ON THE APPROACHING TIME TOWARDS THE ATTRACTOR OF DIFFERENTIAL EQUATIONS PERTURBED BY SMALL NOISE

ISABELL VORKASTNER
Institut für Mathematik, Technische Universität Berlin
Strasse des 17. Juni 136
10623 Berlin, Germany

(Communicated by Martin Rasmussen)

Abstract. We estimate the time that a point or set, respectively, requires to approach the attractor of a radially symmetric gradient type stochastic differential equation driven by small noise. Here, both of these times tend to infinity as the noise gets small. However, the rates at which they go to infinity differ significantly. In the case of a set approaching the attractor, we use large deviation techniques to show that this time increases exponentially. In the case of a point approaching the attractor, we apply a time change and compare the accelerated process to a process on the sphere and obtain that this time increases merely linearly.

1. Introduction. One way to describe (random) dynamical systems and the asymptotic behavior of its trajectories is to examine the (random) attractors of the system. We differ between (random) point and set attractors. While we require (random) point attractors to attract any single point, (random) set attractors are even required to attract compact sets uniformly. If an (random) attractor is a single (random) point, the long-time dynamics are asymptotically globally stable.

As shown in [8], the addition of noise to a deterministic system can stabilize the system. Here, the deterministic system is not asymptotically globally stable and the attractor of the deterministic system is not a single point. However, the random attractor collapses to a single random point under the addition of noise. Since the random system behaves similarly to its deterministic system on a finite time interval for small noise, we can anticipate that the time until the trajectories approach the attractor goes to infinity as the noise gets small. In the following we differ between the time required for a point or set, respectively, to approach the attractor, where the set is chosen sufficiently large.

Our aim is to provide an example for which the approaching time of a point and the approaching time of a set towards the attractor differ significantly. Further we want to estimate these times and provide the rates at which they tend to infinity.

2010 Mathematics Subject Classification. Primary: 37G35, 37H99; Secondary: 60F10, 60H10.

Key words and phrases. random attractor, stochastic differential equation, synchronization, large deviation principle.
We consider a radially symmetric gradient type stochastic differential equation, i.e.,

\[ dX_t^\varepsilon = -\nabla U(X_t^\varepsilon) \, dt + \sqrt{\varepsilon} \, dW_t \quad \text{on} \quad \mathbb{R}^d, \quad (1) \]

where \( d \geq 2, \varepsilon > 0, (W_t : t \geq 0) \) is a \( d \)-dimensional Brownian motion, \( U : \mathbb{R}^d \to \mathbb{R} \) is differentiable and \( U(x) = u(|x|^2) \) for all \( x \in \mathbb{R}^d \) and some twice differentiable convex function \( u : [0, \infty) \to \mathbb{R} \) attaining its unique minimum in \((0, \infty)\).

In the absence of noise, the solution of the differential equation

\[ dx_t = -\nabla U(x_t) \, dt \quad \text{on} \quad \mathbb{R}^d. \quad (2) \]

has a stable sphere, meaning that any point on this sphere is a fixed point of (2) and any point except 0 converges towards the sphere under the dynamics of (2). The point 0 is also a fixed point of (2). In terms of attractors this means that the point attractor of (2) is the union of 0 and the stable sphere while the set attractor (2) is the closed ball of the same radius as the stable sphere centered at 0.

In contrast, the random set attractor and minimal point attractor of (1) collapses to a single random point under the addition of noise, see [8, Theorem 3.15 and 3.16]. This phenomenon is called synchronization by noise. We say that a random dynamical system synchronizes weakly if it has a random point attractor being a single random point. If such a random single point is additionally a random set attractor we say that synchronization occurs. In the literature synchronization has been investigated, for example, in [3, 9, 11] for order-preserving random dynamical systems and in [12] for iterated function systems.

Using large deviation techniques, this phenomenon is also analyzed in [14] and [15]. There, it is assumed that \( \nabla U \) has finitely many fixed points and they proved synchronization for small noise. In particular, our equation (1) is covered for \( d = 1 \). Moreover in [14], the authors provide an upper bound for the time until trajectories started in a bounded ball approach each other. This upper bound increases exponentially in \( \varepsilon^{-1} \) which is the same time scale that we get in Section 3.

In the one-dimensional case, \( d = 1 \), one can estimate the approaching time of a point and a set towards the attractor by computing the time in which a process started in a point requires to exit a domain using Freidlin-Wentzel theory (see [10, Chapter 2] or [6, Chapter 5]) or solving the Poisson problem (see [13, Section 5.5]). These methods show that both approaching times, starting in a point and in a set, increase exponentially in \( \varepsilon^{-1} \).

In contrast to the one-dimensional case, the rates of both approaching times differ in the two-dimensional case. By the above observations, both approaching times should go to infinity as the noise gets small. We estimate the rates at which these times tend to infinity and show that a point approaches the attractor significantly faster than a set. This result shows that the effects of weak synchronization can be seen long before synchronization is observable.

The difference in the approaching rates is due to the fact that a point can approach the attractor moving to the stable sphere and then along the sphere while a sufficiently large set can just approach the attractor if a point of the stable sphere moves close to zero.

Another way to explain the difference is to consider the behavior of the process \( X_t^\varepsilon \) as \( \varepsilon \to 0 \). Perturbed by small noise, the process behaves similar to a process on the sphere with one stable and one unstable random point (see Section 4). If one now starts in a set containing this unstable point, the process (1) requires a lot of
time to approach the stable point while the process started in point with positive distance to the unstable point approaches the stable point comparatively fast.

In Section 3, we show that the time until a set approaches the attractor increases exponentially in $\varepsilon^{-1}$ using large deviation techniques similar to [6, Section 5.7]. We obtain a lower bound by taking the difference between the potential at the stable sphere and 0, since the potential describes the costs of a point on the stable sphere to approach zero. Assuming that the differential equation (2) pushes all mass into a bounded set in finite time, we get an upper bound for the time. This estimate in particular demonstrates the sharpness of our lower bound.

In Section 4, we prove that the time until a point approaches the attractor is of order $\varepsilon^{-1}$ for dimension $d = 2$. We accelerate the process and compare the accelerated process to a process on the sphere that is known to synchronize weakly.

2. Preliminaries. We consider stochastic differential equation (SDE) (1) and assume that $-\nabla U$ satisfies a one-sided-Lipschitz condition, i.e., there exists some $C > 0$ such that

$$\langle x - y, -\nabla U(x) + \nabla U(y) \rangle \leq C|x - y|^2$$

for all $x,y \in \mathbb{R}^d$. Then, SDE (1) has a unique solution. We denote by $X^\varepsilon : [0,\infty) \times \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ the solution of (1).

We say that SDE (1) is strongly contracting if there exist $r, t > 0$ such that $|x_t(y)| \leq r$ for all $y \in \mathbb{R}^d$, where $x_t(y)$ is the solution (2) started in $y$ at time $t$. Therefore, SDE (1) is strongly contracting if and only if $\int_0^\infty |\nabla U(x)|^{-1} \, dx$ exists for some $R > 0$.

Let $R^* \in (0,\infty)$ be the point where $u$ attains its minimum, i.e. $u(R^*) < u(x)$ for any $x \neq R^*$. We restrict the proofs in the following sections to the case $R^* = 1$. However, all results are extendable to general $R^* \in (0,\infty)$ since $X^\varepsilon_t/R^*$ is of the postulated form.

Attractors and synchronization are defined for random dynamical systems. We restrict our definitions to random dynamical system on $\mathbb{R}^d$, see [1] for a more general setting.

**Definition 2.1** (Metric Dynamical System). Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\theta = (\theta_t)_{t \in \mathbb{R}}$ be a group of maps $\theta_t : \Omega \rightarrow \Omega$ satisfying

(i) $(\omega, t) \mapsto \theta_t(\omega)$ is $(\mathcal{F} \otimes \mathcal{B}(\mathbb{R}), \mathcal{F})$-measurable,

(ii) $\theta_0(\omega) = \omega$ for all $\omega \in \Omega$,

(iii) $\theta_{s+t} = \theta_s \circ \theta_t$ for all $s, t \in \mathbb{R}$,

(iv) $\theta_t$ has ergodic invariant measure $\mathbb{P}$ for any $t \in \mathbb{R}$.

The collection $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ is called a metric dynamical system.

**Definition 2.2** (Random Dynamical System). Let $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ be a metric dynamical system and $\varphi : \mathbb{R}_+ \times \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ be such that

(i) $\varphi$ is $(\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F} \otimes \mathcal{B}(\mathbb{R}^d), \mathcal{B}(\mathbb{R}^d))$-measurable,

(ii) $\varphi_0(\omega, x) = x$ for all $x \in \mathbb{R}^d$, $\omega \in \Omega$,

(iii) $\varphi_{t+s}(\omega, x) = \varphi_t(\theta_s(\omega), \varphi_s(\omega, x))$ for all $x \in \mathbb{R}^d$, $t, s \geq 0$, $\omega \in \Omega$,

(iv) $x \mapsto \varphi_s(\omega, x)$ is continuous for each $s \geq 0$ and $\omega \in \Omega$.

The collection $(\Omega, \mathcal{F}, \mathbb{P}, \theta, \varphi)$ is called a random dynamical system (RDS).

**Definition 2.3.** A family $\{D(\omega)\}_{\omega \in \Omega}$ of non-empty subsets of $\mathbb{R}^d$ is said to be
(i) a random compact set if it is \(\mathbb{P}\)-almost surely compact and \(\omega \mapsto \sup_{y \in \mathcal{D}(\omega)} |x - y|\) is \(\mathcal{F}\)-measurable for each \(x \in \mathbb{R}^d\).

(ii) \(\varphi\)-invariant if for all \(t \geq 0\)

\[
\varphi_t(\omega, D(\omega)) = D(\theta_t \omega)
\]

for almost all \(\omega \in \Omega\).

**Definition 2.4** (Attractor). Let \((\Omega, \mathcal{F}, \mathbb{P}, \theta, \varphi)\) be a RDS and \(A\) be \(\varphi\)-invariant random compact set \(A\).

(i) \(A\) is called a weak point attractor if for every \(x \in \mathbb{R}^d\)

\[
\lim_{t \to \infty} \inf_{a \in A(\omega)} |\varphi_t(\theta^{-t} \omega, x) - a| = 0 \quad \text{in probability.}
\]

(ii) \(A\) is called a weak attractor if for every compact set \(B \subset \mathbb{R}^d\)

\[
\lim_{t \to \infty} \sup_{x \in B} \inf_{a \in A(\omega)} |\varphi_t(\theta^{-t} \omega, x) - a| = 0 \quad \text{in probability.}
\]

By [7], SDE (1) generates a RDS \((\Omega, \mathcal{F}, \mathbb{P}, \theta, X)\) with respect to the canonical setup and this RDS has a weak attractor. Note that every weak attractor is a weak point attractor. The converse is not true.

Here, the space \(\Omega\) is \(\mathcal{C}(\mathbb{R}, \mathbb{R}^d)\), \(\mathcal{F}\) is the Borel \(\sigma\)-field, \(\mathbb{P}\) is the two-sided Wiener measure, \(\mathcal{F}\) is the \(\sigma\)-algebra generated by \(W_u - W_v\) for \(v \leq u\), where \(W_v : \Omega \to \mathbb{R}^d\) is defined as \(W_v(\omega) = \omega(s)\), and \(\theta_t\) is the shift \((\theta_t \omega)(s) = \omega(s + t) - \omega(t)\). Further, define \(\mathcal{F}^+\) as the \(\sigma\)-algebra generated by \(W_u - W_v\) for \(0 \leq v \leq u\) and \(\mathcal{F}^-\) as the \(\sigma\)-algebra generated by \(W_u - W_v\) for \(v \leq u \leq 0\).

**Definition 2.5** (Synchronization). Synchronization occurs if there is a weak attractor \(A(\omega)\) being a singleton for \(\mathbb{P}\)-almost every \(\omega \in \Omega\). Weak synchronization is said to occur if there is a weak point attractor \(A(\omega)\) being a singleton for \(\mathbb{P}\)-almost every \(\omega \in \Omega\).

We do not require that the RDS to synchronize (weakly) in order to get lower and upper bounds on the time required to approach the attractor. However, we differ between the smallest and largest distance to the attractor. Both quantities coincide if the RDS synchronizes (weakly).

The paper [8] provides general conditions for the RDS associated to (1) to synchronize (weakly). If \(U(x) = u(|x|^2)\) in SDE (1) additionally satisfies \(u \in \mathcal{C}_\text{loc}^3\), \(\log^+ |x| \exp(-2u(|x|^2)/\varepsilon) \in L^1(\mathbb{R}^d)\) and \(|u'''(x)| \leq C(|x|^m + 1)\) for some \(m \in \mathbb{N}\), \(C \geq 0\) and where \(u''\) is the third derivative of \(u\), then the associated RDS synchronizes by [8, Theorem 3.15 and 3.16]. The assumption \(\log^+ |x| \exp(-2u(|x|^2)/\varepsilon) \in L^1(\mathbb{R}^d)\) is in particular satisfied for a strongly contracting SDE (1).

Note that synchronization implies weak synchronization. It is left as an open problem in [8] whether any RDS associated to SDE (1) satisfying \(\exp(-2u(|x|^2)/\varepsilon) \in L^1(\mathbb{R}^d)\) synchronizes weakly.

Denote by

\[
B_r := \{x \in \mathbb{R}^d : |x| < r\}
\]

the open ball of radius \(r > 0\) centered at 0 and by

\[
S_r := \{x \in \mathbb{R}^d : |x| = r\}
\]

the sphere of radius \(r > 0\) centered at 0. For a set \(M \subset \mathbb{R}^d\) denote by \(M^\circ\) the interior and by \(\overline{M}\) the closure of the set \(M\).
3. Time required for a set to approach the attractor.

3.1. Large deviation principle. We use the large deviation principle (LDP) to describe the behavior of $X^\varepsilon_t$ for small $\varepsilon > 0$. The semi-flow $(X^\varepsilon_t)_{t \geq 0}$ for some $T > 0$ satisfies the LDP with rate function \cite[(6, page 185)]{6} \(I : C([0, T] \times \mathbb{R}^d, \mathbb{R}^d) \to \mathbb{R}\) if for all $\Gamma \subset X$ for any $\delta > 0$ and a set $M \subset \mathbb{R}^d$.

\[
\inf_{\phi \in \Gamma} I(\phi) = \lim_{\varepsilon \to 0} \inf \varepsilon \log \mathbb{P}(X_t \in \Gamma) \leq \lim_{\varepsilon \to 0} \sup \varepsilon \log \mathbb{P}(X_t \in \Gamma) \leq -\inf_{\phi \in \Gamma} I(\phi).
\]

We aim to give an estimate on the time a set needs to approach the weak attractor. Observe that by \cite[Theorem 3.1]{7} there exists a weak attractor of the RDS associated to (1) and that the weak attractor is \(\mathbb{P}\)-almost surely uniquely by \cite[Lemma 1.3]{8}. We denote by $A^{X,\varepsilon}$ the weak attractor.

Let $\mu^\varepsilon_t$ be the probability measure induced by $\sqrt{\varepsilon}W_t$ on $C_0([0, T])$, the space of all continuous functions $\phi : [0, T] \to \mathbb{R}^d$ such that $\phi(0) = 0$ equipped with the supremum norm topology. By Schilder’s theorem \cite[(6, page 185)]{6} $\mu^\varepsilon_t$ satisfies an LDP with good rate function

\[
\hat{I}_T(g) = \begin{cases} \frac{1}{2} \int_0^T |\dot{g}(t)|^2 \, dt, & g \in \left\{ t \mapsto \int_0^t h(s) \, ds : h \in L^2([0, T]) \right\}, \\ \infty, & \text{otherwise} \end{cases}
\]

for $g \in C_0([0, T])$ and $\dot{g}$ denotes the derivative of $g$. The deterministic map $F_T : C_0([0, T]) \to C([0, T] \times \mathbb{R}^d, \mathbb{R}^d)$ is defined by $F_T(g)(t, f) = f(t)$, where $f$ is the semi-flow associated to

\[
f(t) = f_0 - \int_0^t \nabla U(f(s)) \, ds + g(t), \quad t \in [0, T].
\]

Hence, the LDP associated to the semi-flow $(X^\varepsilon_t)_{t \geq 0}$ is a direct application of the contraction principle with respect to the continuous map $F_T$. Therefore, $X^\varepsilon_t$ satisfies the LDP in $C([0, T] \times \mathbb{R}^d, \mathbb{R}^d)$ with good rate function

\[
I_T(f) = \inf \left\{ \hat{I}_T(g) : g \in C_0([0, T]) \right\}.
\]

for $f \in C([0, T] \times \mathbb{R}^d, \mathbb{R}^d)$. In particular, for suitable $f \in C([0, T] \times \mathbb{R}^d, \mathbb{R}^d)$ the good rate function is equal to

\[
I_T(f) = \frac{1}{2} \int_0^T |\dot{f}(s, x) + \nabla U(f(s, x))|^2 \, dt.
\]

for any $x \in \mathbb{R}^d$ (otherwise $I(f) = \infty$). Define the stopping times

\[
\tau^{\varepsilon}_{1,\delta} := \inf \left\{ t \geq 0 : |X^\varepsilon_t(x) - X^\varepsilon_t(y)| \leq \delta \text{ for all } x, y \in S_1 \right\},
\]

\[
\tau^{\varepsilon}_{2,\delta,M} := \inf \left\{ t \geq 0 : \sup_{a \in A^{X,\varepsilon}(\theta_t)} |X^\varepsilon_t(x) - a| \leq \delta \text{ for all } x \in M \right\}
\]

\[
\tau^{\varepsilon}_{3,\delta} := \inf \left\{ t \geq 0 : |X^\varepsilon_t(x) - X^\varepsilon_t(y)| \leq \delta \text{ for all } x, y \in \mathbb{R}^d \right\}
\]

for $\delta > 0$ and a set $M \subset \mathbb{R}^d$. The first stopping time $\tau^{\varepsilon}_{1,\delta}$ describes the time until the process started in the unit sphere contracts, while the third stopping time $\tau^{\varepsilon}_{3,\delta}$ describes the time until the process started in $\mathbb{R}^d$ contracts. Obviously, $\tau^{\varepsilon}_{1,\delta} \leq \tau^{\varepsilon}_{3,\delta}$ for any $\delta > 0$. The second stopping time $\tau^{\varepsilon}_{2,\delta,M}$ relates to the attractor. Here, $\tau^{\varepsilon}_{2,\delta,M}$ describes the time the set $M$ needs to approach the attractor $A^{X,\varepsilon}$. Observe that $\tau^{\varepsilon}_{1,\delta} \leq \tau^{\varepsilon}_{2,\delta,M} \leq \tau^{\varepsilon}_{3,\delta}$ for any $\delta > 0$ and $S_1 \subset M \subset \mathbb{R}^d$. 
In the next subsection we use the LDP to show a lower bound for \( \tau_{1,\delta} \) and an upper bound for \( \tau_{3,\delta} \). We then conclude this section combining these estimates and showing that \( \tau_{1,\delta} \), \( \tau_{2,\delta,M} \) and \( \tau_{3,\delta} \) are roughly of order \( \exp(\sqrt{\varepsilon}) \) for some \( V > 0 \).

3.2. Lower bound for \( \tau_{1,\delta} \). In this subsection we show a lower bound for \( \tau_{1,\delta} \) (see (5)). Using the gradient type form of SDE (1), we provide an upper bound for the probability that this stopping time is smaller than some deterministic time. Afterwards, we use a similar approach as in [6, Section 5.7] to deduce that \( \tau_{1,\delta} \) is roughly greater than \( \exp(\sqrt{\varepsilon}) \) where \( V > 0 \) is determined by the potential \( U \).

Define the annulus

\[ D_{r,R} := \{ x \in \mathbb{R}^d : r < |x| < R \} \]

for \( 0 \leq r < R \leq \infty \). Moreover, denote by

\[ \tau^\varepsilon(M,D) := \inf \{ t \geq 0 : X^\varepsilon_t(x) \notin D \text{ for some } x \in M \} \]

the time until the semi-flow started in \( M \subset \mathbb{R}^d \) leaves \( D \subset \mathbb{R}^d \). Observe that \( \tau^\varepsilon(M,D) = \tau^\varepsilon(\partial M,D) \) for any closed set \( M \subset \mathbb{R}^d \), where \( \partial M \) is the boundary of \( M \). For \( 0 \leq r_1 < r_2 < r_3 \leq \infty \) with \( r_1 < 1 < r_3 \) set

\[ V(r_1, r_2, r_3) := 2 \min \{ u(r_1^2) - u(\min \{ r_2^2, 1 \}), u(r_2^2) - u(\max \{ r_2^2, 1 \}) \} \]

where \( u(\infty) := \lim_{x \to \infty} u(x) = \infty \). We show that \( V \) represents the cost of forcing the system (1) started on sphere \( S_{r_2} \) to leave the annulus \( D_{r_1,r_3} \).

**Lemma 3.1.** Let \( 0 \leq r_1 < r_2 < r_3 \leq \infty \) with \( r_1 < 1 < r_3 \) and let \( T > 0 \). Then,

\[ \lim_{\varepsilon \to 0} \varepsilon \log \mathbb{P} (\tau^\varepsilon(S_{r_2}, D_{r_1,r_3}) \leq T) \leq -V(r_1, r_2, r_3) \]

**Proof.** For any \( \phi \in C([0,T] \times \mathbb{R}^d, \mathbb{R}^d), x \in \mathbb{R}^d \) and \( 0 \leq s < t \leq T, I(\phi) = \infty \) or

\[ I_T(\phi) \geq \frac{1}{2} \int_s^t \left| \dot{\phi}(u,x) + \nabla U(\phi(u,x)) \right|^2 \, du \]

\[ = \frac{1}{2} \int_s^t \left| \phi(u,x) - \nabla U(\phi(u,x)) \right|^2 \, du + 2 \int_s^t \langle \dot{\phi}(u,x), \nabla U(\phi(u,x)) \rangle \, du \]

\[ \geq 2 \left( U(\phi(t,x)) - U(\phi(s,x)) \right) = 2 \left( u(|\phi(t,x)|^2) - u(|\phi(s,x)|^2) \right). \]

Define

\[ \Phi_i := \{ \phi \in C([0,T] \times \mathbb{R}^d, \mathbb{R}^d) : \phi(0, \cdot) = Id \text{ and } |\phi(t,x)| = r_i \text{ for some } x \in S_{r_2}, t \in [0,T] \} \]

for \( i = 1, 3 \). By the LDP it follows that

\[ \lim_{\varepsilon \to 0} \varepsilon \log \mathbb{P} (\tau^\varepsilon(S_{r_2}, D_{r_1,r_3}) \leq T) \leq - \inf_{\phi \in \Phi_1 \cup \Phi_3} I_T(\phi). \]

(7)

We consider the case \( r_2 \leq 1 \). If \( \phi \in \Phi_1 \), there exists \( x \in S_{r_2} \) and \( t \in [0,T] \) such that \( |\phi(t,x)| = r_1 \). By (6), \( I_T(\phi) \geq 2(u(r_1^2) - u(r_2^2)). \) If \( \phi \in \Phi_3 \), there exists \( x \in S_{r_2} \) and \( 0 \leq s < t \leq T \) such that \( |\phi(s,x)| = 1 \) and \( |\phi(t,x)| = r_3 \). Using (6), it follows that \( I_T(\phi) \geq 2(u(r_3^2) - u(1)). \) Taking the Infimum over \( \Phi_1 \) and \( \Phi_3 \) as in (7) the statement follows for \( r_2 \leq 1 \). Repeating the same arguments, the statement follows for \( r_2 > 1 \).
Denote by
\[ \sigma^\varepsilon(M, D) := \inf \{ t \geq 0 : X^\varepsilon_t(x) \in D \text{ for all } x \in M \} \]
the time until \( D \subset \mathbb{R}^d \) contains the semi-flow started in \( M \subset \mathbb{R}^d \).

The next lemma shows that after some time \( t \), which does not depend on the strength of the noise, the semi-flow started in an annulus is contained in a neighborhood of the stable sphere for small noise. This time \( t \) roughly coincides with the time the semi-flow of the ODE (2) started in the annulus requires to be contained in the neighborhood since the semi-flow of SDE (1) behaves similar to the semi-flow of the ODE (2) for small noise on a fixed time scale.

**Lemma 3.2.** Let \( 0 < r_1 < r_2 < \infty \) and \( 0 \leq r_3 < 1 < r_4 \leq \infty \). Then
\[
\lim_{t \to \infty} \limsup_{\varepsilon \to 0} \varepsilon \log \mathbb{P} (\sigma^\varepsilon(D_{r_1, r_2}, D_{r_3, r_4}) > t) \leq -V(0, r_1, \infty).
\]

**Proof.** Note that \( V(0, r_1, \infty) \leq V(0, r_2, \infty) \) and set \( V := V(0, r_1, \infty) > 0 \) and \( 0 < \delta < V/2 \).

We choose \( 0 < \alpha < r_1 < r_2 < \beta \) such that \( V(\alpha, r_1, \beta) \geq V - \delta/2 \) and \( V(\alpha, r_2, \beta) \geq V - \delta/2 \).

Set \( M := D_{r_1, r_2} \) and \( N := D_{\alpha, \beta} \). It holds that
\[
\mathbb{P} (\sigma^\varepsilon(M, D_{r_3, r_4}) > t) \leq \mathbb{P} (\tau^\varepsilon(M, N) > t) + \mathbb{P} (\tau^\varepsilon(M, N) \leq t).
\]

By Lemma 3.1 for any \( t > 0 \) there exists \( \varepsilon_0 > 0 \) such that
\[
\mathbb{P} (\tau^\varepsilon(M, N) \leq t) \leq \mathbb{P} (\tau^\varepsilon(S_1 \cup S_2, N) \leq t) \leq \mathbb{P} (\tau^\varepsilon(S_{r_1}, N) \leq t) + \mathbb{P} (\tau^\varepsilon(S_{r_2}, N) \leq t) \leq 2 \exp(-V - \delta)/\varepsilon
\]
for all \( \varepsilon \leq \varepsilon_0 \). We consider the closed sets
\[
\Psi_t := \{ \phi \in C([0, t] \times \mathbb{R}^d, \mathbb{R}^d) : \phi(r, y) \in N \text{ for all } r \in [0, t] \text{ and } y \in M \text{ and for each } s \in [0, t] \text{ there exists an } x \in M \text{ such that } \phi(s, x) \notin D_{r_3, r_4}, \}
\]
\[
\tilde{\Psi}_t := \{ \phi \in C([0, t] \times \mathbb{R}^d, \mathbb{R}^d) : \text{ for each } s \in [0, t] \text{ there exists an } x \in N \text{ such that } \phi(s, x) \notin D_{r_3, r_4} \}.
\]

The event \( \{ \tau^\varepsilon(M, N) > t \} \cap \{ \tau^\varepsilon(M, D_{r_3, r_4}) > t \} \) is contained in \( \{ X^\varepsilon_t \in \Psi_t \} \). By the LDP
\[
\limsup_{\varepsilon \to 0} \varepsilon \log \mathbb{P} (\tau^\varepsilon(M, N) > t) + \mathbb{P} (\tau^\varepsilon(M, D_{r_3, r_4}) > t) \leq -\inf_{\phi \in \tilde{\Psi}_t} I_t(\phi).
\]

It suffices to show that
\[
\lim_{t \to \infty} \inf_{\phi \in \Psi_t} I_t(\phi) = \infty.
\]

There exists a \( T > 0 \) such that the semi-flow associated to the deterministic ODE (2) started in \( N \) in \( D_{r_3, r_4} \) at time \( T \). Assume that (8) is false. Then, there exists a \( c > 0 \) such that for every \( n \in \mathbb{N} \) there exists \( \tilde{\phi}_n \in \tilde{\Psi}_{nT} \) with \( I_{nT}(\tilde{\phi}_n) \leq c \).

Hence, there exists a \( g_n \in C_0([0, nT]) \) with \( F_{nT}(g_n) = \phi_n \) and \( I_{nT}(g_n) \leq 2c \).

Set \( g_{n, k}(t) := g_n(t + kT) - g_n(kT) \) for \( 0 \leq k \leq n - 1 \) and \( 0 \leq t \leq T \). We define \( \phi_{n,k} := F_T(g_{n,k}) \). Observe that \( \phi_{n,k} \in \Psi_T \) since for each \( s \in [0, T] \) there exists an
By Lemma 3.2 there exists a $T > 0$ such that $\phi_n(kT, x) = \phi_n(kT + t, x) \notin D_{r_3, r_4}$ and $\phi_n(kT, x) \in N$. By the definition of $g_n$, it follows that

$$\sum_{k=0}^{n-1} \hat{I}_T(g_{n,k}) = \frac{1}{2} \int_0^T (\hat{g}(t+kT))^2 \, dt = \frac{1}{2} \int_0^n (\hat{g}(t))^2 \, dt = \hat{I}_n T(g_n) \leq 2c$$

for all $n \in \mathbb{N}$. Then, for each $n \in \mathbb{N}$ there exists a $0 \leq k \leq n - 1$ such that $\hat{I}_T(g_{n,k}) \leq \frac{2c}{n}$. Hence there exists a sequence $h_n \in C_0([0, T])$ with $\lim_{n \to \infty} \hat{I}_T(h_n) = 0$ and $F_T(h_n) \in \tilde{Ψ}_T$ for all $n \in \mathbb{N}$. Arzelà-Ascoli Theorem implies that $\{ h \in C_0([0, T]) : \hat{I}_T(h) \leq 2c \}$ is a compact subset of $C_0([0, T])$. Therefore, the sequence $h_n$ has a limit point $h$ in $C_0([0, T])$. Continuity of $F_T$ implies that $\psi := F_T(h) \in \tilde{Ψ}_T$. By lower semi-continuity of $\hat{I}_T$, $\hat{I}_T(\psi) = 0$ and $\psi$ describes the flow of the deterministic ODE (2). By definition of $T$, for all $x \in N$ it holds that $\psi(T, x) \in D_{r_3, r_4}$ which is a contradiction to $\psi \in \tilde{Ψ}_T$.

**Proposition 1.** Let $0 \leq r_1 < r_2 < r_3 \leq \infty$ with $r_1 < 1 < r_3$. Set $V := V(r_1, 1, r_3)$. For any $\beta > 0$ it holds that

$$\lim_{\varepsilon \to 0} \mathbb{P}(\tau^\varepsilon(S_{r_2}, D_{r_1, r_3}) > \exp((V - \beta)/\varepsilon)) = 1$$

and

$$\lim_{\varepsilon \to 0} \mathbb{E} \log \mathbb{E} \tau^\varepsilon(S_{r_2}, D_{r_1, r_3}) \geq V.$$

**Proof.** Let $\beta < V(r_1, r_2, r_3)$ and $\eta > 0$ be small enough such that $r_1 < 1 - 2\eta$, $r_3 > 1 + 2\eta$, $V(r_1, 1 - 2\eta, r_3) > V - \beta/2$ and $V(r_1, 1 + 2\eta, r_3) > V - \beta/2$. Let $\rho_0 = 0$ and for $n \in \mathbb{N}_0$ define the stopping times

$$\sigma_n := \inf \{ t \geq \rho_n : |X_t^\varepsilon(x)| \notin (1 - \eta, 1 + \eta) \text{ for all } x \in S_{r_2} \}
$$

or

$$|X_t^\varepsilon(x)| \notin (r_1, r_3) \text{ for some } x \in S_{r_2},$$

and

$$\rho_{n+1} := \inf \{ t \geq \sigma_n : |X_t^\varepsilon(x)| \notin (1 - 2\eta, 1 + 2\eta) \text{ for some } x \in S_{r_2} \}$$

with convention that $\rho_{n+1} = \infty$ if $\sigma_n = \tau^\varepsilon(S_{r_2}, D_{r_1, r_3})$. During each time interval $[\rho_n, \sigma_n]$ one point of the semi-flow either leaves the annulus $D_{r_1, r_3}$ or the semi-flow reenters the smaller annulus $D_{1-\eta, 1+\eta}$. Note that necessarily $\tau^\varepsilon(S_{r_2}, D_{r_1, r_3}) = \sigma_n$ for some $n \in \mathbb{N}_0$.

![Figure 1. Outline of the set $|X_t^\varepsilon(S_{r_2})|$ and the stopping times $\sigma_n$ and $\rho_n$.](image)

By Lemma 3.2 there exists a $T > 0$ and $\varepsilon_1 > 0$ such that

$$\mathbb{P}(\sigma_0 > T) \leq \mathbb{P}(\sigma^\varepsilon(S_{r_2}, D_{1-\eta, 1+\eta}) > T) \leq \exp(-(V(r_1, r_2, r_3) - \beta/2)/\varepsilon)$$
and
\[ \mathbb{P}(\sigma_n - \rho_n > T) \leq \mathbb{P}(\sigma^\varepsilon(D_{1-2\eta,1+2\eta}, D_{1-\eta,1+\eta}) > T) \]
\[ \leq \exp(-(V - \beta/2)/\varepsilon) \]

for all \( n \in \mathbb{N} \) and \( \varepsilon \leq \varepsilon_1 \). Using Lemma 3.1 and the estimates above, there exists \( \varepsilon_2 > 0 \) such that
\[ \mathbb{P}(\tau^\varepsilon(S_r, D_{r_1,r_3}) = \sigma_0) \leq \mathbb{P}(\sigma_0 > T) + \mathbb{P}(\tau^\varepsilon(S_r, D_{r_1,r_3}) \leq T) \]
\[ \leq 2\exp(-(V(r_1, r_2, r_3) - \beta/2)/\varepsilon) \] (9)

and
\[ \mathbb{P}(\tau^\varepsilon(S_r, D_{r_1,r_3}) = \sigma_n) \leq \mathbb{P}(\sigma_n - \rho_n > T) + \mathbb{P}(\tau^\varepsilon(S_{1-2\eta}, D_{r_1,r_3}) \leq T) \]
\[ + \mathbb{P}(\tau^\varepsilon(S_{1+2\eta}, D_{r_1,r_3}) \leq T) \]
\[ \leq 3\exp(-(V - \beta/2)/\varepsilon) \] (10)

for all \( n \in \mathbb{N} \) and \( \varepsilon \leq \varepsilon_2 \). Choose \( T_0 > 0 \) such that \( 2dT_0(V - \beta/2) \leq \eta^2 \). By [6, Lemma 5.2.1], for all \( n \in \mathbb{N} \)
\[ \mathbb{P}(\rho_n - \sigma_{n-1} \leq T_0) \leq \mathbb{P}\left(\sup_{t \in [0,T_0]} \sqrt{\varepsilon} |W_t| \geq \eta\right) \leq 4d\exp(-\eta^2/(2dT_0\varepsilon)) \]
\[ \leq 4d\exp(-(V - \beta/2)/\varepsilon). \] (11)

The event \( \{\tau^\varepsilon(S_r, D_{r_1,r_3}) \leq kT_0\} \) implies that either \( \{\tau^\varepsilon(S_r, D_{r_1,r_3}) = \sigma_n\} \) for some \( 0 \leq n \leq k \) or that at least one of the interval \([\sigma_n, \sigma_{n+1}]\) for \( 0 \leq n < k \) is at most of length \( T_0 \). Combining the estimates (10) and (11), it follows that
\[ \mathbb{P}(\tau^\varepsilon(S_r, D_{r_1,r_3}) \leq kT_0) \leq \sum_{n=0}^{k} \mathbb{P}(\tau^\varepsilon(S_r, D_{r_1,r_3}) = \sigma_n) + \sum_{n=1}^{k} \mathbb{P}(\rho_n - \sigma_{n-1} \leq T_0) \]
\[ \leq \mathbb{P}(\tau^\varepsilon(S_r, D_{r_1,r_3}) = \sigma_0) + (3 + 4d)k\exp(-(V - \beta/2)/\varepsilon) \]

for all \( k \in \mathbb{N} \) and \( \varepsilon \leq \varepsilon_0 := \min\{\varepsilon_1, \varepsilon_2\} \). Choose \( k \) to be \( T_0^{-1}\exp((V - \beta)/\varepsilon) \)
rounded up to integers. Hence,
\[ \mathbb{P}(\tau^\varepsilon(S_r, D_{r_1,r_3}) \leq \exp((V - \beta)/\varepsilon)) \]
\[ \leq \mathbb{P}(\tau^\varepsilon(S_r, D_{r_1,r_3}) \leq kT_0) \]
\[ \leq \mathbb{P}(\tau^\varepsilon(S_r, D_{r_1,r_3}) = \sigma_0) + 8dT_0^{-1}\exp(-\beta/(2\varepsilon)) \]

for small enough \( \varepsilon \). By estimate (9), the right side of the inequality converges to zero as \( \varepsilon \to 0 \). The lower bound for \( \mathbb{E}\tau^\varepsilon(S_r, D_{r_1,r_3}) \) follows by Markov’s inequality. \( \square \)

**Corollary 1.** Set \( V := V(0, 1, \infty) \). For any \( \beta > 0 \) there exists \( \delta_0 > 0 \) such that
\[ \lim_{\varepsilon \to 0} \mathbb{P}\left(\tau^\varepsilon_{1,\delta} > \exp((V - \beta)/\varepsilon)\right) = 1 \]

and
\[ \lim_{\varepsilon \to 0} \varepsilon \log \mathbb{E}\tau^\varepsilon_{1,\delta} \geq V \]

for any \( 0 < \delta < \delta_0 \).

**Proof.** Observe that \( \tau^\varepsilon_{1,\delta} \geq \tau^\varepsilon(S_1, D_{\delta/2,\infty}) \). \( \square \)
3.3. Upper bound for $\tau_{3,\delta}$. In this subsection, we give an upper bound for $\tau_{3,\delta}$ (see (5)). Here, $\tau_{3,\delta}$ is associated to the solution of (1) where the differential equation (1) is additionally assumed to contract strongly. Since $X_t^\alpha$ satisfies the LDP with good rate function (4), it is sufficient to choose a sample path to get a lower estimate on the probability that $\tau_{3,\delta}$ is smaller than some fixed time. The construction of this sample path and the estimate can be found in Lemma 3.3. In Proposition 2 using this probability as the success probability of a geometric distribution, we get the upper bound for $\tau_{3,\delta}$.

Lemma 3.3. Assume that SDE (1) is strongly contracting. For any $\delta > 0$

$$\lim_{T \to \infty} \liminf_{\varepsilon \to 0} \varepsilon \log P(\tau_{3,\delta} \leq T) \geq -V(0,1,\infty).$$

Proof. In this proof we construct a function $g^\alpha : \mathbb{R} \to \mathbb{R}^d$ such that the semi-flow $F(g^\alpha)$ started in $\mathbb{R}^d$ is contained in a small ball of radius $\delta$ at some time $T^\alpha$ and $\lim_{\alpha \to 0} \|T^\alpha - g^\alpha\| \leq V(0,1,\infty)$. An outline of the semi-flow $F(g^\alpha)$ can be found in Figure 2.

Denote by $u'$ the first derivative of $u$. Let $0 < \alpha < 1$ be small enough such that $u'(4) \geq 2\alpha$. Set $c_1^\alpha := \max\{1 - \alpha, \sup\{0 < \alpha < 1 : u'(|x|^2) \leq -\alpha\}\}$ and $c_2^\alpha := \inf\{x > 1 : u'(|x|^2) \geq 2\alpha\} \leq 2$.

We choose $g^\alpha(t) := \int_0^t h^\alpha(s) \, ds, 0, \ldots, 0) \in \mathbb{R}^d$ with

$$h^\alpha(s) := \begin{cases} 0, & \text{for } 0 \leq s \leq T_1^\alpha \text{ or } T_5^\alpha < s \leq T_6^\alpha \\ 3\alpha, & \text{for } T_1^\alpha < s \leq T_2^\alpha \\ 2\nabla \tilde{U}(\varphi(s - T_2^\alpha)), & \text{for } T_2^\alpha < s \leq T_3^\alpha \\ (-2u'(0) + 1)\alpha, & \text{for } T_3^\alpha < s \leq T_4^\alpha \\ \beta^\alpha, & \text{for } T_4^\alpha < s \leq T_5^\alpha \\ 4\alpha c_2^\alpha, & \text{for } T_5^\alpha < s \leq T_7^\alpha \\ \end{cases}$$

for some $\beta^\alpha > 0$, $0 = T_0^\alpha < T_1^\alpha < T_2^\alpha < \cdots < T_7^\alpha < \infty$ determined in the following and where $\varphi$ is the solution of

$$\dot{\varphi}(s) = \nabla \tilde{U}(\varphi(s)) \quad \text{on } \mathbb{R}$$

started in $\varphi(0) = -c_1^\alpha$, where $\tilde{U} : \mathbb{R} \to \mathbb{R}$ with $\tilde{U}(x) := u(x^2)$. Hence,

$$\dot{I}_{T_4^\alpha - T_3^\alpha}(g^\alpha(\cdot + T_2^\alpha)) = 2 \int_{T_3^\alpha}^{T_2^\alpha} \langle \dot{\varphi}(s), \nabla \tilde{U}(\varphi(s)) \rangle \, ds$$

$$= 2(\tilde{U}(\varphi(T_3^\alpha - T_2^\alpha)) - \tilde{U}(\varphi(0)))$$

$$\leq 2(u(0) - u(1)) = V(0,1,\infty).$$

Moreover, $\dot{I}_{T_{j+1}^\alpha - T_j^\alpha}(g^\alpha(\cdot + T_j^\alpha)) = 0$ for $j = 0, 5$ and

$$\dot{I}_{T_{j+1}^\alpha - T_j^\alpha}(g^\alpha(\cdot + T_j^\alpha)) = \int_{T_j^\alpha}^{T_{j+1}^\alpha} |h^\alpha(s)|^2 \, ds \leq (h^\alpha(T_{j+1}^\alpha))^2 (T_{j+1}^\alpha - T_j^\alpha)$$

for $j = 1, 3, 4, 6$. Denote by $F(g) := F_{T_2^\alpha}(g)$ the semi-flow associated to (3). In the following, we construct $\beta^\alpha$ and $\tilde{T}_i^\alpha$ for $i = 1, 2, \ldots, 7$ such that

$$\lim_{\alpha \to 0} (h^\alpha(T_{j+1}^\alpha))^2 (T_{j+1}^\alpha - T_j^\alpha) = 0$$
for $j = 1, 3, 4, 6$ and

$$|F(g^\alpha (T^\alpha_T, x)) - F(g^\alpha (T^\alpha_T, y))| \leq \delta$$

for all $x, y \in \mathbb{R}^d$. Then

$$\lim_{\alpha \to 0} \lim_{\varepsilon \to 0} \inf \varepsilon \log \mathbb{P} \left( \tau_{3, \delta}^\varepsilon \leq T^\alpha_T \right) \geq - \lim_{\alpha \to 0} \int_{T^\alpha_T} (g^\alpha)$$

$$= - \lim_{\alpha \to 0} \sum_{j=0}^7 \int_{T^\alpha_j} - T^\alpha_j (g^\alpha(\cdot + T^\alpha_j))$$

$$\geq -V(0, 1, \infty)$$

by the LDP and the statement follows as soon as we complete the construction.

**Figure 2.** Outline of the semi-flow $F(g^\alpha)$ in $\mathbb{R}^2$ at time $t$

**Step 1:** Since (1) is strongly contracting, we can choose $T^\alpha_1$ such that $|F(g^\alpha (T^\alpha_1, x))| \leq 1 + \alpha$ for all $x \in \mathbb{R}^d$.

Define $Y(t, y) := F(g^\alpha (t + T^\alpha_1))(t, y)$ for $y \in \mathbb{R}^d$ and write $T^\alpha_k := T^\alpha_k - T^\alpha_1$ for $k = 2, \ldots, 7$. Observe that it is sufficient to restrict the analysis to $Y(t, y)$ on the set $B_{1+\alpha}$ since $Y(t, y)$ describes the dynamics of $F(g^\alpha)$ after time $T^\alpha_1$. Denote by $\Pi_k$ the projection on the $k$-th component in $\mathbb{R}^d$.

In the steps 2 to 4, we concentrate on the movement of the point $y_1 := (-1 - \alpha, 0, \ldots, 0) \in \mathbb{R}^d$, choose $t^\alpha_2$, $t^\alpha_3$ and $t^\alpha_4$ and show that $\Pi_1(\alpha) > 0$. We extend this behavior to the set $B_{1+\alpha}$ in step 5 by choosing $t^\alpha_2$ and $t^\alpha_4$ suitably and showing that $\Pi_1(\alpha)$ for all $y \in B_{1+\alpha}$. In the steps 6 and 7, we choose $t^\alpha_6$ and $t^\alpha_7$ and show the contraction.

**Step 2:** Set $y_1 := (-1 - \alpha, 0, \ldots, 0) \in \mathbb{R}^d$. Observe that $\Pi_1(F(t \mapsto 3\alpha t)(2, y_1)) \geq -c\alpha^2$. Choose $t^\alpha_2 = 2$.

**Step 3:** The function $\varphi$ as defined above describes the movement of $Y(\cdot + t^\alpha_2, y_2)$ started in $y_2 := (-c\alpha, 0, \ldots, 0) \in \mathbb{R}^d$. Choose $t^\alpha_3$ such that $\varphi(t^\alpha_3 - t^\alpha_2) \geq -\alpha$. Then, $\Pi_1(\alpha)$.

**Step 4:** Let $y_3 := (-\alpha, 0, \ldots, 0) \in \mathbb{R}^d$. Observe that $\Pi_1(F(t \mapsto (-2u(0) + 1\alpha t)(2, y_3)) \geq \alpha$. Choose $t^\alpha_7 = t^\alpha_3 + 2$. Then, $\Pi_1(\alpha)$.

**Step 5:** Since $y \mapsto Y(t^\alpha_7, y)$ is continuous, there exists a neighborhood of $y_1$ such that $\Pi_1(Y(t^\alpha_7, y)) > 0$ for all $y$ in this neighborhood. Hence, there exists an $0 < \eta^\alpha < 1$ such that $\Pi_1(Y(t^\alpha_7, y)) < 0$ for some $y \in S_{1+\alpha}$ implies that $|y - \Pi_1(y)| \geq \eta^\alpha$. Observe that

$$\frac{d}{dt} \frac{\Pi_1(F(g^\alpha)(t, y))}{\sqrt{\sum_{k=2}^d (\Pi_k(F(g^\alpha)(t, y)))^2}} = \frac{1}{\sqrt{\sum_{k=2}^d (\Pi_k(F(g^\alpha)(t, y)))^2}} g^\alpha(t)$$
Proposition 2. Assume that SDE (1) is strongly contracting. Set \( V := V(0, 1, \infty) \). Then, for any \( \delta > 0 \) and \( \beta > 0 \) it holds that

\[
\lim_{\varepsilon \to 0} \mathbb{P} \left( \tau_{3,\delta} < \exp((V + \beta)/\varepsilon) \right) = 1
\]

and

\[
\lim_{\varepsilon \to 0} \varepsilon \log \mathbb{E} \tau_{3,\delta} \leq V.
\]

Proof. Let \( 0 < \eta < \beta/2 \). By Lemma 3.3 there exists \( \varepsilon_0 > 0 \) and \( T > 0 \) such that

\[
\mathbb{P}(\tau_{3,\delta} \leq T) \geq \exp((-V - \eta)/\varepsilon).
\]

for all \( \varepsilon \leq \varepsilon_0 \). Conditioning on the event \( \{\tau_{3,\delta} > (k - 1)T\} \) for \( k = 2, 3, \ldots \) yields

\[
\mathbb{P}(\tau_{3,\delta} > kT) = \mathbb{P}(\tau_{3,\delta} > kT|\tau_{3,\delta} > (k - 1)T) \mathbb{P}(\tau_{3,\delta} > (k - 1)T)
\]

\[
\leq \mathbb{P}(\tau_{3,\delta} > T) \mathbb{P}(\tau_{3,\delta} > (k - 1)T)
\]

\[
\leq \mathbb{P}(\tau_{3,\delta} > T)^k
\]
Therefore,
\[
\mathbb{E} \tau^\varepsilon_{3,\delta} \leq T \left( 1 + \sum_{k=1}^{\infty} \mathbb{P} \left( \tau^\varepsilon_{3,\delta} > kT \right) \right) \leq T \left( 1 + \sum_{k=1}^{\infty} \left( 1 - \exp((-V - \eta) / \varepsilon) \right)^k \right)
\]
for all \( \varepsilon \leq \varepsilon_0 \). Using Markov’s inequality it follows that
\[
\mathbb{P} \left( \tau^\varepsilon_{3,\delta} \geq \exp((V + \beta) / \varepsilon) \right) \leq T \exp(-\beta / (2\varepsilon))
\]
for all \( \varepsilon \leq \varepsilon_0 \).

**Remark 1.** We require strong contraction mainly to obtain Proposition 2. Without assuming strong contraction, we can get an equivalent statement to Lemma 3.3 for \( \tau^\varepsilon_{4,\delta,R} \). Starting in the set \( B_R \) for large enough \( R > 0 \), one could try to get a similar result as in Proposition 2 for \( \tau^\varepsilon_{4,\delta,R} \) using \( \mathbb{P}(\tau^\varepsilon_{4,\delta,R} < T) \) as success probability for some \( \hat{R} > 0 \). However, if a trial fails one cannot directly start a new trial since one does not know whether the process is contained in \( B_R \) at time \( T \). The probability to leave \( B_R \) is close to 0 for small noise on a fixed time interval. Nonetheless, this term causes the proof of Proposition 2 to fail since the analyzed time interval grows for \( \varepsilon \to 0 \) and the estimate of the expected value requires to sum up the probability infinitely often.

**Remark 2.** Observe that the upper bound for \( \tau^\varepsilon_{3,\delta} \) as in Proposition 2 holds even for some RDS that does not synchronize. In [16], an example of an SDE is presented which does not synchronize for small noise. The drift of this SDE is of the same form as in SDE (1) while the noise merely acts in the first component. Hence, the arguments in Lemma 3.3 and Proposition 2 extend to this SDE since \( g^\alpha \) in Lemma 3.3 is chosen to be 0 in all components except for the first one.

### 3.4. Approaching the set attractor

Combining the estimates from the previous subsections, we get a lower and an upper meaningful bounds. These bounds show that the time a set requires to approach the attractor is roughly \( \exp(V(0,1,\infty) / \varepsilon) \).

**Theorem 3.4.** Assume that SDE (1) is strongly contracting. Set \( V := V(0,1,\infty) \) and let \( S_1 \subset M \subset \mathbb{R}^d \). For any \( \beta > 0 \) there exists \( \delta_0 > 0 \) such that for all \( 0 < \delta \leq \delta_0 \) it holds that
\[
\lim_{\varepsilon \to 0} \mathbb{P} \left( \exp((V - \beta) / \varepsilon) < \tau^\varepsilon_{1,2\delta} \leq \tau^\varepsilon_{2,\delta,M} \leq \tau^\varepsilon_{3,\delta} < \exp((V + \beta) / \varepsilon) \right) = 1
\]
and
\[
\lim_{\varepsilon \to 0} \varepsilon \log \mathbb{E} \tau^\varepsilon_{1,2\delta} = \lim_{\varepsilon \to 0} \varepsilon \log \mathbb{E} \tau^\varepsilon_{2,\delta,M} = \lim_{\varepsilon \to 0} \varepsilon \log \mathbb{E} \tau^\varepsilon_{3,\delta} = V.
\]

**Proof.** Using \( \tau^\varepsilon_{1,2\delta} \leq \tau^\varepsilon_{2,\delta,M} \leq \tau^\varepsilon_{3,\delta} \), the statement follows by Corollary 1 and Proposition 2. \( \Box \)
4. Time required for a point to approach the attractor.

4.1. Convergence to a process on the unit sphere. In this section, we show that the time required for a point to approach the attractor under the dynamics of (1) in dimension \(d = 2\) is exactly of order \(\varepsilon^{-1}\). In particular, we give an estimate on the rate of convergence of a point under the dynamics of (1) towards the attractor.

Here, we consider the minimal weak point attractor \(A^{X,\varepsilon}_{\text{point}}\) since this attractor is associated with weak synchronization. A minimal weak point attractor is a weak point attractor that is contained in any other weak point attractor. By [7, Theorem 3.1] and [5, Theorem 23] such a minimal weak point attractor exists. Instead of the minimal weak point attractor one could also use the weak attractor. However, one then needs to assume synchronization instead of weak synchronization to get the second statement of Corollary 3.

We concentrate on the two-dimensional case, \(d = 2\), where the process behaves similarly to a process on a unit sphere which is known to synchronize weakly. We obtain this process on the unit sphere by performing a time change and writing the accelerated process in polar coordinates, more precisely, we consider

\[
(R_t^\varepsilon \cos \phi_t^\varepsilon, R_t^\varepsilon \sin \phi_t^\varepsilon) = X_{t/\varepsilon}.
\]

Using Itô’s formula, we obtain

\[
d(R_t^\varepsilon)^2 = -\frac{4}{\varepsilon} (R_t^\varepsilon)^2 u'(R_t^\varepsilon)^2 \, dt + 2R_t^\varepsilon \cos \phi_t^\varepsilon \, d\tilde{W}_t^1 + 2R_t^\varepsilon \sin \phi_t^\varepsilon \, d\tilde{W}_t^2 + 2 \, dt \tag{12}
\]

where \(u'\) is the first derivative of \(u\) and

\[
d\phi_t^\varepsilon = \frac{1}{R_t^\varepsilon} \left( -\sin \phi_t^\varepsilon \, d\tilde{W}_t^1 + \cos \phi_t^\varepsilon \, d\tilde{W}_t^2 \right) \tag{13}
\]

where \((\tilde{W}_t^1, \tilde{W}_t^2) = \sqrt{\varepsilon} W_{t/\varepsilon}\) and \((\tilde{W}_t^1, \tilde{W}_t^2)\) is a 2-dimensional Brownian motion. As will be shown in the sequel when \(\varepsilon \to 0\), the drift of \(R_t^\varepsilon\) will move the radius close to 1. By substituting in (13) \(R_t^\varepsilon\) by 1, we aim to compare \(\phi_t^\varepsilon\) to the process

\[
dZ_t = -\sin Z_t \, d\tilde{W}_t^1 + \cos Z_t \, d\tilde{W}_t^2 \tag{14}
\]

on the limit cycle \(S = \mathbb{R}/2\pi \mathbb{Z}\). After we show that \(R_t^\varepsilon\) is close to 1 and \(\phi_t^\varepsilon\) is close to \(Z_t\), we will use that the RDS associated to (14) synchronizes weakly, i.e. every point in \(S\) converges to a single random point.

Remark 3. In higher dimensional cases, \(d > 2\), more complicated processes on the sphere need to be analyzed. For example in the three-dimensional case, \(d = 3\), we can write

\[
\left( \hat{R}_t^\varepsilon \cos \hat{\phi}_{1,t}^\varepsilon \sin \hat{\phi}_{2,t}^\varepsilon, \hat{R}_t^\varepsilon \sin \hat{\phi}_{1,t}^\varepsilon \sin \hat{\phi}_{2,t}^\varepsilon, \hat{R}_t^\varepsilon \cos \hat{\phi}_{2,t}^\varepsilon \right) = X_{t/\varepsilon}^\varepsilon.
\]

Then the SDE on the sphere should behave similarly to

\[
d\hat{\phi}_{1,t}^\varepsilon = \frac{1}{\hat{R}_t^\varepsilon} \left( -\sin \hat{\phi}_{1,t}^\varepsilon \, d\tilde{W}_t^1 + \cos \hat{\phi}_{1,t}^\varepsilon \, d\tilde{W}_t^2 \right)
\]

\[
d\hat{\phi}_{2,t}^\varepsilon = \frac{1}{\hat{R}_t^\varepsilon} \left( \cos \hat{\phi}_{1,t}^\varepsilon \cos \hat{\phi}_{2,t}^\varepsilon \, d\tilde{W}_t^1 + \sin \hat{\phi}_{1,t}^\varepsilon \cos \hat{\phi}_{2,t}^\varepsilon \, d\tilde{W}_t^2 + \sin \hat{\phi}_{2,t}^\varepsilon \, d\tilde{W}_t^3 \right)
\]

where \((\tilde{W}_t^1, \tilde{W}_t^2, \tilde{W}_t^3) = \sqrt{\varepsilon} W_{t/\varepsilon}\) and \((\tilde{W}_t^1, \tilde{W}_t^2, \tilde{W}_t^3)\) is a three-dimensional Brownian motion. Even though we do not compute the rates in which a point approaches
the attractor in higher dimensions, we expect these rates to be the same as in the two-dimensional case.

Returning to the two-dimensional case, we show that the radial component of the accelerated process $R^\varepsilon_t$ is close to 1 for $t > 0$ and small noise intensities $\varepsilon$.

**Lemma 4.1.** Let $0 < \alpha < \beta < 1$, $T > 0$ and $0 < r_1 < 1 < r_2 < r_3 < \infty$. Then, there exists an $\varepsilon_0 > 0$ such that

$$\mathbb{P}(r_1 < R^\varepsilon_T < r_2) \geq 1 - \beta$$

for all $\varepsilon \leq \varepsilon_0$ and any $F^\varepsilon$-measurable $X^\varepsilon_0$ satisfying

$$\mathbb{P}(R^\varepsilon_0 \leq r_3) \geq 1 - \alpha.$$  \hspace{1cm} (15)

**Proof.** Choose $k \in \mathbb{N}$ such that $2^{-k+1} \leq \beta - \alpha$ and set $t = \min\{1/2, T/(2k)\}$. Using (12),

$$\int_0^t R^\varepsilon_t \cos \phi^\varepsilon_t \, d\tilde{W}^1_t + \int_0^t R^\varepsilon_t \sin \phi^\varepsilon_t \, d\tilde{W}^2_t > 0$$

implies that $(R^\varepsilon_s)^2 \geq 2t$ for some $s \leq t$. Set $r_4 = \sqrt{2t} \leq 1$. Then

$$\mathbb{P}(R^\varepsilon_s \geq r_4 \text{ for some } s \leq t) \geq 1/2.$$  \hspace{1cm} (16)

Conditioning on the event $\{R^\varepsilon_s < r_4 \text{ for all } s \leq (j-1)t\}$ for $j = 2, 3, \ldots, k$ yields to

$$\mathbb{P}(R^\varepsilon_s < r_4 \text{ for all } s \leq T/2) \leq \mathbb{P}(R^\varepsilon_s < r_4 \text{ for all } s \leq kt) \leq 1/2 \mathbb{P}(R^\varepsilon_s < r_4 \text{ for all } s \leq (k-1)t) \leq 2^{-k} \leq (\beta - \alpha)/2.$$  \hspace{1cm} (17)

Combining this estimate and the assumption (15), it follows that

$$\mathbb{P}(r_4 \leq R^\varepsilon_s \leq r_3 \text{ for some } s \leq T/2) \geq 1 - (\alpha + \beta)/2.$$  \hspace{1cm} (18)

Let $r_1 < r_5 < 1 < r_6 < r_2$. By Lemma 3.2, there exists $C, \varepsilon_1 > 0$ such that

$$\mathbb{P}(r_5 < R^\varepsilon_s < r_6 \text{ for some } s \leq T/2 + \varepsilon C) \geq 1 - (\alpha + 2\beta)/3.$$  \hspace{1cm} (19)

for all $\varepsilon \leq \varepsilon_1$. Hence, for all $\varepsilon \leq \min\{\varepsilon_1, T/(2C)\}$

$$\mathbb{P}(r_5 < R^\varepsilon_s < r_6 \text{ for some } s \leq T) \geq 1 - (\alpha + 2\beta)/3.$$  \hspace{1cm} (20)

Using Proposition 1, the statement follows. \hfill \Box

**Lemma 4.2.** Let $0 < \alpha < \beta < 1$ and $\delta, T > 0$. Then, there exists $\varepsilon_0, \eta > 0$ such that

$$\mathbb{P}\left(\max_{t \leq T} |R^\varepsilon_t - 1| < \delta \text{ and } \max_{t \leq T} |\phi^\varepsilon_t - Z_t| < \delta \right) \geq 1 - \beta$$

for all $\varepsilon \leq \varepsilon_0$ and any $F^\varepsilon$-measurable $X^\varepsilon_0$ and $Z_0$ satisfying

$$\mathbb{P}(|R^\varepsilon_0 - 1| < \eta \text{ and } |\phi^\varepsilon_0 - Z_0| < \eta) \geq 1 - \alpha.$$  \hspace{1cm} (21)

**Proof.** Choose $0 < \eta < \frac{1}{4} \min\{\delta, 1\}$ such that $(4\eta^2 + 128T\eta^2(1 - 2\eta)^{-2}) e^{16T} < (\beta - \alpha)\delta^2$. Define

$$B^\varepsilon_t := \left\{ \max_{s \leq t} |R^\varepsilon_s - 1| < 2\eta \right\} \cap \{ |\phi^\varepsilon_0 - Z_0| < \eta \}.$$
for all \( t \leq T \). Using Proposition 1 and the assumption (17), there exists \( \varepsilon_0 > 0 \) such that
\[
\mathbb{P}(B_T^\varepsilon) \geq 1 - (\alpha + \beta)/2
\]
for all \( \varepsilon \leq \varepsilon_0 \). We use Doob’s inequality and Itô’s isometry to estimate
\[
\mathbb{E} \max_{t \leq T} |\phi_t^\varepsilon - Z_t|^2 \mathbb{1}_{B_T^\varepsilon} \\
\leq 2 \mathbb{E}|\phi_0^\varepsilon - Z_0|^2 \mathbb{1}_{B_T^\varepsilon} + 2 \mathbb{E} \max_{t \leq T} \left( \int_0^t \left( \sin Z_s - \frac{1}{R_s} \sin \phi_s^\varepsilon \right) \mathbb{1}_{B_T^\varepsilon} \ dW_s^1 \\
+ \int_0^t \left( -\cos Z_s + \frac{1}{R_s} \cos \phi_s^\varepsilon \right) \mathbb{1}_{B_T^\varepsilon} \ dW_s^2 \right)^2 \\
\leq 2 \eta^2 + 4 \mathbb{E} \int_0^T \left( \left( \sin Z_t - \frac{1}{R_t} \sin \phi_t^\varepsilon \right)^2 + \left( -\cos Z_t + \frac{1}{R_t} \cos \phi_t^\varepsilon \right)^2 \right) \mathbb{1}_{B_T^\varepsilon} \ dt \\
\leq 2 \eta^2 + 16 \mathbb{E} \int_0^T \left( |Z_t - \phi_t^\varepsilon|^2 + \left| 1 - \frac{1}{R_t} \right|^2 \right) \mathbb{1}_{B_T^\varepsilon} \ dt \\
\leq 2 \eta^2 + 64T \frac{\eta^2}{(1 - 2\eta)^2} + 16 \int_0^T \mathbb{E} \max_{t \leq s} |Z_s - \phi_s^\varepsilon|^2 \mathbb{1}_{B_T^\varepsilon} \ dt.
\]
The third inequality in the previous estimate uses that \( (x - \frac{1}{R}y)^2 \leq 2(x - y)^2 + 2y^2 \left( 1 - \frac{1}{R} \right)^2 \) for any \( x, y \in \mathbb{R} \), \( R > 0 \). Using Gronwall’s inequality, it follows that
\[
\mathbb{E} \max_{t \leq T} |\phi_t^\varepsilon - Z_t|^2 \mathbb{1}_{B_T^\varepsilon} \leq \left( 2 \eta^2 + 64T \frac{\eta^2}{(1 - 2\eta)^2} \right) e^{16T} < (\beta - \alpha)\delta^2/2.
\]
Using Markov’s inequality, we get
\[
\mathbb{P} \left( \max_{t \leq T} |R_t^\varepsilon - 1| < \delta \text{ and } \max_{t \leq T} |\phi_t^\varepsilon - Z_t| < \delta \right) \\
\geq \mathbb{P}(B_T^\varepsilon) - \mathbb{P} \left( B_T^\varepsilon \text{ and } \max_{t \leq T} |\phi_t^\varepsilon - Z_t| \geq \delta \right) \\
\geq 1 - (\alpha + \beta)/2 - \delta^{-2} \mathbb{E} \max_{t \leq T} |\phi_t^\varepsilon - Z_t|^2 \mathbb{1}_{B_T^\varepsilon} \geq 1 - \alpha
\]
for all \( \varepsilon \leq \varepsilon_0 \).

4.2. Asymptotic stability of the process on the unit sphere. The SDE (14) has a random stable point whose Lyapunov exponent is negative, see [2, Example 2] for further details. This random point is the minimal weak point attractor of the RDS associated to (14) which we in the following denote by \( A^Z \). Observe that due to the time change the minimal weak point attractor \( A^Z \) of the RDS associated to (14) at time \( t \) is \( A^Z(\theta_{t/\varepsilon}, \omega) \). When we consider the distance of \( A^Z \) to a point in \( \mathbb{R}^2 \), we identify with \( A^Z \) the point \( (\cos A^Z, \sin A^Z) \) on the unit sphere.

Denote by \( Z_t(Z_0) \) the solution of (14) started in \( Z_0 \). We now show the rate of convergence of \( Z_t(Z_0) \) to \( A^Z \), first for deterministic \( Z_0 \) and then for \( F^- \)-measurable \( Z_0 \).

**Lemma 4.3.** For any \( \alpha > 0 \) and \( 0 < \mu < 1/2 \) there exists \( C > 0 \) such that
\[
\mathbb{P}(\|Z_t(Z_0) - A^Z(\theta_{t/\varepsilon})\| \leq C \epsilon^{-\mu t} \text{ for all } t \geq 0) \geq 1 - \alpha
\]
for all \( Z_0 \in [0, 2\pi) \).
for all \( 0 \leq F \).

Proof. By [2, Example 2], the top Lyapunov exponent of (14) is \(-1/2\). Stable manifold theorem implies that for all \( 0 < \mu < 0.5 \) there exist a measurable \( c(\omega) > 0 \) and a measurable neighborhood \( U(\omega) \) of \( A^Z(\omega) \) such that

\[
|Z_t(x) - A^Z(\theta_t/\omega)| < c(\omega)e^{-\mu t}
\]

for all \( x \in U(\omega) \) and \( t \geq 0 \). Hence, for any \( \alpha > 0 \) there exists some \( \tilde{c}, \delta > 0 \) such that

\[
\mathbb{P} \left( |Z_t(x) - A^Z(\theta_t/\omega)| < \tilde{c}e^{-\mu t} \text{ for all } x \in A^Z(\omega) \delta \text{ and } t \geq 0 \right) \geq 1 - \alpha/2.
\]

Since \( A^Z(\omega) \) is the attractor of the RDS associated to (14), there exists a time \( T > 0 \) such that

\[
\mathbb{P} \left( |Z_T(x) - A^Z(\theta_T/\omega)| < \delta \right) \geq 1 - \alpha/2
\]

for all \( x \in [0, 2\pi) \). Combining these two estimates yields to

\[
\mathbb{P} \left( |Z_t(x) - A^Z(\theta_t/\omega)| < \tilde{c}e^{-\mu(t-T)} \text{ for all } t \geq T \right) \geq 1 - \alpha
\]

for all \( x \in [0, 2\pi) \) and \( t \geq 0 \). \( \square \)

Proposition 3. For any \( \alpha > 0 \) and \( 0 < \mu < 0.5 \) there exists \( C > 0 \) such that

\[
\mathbb{P} \left( |Z_t - A^Z(\theta_t/\epsilon)| \leq C e^{-\mu t} \text{ for all } t \geq 0 \right) \geq 1 - \alpha
\]

for all \( \mathcal{F}^- \)-measurable \( Z_0 \).

Proof. The weak point attractor \( A^Z(\omega) \) is an \( \mathcal{F}^- \)-measurable stable point. Reverting the time, one obtains an \( \mathcal{F}^+ \)-measurable unstable point \( U^Z(\omega) \). Hence, \( U^Z(\omega) \) and \( A^Z(\omega) \) are independent. Under the dynamics of (14) every single deterministic point converges to the attractor. However, the unstable (random) point does not converge to the attractor. Observe that \( A^Z \) and \( U^Z \) are uniformly distributed on the limit cycle \( S \) since \( Z_t \) is rotational invariant.

If the unstable point is in an interval and the attractor is not, then the time the endpoints of this interval require to approach the attractor is an upper bound for the time any point outside the interval requires to approach the attractor.

Choose \( n \in \mathbb{N} \) such that \( \alpha n \geq 4 \). We define

\[
I_k := \left[ \frac{2\pi}{n}, \frac{(k+1)2\pi}{n} \right), \quad P_k = \frac{2\pi}{n} \text{ and } P_n = P_0
\]

for \( 0 \leq k < n \). By Lemma 4.3 there exists \( C > 0 \) such that

\[
\mathbb{P} \left( |Z_t(P_k) - A^Z(\theta_t/\epsilon)| > C e^{-\mu t} \text{ for some } t \geq 0 \right) \leq \frac{\alpha}{4n}
\]

for all \( 0 \leq k \leq n \). If \( U^Z(\omega) \in I_k \) and \( A(\omega) \notin I_k \) for some \( 0 \leq k < n \), then

\[
\sup_{z \notin I_k} |Z_t(z) - A^Z(\theta_t/\omega)| = \min \{ |Z_t(P_k) - A^Z(\theta_t/\omega)|, |Z_t(P_{k+1}) - A^Z(\theta_t/\omega)| \}
\]
for all $t \geq 0$. Therefore,
\[
P \left( \left| Z_t(Z_0) - A^Z(\theta_t/\epsilon) \right| \leq C e^{-\mu t} \text{ for all } t \geq 0 \right)
\geq \sum_{k=0}^{n-1} \mathbb{P}(U^Z(\cdot) \in I_k, A^Z(\cdot) \notin I_k, Z_0 \notin I_k, |Z_t(P_k) - A^Z(\theta_t/\epsilon)| \leq C e^{-\mu t}
\quad \text{and } |Z_t(P_{k+1}) - A^Z(\theta_t/\epsilon)| \leq C e^{-\mu t} \text{ for all } t \geq 0)
\geq \sum_{k=0}^{n-1} \mathbb{P}(U^Z(\cdot) \in I_k) \mathbb{P}(A^Z(\cdot) \notin I_k, Z_0 \notin I_k)
\quad - \mathbb{P}(|Z_t(P_k) - A^Z(\theta_t/\epsilon)| > C e^{-\mu t} \text{ for some } t \geq 0)
\quad - \mathbb{P}(|Z_t(P_{k+1}) - A^Z(\theta_t/\epsilon)| > C e^{-\mu t} \text{ for some } t \geq 0)
\geq \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{P}(A^Z(\cdot) \notin I_k, Z_0 \notin I_k) - \alpha/2
\geq \frac{n-1}{n} - \alpha/2 \geq 1 - \alpha
\]
for all $F^\sigma$-measurable $Z_0$. \hfill $\square$

4.3. Approaching the point attractor. Combining the estimates from the previous subsections, we are able to show the rate of convergence of $X^\varepsilon_{t/\varepsilon}$ to $A^Z$. As a direct consequence, we get that $A^Z$ and $A^X_{\text{point}_t}$ are close for small $\varepsilon$ and the upper bound for the rate of convergence of $X^\varepsilon_{t/\varepsilon}$ to $A^X_{\text{point}_t}$. Moreover, we show that $X^\varepsilon_t$ does not approach its attractor on a faster time scale.

**Proposition 4.** Let $0 < \alpha < \beta < 1$, $r > 0$ and $0 < \mu < 0.5$. Then, there exists $C > 0$ such that for all $T_1, T_2 > 0$ there exists an $\varepsilon_0 > 0$ such that
\[
P \left( \left| X^\varepsilon_{t/\varepsilon} - A^Z(\theta_t/\varepsilon) \right| \leq C e^{-\mu(t-T_2)} \text{ for all } T_2 \leq t \leq T_1 + T_3 \right) \geq 1 - \beta
\]
for all $0 < \varepsilon \leq \varepsilon_0$, $T_2 \geq T_1$ and all $F^\sigma$-measurable $X^\varepsilon_0$ satisfying
\[
P(R^\varepsilon_0 \leq r) \geq 1 - \alpha.
\]
**Proof.** Let $\varepsilon > 0$. We start SDE (14) in $Z^\varepsilon_{T_1} = \phi^\varepsilon_{T_1}$. By Proposition 3 there exists $c > 0$ such that
\[
P \left( \left| Z_t - A^Z(\theta_t/\varepsilon) \right| \leq c e^{-\mu(t-T_1)} \text{ for all } t \geq T_1 \right) \geq 1 - \alpha/2.
\]
for all $\varepsilon > 0$. Using Lemma 4.1 and 4.2, there exists $\varepsilon_0 > 0$ such that
\[
P \left( \max_{t_1 \leq t \leq t_1 + T_3} |R^\varepsilon_t - 1| < e^{-\mu T_3} \text{ and } \max_{t_1 \leq t \leq t_1 + T_3} |\phi^\varepsilon_t - Z^\varepsilon_t| < e^{-\mu T_3} \right) \geq 1 - \alpha/2
\]
for all $\varepsilon \leq \varepsilon_0$. Setting $C := c + 2$ it follows that
\[
P \left( \left| X^\varepsilon_{t/\varepsilon} - A^Z(\theta_t/\varepsilon) \right| \leq C e^{-\mu(t-T_1)} \text{ for all } T_1 \leq t \leq T_1 + T_3 \right) \geq 1 - \alpha.
\]
Using the same arguments for the process starting in $X^\varepsilon_{(T_2-T_1)/\varepsilon}$ at time $(T_2-T_1)/\varepsilon$, the statement follows. \hfill $\square$
Remark 4. Observe that the statement of Proposition 4 is not true for $T_3 = \infty$. Precisely, for all $\delta, \varepsilon, T > 0$
$$
\mathbb{P} \left( \sup_{t \geq T} \left| X^x_t / \varepsilon - A^Z(\theta_{t / \varepsilon}) \right| \leq \delta \right) = 0
$$
since the process $X^x_t$ leaves any neighborhood of the unit sphere for some $t \geq T / \varepsilon$ almost surely.

Corollary 2. For all $\alpha, \delta, T > 0$ there exists an $\varepsilon_0 > 0$ such that
$$
\mathbb{P} \left( \inf_{a \in A^X_{\text{point}}(\theta \cdot)} \left| A^Z(\theta \cdot) - a \right| \leq \delta \text{ for all } 0 \leq t \leq T / \varepsilon \right) \geq 1 - \alpha
$$
for all $0 < \varepsilon \leq \varepsilon_0$.

Proof. By the construction of the minimal weak point attractor in [5, Theorem 23], the minimal weak point attractor of (1) has a $\mathcal{F}^-$-measurable version. We denote this version also by $A^X_{\text{point}}$. Using [4, Theorem III.9], we can select an $\mathcal{F}^-$-measurable $x^\varepsilon(\omega)$ where
$$
x^\varepsilon(\omega) \in \begin{cases} A^X_{\text{point}}(\omega) \cap B_2, & \text{if } A^X_{\text{point}}(\omega) \cap B_2 \neq \emptyset; \\
\mathbb{R}^2, & \text{else.}
\end{cases}
$$
and $B_2$ denotes the ball of radius 2 centered at 0. Since the drift of (1) pushes any point outside the unit ball towards the unit ball, it holds that
$$
\lim_{\varepsilon \to 0} \mathbb{P} \left( A^X_{\text{point}}(\omega) \cap B_2 = \emptyset \right) = 0.
$$
Applying Proposition 4, there exist some $\varepsilon_1, s > 0$ such that
$$
\mathbb{P} \left( \left| X^x_{t / \varepsilon}(x^\varepsilon(\cdot)) - A^Z(\theta_{t / \varepsilon}) \right| \leq \delta \text{ for all } s \leq t \leq s + T \right) \geq 1 - \alpha / 2
$$
for all $\varepsilon \leq \varepsilon_1$. Since $x^\varepsilon(\omega) \in A^X_{\text{point}}(\omega)$ implies that $X^x_t(x^\varepsilon(\omega)) \in A^X_{\text{point}}(\theta_{t \cdot})$, there exists $\varepsilon_2 > 0$ such that
$$
\mathbb{P} \left( \inf_{a \in A^X_{\text{point}}(\theta \cdot)} \left| A^Z(\theta \cdot) - a \right| \leq \delta \text{ for all } s / \varepsilon \leq t \leq (s + T) / \varepsilon \right) \geq 1 - \alpha
$$
for all $\varepsilon \leq \varepsilon_2$. Using $\theta_{s / \varepsilon}$-invariance of $\mathbb{P}$, the statement follows. \qed

Theorem 4.4. Let $0 < \alpha < \beta < 1$, $r > 0$ and $0 < \mu < 0.5$ Then, there exists $C > 0$ such that for all $T_1, T_3 > 0$ there exists an $\varepsilon_0 > 0$ such that
$$
\mathbb{P} \left( \inf_{a \in A^X_{\text{point}}(\theta_{t / \varepsilon})} \left| X^x_t / \varepsilon - a \right| \leq C e^{-\mu(t - T_2)} \text{ for all } T_2 \leq t \leq T_2 + T_3 \right) \geq 1 - \beta
$$
for all $0 < \varepsilon \leq \varepsilon_0$, $T_2 \geq T_1$ and all $\mathcal{F}^-$-measurable $X^x_0$ satisfying
$$
\mathbb{P} (R_0^r \leq r) \geq 1 - \alpha.
$$
Proof. Apply Proposition 4 and Corollary 2 and use the triangle inequality. \qed

Theorem 4.5. For any $\alpha > 0$ there exist $\varepsilon_0, \delta, T > 0$ such that
$$
\mathbb{P} \left( \sup_{a \in A^X_{\text{point}}(\theta \cdot)} \left| X^x_t - a \right| > \delta \text{ for all } 0 \leq t \leq T / \varepsilon \right) \geq 1 - \alpha
$$
for all $\varepsilon \leq \varepsilon_0$ and all deterministic $X^x_0 \in \mathbb{R}^2$.\qed
Proof. Choose $\gamma, T > 0$ such that $20\gamma \leq \alpha \pi$ and $10T \leq \alpha \gamma^2$. By Lemma 4.2 there exists $0 < 2\delta < \sin \gamma$ and

$$\mathbb{P} \left( |\phi_t^\varepsilon - Z_t^\varepsilon| \leq \gamma \text{ for all } \sigma^\delta \leq t \leq T \right) \geq 1 - \alpha/5$$

where $\sigma^\delta := \inf \{ t \geq 0 : |R_t^\varepsilon - 1| \leq 2\delta \}$ and $Z_t^\varepsilon$ is the solution to (14) started in $Z_t^\varepsilon = \phi_t^\varepsilon$. Then,

$$\mathbb{P} \left( |X_{t/\varepsilon} - A^Z(\theta_t/\varepsilon)| > 2\delta \text{ for all } t \leq T \right)$$

$$\geq \mathbb{P} \left( |X_{t/\varepsilon} - A^Z(\theta_t/\varepsilon)| > \sin \gamma \text{ for all } \sigma^\delta \leq t \leq T \right)$$

$$\geq \mathbb{P} \left( |\phi_t^\varepsilon - A^Z(\theta_t/\varepsilon)| > \gamma \text{ for all } \sigma^\delta \leq t \leq T \right)$$

$$\geq \mathbb{P} \left( |\phi_t^\varepsilon - Z_t^\varepsilon| \leq \gamma \text{ and } |Z_t^\varepsilon - A^Z(\theta_t/\varepsilon)| > 2\gamma \text{ for all } \sigma^\delta \leq t \leq T \right)$$

$$\geq \mathbb{P} \left( |Z_t^\varepsilon - A^Z(\theta_t/\varepsilon)| > 2\gamma \text{ for all } \sigma^\delta \leq t \leq T \right) - \alpha/5.$$  \hspace{1cm} (18)

Independence of $A^Z(\cdot)$ and $Z_{\sigma^\delta}^{\varepsilon}$ implies

$$\mathbb{P} \left( |Z_{\sigma^\delta}^{\varepsilon} - A^Z(\cdot)| \leq 4\gamma \right) = \frac{8\gamma}{2\pi} \leq \alpha/5.$$  

Let $\tilde{Z}_t(\tilde{Z}_0)$ be the solution to (14) started in $\tilde{Z}_0$. We use Doob’s inequality and Itô’s isometry to estimate

$$\mathbb{E} \max_{0 \leq t \leq T} |\tilde{Z}_t(\tilde{Z}_0) - \tilde{Z}_0|^2 \leq 2 \left( \int_0^T \sin \tilde{Z}_s(\tilde{Z}_0) \, d\tilde{W}_s^1 + \int_0^T \cos \tilde{Z}_s(\tilde{Z}_0) \, d\tilde{W}_s^2 \right)^2 \leq 2T.$$  

Using the previous two estimates, Markov inequality and $10T \leq \alpha \gamma^2$, it follows that

$$\mathbb{P} \left( |Z_t^\varepsilon - A^Z(\theta_t/\varepsilon)| > 2\gamma \text{ for all } \sigma^\delta \leq t \leq T \right)$$

$$\geq \mathbb{P} \left( |Z_{\sigma^\delta}^{\varepsilon} - A^Z(\cdot)| > 4\gamma, |Z_t^\varepsilon - Z_{\sigma^\delta}^{\varepsilon}| \leq \gamma \text{ and } |A^Z(\cdot) - A^Z(\theta_t/\varepsilon)| \leq \gamma \text{ for all } \sigma^\delta \leq t \leq T \right)$$

$$\geq 1 - \alpha/5 - \mathbb{P} \left( \max_{\sigma^\delta \leq t \leq T} |Z_t^\varepsilon - Z_{\sigma^\delta}^{\varepsilon}| > \gamma \right)$$

$$- \mathbb{P} \left( \max_{\sigma^\delta \leq t \leq T} |A^Z(\cdot) - A^Z(\theta_t/\varepsilon)| > \gamma \right)$$

$$\geq 1 - 3\alpha/5.$$  \hspace{1cm} (19)

Combining the estimates (18) and (19), we obtain

$$\mathbb{P} \left( |X_{t/\varepsilon} - A^Z(\theta_t/\varepsilon)| > 2\delta \text{ for all } t \leq T \right) \geq 1 - 4\alpha/5.$$  

Applying Corollary 2, the statement follows. \hfill \Box

For small $\delta > 0$ denote by

$$\tau_{0,\delta,\varepsilon}^\gamma := \inf \left\{ t \geq 0 : \inf_{a \in A_{\text{point}}^{X^{\varepsilon}}(\theta_{t/\varepsilon})} |X_t^\varepsilon(x) - a| \leq \delta \right\}$$

and

$$\tau_{0,\delta,\varepsilon}^{\sup} := \inf \left\{ t \geq 0 : \sup_{a \in A_{\text{point}}^{X^{\varepsilon}}(\theta_{t/\varepsilon})} |X_t^\varepsilon(x) - a| \leq \delta \right\}$$
the time the process \( X_t^\varepsilon \) started in \( x \in \mathbb{R}^2 \) requires to approach some point respectively all points of the minimal weak point attractor \( A_{\text{point}}^{X^\varepsilon} \). Observe that \( \tau_{0,\delta,x}^\varepsilon \leq \tau_{0,\delta,x}^\varepsilon \). If the RDS associated to (1) synchronize both quantities coincide.

**Corollary 3.** For any \( \alpha > 0 \) there exists some \( \delta_0 > 0 \) such that for all \( 0 < \delta \leq \delta_0 \) there exist \( \varepsilon_0, T_1, T_2 > 0 \) such that
\[
P\left( \tau_{0,\delta,x}^\varepsilon < T_2/\varepsilon \text{ and } \tau_{0,\delta,x}^\varepsilon > T_1/\varepsilon \right) \geq 1 - \alpha
\]
for all \( 0 < \varepsilon \leq \varepsilon_0 \) and \( x \in \mathbb{R}^2 \). In particular, if the RDS associated to (1) synchronize weakly, then
\[
P\left( T_1/\varepsilon < \tau_{0,\delta,x}^\varepsilon = \tau_{0,\delta,x}^\varepsilon < T_2/\varepsilon \right) \geq 1 - \alpha.
\]

**Proof.** The lower bound follows by Theorem 4.5 on \( \tau_{0,\delta,x}^\varepsilon \) and the upper bound on \( \tau_{0,\delta,x}^\varepsilon \) by Theorem 4.4. \( \square \)

**Remark 5.** The dependence of \( T_1 \) and \( T_2 \) on \( \alpha \) in Corollary 3 is passed on to \( X_t^\varepsilon \) by the process \( Z_t \) on the unit sphere. For small \( \alpha \), starting close to the random unstable point cannot be excluded and the process \( Z_t \) may require a lot of time \( T_2 \) to approach its attractor. Similarly, for small \( \alpha \), one may start close to the random stable point, which is the attractor of \( Z_t \), then \( T_1 \) needs to be chosen small.

**Remark 6.** In contrast to Corollary 3, if \( u \) has more than one local minima the time until a point approaches the attractor under the dynamics of (1) can increase exponentially in \( \varepsilon^{-1} \). For this purpose, observe that one can find a lower bound for the time until the paths of the solution started in different minima approach each other using the difference of the potential \( U \) in the minima and similar arguments as in Section 3.2.

Hence, in the case of \( u \) having multiple minima, the difference between the time a point and a set requires to approach the attractor is not as significant as in the case where \( u \) has exactly one minimum.

**Acknowledgment.** The author would like to thank Michael Scheutzow and Anthony Quas for drawing her attention to this problem.

**REFERENCES**

[1] L. Arnold, *Random Dynamical Systems*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 1998.

[2] P. H. Baxendale, Asymptotic behaviour of stochastic flows of diffeomorphisms, in *Stochastic Processes and Their Applications* (Nagoya, 1985), Lecture Notes in Math., Vol. 1203, Springer, Berlin, 1986, 1–19.

[3] T. Caraballo, I. D. Chueshov and P. E. Kloeden, Synchronization of a stochastic reaction-diffusion system on a thin two-layer domain, *SIAM J. Math. Anal.*, 38 (2006/07), 1489–1507.

[4] C. Castaing and M. Valadier, *Convex Analysis and Measurable Multifunctions*, Lecture Notes in Math., Vol. 580, Springer-Verlag, Berlin-New York, 1977.

[5] H. Crauel and M. Scheutzow, Minimal random attractors, *J. Differential Equations*, 265 (2018), 702–718.

[6] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd edition, Applications of Mathematics, Vol. 38, Springer-Verlag, New York, 1998.

[7] G. Dimitroff and M. Scheutzow, Attractors and expansion for Brownian flows, *Electron. J. Probab.*, 16 (2011), 1193–1213.

[8] F. Flandoli, B. Gess and M. Scheutzow, Synchronization by noise, *Probab. Theory Related Fields*, 168 (2017), 511–556.
F. Flandoli, B. Gess and M. Scheutzow, *Synchronization by noise for order-preserving random dynamical systems*, Ann. Probab., 45 (2017), 1325–1350.

M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, Grundlehren der Mathematischen Wissenschaften, Vol. 260, Springer-Verlag, New York, 1984.

B. Gess, *Random attractors for singular stochastic evolution equations*, J. Differential Equations, 255 (2013), 524–559.

A. J. Homburg, *Synchronization in minimal iterated function systems on compact manifolds*, Bull. Braz. Math. Soc. (N.S.), 49 (2018), 615–635.

I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, 2nd edition, Graduate Texts in Mathematics, Vol. 113, Springer-Verlag, New York, 1991.

F. Martinelli and E. Scoppola, *Small random perturbations of dynamical systems: Exponential loss of memory of the initial condition*, Comm. Math. Phys., 120 (1988), 25–69.

O. M. Tearne, *Collapse of attractors for ODEs under small random perturbations*, Probab. Theory Related Fields, 141 (2008), 1–18.

I. Vorkastner, *Noise dependent synchronization of a degenerate SDE*, Stoch. Dyn., 18 (2018), 1850007, 21pp.

Received August 2019; revised November 2019.

E-mail address: vorkastn@math.tu-berlin.de