

# SYNCHRONIZATION BY NOISE

FRANCO FLANDOLI, BENJAMIN GESS, AND MICHAEL SCHEUTZOW

ABSTRACT. We provide sufficient conditions for synchronization by noise, i.e. under these conditions we prove that weak random attractors for random dynamical systems consist of single random points. In the case of SDE with additive noise, these conditions are also essentially necessary. In addition, we provide sufficient conditions for the existence of a minimal weak point random attractor consisting of a single random point. As a result, synchronization by noise is proven for a large class of SDE with additive noise. In particular, we prove that the random attractor for an SDE with drift given by a (multidimensional) double-well potential and additive noise consists of a single random point. All examples treated in [46] are also included.

## 1. INTRODUCTION

In this paper we introduce new, checkable conditions for synchronization by noise for general white noise, random dynamical systems (RDS)  $\varphi$  on complete, separable metric spaces  $E$ . Here, synchronization by noise means that there is a (weak) random attractor<sup>1</sup>  $A$  for  $\varphi$  consisting of a single random point, i.e.  $A(\omega) = \{a(\omega)\}$  a.s. and thus the long-time dynamics are asymptotically globally stable. In particular, for each  $x, y \in E$  it follows that

$$\lim_{t \rightarrow \infty} d(\varphi_t(\omega, x), \varphi_t(\omega, y)) = 0$$

in probability.

We are especially interested in SDE with additive noise

$$dX_t = b(X_t)dt + \sigma dW_t \quad \text{on } \mathbb{R}^d \tag{1.1}$$

---

*Date:* April 19, 2016.

*2010 Mathematics Subject Classification.* 37B25; 37G35, 37H15.

*Key words and phrases.* synchronization, random dynamical system, random attractor, Lyapunov exponent, stochastic differential equation, statistical equilibrium.

B.G. has been partially supported by the research project “Random dynamical systems and regularization by noise for stochastic partial differential equations” funded by the German Research Foundation.

<sup>1</sup>For notation and background on RDS see Section 1.2 below.

with  $\sigma > 0$ , for choices of  $b$  such that the deterministic dynamics corresponding to  $\sigma = 0$  are *not* asymptotically globally stable. We provide general conditions on the coefficients  $b, \sigma$  that lead to synchronization by noise. Hence, in these cases the inclusion of additive noise in (1.1) stabilizes the long-time dynamics.

As a model example, one may consider the multidimensional double-well potential with additive noise, that is

$$dX_t = (X_t - |X_t|^2 X_t) dt + \sigma dW_t \quad \text{on } \mathbb{R}^d. \quad (1.2)$$

In this case, for  $\sigma = 0$  the long-time dynamics are not asymptotically globally stable, but the attractor is given by the closed unit ball  $\bar{B}(0, 1)$ . We shall also analyze the associated point attractor, which consists of all invariant points, i.e.  $S^{d-1} \cup \{0\}$ , where  $S^{d-1}$  is the  $(d-1)$ -dimensional unit sphere. It follows from the general conditions developed in this paper, that for  $\sigma > 0$  synchronization occurs, that is, the random attractor collapses into a single (random) point.

The paper is split into two main parts. In the first part, Section 2, we provide general and new sufficient conditions for synchronization by noise for RDS on separable metric spaces. In the second part, Section 3, these general conditions are verified for classes of SDE of the type (1.1), thus proving synchronization by noise for SDE with additive noise on  $\mathbb{R}^d$ .

The sufficient conditions for synchronization by noise developed in Section 2.1 are essentially sharp in the case of SDE driven by additive noise, i.e. sufficient and necessary. On the other hand, their verification in applications may rely on stronger but easily checkable assumptions. In particular, we verify the general conditions under an *eventual monotonicity* condition on the drift  $b$ . In particular, this yields synchronization for (1.2). However, this eventual monotonicity condition is not necessary for synchronization. This issue is resolved, in a second step, in Section 2.2 by concentrating on a weaker concept of synchronization, so-called weak synchronization. Weak synchronization means that there is a minimal weak *point* attractor  $A$  consisting of a single random point. Our results on weak synchronization are particularly complete in the case of gradient-type SDE, i.e. for

$$dX_t = -\nabla V(X_t) dt + \sigma dW_t \quad \text{on } \mathbb{R}^d, \quad (1.3)$$

with  $V \in C^2(\mathbb{R}^d, \mathbb{R})$ ,  $\sigma > 0$  and  $b := -\nabla V$  satisfying a one-sided Lipschitz condition (among other assumptions). In particular, no eventual monotonicity condition has to be assumed in this case.

In fact, the concept of weak synchronization turns out to be of independent interest with intriguing relations to strong mixing properties of the associated Markovian semigroup (cf. in particular Proposition 2.20 below).

The proof of weak synchronization is based on an analysis of the support properties of the statistical equilibrium, which leads us to (partial) generalizations of results developed in [31, 32].

The main results for RDS on separable metric spaces are given in Theorem 2.14 concerning synchronization and Theorem 2.23 concerning weak synchronization by noise. The main result on weak synchronization for gradient-type SDE is given in Theorem 3.12. The results on general classes of SDE of the type (1.1) are given in Section 3.4.

**1.1. Comments on the existing literature.** There are several distinct approaches to synchronization by noise to be found in the literature. We distinguish three main types of arguments (without aiming for completeness here): Order-preserving RDS, local stability and transitivity of the two point motion, perturbation techniques based on large deviation results.

Synchronization by noise for order-preserving, strongly mixing RDS  $\varphi$  has been analyzed, for example, in [2, 6, 7, 12, 13, 22, 23] and rather general results on (weak) synchronization have been obtained. However, assuming  $\varphi$  to be order-preserving is a significant restriction, leading to stringent assumptions on the drift  $b$  for (1.1) in dimensions larger than one (cf. [12]). In particular, our model example (1.2) is covered for  $d = 1$  only.

In [5], Baxendale proves synchronization for SDE on compact manifolds, assuming ergodicity, local stability, in the sense that the top Lyapunov exponent is supposed to be negative, and assuming transitivity of the two point motion (condition (4.1) in [5]). As compared to Baxendale's work, we focus on the case of RDS on not necessarily compact separable metric spaces  $E$ . In particular, it is one of the aims of this paper to provide conditions for synchronization by noise that are easily checkable for SDE with additive noise on  $\mathbb{R}^d$ . The resulting conditions are rather different and not easily comparable to those developed in [5], which are quite specific to the compact case.

Another approach, based on large deviation techniques, has been introduced in [34, 35, 46]. Besides several technical assumptions, assuming for (1.1) that  $b$  has only finitely many fixed points and that  $\sigma$  is small enough, these works prove synchronization by noise. Again, we note that the model example (1.2) is covered for  $d = 1$  only. In contrast, all examples treated in [46] are easily seen to be included in our results.

Synchronization by linear multiplicative noise has been investigated in [3, 8]. For the related effect of synchronization in master-slave systems we refer to [14] and the references therein. Synchronization for discrete time random dynamical systems

(iterated function systems) has also been investigated and the recent results are deep and advanced, see [25–27, 37] and references therein.

Synchronization has been advocated as a relevant property for certain applications. From the theoretical physics literature let us mention [27, 39, 40, 42]. In climate dynamics it has been mentioned as an indication of the possibility to reduce variability of predictions, see [10, 20, 24]. In neurophysiology, synchronous firing of neurons subject to the same input, which may be seen as a dynamical system driven by the same noise path but different initial conditions, is a phenomenon of interest, see [45] and the references therein. Finally, synchronization plays a role in Richardson-Romberg extrapolation numerical method, see [33].

**1.2. Preliminaries and notation.** Let  $(E, d)$  be a complete separable metric space with Borel  $\sigma$ -algebra  $\mathcal{E}$  and  $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$  be an *ergodic metric dynamical system*, i.e.  $(\Omega, \mathcal{F}, \mathbb{P})$  is a (not necessarily complete) probability space and  $\theta := (\theta_t)_{t \in \mathbb{R}}$  is a group of jointly measurable maps on  $(\Omega, \mathcal{F}, \mathbb{P})$  with ergodic invariant measure  $\mathbb{P}$ .

Further, let  $\varphi : \mathbb{R}_+ \times \Omega \times E \rightarrow E$  be a *perfect cocycle*: i.e.  $\varphi$  is measurable,  $\varphi_0(\omega, x) = x$  and  $\varphi_{t+s}(\omega, x) = \varphi_t(\theta_s \omega, \varphi_s(\omega, x))$  for all  $x \in E$ ,  $t, s \geq 0$ ,  $\omega \in \Omega$ . We will assume that  $\varphi_s(\omega, \cdot)$  is continuous for each  $s \geq 0$  and  $\omega \in \Omega$ . The collection  $(\Omega, \mathcal{F}, \mathbb{P}, \theta, \varphi)$  is then called a *random dynamical system* (in short: RDS), see [1] for a comprehensive treatment.

By definition,  $(\Omega, \mathcal{F}, \mathbb{P}, \theta, \varphi)$  is a *local* RDS if  $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$  is as above and  $\varphi : \mathbb{R}_+ \times \Omega \times \bar{E} \rightarrow \bar{E}$  is measurable, where  $\bar{E} := E \cup \{\partial\}$  and  $\partial$  is some adjoined state with the following properties:  $D := \varphi^{-1}(E) \subseteq \mathbb{R}_+ \times \Omega \times E$  and for each  $\omega \in \Omega$  the set  $D(\omega) := \{(t, x) \in \mathbb{R}_+ \times E : (t, \omega, x) \in D\}$  is open,  $(t, x) \in D(\omega)$  and  $0 \leq s \leq t$  imply  $(s, x) \in D(\omega)$ ,  $x \mapsto \varphi_t(\omega, x)$  is continuous at  $x_0 \in E$  whenever  $\varphi_t(\omega, x_0) \in E$ ,  $\varphi_0(\omega, \cdot) = \text{Id}$  and  $\varphi$  has the perfect cocycle property (as above). Note that a local RDS is an RDS iff  $D = \mathbb{R}_+ \times \Omega \times E$ . Given a (local) RDS  $(\Omega, \mathcal{F}, \mathbb{P}, \theta, \varphi)$  we may define the *skew-product* flow  $\Theta$  on  $\Omega \times \bar{E}$  by  $\Theta_t(\omega, x) = (\theta_t \omega, \varphi_t(\omega, x))$ . In the following we will often omit the qualifier *local*. We say that a local RDS is *weakly complete* if  $\varphi_t(\cdot, x) \in E$ ,  $\mathbb{P}$ -a.s. for all  $t \geq 0, x \in E$ .

Since our main applications are RDS generated by SDE driven by Brownian motion, we will assume that the RDS  $\varphi$  is suitably adapted to a filtration and is of white noise type. More precisely, we will assume that we have a family  $\mathbb{F} = (\mathcal{F}_{s,t})_{-\infty < s \leq t < \infty}$  of sub- $\sigma$  algebras of  $\mathcal{F}$  such that  $\mathcal{F}_{t,u} \subseteq \mathcal{F}_{s,v}$  whenever  $s \leq t \leq u \leq v$ ,  $\theta_r^{-1}(\mathcal{F}_{s,t}) = \mathcal{F}_{s+r, t+r}$  for all  $r, s, t$  and  $\mathcal{F}_{s,t}$  and  $\mathcal{F}_{u,v}$  are independent whenever  $s \leq t \leq u \leq v$ . For each  $t \in \mathbb{R}$ , let us denote the smallest  $\sigma$ -algebra containing all  $\mathcal{F}_{s,t}$ ,  $s \leq t$  by  $\mathcal{F}_t$  and the smallest  $\sigma$ -algebra containing all  $\mathcal{F}_{t,u}$ ,  $t \leq u$  by  $\mathcal{F}_{t,\infty}$ . Note that for each  $t \in \mathbb{R}$ , the  $\sigma$ -algebras  $\mathcal{F}_t$  and  $\mathcal{F}_{t,\infty}$  are independent. We

will further assume that  $\varphi_s(\cdot, x)$  is  $\mathcal{F}_{0,s}$ -measurable for each  $s \geq 0$ . The collection  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}, \theta, \varphi)$  is then called a *white noise (filtered) random dynamical system*.

An invariant measure for an RDS  $\varphi$  is a probability measure on  $\Omega \times E$  with marginal  $\mathbb{P}$  on  $\Omega$  that is invariant under  $\Theta_t$  for  $t \geq 0$ . For each probability measure  $\mu$  on  $\Omega \times E$  with marginal  $\mathbb{P}$  on  $\Omega$  there is a unique disintegration  $\omega \mapsto \mu_\omega$  and  $\mu$  is an invariant measure for  $\varphi$  iff  $\varphi_t(\omega)\mu_\omega = \mu_{\theta_t\omega}$  for all  $t \geq 0$ , almost all  $\omega \in \Omega$ . Here  $\varphi_t(\omega)\mu_\omega$  denotes the push-forward of  $\mu_\omega$  under  $\varphi_t(\omega)$ . An invariant measure  $\mu_\omega$  is said to be a Markov measure, if  $\omega \mapsto \mu_\omega$  is measurable with respect to the past  $\mathcal{F}_0$ . In case of a weakly complete, white noise RDS  $\varphi$  we may define the associated Markovian semigroup by

$$P_t f(x) := \mathbb{E} f(\varphi_t(\cdot, x)),$$

for  $f$  being measurable, bounded. There is a one-to-one correspondence between invariant measures for  $P_t$  and Markov invariant measures for  $\varphi$ : If  $\rho$  is  $P_t$ -invariant, then for every sequence  $t_k \rightarrow \infty$  the weak\* limit

$$\mu_\omega := \lim_{k \rightarrow \infty} \varphi_{t_k}(\theta_{-t_k}\omega)\rho \tag{1.4}$$

exists  $\mathbb{P}$ -a.s. The weak\* limit  $\mu_\omega$  does not depend on the sequence  $t_k$ ,  $\mathbb{P}$ -a.s. and defines a Markov invariant measure for  $\varphi$ . Vice versa,  $\rho := \mathbb{E}\mu_\omega$  defines an invariant measure for  $P_t$ . Note that the proof of these facts given in [15] applies without change to local RDS (cf. also [11] and [30]).

We say that a Markovian semigroup  $P_t$  with invariant measure  $\rho$  is *strongly mixing* if

$$P_t f(x) \rightarrow \int_E f(y) d\rho(y) \quad \text{for } t \rightarrow \infty$$

for each continuous, bounded  $f$  and all  $x \in E$ . Similarly, we say that an RDS  $\varphi$  is strongly mixing if the law of  $\varphi_t(\cdot, x)$  converges to  $\rho$  for  $t \rightarrow \infty$  for all  $x \in E$ .

As a notational convention, we let

$$B(x, r) := \{y \in E : d(x, y) < r\}$$

be the open ball of radius  $r$  centered at  $x$  and  $\bar{B}(x, r)$  the respective closed ball. For a set  $A \subseteq E$  we let

$$A^\varepsilon := \{y \in E : d(y, A) = \inf_{a \in A} d(y, a) < \varepsilon\}$$

and

$$\text{diam}(A) := \sup_{a, b \in A} d(a, b).$$

**Definition 1.1.** A family  $\{D(\omega)\}_{\omega \in \Omega}$  of non-empty subsets of  $E$  is said to be

- (1) a random closed (resp. compact) set if it is  $\mathbb{P}$ -a.s. closed (resp. compact) and  $\omega \mapsto d(x, D(\omega))$  is  $\mathcal{F}$ -measurable for each  $x \in E$ . In this case we also call  $D$ ,  $\mathcal{F}$ -measurable.

(2)  $\varphi$ -invariant, if for all  $t \geq 0$

$$\varphi_t(\omega, D(\omega)) = D(\theta_t\omega),$$

for almost all  $\omega \in \Omega$ .

Next, we recall the definition of a pullback attractor and a weak (random) attractor (cf. [19, 38]).

**Definition 1.2.** Let  $(\Omega, \mathcal{F}, \mathbb{P}, \theta, \varphi)$  be an RDS. A random, compact set  $A$  is called a *pullback attractor*, if

- (1)  $A$  is  $\varphi$ -invariant, and
- (2) for every compact set  $B$  in  $E$ , we have

$$\limsup_{t \rightarrow \infty} \sup_{x \in B} d(\varphi_t(\theta_{-t}\omega, x), A(\omega)) = 0, \text{ almost surely.}$$

The map  $A$  is called a *weak attractor*, if it satisfies the properties above with almost sure convergence replaced by convergence in probability in (2). It is called a *(weak) point attractor*, if it satisfies the properties above with compact sets  $B$  replaced by single points in (2).

A (weak) point attractor is said to be *minimal* if it is contained in each (weak) point attractor.

Clearly, every pullback attractor is a weak attractor but the converse is not true (see e.g. [44] for examples).

**Lemma 1.3.** Weak attractors (and hence pullback attractors) are unique in the sense that if an RDS has two weak attractors, then they agree almost surely.

*Proof.* Let  $A, \tilde{A}$  be two weak random attractors. Since  $\tilde{A}$  is a random compact set, by [17, Proposition 3.15] for each  $\varepsilon > 0$  there is a compact, deterministic set  $K_\varepsilon$  and such that

$$\mathbb{P}[\tilde{A} \subseteq K_\varepsilon] \geq 1 - \varepsilon.$$

Since  $A$  weakly attracts compact sets, for all  $\delta, \varepsilon > 0$  there is a  $t_0(\delta, \varepsilon)$  such that

$$\mathbb{P}[d(\varphi_t(\omega, K_\varepsilon), A(\theta_t\omega)) > \delta] \leq \varepsilon, \quad \forall t \geq t_0.$$

Hence, also

$$\mathbb{P}[d(\varphi_t(\omega, \tilde{A}(\omega)), A(\theta_t\omega)) > \delta] \leq 2\varepsilon, \quad \forall t \geq t_0.$$

By invariance  $\varphi_t(\omega, \tilde{A}(\omega)) = \tilde{A}(\theta_t\omega)$ ,  $\mathbb{P}$ -a.s.. Thus,

$$\mathbb{P}[d(\tilde{A}(\omega), A(\omega)) > \delta] = \mathbb{P}[d(\tilde{A}(\theta_t\omega), A(\theta_t\omega)) > \delta] \leq 2\varepsilon, \quad \forall t \geq t_0.$$

Since  $\varepsilon$  is arbitrary we conclude

$$\mathbb{P}[d(\tilde{A}(\omega), A(\omega)) > \delta] = 0 \quad \forall \delta > 0,$$

which implies the claim.  $\square$

We point out that the previous lemma does not hold for weak point attractors (see e.g. Example 2.21). If an RDS has a weak attractor  $A$ , then  $A$  admits an  $\mathcal{F}_0$ -measurable version by Lemma 1.3 and [18, Corollary 4.5], that is, there exists an  $\mathcal{F}_0$ -measurable weak attractor  $\tilde{A}$  such that  $A = \tilde{A}$ ,  $\mathbb{P}$ -a.s. (note that we did not assume  $\mathcal{F}_0$  to be complete).

When discussing (weak or pullback) attractors we will always assume that the underlying RDS is global. In contrast, we allow the RDS to be local when we discuss invariant measures and (weak) point attractors. The existence of an invariant measure does not guarantee that the RDS is global but it does impose some obvious constraints on the set  $D$  in the definition of a local RDS.

## 2. SYNCHRONIZATION AND WEAK SYNCHRONIZATION FOR RDS ON COMPLETE SEPARABLE METRIC SPACES

**2.1. Synchronization.** In this section we introduce general sufficient conditions for synchronization by noise for RDS on complete separable metric spaces. More precisely, we show that asymptotic stability (a local stability condition), swift transitivity (an irreducibility condition) and contraction on large sets imply synchronization by noise. If  $E$  is Heine-Borel, then asymptotic stability and contraction on large sets are also necessary conditions. Moreover, swift transitivity is satisfied by SDE of the type (1.1) with locally Lipschitz drift satisfying a one-sided Lipschitz condition (cf. Proposition 3.10 below).

We can now define formally what we mean by synchronization for a given RDS  $\varphi$ .

**Definition 2.1.** We say that *synchronization* occurs if there is a weak attractor  $A(\omega)$  being a singleton, for  $\mathbb{P}$ -a.e.  $\omega \in \Omega$ .

The problems of existence of a weak attractor and synchronization, i.e. it consisting of a single random point, are of quite different nature. Indeed, existence of weak attractors for SDE of the type (1.1) essentially relies on coercivity conditions for the drift  $b$ . For example, the existence of weak attractors for (1.1) under a mild coercivity condition has been shown in [21]. In this work we shall concentrate on the problem of synchronization and thus assume the existence of a weak attractor.

We will now formulate sufficient conditions for synchronization to occur.

**Definition 2.2.** Let  $U \subset E$  be a (deterministic) non-empty open set. We say that  $\varphi$  is *asymptotically stable on  $U$*  if there exists a (deterministic) sequence  $t_n \uparrow \infty$  such that

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} \text{diam}(\varphi_{t_n}(\cdot, U)) = 0\right) > 0. \tag{2.1}$$

**Remark 2.3** (Necessity of asymptotic stability). Assume that synchronization holds and that there is at least one non-empty, open set  $U \subseteq E$  that is attracted by  $A(\omega) = \{a(\omega)\}$  (this is always true if  $E$  is locally compact), i.e.

$$d(\varphi_t(\omega, U), A(\theta_t\omega)) \rightarrow 0 \quad \text{for } t \rightarrow \infty,$$

in probability. Then,  $\varphi$  is asymptotically stable on  $U$ . Indeed:

$$\begin{aligned} \text{diam}(\varphi_t(\omega, U)) &= \sup_{x, y \in U} d(\varphi_t(\omega, x), \varphi_t(\omega, y)) \\ &\leq \sup_{x, y \in U} d(\varphi_t(\omega, x), a(\theta_t\omega)) + d(a(\theta_t\omega), \varphi_t(\omega, y)) \\ &\rightarrow 0 \quad \text{for } t \rightarrow \infty \end{aligned}$$

in probability.

Clearly, property (2.1) follows from the stronger assumption

$$\mathbb{P}\left(\lim_{t \rightarrow +\infty} \text{diam}(\varphi_t(\cdot, U)) = 0\right) > 0, \quad (2.2)$$

but there are a number of interesting cases in which (2.1) holds but (2.2) does not (cf. also Remark 3.2 below):

**Example 2.4.** We provide an example of an RDS  $\psi$  satisfying asymptotic stability, i.e. (2.1), but not satisfying (2.2) regardless of the choice of  $U$ . Consider the one-dimensional SDE

$$dX_t = -X_t dt + dW_t$$

with associated RDS  $\varphi$ . Obviously,  $\varphi_t(\omega, x) - \varphi_t(\omega, y) = (x - y)e^{-t}$ . Let now  $t_n, x_n \uparrow \infty$  be such that

$$\mathbb{P}\left(\sup_{t \in [t_{n-1}, t_n]} \varphi_t(\cdot, x) \geq x_n\right) \rightarrow 1 \quad \text{for } n \rightarrow \infty, \quad (2.3)$$

for all  $x \in \mathbb{R}$ . We choose  $f : \mathbb{R} \rightarrow \mathbb{R}$  smooth, strictly increasing with  $\text{range}(f) = \mathbb{R}$  such that

$$f'(x) \geq ne^{t_n} \quad \forall x \in [x_n, x_{n+1}]$$

and set  $\psi_t(\omega, x) := f(\varphi_t(\omega, f^{-1}(x)))$ . Let  $y > x$ . Then

$$\psi_t(\omega, y) - \psi_t(\omega, x) \geq n(f^{-1}(y) - f^{-1}(x))$$

if  $\varphi_t(\omega, f^{-1}(x)) \geq x_n$  and  $t \in [t_{n-1}, t_n]$ . Due to (2.3) this happens i.o.  $\mathbb{P}$ -a.s.. Hence, for all  $y > x$  we have

$$\limsup_{t \rightarrow \infty} |\psi_t(\cdot, x) - \psi_t(\cdot, y)| = \infty \quad \mathbb{P}\text{-a.s.}$$

and thus (2.2) does not hold. In contrast, (2.1) is easily verified for  $\psi$  and  $U \subseteq \mathbb{R}$  bounded.

Let us first state Lemma 2.5, a very general and almost obvious criterion for synchronization.

**Lemma 2.5.** Let  $\varphi$  be asymptotically stable on  $U$  and  $A$  be an  $\mathcal{F}_0$ -measurable,  $\varphi$ -invariant, random closed set with

$$\mathbb{P}(A \subset U) > 0. \tag{2.4}$$

Then  $A$  is a singleton  $\mathbb{P}$ -a.s..

*Proof.* By property (2.1) there exists a sequence  $t_n \uparrow \infty$  such that

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} \text{diam}(\varphi_{t_n}(\cdot, U)) = 0\right) > 0.$$

Since  $\{A \subset U\}$  is  $\mathcal{F}_0$ -measurable,  $\{\lim_{n \rightarrow \infty} \text{diam}(\varphi_{t_n}(\cdot, U)) = 0\}$  is  $\mathcal{F}_{0, \infty}$ -measurable and  $\mathcal{F}_0$  and  $\mathcal{F}_{0, \infty}$  are independent, we obtain

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} \text{diam}(\varphi_{t_n}(\cdot, A)) = 0\right) > 0.$$

In particular, since  $\text{diam}(\varphi_{t_n}(\cdot, A))$  has the same law as  $\text{diam}(A)$ , we get

$$\mathbb{P}(\text{diam}(A) = 0) > 0.$$

We need to show that this probability is in fact 1. We observe that for each  $t \geq 0$  we have

$$\{\text{diam}(A(\theta_{-t}\omega)) = 0\} \subseteq \{\text{diam}(A(\omega)) = 0\}$$

up to a set of measure 0. Since  $\theta_t$  is  $\mathbb{P}$  invariant these events have the same  $\mathbb{P}$ -mass and thus coincide almost surely. Note that  $\{\text{diam}(A(\theta_{-t}\omega)) = 0\}$  is  $\mathcal{F}_{-t}$ -measurable. Hence,  $\{\text{diam}(A(\omega)) = 0\}$  is measurable with respect  $\cap_{t < 0} \bar{\mathcal{F}}_t$  which is trivial by Kolmogorov's 0-1 law. Here,  $\bar{\mathcal{F}}_t$  is the  $\mathbb{P}$ -completion of  $\mathcal{F}_t$ . Therefore, we get  $\text{diam}(A(\omega)) = 0$  almost surely and the proof of the lemma is complete.  $\square$

In applications to SDE, assumption (2.1) will be a consequence of the property that the top Lyapunov exponent  $\lambda_{top}$  is negative and a regularity estimate, although being more general (cf. Section 3.1 below). Example 2.4 provides an RDS satisfying (2.1), but the top Lyapunov exponent does not exist.

Let us come to assumption (2.4) of Lemma 2.5. We can view it as an obvious consequence of the following condition.

**Definition 2.6.** We say that a random closed set  $A$  has *full support* if

$$\mathbb{P}(A \subset U) > 0 \tag{2.5}$$

for every non-empty (deterministic) open set  $U \subset E$ .

Let us give a sufficient condition for full support.

**Definition 2.7.** We say that  $\varphi$  is *swift transitive* if, for every (starting) ball  $B(x, r)$  and every (arrival) point  $y$ , there is a time  $t > 0$  such that

$$\mathbb{P}(\varphi_t(\cdot, B(x, r)) \subset B(y, 2r)) > 0.$$

**Definition 2.8.** A random closed set  $A$  is said to have *small diameter* if

$$\text{ess inf } \{\text{diam}(A(\omega)); \omega \in \Omega\} = 0. \quad (2.6)$$

Condition (2.6) means that

$$\mathbb{P}(\text{diam}(A) < \varepsilon) > 0$$

for every  $\varepsilon > 0$  and we have the following equivalence:

**Lemma 2.9.** Let  $A$  be a closed random set. Then  $A$  has small diameter iff for each  $\varepsilon > 0$  there is an  $x_0 \in E$  such that

$$\mathbb{P}(A \subset B(x_0, \varepsilon)) > 0.$$

*Proof.* Assume that  $A$  has small diameter and consider a countable family of balls of the form  $B(x_n, \varepsilon)$  where  $\{x_n, n \in \mathbb{N}\}$  is a dense countable set in  $E$ . We know that

$$\mathbb{P}(\text{diam}(A) < \varepsilon) > 0.$$

We have

$$\{\text{diam}(A) < \varepsilon\} \subset \{A \subset B(x_n, \varepsilon) \text{ for some } n \in \mathbb{N}\}$$

hence

$$\begin{aligned} 0 &< \mathbb{P}(A \subset B(x_n, \varepsilon) \text{ for some } n \in \mathbb{N}) \\ &= \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} \{A \subset B(x_n, \varepsilon)\}\right) \\ &\leq \sum_{n \in \mathbb{N}} \mathbb{P}(A \subset B(x_n, \varepsilon)) \end{aligned}$$

and thus  $\mathbb{P}(A \subset B(x_n, \varepsilon)) > 0$  for some  $n \in \mathbb{N}$ , proving the claim. The reverse implication is obvious.  $\square$

**Lemma 2.10.** If  $\varphi$  is swift transitive and  $A$  is an  $\mathcal{F}_0$  measurable,  $\varphi$ -invariant random closed set with small diameter, then  $A$  has full support.

*Proof.* Let  $U$  be a non-empty open set and  $B(y, R) \subset U$ . By Lemma 2.9 there is an  $x_0 \in E$  such that  $\mathbb{P}(A \subset B(x_0, \frac{R}{2})) > 0$ . By Definition 2.7 with the starting ball  $B(x_0, \frac{R}{2})$  and the arrival point  $y$ , there is a time  $t > 0$  such that

$$\mathbb{P}\left(\varphi_t\left(\cdot, B\left(x_0, \frac{R}{2}\right)\right) \subset B(y, R)\right) > 0.$$

By the independence of  $\mathcal{F}_{0,t}$  and  $\mathcal{F}_0$  and the fact that  $\varphi_t$  is  $\mathcal{F}_{0,t}$ -measurable and  $A$  is  $\mathcal{F}_0$ -measurable, it follows that

$$\mathbb{P}\left(A \subset B\left(x_0, \frac{R}{2}\right), \varphi_t\left(\cdot, B\left(x_0, \frac{R}{2}\right)\right) \subset B(y, R)\right) > 0$$

and thus

$$\mathbb{P}(\varphi_t(\cdot, A) \subset B(y, R)) > 0.$$

We have  $\varphi_t(\omega, A(\omega)) = A(\theta_t\omega)$ , hence  $\mathbb{P}(A(\theta_t\cdot) \subset B(y, R)) > 0$ . By  $\theta_t$  invariance of  $\mathbb{P}$  we get

$$\mathbb{P}(A \subset B(y, R)) > 0$$

hence  $\mathbb{P}(A \subset U) > 0$ . The proof is complete.  $\square$

The property of swift transitivity is generally true for SDE with additive noise and drift satisfying a local one-sided Lipschitz condition, see Section 3. Concerning the small diameter property, it looks also very general; we proceed to provide a sufficient condition.

The examples we have in mind which fulfill the small diameter property have the following features. With some (presumably very small) probability, their attractors are driven to regions of strong contraction, where the size of the attractor strictly decreases (cf. Section 3.2 below for examples). Possibly this procedure has to be iterated, until we reach a specified small value of the diameter. Let us formalize one of these steps in a definition.

**Definition 2.11.** We say that  $\varphi$  is *contracting on large sets* if for every  $R > 0$ , there is a ball  $B(y, R)$  and a time  $t > 0$  such that

$$\mathbb{P}\left(\text{diam}(\varphi_t(\cdot, B(y, R))) \leq \frac{R}{4}\right) > 0.$$

**Remark 2.12.** The definition requires contraction, of some ball, for *every* radius  $R$ , not only for large  $R$ . However, contraction of some ball of small radius is a consequence of a suitable local stability assumption, similar to those we already assume. Hence the distinguished feature of this new condition is the fact that some ball of large radius is contracted. The name of the property has been chosen with this idea in mind, although it is not comprehensive of the full power of the definition.

**Remark 2.13** (Necessity of contraction on large sets). Assume that synchronization holds and that  $A(\omega) = \{a(\omega)\}$  weakly attracts all closed, bounded sets (which is always true if  $E$  is Heine-Borel<sup>2</sup>), then  $\varphi$  is contracting on large sets. This follows as in Remark 2.3.

<sup>2</sup>For example, every metric space  $(E, d)$  that is both locally compact and  $\sigma$ -compact allows an equivalent metric  $d'$  such that  $(E, d')$  is Heine-Borel [49].

Let us state the main abstract result of this section.

**Theorem 2.14.** Assume that  $\varphi$  is swift transitive. Then:

- (1) Assume that  $\varphi$  is contracting on large sets. Then every  $\mathcal{F}_0$ -measurable,  $\varphi$ -invariant random compact set  $A$  has small diameter.
- (2) Assume that  $\varphi$  is asymptotically stable on some non-empty open set  $U \subset E$ . Let  $A$  be an  $\mathcal{F}_0$  measurable,  $\varphi$ -invariant random closed set with small diameter. Then  $A$  is a singleton. In particular, if there is a weak attractor  $A$  with small diameter, then synchronization occurs.

*Proof.* (1): *Step 1:* We start by proving the following claim: If

$$\mathbb{P}(A \subset B(x_0, r_0)) > 0$$

for some  $r_0 > 0$ ,  $x_0 \in E$  then

$$\mathbb{P}\left(A \subset B\left(x_1, \frac{2}{3}r_0\right)\right) > 0$$

for some  $x_1 \in E$ .

Indeed: apply Definition 2.11 with  $R = 2r_0$ : there is  $y_1 \in E$ ,  $t_1 > 0$  such that

$$\mathbb{P}\left(\text{diam}(\varphi_{t_1}(\cdot, B(y_1, 2r_0))) \leq \frac{r_0}{2}\right) > 0.$$

For every  $t_0 > 0$ , since  $\mathbb{P}$  is invariant under  $\theta_{t_0}$ , we also have

$$\mathbb{P}\left(\text{diam}(\varphi_{t_1}(\theta_{t_0}\cdot, B(y_1, 2r_0))) \leq \frac{r_0}{2}\right) > 0.$$

Apply Definition 2.7 with the starting ball equal to  $B(x_0, r_0)$  and the arrival point equal to  $y_1$ : there is a time  $t_0 > 0$  such that

$$\mathbb{P}(\varphi_{t_0}(\cdot, B(x_0, r_0)) \subset B(y_1, 2r_0)) > 0.$$

We have  $\{\varphi_{t_0}(\cdot, B(x_0, r_0)) \subset B(y_1, 2r_0)\} \in \mathcal{F}_{0, t_0}$  and

$$\left\{\text{diam}(\varphi_{t_1}(\theta_{t_0}\cdot, B(y_1, 2r_0))) \leq \frac{r_0}{2}\right\} \in \mathcal{F}_{t_0, t_0+t_1},$$

since  $\{\text{diam}(\varphi_{t_1}(B(y_1, 2r_0))) \leq \frac{r_0}{2}\} \in \mathcal{F}_{0, t_1}$  and  $\theta_{t_0}^{-1}\mathcal{F}_{0, t_1} = \mathcal{F}_{t_0, t_0+t_1}$ . Since  $\mathcal{F}_{0, t_0}$  and  $\mathcal{F}_{t_0, t_0+t_1}$  are independent, and

$$\varphi_{t_1}(\theta_{t_0}\omega, \varphi_{t_0}(\omega, B(x_0, r_0))) = \varphi_{t_1+t_0}(\omega, B(x_0, r_0))$$

we deduce

$$\mathbb{P}\left(\text{diam}(\varphi_{t_1+t_0}(\cdot, B(x_0, r_0))) \leq \frac{r_0}{2}\right) > 0.$$

This implies

$$\mathbb{P}\left(\varphi_{t_1+t_0}(\cdot, B(x_0, r_0)) \subseteq \bar{B}(\varphi_{t_1+t_0}(\cdot, x_0), \frac{r_0}{2})\right) > 0.$$

Let  $\{z_m\}_{m \in \mathbb{N}}$  be dense in  $\mathbb{R}^d$ . Then

$$\begin{aligned} & \left\{ \omega \in \Omega : \varphi_{t_1+t_0}(\omega, B(x_0, r_0)) \subseteq \bar{B}(\varphi_{t_1+t_0}(\omega, x_0), \frac{r_0}{2}) \right\} \\ & \subseteq \left\{ \omega \in \Omega : \varphi_{t_1+t_0}(\omega, B(x_0, r_0)) \subseteq B(z_m, \frac{2}{3}r_0) \text{ for some } m \in \mathbb{N} \right\}. \end{aligned}$$

Hence there is an  $x_1 \in E$  such that

$$\mathbb{P} \left( \varphi_{t_1+t_0}(\cdot, B(x_0, r_0)) \subset B \left( x_1, \frac{2}{3}r_0 \right) \right) > 0.$$

By the independence of  $\mathcal{F}_{0, t_1+t_0}$  and  $\mathcal{F}_0$  and the fact that  $\varphi_{t_1+t_0}$  is  $\mathcal{F}_{0, t_1+t_0}$ -measurable and  $A$  is  $\mathcal{F}_0$ -measurable, it follows that

$$\mathbb{P} \left( A \subset B(x_0, r_0), \varphi_{t_1+t_0}(\cdot, B(x_0, r_0)) \subset B \left( x_1, \frac{2}{3}r_0 \right) \right) > 0.$$

Hence

$$\mathbb{P} \left( \varphi_{t_1+t_0}(\cdot, A) \subset B \left( x_1, \frac{2}{3}r_0 \right) \right) > 0.$$

We have  $\varphi_{t_1+t_0}(\omega, A(\omega)) = A(\theta_{t_1+t_0}\omega)$ , hence  $\mathbb{P}(A(\theta_{t_1+t_0}\cdot) \subset B(x_1, \frac{2}{3}r_0)) > 0$ . By  $\theta_{t_1+t_0}$  invariance of  $\mathbb{P}$  we get

$$\mathbb{P} \left( A \subset B \left( x_1, \frac{2}{3}r_0 \right) \right) > 0.$$

*Step 2:* The proof is now obvious. Since  $A$  is a random compact set we can choose  $r_0 > 0$ ,  $x_0 \in E$  such that

$$\mathbb{P}(A \subset B(x_0, r_0)) > 0.$$

Given any  $\varepsilon > 0$ , we may apply step one iteratively until we get

$$\mathbb{P}(A \subset B(x, \varepsilon)) > 0$$

for some  $x \in E$  and the proof is complete.

(2): Simple consequence of Lemma 2.10 and Lemma 2.5. □

We finish this section with a simple example which illustrates the concepts introduced above.

**Example 2.15.** Consider the one-dimensional SDE

$$dX_t = X_t dW_t, \quad X_0 = x.$$

where  $W$  is standard Brownian motion. The RDS generated by the solution is given by

$$\varphi_t(\omega, x) = xe^{-\frac{t}{2} + W_t(\omega)}.$$

$A(\omega) = \{0\}$  is the weak attractor of  $\varphi$ , so synchronization occurs. The RDS  $\varphi$  is asymptotically stable on any bounded open set  $U \subset \mathbb{R}$  and is contracting on large

sets but  $\varphi$  is not swift transitive. Lemma 2.5 can be applied but Lemma 2.10 and Theorem 2.14 cannot.

**2.2. Weak synchronization.** Although contraction on large sets is a necessary condition for synchronization, it is not always easy to check for SDE. In Section 3 below, for (1.1) we prove that  $b$  being monotone on large sets implies contraction on large sets (cf. Proposition 3.10). However, monotonicity on large sets is not necessary for synchronization.

Therefore, in this section we investigate a weaker form of synchronization, so-called weak synchronization. The main improvement is that we are able to prove weak synchronization without assuming contraction on large sets, which in turn allows us to consider drifts  $b$  not necessarily monotone on large sets.

More precisely, for strongly mixing, white noise RDS we prove that the weak asymptotic stability condition (2.10) (a pointwise local stability condition), pointwise strong swift transitivity and a global pointwise stability condition imply weak synchronization. Again, weak asymptotic stability and the pointwise stability condition are also necessary for weak synchronization, while pointwise strong swift transitivity is easily checked for (1.1) under mild conditions as above.

We will assume throughout this subsection that  $\varphi$  is a *local*, white noise RDS.

**Definition 2.16.** We say that *weak synchronization* occurs if there is a minimal weak point attractor  $A(\omega)$  being a singleton, for  $\mathbb{P}$ -a.e.  $\omega \in \Omega$ .

If there is a weak attractor  $A$ , then  $A$  contains each minimal weak point attractor. In particular, synchronization implies weak synchronization.

We now introduce a weaker concept of asymptotic stability. The point is, that asymptotic stability in the sense of Definition 2.2 is not necessary for weak synchronization, while the following concept of weak asymptotic stability obviously is:

**Definition 2.17.** Let  $U \subset E$  be a (deterministic) non-empty open set. We say that  $\varphi$  is *weakly asymptotically stable on  $U$*  if there exists a (deterministic) sequence  $t_n \uparrow \infty$  and a set  $\mathcal{M} \subseteq \Omega$  of positive  $\mathbb{P}$ -measure, such that, for all  $x, y \in U$

$$1_{\mathcal{M}}(\cdot) d(\varphi_{t_n}(\cdot, x), \varphi_{t_n}(\cdot, y)) \rightarrow 0 \quad \text{for } n \rightarrow \infty, \quad (2.7)$$

in probability.

**Remark 2.18.** If weak synchronization occurs, then weak asymptotic stability is satisfied with  $U = E$ ,  $\mathcal{M} = \Omega$  and every sequence  $t_n \rightarrow \infty$  since, for all  $x, y \in E$  we have

$$d(\varphi_t(\cdot, x), \varphi_t(\cdot, y)) \rightarrow 0 \quad \text{for } t \rightarrow \infty,$$

in probability.

Assume that the Markov semigroup corresponding to  $\varphi$  has an ergodic invariant measure  $\rho$  with corresponding statistical equilibrium  $\mu_\omega$  given by (1.4).

Local stability in terms of weak asymptotic stability can be nicely captured in terms of the support of the statistical equilibrium  $\mu_\omega$ , i.e. if  $\varphi$  is weakly asymptotically stable then the support has to consist of finitely many random points. For RDS with negative top Lyapunov exponent and on compact manifolds this goes back to [32].

**Lemma 2.19.** (1) The statistical equilibrium  $\mu_\omega$  is either discrete or diffuse. More precisely, either  $\mu_\omega$  consists of finitely many atoms of the same mass  $\mathbb{P}$ -a.s., i.e. there is an  $N \in \mathbb{N}$  and  $\mathcal{F}_0$ -measurable random variables  $a_1, \dots, a_N$  such that

$$\mu_\omega = \frac{1}{N} \sum_{i=1}^N \delta_{a_i(\omega)}$$

or  $\mu_\omega$  does not have point masses  $\mathbb{P}$ -a.s..

(2) Assume that  $\varphi$  is weakly asymptotically stable on  $U$  with  $\rho(U) > 0$ . Then  $\mu_\omega$  is discrete.

*Proof.* The proof uses modified arguments from [32].

(1): Assume that  $\mu_\omega$  is not diffuse, that is,  $\mu_\omega$  has point masses with positive probability. We let  $\mu$  be the probability measure on  $\Omega \times E$  with marginal  $\mathbb{P}$  on  $\Omega$  and disintegration  $\mu_\omega$ . Let  $g(\omega, x) := \mu_\omega(\{x\})$ . Then

$$\begin{aligned} g(\Theta_t(\omega, x)) &= \mu_{\theta_t \omega}(\{\varphi_t(\omega, x)\}) \\ &= \mu_\omega(\varphi_t(\omega, \cdot)^{-1}\{\varphi_t(\omega, x)\}) \\ &\geq g(\omega, x), \quad \mu - \text{a.s.} \end{aligned}$$

Since  $\Theta_t$  is  $\mu$  ergodic (cf. [9]), this implies that  $g$  is constant  $\mu$ -a.s.. Hence, all point masses of  $\mu_\omega$  have the same mass  $m \in \mathbb{R}_+$ ,  $\mathbb{P}$ -a.s.. Since  $\mu_\omega$  is not diffuse, we have  $m > 0$ . Hence,  $\mathbb{P}$ -a.s.,

$$\begin{aligned} m &= \int_E g(\omega, x) d\mu_\omega(x) \\ &= \int_E \mu_\omega(\{x\}) d\mu_\omega(x) \\ &= N(\omega) m^2, \end{aligned}$$

where  $N(\omega)$  denotes the number of point masses of  $\mu_\omega$ . This implies  $N(\omega) = \frac{1}{m}$ ,  $\mathbb{P}$ -a.s., which finishes the proof.

(2): Due to (1) we only have to show that  $\mu_\omega$  has a point mass with positive probability.

Let  $\Delta := \{(x, x) : x \in E\} \subseteq E \times E$  be the diagonal in  $E \times E$  and let  $\psi : (E \times E) \setminus \Delta \rightarrow [0, \infty)$  be measurable such that  $\psi(x, y) \rightarrow \infty$  for  $d(x, y) \rightarrow 0$  and

$$\mathbb{E} \int_{(E \times E) \setminus \Delta} \psi(x, y) d\mu_\omega(x) d\mu_\omega(y) < \infty.$$

In order to prove the existence of such a function  $\psi$ , define the probability measure  $\nu := \mathbb{E}\mu \otimes \mu$  on  $E \times E$ . Then,  $\nu(\Delta^\varepsilon \setminus \Delta) \rightarrow 0$  for  $\varepsilon \rightarrow 0$ . Hence, we can choose  $\varepsilon_k \downarrow 0$  with  $\varepsilon_k \leq \varepsilon_0 = 1$  such that  $\nu(\Delta^{\varepsilon_k} \setminus \Delta) \leq e^{-k}$ . We then set

$$\psi(x, y) = \begin{cases} k, & \text{if } (x, y) \in \Delta^{\varepsilon_k} \setminus \Delta^{\varepsilon_{k+1}} \\ 0, & \text{if } |x - y| \geq 1. \end{cases}$$

Let  $U$  be as in the assumption of weak asymptotic stability. By invariance of  $\mu_\omega$  we have

$$\begin{aligned} & \mathbb{E} \int_{(E \times E) \setminus \Delta} \psi(x, y) d\mu_\omega(x) d\mu_\omega(y) \\ & \geq \mathbb{E} \int_{(E \times E) \setminus \Delta} \psi(\varphi_t(\omega, x), \varphi_t(\omega, y)) d\mu_\omega(x) d\mu_\omega(y) \\ & \geq \mathbb{E} \int_{(E \times E) \setminus \Delta} 1_U(x) 1_U(y) \psi(\varphi_t(\omega, x), \varphi_t(\omega, y)) d\mu_\omega(x) d\mu_\omega(y). \end{aligned} \tag{2.8}$$

By weak asymptotic stability there is a set  $\mathcal{M} \subseteq \Omega$  with positive  $\mathbb{P}$ -measure and a sequence  $t_n \rightarrow \infty$  such that, for all  $x, y \in U$

$$1_{\mathcal{M}}(\cdot) d(\varphi_{t_n}(\cdot, x), \varphi_{t_n}(\cdot, y)) \rightarrow 0 \quad \text{for } n \rightarrow \infty,$$

in probability. We define  $C(n, x, y, R) := \{\omega \in \Omega : \psi(\varphi_{t_n}(\omega, x), \varphi_{t_n}(\omega, y)) \geq R\}$  and observe

$$\liminf_{n \rightarrow \infty} \mathbb{P}(C(n, x, y, R)) \geq \mathbb{P}(\mathcal{M}),$$

for all  $x, y \in U$ . From (2.8) we obtain

$$\begin{aligned} & \mathbb{E} \int_{(E \times E) \setminus \Delta} \psi(x, y) d\mu_\omega(x) d\mu_\omega(y) \\ & \geq R \mathbb{E} \int_{(E \times E) \setminus \Delta} 1_U(x) 1_U(y) 1_{C(n, x, y, R)}(\omega) d\mu_\omega(x) d\mu_\omega(y). \end{aligned}$$

Since  $\mu_\omega$  is  $\mathcal{F}_0$ -measurable,  $C(n, x, y, R)$  is  $\mathcal{F}_{0,\infty}$ -measurable and  $\mathcal{F}_0, \mathcal{F}_{0,\infty}$  are independent, we conclude

$$\begin{aligned} & \mathbb{E} \int_{(E \times E) \setminus \Delta} 1_U(x) 1_U(y) 1_{C(n,x,y,R)}(\omega) d\mu_\omega(x) d\mu_\omega(y) \\ &= \mathbb{E} \mathbb{E} \left[ \int_{(E \times E) \setminus \Delta} 1_U(x) 1_U(y) 1_{C(n,x,y,R)}(\omega) d\mu_\omega(x) d\mu_\omega(y) \middle| \mathcal{F}_0 \right] \\ &= \mathbb{E} \tilde{\mathbb{E}} \int_{(E \times E) \setminus \Delta} 1_U(x) 1_U(y) 1_{C(n,x,y,R)}(\tilde{\omega}) d\mu_\omega(x) d\mu_\omega(y) \\ &= \mathbb{E} \int_{(E \times E) \setminus \Delta} 1_U(x) 1_U(y) \mathbb{P}[C(n, x, y, R)] d\mu_\omega(x) d\mu_\omega(y) \end{aligned}$$

Using this above, taking  $\liminf_{n \rightarrow \infty}$  and using Fatou's Lemma yields

$$\begin{aligned} & \mathbb{E} \int_{(E \times E) \setminus \Delta} \psi(x, y) d\mu_\omega(x) d\mu_\omega(y) \\ & \geq R \mathbb{E} \int_{(E \times E) \setminus \Delta} 1_U(x) 1_U(y) \liminf_{n \rightarrow \infty} \mathbb{P}[C(n, x, y, R)] d\mu_\omega(x) d\mu_\omega(y) \\ & \geq \mathbb{P}[\mathcal{M}] R \mathbb{E} \int_{(E \times E) \setminus \Delta} 1_U(x) 1_U(y) d\mu_\omega(x) d\mu_\omega(y) \end{aligned}$$

If  $\mu_\omega$  has no point masses, then  $(\mu_\omega \otimes \mu_\omega)(\Delta) = 0$  and thus

$$\begin{aligned} \mathbb{E} \int_{(E \times E) \setminus \Delta} 1_U(x) 1_U(y) d\mu_\omega(y) d\mu_\omega(x) &= \mathbb{E} \int_{E \times E} 1_U(x) 1_U(y) d\mu_\omega(y) d\mu_\omega(x) \\ &= \mathbb{E}(\mu_\omega(U)^2) \\ &\geq \rho(U)^2 \\ &> 0. \end{aligned}$$

Since  $R > 0$  is arbitrary we obtain a contradiction. This concludes the proof.  $\square$

Let

$$E_0 := \{x \in E : \lim_{t \rightarrow \infty} P_t(x, \cdot) = \rho\},$$

where  $P_t(x, \cdot)$  denotes the transition probability and convergence is to be understood in the weak\* sense. Note that if the support of  $\mu_\omega$  is compact with strictly positive probability then it is compact with probability one.

**Proposition 2.20.** (1) Assume that  $A(\omega) := \text{supp}(\mu_\omega)$  is (almost surely) compact. Then  $A$  is a weak point attractor of the set  $E_0$ . In particular, if  $\varphi$  is strongly mixing then  $A$  is a minimal weak point attractor.

(2) If  $\varphi$  is strongly mixing and weakly asymptotically stable on  $U$  with  $\rho(U) > 0$ , then there is an  $N \in \mathbb{N}$  and  $\mathcal{F}_0$ -measurable random variables  $a_1, \dots, a_N$  such that

$$A(\omega) = \text{supp}(\mu_\omega) = \{a_i(\omega) : i = 1, \dots, N\}$$

is a minimal weak point attractor.

*Proof.* Under a compact absorption assumption, Proposition 2.20 (1) corresponds to [31, Theorem 2.4]. The more general setting treated here, however, requires a quite different proof.

(1): We show that  $A$  attracts each  $x \in E_0$  in probability.

Fix  $\varepsilon > 0$ . Let  $x_1, x_2, \dots$  be countable dense in  $E$  and define  $B_i := B(x_i, \varepsilon/3)$  and

$$I(\omega) := \min \left\{ i : A(\omega) \subseteq \bigcup_{j=1}^i B_j \right\}.$$

Note that  $I(\omega)$  is almost surely finite. Let

$$\tilde{A}(\omega) := \bigcup_{j \in J(\omega)} B_j,$$

where

$$J(\omega) := \{j \leq I(\omega) : B_j \cap A(\omega) \neq \emptyset\}.$$

Then  $\tilde{A}(\omega)$  is a random bounded open set and

$$\beta(\omega) := \min \{ \mu_\omega(B_j) : j \in J(\omega) \}$$

is measurable and almost surely positive.

For  $b > 0$  and  $n \in \mathbb{N}$ , we define

$$A_n^{(b)}(\omega) = \bigcup_{x \in D_b(n, \omega)} B(x, \frac{\varepsilon}{3}),$$

where  $D_b(n, \omega)$  is the set of all  $x \in E$  for which  $(\varphi_n(\theta_{-n}\omega)\rho)(B(x, \frac{\varepsilon}{3})) \geq b$ . Note that  $A_n^{(b)}$  is  $\mathcal{F}_{-n,0}$ -measurable (unlike  $A$  and  $\tilde{A}$  which are only  $\mathcal{F}_0$ -measurable).

Since  $(\varphi_n(\theta_{-n}\omega)\rho) \rightarrow \mu_\omega$  weakly\* for  $n \rightarrow \infty$ ,  $\mathbb{P}$ -a.s. and  $\mu_\omega(A(\omega)^{\frac{\varepsilon}{3}}) = 1$  we have,  $\mathbb{P}$ -a.s.,

$$(\varphi_n(\theta_{-n}\omega)\rho)(A(\omega)^{\frac{\varepsilon}{3}}) \geq 1 - \frac{b}{2}$$

for  $n$  large enough. Hence,  $\mathbb{P}$ -a.s. for  $x \in D_b(n, \omega)$  we have  $B(x, \frac{\varepsilon}{3}) \cap A(\omega)^{\frac{\varepsilon}{3}} \neq \emptyset$  for  $n$  large enough. Thus,

$$\lim_{n \rightarrow \infty} \mathbb{P}(A_n^{(b)}(\omega) \subseteq A(\omega)^\varepsilon) = 1.$$

Further, we have

$$\liminf_{n \rightarrow \infty} \mathbb{P}(\tilde{A} \subseteq A_n^{(b)}) \geq \mathbb{P}(\beta(\omega) > b).$$

For given  $\delta > 0$ , we find  $b$  so small and  $n_0$  so large that

$$\mathbb{P}(\tilde{A} \subseteq A_n^{(b)} \subseteq A^\varepsilon) \geq 1 - \frac{\delta}{4}, \tag{2.9}$$

for all  $n \geq n_0$ .

Since  $\tilde{A}(\omega) \supseteq A(\omega)$  we have  $\mu_\omega(\tilde{A}(\omega)) = 1$ . Thus, since  $\tilde{A}(\omega)$  is an open set, weak\* convergence of  $(\varphi_n(\theta_{-n}\omega)\rho)$  to  $\mu_\omega$  implies  $(\varphi_n(\theta_{-n}\omega)\rho)(\tilde{A}(\omega)) \geq 1 - \frac{\delta}{3}$  for all  $n$  large enough  $\mathbb{P}$ -a.s.. Hence, (2.9) yields that there exists some  $n_1 \geq n_0$  such that

$$\mathbb{P}\left((\varphi_n(\theta_{-n}\cdot)\rho)(A_n^{(b)}(\cdot)) \geq 1 - \frac{\delta}{3}\right) \geq 1 - \frac{\delta}{3}$$

for all  $n \geq n_1$ .

Fix  $x \in E_0$ . Then, for all  $n \geq n_1$ ,

$$\begin{aligned} & \liminf_{s \rightarrow \infty} \mathbb{P}(\varphi_{s+n}(\theta_{-(s+n)}\omega, x) \in A^\varepsilon) \\ & \geq \liminf_{s \rightarrow \infty} \mathbb{P}(\varphi_{s+n}(\theta_{-(s+n)}\omega, x) \in A_n^{(b)}) - \frac{\delta}{4} \\ & = \liminf_{s \rightarrow \infty} \mathbb{P}(\varphi_s(\theta_{-(s+n)}\omega, x) \in \varphi_n^{-1}(\theta_{-n}\omega)(A_n^{(b)})) - \frac{\delta}{4} \\ & \geq \mathbb{E}((\varphi_n(\theta_{-n}\omega)\rho)(A_n^{(b)})) - \frac{\delta}{4} \\ & \geq \left(1 - \frac{\delta}{3}\right)^2 - \frac{\delta}{4} \geq 1 - \delta, \end{aligned}$$

where we used the independence of  $\mathcal{F}_{-n,0}$  and  $\mathcal{F}_{-\infty,-n}$ , the fact that  $A_n^{(b)}$  is  $\mathcal{F}_{-n,0}$ -measurable and the fact that  $x \in E_0$  in the step from the third to the fourth line. Since  $\delta > 0$  and  $\varepsilon > 0$  are arbitrary, the claim follows.

Let now  $\varphi$  be strongly mixing, i.e.  $E_0 = E$ . Minimality of  $A$  follows from the fact that every  $\varphi$ -invariant Markov measure is supported by every weak point attractor  $A'$  (cf. [16]), i.e.  $\mu_\omega(A'(\omega)) = 1$  a.s.. Hence,  $A' \subseteq A$  a.s..

(2): Follows from Lemma 2.19 and (1).  $\square$

By Proposition 2.20, without any assumption on  $A$  having full support, asymptotic stability of  $\varphi$  implies that the minimal weak point attractor consists of finitely many points. We note that if  $E$  is connected, then this is true for a weak attractor iff synchronization occurs. Indeed, if  $E$  is connected then so are weak attractors (which follows from the same proof as for [19, Proposition 3.13]). The following example shows that Proposition 2.20 is not true for weak attractors.

**Example 2.21.** Consider

$$d\alpha_t = \cos(2\alpha_t) \circ dW_t^1 + \sin(2\alpha_t) \circ dW_t^2$$

on the one-dimensional sphere  $S^1$ . Then the weak attractor is the whole sphere  $S^1$  while the minimal weak point attractor consists of two (antipodal) random points  $\mathbb{P}$ -a.s. (cf. [5, Remark 4.11]).

**Definition 2.22.** We say that  $\varphi$  is *pointwise strongly swift transitive* if there is a time  $t > 0$  such that for every  $x^1, x^2 \in E$  and every (arrival) point  $y$ ,

$$\mathbb{P}(\varphi_t(\cdot, \{x^1, x^2\}) \subset B(y, 2d(x^1, x^2))) > 0.$$

We obtain

**Theorem 2.23.** Assume that  $\varphi$  has right-continuous trajectories, is strongly mixing, weakly asymptotically stable on  $U$  with  $\rho(U) > 0$ , pointwise strongly swift transitive and

$$\liminf_{t \rightarrow \infty} d(\varphi_t(\omega, x), \varphi_t(\omega, y)) = 0 \quad \text{a.s., } \forall x, y \in E. \quad (2.10)$$

Then, there is a minimal weak point attractor  $A$  consisting of a single random point  $a(\omega)$  and

$$A(\omega) = \text{supp}(\mu_\omega) = \{a(\omega)\} \quad \mathbb{P}\text{-a.s.},$$

i.e. weak synchronization holds.

*Proof.* Since  $\varphi$  is strongly mixing and weakly asymptotically stable on some non-empty open set with positive  $\rho$ -measure, by Proposition 2.20 there are  $\mathcal{F}_0$ -measurable random variables  $a_i(\omega)$ ,  $i = 1, \dots, N$  such that

$$A(\omega) := \{a_i(\omega) : i = 1, \dots, N\}$$

is a minimal weak point attractor.

*Step 1:* By weak asymptotic stability there is an open set  $U$ , a sequence  $t_n \rightarrow \infty$  and a  $\delta > 0$  such that for all  $x, y \in U$  and all  $\eta > 0$

$$\liminf_{n \rightarrow \infty} \mathbb{P}(d(\varphi_{t_n}(\cdot, x), \varphi_{t_n}(\cdot, y)) \leq \eta) \geq \delta > 0. \quad (2.11)$$

Without loss of generality we may assume  $U = B(\varepsilon, x_0)$  for some  $x_0 \in E$ ,  $\varepsilon > 0$ . Let  $x, y \in E$ . By assumption, the stopping time

$$\tau_\varepsilon^{x,y}(\omega) := \inf\{t \geq 0 : d(\varphi_t(\omega, x), \varphi_t(\omega, y)) \leq \frac{\varepsilon}{4}\}$$

is finite  $\mathbb{P}$ -almost surely.

Let now  $a(\omega), b(\omega) \in A(\omega)$  be two  $\mathcal{F}_0$ -measurable selections and let  $\tau_\varepsilon(\omega) := \tau_\varepsilon^{a(\omega), b(\omega)}$ , where  $\tau_\varepsilon^{x,y}$  is defined as above. Due to independence of  $\mathcal{F}_0$  and  $\mathcal{F}_{0,\infty}$ ,  $\tau_\varepsilon$  is finite a.s.. Right-continuity of the trajectories implies that there is a  $\iota : \Omega \rightarrow \mathbb{R}_+ \setminus \{0\}$  such that

$$d(\varphi_{\tau_\varepsilon(\omega)+t}(\omega, a(\omega)), \varphi_{\tau_\varepsilon(\omega)+t}(\omega, b(\omega))) \leq \frac{\varepsilon}{3} \quad (2.12)$$

for all  $t \in [0, \iota(\omega)]$ ,  $\mathbb{P}$ -a.s.. Hence, there is a  $\bar{t}_0 \geq 0$  such that

$$\mathbb{P}\left(d(\varphi_{\bar{t}_0}(\cdot, a(\cdot)), \varphi_{\bar{t}_0}(\cdot, b(\cdot))) \leq \frac{\varepsilon}{3}\right) > 0.$$

Indeed: Assume not. Then

$$\mathbb{P}\left(d(\varphi_t(\cdot, a(\cdot)), \varphi_t(\cdot, b(\cdot))) > \frac{\varepsilon}{3}, \text{ for all } t \in \mathbb{Q}_+\right) = 1,$$

in contradiction to (2.12).

*Step 2:* By pointwise strong swift transitivity and using that  $\varphi$  is a white noise RDS there is a time  $\bar{t}_1 \geq 0$  such that

$$\mathbb{P}(\varphi_{\bar{t}_0 + \bar{t}_1}(\cdot, \{a(\cdot), b(\cdot)\}) \subseteq U) > 0.$$

Again using that  $\varphi$  is a white noise RDS we conclude

$$\liminf_{n \rightarrow \infty} \mathbb{P}(d(\varphi_{\bar{t}_0 + \bar{t}_1 + t_n}(\cdot, a(\cdot)), \varphi_{\bar{t}_0 + \bar{t}_1 + t_n}(\cdot, b(\cdot))) \leq \eta) \geq \frac{\delta}{2} > 0. \quad (2.13)$$

*Step 3:* Assume  $A(\omega)$  is not a singleton  $\mathbb{P}$ -a.s.. Then

$$F(\omega) := \min_{i, j=1, \dots, N, i \neq j} d(a_i(\omega), a_j(\omega)) > 0, \quad (2.14)$$

$\mathbb{P}$ -a.s.. Moreover, since  $\varphi_t(\omega, A(\omega)) = A(\theta_t \omega)$  we have

$$\begin{aligned} F(\theta_t \omega) &= \min_{i, j=1, \dots, N, i \neq j} d(a_i(\theta_t \omega), a_j(\theta_t \omega)) \\ &= \min_{i, j=1, \dots, N, i \neq j} d(\varphi_t(\omega, a_i(\omega)), \varphi_t(\omega, a_j(\omega))) \\ &\leq d(\varphi_t(\omega, a_1(\omega)), \varphi_t(\omega, a_2(\omega))). \end{aligned}$$

Hence, for all  $\eta > 0$

$$\begin{aligned} \mathbb{P}(F(\cdot) \leq \eta) &= \mathbb{P}(F(\theta_{\bar{t}_0 + \bar{t}_1 + t_n} \cdot) \leq \eta) \\ &\geq \mathbb{P}(d(\varphi_{\bar{t}_0 + \bar{t}_1 + t_n}(\cdot, a_1(\cdot)), \varphi_{\bar{t}_0 + \bar{t}_1 + t_n}(\cdot, a_2(\cdot))) \leq \eta). \end{aligned}$$

Taking  $\liminf_{n \rightarrow \infty}$  and using (2.13) we conclude:

$$\mathbb{P}(F(\cdot) \leq \eta) \geq \frac{\delta}{2} > 0, \quad (2.15)$$

for all  $\eta$  in contradiction to (2.14).  $\square$

As we will observe in Section 3.3 below, condition (2.10) is satisfied for a large class of gradient-type SDE.

### 3. SYNCHRONIZATION AND WEAK SYNCHRONIZATION FOR SDE

In this section we provide classes of SDE satisfying asymptotic stability, swift transitivity and contraction on large sets and, thus, synchronization by noise.

The section is divided into four parts. In the first part we will focus on asymptotic stability. We first develop a local stable manifold theorem for general, differentiable RDS and prove that a negative top Lyapunov exponent plus a regularity condition

lead to asymptotic stability. We then provide sufficient conditions for SDE to have a negative top Lyapunov exponent and to satisfy this regularity condition. In the second part we will prove swift transitivity and contraction on large sets for SDE with additive noise. The third part concentrates on gradient-type SDE, proving weak synchronization under weak assumptions. Finally, these results are summarized and applied to SDE in the fourth part.

In this section we consider finite dimensional SDE driven by  $d$ -dimensional Brownian motion, i.e.

$$dX_t = b(X_t) dt + \sigma dW_t, \quad (3.1)$$

with  $\sigma > 0$ ,  $b$  locally Lipschitz continuous and satisfying the following one-sided Lipschitz condition

$$(b(x) - b(y), x - y) \leq \lambda |x - y|^2, \quad (3.2)$$

for all  $x, y \in \mathbb{R}^d$  and some  $\lambda > 0$ .

By [21] there is a white noise RDS  $\varphi$  associated to (3.1), with respect to the canonical setup: The space  $\Omega$  is  $C(\mathbb{R}; \mathbb{R}^d)$ ,  $\mathcal{F}$  is the (not completed) Borel  $\sigma$ -field,  $\mathbb{P}$  is the two-sided Wiener measure,  $\mathcal{F}_{s,t}$  is the  $\sigma$ -algebra generated by  $W_u - W_v$  for  $s \leq v \leq u \leq t$ , where  $W_s : \Omega \rightarrow \mathbb{R}^d$  is defined as  $W_s(\omega) = \omega(s)$ , and  $\theta_t$  is the shift

$$(\theta_t \omega)(s) = \omega(s + t) - \omega(t)$$

which is ergodic.

**3.1. Asymptotic stability and top Lyapunov exponent.** In this section we provide sufficient conditions for asymptotic stability for certain diffusions. We start by considering general RDS and proving that a negative top Lyapunov exponent implies asymptotic stability. Then we provide sufficient conditions for SDE to have negative top Lyapunov exponents.

**3.1.1. A time-discrete, local stable manifold theorem and asymptotic stability.** Let  $\varphi$  be a white-noise RDS on  $\mathbb{R}^d$  with respect to an ergodic metric dynamical system  $(\Omega, \mathbb{P}, \theta)$  and let  $P_t$  be the associated Markovian semigroup. In this section we will introduce the associated Lyapunov spectrum under appropriate assumptions on  $\varphi$  and provide a local stable manifold theorem for discrete time and negative top Lyapunov exponent. We then prove that this implies asymptotic stability.

**Lemma 3.1.** Assume that  $\varphi_t(\omega, \cdot) \in C_{\text{loc}}^{1,\delta}$  for some  $\delta \in (0, 1)$  and all  $t \geq 0$ . Further assume that  $P_1$  has an ergodic invariant measure  $\rho$  such that

$$\mathbb{E} \int_{\mathbb{R}^d} \log^+ \|D\varphi_1(\omega, x)\| d\rho(x) < \infty. \quad (3.3)$$

Then

(1) There are constants  $\lambda_N < \dots < \lambda_1$  (the Lyapunov spectrum) such that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log |D\varphi_n(\omega, x)v| \in \{\lambda_i\}_{i=1}^N,$$

for all  $v \in \mathbb{R}^d \setminus \{0\}$  and  $\mathbb{P} \otimes \rho$ -a.a.  $(\omega, x) \in \Omega \times \mathbb{R}^d$ .

We define the top Lyapunov exponent by  $\lambda_{top} := \lambda_1$ . Assume

$$\mathbb{E} \int_{\mathbb{R}^d} \log^+(\|\varphi_1(\omega, \cdot + x) - \varphi_1(\omega, x)\|_{C^{1,\delta}(\bar{B}(1,0))}) d\rho(x) < \infty \quad (3.4)$$

and  $\lambda_{top} < 0$ . Then

(2) For every  $\varepsilon \in (\lambda_{top}, 0)$  there is a measurable map  $\beta : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}_+ \setminus \{0\}$  such that for  $\rho$ -a.a.  $x \in \mathbb{R}^d$

$$\mathcal{S}(\omega, x) := \{y \in \mathbb{R}^d : |\varphi_n(\omega, y) - \varphi_n(\omega, x)| \leq \beta(\omega, x)e^{\varepsilon n}, \forall n \in \mathbb{N}\}$$

is an open neighborhood of  $x$ ,  $\mathbb{P}$ -a.s..

*Proof.* (1): The introduction of the Lyapunov spectrum and the time-discrete stable manifold theorem will be based on [41]. In order to do so, we need to rewrite the dynamics in an appropriate form. This essentially follows the setup put forward in [36].

We define the following extension of the probability space (cf. e.g. [36, p. 626 and Corollary 3.1.1, Remark (iii)]): Let  $M := \Omega \times \mathbb{R}^d$ ,  $\tilde{\mathcal{F}} := \mathcal{F}_{0,\infty} \otimes \mathcal{B}(\mathbb{R}^d)$ ,  $\mu = \mathbb{P} \otimes \rho|_{\tilde{\mathcal{F}}}$  and  $\tau : M \rightarrow M$  be defined by

$$\tau(m) := (\theta_1\omega, \varphi_1(\omega, x)) \quad \text{for } m = (\omega, x) \in M.$$

By [9]  $\tau$  is ergodic. We then obtain a perfect (time-discrete) cocycle on  $(M, \mu, \tau)$  by

$$Z_n(m, y) := \varphi_n(\omega, y + x) - \varphi_n(\omega, x) \quad \text{for } m = (\omega, x) \in M.$$

Note that  $Z_n(m, 0) = 0$ . We further set  $F_m(y) := Z_1(m, y)$ ,

$$F_m^n := F_{\tau^{n-1}m} \circ \dots \circ F_m$$

and observe  $F_m^n = Z_n(m, \cdot)$ . Obviously,  $F_m \in C^{1,\delta}$  and

$$T_m = T(m) := DF_m(0) = D\varphi_1(\omega, x) \quad \text{for } m = (\omega, x) \in M.$$

Setting  $T_m^n := T_{\tau^{n-1}m} \circ \dots \circ T_m$  we have  $T_m^n = DF_m^n(0) = D\varphi_n(\omega, x)$  for  $m = (\omega, x) \in M$ . By assumption

$$\int \log^+ \|T(m)\| d\mu(m) = \mathbb{E} \int_{\mathbb{R}^d} \log^+ \|D\varphi_1(\omega, x)\| d\rho(x) < \infty.$$

Since  $\mu$  is ergodic, by the multiplicative ergodic theorem [41, Theorem 1.6], there are constants  $\{\lambda_i\}_{i=1}^N$  (the Lyapunov spectrum) such that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log |T_m^n v| \in \{\lambda_i\}_{i=1}^N$$

and the limit exists for  $\mu$ -a.a.  $m \in M$ . Since  $T_m^n v = D\varphi_n(\omega, x)v$  this finishes the proof.

(2): By [41, Theorem 5.1] (a), there are measurable maps  $\beta > \alpha > 0$  such that, for a.a.  $m = (\omega, x) \in M$ ,

$$\begin{aligned} & \{y \in B(0, \alpha(m)) : |F_m^n y| \leq \beta(m)e^{n\varepsilon}, \forall n \in \mathbb{N}\} \\ & = \{y \in B(0, \alpha(\omega, x)) : |\varphi_n(\omega, y+x) - \varphi_n(\omega, x)| \leq \beta(\omega, x)e^{n\varepsilon}, \forall n \in \mathbb{N}\} \end{aligned}$$

is an open neighborhood of  $0 \in \mathbb{R}^d$ , which implies (2). □

**Remark 3.2.** In contrast to the time-continuous local stable manifold theorem developed in [36], Lemma 3.1 (2) only yields local stability along the natural numbers  $n \in \mathbb{N}$ . On the other hand, the assumptions of [36] do not cover our model example of a double-well potential. At this point we make use of the weaker form, as compared to (2.2), of asymptotic stability introduced in Definition 2.2. In fact, as we will see below, Lemma 3.1 (2) will be sufficient to deduce asymptotic stability, which significantly simplifies the proof of asymptotic stability in cases for which no local stable manifold theorem has (yet) been established.

From Lemma 3.1 (2) we obtain the existence of random neighborhoods of points, that are contracted under the (time-discrete) flow. The following Lemma clarifies the relation to asymptotic stability in the sense of Definition 2.2.

**Lemma 3.3.** Let  $U_1$  be a random, non-empty, open set and assume that there is a sequence  $t_n \rightarrow \infty$  such that

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} \text{diam}(\varphi_{t_n}(\cdot, U_1(\cdot))) = 0\right) > 0. \quad (3.5)$$

Then there is a (deterministic) non-empty, open set  $U$  such that

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} \text{diam}(\varphi_{t_n}(\cdot, U)) = 0\right) > 0.$$

In particular,  $\varphi$  is asymptotically stable in the sense of Definition 2.2.

*Proof.* Consider the countable family of balls of the form  $B(x_m, r_m)$  where  $(x_m, r_m)$  is an enumeration of pairs of points  $x_m$  of  $\mathbb{R}^d$  with rational coordinates and positive rational radii  $r_m$ . We have

$$\left\{ \lim_{n \rightarrow \infty} \text{diam}(\varphi_{t_n}(\cdot, U_1(\cdot))) = 0 \right\} \subset \left\{ \lim_{n \rightarrow \infty} \text{diam}(\varphi_{t_n}(\cdot, B(x_m, r_m))) = 0, \text{ for some } m \in \mathbb{N} \right\}.$$

Hence, there exists  $m \in \mathbb{N}$  such that

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} \text{diam}(\varphi_{t_n}(\cdot, B(x_m, r_m))) = 0\right) > 0.$$

The ball  $B(x_m, r_m)$  is the set  $U$  of the definition of asymptotic stability. The proof is complete.  $\square$

As immediate consequence we obtain

**Corollary 3.4.** Let  $\varphi$  be as in Lemma 3.1 and assume (3.4),  $\lambda_{top} < 0$ . Then  $\varphi$  is asymptotically stable in the sense of Definition 2.2.

3.1.2. *Examples with negative top Lyapunov exponent.* In this section we provide three main classes of SDE for which we prove the top Lyapunov exponent to be negative. The first class of examples will be SDE with eventually monotone drifts and large noise. The second class consists of SDE with gradient structure and small noise. The third class are gradient-type SDE with certain symmetric potentials and all noise intensities.

Consider the following SDE with additive noise

$$dX_t = b(X_t) dt + \sigma dW_t \quad \text{on } \mathbb{R}^d, \quad (3.6)$$

where  $\sigma > 0$ ,  $b \in C_{loc}^{1,\delta}(\mathbb{R}^d)$  for some  $\delta \in (0, 1)$  and  $b$  satisfies (3.2). Hence, there is a corresponding white noise RDS  $\varphi$  with<sup>3</sup>  $\varphi_t(\omega, \cdot) \in C_{loc}^{1,\delta}$  and  $D\varphi_t(\omega, x)$  satisfies the equation

$$\frac{d}{dt} D\varphi_t(\omega, x) = Db(\varphi_t(\omega, x)) D\varphi_t(\omega, x), \quad D\varphi_0(\omega, x) = \text{Id}.$$

In particular, by (3.2), given any  $v \in \mathbb{R}^d \setminus \{0\}$ ,

$$\begin{aligned} \frac{d}{dt} |D\varphi_t(\omega, x) v|^2 &= 2(Db(\varphi_t(\omega, x)) D\varphi_t(\omega, x) v, D\varphi_t(\omega, x) v) \\ &\leq 2\lambda |D\varphi_t(\omega, x) v|^2. \end{aligned} \quad (3.7)$$

Gronwall's inequality yields

$$\|D\varphi_t(\omega, x)\| \leq e^{\lambda t}. \quad (3.8)$$

**Definition 3.5.** A vector field  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is said to be *eventually strictly monotone* if there exists an  $R > 0$  such that

$$(b(x) - b(y), x - y) \leq -\lambda_1 |x - y|^2 \quad \text{for all } |x|, |y| > R$$

for some  $\lambda_1 > 0$ .

---

<sup>3</sup>Note that we have  $\varphi_t(\omega, \cdot) \in C_{loc}^{1,\delta}$  and not only  $\varphi_t(\omega, \cdot) \in C_{loc}^{1,\beta}$  for every  $\beta < \delta$  due to the additive noise in (3.6). This may be seen by considering the transformation  $\tilde{\varphi} = \varphi - \sigma W$  and studying the regularity of the corresponding pathwise ODE.

Assume that  $b$  is eventually strictly monotone and  $b \in C_{loc}^{1,\delta}(\mathbb{R}^d)$  for some  $\delta \in (0, 1)$ . Clearly, this implies (3.2) as well as the existence of some  $c > 0$ ,  $C > 0$  such that

$$(b(x), x) \leq -c|x|^2 + C \quad \text{for all } x \in \mathbb{R}^d. \quad (3.9)$$

Indeed, let  $D := \sup_{|x| \leq R} (b(x), x)$  and for  $|x| > R$  let  $y = \frac{Rx}{|x|}$ . Then

$$\begin{aligned} (b(x), x) &= (b(x) - b(y), x) + (b(y), x) \\ &= \left( \frac{|x|}{|x| - R} \right) (b(x) - b(y), x - y) + (b(y), x) \\ &\leq -\lambda_1 \left( \frac{|x|}{|x| - R} \right) |x - y|^2 + \frac{|x|}{R} D \\ &= -\lambda_1 |x| (|x| - R) + \frac{D}{R} |x| \\ &= -\lambda_1 |x|^2 + \left( \lambda_1 R + \frac{D}{R} \right) |x|. \end{aligned}$$

Hence, using Young's inequality we obtain that

$$(b(x), x) \leq -\frac{\lambda_1}{2} |x|^2 + C, \quad \forall |x| > R,$$

which implies (3.9) by local boundedness of  $b$ .

By [28, Theorem 4.3] we obtain that  $\varphi$  is strongly mixing with invariant probability measure  $\rho$ . Due to (3.8) this implies (3.3). Hence, an application of Lemma 3.1 implies the existence of a corresponding (deterministic) Lyapunov spectrum with top Lyapunov exponent  $\lambda_{top}$ .

**Example 3.6.** Let  $b \in C_{loc}^{1,\delta}$  for some  $\delta \in (0, 1)$  be eventually strictly monotone and consider the SDE

$$dX_t = b(X_t) dt + \sigma dW_t \quad \text{on } \mathbb{R}^d$$

with  $\sigma > 0$ . If  $\sigma$  is large enough, then  $\lambda_{top} < 0$ .

*Proof. Step 1:* By (3.7), for any  $v \in \mathbb{R}^d \setminus \{0\}$ ,

$$\begin{aligned} \frac{d}{dt} |D\varphi_t(\omega, x)v|^2 &= 2(Db(\varphi_t(\omega, x)) D\varphi_t(\omega, x)v, D\varphi_t(\omega, x)v) \\ &= 2(Db(\varphi_t(\omega, x)) r_t(\omega, x, v), r_t(\omega, x, v)) |D\varphi_t(\omega, x)v|^2 \end{aligned}$$

where  $r_t(\omega, x, v) = \frac{D\varphi_t(\omega, x)v}{|D\varphi_t(\omega, x)v|}$ . Hence,

$$|D\varphi_t(\omega, x)v|^2 = |v|^2 e^{2 \int_0^t (Db(\varphi_s(\omega, x)) r_s(\omega, x, v), r_s(\omega, x, v)) ds}.$$

Recall that there exists a  $v \in \mathbb{R}^d \setminus \{0\}$  such that

$$\lambda_{top} = \lim_{n \rightarrow \infty} \frac{1}{n} \log |D\varphi_n(\omega, x)v|.$$

Hence,

$$\lambda_{top} = \lim_{n \rightarrow \infty} \frac{1}{n} \int_0^n (Db(\varphi_s(\omega, x)) r_s(\omega, x, v), r_s(\omega, x, v)) ds.$$

With

$$\begin{aligned} \lambda^+(x) &:= \max_{|r|=1} (Db(x)r, r) \\ \lambda^-(x) &:= \min_{|r|=1} (Db(x)r, r) \end{aligned}$$

we thus have

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \int_0^n \lambda^-(\varphi_s(\omega, x)) ds \leq \lambda_{top} \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \int_0^n \lambda^+(\varphi_s(\omega, x)) ds.$$

Since  $b$  satisfies (3.2), we have that  $\lambda^+(x) \leq C$  for all  $x \in \mathbb{R}^d$  and some constant  $C > 0$ . Ergodicity and monotone convergence then yield

$$\lambda_{top} \leq \int_{\mathbb{R}^d} \lambda^+(x) d\rho(x). \quad (3.10)$$

By eventual strict monotonicity of  $b$  we have

$$\begin{aligned} \int_{\mathbb{R}^d} \lambda^+(x) d\rho(x) &= \int_{B_R} \lambda^+(x) d\rho(x) + \int_{B_R^c} \lambda^+(x) d\rho(x) \\ &\leq \|Db\|_{C^0(B_R)} \rho(B_R) - \lambda_1 \rho(B_R^c). \end{aligned} \quad (3.11)$$

Next, we will prove that for  $\sigma \gg 1$  the invariant measure  $\rho$  “flattens”, i.e. for each  $\tilde{R} \geq 0$ ,  $\rho(B_{\tilde{R}}) \rightarrow 0$  for  $\sigma \rightarrow \infty$ . Thus, the right hand side in (3.11) becomes negative for  $\sigma$  large enough, which finishes the proof.

*Step 2:* For each  $\tilde{R} \geq 0$ ,  $\rho(B_{\tilde{R}}) \rightarrow 0$  for  $\sigma \rightarrow \infty$ .

Indeed: Given  $\sigma > 0$  let  $\rho^\sigma$  be the corresponding invariant measure, thus solving the Fokker-Planck equation

$$\frac{\sigma^2}{2} \Delta \rho^\sigma - \operatorname{div}(b \rho^\sigma) = 0$$

in distributional sense. Consequently, for all  $\varphi \in C_c^\infty$  we have

$$\int (\Delta \varphi + \frac{2}{\sigma^2} b \cdot \nabla \varphi) d\rho^\sigma = 0.$$

Since  $\rho^\sigma(\mathbb{R}^d) = 1$ , there is a weakly\* convergent subsequence  $\rho^{\sigma_n} \rightharpoonup^* \rho$  in the space of all signed measures of total variation on  $\mathbb{R}^d$ . Clearly,  $\rho(\mathbb{R}^d) \leq 1$ . Since,

$$- \int (\frac{2}{\sigma^2} b \cdot \nabla \varphi) d\rho^\sigma \leq \frac{2}{\sigma^2} \|b \cdot \nabla \varphi\|_{C^0} \rho^\sigma(\mathbb{R}^d) = \frac{2}{\sigma^2} \|b \cdot \nabla \varphi\|_{C^0}$$

we have

$$\int \Delta \varphi d\rho^{\sigma_n} \leq \frac{2}{\sigma_n^2} \|b \cdot \nabla \varphi\|_{C^0}.$$

Taking the limit yields  $\int \Delta \varphi d\rho \leq 0$ , for all  $\varphi \in C_c^\infty$ . Thus, also  $\int \Delta \varphi d\rho = 0$ , for all  $\varphi \in C_c^\infty$ . We next show that this implies  $\rho = 0$ . Let  $\varphi_\lambda(x) = e^{-\lambda|x|^2}$  and note

$$\Delta \varphi_\lambda(x) = 2\lambda\varphi_\lambda(x)(2\lambda|x|^2 - d).$$

A simple approximation/cut-off argument implies  $\int \Delta \varphi_\lambda d\rho = 0$ . Given  $\tilde{R} > 0$  we can choose  $\lambda$  small enough such that  $-\Delta \varphi_\lambda(x) \geq \lambda$  for all  $x \in B_{\tilde{R}}$ . Let  $\tilde{R}(\lambda) = \sqrt{\frac{d}{2\lambda}}$  and note  $\tilde{R}(\lambda) \geq \tilde{R}$  for all  $\lambda$  small enough. Then

$$\begin{aligned} \rho(B_{\tilde{R}}) &\leq \frac{1}{\lambda} \int_{B_{\tilde{R}}} -\Delta \varphi_\lambda d\rho \\ &\leq \frac{1}{\lambda} \int_{B_{\tilde{R}(\lambda)}} -\Delta \varphi_\lambda d\rho \\ &= \frac{1}{\lambda} \int_{B_{\tilde{R}(\lambda)}^c} \Delta \varphi_\lambda d\rho. \end{aligned}$$

Note that  $\frac{1}{\lambda} \Delta \varphi_\lambda(x) \leq 4e^{-\lambda|x|^2} \lambda|x|^2 \leq C$  for some constant  $C$  independent of  $\lambda \geq 0$ . Since  $\tilde{R}(\lambda) \rightarrow \infty$  for  $\lambda \rightarrow 0$  and  $\rho(\mathbb{R}^d) \leq 1$  we obtain

$$\rho(B_{\tilde{R}}) \leq C\rho(B_{\tilde{R}(\lambda)}^c) \rightarrow 0$$

for  $\lambda \rightarrow 0$ . Hence,  $\rho = 0$  and thus  $\rho^{\sigma_n} \rightarrow^* 0$ , which finishes the proof.  $\square$

We next consider the case of SDE with gradient structure, i.e.

$$dX_t = -\nabla V(X_t) dt + \sigma dW_t \quad \text{on } \mathbb{R}^d, \quad (3.12)$$

with  $\sigma > 0$ ,  $V \in C_{loc}^{2,\delta}(\mathbb{R}^d)$  for some  $\delta > 0$  and  $b := -\nabla V$  satisfying (3.2). By [21] there is an associated white noise RDS  $\varphi$  to (3.12).

If  $\varrho(x) := e^{-\frac{2}{\sigma^2}V(x)} \in L^1(\mathbb{R}^d)$ , then by [48, Theorem, p.243]<sup>4</sup>, the Markovian semi-group defined by  $P_t f(x) := \mathbb{E}f(\varphi_t(\cdot, x))$  has

$$\rho = \frac{1}{Z_\sigma} e^{-\frac{2}{\sigma^2}V(x)} dx$$

as an invariant probability measure, where  $Z_\sigma := \int_{\mathbb{R}^d} e^{-\frac{2}{\sigma^2}V(x)} dx$ . In this case, by [43, Theorem 3],  $P_t$  is strongly mixing with ergodic measure  $\rho$ . Thus, by Lemma 3.1 and (3.8) the top Lyapunov exponent  $\lambda_{top}$  is well-defined and it only remains to show  $\lambda_{top} < 0$ .

<sup>4</sup>In fact, [48, Theorem, p.243] assumes  $b$  to be smooth. However, it is an easy exercise to see that only  $b \in C^\delta$  for some  $\delta > 0$  is required for the proof (cf. also [29, Theorem 10.4.1] for an according regularity result for linear, non-degenerate second order PDE with Hölder coefficients).

**Example 3.7.** Consider the SDE in  $\mathbb{R}^d$

$$dX_t = -\nabla V(X_t) dt + \sigma dW_t,$$

with  $\sigma > 0$ ,  $V \in C_{loc}^{2,\delta}(\mathbb{R}^d)$  for some  $\delta > 0$  and  $b := -\nabla V$  satisfying (3.2). Further assume that

$$\begin{aligned} V(x) &\geq C_0 \log|x| \\ \|D^2V(x)\| &\leq C_0 |x|^N, \end{aligned} \tag{3.13}$$

for all  $|x| \geq R_0$  and some  $C_0, R_0 > 1, N \geq 0$  and that

$$\inf \left\{ \min_{|r|=1} (D^2V(x)r, r) : x \text{ global minimum of } V \right\} > 0.$$

Then  $\lambda_{top} < 0$  for  $\sigma$  small enough.

*Proof.* By (3.13),  $\varrho(x) := e^{-\frac{2}{\sigma^2}V(x)} \in L^1(\mathbb{R}^d)$  for  $\sigma$  small enough. Hence, as we have seen in Section 2.2, there is a corresponding white noise RDS  $\varphi$  with strongly mixing invariant probability measure  $\rho$ . Recall that

$$d\rho(x) = \frac{1}{Z_\sigma} e^{-\frac{2}{\sigma^2}V(x)} dx,$$

where  $Z_\sigma = \int e^{-\frac{2}{\sigma^2}V(x)} dx$ . This integral is finite for  $\sigma$  small enough, because  $V(x) \geq C_0 \log|x|$  for large  $x$ . Let  $\mathcal{M}$  denote the set of global minima of  $V$ . Without loss of generality, we may assume  $V = 0$  on  $\mathcal{M}$  (hence  $V \geq 0$  on  $\mathbb{R}^d$ ) and  $0 \in \mathcal{M}$ . We also have  $DV = 0$  on  $\mathcal{M}$ .

*Step 1:* We prove that, for some constant  $C > 0$ , we have

$$Z_\sigma \geq C\sigma^d \quad \forall \sigma \in (0, 1].$$

Let  $C_1 := \sup_{B(0,1)} \|D^2V\|$ . For  $x \in B(0, 1)$  we have

$$V(x) = \frac{1}{2} (D^2V(\theta_x x)x, x) \leq C_1 |x|^2$$

for some  $\theta_x \in (0, 1)$ . Hence, for  $x \in B(0, \sigma)$ , we have

$$\begin{aligned} V(x) &\leq C_1 \sigma^2 \\ e^{-\frac{2}{\sigma^2}V(x)} &\geq e^{-2C_1} \end{aligned}$$

and therefore

$$Z_\sigma \geq \int_{B(0,\sigma)} e^{-2C_1} dx = C\sigma^d$$

for a suitable constant  $C > 0$ .

*Step 2:* We prove that, for every  $R \geq R_0$ ,

$$\lim_{\sigma \rightarrow 0} \frac{1}{Z_\sigma} \int_{B(0,R)^c} (1 + \|D^2V(x)\|) e^{-\frac{2}{\sigma^2}V(x)} dx = 0.$$

We have (using Step 1)

$$\begin{aligned} & \frac{1}{Z_\sigma} \int_{B(0,R)^c} (1 + \|D^2V(x)\|) e^{-\frac{2}{\sigma^2}V(x)} dx \\ & \leq \frac{1}{C\sigma^d} \int_{B(0,R)^c} (1 + C_0|x|^N) |x|^{-\frac{2C_0}{\sigma^2}} dx \end{aligned}$$

and the result follows by dominated convergence.

*Step 3:* Let  $\mathcal{U}$  be an open neighborhood of  $\mathcal{M}$ . We prove that

$$\lim_{\sigma \rightarrow 0} \frac{1}{Z_\sigma} \int_{\mathcal{U}} e^{-\frac{2}{\sigma^2}V(x)} dx = 1$$

and that, for every  $R \geq R_0$  such that  $\mathcal{U} \subset B(0, R)$ ,

$$\lim_{\sigma \rightarrow 0} \frac{1}{Z_\sigma} \int_{B(0,R) \setminus \mathcal{U}} e^{-\frac{2}{\sigma^2}V(x)} dx = 0.$$

We have  $k = \inf_{\mathcal{U}^c} V > 0$ . Let  $\mathcal{U}' \subset \mathcal{U}$  be such that  $V(x) \leq \frac{k}{2}$  for all  $x \in \mathcal{U}'$ . Then

$$\begin{aligned} \int_{\mathcal{U}} e^{-\frac{2}{\sigma^2}V(x)} dx & \geq |\mathcal{U}'| e^{-\frac{k}{\sigma^2}} \\ \int_{B(0,R) \setminus \mathcal{U}} e^{-\frac{2}{\sigma^2}V(x)} dx & \leq |B(0, R)| e^{-\frac{2k}{\sigma^2}}. \end{aligned}$$

Hence

$$\int_{B(0,R) \setminus \mathcal{U}} e^{-\frac{2}{\sigma^2}V(x)} dx \leq g(\sigma) \int_{\mathcal{U}} e^{-\frac{2}{\sigma^2}V(x)} dx$$

where  $g(\sigma) := \frac{|B(0,R)| e^{-\frac{2k}{\sigma^2}}}{|\mathcal{U}'| e^{-\frac{k}{\sigma^2}}} \rightarrow 0$  as  $\sigma \rightarrow 0$ . Moreover, we have seen in Step 2 that

$$\lim_{\sigma \rightarrow 0} \frac{1}{Z_\sigma} \int_{B(0,R)^c} e^{-\frac{2}{\sigma^2}V(x)} dx = 0.$$

Therefore

$$\begin{aligned} & \frac{1}{Z_\sigma} \int_{\mathcal{U}} e^{-\frac{2}{\sigma^2}V(x)} dx \\ & = 1 - \frac{1}{Z_\sigma} \int_{B(0,R) \setminus \mathcal{U}} e^{-\frac{2}{\sigma^2}V(x)} dx - \frac{1}{Z_\sigma} \int_{B(0,R)^c} e^{-\frac{2}{\sigma^2}V(x)} dx \\ & \geq 1 - g(\sigma) \frac{1}{Z_\sigma} \int_{\mathcal{U}} e^{-\frac{2}{\sigma^2}V(x)} dx - \frac{1}{Z_\sigma} \int_{B(0,R)^c} e^{-\frac{2}{\sigma^2}V(x)} dx \end{aligned}$$

and the result follows by dominated convergence. The proof of the second claim is similar.

*Step 4:* We may now complete the proof. Under our assumptions, there exists an open neighborhood  $\mathcal{U}$  of  $\mathcal{M}$  such that  $c := \inf \{ \min_{|r|=1} (D^2V(x)r, r) : x \in \mathcal{U} \} > 0$ . Since  $\lambda^+(x) = -\min_{|r|=1} (D^2V(x)r, r)$ , we have

$$\lambda^+(x) \leq \|D^2V(x)\| \quad \text{for all } x \in \mathbb{R}^d$$

$$\lambda^+(x) \leq -c \quad \text{for all } x \in \mathcal{U}.$$

Hence, for  $R \geq R_0$  such that  $\mathcal{U} \subset B(0, R)$ , we have

$$\begin{aligned} \int \lambda^+(x) d\rho(x) &\leq \frac{1}{Z_\sigma} \int_{B(0,R)^c} \|D^2V(x)\| e^{-\frac{2}{\sigma^2}V(x)} dx \\ &\quad + \left( \sup_{B(0,R)} \|D^2V(x)\| \right) \frac{1}{Z_\sigma} \int_{B(0,R) \setminus \mathcal{U}} e^{-\frac{2}{\sigma^2}V(x)} dx \\ &\quad - \frac{c}{Z_\sigma} \int_{\mathcal{U}} e^{-\frac{2}{\sigma^2}V(x)} dx. \end{aligned}$$

Form the results of the previous steps we get

$$\int \lambda^+(x) d\rho(x) < 0$$

for  $\sigma$  small enough, hence  $\lambda_{top} < 0$  by the same arguments as used in the proof of Example 3.6.  $\square$

We next consider SDE of the type (3.12) with radially symmetric potential. Note that we neither need to assume  $\sigma$  small nor an assumption of the type (3.13) here.

**Example 3.8.** Consider the SDE

$$dX_t = -\nabla V(X_t) dt + \sigma dW_t \quad \text{on } \mathbb{R}^d,$$

with  $\sigma > 0$  and  $b := -\nabla V$  satisfying (3.2). Further assume that  $V$  is radially symmetric with  $V(x) = g(|x|^2)$ ,  $g \in C_{loc}^{2,\delta}$  being a convex function and  $\varrho(x) = e^{-\frac{2}{\sigma^2}V(x)} \in L^1(\mathbb{R}^d)$ . Then  $\lambda_{top} < 0$ .

*Proof.* Case  $d = 1$ : Let  $x_0 := 0$  and  $x_n := \inf \{ x \geq x_{n-1} : e^{-\frac{2}{\sigma^2}V(x)} \leq \frac{1}{n} \}$ . Since  $e^{-\frac{2}{\sigma^2}V(x)} \in L^1(\mathbb{R}^d)$  we have  $x_n < \infty$  for all  $n \in \mathbb{N}$ . Moreover,  $x_n \rightarrow \infty$  and  $V'(x_n) \geq 0$ .

By symmetry  $V'(0) = 0$ . We conclude that

$$\begin{aligned}
\int \lambda^+(x) d\rho(x) &= -\frac{1}{Z_\sigma} \int_{\mathbb{R}} V''(x) e^{-\frac{2}{\sigma^2} V(x)} dx \\
&= -\frac{2}{Z_\sigma} \lim_{n \rightarrow \infty} \int_0^{x_n} V''(x) e^{-\frac{2}{\sigma^2} V(x)} dx \\
&= -\frac{2}{Z_\sigma} \lim_{n \rightarrow \infty} \left( V'(x) e^{-\frac{2}{\sigma^2} V(x)} \Big|_0^{x_n} + \frac{2}{\sigma^2} \int_0^{x_n} |V'(x)|^2 e^{-\frac{2}{\sigma^2} V(x)} dx \right) \\
&\leq -\frac{4}{Z_\sigma \sigma^2} \int_0^\infty |V'(x)|^2 e^{-\frac{2}{\sigma^2} V(x)} dx \\
&< 0,
\end{aligned}$$

which implies  $\lambda_{top} < 0$  by the same arguments as used in the proof of Example 3.6.

It is, in fact, well-known that in case  $d = 1$  the claim is true under much weaker assumptions. Synchronization in this *monotone* case is discussed in [22], for example.

*Case  $d \geq 2$ :* Since  $V(x) = g(|x|^2)$  we compute

$$\begin{aligned}
\nabla V(x) &= 2g'(|x|^2)x \\
D^2V(x) &= 4g''(|x|^2)x \otimes x + 2g'(|x|^2)Id.
\end{aligned}$$

Thus (using  $g'' \geq 0$ )

$$\begin{aligned}
(D^2V(x)r, r) &= 4g''(|x|^2)(x, r)^2 + 2g'(|x|^2)|r|^2 \\
\min_{|r|=1} (D^2V(x)r, r) &= 2g'(|x|^2).
\end{aligned}$$

We note that

$$\int e^{-\frac{2}{\sigma^2} V(x)} dx = \int e^{-\frac{2}{\sigma^2} g(|x|^2)} dx = C \int_{\mathbb{R}_+} r^{d-1} e^{-\frac{2}{\sigma^2} g(r^2)} dr < \infty.$$

Hence, there is a sequence  $t_n \uparrow \infty$  such that  $t_n^{d-1} e^{-\frac{2}{\sigma^2} g(t_n^2)} \rightarrow 0$  for  $n \rightarrow \infty$ . We conclude,

$$\begin{aligned}
\int \lambda^+(x) d\rho(x) &= -\frac{2}{Z} \int g'(|x|^2) e^{-\frac{2}{\sigma^2} g(|x|^2)} dx \\
&= -c \int_{\mathbb{R}_+} r^{d-1} g'(r^2) e^{-\frac{2}{\sigma^2} g(r^2)} dr \\
&= c \int_{\mathbb{R}_+} r^{d-2} \frac{d}{dr} e^{-\frac{2}{\sigma^2} g(r^2)} dr \\
&= c \lim_{n \rightarrow \infty} r^{d-2} e^{-\frac{2}{\sigma^2} g(r^2)} \Big|_0^{t_n} - c(d-2) \int_{\mathbb{R}_+} r^{d-3} e^{-\frac{2}{\sigma^2} g(r^2)} dr \\
&< 0,
\end{aligned}$$

which implies  $\lambda_{top} < 0$  by the same arguments as used in the proof of Example 3.6.  $\square$

**3.1.3. Deducing asymptotic stability.** As in the previous section we consider SDE with additive noise of the type (3.6) with  $\sigma > 0$ ,  $b \in C_{loc}^2(\mathbb{R}^d)$  and  $b$  satisfying (3.2). Assume that  $\varphi$  is strongly mixing with invariant measure  $\rho$  (cf. the previous section for appropriate conditions). In the last section we have introduced sufficient conditions for the corresponding RDS  $\varphi$  to have negative top Lyapunov exponent. According to Corollary 3.4 this implies asymptotic stability for  $\varphi$  if condition (3.4) holds. Thus, we next present sufficient conditions implying (3.4).

**Lemma 3.9.** Assume that  $b \in C_{loc}^2(\mathbb{R}^d)$  satisfies (3.2) and

$$\|D^2b(x)\| \leq C(|x|^M + 1) \quad \forall x \in \mathbb{R}^d, \quad (3.14)$$

for some  $M \in \mathbb{N}, C \geq 0$ . Further assume

$$\int_{\mathbb{R}^d} \log^+(|x|) d\rho(x) < \infty. \quad (3.15)$$

Then

$$\mathbb{E} \int_{\mathbb{R}^d} \log^+(\|\varphi_1(\omega, \cdot + x) - \varphi_1(\omega, x)\|_{C^{1,\delta}(\bar{B}(0,1))}) d\rho(x) < \infty,$$

for every  $\delta \in (0, 1)$ .

*Proof.* We first note that  $\varphi_t(\omega, \cdot) \in C_{loc}^2$  since  $b \in C_{loc}^2$ . Indeed, this follows by considering the transformation  $\tilde{\varphi}_t(\omega, x) := \varphi_t(\omega, x) - \sigma W_t(\omega)$  satisfying

$$\frac{d}{dt} \tilde{\varphi}_t = b(\tilde{\varphi}_t + \sigma W_t).$$

Then,  $\tilde{\varphi}_t(\omega, \cdot) \in C_{loc}^2$  follows by arguments similar to [47, Theorem 2.10].

Let

$$F^x(\omega, y) := \varphi_1(\omega, y + x) - \varphi_1(\omega, x).$$

We aim to estimate  $\|F^x(\omega, \cdot)\|_{C^{1,\delta}(\bar{B}(0,1))}$ . Due to (3.8) we have that

$$|F^x(\omega, y)| = |\varphi_1(\omega, y + x) - \varphi_1(\omega, x)| \leq e^\lambda, \quad \forall y \in \bar{B}(0, 1)$$

and since  $DF^x(\omega, y)(v) := D\varphi_1(\omega, y + x)(v)$  we obtain that

$$\|DF^x(\omega, y)\| = \|D\varphi_1(\omega, y + x)\| \leq e^\lambda.$$

It remains to estimate  $\|DF^x(\omega, \cdot)\|_{C^\delta(\bar{B}(0,1))}$ . First note that (here and in the following  $C$  denotes a generic constant that may change value from line to line)

$$\begin{aligned} \|DF^x(\omega, \cdot)\|_{C^\delta(\bar{B}(0,1))} &\leq \|DF^x(\omega, \cdot)\|_{C(\bar{B}(0,1))} + C\|D^2F^x(\omega, \cdot)\|_{C(\bar{B}(0,1))} \\ &\leq e^\lambda + C\|D^2\varphi_1(\omega, \cdot)\|_{C(\bar{B}(x,1))}. \end{aligned}$$

In the following, for simplicity, we suppress the  $\omega$ -dependence in the notation. Since, for all  $z \in \bar{B}(x, 1)$  and all  $v, w \in \mathbb{R}^d$  with  $|v|, |w| \leq 1$ ,

$$\frac{d}{dt} D^2 \varphi_t(z)(v, w) = D^2 b(\varphi_t(z))(D\varphi_t(z)v, D\varphi_t(z)w) + Db(\varphi_t(z)) D^2 \varphi_t(z)(v, w),$$

we have that

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} |D^2 \varphi_t(z)(v, w)|^2 &\leq \|D^2 b(\varphi_t(z))\| \|D\varphi_t(z)\|^2 |D^2 \varphi_t(z)(v, w)| \\ &\quad + \langle Db(\varphi_t(z)) D^2 \varphi_t(z)(v, w), D^2 \varphi_t(z)(v, w) \rangle \\ &\leq e^{2\lambda t} C (|\varphi_t(z)|^M + 1) |D^2 \varphi_t(z)(v, w)| + \lambda |D^2 \varphi_t(z)(v, w)|^2 \\ &\leq e^{4\lambda t} C (|\varphi_t(z)|^M + 1)^2 + (\lambda + 1) |D^2 \varphi_t(z)(v, w)|^2. \end{aligned}$$

Hence,

$$\begin{aligned} |D^2 \varphi_t(z)(v, w)|^2 &\leq \int_0^t e^{2(\lambda+1)(t-s)} e^{4\lambda s} C (|\varphi_s(z)|^M + 1)^2 ds \\ &\leq C \int_0^t |\varphi_s(z)|^{2M} ds + C, \end{aligned}$$

for all  $t \in [0, 1]$  and thus

$$\|D^2 \varphi_1(z)\|^2 \leq C \int_0^1 |\varphi_s(z)|^{2M} ds + C.$$

Since

$$|\varphi_s(\omega, z)| \leq |\varphi_s(\omega, x)| + e^\lambda,$$

for all  $z \in \bar{B}(x, 1)$  and  $s \in [0, 1]$ , we have

$$\|D^2 \varphi_1(\omega)\|_{C(\bar{B}(x, 1))}^2 \leq C \int_0^1 |\varphi_s(\omega, x)|^{2M} ds + C.$$

In conclusion,

$$\|F^x(\omega, \cdot)\|_{C^{1, \delta}(\bar{B}(0, 1))} \leq C \int_0^1 |\varphi_s(\omega, x)|^{2M} ds + C.$$

Hence, using Fubini's theorem and Jensen's inequality (taking  $C \geq 1$ ),

$$\begin{aligned} &\mathbb{E} \int_{\mathbb{R}^d} \log^+ (\|F^x(\omega, \cdot)\|_{C^{1, \delta}(\bar{B}(0, 1))}) d\rho(x) \\ &\leq \mathbb{E} \int_{\mathbb{R}^d} \log^+ \left( C \int_0^1 |\varphi_s(\omega, x)|^{2M} ds + C \right) d\rho(x) \\ &\leq \int_{\mathbb{R}^d} \log^+ \left( C \int_0^1 \mathbb{E} [|\varphi_s(\omega, x)|^{2M}] ds + C \right) d\rho(x). \end{aligned}$$

We note that (3.2) implies that there is a constant  $C > 0$  such that

$$(b(x), x) \leq C(|x|^2 + 1) \quad \forall x \in \mathbb{R}^d.$$

Hence, an application of Itô's formula yields

$$\sup_{r \in [0,1]} \mathbb{E} |\varphi_r(\omega, x)|^{2p} \leq C(|x|^{2p} + 1),$$

for all  $p \geq 1$  (where  $C$  depends on  $p$ ). Thus, using (3.15) we conclude that

$$\begin{aligned} \mathbb{E} \int_{\mathbb{R}^d} \log^+ \left( \|F^x(\omega, \cdot)\|_{C^{1,\delta}(\bar{B}(0,1))} \right) \rho(dx) &\leq \int_{\mathbb{R}^d} \log^+ (C|x|^{2M} + C) d\rho(x) \\ &< \infty, \end{aligned}$$

which finishes the proof.  $\square$

**3.2. Properties of dissipativity, contraction and swift transitivity.** This section is devoted to the proof of contraction on large sets and swift transitivity. We consider SDE with additive noise

$$dX_t = b(X_t)dt + \sigma dW_t, \tag{3.16}$$

where  $b$  is locally Lipschitz and satisfies (3.2),  $\sigma > 0$ .

**Proposition 3.10.** Let  $\varphi$  be the RDS associated to (3.16). Then, for all balls  $B(x, r)$  and all  $\delta > 0$ ,  $z \in \mathbb{R}^d$ , there is a time  $T_0 > 0$  such that for all  $t_0 \in (0, T_0]$  one has

$$\mathbb{P}(\varphi_{t_0}(\cdot, x') \in B(x' + z, \delta), \forall x' \in B(x, r)) > 0.$$

In particular, the swift transitivity property holds.

Assume, in addition, that  $b$  is monotone on large sets, i.e. for each  $r > 0$  there exists some  $z \in \mathbb{R}^d$  such that

$$(b(x) - b(y), x - y) < 0 \quad \text{for all } x \neq y, x, y \in B(z, r).$$

Then the property of contraction on large sets holds.

*Proof. Part 1* (swift transitivity). Fix  $x, z \in \mathbb{R}^d$  and  $r, \delta > 0$ .

Let  $B := \sup_{|v| \leq r+|z|+1} |b(x+v)|$ , and  $T_0 := \frac{\delta \wedge 1}{4B} \wedge 1$ .

Fix  $t_0 \in (0, T_0]$  and define

$$\psi(t) := x + \frac{t}{t_0}z, \quad t \in [0, t_0],$$

and

$$f(t) := \frac{1}{\sigma} \left( \psi(t) - x - \int_0^t b(\psi(s)) ds \right), \quad t \in [0, t_0].$$

Abusing notation we write  $\varphi_t(g, y)$ ,  $t \in [0, t_0]$  for the solution of (3.16) with initial condition  $y$  driven by  $g$  instead of  $W$  where  $g \in C_0 := \{h \in C([0, t_0], \mathbb{R}^d) : g(0) = 0\}$ .

Then  $\varphi_t(f, x) = \psi(t)$ ,  $t \in [0, t_0]$  and, in particular,  $\varphi_{t_0}(f, x) = x + z$ .

The map  $g \mapsto \varphi_t(g, y)$  is continuous from  $C_0$  to  $C([0, t_0], \mathbb{R}^d)$  with respect to the supremum norm  $\|\cdot\|$ . Therefore, there exists some  $\varepsilon \in (0, (\delta \wedge 1)/(2\sigma))$  such that for  $g \in C_\varepsilon := \{\bar{g} \in C_0 : \|\bar{g} - f\| \leq \varepsilon\}$  we have

$$\|\varphi_t(g, x) - \varphi_t(f, x)\| \leq \delta/2.$$

Let  $x' \in B(x, r)$  and

$$\tau(g) := \inf\{t \geq 0 : |\varphi_t(g, x') - x'| = |z| + 1\}$$

and assume that there exists some  $g \in C_\varepsilon$  for which  $\tau := \tau(g) < t_0$ . Then

$$\begin{aligned} |z| + 1 &= |\varphi_\tau(g, x') - x'| \leq Bt_0 + \sigma|g(\tau)| \leq Bt_0 + \sigma(|f(\tau)| + \varepsilon) \\ &\leq Bt_0 + \sigma \left( \sup_{0 \leq s \leq t_0} |f(s)| + \varepsilon \right) \\ &\leq Bt_0 + |z| + Bt_0 + \sigma\varepsilon < \frac{1}{4} + |z| + \frac{1}{4} + \frac{1}{2} = |z| + 1, \end{aligned}$$

which is a contradiction. Hence, for  $g \in C_\varepsilon$ , we have

$$\begin{aligned} |\varphi_{t_0}(g, x') - (x' + z)| &\leq |\varphi_{t_0}(g, x') - x' - \sigma f(t_0)| + |\sigma f(t_0) - z| \\ &\leq Bt_0 + \sigma|g(t_0) - f(t_0)| + |\sigma f(t_0) - z| \\ &\leq Bt_0 + \sigma\varepsilon + Bt_0 \leq \delta. \end{aligned}$$

Since  $\mathbb{P}(W|_{[0, t_0]} \in C_\varepsilon) > 0$  the first claim in the proposition follows.

**Part 2** (contraction on large sets).

Let  $R > 0$ . By assumption there is a  $z \in \mathbb{R}^d$  such that

$$(b(x) - b(y), x - y) < 0 \quad \text{for all } x \neq y, x, y \in B(z, 3R). \quad (3.17)$$

Let

$$\omega^0(t) := -\frac{tb(z)}{\sigma} \quad t \geq 0.$$

Then  $\varphi_t(\omega^0, z) = z$  for all  $t \geq 0$ . Due to (3.17),

$$t \mapsto |\varphi_t(\omega^0, x) - \varphi_t(\omega^0, z)| = |\varphi_t(\omega^0, x) - z|$$

is non-increasing for all  $x \in B(z, R)$ . Hence,

$$\varphi_t(\omega^0, B(z, R)) \subseteq B(z, R) \quad \forall t \geq 0.$$

By (3.17), there is a  $c > 0$  such that

$$(b(x) - b(y), x - y) < -c|x - y|^2 \quad \forall x \neq y, x, y \in B(z, 2R), |x - y| \geq \frac{R}{9}. \quad (3.18)$$

Let  $T_0 > 0$  such that  $e^{-cT_0} \leq \frac{1}{9}$ . Further, let  $x \in B(z, R)$  and define

$$\tau := \inf\{t \geq 0 : |\varphi_t(\omega^0, x) - z| \leq \frac{R}{9}\}.$$

Due to (3.18) we have that

$$d|\varphi_t(\omega^0, x) - \varphi_t(\omega^0, z)|^2 \leq -2c|\varphi_t(\omega^0, x) - \varphi_t(\omega^0, z)|^2 dt \quad \text{on } [0, \tau].$$

By Gronwall's Lemma this implies

$$|\varphi_t(\omega^0, x) - z| = |\varphi_t(\omega^0, x) - \varphi_t(\omega^0, z)| \leq e^{-ct} R, \quad \forall t \in [0, \tau]$$

and thus  $\tau \leq T_0$ . Since  $t \mapsto |\varphi_t(\omega^0, x) - z|$  is non-increasing, we conclude that

$$|\varphi_{T_0}(\omega^0, x) - z| \leq \frac{R}{9}.$$

By continuity of  $\omega \mapsto \varphi_{T_0}(\omega, \cdot) \in C(\bar{B}(z, R); \mathbb{R}^d)$  this implies

$$\mathbb{P}\left(\varphi_{T_0}(\cdot, B(z, R)) \subseteq B\left(z, \frac{R}{8}\right)\right) > 0$$

and thus

$$\mathbb{P}\left(\text{diam}(\varphi_{T_0}(\cdot, B(z, R))) \leq \frac{R}{4}\right) > 0.$$

□

**Remark 3.11.** Note that we did not use the one-sided Lipschitz property (3.2) in the proof of the first part of Proposition 3.10, so that the statement even holds assuming only that the drift  $b$  is locally Lipschitz continuous.

**3.3. Weak synchronization for gradient-type SDE.** In the case of gradient-type SDE we can use the results from Section 2.2 in order to deduce weak synchronization without assuming contraction of large balls (as compared to Section 2.1). In fact, we will prove that (2.10) is always satisfied as soon as there is an invariant measure.

More precisely, consider the SDE

$$dX_t = -\nabla V(X_t)dt + \sigma dW_t \quad \text{on } \mathbb{R}^d, \quad (3.19)$$

with  $V \in C^2(\mathbb{R}^d, \mathbb{R})$ ,  $\sigma > 0$  and  $b := -\nabla V$  satisfying assumption (3.2). Further assume  $\varrho(x) := e^{-\frac{2}{\sigma^2}V(x)} \in L^1(\mathbb{R}^d)$ . As seen in Section 3.1.2, there is an associated white noise RDS  $\varphi$  to (3.19) and  $P_t f(x) := \mathbb{E}f(\varphi_t(\cdot, x))$  is strongly mixing with ergodic measure

$$\rho = \frac{1}{Z_\sigma} e^{-\frac{2}{\sigma^2}V(x)} dx,$$

where  $Z_\sigma := \int_{\mathbb{R}^d} e^{-\frac{2}{\sigma^2}V(x)} dx$ .

**Theorem 3.12.** Assume that  $\varrho(x) := e^{-\frac{2}{\sigma^2}V(x)} \in L^1(\mathbb{R}^d)$  and that  $\varphi$  is weakly asymptotically stable on  $U$  with  $\rho(U) > 0$ . Then, there is a minimal weak point attractor  $A$  consisting of a single random point  $a(\omega)$  and

$$A(\omega) = \text{supp}(\mu_\omega) = \{a(\omega)\} \quad \mathbb{P}\text{-a.s.}$$

The proof of Theorem 3.12 is a simple consequence of the following Lemma and Theorem 2.23. Note that by Proposition 3.10  $\varphi$  is pointwise strongly swift transitive.

**Lemma 3.13.** (1) Assume that  $\varrho(x) := e^{-\frac{2}{\sigma^2}V(x)} \in L^1(\mathbb{R}^d)$ . Then, for each  $v \in \mathbb{R}^d \setminus \{0\}$ , there exists some  $z \in \mathbb{R}^d$  such that

$$(b(z) - b(z - v), v) < 0. \quad (3.20)$$

(2) Let  $b$  be locally Lipschitz continuous and satisfy (3.2). Further, assume that for each  $v \in \mathbb{R}^d \setminus \{0\}$ , there exists some  $z \in \mathbb{R}^d$  such that  $b$  satisfies (3.20) and that the SDE (3.16) admits an invariant probability measure  $\rho$ . Then, for each pair  $x, y \in \mathbb{R}^d$ , we have

$$\liminf_{t \rightarrow \infty} |\varphi_t(x) - \varphi_t(y)| = 0,$$

almost surely.

*Proof.* (1): *Step 1:* We claim that for each  $v \in \mathbb{R}^d \setminus \{0\}$  there exists some  $z \in \mathbb{R}^d$ , such that

$$V(z) < \frac{1}{2}(V(z + v) + V(z - v)). \quad (3.21)$$

Assume this is wrong for some particular  $v \in \mathbb{R}^d \setminus \{0\}$ , then for every  $z \in \mathbb{R}^d$  we have

$$V(z) \geq \frac{1}{2}(V(z + v) + V(z - v)).$$

Therefore, for each  $z \in \mathbb{R}^d$ , one of the functions  $n \mapsto V(z + nv), V(z - nv)$  is non-increasing for  $n \in \mathbb{N}_0$ . Fix  $z \in \mathbb{R}^d$ . Without loss of generality let  $n \mapsto V(z + nv)$  be non-increasing. Let  $n \in \mathbb{N}_0$ . Due to the one-sided Lipschitz condition on  $b$  the function  $g(h) := V(z + nv + hv) + \frac{\lambda}{2}(h|v|)^2$  is convex on  $h \in [0, 1]$ . Moreover,  $g(0) \leq V(z)$  and  $g(1) \leq V(z) + \frac{\lambda}{2}|v|^2$ . Since  $g$  is convex this implies  $\sup_{h \in [0, 1]} g(h) \leq V(z) + \frac{\lambda}{2}|v|^2$ . Therefore,  $\sup_{h \in [0, 1]} V(z + nv + hv) \leq V(z) + \frac{\lambda}{2}|v|^2$  for all  $n \in \mathbb{N}_0$ . Hence,

$$\gamma \mapsto V(z + \gamma v)$$

where  $\gamma \in [0, \infty)$ , is bounded from above. In particular,  $\int_{\mathbb{R}} \varrho(z + hv) dh = \infty$  holds for each  $z \in \mathbb{R}^d$ , and therefore  $\varrho$  cannot be integrable.

*Step 2:* To see the claim in the first part of the lemma, let  $f(h) := V(z + hv) - V(z)$  with  $z$  chosen such that (3.21) holds. Due to (3.21) we have

$$f(0) = 0 < \frac{1}{2}(f(1) + f(-1)). \quad (3.22)$$

Assume that  $f'(\alpha) \leq f'(\alpha - 1)$  for all  $\alpha \in [0, 1]$ . Integrating over  $[0, 1]$  yields  $f(1) \leq -f(-1)$  contradicting (3.22). Therefore, there exists some  $\alpha \in [0, 1]$  such

that  $f'(\alpha) > f'(\alpha - 1)$ . The first claim in the lemma follows after replacing  $z + \alpha v$  by  $z$ .

(2): The proof is inspired by the *controllability* approach which was used, for example, in [4].

For  $x, y \in \mathbb{R}^d$ , let  $C(x, y)$  be the closure of the set of all  $(u, v) \in \mathbb{R}^d \times \mathbb{R}^d$  for which there exists some  $T \geq 0$  and  $f \in C([0, T], \mathbb{R}^d)$  such that  $f(0) = 0$  and the (unique) solution of the pair

$$X_t = x + \int_0^t b(X_s) ds + \sigma f_t, \quad Y_t = y + \int_0^t b(Y_s) ds + \sigma f_t$$

satisfies  $(X_T, Y_T) = (u, v)$ .

*Step 1:* Fix  $x \neq y \in \mathbb{R}^d$  and let  $C := C(x, y)$ . We show that  $C \cap \Delta \neq \emptyset$ , where  $\Delta$  is the diagonal in  $\mathbb{R}^d \times \mathbb{R}^d$  as before. First, note that the first part of Proposition 3.10 shows that  $(u, v) \in C$  implies  $(u + z, v + z) \in C$  for every  $z \in \mathbb{R}^d$ . Therefore the infimum in  $\delta := \inf\{|u - v| : (u, v) \in C\}$  is actually attained, say at  $(u_0, v_0)$ . Assume that  $\delta > 0$  and let  $z$  satisfy the assumption in the lemma for  $v := v_0 - u_0$ . Then  $(z - v, z) \in C$  and (3.20) holds which shows (using the control  $f = 0$ ) that there exists  $(\tilde{u}, \tilde{v}) \in C$  for which  $|\tilde{v} - \tilde{u}| < \delta$  contradicting the definition of  $\delta$ . Therefore,  $\delta = 0$  and  $C \cap \Delta \neq \emptyset$ .

*Step 2:* Fix  $\varepsilon > 0$ . We show that for each  $x, y \in \mathbb{R}^d$  we have that

$$\liminf_{t \rightarrow \infty} |\varphi_t(x) - \varphi_t(y)| \leq \varepsilon, \text{ almost surely,}$$

which obviously implies the statement in the second part of the lemma.

For  $x, y \in \mathbb{R}^d$  let

$$\kappa(x, y) := \mathbb{P}\left(\inf_{t \geq 0} |\varphi_t(\cdot, x) - \varphi_t(\cdot, y)| < \varepsilon\right),$$

which is strictly positive by Step 1 and the fact that the function which maps the control  $f$  to the solution is continuous in the supremum norm on  $C([0, T], \mathbb{R}^d)$  for each  $T > 0$ . Further, the map  $(x, y) \mapsto \kappa(x, y)$  is lower semicontinuous since  $\varphi$  is continuous. Therefore,  $\kappa$  is bounded away from 0 uniformly on compact subsets of  $\mathbb{R}^d \times \mathbb{R}^d$ . Now we use the fact that the RDS admits an invariant probability measure  $\rho$ . Let  $K \subset \mathbb{R}^d$  be a compact set satisfying  $\rho(K) > \frac{1}{2}$ . By [43, Lemma 2, Theorem 3] we know that  $\rho$  and all transition probabilities for positive times are absolutely continuous with respect to the Lebesgue measure on  $\mathbb{R}^d$  and therefore

Birkhoff's ergodic theorem shows that for *every* pair  $x, y \in \mathbb{R}^d$  we have that

$$\begin{aligned} & \liminf_{T \rightarrow \infty} \frac{1}{T} \int_0^T 1_{K \times K}(\varphi_t(x), \varphi_t(y)) dt \\ & \geq \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T 1_K(\varphi_t(x)) dt - \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T 1_{K^c}(\varphi_t(y)) dt \\ & = \rho(K) - \rho(K^c) > 0, \quad \mathbb{P}\text{-a.s.} \end{aligned}$$

Hence, for all  $x, y \in \mathbb{R}^d$  and  $n \in \mathbb{N}$  there exists some (random)  $t \geq n$  such that  $(\varphi_t(x), \varphi_t(y)) \in K \times K$  almost surely. Define

$$\kappa := \inf\{\kappa(x, y) : x, y \in K\},$$

which is strictly positive and

$$T(x, y) := \inf\left\{t \geq 0 : \mathbb{P}\left(\inf_{s \in [0, t]} |\varphi_s(\cdot, x) - \varphi_s(\cdot, y)| < \varepsilon\right) \geq \frac{\kappa}{2}\right\}.$$

For fixed  $x, y \in \mathbb{R}^d$ ,  $x \neq y$ , we define the stopping times  $S_n, T_n$  and the random variables  $X_n, Y_n$ ,  $n \in \mathbb{N}_0$  as follows:

$$\begin{aligned} S_0 &:= 0, \quad X_0 := x, \quad Y_0 := y \\ T_n &:= \inf\{t \geq S_n + 1 : (\varphi_t(x), \varphi_t(y)) \in K \times K\} \\ X_n &:= \varphi_{T_n}(x), \quad Y_n := \varphi_{T_n}(y) \\ S_{n+1} &:= \inf\{t \geq T_n : |\varphi_t(x) - \varphi_t(y)| < \varepsilon\} \wedge T(X_n, Y_n). \end{aligned}$$

Note that by the strong Markov property we have

$$\mathbb{P}\left(|\varphi_{S_{n+1}}(x) - \varphi_{S_{n+1}}(y)| \leq \varepsilon \mid \mathcal{F}_{T_n}\right) \geq \kappa/2,$$

for every  $n \in \mathbb{N}_0$  almost surely, so the assertion follows and the proof is complete.  $\square$

**3.4. Summary, explicit examples and open problems.** We close the paper by giving a short summary of general conditions on SDE for synchronization by noise and providing some explicit examples of our general results.

**Theorem 3.14** (Synchronization for general drift). For SDE of the form (1.1), with drift  $b \in C_{loc}^2(\mathbb{R}^d)$  satisfying (3.2), (3.14) and assuming eventual strict monotonicity (cf. Definition 3.5), we have synchronization by noise for sufficiently large noise intensity  $\sigma$ .

*Proof.* First, notice that eventual strict monotonicity implies monotonicity on large sets (cf. Proposition 3.10) as well as (3.9) as observed in Section 3.1.2. By [21] this implies the existence of a (weak) random attractor. Condition (3.2) and the local Lipschitz property, implied by  $b \in C_{loc}^2(\mathbb{R}^d)$ , yield the existence of an RDS  $\varphi$ . By (3.9) and [28, Theorem 4.3],  $\varphi$  is strongly mixing. Assumption (3.15) is a

consequence of (3.9) since, by Itô's formula applied to  $e^{\gamma|X_t|^2}$ , one can even show  $\int_{\mathbb{R}^d} e^{\gamma|x|^2} d\rho(x) < \infty$  for  $\gamma$  small enough. Assumption (3.14), (3.15),  $b \in C_{loc}^2(\mathbb{R}^d)$  and (3.9) imply, via Lemma 3.9, the assumptions of Lemma 3.1 and thus the existence of the Lyapunov spectrum, in particular  $\lambda_{top}$ . By Example 3.6 and the property of eventual strict monotonicity we deduce  $\lambda_{top} < 0$  for large noise intensity  $\sigma$ . By Corollary 3.4 this implies asymptotic stability.

The local Lipschitz property and the monotonicity on large sets guarantee swift transitivity and contraction on large sets, by Proposition 3.10. Then, by Theorem 2.14 we deduce synchronization.  $\square$

For gradient systems,  $b = -\nabla V$ , condition (3.2) on  $b$  can be expressed more naturally as

$$(D^2V(x)\xi, \xi) \geq -\lambda|\xi|^2 \tag{3.23}$$

for all  $x, \xi \in \mathbb{R}^d$  and some  $\lambda \geq 0$ .

**Theorem 3.15** (Synchronization for gradient systems). Consider an SDE of the form (1.3) with gradient drift  $b = -\nabla V$ . Then:

- (1) If  $V \in C^2(\mathbb{R}^d)$  satisfies condition (3.23) and  $e^{-\frac{2}{\sigma^2}V} \in L^1(\mathbb{R}^d)$ , then weak asymptotic stability on  $U$  with  $\rho(U) > 0$  implies weak synchronization.
- (2) Let  $V \in C_{loc}^{2,\delta}(\mathbb{R}^d)$  be such that there exists an  $R > 0$  for which<sup>5</sup>

$$(D^2V(x)\xi, \xi) > 0 \tag{3.24}$$

for all  $|x| > R$  and  $\xi \in \mathbb{R}^d$ . Assume that there is a weak random attractor<sup>6</sup>. Then, asymptotic stability on some non-empty open set  $U$  implies synchronization.

*Proof.* (1): Follows by Theorem 3.12. (2): Condition (3.24) plus local boundedness of  $\|D^2V(x)\|$  imply condition (3.23). Condition (3.23), which implies (3.2) for  $b = -\nabla V$ , and the local Lipschitz property of  $b$ , implied by  $V \in C_{loc}^2(\mathbb{R}^d)$ , yield the existence of an RDS. Then Proposition 3.10 applies because  $b = -\nabla V$  is locally Lipschitz and (3.24) implies monotonicity on large sets. By Theorem 2.14 we deduce synchronization.  $\square$

**Theorem 3.16** (Sufficient conditions for asymptotic stability of gradient systems). Consider an SDE of the form (1.3) with gradient drift  $b = -\nabla V$ . Assume that  $V \in C_{loc}^3(\mathbb{R}^d)$  satisfies conditions (3.14) and (3.23).

<sup>5</sup>Note that this property is implied by (3.25).

<sup>6</sup>The existence of (weak) random attractors under a weak coercivity condition has been obtained in [21].

(1) If  $V$  satisfies condition (3.13) and the positivity assumption on the infimum in Example 3.7, then asymptotic stability holds for small noise intensity  $\sigma > 0$ .

(2) If there exists an  $R > 0$  such that

$$(D^2V(x)\xi, \xi) \geq \lambda_1 |\xi|^2 \quad (3.25)$$

for all  $|x| > R$ ,  $\xi \in \mathbb{R}^d$  and some  $\lambda_1 > 0$ , then asymptotic stability holds for large noise intensity  $\sigma > 0$ .

(3) If  $V(x) = g(|x|^2)$ ,  $g \in C_{loc}^3$  is convex and  $\log^+(|x|)e^{-\frac{2}{\sigma^2}V(x)} \in L^1(\mathbb{R}^d)$ , then asymptotic stability holds.

*Proof.* Conditions (3.23) and  $V \in C_{loc}^3(\mathbb{R}^d, \mathbb{R})$  give us the existence of an RDS, as for (2) of the previous theorem. Assumption (3.15) holds in case (1) by (3.13) (for small  $\sigma$ ); it holds in case (2) since (3.25) implies eventual strict monotonicity and thus (3.9) and we have already remarked in the proof of Theorem 3.14 that this implies (3.15); and finally it holds in case (3) since we assume  $\log^+(|x|)e^{-\frac{2}{\sigma^2}V(x)} \in L^1(\mathbb{R}^d)$ . The assumptions of Lemma 3.9 hold and thus Lemma 3.1 and Corollary 3.4 apply. In particular,  $\lambda_{top}$  exists. In case (3),  $V \in C_{loc}^3(\mathbb{R}^d, \mathbb{R})$  follows from  $g \in C_{loc}^3$  and see the proof of Example 3.8 to realize that Lemma 3.1 applies without assumption (3.13). Asymptotic stability follows by Corollary 3.4 as soon as we prove  $\lambda_{top} < 0$ . Let us recall the proof in three cases. For (1) see Example 3.7 and for (3) see Example 3.8. Finally, for (2), condition (3.25) implies eventual strict monotonicity of  $b$ , hence  $\lambda_{top} < 0$  for large  $\sigma$  by Example 3.6.  $\square$

As in [46] our results apply to

$$\begin{aligned} V_E(x) &:= (0.5 - 10e^{-|x-p_1|^2} - 10e^{-|x-p_2|^2})|x|^2, \\ V_S(x) &:= (2 - 5e^{-|x-p_3|^2} - 6e^{-|x-p_4|^2} - 7e^{-|x-p_5|^2})|x|^2, \end{aligned}$$

where  $p_1 = (0, 1)$ ,  $p_2 = (0, -1)$ ,  $p_3 = (0, 2)$ ,  $p_4 = (2, -2)$ ,  $p_5 = (-2, -2)$ . In contrast to [46], where only small noise  $\sigma$  can be treated, our results also yield synchronization for large noise  $\sigma$ .

As pointed out in the introduction, the model example of a double-well potential

$$V_D(x) = \frac{1}{4}|x|^4 - \frac{1}{2}|x|^2,$$

is not covered by the techniques in [46] for  $d \geq 2$ . In contrast, our results imply synchronization in this case for all  $\sigma > 0$ . In particular, no restriction to small or large noise  $\sigma$  is required here.

We close the paper by pointing out some open problems: In Theorem 3.12 we assumed that weak asymptotic stability holds. We leave it as an open problem whether this condition is always satisfied for gradient type SDE with additive noise

(1.3). Our general results may also be applied to infinite dimensional examples. In particular, synchronization for SPDE could be investigated. This will be subject of subsequent work. Numerical evidence suggests that the top Lyapunov exponent for the Lorentz system perturbed by strong noise (i.e. for  $\sigma$  large) is negative and (weak) synchronization occurs. The Lorentz system, however, is not covered by the techniques put forward in Section 3. We leave this as an open problem. We prove swift transitivity for a large class of SDE with (non-degenerate) additive noise in Section 3. It is left as an open problem to establish swift transitivity in other situations, such as degenerate additive or multiplicative noise.

**Acknowledgement.** We thank the anonymous referees for their careful reading of our manuscript and their many insightful comments and suggestions that helped to improve the presentation of the manuscript and simplify several proofs.

#### REFERENCES

- [1] Ludwig Arnold. *Random dynamical systems*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998.
- [2] Ludwig Arnold and Igor Chueshov. Order-preserving random dynamical systems: equilibria, attractors, applications. *Dynam. Stability Systems*, 13(3):265–280, 1998.
- [3] Ludwig Arnold, Hans Crauel, and Volker Wihstutz. Stabilization of linear systems by noise. *SIAM J. Control Optim.*, 21(3):451–461, 1983.
- [4] Ludwig Arnold and Wolfgang Kliemann. On unique ergodicity for degenerate diffusions. *Stochastics*, 21(1):41–61, 1987.
- [5] Peter H. Baxendale. Statistical equilibrium and two-point motion for a stochastic flow of diffeomorphisms. In *Spatial stochastic processes*, volume 19 of *Progr. Probab.*, pages 189–218. Birkhäuser Boston, Boston, MA, 1991.
- [6] Tomás Caraballo, Igor D. Chueshov, and Peter E. Kloeden. Synchronization of a stochastic reaction-diffusion system on a thin two-layer domain. *SIAM J. Math. Anal.*, 38(5):1489–1507, 2006/07.
- [7] Tomás Caraballo, Hans Crauel, José A. Langa, and James C. Robinson. The effect of noise on the Chafee-Infante equation: a nonlinear case study. *Proc. Amer. Math. Soc.*, 135(2):373–382 (electronic), 2007.
- [8] Tomás Caraballo and James C. Robinson. Stabilisation of linear PDEs by Stratonovich noise. *Systems Control Lett.*, 53(1):41–50, 2004.
- [9] Andrew Carverhill. Flows of stochastic dynamical systems: ergodic theory. *Stochastics*, 14(4):273–317, 1985.
- [10] M. D. Chekroun, E. Simonnet, and M. Ghil. Stochastic climate dynamics: random attractors and time-dependent invariant measures. *Physica D: Nonlinear Phenomena*, 240(21):1685–1700, 2011.
- [11] Xiaopeng Chen, Jinqiao Duan, and Michael Scheutzow. Evolution systems of measures for stochastic flows. *Dyn. Syst.*, 26(3):323–334, 2011.
- [12] Igor Chueshov. *Monotone random systems theory and applications*, volume 1779 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2002.
- [13] Igor Chueshov and Michael K. R. Scheutzow. On the structure of attractors and invariant measures for a class of monotone random systems. *Dyn. Syst.*, 19(2):127–144, 2004.

- [14] Igor Chueshov and Björn Schmalfuß. Master-slave synchronization and invariant manifolds for coupled stochastic systems. *J. Math. Phys.*, 51(10):102702, 23, 2010.
- [15] Hans Crauel. Markov measures for random dynamical systems. *Stochastics Stochastics Rep.*, 37(3):153–173, 1991.
- [16] Hans Crauel. Random point attractors versus random set attractors. *J. London Math. Soc. (2)*, 63(2):413–427, 2001.
- [17] Hans Crauel. *Random probability measures on Polish spaces*, volume 11 of *Stochastics Monographs*. Taylor & Francis, London, 2002.
- [18] Hans Crauel, Georgi Dimitroff, and Michael K. R. Scheutzow. Criteria for strong and weak random attractors. *J. Dynam. Differential Equations*, 21(2):233–247, 2009.
- [19] Hans Crauel and Franco Flandoli. Attractors for random dynamical systems. *Probab. Theory Related Fields*, 100(3):365–393, 1994.
- [20] M. Crucifix. Why could ice ages be unpredictable? *Climate of the Past*, 9:2253–2267, 2013.
- [21] Georgi Dimitroff and Michael K. R. Scheutzow. Attractors and expansion for Brownian flows. *Electron. J. Probab.*, 16:no. 42, 1193–1213, 2011.
- [22] Franco Flandoli, Benjamin Gess, and Michael Scheutzow. Synchronization by noise for order-preserving random dynamical systems. *arXiv:1503.08737*, to appear in: *Ann. Probab.*, 2015. doi:10.1214/16-AOP1088.
- [23] Benjamin Gess. Random attractors for singular stochastic evolution equations. *J. Differential Equations*, 255(3):524–559, 2013.
- [24] Michael Ghil, Mickaël D. Chekroun, and Eric Simonnet. Climate dynamics and fluid mechanics: Natural variability and related uncertainties. *Physica D: Nonlinear Phenomena*, 237(14-17):2111–2126, 2008.
- [25] A. J. Homburg. Synchronization in iterated function systems. *arXiv:1303.6054*, 2013.
- [26] T. Jäger and G. Keller. Random minimality and continuity of invariant graphs in random dynamical systems. *arXiv:1211.5885*, to appear in *Trans. Amer. Math. Soc.*, 2013.
- [27] Gerhard Keller, Haider H. Jafri, and Ram Ramaswamy. Nature of weak generalized synchronization in chaotically driven maps. *Phys. Rev. E*, 87, 2013.
- [28] Rafail Khasminskii. *Stochastic stability of differential equations*, volume 66 of *Stochastic Modelling and Applied Probability*. Springer, Heidelberg, second edition, 2012. With contributions by G. N. Milstein and M. B. Nevelson.
- [29] Nicolai V. Krylov. *Lectures on elliptic and parabolic equations in Hölder spaces*, volume 12 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1996.
- [30] Sergei Kuksin and Armen Shirikyan. *Mathematics of Two-Dimensional Turbulence*. Cambridge University Press, 2012. Cambridge Books Online.
- [31] Sergei B. Kuksin and Armen Shirikyan. On random attractors for systems of mixing type. *Funktsional. Anal. i Prilozhen.*, 38(1):34–46, 2004.
- [32] Yves Le Jan. Équilibre statistique pour les produits de difféomorphismes aléatoires indépendants. *Ann. Inst. H. Poincaré Probab. Statist.*, 23(1):111–120, 1987.
- [33] Vincent Lemaire, Gilles Pagès, and Fabien Panloup. Invariant distribution of duplicated diffusions and application to richardson-romberg extrapolation. *arXiv:1302.1651*, to appear in *Ann. Institut. Henri Poincaré, Probab. Stat.*, 2014.
- [34] Fabio Martinelli, Luca Sbano, and Elisabetta Scoppola. Small random perturbation of dynamical systems: recursive multiscale analysis. *Stochastics Stochastics Rep.*, 49(3-4):253–272, 1994.
- [35] Fabio Martinelli and Elisabetta Scoppola. Small random perturbations of dynamical systems: exponential loss of memory of the initial condition. *Comm. Math. Phys.*, 120(1):25–69, 1988.

- [36] Salah-Eldin A. Mohammed and Michael K. R. Scheutzow. The stable manifold theorem for stochastic differential equations. *Ann. Probab.*, 27(2):615–652, 1999.
- [37] Julian Newman. Necessary and sufficient conditions for stable synchronisation in random dynamical systems. *arXiv:1408.5599*, 2014.
- [38] Gunter Ochs. Weak random attractors. *Report 449, Institut für Dynamische Systeme, Universität Bremen*,, pages 1–18, 1999.
- [39] Arkady Pikovsky, Michael Rosenblum, Jürgen Kurths, and Robert C. Hilborn. Synchronization: A universal concept in nonlinear science. *American Journal of Physics*, 70(6), 2002.
- [40] K. Pyragas. Weak and strong synchronization of chaos. *Phys. Rev. E*, 54:R4508–R4511, 1996.
- [41] David Ruelle. Ergodic theory of differentiable dynamical systems. *Inst. Hautes Études Sci. Publ. Math.*, (50):27–58, 1979.
- [42] Nikolai F. Rulkov, Mikhail M. Sushchik, Lev S. Tsimring, and Henry D. I. Abarbanel. Generalized synchronization of chaos in directionally coupled chaotic systems. *Phys. Rev. E*, 51:980–994, 1995.
- [43] Michael K. R. Scheutzow. Qualitative behaviour of stochastic delay equations with a bounded memory. *Stochastics*, 12(1):41–80, 1984.
- [44] Michael K. R. Scheutzow. Comparison of various concepts of a random attractor: a case study. *Arch. Math. (Basel)*, 78(3):233–240, 2002.
- [45] R. Sirovich, L. Sacerdote, and A. E. P. Villa. Cooperative behavior in a jump diffusion model for a simple network of spiking neurons. *Math. Biosci. Eng.*, 11(2):385–401, 2014.
- [46] Oliver M. Tearne. Collapse of attractors for ODEs under small random perturbations. *Probab. Theory Related Fields*, 141(1-2):1–18, 2008.
- [47] Gerald Teschl. *Ordinary differential equations and dynamical systems*, volume 140 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.
- [48] Srinivasa R. S. Varadhan. *Lectures on diffusion problems and partial differential equations*, volume 64 of *Tata Institute of Fundamental Research Lectures on Mathematics and Physics*. Tata Institute of Fundamental Research, Bombay, 1980. With notes by Pl. Muthuramalingam and Tara R. Nanda.
- [49] Robert Williamson and Ludvik Janos. Constructing metrics with the Heine-Borel property. *Proc. Amer. Math. Soc.*, 100(3):567–573, 1987.

DIPARTIMENTO DI MATEMATICA, LARGO BRUNO PONTECORVO 5, 56127 PISA, ITALY

*E-mail address:* `flandoli@dma.unipi.it`

MAX-PLANCK INSTITUTE FOR MATHEMATICS IN THE SCIENCES, 04103 LEIPZIG, GERMANY

*E-mail address:* `bgess@mis.mpg.de`

INSTITUT FÜR MATHEMATIK, MA 7-5, TECHNISCHE UNIVERSITÄT BERLIN, 10623 BERLIN, GERMANY

*E-mail address:* `ms@math.tu-berlin.de`