsh diffusions.

1.2. Historical review

Walsh Brownian motion was introduced in Walsh [34] and further studied in Barlow, Pitman and Yor [2], in the two-dimensional context (but the results are immediately transferred to $\mathbb{R}^d$ for $d > 2$). In much of the existing literature, the support of the spinning measure μ is finite, that is, supp $\mu = \{\theta_1, \ldots, \theta_k\} \subset \mathbb{S}$, where $\theta_1, \ldots, \theta_k$ are k distinct points in $\mathbb{S}$. In this case, Walsh Brownian motion or Walsh diffusion has effective state space $\{r\theta_i, i = 1, \ldots, k, r \geq 0\}$, and is sometimes called a *spider*. Filtrations generated by Walsh Brownian motion on a spider were studied in Tsirelson [32], Watanabe [35]. A construction of Walsh Brownian motion and, more generally, a Walsh diffusion, on a spider via pinching points together was done in Evans and Sowers [9]. Itô's formula for Walsh Brownian motion with general spinning measure was proved in Hajri and Touhami [15]. Stochastic calculus for general tree-valued diffusion processes is developed in Freidlin and Sheu [12], Freidlin and Wentzell [13], Picard [23]. A Dirichlet form approach was used in Chen and Fukushima [4] to construct Walsh Brownian motion with spinning measure μ, and in Chen and Fukushima [3], Section 7.5, Section 7.6, for the spider. Here, we apply this method of Dirichlet forms to find stationary measures of Walsh diffusions with general spinning measure. We also use Dirichlet forms to construct reflected Walsh diffusions. Stochastic flows and harmonic functions for Walsh Brownian motion were studied in Hajri [14] and Fitzsimmons and Kuter [11], respectively. Walsh semimartingales and diffusions, with arbitrary spinning measure μ, were introduced in Ichiba *et al.* [16] and further studied in Karatzas and Yan [18] with

control problems. Exponential ergodicity for a related class of Markov processes with *random switching* was recently studied in Cloez and Hairer [5].

1.3. Overview of the paper

In Section 2, we introduce notation, define Walsh Brownian motions as well as Walsh diffusions, and study some of their elementary properties. Section 2, for the most part, does not contain new results; it is a review of Karatzas and Yan [18]. In Section 3, we construct Walsh diffusions using a Dirichlet form and the method of *one-point reflection* from the paper Chen and Fukushima [4] and the book Chen and Fukushima [3], Section 7.5, Section 7.6. This alternative construction helps us to find a stationary measure for a Walsh diffusion (which is, however, not necessarily a stationary distribution). In Section 4, we discuss continuous dependence of the law of a Walsh diffusion $X = (X(t), t \geq 0)$ on the spinning measure μ and the initial condition $X(0) = x \in \mathbb{R}^d$. For the case of a Walsh Brownian motion, we quantify this continuity, effectively saying that the law of a Walsh Brownian motion is a Hölder continuous function of μ with respect to a Wasserstein distance on $\mathbb{R}^d$.

In Section 5, we study additional properties of Walsh diffusions, which were not considered in the previous paper Karatzas and Yan [18]: positivity of the transition kernel with respect to a reference measure on $\mathbb{R}^d$, and Feller continuity. In Section 6, we construct Lyapunov functions for Walsh diffusions to show *ergodicity*: existence and uniqueness of a stationary distribution π, and convergence to π in the total variation norm (or even stronger norms) as $t \to \infty$. Under some more restrictive conditions, we also prove *uniform ergodicity*: exponentially fast convergence to π as $t \to \infty$.

Our main contribution is to find explicit estimates of the rate of exponential convergence for Walsh diffusions, extending the ones in Lund, Meyn and Tweedie [19], Sarantsev [29] for reflected diffusions and jump-diffusions on a positive half-line. These results are then applied to a *non-reflected* diffusion on the whole real line. We would like to stress that often, it is relatively easy to prove that a diffusion process (or a discrete-time Markov chain) on $\mathbb{R}^d$ converges to its stationary distribution exponentially fast, but difficult to find or estimate an explicit rate of exponential convergence. Some partial results in this direction are provided in the papers Bakry, Cattiaux and Guillin [1], Davies [6], Meyn and Tweedie [22], Roberts and Rosenthal [26], Roberts and Tweedie [27,28].

2. Background and definitions

2.1. Notation

Recall that in $\mathbb{R}^d$, the Euclidean norm is defined by $\|x\| := (x_1^2 + \cdots + x_d^2)^{1/2}$ for $x = (x_1, \ldots, x_d)$. We shall denote the origin in $\mathbb{R}^d$ by $\mathbf{0} = (0, \ldots, 0)$. Let $\mathbb{S} := \{x \in \mathbb{R}^d \mid \|x\| = 1\}$ and $\mathbb{B} := \{x \in \mathbb{R}^d \mid \|x\| \leq 1\}$ be the unit sphere and the unit ball in $\mathbb{R}^d$. For every $x \in \mathbb{R}^d \setminus \{\mathbf{0}\}$, we write $x = (r, \theta)$ or simply $x = r\theta$ if $r = \|x\| > 0$ and $\theta = \arg(x) := x/r \in \mathbb{S}$ (polar coordinates). We denote by $(x)_-$ the non-positive part of a real number $x \in \mathbb{R}$. We define the *tree-metric*

as follows: for every $x_1, x_2 \in \mathbb{R}^d$ with $r_i := \|x_i\|$, $i = 1, 2$,

$$
\text{dist}(x_1, x_2) := \begin{cases} r_1, & \text{if } r_2 = 0; \\ r_2, & \text{if } r_1 = 0; \\ |r_1 - r_2|, & \text{if } r_1 r_2 \neq 0 \text{ and } \arg(x_1) = \arg(x_2); \\ r_1 + r_2, & \text{if } r_1 r_2 \neq 0 \text{ and } \arg(x_1) \neq \arg(x_2). \end{cases} \tag{2.1}
$$

This essentially means that we have the usual Euclidean distance on each ray $\mathcal{R}_\theta$ but a continuous movement cannot jump between rays, except through the origin. One can think of it as railroads converging to the central city; this is why it is sometimes also called the *railway metric*. The corresponding topology is called the *tree-topology*. This topology is stronger than the usual Euclidean topology. That is, convergence in the tree-topology means also convergence in the Euclidean sense, but the converse is not true. Being an open, closed, or Borel set in the Euclidean metric implies being, respectively, open, closed, or Borel in the tree metric, but not vice versa. When we refer to Borel subsets of $\mathbb{R}^d$ below, we mean "Borel in the Euclidean topology". The property of boundedness is equivalent in these two metrics; but the property of compactness is not, as described in the following remark.

Remark 2.1. In the Euclidean topology in $\mathbb{R}^d$, a closed bounded set is compact. In the tree-topology, this is no longer true in the general case. Here is a counterexample. $\{(1, \theta) \mid \theta \in \Theta\}$ for an infinite subset $\Theta \subseteq \mathbb{S}$ is closed and bounded, but not compact in the tree topology. However, if a set $A \subseteq \mathbb{R}^d$ is bounded and closed in the tree topology, and the set $\{\theta \in \mathbb{S} \mid \exists r > 0 : r\theta \in A\}$ is finite, then it can be shown that the set A is compact in the tree topology.

We can define two concepts and spaces of continuity of function $x : [0, T] \to \mathbb{R}^d$ for every $T > 0$.

(a) Continuity in the Euclidean norm $\|\cdot\|$; this space is denoted by $C([0, T], \mathbb{R}^d)$, with the norm

$$
\|x\|_T := \max_{t \in [0, T]} \|x(t)\|, \qquad x \in C([0, T], \mathbb{R}^d). \tag{2.2}
$$

For $d = 1$ we simply write $C[0, T]$, instead of $C([0, T], \mathbb{R}^d)$.

(b) Continuity in the tree-metric (2.1); this space is denoted by $C_t([0, T], \mathbb{R}^d)$, with the metric

$$
\text{dist}_T(x, y) := \max_{t \in [0, T]} \text{dist}(x(t), y(t)); \qquad x, y \in C_t([0, T], \mathbb{R}^d).
$$

Fix a Borel subset $B \subseteq \mathbb{R}^d$; then int B denotes the interior of B. For any Borel (signed) measure ν on B and any function $f : B \to \mathbb{R}$, we denote by (ν, f) the integral of f over B with respect to ν. Given a Borel measurable function $V : B \to [1, \infty)$, a finite, signed Borel measure ν on B has the following *V-norm*:

$$
\|\nu\|_V := \sup_{\substack{f : B \to \mathbb{R} \\ |f| \leq V}} |(\nu, f)|.
$$

When $V \equiv 1$, this norm is called the *total variation norm* and is denoted by $\|\cdot\|_{TV}$.

For a continuous function $f : [0, T] \to \mathbb{R}$, and $\delta > 0$, we define the *modulus of continuity*:

$$\omega\big(f, \delta, [0, T]\big) := \max_{\substack{s,t\in[0,T] \\ |s-t|\leq\delta}} \big|f(t) - f(s)\big|.$$

For every $f : \mathbb{R}^d \to \mathbb{R}$, the radial derivative $f'(r, \theta)$ at (r, θ) along the ray $\mathcal{R}_\theta$ is defined by

$$f'(r, \theta) := \lim_{\varepsilon\downarrow 0} \frac{1}{\varepsilon}\big[f(r + \varepsilon, \theta) - f(r, \theta)\big], \qquad r > 0, \theta \in \mathbb{S}.$$

For $r = 0$, we can also define such (one-sided) derivative in the direction of θ at the origin:

$$f'(0, \theta) \equiv f'(0+, \theta) := \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\big[f(\varepsilon, \theta) - f(0, \theta)\big]. \qquad (2.3)$$

Similarly, we can define f'', the second-order radial derivative. For every Borel subset $A \subseteq \mathbb{S}$, we define the function $\chi_A : \mathbb{R}^d \to \mathbb{R}$ as follows:

$$\chi_A(r, \theta) := \begin{cases} r, & \text{if } r > 0, \text{ and } \theta \in A, \\ 0, & \text{otherwise} \end{cases} = 1_A(\theta)r. \qquad (2.4)$$

Throughout this article, we operate on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, with the filtration satisfying the usual conditions. The arrow $\Rightarrow$ stands for weak convergence of probability measures or random variables. For example, we write $X_n \Rightarrow X_0$ as $n \to \infty$ for random variables $X_n, n = 0, 1, 2, \ldots$ The symbol mes stands for the Lebesgue measure on the real line.

For an $\mathbb{R}_+$-valued continuous semimartingale $Y = (Y(t), t \geq 0)$ its *local time at zero* is

$$\Lambda^Y := \big(\Lambda^Y(t), t \geq 0\big), \qquad \Lambda^Y(t) := \lim_{\varepsilon\downarrow 0} \frac{1}{2\varepsilon} \int_0^t 1_{[0,\varepsilon)}\big(Y(s)\big) \, d\langle Y\rangle(s), \qquad t \geq 0.$$

2.2. Definitions of Walsh semimartingales and Walsh diffusions

Now, let us take a real-valued continuous semimartingale $U = (U(t), t \geq 0)$ with $\mathrm{mes}\{t \geq 0 \mid U(t) = 0\} = 0$ a.s.

Definition 2.1. An adapted, continuous (in the tree-topology), $\mathbb{R}^d$-valued process $X = (X(t), t \geq 0)$ is called a *semimartingale on rays driven by U*, if

$$\mathrm{mes}\big\{t \geq 0 \mid X(t) = \mathbf{0}\big\} = 0 \qquad \text{a.s.,}$$

and the norm $\|X(\cdot)\|$ of $X(\cdot)$ has the following Skorohod decomposition:

$$\big\|X(t)\big\| = U(t) + \Lambda^{\|X\|}(t), \qquad \text{where } \Lambda^{\|X\|}(t) = \max_{0\leq s\leq t}\big(U(s)\big)_-.$$

This process $\Lambda^{\|X\|} = (\Lambda^{\|X\|}(t), t \geq 0)$, which is the semimartingale local time of $\|X\|$ at zero, will be also called the *local time of X accumulated at the origin*.

Assume X is a semimartingale on rays driven by U. For every Borel subset $A \subseteq \mathbb{S}$, we can consider the following real-valued process: $\chi_A(X) = (\chi_A(X(t)), t \geq 0)$, where χ_A is defined in (2.4). It follows from Karatzas and Yan [18], Theorem 2.12(ii), that $\chi_A(X)$ is a real-valued, continuous semimartingale. In Karatzas and Yan [18], Definition 2.13, a Walsh semimartingale is defined as follows.

Definition 2.2 (Walsh semimartingale). Take a semimartingale X on rays. Assume in the sense of Definition 2.1, there exists a Borel probability measure μ on $\mathbb{S}$ such that, for every Borel $A \subseteq \mathbb{S}$, the local time of $\Lambda^{\chi_A(X)}$ at the origin satisfies the "partition of local time" property a.s.

$$\Lambda^{\chi_A(X)}(t) \equiv \mu(A)\Lambda^{\|X\|}(t), \qquad t \geq 0. \tag{2.5}$$

Then the semimartingale on rays X is called a *Walsh semimartingale* with *spinning measure* μ. See Ichiba *et al.* [16], Theorem 2.1, for a construction of such Walsh semimartingale with (2.5).

For example, as in the Introduction, if **N** and **S** are North and South Poles, respectively, then a Walsh semimartingale with $\mu = p\delta_{\mathbf{N}} + (1-p)\delta_{\mathbf{S}}$ corresponds to the skew Brownian motion on the real line for $p \in [0,1]$, where δ. is a Dirac measure; and the case $p = 1/2$ corresponds to the usual Brownian motion; and the case $p = 0$ or 1 corresponds to a reflected Brownian motion.

Now, let us fix a measurable function $\ell : \mathbb{S} \to (0, \infty]$ with $\inf_{\mathbb{S}} \ell > 0$, and define the set

$$\mathcal{I} := \left\{ r\theta \mid 0 < r < \ell(\theta), \theta \in \mathbb{S} \right\} \cup \{\mathbf{0}\}. \tag{2.6}$$

This includes the case when some or all of the values of $\ell(\theta)$ are infinite. For example, if $\ell(\cdot) \equiv +\infty$, then $\mathcal{I} = \mathbb{R}^d$. The set $\mathcal{I}$ from (2.6) is open in the tree-topology, with the boundary

$$\partial \mathcal{I} := \left\{ \ell(\theta)\theta \mid \theta \in \mathbb{S}, \ell(\theta) < \infty \right\}. \tag{2.7}$$

Take Borel measurable functions $g : \mathcal{I} \to \mathbb{R}$ and $\sigma : \mathcal{I} \to (0, \infty)$, and a Borel probability measure μ on $\mathbb{S}$. Let $W = (W(t), t \geq 0)$ be an $(\mathcal{F}_t)_{t \geq 0}$-Brownian motion in one dimension.

Definition 2.3. An $\mathcal{I}$-valued continuous adapted process X is called a *Walsh diffusion associated with the triple* (g, σ, μ), if this is a Walsh semimartingale with spinning measure μ, driven by

$$U(t) := \|X(0)\| + \int_0^t \left[g\big(X(s)\big)\, \mathrm{d}s + \sigma\big(X(s)\big)\, \mathrm{d}W(s) \right], \qquad t \geq 0. \tag{2.8}$$

In this case, we say that g is the *drift coefficient*, σ^2 is the *diffusion coefficient*, and μ is the *spinning measure* for X. For the case $g \equiv 0$ and $\sigma \equiv 1$, this is called *Walsh Brownian motion*, associated with $(0, 1, \mu)$, or simply *Walsh Brownian motion with spinning measure* μ.

Remark 2.2. The *effective state space* of a Walsh diffusion in $\mathcal{I}$ associated with (g, σ, μ), or any Walsh semimartingale with spinning measure μ, is the following set:

$$\mathcal{I}^{\mu} := \left\{ (r, \theta) \mid 0 < r < \ell(\theta), \theta \in \mathrm{supp}\, \mu \right\} \cup \{\mathbf{0}\}. \tag{2.9}$$

In other words, we need to consider only rays $\mathcal{R}_\theta$ which correspond to θ in the support of measure μ. We always start this Walsh diffusion from $X(0) = x \in \mathcal{I}^\mu$, and define the function ℓ only on supp μ, and the coefficients g and σ only on the set (2.9). This distinction becomes important in Section 5 of the current paper. In Sections 2–4, we just assume that $\mathcal{I}$ is the state space.

2.3. Existence and uniqueness of Walsh diffusions

Definitions 2.2 and 2.3 are adapted from Karatzas and Yan [18], and extended to $\mathbb{R}^d$. We shall impose some assumptions.

Definition 2.4. A function $\varphi : \mathcal{I} \to \mathbb{R}$ is called *locally bounded* if $\sup_K |\varphi| < \infty$ for every measurable function $R : \mathbb{S} \to (0, \infty)$ such that

$$K := \left\{ r\theta \mid 0 < r \le R(\theta) \right\} \cup \{\mathbf{0}\} \subseteq \mathcal{I},$$

or, equivalently, $R(\theta) < \ell(\theta)$ for $\theta \in \mathbb{S}$.

Assumption 2.1. The functions g, σ, σ^{-1} are locally bounded on $\mathcal{I}$.

Under Assumption 2.1, it was proved in Karatzas and Yan [18], Section 3, that there exists a weak version of the Walsh diffusion on $\mathcal{I}$, associated with the triple (g, σ, μ), up to the *explosion time*, that is, the first passage time of $\partial\mathcal{I}$ in (2.7). To simplify exposition, we make the following assumption.

Assumption 2.2. For every initial condition $X(0) = x \in \mathcal{I}$, there exists a weak version, unique in law, of the Walsh diffusion X in $\mathcal{I}$, associated with (g, σ, μ). That is, the moment τ of hitting the boundary $\partial\mathcal{I}$ is a.s. infinite, that is, $\mathbb{P}(\tau = \infty) = 1$.

As mentioned earlier, a Walsh diffusion associated with (g, σ, μ) behaves on any ray $\mathcal{R}_\theta$ as a solution of a one-dimensional SDE with drift $g(\cdot, \theta)$ and diffusion $\sigma^2(\cdot, \theta)$, as long as it does not hit the origin. When this process hits the origin, it instantaneously chooses the new ray according to the spinning measure μ (or, more precisely, according to the formula (2.5), which is apportioning the local time at the origin between the rays), independently of the past behavior.

It was shown in Karatzas and Yan [18], Proposition 4.2, that under Assumptions 2.1 and 2.2, this Walsh diffusion $X = (X(t), t \ge 0)$ is, indeed, a Markov process. Denote its transition kernel by $P^t(x, \cdot), t \ge 0, x \in \mathbb{R}^d$. Define a family $\mathcal{D}_\mathcal{I}$ of measurable functions $f : \mathcal{I} \to \mathbb{R}$ which satisfy the following (a)–(c):

(a) For every $\theta \in \mathbb{S}$, the function $f(\cdot, \theta)$ is $C^2((0, \ell(\theta)))$, and continuous at $r = 0$;

(b) For every $\theta \in \mathbb{S}$, the derivative $f'(0+, \theta)$ from (2.3) is well-defined. Moreover, the function $\theta \mapsto f'(0+, \theta)$ is measurable and bounded on $\mathbb{S}$;

(c) There exists an $\varepsilon > 0$ such that $\sup_{\theta \in \mathbb{S}, r \le \varepsilon} |f''(r, \theta)| < \infty$.

This is a more restrictive class of functions than the one described in Karatzas and Yan [18], Definition 2.6; however, it will suffice for our purposes.

From Karatzas and Yan [18], Theorem 2.12 (suitably adapted to the context of the sphere $\mathbb{S}$ instead of a unit circle), and Definition 2.3, we get the following version of the Itô–Tanaka formula

$$f\big(X(t)\big) = f\big(X(0)\big) + \int_0^t \left[g\big(X(s)\big) f'\big(X(s)\big) + \frac{1}{2}\sigma^2\big(X(s)\big) f''\big(X(s)\big) \right] ds$$

$$+ \int_0^t \sigma\big(X(s)\big) f'\big(X(s)\big) \, dW(s) \tag{2.10}$$

$$+ \left[\int_{\mathbb{S}} f'(0+,\theta)\mu(d\theta) \right] \Lambda^{\|X\|}(t), \qquad t \ge 0$$

for every function $f \in \mathcal{D}_{\mathcal{I}}$. Here, the one-dimensional Brownian motion W is taken from (2.8). Using this version of the Itô–Tanaka formula, one can prove the following statement.

Proposition 2.1. *Under Assumptions* 2.1 *and* 2.2, *the generator* $\mathcal{L}$ *of the Walsh diffusion associated with* (g,σ,μ) *is given by*

$$\mathcal{L}f(r,\theta) = g(r,\theta)f'(r,\theta) + \frac{1}{2}\sigma^2(r,\theta)f''(r,\theta), \qquad x = (r,\theta) \in \mathbb{R}^d \tag{2.11}$$

for the following class of functions f *in*

$$\mathcal{D}_{\mathcal{I},\mu} := \left\{ f \in \mathcal{D}_{\mathcal{I}} \,\Big|\, \int_{\mathbb{S}^d} f'(0+,\theta)\mu(d\theta) = 0 \right\}. \tag{2.12}$$

Remark 2.3. We can also consider the generator $\mathcal{L}$ in (2.11) with respect to Euclidean topology; then we need to take functions $f \in \mathcal{D}_{\mathcal{I},\mu}$ on $\mathbb{R}^d$ which are continuous in Euclidean topology.

2.4. Digression into one-dimensional theory

The content of this section is taken from Karatzas and Shreve [17], Section 5.5. Consider the one-dimensional stochastic differential equation (SDE)

$$dZ(t) = g\big(Z(t)\big) dt + \sigma\big(Z(t)\big) dW(t), \qquad t \ge 0, \tag{2.13}$$

where the coefficients $g : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to (0,\infty)$ are locally bounded. From Engelbert-Schmidt theory, this type of SDE has a unique in law weak solution up to the explosion time, for every initial condition $Z(0) = z$. One powerful tool to study this type of SDE is the *scale function*

$$s(x) := \int_0^x \exp\left(-2\int_0^u \frac{g(z)}{\sigma^2(z)} \, dz \right) du,$$

for $x \in \mathbb{R}$. This scale function is strictly increasing on the whole real line, and so we can define its *inverse* $s^{-1} : (s(-\infty), s(\infty)) \to \mathbb{R}$. If we apply the scale function to a solution Z of (2.13),

we remove the drift coefficient from (2.13) and get a continuous local martingale $\tilde{Z} := s(Z) = (s(Z(t)), t \geq 0)$. More precisely, the process $\tilde{Z} = s(Z)$ is a solution of the following SDE

$$d\tilde{Z}(t) = v\big(\tilde{Z}(t)\big) \, dW(t), \qquad t \geq 0, \tag{2.14}$$

where we define the *speed function* $v(x) := \sigma(s^{-1}(x))s'(s^{-1}(x))$, $x \in \mathbb{R}$.

Next, recall the concept of a time-change to make a Brownian motion from this local martingale $\tilde{Z}$. For simplicity of notation, let us assume that $g \equiv 0$, and hence $s(x) \equiv x$, and $v(x) \equiv \sigma(x)$, $x \in \mathbb{R}$. The time-change is defined as $\mathcal{T}(t) := \int_0^t \sigma^2(Z(s)) \, ds$, $t \geq 0$. By definition of σ this is a strictly increasing function, and one can find a one-dimensional Brownian motion $B = (B(t), t \geq 0)$ such that $Z(t) = B(\mathcal{T}(t))$. Thus, we make a linear Brownian motion from a solution of the one-dimensional SDE (2.13) in two steps: (a) removal of drift coefficient by applying the scale function; (b) standardization of diffusion coefficient by applying the time-change.

2.5. Scale functions and time-change for Walsh diffusions

Same techniques as described in Section 2.4 can be used for Walsh diffusions, in principle. However, we need to adjust for dependency of drift and diffusion coefficients on the angular coordinate $\theta \in \mathbb{S}$.

First, let us recall the theory of scale functions for Walsh diffusions, developed in Karatzas and Yan [18], Section 3.3. Take a Walsh diffusion on $\mathcal{I}$ associated with (g, σ, μ), which satisfies Assumptions 2.1 and 2.2. For $(r, \theta) \in \mathcal{I}$, define the *scale function*:

$$s(r, \theta) := \int_0^r \exp\left(-2 \int_0^u \frac{g(z, \theta)}{\sigma^2(z, \theta)} \, dz\right) du. \tag{2.15}$$

Under Assumptions 2.1 and 2.2, the expression (2.15) is well defined. Moreover, $s(\cdot, \theta)$ is strictly increasing for every $\theta \in \mathbb{S}$. Thus for every $\theta \in \mathbb{S}$, there exists an inverse function $s^{-1}(\cdot, \theta)$ such that

$$s\big(s^{-1}(r, \theta), \theta\big) \equiv r, \qquad r \geq 0, \theta \in \mathbb{S}. \tag{2.16}$$

Then the function

$$\mathcal{P} : \mathcal{I} \ni (r, \theta) \mapsto \big(s(r, \theta), \theta\big) \in \tilde{\mathcal{I}} := \big\{(r, \theta) \mid 0 < r < s\big(\ell(\theta), \theta\big), \theta \in \mathbb{S}\big\} \cup \{\mathbf{0}\} \tag{2.17}$$

is a one-to-one mapping. The function (2.17) maps the Walsh diffusion on $\mathcal{I}$ associated with (g, σ, μ) into the Walsh diffusion associated with $(0, \tilde{\sigma}, \mu)$, where the new coefficient $\tilde{\sigma}$ is given by

$$\tilde{\sigma}(r, \theta) = s'\big(s^{-1}(r, \theta), \theta\big)\sigma\big(s^{-1}(r, \theta), \theta\big) \qquad \text{for } (r, \theta) \in \tilde{\mathcal{I}} \tag{2.18}$$

analogous to the speed function. In other words, just like for the one-dimensional SDE in (2.13), applying the scale function (2.15) to the drifted Walsh diffusion would remove the drift coefficient.

Next, let us make a time-change, as in Karatzas and Yan [18], Section 3.2. Assume for nota-tional convenience that the Walsh diffusion already had zero drift coefficient, that is, it is asso-ciated with the triple $(0, \sigma, \mu)$. Then $s(r, \theta) \equiv r$, and $\tilde{\sigma}(r, \theta) \equiv \sigma(r, \theta)$ for $(r, \theta) \in \mathcal{I}$. Define the time-change

$$\mathcal{T}(t) = \int_0^t \sigma^2\big(X(s)\big) \, ds, \qquad t \geq 0.$$

This is a strictly increasing function, and there exists a Walsh Brownian motion $B = (B(t), t \geq 0)$ with spinning measure μ such that $X(t) \equiv B(\mathcal{T}(t))$, $t \geq 0$.

3. Dirichlet forms approach and stationary measures

Another way to define a Walsh diffusion is using Dirichlet forms, via *one-point reflection*. This method was designed in Chen and Fukushima [4], Section 4, to construct Walsh Brownian mo-tion in Chen and Fukushima [4], Section 5. It is also developed in Chen and Fukushima [3], Section 7.5, and used in Chen and Fukushima [3], Section 7.6, Example 3, to construct general Walsh diffusions with finitely supported spinning measure μ. With minor changes, it is applicable to general Walsh diffusions. We shall merely outline the construction here, referring the reader to Chen and Fukushima [3], Section 7.6, Example 3, for all details. A benefit of this method is that it gives us a stationary distribution. Note that this method allows us to define Walsh diffusions starting from mes $\otimes \mu$-a.e. point $x \in \mathcal{I}$. Assume we have the same parameters (μ, g, σ) and the domain $\mathcal{I}$, as before.

3.1. Construction of a Walsh diffusion using Dirichlet forms

For μ-a.a. $\theta \in \mathbb{S}$, define the process $\tilde{X}_\theta$ on $[0, \ell(\theta))$ which behaves as a solution of an SDE with drift coefficient $g(\cdot, \theta)$ and diffusion coefficient $\sigma^2(\cdot, \theta)$, absorbed at $x = 0$.

Assumption 3.1. For μ-a.a. $\theta \in \mathbb{S}$, the process $\tilde{X}_\theta$ is conservative.

Under Assumption 2.1, this Assumption 3.1 is equivalent to the assumption that $\tilde{X}_\theta$ does not reach $\ell(\theta)$ in a finite time a.s. Under Assumption 3.1 define a process $\tilde{X}$ on $\mathcal{I}$, as follows: if $\tilde{X}(0) \neq \mathbf{0}$, and if $\theta := \arg(\tilde{X}(0))$, then $\tilde{X}(t) = \tilde{X}_\theta(t)\theta$ for $0 \leq t < \inf\{s : \tilde{X}_\theta(s) = \mathbf{0}\}$. In words: the process $\tilde{X}$ stays on the same ray $\mathcal{R}_\theta$, and evolves there as a one-dimensional diffusion process with drift coefficient $g(\cdot, \theta)$ and diffusion coefficient $\sigma^2(\cdot, \theta)$, killed at the origin. For notational convenience, we identify the origin $\mathbf{0}$ with the cemetery state Δ, as long as we talk about $\tilde{X}$.

Fix a $\theta \in \mathbb{S}$. Let $\tilde{P}_\theta = (\tilde{P}_\theta^t)_{t \geq 0}$ be the transition semigroup of the process $\tilde{X}_\theta$. From the the-ory of one-dimensional SDE, it is known that this semigroup is symmetric with respect to the measure π_θ on $[0, \ell(\theta))$, given by

$$\pi_\theta(dr) = \sigma^{-2}(r, \theta) \exp\left(2 \int_0^r \frac{g(\rho, \theta)}{\sigma^2(\rho, \theta)} \, d\rho\right) dr. \tag{3.1}$$

This means that the semigroup $\tilde{P}_\theta$ satisfies

$$\int_0^{\ell(\theta)} \tilde{P}_\theta^t f(r)g(r)\pi_\theta(\mathrm{d}r) = \int_0^{\ell(\theta)} \tilde{P}_\theta^t g(r)f(r)\pi_\theta(\mathrm{d}r) \tag{3.2}$$

for all bounded measurable functions $f, g : [0, \ell(\theta)) \to \mathbb{R}$ with compact support. This additional condition on f and g is introduced to make the integrals in (3.2) finite. Indeed, the measure π_θ might be infinite, as for the case of Walsh Brownian motion, $g \equiv 0$ and $\sigma \equiv 1$.

Let $\tilde{P} = (\tilde{P}^t)_{t\geq 0}$ be the transition semigroup of the process $\tilde{X}$. Define $f_\theta(r) := f(r, \theta)$. Then

$$\tilde{P}^t f(r, \theta) = \tilde{P}_\theta^t f_\theta(r). \tag{3.3}$$

Indeed, as mentioned above, the process $\tilde{X}$ always stays on the same ray $\mathcal{R}_\theta$ and behaves there as the process $\tilde{X}_\theta$. Integrating (3.2) with respect to the measure μ and using (3.3), we get: for any bounded measurable functions $f, g : \mathcal{I} \to \mathbb{R}$ with bounded support,

$$\int_{\mathcal{I}} \tilde{P}^t f(x)g(x)\pi(\mathrm{d}x) = \int_{\mathbb{S}} \int_0^{\ell(\theta)} \tilde{P}_\theta^t f_\theta(r)g(r, \theta)\pi_\theta(\mathrm{d}r)\mu(\mathrm{d}\theta)$$

$$= \int_{\mathbb{S}} \int_0^{\ell(\theta)} \tilde{P}_\theta^t g_\theta(r)f(r, \theta)\pi_\theta(\mathrm{d}r)\mu(\mathrm{d}\theta) = \int_{\mathcal{I}} \tilde{P}^t g(x)f(x)\pi(\mathrm{d}x), \tag{3.4}$$

where the new measure π on $\mathcal{I}$ is defined as

$$\pi(\mathrm{d}r, \mathrm{d}\theta) = \pi_\theta(\mathrm{d}r)\mu(\mathrm{d}\theta)$$

$$= \sigma^{-2}(r, \theta) \exp\left(2 \int_0^r \frac{g(\rho, \theta)}{\sigma^2(\rho, \theta)} \,\mathrm{d}\rho\right) \mathrm{d}r\,\mu(\mathrm{d}\theta), \qquad \pi(\{\mathbf{0}\}) = 0. \tag{3.5}$$

Therefore, $\tilde{X}$ is symmetric w.r. to π. Let us translate the notation from Chen and Fukushima [4] into our notation. The state space is $E := \mathcal{I}$. The lifetime ζ of $\tilde{X}$ is the first hitting time of the origin:

$$\zeta := \inf\{t \geq 0 \mid \tilde{X}(t) = \mathbf{0}\}.$$

Indeed, the process $\tilde{X}$ can get killed only by reaching the origin. The reason for this is that for every $\theta \in \mathbb{S}$, the process $\tilde{X}_\theta$ is conservative, that is, it does not reach $\ell(\theta)$ in finite time a.s. The new point a is the origin: $a = \mathbf{0}$. We denote Markov transition probabilities for the process $\tilde{X}$ by $\tilde{P}_x(A) = \mathbb{P}(A \mid \tilde{X}(0) = x)$ for $x \in \mathcal{I}$ and Borel subsets $A \subseteq \mathcal{I}$. Applying the one-point reflection construction from Chen and Fukushima [3], Section 7.5, we get a Walsh diffusion associated with the triple (g, σ, μ), as a Markov process with an infinitesimal generator (2.11). It is conservative (i.e., non-exploding), and its transition semigroup $(P^t)_{t\geq 0}$ satisfies $P^t 1 = 1, t \geq 0$.

Under Assumption 2.1, let us now present Assumption 2.2 in a slightly weakened form.

Assumption 3.2. For mes $\otimes\mu$-a.e. starting point $x \in \mathcal{I}$, the Walsh diffusion on $\mathcal{I}$ associated with (g, σ, μ) is conservative.

Lemma 3.1. *Under Assumption* 2.1, *Assumption* 3.1 *is equivalent to Assumption* 3.2.

Proof. Under Assumptions 2.1 and 3.1, non-explosiveness of the Walsh diffusion follows from the construction above. Conversely, suppose Assumption 3.1 were not satisfied for some $\theta \in A$, where $A \subseteq \mathbb{S}$ is a Borel subset with $\mu(A) > 0$. Start the Walsh diffusion associated with (g, σ, μ) from any point $x = (r, \theta) \in (0, \infty) \times A$. Then with positive probability it does not exit the ray $\mathcal{R}_\theta$, and hence with positive probability it would reach the point $\ell(\theta)\theta \in \partial \mathcal{I}$. Thus, this Walsh diffusion would not be conservative for this starting point, although the set $(0, \infty) \times A$ of such points has positive mes $\otimes \mu$-measure, which contradicts Assumption 3.2. Thus, these assumptions are equivalent. $\qquad\square$

3.2. Stationary distributions and measures

Take an $\mathbb{R}^d$-valued continuous-time Markov process $X = (X(t), t \geq 0)$, with transition kernel $P^t(x, \cdot)$. We shall distinguish two measures below.

Definition 3.1 (Stationary distribution). We say that a Borel probability measure π on $\mathbb{R}^d$ is a *stationary distribution* for X if the process X starting from the initial distribution π, forever remains at the same distribution π (we write $X(t) \sim \pi$), that is, $X(t) \sim \pi$ for every $t \geq 0$, if $X(0) \sim \pi$; Or equivalently for every bounded measurable function $f : \mathbb{R}^d \to \mathbb{R}$, and every $t > 0$

$$\left(\pi, P^t f\right) = (\pi, f), \tag{3.6}$$

where we use the notation $(\mu, f) = \int_{\mathbb{R}^d} f(x) \mu(\mathrm{d}x)$ of f with respect to the measure μ.

Definition 3.2 (Stationary measure). A σ-finite Borel measure π on $\mathbb{R}^d$ with $\pi(K) < \infty$ for bounded subset K is called a *stationary measure* for the Markov process X if the equality (3.6) holds for every bounded measurable function $f : \mathbb{R}^d \to \mathbb{R}$ with bounded support.

Now take $X = (X(t), t \geq 0)$ to be the Walsh diffusion on $\mathcal{I}$ associated with (g, σ, μ). By construction in Section 3, the new process X is symmetric with respect to the measure π from (3.5). That is, if $(P^t)_{t \geq 0}$ is its transition semigroup, then for any bounded measurable function $f, g : \mathcal{I} \to \mathbb{R}$ with bounded support, and any $t > 0$,

$$\int_{\mathcal{I}} \left(P^t f\right)(x) g(x) \pi(\mathrm{d}x) = \int_{\mathcal{I}} f(x) \left(P^t g\right)(x) \pi(\mathrm{d}x). \tag{3.7}$$

Since X is conservative and it does not explode, that is, $P^t 1 = 1$, letting $g \equiv 1$ in (3.7), we obtain the following theorem.

Theorem 3.1. *The measure π from* (3.5) *is a stationary measure for the Walsh diffusion on $\mathcal{I}$ associated with the triple (g, σ, μ).*

The following discussion clarifies a "physical sense" of stationary measure π with $\pi(\mathcal{I}) < \infty$. We may normalize π to be a stationary distribution. For μ-almost all $\theta \in \mathbb{S}$, $\pi_\theta([0, \ell(\theta))) < \infty$. Now, fix a $\theta \in \mathbb{S}$. Take a reflected one-dimensional diffusion on $[0, \ell(\theta))$ with drift coefficient $g(\cdot, \theta)$ and diffusion coefficient $\sigma^2(\cdot, \theta)$, reflected at 0, which, by Assumption 3.1, never hits $\ell(\theta)$. Then this process has a unique stationary distribution

$$\pi_\theta(\mathrm{d}r) = C^{-1}(\theta)\sigma^{-2}(r, \theta)\exp\left(2\int_0^r \frac{g(\rho, \theta)}{\sigma^2(\rho, \theta)}\,\mathrm{d}\rho\right)\mathrm{d}r$$

$$=: C^{-1}(\theta)p(r, \theta)\,\mathrm{d}r, \qquad (r, \theta) \in \mathcal{I}, \tag{3.8}$$

with the normalizing constant $C(\theta) := \int_0^{\ell(\theta)} p(r, \theta)\,\mathrm{d}r$. The stationary distribution can then be

$$\pi(\mathrm{d}r, \mathrm{d}\theta) = C(\theta)\pi_\theta(\mathrm{d}r)\mu(\mathrm{d}\theta), \qquad (r, \theta) \in \mathcal{I}. \tag{3.9}$$

Remark 3.1. An informal interpretation of (3.9) is as follows. The stationary distribution of the given Walsh diffusion is a combination of all radial one-dimensional stationary distributions of the radial processes, weighted by the spinning measure μ (governing how often at the origin the process chooses a given direction), and by $C(\theta)$, the average excursion time for the ray $\mathcal{R}_\theta$.

To simplify, let us assume the support $\operatorname{supp}\mu = \{\theta_1, \ldots, \theta_p\}$ is finite. Then (3.9) becomes

$$\pi(\mathrm{d}r, \{i\}) = C_i\mu_i\pi_i(\mathrm{d}r), \qquad i = 1, \ldots, p, r \in [0, \ell_i]. \tag{3.10}$$

Here, $C_i := C(\theta_i)$, $\pi_i := \pi_{\theta_i}$, $\ell_i := \ell(\theta_i)$ and $\mu_i := \mu(\theta_i)$ for $i = 1, \ldots, p$. Let us interpret the stationary distribution as a long-term average occupation time. Then for $A \subseteq [0, R_i]$

$$\pi(A \times \{i\}) = \lim_{T\to\infty}\frac{1}{T}\int_0^T \mathbf{1}\big(\arg X(s) = \theta_i, \|X(s)\| \in A\big)\,\mathrm{d}s, \qquad i = 1, \ldots, p,$$

thanks to the average occupation times formula. In particular, letting $A = [0, \ell_i]$, we get

$$C_i\mu_i = \pi([0, \ell_i] \times \{i\}) = \lim_{T\to\infty}\frac{1}{T}\operatorname{mes}\{s \in [0, T] \mid \arg X(s) = \theta_i\}.$$

Here, μ_i is the long-term proportion of the times this Walsh diffusion chooses the ray $\ell_i := \ell_{\theta_i}$, and C_i is the factor corresponding to the average time spent on this ray ℓ_i, for each such excursion.

Example 3.1. Let $\mathcal{I} := \mathbb{R}^d$, $g(r, \theta) \equiv g(\theta) < 0$, $\sigma(r, \theta) \equiv \sigma(\theta) > 0$ for every $r > 0$, $\theta \in \mathbb{S}$. Denote $\lambda(\theta) := -2g(\theta)\sigma^{-2}(\theta) > 0$ for $\theta \in \mathbb{S}$. Then direct calculation gives us the scale function $s(\cdot)$ and the stationary measure $\pi(\cdot)$

$$s(r, \theta) = \frac{1}{\lambda(\theta)}\big(e^{\lambda(\theta)r} - 1\big), \qquad \pi(\mathrm{d}r, \mathrm{d}\theta) = \sigma^{-2}(\theta)\exp\big(-\lambda(\theta)r\big)\,\mathrm{d}r\,\mu(\mathrm{d}\theta), \qquad r > 0, \theta \in \mathbb{S}.$$

4. Continuous dependence on spinning measure and initial condition

4.1. Main results

In this section, using Euclidean distance $\|\cdot\|_T$ from (2.2) for $C([0, T], \mathbb{R}^d)$, we show that the law of the Walsh diffusion on $\mathcal{I}$ associated with (g, σ, μ), starting from $X(0) = x$, continuously (in the weak sense) depends on the measure μ, and on the initial condition x.

Assumption 4.1. The functions $g, \sigma : \mathcal{I} \to \mathbb{R}$ are continuous in the Euclidean topology. Moreover, the function $\ell : \mathbb{S} \to (0, \infty]$ is lower semicontinuous, that is, for every $\theta \in \mathbb{S}$,

$$\varliminf_{\theta \to \theta_0} \ell(\theta) \geq \ell(\theta_0).$$

Remark 4.1. The assumption on ℓ can be equivalently stated as follows: the function ℓ can be pointwise approximated by an increasing sequence $(R_n : \mathbb{S} \to (0, \infty))_{n \geq 1}$ of continuous functions

$$R_n(\theta) \uparrow \ell(\theta) \qquad \text{as } n \to \infty \text{ for every } \theta \in \mathbb{S}.$$

In this case, define

$$D_m := \left\{ (r, \theta) \mid 0 < r \leq R_m(\theta), \theta \in \mathbb{S} \right\} \cup \{\mathbf{0}\}, \qquad m \geq 1. \tag{4.1}$$

Then $(D_m)_{m \geq 1}$ is an increasing sequence of compact (in a Euclidean topology) subsets of $\mathcal{I}$ with each interior $\mathrm{int}\, D_m := \{(r, \theta) \mid 0 < r < R_m(\theta), \theta \in \mathbb{S}\} \cup \{\mathbf{0}\}$. By construction, this sequence satisfies the properties

$$\mathrm{int}\, D_m \subseteq \mathrm{int}\, D_{m+1}, \quad \text{and} \quad \bigcup_{m=1}^{\infty} \mathrm{int}(D_m) = \mathcal{I}.$$

Assumption 4.1 holds, for example, when $\ell(\cdot) \equiv \infty$, or when $\ell : \mathbb{S} \to (0, \infty)$ is continuous. Assumption 4.1 does not hold, for example, $\ell(\theta) := 1 + \mathbf{1}_{\{\theta = \theta_0\}}$ for $d \geq 2$ and a fixed point $\theta_0 \in \mathbb{S}$.

Remark 4.2. Assumption 2.1 actually follows from Assumption 4.1. Indeed, take a function R as in Definition 2.4, and approximate the function ℓ by R_n, $n \geq 1$ as in Remark 4.1. Then we have

$$R(\theta) < \ell(\theta) = \lim_{n \to \infty} R_n(\theta), \qquad \forall \theta \in \mathbb{S}.$$

Thus the open sets $O_n := \{\theta \in \mathbb{S} \mid R(\theta) < R_n(\theta)\}$, $n \geq 1$ form an open cover of $\mathbb{S}$. By compactness, we can extract a finite subcover $O_{n_1}, \ldots, O_{n_j}$ of $\mathbb{S}$. Then there exists an $m = \max(n_1, \ldots, n_j)$ such that for all $\theta \in \mathbb{S}$ we have $R(\theta) < R_m(\theta)$. Therefore, $K \subseteq D_m$, where K is defined in Definition 2.4, and D_m is taken from (4.1). Since g is continuous and D_m is compact

in Euclidean topology, we conclude g is bounded on D_m, and so is on K. The same argument applies to σ and σ^{-1}.

Theorem 4.1. *Let us consider the Walsh diffusion X_n on $\mathcal{I}$ associated with (g, σ, μ_n), starting from $X_n(0) = x_n$, $n \geq 0$. For every $n = 0, 1, 2, \ldots$, take a point $x_n \in \mathcal{I}$ in (2.6) and a spinning measure μ_n on $\mathbb{S}$. Suppose that the functions g and σ, and the domain $\mathcal{I}$, satisfy Assumption 4.1. Suppose also that every Walsh diffusion X_n satisfies Assumption 2.2. If $\mu_n \Rightarrow \mu_0$, and $x_n \to x_0$, then $X_n \Rightarrow X_0$ in $C([0, T], \mathbb{R}^d)$.*

Let us state separately this convergence result for Walsh Brownian motions in $\mathbb{R}^d$. For this case, $g \equiv 0$, $\sigma \equiv 1$, and $\ell \equiv \infty$, and Theorem 4.1 takes the following form.

Corollary 4.1. *For every $n = 0, 1, 2, \ldots$, take a Walsh Brownian motion W_n with spinning measure μ_n, starting from $W_n(0) = x_n$. If $x_n \to x_0$, and $\mu_n \Rightarrow \mu_0$, then $W_n \Rightarrow W_0$ in $C([0, T], \mathbb{R}^d)$.*

Consider the case when all these Walsh Brownian motions in $\mathbb{R}^d$ start from the origin: $x_n = \mathbf{0}$, $n = 0, 1, 2, \ldots$. Then we can actually quantify the rate of convergence in the following Theorem 4.2. For two Borel probability measures ν_1 and ν_2 on a metric space $(\mathcal{X}, \mathbf{d})$ with metric $\mathbf{d}$, the *Wasserstein distance of order $p \geq 1$* is defined as follows:

$$W_p(\nu_1, \nu_2) := \inf_{\gamma \in \Gamma} \left[\int_{\mathcal{X} \times \mathcal{X}} \big(\mathbf{d}(x_1, x_2)\big)^p \, d\gamma(x_1, x_2) \right]^{1/p}, \qquad (4.2)$$

where the infimum is taken over the family Γ of probability measures on $\mathcal{X} \times \mathcal{X}$ with marginals ν_1 and ν_2 for which the integral inside the bracket is finite. It is known from Rachev [24], Villani [33] that convergence in the Wasserstein distance implies the weak convergence. Thus, we shall estimate the Wasserstein distance between the distributions of X_n and X_0, as random elements of $C([0, T], \mathbb{R}^d)$, using the Wasserstein distance between μ_n and μ_0 on $\mathbb{S}$.

Denote by $\mathbb{Q}_T(\mu)$ the law of the Walsh Brownian motion starting from the origin with spinning measure μ in the space $C([0, T], \mathbb{R}^d)$. Theorem 4.2 provides the upper estimate of the Wasserstein distance between $\mathbb{Q}_T(\mu)$ and $\mathbb{Q}_T(\overline{\mu})$, for two spinning measures μ and $\overline{\mu}$ on $\mathbb{S}$.

Theorem 4.2. *Take two Borel probability measures μ and $\overline{\mu}$ on $\mathbb{S}$. For all positive constants p, q, ρ, T with*

$$1 \leq q < p, \qquad \rho \in \left(0, \frac{p}{p+1}\right), \qquad (4.3)$$

there exists a positive constant C^* dependent on T, p, q, ρ, such that

$$W_q\big(\mathbb{Q}_T(\mu), \mathbb{Q}_T(\overline{\mu})\big) \leq C^* \big[W_p(\mu, \overline{\mu})\big]^\rho. \qquad (4.4)$$

The rest of this section is organized as follows. We shall prove Theorem 4.2 in Section 4.2, Corollary 4.1 in Section 4.3 and then complete the proof of Theorem 4.1 in Section 4.4.

4.2. Proof of Theorem 4.2

Step 1: Efficient coupling. Let us consider the following Walsh Brownian motions $X = (X(t), t \geq 0)$ and $\overline{X} = (\overline{X}(t), t \geq 0)$ with spinning measures μ and $\overline{\mu}$, respectively, coupled by the following procedure.

On a filtered probability space, let us take a reflected Brownian motion $S = (S(t), t \geq 0)$ with values in $[0, \infty)$, starting from zero, instantaneously reflected at zero. Let $e = (e(t), t \geq 0)$ be the excursion process, so that $e(t)$ is the excursion of the reflected Brownian motion S at time t, and denote by $\mathcal{J}$ the (countable) set of its distinct elements. This set depends on $\omega \in \Omega$. The notation is taken from Revuz and Yor [25], Chapter 12. Take a certain coupling Θ of marginal probability measures μ and $\overline{\mu}$. Generate a sequence $(\theta_j, \overline{\theta}_j)$ of $\mathbb{S} \times \mathbb{S}$-valued, independently, identically distributed random variables jointly distributed in Θ, indexed by $j \in \mathcal{J}$. Define the Walsh Brownian motions X and $\overline{X}$ as follows: for each $t \geq 0$, if $S(t) = 0$, then $X(t) := 0 =: \overline{X}(t)$, and if $S(t) > 0$, then there exists a unique index $j \in \mathcal{J}$, such that $e(t) = j$, and we let

$$X(t) := \theta_j S(t), \qquad \overline{X}(t) := \overline{\theta}_j S(t) \tag{4.5}$$

on an extended, filtered probability space. Such construction of Walsh Brownian motions (and semimartingales) is recently examined and described in Ichiba *et al.* [16].

Note that we use the common reflected Brownian motion S in the construction, and so the resulting Walsh Brownian motions X and $\overline{X}$ have the same time intervals for excursions, that is, $\{t : X(t) = \mathbf{0}\} = \{t : \overline{X}(t) = \mathbf{0}\}$ a.s. This procedure creates two Walsh Brownian motions X and $\overline{X}$ with spinning measures μ and $\overline{\mu}$, respectively. In other words, the probability measure induced by this pair $(X, \overline{X})$ is a coupling $\Pi(\Theta)$ of marginal probability distributions $\mathbb{Q}_T(\mu)$ and $\mathbb{Q}_T(\overline{\mu})$ in $C([0, T], \mathbb{R}^d)$, where μ and $\overline{\mu}$ are the marginal of the coupling Θ. To achieve an efficient coupling $\Pi(\Theta)$ of X and $\overline{X}$, in the sense of smaller Wasserstein distance, we take an efficient coupling Θ of μ and $\overline{\mu}$.

Let us denote by Γ_0 all couplings of marginal probability measures μ and $\overline{\mu}$. Also, let us denote by Γ_1 the family of probability measures, each of which is induced by the coupling $(X, \overline{X})$ of distributions $\mathbb{Q}_T(\mu)$ and $\mathbb{Q}_T(\overline{\mu})$, constructed in the above procedure from a coupling Θ of μ and $\overline{\mu}$ in Γ_0, that is, $\Gamma_1 := \{\Pi(\Theta) : \Theta \in \Gamma_0\}$. Then by definition (4.2) for $p, q \geq 1$ here we shall evaluate the Wasserstein distances

$$W_p(\mu, \overline{\mu}) := \left(\inf_{\Theta \in \Gamma_0} \mathbb{E}^\Theta \left[\|\theta - \overline{\theta}\|^p \right] \right)^{1/p},$$

$$W_q\big(\mathbb{Q}_T(\mu), \mathbb{Q}_T(\overline{\mu})\big) \leq \left(\inf_{P \in \Gamma_1} \mathbb{E}^P \left[\|X - \overline{X}\|_T^q \right] \right)^{1/q} = \left(\inf_{\Theta \in \Gamma_0} \mathbb{E}^{\Pi(\Theta)} \left[\|X - \overline{X}\|_T^q \right] \right)^{1/q}, \tag{4.6}$$

where $\mathbb{E}^\Theta$ and $\mathbb{E}^P = \mathbb{E}^{\Pi(\Theta)}$ are expectations under $\Theta \in \Gamma_0$, $P = \Pi(\Theta) \in \Gamma_1$, respectively. To this end, given the constants p, q, T in (4.3), we shall estimate the upper bound of

$$\mathbb{E}\big[\|X - \overline{X}\|_T^q \big] = \mathbb{E}^{\Pi(\Theta)} \left[\max_{0 \leq t \leq T} \|X(t) - \overline{X}(t)\|^q \right], \tag{4.7}$$

in terms of $\mathbb{E}^\Theta[\|\theta - \overline{\theta}\|^p]$, for the coupling $\Pi(\Theta)$ of the pair X and $\overline{X}$ of Walsh Brownian motions, and for the coupling Θ of the spinning measures μ and $\overline{\mu}$, described in (4.5).

The idea of the proof is as follows. Take a constant $\delta(> 0)$. There are two kinds of excursions of S: the first kind consists of excursions with height greater than δ, and the second kind consists of excursions with height less than or equal to δ. There are at most finite number of excursions of the first kind. For them, we have

$$\left\| X(t) - \overline{X}(t) \right\| \le \| \theta_j - \overline{\theta}_j \| S(t), \qquad t \ge 0, \tag{4.8}$$

where we can estimate the running maximum of S from above. Take a time moment t corresponding to the second kind of excursions. By the triangle inequality and the construction in (4.5), we have

$$\left\| X(t) - \overline{X}(t) \right\| \le \left\| X(t) \right\| + \left\| \overline{X}(t) \right\| = S(t) + S(t) \le 2\delta. \tag{4.9}$$

With this idea of separating excursions in two kinds, we estimate the upper bound of (4.7), and then minimize it by choosing δ.

Step 2: Excursions. Define the set of excursions of S

$$\mathcal{J}_T := \left\{ j \in \mathcal{J} \mid \exists t \in [0, T] : e(t) = j \right\}$$

restricted to the time interval $[0, T]$, including the last excursion (which is sometimes called a *Brownian meander*). This last excursion can start at a time moment $t' \le T$, but end at $t'' > T$; this excursion is included if $\max_{t \in [t', T]} S(t) \ge \delta$. Let us classify it into two kinds of excursions

$$\mathcal{J}_{T,\delta} := \left\{ j \in \mathcal{J}_T \mid H(j) > \delta \right\}, \qquad \mathcal{J}_{T,\delta}^c := \mathcal{J}_T \setminus \mathcal{J}_{T,\delta}, \tag{4.10}$$

where $H(j)$ is the height of the excursion $j \in \mathcal{J}$. Since $\mathcal{J}_{T,\delta}$ is a finite set, by (4.9)–(4.8),

$$\max_{0 \le t \le T} \left\| X(t) - \overline{X}(t) \right\|^q \le \max_{\{t : e(t) \in \mathcal{J}_{T,\delta}\}} \left\| X(t) - \overline{X}(t) \right\|^q + \max_{\{t : e(t) \in \mathcal{J}_{T,\delta}^c\}} \left\| X(t) - \overline{X}(t) \right\|^q$$

$$\le \left(\max_{0 \le t \le T} S(t) \right)^q \cdot \max_{j \in \mathcal{J}_{T,\delta}} \| \theta_j - \overline{\theta}_j \|^q + (2\delta)^q.$$

If the set $\mathcal{J}_{T,\delta}$ is empty, we let the maximum of zero numbers to be zero. Taking the expected values and applying Hölder's inequality to the product in the right-hand side with

$$r_1 := \frac{p}{p - q}, \qquad r_2 := \frac{p}{q}, \qquad \frac{1}{r_1} + \frac{1}{r_2} = 1,$$

we obtain the first upper bound of (4.7)

$$\mathbb{E}^{\Pi(\Theta)} \left[\max_{0 \le t \le T} \left\| X(t) - \overline{X}(t) \right\|^q \right] \le (2\delta)^q + C_1 \cdot \left[\mathbb{E} \left[\max_{j \in \mathcal{J}_{T,\delta}} \| \theta_j - \overline{\theta}_j \|^p \right] \right]^{q/p}, \tag{4.11}$$

where with the Gamma function $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} \, dx$, $a > 0$,

$$C_1 := \left[\mathbb{E} \left[\left(\max_{0 \le t \le T} S(t) \right)^{q r_1} \right] \right]^{1/r_1} = \frac{(2T)^{q/2}}{\pi^{1/(2r_1)}} \cdot \left[\Gamma \left(\frac{1 + q r_1}{2} \right) \right]^{1/r_1}. \tag{4.12}$$

Here we used the density function of the running maximum of S to compute C_1 (see Karatzas and Shreve [17]).

For the second term in (4.11) it is not easy to estimate the number of elements in $\mathcal{J}_{T,\delta}$ directly. Instead, we can estimate it indirectly by the number of elements in a set which is (usually) larger than $\mathcal{J}_{T,\delta}$. Take a number $D(> 0)$ large enough, to be determined later. Define $L = (L(t), t \geq 0)$ to be the local time process of the reflected Brownian motion S at zero, i.e., $L(\cdot) := \int_0^\cdot \mathbf{1}_{\{S(t)=0\}} \, dS(t)$. Define also $L^{-1}(s) := \inf\{t \geq 0 \mid L(t) = s\}$ to be the inverse local time of S. Then the probability $\mathbb{P}(L(T) > D\sqrt{T})$ is very small. And if $L(T) \leq D\sqrt{T}$, then

$$\mathcal{J}_{T,\delta} \subseteq \mathcal{J}_{L^{-1}(D\sqrt{T}),\delta}, \tag{4.13}$$

From Revuz and Yor [25], Chapter 12, the number $|\mathcal{J}_{L^{-1}(D\sqrt{T}),\delta}|$ of elements in the set $\mathcal{J}_{L^{-1}(D\sqrt{T}),\delta}$ has Poisson distribution with parameter $\lambda := D\sqrt{T}/\delta$, that is,

$$\mathbb{E}\big[|\mathcal{J}_{L^{-1}(D\sqrt{T}),\delta}|\big] = \frac{D\sqrt{T}}{\delta}. \tag{4.14}$$

Consider two cases $\{L(T) \leq D\sqrt{T}\}$ and $\{L(T) > D\sqrt{T}\}$. By (4.13),

$$\mathbb{E}\Big[\max_{j \in \mathcal{J}_{T,\delta}} \|\theta_j - \overline{\theta}_j\|^p\Big] = \mathbb{E}\Big[\max_{j \in \mathcal{J}_{T,\delta}} \|\theta_j - \overline{\theta}_j\|^p (\mathbf{1}_{\{L(T) \leq D\sqrt{T}\}} + \mathbf{1}_{\{L(T) > D\sqrt{T}\}})\Big]$$
$$\leq \mathbb{E}\Big[\max_{j \in \mathcal{J}_{L^{-1}(D\sqrt{T}),\delta}} \|\theta_j - \overline{\theta}_j\|^p\Big] + (2\pi)^p \cdot \mathbb{P}\big(L(T) > D\sqrt{T}\big). \tag{4.15}$$

Since by the construction of the pair $(X, \overline{X})$ in (4.5) the random variables θ_j and $\overline{\theta}_j$ are independent of S, it follows from Wald's identity and (4.14) that

$$\mathbb{E}\Big[\max_{j \in \mathcal{J}_{L^{-1}(D\sqrt{T}),\delta}} \|\theta_j - \overline{\theta}_j\|^p\Big] \leq \mathbb{E}\Big[\sum_{j \in \mathcal{J}_{L^{-1}(D\sqrt{T}),\delta}} \|\theta_j - \overline{\theta}_j\|^p\Big]$$
$$= \mathbb{E}\big[|\mathcal{J}_{L^{-1}(D\sqrt{T}),\delta}|\big] \cdot \mathbb{E}^\Theta\big[\|\theta - \overline{\theta}\|^p\big]. \tag{4.16}$$

The classical Lévy theorem states that $L(T)$ has the same distribution as $S(T)$. Then by the Mills ratio of the Gaussian tail probability Feller [10], Chapter 7, for every $r \geq 1$ there exists a constant C_2 (which does not depend on D), such that

$$\mathbb{P}\big(L(T) > D\sqrt{T}\big) = \mathbb{P}\big(S(1) > D\big) \leq \frac{2}{\sqrt{2\pi}D} e^{-D^2/2} \leq \frac{C_2}{D^r}; \qquad D > 0. \tag{4.17}$$

Combining (4.14)–(4.17) together, we have

$$\mathbb{E}\Big[\max_{j \in \mathcal{J}_{T,\delta}} \|\theta_j - \overline{\theta}_j\|^p\Big] \leq \frac{D\sqrt{T}}{\delta} \cdot \mathbb{E}^\Theta\big[\|\theta - \overline{\theta}\|^p\big] + \frac{C_3}{D^r} \tag{4.18}$$

for every $D > 0$, $\delta > 0$ where $C_3 = (2\pi)^p C_2$. Thus the right-hand side of (4.11) is evaluated as

$$\mathbb{E}^{\Pi(\Theta)}\left[\max_{0 \leq t \leq T} \|X(t) - \overline{X}(t)\|^q\right] \leq C_4 \cdot \left[\delta^q + \left(\frac{D}{\delta} \cdot \mathbb{E}^{\Theta}[\|\theta - \overline{\theta}\|^p] + \frac{1}{D^r}\right)^{q/p}\right], \quad (4.19)$$

where C_4 does not depend on $\delta(> 0)$, $D(> 0)$ but on $r(\geq 1)$, $T(> 0)$, $1 \leq q < p$.

Step 3: Minimization. Since the left-hand side of (4.19) does not depend on (D, δ), let us minimize the right-hand side of (4.19) with respect to D and δ by applying twice an inequality

$$f_0(x) := a_1 x^{c_1} + a_2 x^{-c_2}$$

$$\geq \left[\left(\frac{c_2}{c_1}\right)^{\frac{c_1}{c_1+c_2}} + \left(\frac{c_1}{c_2}\right)^{\frac{c_2}{c_1+c_2}}\right] \cdot a_1^{\frac{c_2}{c_1+c_2}} \cdot a_2^{\frac{c_1}{c_1+c_2}} \quad (4.20)$$

$$= f_0(x^*)$$

for every $x > 0$, where a_i, c_i, $i = 1, 2$ are fixed positive constants and $x^* := (\frac{a_2 c_2}{a_1 c_1})^{1/(c_1+c_2)}$ is a unique minimizer of the function $f_0(\cdot)$. Applying (4.20) with

$$(x, a_1, a_2, c_1, c_2) = \left(D, \mathbb{E}^{\Theta}[\|\theta - \overline{\theta}\|^p]/\delta, 1, 1, r\right)$$

and $D^* := (\delta r/\mathbb{E}^{\Theta}[\|\theta - \overline{\theta}\|^p])^{1/(1+r)}$, we obtain

$$\left[f_0(D)\right]^{q/p} = \left(\frac{D}{\delta} \cdot \mathbb{E}^{\Theta}[\|\theta - \overline{\theta}\|^p] + \frac{1}{D^r}\right)^{q/p}$$

$$\geq C_5 \cdot \left(\frac{\mathbb{E}^{\Theta}[\|\theta - \overline{\theta}\|^p]}{\delta}\right)^{qr/(p(1+r))} = \left[f_0(D^*)\right]^{q/p},$$

and then applying (4.20) with

$$(x, a_1, a_2, c_1, c_2) = \left(\delta, 1, C_5\left(\mathbb{E}^{\Theta}[\|\theta - \overline{\theta}\|^p]\right)^{qr/(p+pr)}, q, qr/(p(1+r))\right),$$

we obtain

$$C_4 \cdot \left(\delta^q + C_5 \cdot \left(\frac{\mathbb{E}^{\Theta}[\|\theta - \overline{\theta}\|^p]}{\delta}\right)^{qr/(p(1+r))}\right) \geq C_6\left(\mathbb{E}^{\Theta}[\|\theta - \overline{\theta}\|^p]\right)^{\rho q/p} \quad (4.21)$$

with some constants C_i, $i = 5, 6$, where

$$\rho := \frac{pr}{(1+p)r + p} \in \left[\frac{p}{2p+1}, \frac{p}{1+p}\right), \qquad \text{with } r = \frac{p\rho}{p - (1+p)\rho}. \quad (4.22)$$

Now for every $\rho \in [p/(2p+1), p/(1+p))$ given, we may choose the corresponding constants $r(\geq 1)$ from (4.22), C_2 in (4.17) and then the resulting constant $C^* := C_6$ in (4.21) to obtain

$$\left[\mathbb{E}^{\Pi(\Theta)}\left[\max_{0 \leq t \leq T} \|X(t) - \overline{X}(t)\|^q\right]\right]^{1/q} \leq C^*\left(\mathbb{E}^{\Theta}[\|\theta - \overline{\theta}\|^p]\right)^{\rho/p}. \quad (4.23)$$

Taking the infimum on both sides of (4.23) over the family Γ_0 of measures for the coupling Θ on $\mathbb{S} \times \mathbb{S}$, we achieve the desired upper bound (4.4) for $\rho \in [p/(2p + 1), p/(1 + p))$. For every $\rho_1 \in (0, p/(2p + 1))$ and $\rho_2 \in [p/(2p + 1)), p/(1 + p))$, because of boundedness of $0 \leq W_p(\mu, \overline{\mu}) \leq \sup_{\theta, \overline{\theta} \in \mathbb{S}} \|\theta - \overline{\theta}\| = 2$, there exists a constant C_7 such that $[W_p(\mu, \overline{\mu})]^{\rho_2} \leq C_7 [W_p(\mu, \overline{\mu})]^{\rho_1}$. Using this equality and modifying the constant C^* for ρ_2, we obtain the desired upper bound (4.4) also for $\rho \in (0, p/(2p + 1))$. Therefore, we conclude the proof of (4.4) for every $\rho \in (0, p/(1 + p))$.

4.3. Proof of Corollary 4.1

For the case $x_n = \mathbf{0}$ for all n, this follows immediately: Convergence in the Wasserstein distance of order p is equivalent to weak convergence plus uniform boundedness of the pth moment, see Villani [33]. For measures $(\mu_n)_{n \geq 0}$ on $\mathbb{S}$, their pth moments are trivially uniformly bounded, since $\mathbb{S}$ is a bounded set. Theorem 4.2 makes the rest of the proof trivial.

Consider the general case of Corollary 4.1, with arbitrary initial conditions. By the Skorohod representation theorem, we can create a probability space with copies $\tilde{W}_n$ of Walsh Brownian motions starting from the origin, such that $\tilde{W}_n \to \tilde{W}_0$ a.s. uniformly on every $[0, T]$. For $n \geq 0$, let $\theta_n := \arg(x_n)$. Then we can couple W_n for $n = 0, 1, 2, \ldots$ as follows.

Let us take a standard Brownian motion $B = (B(t), t \geq 0)$ on the real line with $B(0) = 0$. Since $x_n \to x_0$ as $n \to \infty$ in Euclidean topology, for n large enough (and therefore without loss of generality for all n) we have: $x_n \neq \mathbf{0}$. Since $x_n \to x_0$, we have $\|x_n\| \to \|x_0\|$, $\theta_n \to \theta_0$.

Now, define stopping times $\tau_n := \inf\{t \geq 0 \mid B(t) = \|x_n\|\}$. We construct copies of Walsh Brownian motions W_n starting from $W_n(0) = x_n$, as follows:

$$W_n(t) = \begin{cases} x_n - \theta_n B(t) = \theta_n\big(\|x_n\| - B(t)\big), & t \leq \tau_n; \\ \tilde{W}_n(t - \tau_n), & t \geq \tau_n. \end{cases} \qquad (4.24)$$

Consider the inverse $\mathfrak{m}^{-1} = (\mathfrak{m}^{-1}(s), s \geq 0)$ of running maximum of a Brownian motion B starting from zero. This process is a.s. continuous at every fixed time (although it has a.s. discontinuous trajectories). We can express $\tau_n := \mathfrak{m}^{-1}(\|x_n\|)$. Therefore, a.s. $\tau_n \to \tau_0$ as $n \to \infty$. Applying Lemma A.1 in the Appendix, we conclude Corollary 4.1.

4.4. Proof of Theorem 4.1 in the general case

We split the proof into four steps. In the first three steps, we consider the driftless case: $g \equiv 0$. Step 1 is devoted to time-change and localization. That is, we consider an exhaustion of the state space $\mathcal{I}$ by an increasing sequence of compact domains D_m, $m \geq 1$. Then we fix one of these domains and stop *all* Walsh diffusions X_n when they exit the interior of the domain. In Step 2, we prove the convergence result for these stopped Walsh diffusions. In Step 3, we switch from int D_m to $\mathcal{I}$ and show that for the driftless case the convergence takes place not only for those stopped, but also for the original Walsh diffusions. Finally, in Step 4, we use the scale mapping (2.15) to extend this result to the general case with non zero drift function g.

Step 1: Time-change and localization. Assume $g \equiv 0$. We apply the time-change in Section 2.5 to the Walsh diffusions X_n on $\mathcal{I}$ associated with $(0, \sigma, \mu_n)$:

$$X_n(t) = W_n\big(T_n(t)\big), \qquad T_n(t) := \int_0^t \sigma^2\big(X_n(s)\big)\,\mathrm{d}s, \qquad t \geq 0, \tag{4.25}$$

where $W_n = (W_n(t), t \geq 0)$ is a Walsh Brownian motion with spinning measure μ_n, starting from $W_n(0) = x_n$, $n \geq 0$. By Assumption 4.1 and Remark 4.1, we may take a sequence of continuous functions $R_n : \mathbb{S} \to (0, \infty)$ such that $R_n(\theta) \to \ell(\theta)$ pointwise on $\mathbb{S}$. Recall the definition of D_m in (4.1). For $n \geq 0$ and $m \geq 1$, define the stopping times

$$\tau_n^{(m)} := \inf\big\{t \geq 0 \mid X_n(t) \notin D_m\big\} = \inf\big\{t \geq 0 \mid W_n\big(T_n(t)\big) \notin D_m\big\},$$

and the corresponding stopped processes $X_n^{(m)}(t) := X_n(t \wedge \tau_n^{(m)})$, $T_n^{(m)}(t) := T_n(t \wedge \tau_n^{(m)})$, $t \geq 0$. Then it follows from (4.25) that for $n \geq 0$ and $m \geq 1$,

$$X_n^{(m)}(t) = W_n\big(T_n^{(m)}(t)\big), \qquad t \geq 0. \tag{4.26}$$

In particular, $X_0^{(m)}(t) = W_0(T_0^{(m)}(t))$, $t \geq 0$.

Step 2: Proof for the stopped processes. The rest of the proof for the driftless case is quite similar to the proof of Sarantsev [30], Theorem 2.2. For the rest of Step 2, fix an $m \geq 1$. We shall show

Lemma 4.1. *Every subsequence of $(X_n^{(m)})_{n \geq 1}$ in (4.26) has a weakly convergent subsequence, and this weak limit behaves as the Walsh (driftless) diffusion $X_0^{(m)}$, at least as long as it stays in* int D_m.

The rest of Step 2 is devoted to proving Lemma 4.1. For every $n \geq 0$ and $t \geq 0$, we have

$$0 \leq \sigma\big(X_n^{(m)}(t)\big) \leq \max_{D_m} \sigma =: \overline{\sigma}_m, \tag{4.27}$$

and hence, by definition of $T_n(t)$ in (4.25),

$$0 \leq T_n^{(m)}(t) \leq \overline{\sigma}_m^2 t. \tag{4.28}$$

By the Arzela–Ascoli criterion combined with (4.27), the sequence $(T_n^{(m)})_{n \geq 1}$ is tight for every $m \geq 1$. Then every subsequence $(n_k)_{k \geq 1}$ of $\mathbb{N}$ has its further subsequence $(n_k')_{k \geq 1}$ for which there exists a random process $\overline{T}^{(m)} = (\overline{T}^{(m)}(t), t \geq 0)$ such that

$$T_{n_k'}^{(m)} \to \overline{T}^{(m)} \qquad \text{in } C[0, T], \text{ as } k \to \infty. \tag{4.29}$$

As we have already proved (in the proof of Corollary 4.1), we have $W_n \Rightarrow W_0$ in $C([0, \overline{\sigma}_m^2 T], \mathbb{R}^d)$, as $\mu_n \Rightarrow \mu_0$ and $x_n \Rightarrow x_0$. Changing the probability space, if necessary, by the virtue of Skorohod representation theorem, we can make this convergence a.s.

$$W_n \to W_0 \qquad \text{in } C\big([0, \overline{\sigma}_m^2 T], \mathbb{R}^d\big). \tag{4.30}$$

Combining (4.26), (4.28), (4.29) and (4.30), we have the limit

$$X_{n'_k}^{(m)}(t) \to \overline{X}^{(m)}(t) := W_0\big(\overline{T}^{(m)}(t)\big) \qquad \text{uniformly on } [0, T]. \tag{4.31}$$

Since the function σ is continuous on D_m by Assumption 4.1, we also get

$$\sigma\big(X_{n'_k}^{(m)}(t)\big) \to \sigma\big(\overline{X}^{(m)}(t)\big), \qquad k \to \infty, \text{ for every } t \in [0, T]. \tag{4.32}$$

By (4.27), (4.32), and the Lebesgue dominated convergence theorem, for every $t \in [0, T]$, we have

$$\int_0^t \sigma^2\big(X_{n'_k}^{(m)}(s)\big)\,\mathrm{d}s \to \int_0^t \sigma^2\big(\overline{X}^{(m)}(s)\big)\,\mathrm{d}s. \tag{4.33}$$

Denote $\tau_\infty^{(m)} := \underline{\lim}_{n \to \infty} \tau_n^{(m)}$, and take a (random) time point $t_0 < \tau_\infty^{(m)} \wedge T$. For every sufficiently large k, and all $s \in [0, t_0] \subseteq [0, \tau_\infty^{(m)}]$, we have $X_{n'_k}^{(m)}(s) = X_{n'_k}(s)$ and $t_0 \wedge \tau_{n'_k}^{(m)} < \tau_\infty^{(m)} \wedge T$. Thus, combining (4.25)–(4.26) with (4.33), for such t_0 we have

$$\begin{aligned}
T_{n'_k}^{(m)}(t_0) = T_{n'_k}\big(t_0 \wedge \tau_{n'_k}^{(m)}\big) &= \int_0^{t_0 \wedge \tau_{n'_k}^{(m)}} \sigma^2\big(X_{n'_k}(s)\big)\,\mathrm{d}s \\
&= \int_0^{t_0 \wedge \tau_{n'_k}^{(m)}} \sigma^2\big(X_{n'_k}^{(m)}(s)\big)\,\mathrm{d}s \to \int_0^{t_0} \sigma^2\big(\overline{X}^{(m)}(s)\big)\,\mathrm{d}s
\end{aligned}$$

as $k \to \infty$. Comparing this observation with (4.29), we get

$$\overline{T}^{(m)}(t) = \int_0^t \sigma^2\big(\overline{X}^{(m)}(s)\big)\,\mathrm{d}s, \quad \text{and} \quad \overline{X}^{(m)}(t) = W_0\big(\overline{T}^{(m)}(t)\big) \qquad \text{for } t < \tau_\infty^{(m)}. \tag{4.34}$$

The system (4.34) of equations implies that every subsequence $(n_k)_{k \geq 1}$ has another subsequence $(n'_k)_{k \geq 1}$ such that $X_{n'_k}^{(m)} \Rightarrow \overline{X}$ in $C([0, T], \mathbb{R}^d)$, where, at least until $\tau_\infty^{(m)}$, the process $\overline{X}$ behaves as a Walsh diffusion starting from $\overline{X}(0) = x_0$, associated with $(0, \sigma, \mu_0)$. For $m \geq 1$ define

$$\overline{\tau}^{(m)} := \inf\big\{t \geq 0 \mid \overline{X}^{(m)}(t) \notin \text{int } D_m\big\} = \inf\big\{t \geq 0 \mid W_0\big(\overline{T}^{(m)}(t)\big) \notin \text{int } D_m\big\}.$$

Then we claim the following inequality $\overline{\tau}^{(m)} \leq \tau_\infty^{(m)}$ a.s. for every $m \geq 1$. Indeed, assume the converse, that is, $\overline{\tau}^{(m)} > \tau_\infty^{(m)}$ with positive probability. Then there would be a positive random variable $t_1 < \overline{\tau}^{(m)}$ such that there exists a sequence $(\tilde{n}_k)_{k \geq 1}$ with $\tau_{\tilde{n}_k}^{(m)} \leq t_1$ for all k. Since $X_{\tilde{n}_k}^{(m)}$ is stopped at $\tau_{\tilde{n}_k}^{(m)}$, we have $X_{\tilde{n}_k}^{(m)}(\tau_{\tilde{n}_k}^{(m)}) = X_{\tilde{n}_k}^{(m)}(t_1) \in \partial D_m$. Since ∂D_m is closed, letting $k \to \infty$, we would have $\overline{X}(t_1) \in \partial D_m$ with positive probability. This, however, contradicts the property $t_1 < \overline{\tau}^{(m)}$, which completes the proof of $\overline{\tau}^{(m)} \leq \tau_\infty^{(m)}$ a.s.

This proves that $\overline{X}$ behaves as a Walsh diffusion starting from $\overline{X}(0) = x_0$, associated with $(0, \sigma, \mu_0)$, at least until it exits int D_m. This completes the proof of Lemma 4.1.

Step 3: Proof for the driftless case. This step is similar to Sarantsev [30], Lemma 3.8. For a given $\eta \in (0, 1)$ we may take an m large enough, so that the set

$$\mathcal{A} := \left\{ x \in C([0, T], \mathbb{R}^d) \mid x(t) \in \text{int } D_m \ \forall t \in [0, T] \right\}$$

has probability greater than $1 - \eta$, i.e., $\mathbb{P}(X_0 \in \mathcal{A}) > 1 - \eta$, since

$$\mathbb{P}\big(X_0(t) \in \mathcal{I} \ \forall t \in [0, T]\big) = 1,$$

and int $D_m \uparrow \mathcal{I}$ as $m \to \infty$.

Now, by Lemma 4.1 every sequence $(n_k)_{k \geq 1}$ has a subsequence $(n'_k)_{k \geq 1}$, such that $X_{n'_k}^{(m)} \Rightarrow \overline{X}$ in $C([0, T], \mathbb{R}^d)$. Then for every Borel subset $\mathcal{B} \subseteq C([0, T], \mathbb{R}^d)$, we have

$$\mathbb{P}(\overline{X} \in \mathcal{A} \cap \mathcal{B}) = \mathbb{P}(X_0 \in \mathcal{A} \cap \mathcal{B}). \tag{4.35}$$

In particular, letting $\mathcal{B} := C([0, T], \mathbb{R}^d)$, we have:

$$\mathbb{P}(\overline{X} \in \mathcal{A}) = \mathbb{P}(X_0 \in \mathcal{A}) > 1 - \eta. \tag{4.36}$$

Moreover, for every Borel subset $\mathcal{B} \subseteq C([0, T], \mathbb{R}^d)$ and $n \geq 1$, we have:

$$\mathbb{P}\big(X_n^{(m)} \in \mathcal{A} \cap \mathcal{B}\big) = \mathbb{P}(X_n \in \mathcal{A} \cap \mathcal{B}). \tag{4.37}$$

For an open subset $\mathcal{G} \subseteq C([0, T], \mathbb{R}^d)$, the subset $\mathcal{A} \cap \mathcal{G}$ is also open, and hence,

$$\mathbb{P}(X_0 \in \mathcal{A} \cap \mathcal{G}) = \mathbb{P}(\overline{X} \in \mathcal{A} \cap \mathcal{G}) \leq \varliminf_{k \to \infty} \mathbb{P}\big(X_{n'_k}^{(m)} \in \mathcal{A} \cap \mathcal{G}\big). \tag{4.38}$$

On the other hand, it follows from (4.35) and (4.36) that

$$\mathbb{P}(\overline{X} \in \mathcal{A} \cap \mathcal{G}) = \mathbb{P}(X_0 \in \mathcal{A} \cap \mathcal{G}) \geq \mathbb{P}(X_0 \in \mathcal{G}) - \mathbb{P}(X_0 \notin \mathcal{A}) \geq \mathbb{P}(X_0 \in \mathcal{G}) - \eta. \tag{4.39}$$

Combining (4.37), (4.38) with (4.39), we obtain

$$\varliminf_{k \to \infty} \mathbb{P}(X_{n'_k} \in \mathcal{G}) \geq \varliminf_{k \to \infty} \mathbb{P}\big(X_{n'_k}^{(m)} \in \mathcal{A} \cap \mathcal{G}\big) \geq \mathbb{P}(X_0 \in \mathcal{G}) - \eta. \tag{4.40}$$

Thus for every sequence $(n_k)_{k \geq 1}$ and every $\eta > 0$ there exists a subsequence $(n'_k)_{k \geq 1}$ such that (4.40) holds. Use the diagonal argument: let $n_k^{(0)} := n_k$, and construct $(n_k^{(l)})_{k \geq 1}$ inductively: $(n_k^{(l)})_{k \geq 1}$ plays the role of $(n'_k)_{k \geq 1}$ for $n_k := n_k^{(l-1)}$, $\eta := l^{-1}$. Then for $\overline{n}_k := n_k^{(k)}$, we have $\varliminf_{k \to \infty} \mathbb{P}(X_{\overline{n}_k} \in \mathcal{G}) \geq \mathbb{P}(X_0 \in \mathcal{G})$. Therefore, we claim that for every sequence (n_k), there exists a subsequence $(\overline{n}_k)$ such that $X_{\overline{n}_k} \Rightarrow X_0$ in $C([0, T], \mathbb{R}^d)$. This completes the proof of Theorem 4.1 for the case $g \equiv 0$.

Step 4: General case. The general case (with an arbitrary drift function g) can be reduced to the driftless case by scale transformation (2.17). Recall that the process $\mathcal{P}(X_n(\cdot))$ is a Walsh diffusion starting from $\mathcal{P}(x_n)$, associated with $(0, \tilde{\sigma}, \mu_n)$, where $\tilde{\mathcal{I}}$ and $\tilde{\sigma}$ are given by (2.17) and

(2.18). Lemma 4.2 below, together with continuity of σ, implies continuity of the function $\tilde{\sigma}$, and of the mappings $\mathcal{P}$ and $\mathcal{P}^{-1}$. Since $x_n \to x_0$, we have $\mathcal{P}(x_n) \to \mathcal{P}(x_0)$. Moreover, because of $\mu_n \Rightarrow \mu_0$, from results just proven, we have $\mathcal{P}(X_n(\cdot)) \Rightarrow \mathcal{P}(X_0(\cdot))$ in $C([0, T], \mathbb{R}^d)$. Finally, because $\mathcal{P}^{-1}$ is continuous, we have $X_n \Rightarrow X_0$ in $C([0, T], \mathbb{R}^d)$. This completes the proof of Theorem 4.1 for the general case, given Lemma 4.2 below.

Lemma 4.2. *Under Assumption* 4.1, *the scale function* $s(r, \theta)$ *from* (2.15), *and the inverse scale function* $s^{-1}(r, \theta)$ *from* (2.16), *are continuous in the Euclidean topology.*

Proof. Under Assumption 4.1, continuity of the scale function $s(\cdot)$ follows from continuity and local boundedness of the function $2g\sigma^{-2}$, together with Lebesgue dominated convergence theorem. Let us take a sequence (r_n, θ_n) that converges to (r_0, θ_0), that is, $(r_n, \theta_n) \to (r_0, \theta_0)$. We shall show that $s^{-1}(r_n, \theta_n) \to s^{-1}(r_0, \theta_0)$, that is, for every $\varepsilon > 0$ and for all n, except finitely many, we have

$$s^{-1}(r_n, \theta_n) < s^{-1}(r_0, \theta_0) + \varepsilon. \tag{4.41}$$

Since $s(r, \theta)$ is strictly increasing in r for every fixed θ, (4.41) would be equivalent to

$$r_n < s\big(s^{-1}(r_0, \theta_0) + \varepsilon, \theta_n\big). \tag{4.42}$$

Note that letting $n \to \infty$ in (4.42) and using continuity of s, we see the left-hand side converges to r_0, and the right-hand side converges to $s(s^{-1}(r_0, \theta_0) + \varepsilon, \theta_0)$. Here since

$$r_0 = s\big(s^{-1}(r_0, \theta_0), \theta_0\big) < s\big(s^{-1}(r_0, \theta_0) + \varepsilon, \theta_0\big),$$

we must get (4.42) for large enough n. This completes the proof of Lemma 4.2. $\qquad\square$

5. Feller and positivity properties

In this section, we study several properties: Feller property; positivity of transition kernel (that mes $\otimes \mu$-positive subsets of $\mathbb{R}^d$ have positive transition measure). They are necessary for the next section, where we find Lyapunov functions for Walsh diffusions to prove existence and uniqueness of a stationary distribution and convergence to this stationary distribution as $t \to \infty$.

5.1. Positivity of transition kernel

Let us first mention intuition for general Markov processes. Very loosely speaking, a Markov process is called *irreducible*, if the state space cannot be separated into two or more parts such that the process cannot move between them; and it is called *aperiodic* if the state space cannot be separated into two or more parts such that the process circulates between them. If the process is tight, then irreducibility and aperiodicity guarantee existence and uniqueness of a probability invariant measure. Since we do not use these particular (very important) concepts in this paper,

we shall not rigorously define them here. But we, however, prove a stronger property: positivity. Let us define the following *reference measure*:

$$\tilde{\mu} := \mu \otimes \text{mes} \tag{5.1}$$

and consider a Walsh diffusion on $\mathcal{I}$ associated with (g, σ, μ) and the transition kernel $P^t(x, \cdot) = \mathbb{P}(X(t) \in \cdot \mid X(0) = x)$ for $x \in \mathcal{I}$, $t > 0$.

Theorem 5.1. *Under Assumptions* 2.1 *and* 2.2, *the transition kernel is positive, that is, for every* $t > 0$, $x \in \mathcal{I}$, *and a Borel subset* $C \subseteq \mathcal{I}$ *with* $\tilde{\mu}(C) > 0$, *we have* $P^t(x, C) > 0$.

Proof. We shall show this theorem in four steps.

Step 1. Let us argue first that it suffices to show the case $x = \mathbf{0}$. Indeed, if the initial value $X(0) = x = (r, \theta) \in \mathcal{I}$ is not the origin $\mathbf{0}$, then until the first hitting time $\tau_{\mathbf{0}} := \inf\{t \geq 0 \mid X(t) = \mathbf{0}\}$, $X(\cdot)$ can be represented as $X(\cdot) = \theta Z(\cdot)$, where $Z = (Z(t), t \geq 0)$ is a diffusion on the half-line with drift $g(\cdot, \theta)$ and diffusion $\sigma^2(\cdot, \theta)$, starting from $Z(0) = r > 0$ and killed at the origin. Then $\tau_{\mathbf{0}} = \inf\{t \geq 0 \mid Z(t) = 0\}$ and hence for such diffusion $Z(\cdot)$ we have $\mathbb{P}(\tau_{\mathbf{0}} < t) > 0$, $t > 0$, since the functions g, σ, σ^{-1} are locally bounded on $\mathbb{R}_+$. Thus if the statement of the theorem is true in the case $x = \mathbf{0}$, then $P^u(\mathbf{0}, C) > 0$ for every $u > 0$, and hence,

$$P^t(x, C) \geq \int_0^t P^{t-s}(\mathbf{0}, C) \mathbb{P}(\tau_{\mathbf{0}} \in ds \mid X(0) = x) > 0; \qquad x \in I,$$

because the Lebesgue integral of a positive function over a set of positive measure is positive, and hence the statement is true for every $x \in I$. Thus it suffices to show the case $x = \mathbf{0}$.

Step 2. Next, consider the case of Walsh Brownian motion $X(\cdot) := W(\cdot)$ associated with $g \equiv 0$ and $\sigma \equiv 1$ starting at the origin, with the set C of the form $C = A \times B \subseteq \mathbb{S} \times (0, \infty)$ with $\tilde{\mu}(C) > 0$. Then by the construction (e.g., Ichiba *et al.* [16], Theorem 2.1), $\arg(W(t))$ is distributed as μ, independent of $\|W(t)\|$, a reflected Brownian motion on the half-line, starting from zero. Thus,

$$P^t(\mathbf{0}, C) = \mathbb{P}(\arg(W(t)) \in A, \|W(t)\| \in B) = \mu(A) \cdot \mathbb{P}(\|W(t)\| \in B) > 0; \qquad t > 0.$$

Step 3. Now consider the case of *driftless* Walsh diffusions with $g \equiv 0$. From Karatzas and Yan [18], Proposition 3.4, we have the Dambis-Dubins-Schwarz-type representation: $X(t) = W(T(t))$, where $W(\cdot)$ is a Walsh Brownian motion in Step 2 starting from $W(0) = \mathbf{0}$, and

$$T(t) = \int_0^t \sigma^2(X(s)) ds, \qquad t \geq 0. \tag{5.2}$$

For the set $C = A \times B$ define $R := \sup B < l_{\min} := \inf_{\theta \in A} \ell(\theta)$, and for a fixed $R' \in (R, l_{\min})$

$$\tau_{R,A}^X := \inf\{t \geq 0 \mid \|X(t)\| = R, \arg X(t) \in A\}, \tag{5.3}$$

$$\tau_{R,A}^W := \inf\{t \geq 0 \mid \|W(t)\| = R, \arg W(t) \in A\}, \tag{5.4}$$

$$\tau_{R'}^X := \inf\{t \geq 0 \mid \|X(t)\| = R'\}, \qquad \tau_{R'}^W := \inf\{s \geq 0 \mid \|W(s)\| = R'\}. \tag{5.5}$$

With Lemma A.2 in the Appendix we claim $\mathbb{P}(\tau_{R,A}^X < t) > 0$. Note that

$$\mathbb{P}\big(X(t) \in C = A \times B\big) \geq \int_0^t \int_A p(t-s,\theta) \mathbb{P}\big(\tau_{R,A}^X \in \mathrm{d}s, \arg(X(s)) \in \mathrm{d}\theta\big), \qquad (5.6)$$

where $p(u,\theta)$ is the probability that a reflected diffusion on the half-line with zero drift and diffusion $\sigma^2(\cdot,\theta)$, starting from R, stays in $(0, R']$ on the time interval $[0, u]$, and hits the set B at time u. From boundedness of σ and σ^{-1} on $[0, R']$, which follows from Assumption 2.1, we have: $p(u,\theta) > 0$. Again using the observation that the Lebesgue integral of a positive function over a set of positive measure is positive, we see that the right hand of (5.6) is positive, which completes the proof of Theorem 5.1 for the case of driftless Walsh diffusions.

Step 4. Finally, let us prove Theorem 5.1 in the general case. It can be reduced via Karatzas and Yan [18], Proposition 3.12, to the driftless case. Using the notation from there, we observe that the one-to-one function $\mathcal{P} : \mathcal{I} \to \tilde{\mathcal{I}}$ from (2.17) maps the Walsh diffusion with nonzero drift to another Walsh diffusion with zero drift. The new Walsh diffusion also has a diffusion coefficient $\tilde{\sigma}$ from (2.18) such that $\tilde{\sigma}$ and $\tilde{\sigma}^{-1}$ are both locally bounded on $\tilde{\mathcal{I}}$. Also, the map $\mathcal{P}$, as well as its inverse $\mathcal{Q}$, maps $\tilde{\mu}$-positive subsets into $\tilde{\mu}$-positive subsets. This follows from the observation that these maps preserve arguments of points: $\arg(\mathcal{P}(x)) = \arg(x)$ for $x \in \mathcal{I} \setminus \{\mathbf{0}\}$, and the radial derivative of $\mathcal{P}$ is everywhere positive. Thus, Theorem 5.1 is a simple corollary for the driftless case. $\qquad\square$

5.2. Feller property for the tree topology

Next, we shall prove the Feller property of Walsh diffusions, that is, the semigroup maps bounded continuous functions into bounded continuous functions. Fix a bounded continuous (in the tree topology) function $f : \mathcal{I} \to \mathbb{R}$.

Theorem 5.2. *Under Assumptions 2.1 and 2.2, for $t > 0$, if $x \to x_0$ in $\mathcal{I}$, then*

$$\mathbb{E}_x\big[f\big(X(t)\big)\big] \to \mathbb{E}_{x_0}\big[f\big(X(t)\big)\big].$$

Proof. We shall first consider the harder case $x_0 \neq \mathbf{0}$, and then discuss the easier case.

Case 1. Assume $x_0 = \theta r_0 \neq \mathbf{0}$ for some θ, that is, $r_0 > 0$; and $x = \theta r$, with $r \uparrow r_0$. Take a copy $X^{(r_0)} = (X^{(r_0)}(t), t \geq 0)$ of this Walsh diffusion starting from $X^{(r_0)}(0) = x_0$. One can construct (on the same probability space as $X^{(r_0)}$) a family $(Z^{(r)})_{r \in (0, r_0]}$ of reflected diffusions on the half-line with drift $g(\cdot,\theta)$ and diffusion $\sigma^2(\cdot,\theta)$, starting from $Z^{(r)}(0) = r$, such that $Z^{(r)}(t) \leq Z^{(r')}(t)$ a.s. for every $t \geq 0$ and $0 < r < r' \leq r_0$. Also, assume that this probability space contains a Walsh diffusion $X^{(0)} = (X^{(0)}(t), t \geq 0)$ with the same drift and diffusion coefficients, starting from the origin, independent of everything else. Let $\tau_{r,a} := \inf\{t \geq 0 \mid Z^{(r)}(t) = a\}$ for $r \in (0, r_0]$ and $a \in \mathbb{R}_+$. Then since $\tau_{r,r_0} \downarrow \tau_{r_0,r_0} = 0$, $\tau_{r,0} \uparrow \tau_{r_0,0} > 0$ a.s., as $r \uparrow r_0$,

$$\lim_{r \uparrow r_0} 1(\tau_{r,r_0} < \tau_{r,0}) = 1, \qquad \lim_{r \uparrow r_0} 1(\tau_{r,r_0} > \tau_{r,0}) = 0,$$

$$\lim_{r \uparrow r_0} 1(t \leq \tau_{r,r_0} < \tau_{r,0}) = 0, \qquad \lim_{r \uparrow r_0} 1(\tau_{r,r_0} < \tau_{r,0} \wedge t) = 1 \qquad \text{a.s.} \qquad (5.7)$$

For every $r \in (0, r_0)$, let us construct a copy $X^{(r)} = (X^{(r)}(t), t \geq 0)$ of this Walsh diffusion starting from $X^{(r)}(0) = \theta r$. The construction of $X^{(r)}$ proceeds as follows.

(a) If $\tau_{r,0} < \tau_{r,r_0}$, that is, the reflected diffusion $Z^{(r)}$ hits zero before r_0, then we let $X^{(r)}$ evolve like this reflected diffusion on the ray $\mathcal{R}_\theta$ before hitting the origin, and then start the independent copy $X^{(0)}$ of the Walsh diffusion from there. Formally, let us define

$$X^{(r)}(t) := \begin{cases} \theta Z^{(r)}(t), & t \leq \tau_{r,0}; \\ X^{(0)}(t - \tau_{r,0}), & t \geq \tau_{r,0}. \end{cases}$$

(b) If $\tau_{r,0} > \tau_{r,r_0}$, that is, $Z^{(r)}$ hits r_0 before zero, then we let $X^{(r)}$ evolve like this reflected diffusion on the ray $\mathcal{R}_\theta$, until it hits x_0. Then we start the copy of $X^{(r_0)}$. Let us define

$$X^{(r)}(t) := \begin{cases} \theta Z^{(r)}(t), & t \leq \tau_{r,r_0}; \\ X^{(r_0)}(t - \tau_{r,r_0}), & t \geq \tau_{r,r_0}. \end{cases}$$

Since $r \uparrow r_0$ case (a) is less likely and case (b) is more likely. Thus, we construct a Walsh diffusion $X^{(r)}$ with initial value $X^{(r)}(0) = (r, \theta)$. We shall evaluate $\mathbb{E}[f(X^{(r)}(t))] =: E(r) + F(r)$, where

$$F(r) := \mathbb{E}[f(X^{(r)}(t))1(\tau_{r,r_0} < \tau_{r,0})], \qquad E(r) := \mathbb{E}[f(X^{(r)}(t))1(\tau_{r,r_0} > \tau_{r,0})]. \quad (5.8)$$

Thanks to (5.7) and boundedness of f, we immediately see

$$E(r) \to 0 \qquad \text{as } r \uparrow r_0. \quad (5.9)$$

Let us decompose the term $F(r)$ into two terms:

$$\begin{aligned} F(r) &= \mathbb{E}[f(X^{(r)}(t))1(t \leq \tau_{r,r_0} < \tau_{r,0})] + \mathbb{E}[f(X^{(r)}(t))1(\tau_{r,r_0} < \tau_{r,0} \wedge t)] \\ &= \mathbb{E}[f(\theta Z^{(r)}(t))1(t \leq \tau_{r,r_0} < \tau_{r,0})] + \mathbb{E}[f(X^{(r_0)}(t - \tau_{r,r_0}))1(\tau_{r,r_0} < \tau_{r,0} \wedge t)] \quad (5.10) \\ &=: F_1(r) + F_2(r). \end{aligned}$$

Next, $F_1(r) \to 0$ as $r \uparrow r_0$, thanks to (5.7) and boundedness of f, and

$$F_2(r) - \mathbb{E}[f(X^{(r_0)}(t - \tau_{r,r_0}))] \to 0 \qquad \text{as } r \uparrow r_0. \quad (5.11)$$

Combine (5.7) with a.s. continuity of trajectories of $X^{(r_0)}$, continuity and boundedness of f, and use the Lebesgue dominated convergence theorem. Then, as $r \uparrow r_0$, we get

$$\mathbb{E}[f(X^{(r_0)}(t - \tau_{r,r_0}))] \to \mathbb{E}[f(X^{(r_0)}(t))]. \quad (5.12)$$

Combining (5.8)–(5.12), we conclude $\lim_{r \uparrow r_0} \mathbb{E}[f(X^{(r)}(t))] = \mathbb{E}[f(X^{(r_0)}(t))]$.

Case 2. Assume $x_0 = \theta r_0 \neq \mathbf{0}$, that is, $r_0 > 0$; and $x = \theta r$, with $r \downarrow r_0$; or $x_0 = \mathbf{0}$, that is, $r_0 = 0$; then also we have $x = \theta r$ with $r \downarrow r_0 = 0$. Then the proof is simpler: there is no case (a). Indeed, a reflected diffusion $Z^{(r)}$ on the half-line, starting from $r > r_0$, must hit r_0 before hitting zero (or at least at the same time when $r_0 = 0$). The details of the proof are left to the reader. □

5.3. Feller property for Euclidean topology

The continuity of the transition kernel in Euclidean topology can be seen as a corollary of Theorem 4.1, if we take $\mu_n \equiv \mu$ for all $n = 0, 1, 2, \ldots$.

Lemma 5.1. *Under Assumptions* 2.1, 2.2, 4.1, *take a bounded continuous* (*in the Euclidean topology*) *function* $f : \mathcal{I} \to \mathbb{R}$. *Fix a* $t > 0$, *and let* $x \to x_0$ *in* $\mathcal{I}$ *in Euclidean topology. Then*

$$\mathbb{E}_x\big[f\big(X(t)\big)\big] \to \mathbb{E}_{x_0}\big[f\big(X(t)\big)\big].$$

6. Lyapunov functions and convergence to the stationary distribution

In this section, we shall find Lyapunov functions for Walsh diffusions to prove existence and uniqueness of a stationary distribution, and convergence to this stationary distribution as $t \to \infty$. For general Markov processes, the application of Lyapunov functions has been widely studied in the last few decades. Without attempting to provide an exhaustive list of references, let us mention the following papers: Bakry, Cattiaux and Guillin [1], Douc, Fort and Guillin [7], Down, Meyn and Tweedie [8], Lund, Meyn and Tweedie [19], Meyn and Tweedie [20,21], Roberts and Tweedie [28]. We have three goals:

(a) Establish the very fact of long-term convergence of the transition kernel $P^t(x, \cdot)$ to the stationary distribution $\pi(\cdot)$ using Lyapunov functions, in a suitable distance;

(b) Prove that the rate of this convergence is exponential; that is, the distance between $P^t(x, \cdot)$ and $\pi(\cdot)$ is estimated from above as a constant (dependent on x) times $e^{-\varkappa t}$, for some $\varkappa > 0$;

(c) Estimate the rate $\varkappa$ of this exponential convergence.

6.1. Definitions and general results

Let us start with general definitions. Consider an $\mathbb{R}^d$-valued continuous-time Markov process $X = (X(t), t \geq 0)$ with transition kernel $P^t(x, \cdot)$.

Definition 6.1. We say that the process X is *ergodic* if there exists a unique stationary distribution $\pi(\cdot)$, and if the transition kernel converges in the total variation norm $\|\cdot\|_{\mathrm{TV}}$ to π, i.e.,

$$\lim_{t \to \infty} \big\| P^t(x, \cdot) - \pi(\cdot) \big\|_{\mathrm{TV}} = 0, \qquad \text{for every } x \in \mathbb{R}^d.$$

Definition 6.2. We say that X is *V-uniformly ergodic* for a function $V : \mathbb{R}^d \to [1, \infty)$, if X is ergodic, and there exist constants $K, \varkappa > 0$ such that for every $t \geq 0, x \in \mathbb{R}^d$, we have:

$$\big\| P^t(x, \cdot) - \pi(\cdot) \big\|_{\mathrm{TV}} \leq K V(x) e^{-\varkappa t}. \tag{6.1}$$

For $V \equiv 1$, we say that X is *exponentially ergodic*.

Assumption 6.1. The spinning measure μ has finite support $\operatorname{supp}\mu = \{\theta_1, \ldots, \theta_p\}$ in $\mathbb{S}$; that is, the effective state space $\mathcal{I}^\mu$ in (2.9) for this Walsh diffusion is a finite union of rays from the origin

$$\mathcal{I}^\mu := \bigcup_{i=1}^p \{r\theta_i \mid 0 \le r < l_i\}, \qquad l_i := \ell(\theta_i), i = 1, \ldots, p.$$

Sometimes, the Walsh diffusion in the finite union $\mathcal{I}^\mu$ of rays is called a *spider*.

Lemma 6.1. *Under Assumptions* 2.1, 2.2, *and either* 4.1 *or* 6.1, *we have the following results.*

(a) *Assume there exist a function* $V : \mathcal{I}^\mu \to \mathbb{R}_+$ *in the domain of the generator* $\mathcal{L}$ *and some positive constants* k, b, r_0, *such that*

$$\mathcal{L}V(x) \le -k + b1_{\mathcal{B}^\mu(r_0)}(x), \qquad x \in \mathcal{I}^\mu, \tag{6.2}$$

$$where \ \mathcal{B}^\mu(r_0) := \big\{(r, \theta) \in \mathcal{I}^\mu \mid 0 < r \le r_0\big\} \cup \{\mathbf{0}\}. \tag{6.3}$$

Then the Walsh diffusion is ergodic. Moreover, for every bounded measurable function $f : \mathcal{I}^\mu \to \mathbb{R}$,

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t f\big(X(u)\big) \, \mathrm{d}u = \int_{\mathcal{I}^\mu} f(x)\pi(\mathrm{d}x). \tag{6.4}$$

(b) *Assume there exist a function* $V : \mathcal{I}^\mu \to [1, \infty)$ *in the domain of the generator* $\mathcal{L}$ *and some positive constants* k, b, r_0, *such that*

$$\mathcal{L}V(x) \le -kV(x) + b1_{\mathcal{B}^\mu(r_0)}(x), \qquad x \in \mathcal{I}^\mu. \tag{6.5}$$

Then the Walsh diffusion is V-*uniformly ergodic, and* $\int_{\mathcal{I}^\mu} V(x)\pi(\mathrm{d}x) < \infty$.

Remark 6.1. A function V which satisfies (6.2) or (6.5) is called a *Lyapunov function* in the literature cited in the beginning of this section.

Proof. *Case 1*. First, we work under Assumptions 2.1, 2.2, 4.1. Then we operate in Euclidean topology. The set $\mathcal{B}^\mu(r_0)$ from (6.3) is compact. The Walsh diffusion is Feller continuous from Lemma 5.1, and has the positivity property from Theorem 5.1.

Case 2. Next, we work under Assumptions 2.1, 2.2, 6.1. Then we operate in the tree-topology. By Remark 2.1, the set $\mathcal{B}^\mu(r_0)$ from (6.3) is compact in the tree-topology. The Walsh diffusion is Feller continuous from Theorem 5.2, and has the positivity property from Theorem 5.1.

Finally, for both cases, the rest of the proof of (b) follows from Sarantsev [31], Lemma 2.3, Theorem 2.6, and the rest of the proof of (a) follows from Sarantsev [31], Proposition 2.2, and Meyn and Tweedie [21], Theorem 5.1. $\qquad\square$

6.2. An example of convergence

Let us provide an example with explicit conditions on the drift and diffusion coefficients g and σ.

Lemma 6.2. *Under Assumptions* 2.1, 2.2, *and either* 4.1 *or* 6.1,
  (a) *the Walsh diffusion is ergodic, if*

$$\varlimsup_{r \to \infty} \sup_{\theta \in \mathbb{S}} g(r, \theta) =: -\overline{g} < 0. \tag{6.6}$$

  (b) *If, in addition, the following condition*

$$\varlimsup_{r \to \infty} \sup_{\theta \in \mathbb{S}} \sigma(r, \theta) =: \overline{\sigma} < \infty, \tag{6.7}$$

holds, then the Walsh diffusion is V-uniformly ergodic with $V = V_\lambda(r, \theta) := e^{\lambda r}$ for some $\lambda > 0$.

Proof. The proof is similar to that from Sarantsev [29], Theorem 3.2. Take a C^∞ nondecreasing function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\varphi(x) = \begin{cases} 0, & x \leq 1; \\ x, & x \geq 2. \end{cases}$$

An example of such function can be found in Sarantsev [29], Section 3.2. For (a), try $V(r, \theta) = \varphi(r)$. This function satisfies condition (2.12), because $V_r'(0+, \theta) = \varphi'(0) = 0$. It is also continuous in the Euclidean topology, which is the additional condition in Remark 2.3. Plug into (2.11) and get

$$\text{for } r \geq 2, \qquad \varphi'(r) = 1, \quad \text{and} \quad \varphi''(r) = 0, \tag{6.8}$$

and hence, $\mathcal{L}V(r, \theta) = g(r, \theta)$. It follows from (6.6) that there exist $r_1, b > 0$ such that $g(r, \theta) \leq -b$ for $r \geq r_1$. Therefore, for $r \geq r_0 := r_1 \vee 2$, we get $\mathcal{L}V(r, \theta) \leq -b$. In addition, the function $\mathcal{L}V$ is continuous and therefore bounded on $\mathcal{B}^\mu(r_0)$. Thus, we have (6.2). Apply Lemma 6.1(a) to complete the proof of Lemma 6.2(a). The proof of (6.4) follows from Meyn and Tweedie [20], Theorem 8.1(a).

For (b), try $V(r, \theta) = \exp(\lambda \varphi(r))$. Similarly, this function satisfies $V_r'(0+, \theta) = 0$ and therefore (2.12); and on top of this, V is also continuous in the Euclidean topology. Using (6.8), for $r \geq 2$, we have: $V'(r, \theta) = \lambda V(r, \theta)$, and $V''(r, \theta) = \lambda^2 V(r, \theta)$. Thus,

$$\mathcal{L}V(r, \theta) = \left[g(r, \theta)\lambda + \frac{1}{2}\sigma^2(r, \theta)\lambda^2 \right] V(r, \theta) \qquad \text{for } r \geq 2. \tag{6.9}$$

Now, from (6.6) and (6.7), there exist constants $\overline{g}, \overline{\sigma} > 0$ and an $r_1 > 0$ such that $g(r, \theta) \leq -\overline{g} < 0$, $\sigma(r, \theta) \leq \overline{\sigma}$, $r \geq r_1$. Then, for $r \geq r_1$, $\lambda = \overline{g} \cdot \overline{\sigma}^{-2}$, we have

$$g(r, \theta)\lambda + \frac{1}{2}\sigma^2(r, \theta)\lambda^2 \leq -\overline{g}\lambda + \frac{\overline{\sigma}^2}{2}\lambda^2 \leq -\frac{\overline{g}^2}{2\overline{\sigma}^2} =: -k < 0. \tag{6.10}$$

Comparing (6.9) with (6.10), we get $\mathcal{L}V(r, \theta) \leq -kV(r, \theta)$ for $\theta \in \operatorname{supp}\mu$ and $r \geq r_1 \vee 2 =: r_0$. Similarly to (a), we get that $\mathcal{L}V$ is continuous and therefore bounded on $\mathcal{B}^\mu(r_0)$. Therefore, we obtain (6.5). Apply Lemma 6.1(b) to complete the proof of Lemma 6.2(b). $\qquad\square$

6.3. Explicit rate of exponential convergence

In a fairly general setting, we can estimate the rate $\varkappa$ of exponential convergence from (6.1). It is hard to estimate this rate for diffusions. Let us informally explain the difficulty: Let $\mathcal{L}$ be the generator of a certain Markov process on $\mathbb{R}^d$. Assume we have a Lyapunov function V in the domain of this generator which satisfies

$$\mathcal{L}V(x) \leq -kV(x) + b1_C(x) \qquad \text{for all } x. \tag{6.11}$$

Here, $k, b > 0$ are some constants, and C is a "small" set. There is actually a precise meaning of the term *small set* in this theory, which was developed in Down, Meyn and Tweedie [8], Meyn and Tweedie [20,21]. For our purposes, it is sufficient to let C be a compact set, as follows from Sarantsev [31], Lemma 2.3, Proposition 2.6. One would like to infer an explicit value of the constant $\varkappa$ in (6.1) from the constants in (6.11). As mentioned in the Introduction, however, it turns out to be very hard, in general, see, for example, Davies [6], Meyn and Tweedie [22], Roberts and Rosenthal [26], Roberts and Tweedie [28], since $\varkappa$ depends in a complicated way on k, b, C, and the transition kernel $P^t(x, \cdot)$.

However, such estimates are much easier if the Markov process is on the half-line $\mathbb{R}_+$, is stochastically ordered, and the "exceptional set" $C = \{0\}$ in the formula (6.11) for a Lyapunov function. Alternatively, the stochastic process itself might not be stochastically ordered, but is stochastically dominated by a stochastically ordered Markov process with a Lyapunov function with $C = \{0\}$. This was done by the coupling method in Lund, Meyn and Tweedie [19] for some processes, including reflected diffusions, and in Sarantsev [29] for reflected jump-diffusions.

Here, we are able to adjust the coupling techniques used in Lund, Meyn and Tweedie [19], Sarantsev [29] for the case of Walsh diffusions, by dominating the radial component of the Walsh diffusion by a stochastically ordered reflected diffusion on $\mathbb{R}_+$. The rest of this section closely follows the ideas of Lund, Meyn and Tweedie [19], Sarantsev [29].

Recall $f : \mathbb{R}_+ \to \mathbb{R}$ is called *locally Lipschitz* if for every $R_0 > 0$ there exists a $C(R_0) > 0$ such that $|f(r_1) - f(r_2)| \leq C(R_0)|r_1 - r_2|$, $r_1, r_2 \in [0, C(R_0)]$. Impose the following assumption.

Assumption 6.2. The drift coefficient $g(r, \theta)$ is dominated by $g(r, \theta) \leq \overline{g}(r)$, $(r, \theta) \in \mathcal{I} \setminus \{0\}$, where $\overline{g}(r)$ is independent of θ, and the diffusion coefficient σ is itself angular-independent:

$$\sigma(r, \theta) = \overline{\sigma}(r), \qquad (r, \theta) \in \mathcal{I} \setminus \{0\}.$$

Here, $\overline{g}, \overline{\sigma} : \mathbb{R}_+ \to \mathbb{R}$ are assumed to be locally Lipschitz continuous, and so is $g(\cdot, \theta)$ for each θ.

Theorem 6.1. (a) *Under Assumption 6.2, suppose there exists a constant $k > 0$ and a nondecreasing C^2 function $V : \mathbb{R}_+ \to [1, \infty)$ such that*

$$\overline{g}(r)V'(r) + \frac{1}{2}\overline{\sigma}^2(r)V''(r) \leq -kV(r), \qquad r > 0. \tag{6.12}$$

Then for every two points $x_1, x_2 \in \mathcal{I}$, and every $t > 0$, we have

$$\left\| P^t(x_1, \cdot) - P^t(x_2, \cdot) \right\|_V \leq \left(V(x_1) + V(x_2)\right)e^{-kt}. \tag{6.13}$$

(b) *If, in addition to* (a), *this Walsh diffusion is ergodic, and π is its stationary distribution, satisfying $(\pi, V) < \infty$, then this Walsh diffusion is V-uniformly ergodic with $\varkappa := k$ in* (6.1).

Remark 6.2. Somewhat abusing the notation, we will refer to V sometimes as a function $V : \mathbb{R}^d \to \mathbb{R}$, and sometimes as a function $V : \mathbb{R}_+ \to \mathbb{R}$.

Proof. Part (a) *Step 1.* Similarly to Lund, Meyn and Tweedie [19], Sarantsev [29], we couple four processes: two copies X_1 and X_2 of the Walsh diffusion associated with (g, σ, μ), starting from $X_i(0) = x_i, i = 1, 2$, and two copies S_1 and S_2 of a reflected diffusion on $\mathbb{R}_+$ with coefficients $\overline{g}(\cdot)$ and $\overline{\sigma}(\cdot)$, starting from $S_i(0) = \|x_i\|$, $i = 1, 2$, so that the following pathwise comparison holds:

$$\|X_i(t)\| \le S_i(t), \qquad i = 1, 2. \tag{6.14}$$

Assume also S_1 and S_2 have the same driving Brownian motion. That is, there exists a standard Brownian motion $B = (B(t), t \ge 0)$, such that

$$dS_i(t) = \overline{g}(S_i(t)) \, dt + \overline{\sigma}(S_i(t)) \, dB(t) + dL_i(t), \qquad i = 1, 2. \tag{6.15}$$

Here, $L_i = (L_i(t), t \ge 0), i = 1, 2$, are some continuous adapted nondecreasing real-valued processes, with $L_i(0) = 0$, such that L_i can increase only when $S_i = 0$, $i = 1, 2$. We shall show at the end of this proof how to construct such coupling.

Step 2. Assume we have already constructed the coupling with all aforementioned properties. Then we can quickly prove the statement of Theorem 6.1(a). Assume without loss of generality that $\|x_1\| \le \|x_2\|$. Then using the standard comparison techniques for diffusions, we get

$$S_1(t) \le S_2(t), \qquad t \ge 0. \tag{6.16}$$

Define the stopping time $\tau := \inf\{t \ge 0 \mid S_2(t) = 0\}$. By (6.14) and (6.16), we have

$$\|X_1(\tau)\| = \|X_2(\tau)\| = 0 \quad \Rightarrow \quad X_1(\tau) = X_2(\tau) = \mathbf{0}.$$

Therefore, τ is a coupling time for X_1 and X_2. Using the standard trick, we assume $X_1(t) = X_2(t)$ a.s. for $t > \tau$. Then for a function $f : \mathcal{I} \to \mathbb{R}$ such that $|f| \le V$, we have

$$
\begin{aligned}
&\left| \mathbb{E}[f(X_1(t))] - \mathbb{E}[f(X_2(t))] \right| \\
&= \left| \mathbb{E}[f(X_1(t)) 1_{\{\tau > t\}}] - \mathbb{E}[f(X_2(t)) 1_{\{\tau > t\}}] \right| \\
&\le \mathbb{E}[|f(X_1(t))| 1_{\{\tau > t\}}] + \mathbb{E}[|f(X_2(t))| 1_{\{\tau > t\}}] \\
&\le \mathbb{E}[V(X_1(t)) 1_{\{\tau > t\}}] + \mathbb{E}[V(X_2(t)) 1_{\{\tau > t\}}] \\
&\le \mathbb{E}[V(S_1(t)) 1_{\{\tau > t\}}] + \mathbb{E}[V(S_2(t)) 1_{\{\tau > t\}}].
\end{aligned}
\tag{6.17}
$$

Combining (6.17) with (6.12), we get as in Sarantsev [29]:

$$\left| \mathbb{E}[f(X_1(t))] - \mathbb{E}[f(X_2(t))] \right| \le (V(x_1) + V(x_2)) e^{-kt}.$$

Taking supremum over all functions $f : \mathcal{I} \to \mathbb{R}$ such that $|f| \le V$, we complete the proof of (a).

Step 3. It remains to show that a coupling of

$$X_1, X_2, S_1, S_2,$$

which satisfies (6.14), (6.15) exists. First, our goal is to construct two Walsh diffusions X_1 and X_2 associated with (g, σ, μ), starting from $X_i(0) = x_i$, with the *same driving Brownian motion B*. That is, X_1 and X_2 need to satisfy

$$\mathrm{d}\|X_i(t)\| = g(X_i(t))\,\mathrm{d}t + \sigma(X_i(t))\,\mathrm{d}B(t) + \mathrm{d}L^{\|X_i\|}(t), \qquad i = 1, 2; t \geq 0. \qquad (6.18)$$

To this end, take a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with infinitely many i.i.d. copies

$$W^{(n)} = (W^{(n)}(t), t \geq 0), \qquad n = 0, 1, 2, \ldots$$

of a Walsh diffusion associated with (g, σ, μ), starting from the origin: $W^{(n)}(0) = \mathbf{0}$; as well as yet another independent standard Brownian motion $\overline{B} = (\overline{B}(t), t \geq 0)$. For each Walsh diffusion $W^{(n)}$, we can write a representation in terms of stochastic differential equation

$$\mathrm{d}\|W^{(n)}(t)\| = g(W^{(n)}(t))\,\mathrm{d}t + \sigma(W^{(n)}(t))\,\mathrm{d}B_n(t) + \mathrm{d}L^{\|W_n\|}(t),$$

where $B_n = (B_n(t), t \geq 0)$, $n = 1, 2, \ldots$ are i.i.d. standard Brownian motions. Let $x_i = \overline{r}_i\overline{\theta}_i$, $i = 1, 2$. If $\overline{r}_1 = 0$ or $\overline{r}_2 = 0$, then let $\tau_0 := 0$. Assume now $\overline{r}_1 > 0$ and $\overline{r}_2 > 0$. For $i = 1, 2$, consider strong solutions $\overline{S}_i$ of a one-dimensional SDE with drift coefficient $g(\cdot, \overline{\theta}_i)$ and diffusion coefficient $\sigma^2(\cdot, \overline{\theta}_i)$, starting from $\overline{S}_i(0) = \overline{r}_i$, driven by Brownian motion $\overline{B}$:

$$\mathrm{d}\overline{S}_i(t) = g(\overline{S}_i, \overline{\theta}_i)\,\mathrm{d}t + \overline{\sigma}(\overline{S}_i)\,\mathrm{d}\overline{B}(t), \qquad t \leq \overline{\tau}_i := \inf\{t \geq 0 \mid \overline{S}_i(t) = 0\}.$$

Because the drift coefficient $g(\cdot, \overline{\theta}_i)$ and the diffusion coefficient $\sigma(\cdot, \overline{\theta}_i)$ are locally Lipschitz continuous, this strong solution exists and is unique. By (6.12) and Assumption 6.2 it follows that this process is non-explosive, at least not until it hits zero. Define

$$X_i(t) = \overline{\theta}_i\overline{S}_i(t), \qquad i = 1, 2; t < \tau_0;$$

where the first stopping time is defined as $\tau_0 := \inf\{t \geq 0 \mid X_1(t) = \mathbf{0} \text{ or } X_2(t) = \mathbf{0}\} = \overline{\tau}_1 \wedge \overline{\tau}_2$. Thus, we defined $X_1(t)$ and $X_2(t)$ for $t \leq \tau_0$ so that (6.18) is satisfied with $B(t) := \overline{B}(t)$ for $t \leq \tau_0$.

Step 4. Next, we construct a sequence $(\tau_n)_{n \geq 0}$ of stopping times such that $\tau_0 \leq \tau_1 \leq \tau_2 \leq \cdots$, and define X_1 and X_2 inductively on each $[\tau_{2k}, \tau_{2k+2}]$ so that (6.18) holds with

$$X_1(\tau_{2k}) = \mathbf{0}, \qquad X_2(\tau_{2k+1}) = \mathbf{0}, \qquad k \geq 0; \qquad \tau_\infty := \lim_{k \to \infty} \tau_k. \qquad (6.19)$$

Therefore, we shall construct X_1 and X_2 on $[0, \tau_\infty]$ so that (6.18) and (6.19) hold. In Step 6, we construct $X_1(t)$ and $X_2(t)$ for $t > \tau_\infty$, if $\tau_\infty < \infty$.

Assume we have already defined X_1 and X_2 on $[0, \tau_{2n}]$, so that (6.18) and (6.19) hold for $k = 0, \ldots, n$. In the case $X_1(\tau_{2n}) = X_2(\tau_{2n}) = \mathbf{0}$, we let $\tau_k := \tau_{2n}$ for $k \geq 2n$. Next, assume

$$X_1(\tau_{2n}) = \mathbf{0}, \qquad X_2(\tau_{2n}) \neq \mathbf{0}.$$

Let $X_2(\tau_{2n}) = \rho_{2n}\theta_{2n}$. Construct a strong solution $S_{2n+1} = (S_{2n+1}(t), t \geq 0)$ of a one-dimensional SDE with coefficients $g(\cdot, \theta_{2n})$ and $\sigma(\cdot)$, starting from $S_{2n+1}(0) = \rho_{2n}$, until it hits zero:

$$dS_{2n+1}(t) = g(S_{2n+1}(t), \theta_{2n}) dt + \overline{\sigma}(S_{2n+1}(t)) dB_{2n+1}(t), \qquad t \leq \tau'_{2n+1}$$
$$:= \inf\{t \geq 0 \mid S_{2n+1}(t) = 0\}.$$

By local Lipschitz continuity of the coefficients $g(\theta_{2n}, \cdot)$ and $\sigma(\cdot)$, this strong solution exists and is unique. From (6.12) and Assumption 6.2, it follows that this process is non-explosive, at least not until it hits zero. Define $\tau_{2n+1} := \tau_{2n} + \tau'_{2n+1}$, and

$$X_1(t + \tau_{2n}) := W^{(2n+1)}(t), \qquad X_2(t + \tau_{2n}) := S_{2n+1}(t), \qquad t \leq \tau'_{2n+1}.$$

This defines X_1 and X_2 on the next time interval $[\tau_{2n}, \tau_{2n+1}]$ such that (6.18) is satisfied. The common driving Brownian motion B is defined as

$$B(t) := B(\tau_{2n}) + B_{2n+1}(t - \tau_{2n}), \qquad t \in [\tau_{2n}, \tau_{2n+1}].$$

If we have $X_1(\tau_{2n+1}) = X_2(\tau_{2n+1}) = \mathbf{0}$, we let $\tau_k := \tau_{2n+1}$ for $k \geq 2n + 1$. Otherwise, we repeat the construction above with X_1 and X_2 swapped, with obvious changes. This allows us to construct X_1 and X_2 on $[\tau_{2n+1}, \tau_{2n+2}]$ so that (6.18) and (6.19) hold for $k \leq 2n + 2$. The common driving Brownian motion B is defined as

$$B(t) := B(\tau_{2n+1}) + B_{2n+2}(t - \tau_{2n+1}), \qquad t \in [\tau_{2n+1}, \tau_{2n+2}].$$

By induction, we have defined $X_1(t)$ and $X_2(t)$ for $t \leq \tau_\infty$.

Step 5. It follows from (6.19) that $X_1(\tau_\infty) = X_2(\tau_\infty) = \mathbf{0}$. It suffices to define X_i, $i = 1, 2$ on $[\tau_\infty, \infty)$ if $\tau_\infty < \infty$. Assume $\tau_\infty < \infty$ and define $X_1(t)$ and $X_2(t)$ for $t \geq \tau_\infty$ by $X_1(t + \tau_\infty) = X_2(t + \tau_\infty) = W^{(0)}(t)$, $t \geq 0$. The common driving Brownian motion B is defined on $[\tau_\infty, \infty)$ as

$$B(t) := B(\tau_\infty) + B_0(t - \tau_\infty).$$

This completes the construction of two copies X_1 and X_2 of the Walsh diffusion, associated with (g, σ, μ), starting from $X_i(0) = x_i$, $i = 1, 2$, such that (6.18) holds.

Step 6. Finally, we construct strong versions S_1 and S_2 of a reflected diffusion with drift coefficient $\overline{g}$ and diffusion coefficient $\overline{\sigma}$, starting from $S_i(0) = \overline{r}_i$, with driving Brownian motion B, as in (6.15). This is possible by classic results, because $\overline{g}$ and $\overline{\sigma}$ are locally Lipschitz continuous. By standard comparison techniques, we have (6.14).

Part (b) Take a function $f : \mathcal{I} \to \mathbb{R}$ with $|f| \leq V$. Then $X = (X(t), t \geq 0)$ satisfies

$$\left| \mathbb{E}_{x_1}\left[f(X(t)) \right] - \mathbb{E}_{x_2}\left[f(X(t)) \right] \right| \leq \left(V(x_1) + V(x_2) \right) e^{-kt}, \qquad x_1, x_2 \geq 0.$$

Integrate over $\mathcal{I}$ with respect to $x_2 \sim \pi$. Then we have:

$$
\begin{aligned}
\left| \mathbb{E}_{x_1} \left[f\left(X(t)\right) \right] - (\pi, f) \right| &= \left| \mathbb{E}_{x_1} \left[f\left(X(t)\right) \right] - \int_{\mathcal{I}} \mathbb{E}_{x_2} \left[f\left(X(t)\right) \right] \pi(\mathrm{d}x_2) \right| \\
&\leq \int_{\mathcal{I}} \left| \mathbb{E}_{x_1} \left[f\left(X(t)\right) \right] - \mathbb{E}_{x_2} \left[f\left(X(t)\right) \right] \right| \pi(\mathrm{d}x_2) \\
&\leq e^{-kt} \left(V(x_1) + \int_{\mathcal{I}} V(x_2) \pi(\mathrm{d}x_2) \right) \\
&\leq e^{-kt} \left(V(x_1) + (\pi, V) \right) \leq e^{-kt} \left(1 + (\pi, V) \right) V(x_1). \qquad \square
\end{aligned}
$$

Remark 6.3. In Assumption 6.2, we imposed Lipschitz continuity assumption only to guarantee strong existence and pathwise uniqueness. It is well known, however, that these existence and uniqueness results hold under weaker conditions. In this case, our result also holds.

Let us present some corollaries. Under assumptions of Theorem 6.1, define

$$
K(x, \lambda) := \overline{g}(x)\lambda + \overline{\sigma}^2(x) \frac{\lambda^2}{2}. \tag{6.20}
$$

The following result is proved similarly to Sarantsev [29], Theorem 4.3, Corollary 5.2.

Corollary 6.1. *Under Assumption 6.2, suppose there exist* $k, \lambda > 0$ *such that*

$$
\sup_{x > 0} K(x, \lambda) =: -k < 0. \tag{6.21}
$$

Then the Walsh diffusion X *is* $V(r) := e^{\lambda r}$-*uniformly ergodic with* $\varkappa := k$, *and the stationary distribution* π *satisfies* $(\pi, V) < \infty$.

Remark 6.4. Under Assumption 6.2, suppose

$$
\sup_{x \in \mathcal{I} \setminus \{0\}} \overline{g}(x) =: -\tilde{g} < 0, \qquad \sup_{x \in \mathcal{I} \setminus \{0\}} \overline{\sigma}(x) =: \tilde{\sigma} < \infty.
$$

Similarly to Sarantsev [29], Corollary 4.4, we can show that (6.21) holds with $\lambda := \tilde{g}/\tilde{\sigma}$. Then, the Walsh diffusion X is $V(r) := e^{\lambda r}$-uniformly ergodic with rate $\varkappa := k$ of exponential convergence from (6.21).

Example 6.1 (Continuing from Example 3.1). Consider a Walsh diffusion X associated with (g, σ, μ), such that the drift and diffusion coefficients are constant:

$$
g(r, \theta) \equiv g < 0, \qquad \sigma(r, \theta) = \sigma > 0.
$$

Then we can take $\overline{g}(r) := g$, and $\overline{\sigma}(r) := \sigma$ in the Assumption 6.2. The expression K in (6.20) is reduced to $K(x, \lambda) = g\lambda + \frac{\sigma^2}{2}\lambda^2$, and is minimized at $\lambda = -g/\sigma^2$, with minimum $-k$, where

$$k := \frac{g^2}{2\sigma^2}. \tag{6.22}$$

This is an estimate of the exponential rate $\varkappa$ of convergence. Actually, this estimate is exact. Indeed, the radial component $\|X(\cdot)\|$ of such Walsh diffusion is a reflected Brownian motion on the half-line with drift g and diffusion σ^2, and such process is known from Lund, Meyn and Tweedie [19] to have exact rate of exponential convergence as in (6.22).

We can apply this result to (non-reflected) diffusions on the real line.

Corollary 6.2. *Take a solution $X = (X(t), t \geq 0)$ of an SDE on the real line $\mathbb{R}$ with drift co-efficient g and diffusion coefficient σ^2. Suppose it is ergodic. Assume that g and σ are locally Lipschitz, and $\sigma(x) = \sigma(-x)$ for all $x \in \mathbb{R}$. Suppose there exists a C^2 nondecreasing function $V : \mathbb{R}_+ \to [1, \infty)$ and a constant $k > 0$ such that*

$$\overline{g}(r)V'(r) + \frac{1}{2}\sigma^2(r)V''(r) \leq -kV(r), \qquad \text{where } \overline{g}(r) := g(r) \vee \left(-g(-r)\right), r > 0;$$

Finally, assume that the stationary distribution π satisfies $(\pi, V) < \infty$. Then X is V-uniformly ergodic with rate of convergence $\varkappa = k$.

Proof. Follows from Theorem 6.1 and the observation that the real line can be thought of as a "spider" with two rays, corresponding to North and South Pole, that is, θ_1 and θ_2. Then the process X becomes a spider associated with $(\tilde{g}, \sigma, \mu)$, where μ is a uniform measure on $\{\theta_1, \theta_2\}$, and $\tilde{g}(r, \theta_i) := g(r) \cdot 1_{\{i=1\}} - g(-r) \cdot 1_{\{i=2\}}, r \geq 0$. $\qquad \square$

Example 6.2 (Bang-bang drifts Karatzas and Shreve [17]). Consider a diffusion on $\mathbb{R}$ with drift and diffusion coefficients $g(x) = -g_1 \cdot 1_{\{x>0\}} + g_2 \cdot 1_{\{x\leq0\}}$, with constants $g_1, g_2 > 0$, and $\sigma(x) \equiv 1$, $x \in \mathbb{R}$. Then we can take $\overline{g}(r) = g_1 \wedge g_2$, and by Corollary 6.2 it is V-uniformly ergodic with $\lambda = g_1 \wedge g_2$ and rate $\varkappa = k = (g_1 \wedge g_2)^2/2$ of exponential convergence.

Appendix

Lemma A.1. *Take two sequences $(f_n)_{n\geq0}$ and $(g_n)_{n\geq0}$ of continuous functions $[0, T] \to \mathbb{R}^d$. Assume $t_n \in [0, T]$ are such that $f_n(t_n) = g_n(0) = \mathbf{0}$ for $n = 0, 1, 2, \ldots$ and $t_n \to t_0$. Assume $f_n \to f_0$ and $g_n \to g_0$ uniformly on $[0, T]$. Define the new sequence of functions $h_n : [0, T] \to \mathbb{R}^d$:*

$$h_n(t) := \begin{cases} f_n(t), & t \in [0, t_n]; \\ g_n(t - t_n), & t \in [t_n, T]. \end{cases}$$

Then $h_n \to h_0$ uniformly on $[0, T]$.

Proof. Without loss of generality, assume $t_n \uparrow t_0$, as $n \to \infty$. Otherwise, we can switch from t_n to $T - t_n$ and from $h_n(t)$ to $h_n(T - t)$. Then we can estimate the maximum difference

$$\max_{t \in [0,T]} \left\| h_n(t) - h_0(t) \right\|$$

$$\leq \max_{t \in [0,t_n]} \left\| h_n(t) - h_0(t) \right\| + \max_{t \in [t_n,t_0]} \left\| h_n(t) - h_0(t) \right\| + \max_{t \in [t_0,T]} \left\| h_n(t) - h_0(t) \right\|$$

$$= \max_{t \in [0,t_n]} \left\| f_n(t) - f_0(t) \right\| + \max_{t \in [t_n,t_0]} \left\| g_n(t - t_n) - f_0(t) \right\| + \max_{t \in [t_0,T]} \left\| g_n(t - t_n) - g_0(t - t_0) \right\|.$$

By uniform convergence $f_n \to f_0$, the first term can be estimated as

$$\max_{t \in [0,t_n]} \left\| f_n(t) - f_0(t) \right\| \leq \max_{t \in [0,t_0]} \left\| f_n(t) - f_0(t) \right\| \to 0. \tag{A.1}$$

The second term can be estimated as

$$\max_{t \in [t_n,t_0]} \left\| g_n(t - t_n) - f_0(t) \right\| \leq \max_{s \in [0,t_0-t_n]} \left\| g_n(s) \right\| + \max_{t \in [t_n,t_0]} \left\| f_0(t) \right\|. \tag{A.2}$$

Note that $\|g_n\| \to \|g_0\|$ uniformly on $[0, T]$, and $t_0 - t_n \to 0$ as $n \to \infty$. Thus,

$$\max_{s \in [0,t_0-t_n]} \left\| g_n(s) \right\| \to \left\| g_0(0) \right\| = 0. \tag{A.3}$$

Since $f_0(t_0) = \mathbf{0}$, $t_n \to t_0$, and $\|f_0\|$ is continuous, the second term in the right-hand side of (A.2) converges to zero, as $n \to \infty$. Combining this with (A.3), we get

$$\max_{t \in [t_n,t_0]} \left\| g_n(t - t_n) - f_0(t) \right\| \to 0. \tag{A.4}$$

Thirdly, the third term can be estimated by

$$\max_{t \in [t_0,T]} \left\| g_n(t - t_n) - g_0(t - t_0) \right\| \leq \max_{t \in [t_0,T]} \left\| g_n(t - t_n) - g_0(t - t_n) \right\|$$

$$+ \max_{t \in [t_0,T]} \left\| g_0(t - t_n) - g_0(t - t_0) \right\|. \tag{A.5}$$

The first term in the right-hand side of (A.5) can be estimated as

$$\max_{t \in [t_0,T]} \left\| g_n(t - t_n) - g_0(t - t_n) \right\| \leq \max_{s \in [0,T]} \left\| g_n(s) - g_0(s) \right\| \to 0. \tag{A.6}$$

The second term in the right-hand side of (A.5) also tends to zero as $n \to \infty$, because g_0 is continuous, and therefore uniformly continuous on $[0, T]$, while $t_n \to t_0$. Combining this observation with (A.6), we get that

$$\max_{t \in [t_0,T]} \left\| g_n(t - t_n) - g_0(t - t_0) \right\| \to 0. \tag{A.7}$$

Finally, combining (A.1), (A.4), (A.7), we complete the proof of Lemma A.1. $\square$

Lemma A.2. *For a Walsh diffusion X without drift from Case* 3 *of the proof of Theorem* 5.1, *we have* $\mathbb{P}(\tau_{R,A}^X < t) > 0$, *where* $\tau_{R,A}^X$ *is defined in* (5.3).

Proof of Lemma A.2. By Assumption 2.1, the functions σ and σ^{-1} are bounded on the (open) ball D with radius R, which, together with its closure $\overline{D}$, lies inside I. This implies

$$0 < c_1 \le \sigma^2(x) \le c_2 < \infty \qquad \text{for all } x \in \overline{D}. \tag{A.8}$$

Next, we can estimate the probability from Lemma A.2 as follows:

$$\mathbb{P}(\tau_{R,A}^X < t) \ge \mathbb{P}(\tau_{R,A}^X < t, \tau_{R'}^X > t). \tag{A.9}$$

It follows from (A.8) and (5.2) that

$$c_1 t \le T(t) \le c_2 t, \qquad \text{for } t \in [0, \tau_{R'}^X]. \tag{A.10}$$

Also, from definitions of hitting times (5.3)–(5.5), it immediately follows that

$$T(\tau_{R'}^X) = \tau_{R'}^W, \qquad T(\tau_{R,A}^X) = \tau_{R,A}^W. \tag{A.11}$$

The estimate (A.10) and the relations (A.11), in turn, imply that the inverted time-change $T^{-1}(s) := \inf\{t \ge 0 \mid T(t) \ge s\}$ satisfies

$$c_2^{-1} s \le T^{-1}(s) \le c_1^{-1} s, \qquad \text{for } s \in [0, \tau_{R'}^W]. \tag{A.12}$$

Let us show that

$$\{\tau_{R,A}^W < c_1 t < c_2 t < \tau_{R'}^W\} \subseteq \{\tau_{R,A}^X < t < \tau_{R'}^X\}. \tag{A.13}$$

Indeed, assume the event in the left-hand side of (A.13) has happened. Then from (A.11), applying the inverted time-change T^{-1}, we get

$$\tau_{R,A}^X < T^{-1}(c_1 t) < T^{-1}(c_2 t) < \tau_{R'}^X. \tag{A.14}$$

Applying (A.12) to (A.14), we have $T^{-1}(c_1 t) \le t$, $T^{-1}(c_2 t) \ge t$. Comparing this with (A.14), we get $\tau_{R,A}^X < t < \tau_{R'}^X$. This completes the proof of (A.13).

To complete the proof of Lemma A.2, we need only to show that

$$\mathbb{P}(\tau_{R,A}^W < c_1 t < c_2 t < \tau_{R'}^W) > 0. \tag{A.15}$$

Indeed, W is Walsh Brownian motion, starting from the origin. As noted earlier before, $\|W(t)\| = Z(t)$ is a reflected Brownian motion on the half-line, starting from zero. Define $\tau_a^Z := \inf\{t \ge 0 \mid Z(t) = a\}$ for $a > 0$. Then $\arg W(\tau_a^Z) \sim \mu$ is independent of Z. Therefore, the probability in the left-hand side of (A.15) is equal to

$$\mu(A) \cdot \mathbb{P}(\tau_R^Z < c_1 t < c_2 t < \tau_{R'}^Z) > 0. \tag{A.16}$$

That the left-hand side of (A.16) is indeed positive follows from $\mu(A) > 0$ and the properties of a reflected Brownian motion on the half-line. This proves (A.15), and with it Lemma A.2. □

Acknowledgements

References

[1] Bakry, D., Cattiaux, P. and Guillin, A. (2008). Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. *J. Funct. Anal.* **254** 727–759. MR2381160

[2] Barlow, M., Pitman, J. and Yor, M. (1989). On Walsh's Brownian motions. In *Séminaire de Probabilités, XXIII. Lecture Notes in Math.* **1372** 275–293. Berlin: Springer. MR1022917

[3] Chen, Z.-Q. and Fukushima, M. (2012). *Symmetric Markov Processes, Time Change, and Boundary Theory. London Mathematical Society Monographs Series* **35**. Princeton, NJ: Princeton Univ. Press. MR2849840

[4] Chen, Z.-Q. and Fukushima, M. (2015). One-point reflection. *Stochastic Process. Appl.* **125** 1368–1393. MR3310351

[5] Cloez, B. and Hairer, M. (2015). Exponential ergodicity for Markov processes with random switching. *Bernoulli* **21** 505–536. MR3322329

[6] Davies, P.L. (1986). Rates of convergence to the stationary distribution for k-dimensional diffusion processes. *J. Appl. Probab.* **23** 370–384. MR0839992

[7] Douc, R., Fort, G. and Guillin, A. (2009). Subgeometric rates of convergence of f-ergodic strong Markov processes. *Stochastic Process. Appl.* **119** 897–923. MR2499863

[8] Down, D., Meyn, S.P. and Tweedie, R.L. (1995). Exponential and uniform ergodicity of Markov processes. *Ann. Probab.* **23** 1671–1691. MR1379163

[9] Evans, S.N. and Sowers, R.B. (2003). Pinching and twisting Markov processes. *Ann. Probab.* **31** 486–527. MR1959800

[10] Feller, W. (1968). *An Introduction to Probability Theory and Its Applications. Vol. I.* New York–London–Sydney: Wiley. MR0228020

[11] Fitzsimmons, P.J. and Kuter, K.E. (2014). Harmonic functions on Walsh's Brownian motion. *Stochastic Process. Appl.* **124** 2228–2248. MR3188354

[12] Freidlin, M. and Sheu, S.-J. (2000). Diffusion processes on graphs: Stochastic differential equations, large deviation principle. *Probab. Theory Related Fields* **116** 181–220. MR1743769

[13] Freidlin, M.I. and Wentzell, A.D. (1993). Diffusion processes on graphs and the averaging principle. *Ann. Probab.* **21** 2215–2245. MR1245308

[14] Hajri, H. (2011). Stochastic flows related to Walsh Brownian motion. *Electron. J. Probab.* **16** 1563–1599. MR2835247

[15] Hajri, H. and Touhami, W. (2014). Itô's formula for Walsh's Brownian motion and applications. *Statist. Probab. Lett.* **87** 48–53. MR3168934

[16] Ichiba, T., Karatzas, I., Prokaj, V. and Yan, M. (2018). Stochastic integral equations for Walsh semimartingales. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** 726–756. MR3795064

[17] Karatzas, I. and Shreve, S.E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. New York: Springer. MR1121940

[18] Karatzas, I. and Yan, M. (2016). Semimartingales on rays, Walsh diffusions, and related problems of control and stopping. Available at arXiv:1603.03124.

[19] Lund, R.B., Meyn, S.P. and Tweedie, R.L. (1996). Computable exponential convergence rates for stochastically ordered Markov processes. *Ann*. *Appl*. *Probab*. **6** 218–237. MR1389838

[20] Meyn, S.P. and Tweedie, R.L. (1993). Stability of Markovian processes. II. Continuous-time processes and sampled chains. *Adv. in Appl*. *Probab*. **25** 487–517. MR1234294

[21] Meyn, S.P. and Tweedie, R.L. (1993). Stability of Markovian processes. III. Foster–Lyapunov criteria for continuous-time processes. *Adv. in Appl*. *Probab*. **25** 518–548. MR1234295

[22] Meyn, S.P. and Tweedie, R.L. (1994). Computable bounds for geometric convergence rates of Markov chains. *Ann*. *Appl*. *Probab*. **4** 981–1011. MR1304770

[23] Picard, J. (2005). Stochastic calculus and martingales on trees. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **41** 631–683. MR2144228

[24] Rachev, S.T. (1991). *Probability Metrics and the Stability of Stochastic Models*. *Wiley Series in Probability and Mathematical Statistics*: *Applied Probability and Statistics*. Chichester: Wiley. MR1105086

[25] Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **293**. Berlin: Springer. MR1725357

[26] Roberts, G.O. and Rosenthal, J.S. (1996). Quantitative bounds for convergence rates of continuous time Markov processes. *Electron*. *J*. *Probab*. **1** no. 9, approx. 21 pp. MR1423462

[27] Roberts, G.O. and Tweedie, R.L. (1999). Bounds on regeneration times and convergence rates for Markov chains. *Stochastic Process*. *Appl*. **80** 211–229. MR1682243

[28] Roberts, G.O. and Tweedie, R.L. (2000). Rates of convergence of stochastically monotone and continuous time Markov models. *J*. *Appl*. *Probab*. **37** 359–373. MR1780996

[29] Sarantsev, A. (2016). Explicit rates of exponential convergence for reflected jump-diffusions on the half-line. *ALEA Lat*. *Am*. *J*. *Probab*. *Math*. *Stat*. **13** 1069–1093. MR3580089

[30] Sarantsev, A. (2017). Penalty method for obliquely reflected diffusions. Available at arXiv:1509.01777.

[31] Sarantsev, A. (2017). Reflected Brownian motion in a convex polyhedral cone: Tail estimates for the stationary distribution. *J*. *Theoret*. *Probab*. **30** 1200–1223. MR3687255

[32] Tsirelson, B. (1997). Triple points: From non-Brownian filtrations to harmonic measures. *Geom*. *Funct*. *Anal*. **7** 1096–1142. MR1487755

[33] Villani, C. (2009). *Optimal Transport*. *Old and New*. *A Series of Comprehensive Studies in Mathematics* **338**. Springer.

[34] Walsh, J. (1978). A diffusion with a discontinuous local time. *Astérisque* **52–53** (53) 37–45.

[35] Watanabe, S. (1999). The existence of a multiple spider martingale in the natural filtration of a certain diffusion in the plane. In *Séminaire de Probabilités*, *XXXIII*. *Lecture Notes in Math*. **1709** 277–290. Berlin: Springer. MR1768001