SCIENCE CHINA
Mathematics

Special Issue on Differentiable Dynamical Systems
• ARTICLES •

September 2020 Vol. 63 No. 9: 1853–1876
https://doi.org/10.1007/s11425-019-1642-5

Towards mesoscopic ergodic theory

Dedicated with Admiration to the Memory of Professor Shantao Liao

Weiwei Qi[1,2,3], Zhongwei Shen[3], Shirou Wang[3] & Yingfei Yi[3,4,*]

[1]*Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing* 100190*, China;*
[2]*School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing* 100049*, China;*
[3]*Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB T*6*G* 2*G*1*, Canada;*
[4]*School of Mathematics, Jilin University, Changchun* 130012*, China*

Email: wqi2@ualberta.ca, zhongwei@ualberta.ca, shirou@ualberta.ca, yingfei@ualberta.ca

Received September 17, 2019; accepted January 9, 2020; published online August 6, 2020

Abstract The present paper is devoted to a preliminary study towards the establishment of an ergodic theory for stochastic differential equations (SDEs) with less regular coefficients and degenerate noises. These equations are often derived as mesoscopic limits of complex or huge microscopic systems. By studying the associated Fokker-Planck equation (FPE), we prove the convergence of the time average of globally defined weak solutions of such an SDE to the set of stationary measures of the FPE under Lyapunov conditions. In the case where the set of stationary measures consists of a single element, the unique stationary measure is shown to be physical. Similar convergence results for the solutions of the FPE are established as well. Some of our convergence results, while being special cases of those contained in Ji et al. (2019) for SDEs with periodic coefficients, have weaken the required Lyapunov conditions and are of much simplified proofs. Applications to stochastic damping Hamiltonian systems and stochastic slow-fast systems are given.

Keywords ergodic theory, stochastic differential equation, Fokker-Planck equation, stationary measure, physical measure, mesoscopic limit

MSC(2010) 37A10, 35Q84, 35J25, 37B25, 60J60

Citation: Qi W W, Shen Z W, Wang S R, et al. Towards mesoscopic ergodic theory. Sci China Math, 2020, 63: 1853–1876, https://doi.org/10.1007/s11425-019-1642-5

1 Introduction

The present paper aims at investigating ergodic properties of mesoscopic stochastic systems described by stochastic differential equations. Such a system arises in many scientific areas as the mesoscopic limit of a large or complex (deterministic) microscopic system involving both structured variables and noises due to uncertainties, lack of mechanisms, high degrees of freedom, dynamical complexities and so on [42].

The traditional statistical theory of large or complex microscopic systems was established within the framework of ergodic theory of measure-preserving dynamical systems. A fundamental result in this theory is the celebrated *Birkhoff ergodic theorem*. Let $\{P^t\}$ be a flow or semi-flow on a probability

* Corresponding author

space $(X, \mathcal{B}, \mu)$ that preserves the probability measure μ. Birkhoff ergodic theorem states that for each observable $\varphi \in L^1(X, \mathcal{B}, \mu)$, the limit of the time average, namely,

$$\bar{\varphi}(x) := \lim_{t \to \infty} \frac{1}{t} \int_0^t \varphi(P^t(x)) dt$$

exists for μ-a.e. $x \in X$ and in $L^1(X, \mathcal{B}, \mu)$, and satisfies $\int_X \bar{\varphi} d\mu = \int_X \varphi d\mu$. Moreover, if μ is ergodic, then $\bar{\varphi}$ is a μ-a.e. constant, and consequently,

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t \varphi(P^t(x)) dt = \int_X \varphi d\mu, \quad \mu\text{-a.e.} \quad x \in X. \tag{1.1}$$

The Birkhoff ergodic theorem is a fundamental result in ergodic theory whose development can be traced back to the pioneering works in thermodynamics made by Boltzmann, Maxwell and Gibbs in the late 19th through the early 20th century. A thermodynamic system at the microscopic level is an isolated but huge particle system consisting of a massive number (usually in the order of Avogadro number 10^{23}) of ordinary differential equations (ODEs). Realizing the limitations of microscopic approaches in studying such a huge mechanical system, Boltzmann proposed in 1870s a macroscopic (statistical) approach and put forward the so-called ergodic hypothesis, which implies in particular that all microstates on a given energy surface should be equally probable, and the measurement of a macroscopic quantity (entropy, temperature, pressure, etc.) as an observable is not sharp in time but the long time average of individual measurements should be equal to its total phase average. This fundamental hypothesis was not able to be rigorously justified until the early 1930s when Birkhoff and von Neumann respectively proved the pointwise ergodic theorem [4] and the mean ergodic theorem [48]. Although both ergodic theorems were originally established for (continuous) measure-preserving flows on a compact manifold, they were later shown to be valid under more general settings including measure preserving maps and semi-flows, random dynamical systems and stochastic processes.

While the Birkhoff ergodic theorem has demonstrated its significance and power in many areas of dynamical systems and probability theory, it is not sufficient to give an affirmative answer to Boltzmann's ergodic hypothesis as the ergodic measure involved usually does not satisfy the properties demanded in Boltzmann's original proposal. An attempt towards the resolution of the ergodic hypothesis is to look for the so-called *physical measures*. A physical measure μ is such that for any continuous and compactly supported observable $\varphi : X \to \mathbb{R}$, the limit (1.1) holds on a set of points with positive Lebesgue measure. This is clearly not the case when (1.1) only holds for μ-a.e. $x \in X$ with μ being Lebesgue singular. The study of physical measures in systems with dissipation motivates the notion of *Sinai-Ruelle-Bowen* (*SRB*) *measures* defined to be physical measures that are Lebesgue singular and are regular on unstable manifolds of an attractor. Initially introduced for Axiom A attractors and further generalized to systems with certain hyperbolicity or partial hyperbolicity, SRB measures have been at the heart of many fundamental studies in the theory of smooth dynamical systems (see [1,3,11,32,33,44,45,50], and the review article [52] and the references therein). However, there is not much understanding, beyond that for SRB measures, of physical measures for systems which are either less regular or lack of hyperbolic structures. Overall, Boltzmann's ergodic hypothesis remains a challenging open problem even after all the fascinating developments of modern theory of dynamical systems and statistical mechanics, especially given the energetic Kolmogorov-Arnold-Moser (KAM) theorem that asserts the general non-ergodicity of energy surfaces in a nearly integrable Hamiltonian system.

A seemingly promising modern treatment of thermodynamic systems uses the so-called *mesoscopic approach* that situates between microscopic and macroscopic ones, and typically deals with the interaction of a fairly small amount of large particles with a huge amount of small particles. The mesoscopic limit of such a thermodynamic system turns out to be a *stochastic differential equation* (SDE) for the positions of large particles in the form

$$dx = V(x)dt + G(x)dW_t, \quad x \in \mathcal{U} \subseteq \mathbb{R}^d, \tag{1.2}$$

where $\mathcal{U}$ is a connected open set, $V : \mathcal{U} \to \mathbb{R}^d$ is the drift field, $G : \mathcal{U} \to \mathbb{R}^{d\times m}$ $(m \geqslant d)$ is the noise coefficient matrix, and $W = (W_t)_{t\in\mathbb{R}}$ is a standard m-dimensional Wiener process representing the fluctuation resulting from the interaction with the small particles [29]. Such a mesoscopic approach has also been used for stochastic modelings or reductions in many other contexts including complex systems with high degree of freedoms, uncertainties or mechanisms lacking [42], slow-fast systems with fast mixing or chaotic dynamics [14, 15, 24, 36], turbulent fluid flows [31, 41, 46] and the data-driven modeling [17,26,37,39]. In all these applications, especially those arising in fluid mechanics, bio-sciences and data sciences, it is well understood that a trajectory-based microscopic approach can only provide little information on the behaviors of the system, and thus, a macroscopic approach by considering distribution of solutions is necessary to adopt.

In this paper, we use the macroscopic approach to study some of the basic ergodic properties of the mesoscopic system (1.2). In particular, we establish some convergence results for globally defined weak solutions of (1.2) with less regular coefficients and degenerate noises. We emphasize that, as either a mesoscopic limit, a stochastic reduction, or an approximation to a complex microscopic system, the SDE (1.2) does not necessarily admit locally Lipschitz coefficients and non-degenerate noises. As a result, in general, neither can (1.2) be converted into a random dynamical system nor can it generate a measurable flow or semi-flow in the path space. Thus, besides the significance in applications, our present work also contains theoretical novelties to study the ergodic properties of systems without flow or semi-flow structures.

The study of ergodic properties of the SDE (1.2) dates back to the seminal work of Khasminskii [27,28] in the 1960s. Assuming locally Lipschitz coefficients and non-degenerate noises and the existence of an unbounded Lyapunov function associated with (1.2) (see Definition 1.3), Khasminskii established the existence of a unique invariant measure of the diffusion process, which must be a Markov process, generated by the solutions of (1.2) and proved the convergence of the solutions of (1.2) to the invariant measure.

When it comes to less regular coefficients, transition probabilities associated with (1.2) can hardly be defined. This invalidates the methods based on Markov processes giving semi-flow properties. Nevertheless, the SDE (1.2) is naturally connected to the following Fokker-Planck equation (FPE):

$$\partial_t u = \mathcal{L}^* u := \partial_{ij}(a^{ij}u) - \partial_i(V^i u), \quad x \in \mathcal{U}, \quad t > 0, \tag{1.3}$$

where $A = (a^{ij}) := \frac{GG^\top}{2}$ is the diffusion matrix, $V = (V^i)$, $\partial_i = \partial_{x_i}$, $\partial^2_{ij} = \partial^2_{x_i x_j}$, and the usual summation convection is used. Not only does the FPE (1.3) govern the distributions of the solutions of (1.2), but also it has been directly used to model the evolution of the probability distributions of many complex stochastic processes [43]. This suggests an alternative approach that focuses on the FPE (1.3). To study the ergodic properties of (1.3), it is natural to look for stationary solutions of the following stationary problem associated to (1.3):

$$\mathcal{L}^* u = 0 \quad \text{in} \quad \mathcal{U}. \tag{1.4}$$

Denote $\mathcal{L} = a^{ij}\partial^2_{ij} + V^i\partial_i$ as the formal L^2-adjoint of the Fokker-Planck operator $\mathcal{L}^*$. It is also the diffusion operator associated with (1.2).

Definition 1.1 (Stationary measure). A Borel measure μ on $\mathcal{U}$ is called a *measure solution* of (1.4) if

$$a^{ij}, V^i \in L^1_{\mathrm{loc}}(\mathcal{U}, d\mu), \quad \forall\, i, j \in \{1, \ldots, d\}$$

and

$$\int_{\mathcal{U}} \mathcal{L}\phi d\mu = 0, \quad \forall\, \phi \in C_0^2(\mathcal{U}).$$

If, in addition, $\mu(\mathcal{U}) = 1$, then μ is called a *stationary measure* of (1.3).

When (1.2) generates a diffusion process, its invariant measures are necessarily stationary measures of (1.3). But stationary measures of (1.3) could exist even when (1.2) fails to generate a diffusion process. This has been clearly demonstrated by the recent studies on the FPE (1.3) with less regular coefficients.

Indeed, the existence of stationary measures of (1.3) with positive definite $A \in W^{1,p}_{\mathrm{loc}}(\mathcal{U})$ and $V \in L^p_{\mathrm{loc}}(\mathcal{U})$ for some $p > d$ has been established in [19] assuming the existence of a Lyapunov function (see [7–10, 47] for related works). The uniqueness follows from [8] if the Lyapunov function is, in addition, unbounded. As for the ergodic properties, it holds that the solutions of (1.3) strongly converge to the unique stationary measure under an unbounded Lyapunov function [22].

In comparison with the developments of ergodic theory of (1.3) with less regular coefficients in the non-degenerate case, there is not much progress in the degenerate case. One of the main purposes of the present paper is to push forward the theory in the degenerate case. To proceed, we make the following assumption on the coefficients of (1.3):

(H) $a^{ij} \in C(\mathcal{U})$ and $V^i \in C(\mathcal{U})$ for each $i, j \in \{1, \dots, d\}$.

We use Lyapunov conditions to quantify the dissipativity of (1.3).

Definition 1.2 (Compact function). A non-negative function $U \in C(\mathcal{U})$ is called a *compact function* if there is $\rho_M \in (0, \infty]$, called *the essential upper bound*, such that

$$U < \rho_M \quad \text{on} \quad \mathcal{U} \quad \text{and} \quad U(x) \to \rho_M \quad \text{as} \quad x \to \partial\mathcal{U}.$$

We refer the reader to [19, Subsection 2.1] for the meaning of the limit $x \to \partial\mathcal{U}$ that appears in the above definition. For a non-negative function $U \in C(\mathcal{U})$, we denote for each $\rho > 0$, the ρ-sublevel set

$$\Omega_\rho = \{x \in \mathcal{U} : U(x) < \rho\}.$$

Definition 1.3 (Lyapunov function). A compact function $U \in C^2(\mathcal{U})$ with the essential upper bound ρ_M is called

(1) a *Lyapunov function* with respect to $\mathcal{L}$ if there are constants $\gamma > 0$, called a *Lyapunov constant* and $\rho_m > 0$, called an *essential lower bound* such that $\mathcal{L}U \leqslant -\gamma$ in $\mathcal{U} \backslash \overline{\Omega}_{\rho_m}$;

(2) a *strong Lyapunov function* with respect to $\mathcal{L}$ if $\sup_{\mathcal{U}\backslash\Omega_\rho} \mathcal{L}U \to -\infty$ as $\rho \to \rho_M^-$.

If $\rho_M = \infty$, the (strong) Lyapunov function is said to be *unbounded*.

Our main result on the existence of stationary measures of (1.3) is stated in the following theorem.

Theorem 1.4. *Assume* (H). *If $\mathcal{L}$ admits a Lyapunov function, then there exists a stationary measure of* (1.3). *If, in addition, the Lyapunov function is strong, then the set of stationary measures of* (1.3) *is compact under the weak*-topology.*

The existence result in Theorem 1.4 generalizes [6, Corollary 2.4.4(i)], where the problem on the whole space $\mathbb{R}^d$ is considered. The compactness of the set of stationary measures (which is convex) under the strong Lyapunov function ensures the existence of extreme points, which, in consideration of the classical ergodic theory, are expected to enjoy special properties that we explore below.

Let $\mathcal{P}(\mathcal{U})$ be the set of Borel probability measures on $\mathcal{U}$. Consider the following initial condition:

$$\mu_0 = \nu \in \mathcal{P}(\mathcal{U}). \tag{1.5}$$

Definition 1.5 (Global probability solution). A Borel measure μ on $\mathcal{U} \times (0, \infty)$ is called a global probability solution of the Cauchy problem (1.3) and (1.5) if there exists a family of Borel measures $(\mu_t)_{t\in(0,\infty)}$ on $\mathcal{U}$ such that $d\mu = d\mu_t dt$, which satisfy the following properties:

(1) $\mu_t \in \mathcal{P}(\mathcal{U})$ for a.e. $t > 0$;

(2) for every Borel set $B \subset \mathcal{U}$, the mapping $t \mapsto \mu_t(B)$ is measurable;

(3) a^{ij}, $V^i \in L^1_{\mathrm{loc}}(\mathcal{U} \times (0, \infty), d\mu_t dt)$ for each $i, j = 1, \dots, d$; and

(4) for any $\phi \in C^2_0(\mathcal{U})$, there exists a subset $J_\phi \subset (0, \infty)$ with $|(0, \infty) \setminus J_\phi| = 0$ such that

$$\int_{\mathcal{U}} \phi d\mu_t = \int_{\mathcal{U}} \phi d\nu + \lim_{J_\phi \ni r \to 0} \int_r^t \int_{\mathcal{U}} \mathcal{L}\phi d\mu_\tau d\tau, \quad \forall\, t \in J_\phi. \tag{1.6}$$

The existence of global probability solutions of the Cauchy problems (1.3) and (1.5) has been investigated in [38] under Lyapunov conditions. More precisely, Manita and Shaposhnikov [38] proved that if there is an unbounded compact function $U \in C^2(\mathcal{U})$ satisfying $\mathcal{L}U \leqslant C_1U + C_2$ on $\mathcal{U}$ for some $C_1, C_2 > 0$, then the Cauchy problems (1.3) and (1.5) admit a global probability solution. The uniqueness under the same conditions remains unknown. Assuming the uniqueness, we prove in the next result that "ergodic stationary measures" must be extreme points of the convex set of stationary measures.

Let $\mathcal{M}$ be the convex set of stationary measures of (1.3). If (H) holds and $\mathcal{L}$ admits a strong Lyapunov function, then Theorem 1.4 ensures that $\mathcal{M}$ is non-empty and compact. In this case, we denote by $\mathcal{M}^{\mathrm{ex}}$ the set of extreme points of $\mathcal{M}$.

Theorem 1.6. *Assume* (H) *and that* $\mathcal{L}$ *admits an unbounded strong Lyapunov function. Suppose, in addition, that the Cauchy problems* (1.3) *and* (1.5) *admit a unique global probability solution for any* $\nu \in \mathcal{P}(\mathcal{U})$*. For given* $\mu^* \in \mathcal{P}(\mathcal{U})$*, consider the following statement*:

(1) *For any* $f \in C_0(\mathcal{U})$*, it holds that*

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \int_{\mathcal{U}} f d\mu_s^x ds = \int_{\mathcal{U}} f d\mu^*, \quad \mu^*\text{-a.s.} \quad x \in \mathcal{U},$$

where $(\mu_t^x)_{t\in(0,\infty)}$ *is the unique global probability solution of the Cauchy problems* (1.3) *and* (1.5) *with the initial condition* $\nu = \delta_x$.

(2) *If* $\mu_0 \in \mathcal{P}(\mathcal{U})$ *is such that* $\mu_0 \ll \mu^*$*, then*

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \int_{\mathcal{U}} f d\mu_s ds = \int_{\mathcal{U}} f d\mu^*, \quad \forall\, f \in C_0(\mathcal{U}),$$

where $(\mu_t)_{t\in(0,\infty)}$ *is the unique global probability solution of the Cauchy problems* (1.3) *and* (1.5) *with the initial condition* $\nu = \mu_0$.

(3) *It holds that* $\mu^* \in \mathcal{M}^{\mathrm{ex}}$.

Then, (1) *implies* (2)*, which implies* (3).

In Theorem 1.6, we establish a partial link between "ergodic stationary measures" and elements in $\mathcal{M}^{\mathrm{ex}}$. We recall from the standard ergodic theory [49] that there is one-to-one correspondence between ergodic invariant measures and extreme points of the set of invariant measures for dynamical systems on a compact metric space. However, we suspect that the three statements in Theorem 1.6 are not equivalent for the stochastic case we are considering.

To understand the ergodic properties of (1.2) or (1.3), it is crucial to investigate the limits of the time average of their solutions. The corresponding result for global probability solutions of the Cauchy problems (1.3) and (1.5) is stated in the next theorem.

Theorem 1.7. *Assume* (H) *and that* $\mathcal{L}$ *admits a strong Lyapunov function* U*. Let* $\mu = (\mu_t)_{t\in(0,\infty)}$ *be a global probability solution of the Cauchy problems* (1.3) *and* (1.5) *with* $\int_{\mathcal{U}} U d\nu < \infty$*. Then for each sequence* $\{t_n\}_{n\in\mathbb{N}} \subset (0,\infty)$ *with* $\lim_{n\to\infty} t_n = \infty$*, there exist a subsequence, still denoted by* $\{t_n\}_{n\in\mathbb{N}}$*, and a stationary measure* $\tilde{\mu}$ *of* (1.3) *such that*

$$\lim_{n\to\infty} \frac{1}{t_n}\int_0^{t_n} \int_{\mathcal{U}} \phi d\mu_\tau d\tau = \int_{\mathcal{U}} \phi d\tilde{\mu}, \quad \forall\, \phi \in C_b(\mathcal{U}).$$

In particular, if (1.3) *admits a unique stationary measure* $\tilde{\mu}$*, then*

$$\lim_{t\to\infty} \frac{1}{t}\int_0^{t} \int_{\mathcal{U}} \phi d\mu_\tau d\tau = \int_{\mathcal{U}} \phi d\tilde{\mu}, \quad \forall\, \phi \in C_b(\mathcal{U}).$$

Theorem 1.7 allows us to derive convergence results for the time average of globally defined weak solutions of (1.2), whose existence has been studied under various conditions (see, e.g., [16,18,25,30]). In the case where (1.3) admits a unique stationary measure, our next result says that the unique stationary measure is in fact physical.

Theorem 1.8. *Assume* V *and* G *are continuous on* $\mathcal{U}$ *and that* $\mathcal{L}$ *admits a strong Lyapunov function* U*. Let* $(X_t)_{t\geqslant 0}$ *be a globally defined weak solution of* (1.2) *with the initial condition* $X_0 \sim \nu$*, where* $\nu \in \mathcal{P}(\mathcal{U})$

satisfies $\int_{\mathcal{U}} U d\nu < \infty$*. Then for each sequence* $\{t_n\}_{n\in\mathbb{N}} \subset (0,\infty)$ *with* $\lim_{n\to\infty} t_n = \infty$*, there exist a subsequence, still denoted by* $\{t_n\}_{n\in\mathbb{N}}$*, and a stationary measure* $\tilde{\mu}$ *of* (1.3) *such that*

$$\lim_{n\to\infty} \frac{1}{t_n} \int_0^{t_n} \mathrm{E}\phi(X_\tau) d\tau = \int_{\mathcal{U}} \phi d\tilde{\mu}, \quad \forall\, \phi \in C_b(\mathcal{U}).$$

In particular, if (1.3) *admits a unique stationary measure* $\tilde{\mu}$*, then for each* $x \in \mathcal{U}$*, it holds that*

$$\lim_{t\to\infty} \frac{1}{t} \int_0^{t} \mathrm{E}\phi(X_\tau^x) d\tau = \int_{\mathcal{U}} \phi d\tilde{\mu}, \quad \forall\, \phi \in C_b(\mathcal{U}),$$

where $(X_t^x)_{t\geqslant 0}$ *is a globally defined weak solution of* (1.2) *with the initial condition* $X_0^x \sim \delta_x$.

Under the conditions of Theorems 1.7 and 1.8, the diffusion matrix (a^{ij}) is allowed to be degenerate in $\mathcal{U}$, in which case the FPE (1.3) can admit multiple stationary measures. This is why the main part in the statement of Theorem 1.7 (resp. Theorem 1.8) only asserts the average attractiveness of global probability solutions of the Cauchy problems (1.3) and (1.5) (resp. globally defined weak solutions of (1.2)) by the set of stationary measures of (1.3). It is worthwhile to point out that noises are more or less essential for (1.3) to admit a unique stationary measure. In fact, if (1.3) is completely degenerate, namely, $A = 0_{d\times d}$, the FPE (1.3) becomes the Liouville equation associated with the ODE $\dot{x} = V(x)$, which in general admits multiple stationary measures.

The proofs of Theorems 1.7 and 1.8 do not require the standard semi-flow property that plays essential roles in the proof of classical convergence results of SDEs or Markov processes. Indeed, under the assumption (H), the uniqueness of the solutions of the Cauchy problems (1.3) and (1.5) is unknown as mentioned earlier. Even if we assume the uniqueness, they are only known to generate a semi-flow under the weak*-topology. Such weak convergence results, that hold in the absence of the semi-flow property, can potentially serve as theoretical foundations for studying ergodic behaviors of complex systems, such as thermodynamic systems, whose mesoscopic limits are often too rough to admit the standard semi-flow property.

Theorems 1.7 and 1.8 are respectively autonomous versions of [21, Theorems B and 6.4], where (1.2) and (1.3) with time-periodic coefficients are considered. Nevertheless, they have at least two advantages over the corresponding ones in [21]. One is that the strong Lyapunov function assumed in Theorems 1.7 and 1.8 does not need to be unbounded, while the unboundedness is required in [21]. The other is that simplified proofs in the present paper are much easier for the reader to catch the main ideas.

This work can be viewed as an attempt towards the development of ergodic theory for general mesoscopic or stochastic systems. There are many rich subjects in modern ergodic theory that are left open in the stochastic context. For example, when the SDE (1.2) generates a linear random dynamical system [2], the multiplicative ergodic theorem (MET) states that Lyapunov exponents can be used to classify the dynamics of trajectories[1)]. There rises the natural and interesting question: whether Lyapunov exponents for general SDEs of the form (1.2) (in particular with less regular coefficients and degenerate noises) can be defined to better characterize the dynamics in the sense of distribution.

The rest of this paper is organized as follows. In Section 2, we study the existence of stationary measures of (1.3) and prove Theorems 1.4 and 1.6. In Section 3, we investigate the convergence of the time average of the solutions of (1.2) and (1.3), and prove Theorems 1.7 and 1.8. Applications to stochastic damping Hamiltonian systems and stochastic slow-fast systems are discussed in Section 4.

2 Stationary measures

In this section, we study the existence of stationary measures and prove Theorems 1.4 and 1.6. We need the following lemma.

1) Besides the well-known work of Oseledet [40] on MET in 1968, we mention that Liao [34] actually implicitly derived the MET in his exploration of ergodic properties for smooth dynamical systems on compact manifolds. In his paper [34] in 1963, Liao derived a set of qualitative functions on a family of frame bundles which turn out to be Lyapunov exponents [35].

Lemma 2.1. *Assume* (H) *and $\mathcal{L}$ admits a Lyapunov function U with the essential lower bound ρ_m and the essential upper bound ρ_M. Then for any $\rho_0 \in (\rho_m, \rho_M)$, there exists a $C_* \geqslant 0$, depending only on ρ_m and ρ_0, such that for any $\rho_1 \in (\rho_0, \rho_M)$ and any measure solution μ of* (1.4)*, it holds that*

$$\mu(\mathcal{U} \setminus \Omega_{\rho_1}) \leqslant \frac{C}{-\sup_{\mathcal{U}\setminus\Omega_{\rho_1}} \mathcal{L}U}\mu(\Omega_{\rho_1}), \tag{2.1}$$

where $C := C_* \sup_{(\Omega_{\rho_0}\setminus\Omega_{\rho_m})} a^{ij}\partial_i U \partial_j U \geqslant 0$.

Proof. As μ is a measure solution of (1.4), it holds that

$$\int_{\mathcal{U}} \mathcal{L}\phi d\mu = 0, \quad \forall\, \phi \in C_0^2(\mathcal{U}). \tag{2.2}$$

We distinguish between two cases: $\rho_M < \infty$ and $\rho_M = \infty$.

Case $\boldsymbol{\rho_M < \infty}$. Given $\rho_0 \in (\rho_m, \rho_M)$, take $\rho_1 \in (\rho_0, \rho_M)$. Let $\{\zeta_\rho\}_{\rho\in(\rho_1,\rho_M)}$ be a family of non-decreasing and smooth functions on $[0, \rho_M)$ such that

$$\zeta_\rho(t) = \begin{cases} 0, & t \in [0, \rho_m], \\ t, & t \in [\rho_0, \rho], \\ \dfrac{\rho + \rho_M}{4}, & t \in \left(\dfrac{\rho+\rho_M}{2}, \rho_M\right), \end{cases} \quad \text{and} \quad \zeta_\rho'' \leqslant 0 \quad \text{on} \quad \left[\rho, \frac{\rho+\rho_M}{2}\right].$$

In addition, let $\{\zeta_\rho\}_{\rho\in(\rho_1,\rho_M)}$ coincide on $[0, \rho_0]$.

Obviously, $\zeta_\rho(U) - \frac{\rho+\rho_M}{4} \in C_0^2(\mathcal{U})$ for each $\rho \in (\rho_0, \rho_M)$. Setting $\phi = \zeta_\rho(U) - \frac{\rho+\rho_M}{4}$ in (2.2), we find from

$$\mathcal{L}\left(\zeta_\rho(U) - \frac{\rho+\rho_M}{4}\right) = \zeta_\rho'(U)\mathcal{L}U + \zeta_\rho''(U)a^{ij}\partial_i U \partial_j U$$

that

$$\int_{\mathcal{U}} [\zeta_\rho'(U)\mathcal{L}U + \zeta_\rho''(U)a^{ij}\partial_i U\partial_j U] d\mu = 0. \tag{2.3}$$

As $\zeta_\rho' = 0$ on $[0, \rho_m]$, $\zeta_\rho' = 1$ on $[\rho_0, \rho]$ and $\zeta_\rho' \geqslant 0$ otherwise, we find from $\mathcal{L}U \leqslant 0$ in $\mathcal{U} \setminus \Omega_{\rho_m}$ that

$$\zeta_\rho'(U)\mathcal{L}U \leqslant \begin{cases} \sup\limits_{\mathcal{U}\setminus\Omega_{\rho_1}} \mathcal{L}U, & \text{in} \quad \Omega_\rho \setminus \Omega_{\rho_1}, \\ 0, & \text{otherwise}. \end{cases} \tag{2.4}$$

Since (a^{ij}) is semi-positive definite, $\zeta_\rho'' \not\equiv 0$ on $[\rho_m, \rho_0]$, $\zeta_\rho'' \leqslant 0$ on $[\rho, \frac{\rho+\rho_M}{2}]$ and $\zeta_\rho'' = 0$ otherwise, we deduce

$$\zeta_\rho''(U)a^{ij}\partial_i U\partial_j U \leqslant \begin{cases} C_*\Big(\max\limits_{\overline{\Omega}_{\rho_0}} a^{ij}\partial_i U \partial_j U\Big), & \text{in} \quad \Omega_{\rho_0} \setminus \Omega_{\rho_m}, \\ 0, & \text{otherwise}, \end{cases} \tag{2.5}$$

where $C_* := \max_{t\in[\rho_m,\rho_0]} \zeta_\rho''(t) < \infty$ is independent of $\rho \in (\rho_1, \rho_M)$ due to the construction of $\{\zeta_\rho\}_{\rho\in(\rho_1,\rho_M)}$. Applying (2.4) and (2.5) to (2.3), we find

$$\begin{aligned} \Big(-\sup_{\mathcal{U}\setminus\Omega_{\rho_1}} \mathcal{L}U\Big)\mu(\Omega_\rho \setminus \Omega_{\rho_1}) &\leqslant -\int_{\mathcal{U}} \zeta_\rho'(U)\mathcal{L}U d\mu \\ &= \int_{\mathcal{U}} \zeta_\rho''(U)a^{ij}\partial_i U \partial_j U d\mu \\ &\leqslant C_*\Big(\sup_{(\Omega_{\rho_0}\setminus\Omega_{\rho_m})} a^{ij}\partial_i U\partial_j U\Big)\mu(\Omega_{\rho_0} \setminus \Omega_{\rho_m}) \\ &\leqslant C\mu(\Omega_{\rho_1}), \end{aligned}$$

where $C := C_* \sup_{(\Omega_{\rho_0}\setminus\Omega_{\rho_m})} a^{ij}\partial_i U\partial_j U \geqslant 0$. By letting $\rho \to \infty$ in the above inequality, (2.1) follows.

Case $\boldsymbol{\rho_M = \infty}$. Given $\rho_0 \in (\rho_m, \infty)$, take $\rho_1 \in (\rho_0, \infty)$. Let $\{\zeta_\rho\}_{\rho \in (\rho_1, \infty)}$ be a family of non-decreasing and smooth functions on $[0, \infty)$ such that

$$\zeta_\rho(t) = \begin{cases} 0, & t \in [0, \rho_m], \\ t, & t \in [\rho_0, \rho], \\ \rho + 1, & t \in (\rho + 2, \infty), \end{cases} \quad \text{and} \quad \zeta_\rho'' \leqslant 0 \quad \text{on} \quad [\rho, \rho + 2].$$

In addition, let $\{\zeta_\rho\}_{\rho \in (\rho_1, \infty)}$ coincide on $[0, \rho_0]$.

Using the function $\phi = \zeta_\rho(U) - (\rho + 1)$, we can follow the arguments in the case $\rho_M < \infty$ to derive the result. We leave the details to the reader. □

In the rest of this section, we prove Theorems 1.4 and 1.6. We recall the definition for convergence of a family of Borel measures on $\mathcal{U}$ under the weak*-topology.

Definition 2.2. A family of Borel measures $\{\mu_n, n \in \mathbb{N}\}$ on $\mathcal{U}$ is said to converge to some Borel measure μ on $\mathcal{U}$ under the *weak***-topology* if

$$\lim_{n \to \infty} \int_{\mathcal{U}} \phi d\mu_n = \int_{\mathcal{U}} \phi d\mu, \quad \forall\, \phi \in C_0(\mathcal{U}).$$

Proof of Theorem 1.4. Let U be a Lyapunov function with respect to $\mathcal{L}$ with the essential lower bound $\rho_m \geqslant 0$, the essential upper bound $\rho_M > 0$ and the Lyapunov constant $\gamma > 0$. To highlight the dependence of $\mathcal{L}$ on A and V, we write $\mathcal{L}_{A,V}$ for $\mathcal{L}$. The proof is broken into four steps.

Step 1. We construct a candidate measure μ.

By a partition of unity (see, e.g., [23]), there exist a locally finite open cover $(\mathcal{V}_\beta)_{\beta \in \mathcal{B}}$ of $\mathcal{U}$ and functions $(f_\beta)_{\beta \in \mathcal{B}} \subset C_c^\infty(\mathcal{U})$ such that

(1) $\mathrm{supp}(f_\beta) \subset \mathcal{V}_\beta$ for all $\beta \in \mathcal{B}$;

(2) $0 \leqslant f_\beta(x) \leqslant 1$ for all $x \in \mathcal{U}$ and $\beta \in \mathcal{B}$;

(3) $\sum_{\beta \in \mathcal{B}} f_\beta(x) = 1$ for all $x \in \mathcal{U}$.

Let $\eta \in C_0^\infty(\mathbb{R}^d)$ be non-negative and satisfy $\int_{\mathbb{R}^d} \eta dx = 1$. For each $n \in \mathbb{N}$, we define the function $\eta_n : \mathbb{R}^d \to \mathbb{R}$ by setting $\eta_n(x) := \frac{1}{n^d} \eta(nx)$, $x \in \mathbb{R}^d$. Clearly, for each $\beta \in \mathcal{B}$, there is a $k_\beta \in \mathbb{N}$ such that

$$\int_{\mathcal{U}} |a^{ij}(y) - a^{ij}(x)| \eta_n(x - y) dy \leqslant \frac{\gamma}{4}, \quad \forall\, x \in \mathcal{V}_\beta \quad \text{and} \quad n \geqslant k_\beta.$$

As $a^{ij} \in C(\mathcal{U})$ for each $i, j = 1, \ldots, d$, it holds that

$$\int_{\mathcal{U}} [a^{ij}(y) - a^{ij}(x)] \eta_{k_\beta + n}(x - y) dy \to 0 \quad \text{as} \quad n \to \infty, \quad \forall\, x \in \mathcal{V}_\beta.$$

Hence, for each $i, j = 1, \ldots, d$, the function

$$\tilde{a}_n^{ij}(x) := \sum_{\beta \in \mathcal{B}} f_\beta(x) \int_{\mathcal{U}} a^{ij}(y) \eta_{k_\beta + n}(x - y) dy, \quad x \in \mathcal{U},$$

belongs to $C^\infty(\mathcal{U})$ and satisfies

$$|\tilde{a}_n^{ij} - a^{ij}| \leqslant \frac{\gamma}{4} \quad \text{in } \mathcal{U} \quad \text{and} \quad \lim_{n \to \infty} \tilde{a}_n^{ij} = a^{ij} \quad \text{locally uniformly in } \mathcal{U}.$$

Moreover, the matrix $(\tilde{a}_n^{ij})$ is semi-positive definite thanks to that of $A = (a^{ij})$.

For each $n \in \mathbb{N}$, define

$$\epsilon_n(x) := \frac{1}{n} \sum_{\beta \in \mathcal{B}} f_\beta(x) \frac{\gamma}{4(1 + \max_{\overline{\mathcal{V}}_\beta} |D^2 U|)}, \quad x \in \mathcal{U},$$

where D^2U denotes the Hessian of U. Clearly, $\epsilon_n \in C^\infty(\mathcal{U})$ for each $n \in \mathbb{N}$. Moreover,

$$\sup_{n\in\mathbb{N}}\sup_{x\in\mathcal{U}}\left[\epsilon_n(x)\sum_{i=1}^d \partial_{ii}^2 U(x)\right] \leqslant \frac{\gamma}{4} \quad \text{and} \quad \lim_{n\to\infty}\epsilon_n = 0 \quad \text{locally uniformly in } \mathcal{U}.$$

For each $n \in \mathbb{N}$, we define $A_n = (a_n^{ij}) := (\tilde{a}_n^{ij}) + \epsilon_n I$, where I denotes the $d\times d$ identity matrix. Obviously, for each $i,j \in \{1,\dots,d\}$, $a_n^{ij} \in C^\infty(\mathcal{U})$ for all $n\in\mathbb{N}$ and $\lim_{n\to\infty}\tilde{a}_n^{ij} = a^{ij}$ locally uniformly in $\mathcal{U}$. For each $n\in\mathbb{N}$, as the matrix $(\tilde{a}_n^{ij})$ is semi-positive definite and $\epsilon_n > 0$ everywhere, A_n is locally uniformly positive definite. Furthermore, it is not hard to verify that $\mathcal{L}_{A_n,V}U \leqslant -\frac{\gamma}{2}$ in $\mathcal{U}\setminus\overline{\Omega}_{\rho_m}$, i.e., U is a Lyapunov function with respect to $\mathcal{L}_{A_n,V}$ for each $n\in\mathbb{N}$ with a uniform essential lower bound ρ_m and a uniform Lyapunov constant $\frac{\gamma}{2}$.

Applying [19, Theorem A], we find that for each $n\in\mathbb{N}$, there exists a stationary measure μ_n of (1.3) with A replaced by A_n. Note that $\sup_n \mu_n(K) < \infty$ for any compact set $K\subset\mathcal{U}$. We apply [13, Corollary A2.6.V] to conclude the existence of a subsequence, still denoted by $\{\mu_n\}_{n\in\mathbb{N}}$, such that μ_n converges to some σ-finite Borel measure μ on $\mathcal{U}$ under the weak*-topology as $n\to\infty$. The measure μ is the candidate.

Step 2. We show that μ is a measure solution of (1.4).

Since μ_n is a stationary measure of (1.3) with A replaced by A_n, it holds that

$$\int_{\mathcal{U}} \mathcal{L}_{A_n,V}\phi d\mu_n = 0, \quad \forall\,\phi\in C_0^2(\mathcal{U}).$$

Fix $\phi\in C_0^2(\mathcal{U})$. The limit

$$\max_{\mathrm{supp}(\phi)}|\mathcal{L}_{A_n,V}\phi - \mathcal{L}_{A,V}\phi| \to 0 \quad \text{as} \quad n\to\infty$$

implies that

$$\left|\int_{\mathcal{U}}(\mathcal{L}_{A_n,V}\phi - \mathcal{L}_{A,V}\phi)d\mu_n\right| \leqslant |\mathcal{L}_{A_n,V}\phi - \mathcal{L}_{A,V}\phi|_\infty \times \sup_n \mu_n(\mathrm{supp}(\phi)) \to 0 \quad \text{as} \quad n\to\infty.$$

Since $\mathcal{L}_{A,V}\phi \in C_0(\mathcal{U})$ and μ_n converges to μ under the weak*-topology as $n\to\infty$, we find

$$\int_{\mathcal{U}}\mathcal{L}_{A,V}\phi d\mu = \lim_{n\to\infty}\int_{\mathcal{U}}\mathcal{L}_{A,V}\phi d\mu_n.$$

Thus,

$$\int_{\mathcal{U}}\mathcal{L}_{A,V}\phi d\mu = \lim_{n\to\infty}\int_{\mathcal{U}}\mathcal{L}_{A_n,V}\phi d\mu_n = 0.$$

Since $\phi\in C_0^2(\mathcal{U})$ is arbitrary, we conclude that μ is a measure solution of (1.4).

Step 3. We show $\mu(\mathcal{U}) > 0$. Thus, $\tilde{\mu} = \frac{\mu}{\mu(\mathcal{U})}$ is a stationary measure of (1.3).

Fix $\rho_0 \in (\rho_m, \rho_M)$. As $\mathcal{L}_{A_n,V}U \leqslant -\frac{\gamma}{2}$ in $\mathcal{U}\setminus\Omega_{\rho_m}$, we find from Lemma 2.1 that

$$\mu_n(\mathcal{U}\setminus\Omega_{\rho_0}) \leqslant \frac{C_n}{\gamma}\mu_n(\Omega_{\rho_0}),$$

where $C_n := C_*\sup_{(\Omega_{\rho_0}\setminus\Omega_{\rho_m})} a_n^{ij}\partial_i U\partial_j U \geqslant 0$. As A_n converges locally uniformly in $\mathcal{U}$ to A as $n\to\infty$, it holds that

$$\max_{\overline{\Omega}_{\rho_0}} a_n^{ij}\partial_i U\partial_j U \to \max_{\overline{\Omega}_{\rho_0}} a^{ij}\partial_i U\partial_j U \quad \text{as} \quad n\to\infty.$$

Thus, there is $C > 0$, independent of n, such that

$$\mu_n(\mathcal{U}\setminus\Omega_{\rho_0}) \leqslant C\mu_n(\Omega_{\rho_0}), \quad \forall\, n\in\mathbb{N}.$$

Since $\mu_n(\mathcal{U}) = 1$ for each $n\in\mathbb{N}$, the above estimate implies that

$$\mu_n(\overline{\Omega}_{\rho_0}) \geqslant \mu_n(\Omega_{\rho_0}) \geqslant \frac{1}{1+C}, \quad \forall\, n\in\mathbb{N}.$$

As μ_n converges to μ on $\mathcal{U}$ under the weak*-topology, we see from the Portmanteau theorem that

$$\mu(\overline{\Omega}_\rho) \geqslant \limsup_{n\to\infty} \mu_n(\overline{\Omega}_{\rho_0}) \geqslant \frac{1}{1+C} > 0.$$

In particular, $\mu(\mathcal{U}) > 0$.

Step 4. We prove that the set of stationary measures of (1.3) is compact when U is a strong Lyapunov function.

By Lemma 2.1, for any $\rho_1 > \rho_0$ and any stationary measure μ of (1.3), it holds that

$$\mu(\mathcal{U} \setminus \Omega_{\rho_1}) \leqslant \frac{C}{-\sup_{\mathcal{U}\setminus\Omega_{\rho_1}} \mathcal{L}U}$$

for some $C \geqslant 0$, independent of ρ_1 and μ. As

$$\sup_{\mathcal{U}\setminus\Omega_\rho} \mathcal{L}_{A,V}U \to -\infty \quad \text{as} \quad \rho \to \rho_M^-,$$

for any $\epsilon > 0$, there exists a $\rho_* \in (\rho_0, \rho_M)$ such that

$$\mu(\mathcal{U} \setminus \Omega_{\rho_1}) \leqslant \epsilon, \quad \forall\, \rho_1 \in (\rho_*, \rho_M)$$

holds for any stationary measure μ of (1.3). It follows the tightness of the set of stationary measures of (1.3).

It remains to show the closedness of the set of stationary measures of (1.3). Let $\{\mu_n\}_{n\in\mathbb{N}}$ be a sequence of stationary measures of (1.3) that converges to some Borel probability measure μ under the weak*-topology as $n \to \infty$. Note that if $\phi \in C_0^2(\mathcal{U})$, then $\mathcal{L}_{A,V}\phi \in C_0(\mathcal{U})$. It follows that

$$\int_{\mathcal{U}} \mathcal{L}_{A,V}\phi d\mu = \lim_{n\to\infty} \int_{\mathcal{U}} \mathcal{L}_{A,V}\phi d\mu_n = 0, \quad \forall\, \phi \in C_0^2(\mathcal{U}).$$

Thus, μ is a stationary measures of (1.3). This completes the proof. □

The rest of this section is devoted to the proof of Theorem 1.6. We need the following result addressing certain measurability related to the solutions of the Cauchy problems (1.3) and (1.5).

Lemma 2.3. *Assume* (H) *and that* $\mathcal{L}$ *admits an unbounded strong Lyapunov function. Suppose, in addition, that the Cauchy problems* (1.3) *and* (1.5) *admit a unique global probability solution for any* $\nu \in \mathcal{P}(\mathcal{U})$. *Then, for each* $f \in C_0(\mathcal{U})$, *the function* $(x,t) \mapsto \int_{\mathcal{U}} f d\mu_t^x$ *is measurable on* $\mathcal{U} \times (0,\infty)$, *where* $(\mu_t^x)_{t\in(0,\infty)}$ *is the unique global probability solution of the Cauchy problems* (1.3) *and* (1.5) *with the initial condition* $\nu = \delta_x$.

Proof. For each $x \in \mathcal{U}$, let $\mu^x := (\mu_t^x)_{t\in(0,\infty)}$ be as in the statement. By [21, Lemma 4.2] (see also Lemma 3.2 and Remark 3.3), we may assume, without loss of generality, that for each $x \in \mathcal{U}$ and $\phi \in C_0^2(\mathcal{U})$, the function $t \mapsto \int_{\mathcal{U}} \phi d\mu_t^x$ is continuous on $(0,\infty)$.

Since $C_0^2(\mathcal{U})$ is dense in $C_0(\mathcal{U})$, the lemma follows if we can show that the function $(x,t) \mapsto \int_{\mathcal{U}} \phi d\mu_t^x$ is continuous on $\mathcal{U} \times (0,\infty)$ for each $\phi \in C_0^2(\mathcal{U})$. To show this continuity, let us fix arbitrary $\varphi \in C_0^2(\mathcal{U})$, $(x_*,t_*) \in \mathcal{U} \times (0,\infty)$ and $\{(x_n,t_n)\}_{n\in\mathbb{N}} \subset \mathcal{U} \times (0,\infty)$ satisfying $\lim_{n\to\infty}(x_n,t_n) = (x_*,t_*)$. We need to show

$$\lim_{n\to\infty} \int_{\mathcal{U}} \varphi d\mu_{t_n}^{x_n} = \int_{\mathcal{U}} \varphi d\mu_{t_*}^{x_*}.$$

For the validity of the above limit, let $\{(x_{n_j},t_{n_j})\}_{j\in\mathbb{N}}$ be an arbitrary subsequence of $\{(x_n,t_n)\}_{n\in\mathbb{N}}$. It suffices to show the existence of a further subsequence, still denoted by $\{(x_{n_j},t_{n_j})\}_{j\in\mathbb{N}}$, such that

$$\lim_{j\to\infty} \int_{\mathcal{U}} \varphi d\mu_{t_{n_j}}^{x_{n_j}} = \int_{\mathcal{U}} \varphi d\mu_{t_*}^{x_*}. \tag{2.6}$$

It remains to prove (2.6). Since $\sup_n \mu^{x_n}(K) < \infty$ for any compact set $K \subset \mathcal{U} \times (0,\infty)$, we apply [13, Corollary A2.6.V] to conclude that the sequence of the measures $\{\mu^{x_n}\}_{n\in\mathbb{N}}$ is pre-compact under the weak*-topology. As a result, $\{\mu^{x_{n_j}}\}_{j\in\mathbb{N}}$ contains a further subsequence, stilled denoted by $\{\mu^{x_{n_j}}\}_{j\in\mathbb{N}}$, that converges under the weak*-topology to some σ-finite Borel measure μ on $\mathcal{U}\times(0,\infty)$.

We claim that there exists a family of Borel probability measures $\{\mu_t : t>0\}$ on $\mathcal{U}$ such that

$$d\mu = d\mu_t dt \tag{2.7}$$

and

$$\lim_{j\to\infty}\int_{\mathcal{U}}\phi d\mu_{t_{n_j}}^{x_{n_j}} = \int_{\mathcal{U}}\phi d\mu_{t_*}, \quad \forall\,\phi\in C_0^2(\mathcal{U}). \tag{2.8}$$

Fix $\phi\in C_0^2(\mathcal{U})$. For each $j\in\mathbb{N}$, we define the function $F_\phi^j:[0,\infty)\to\mathbb{R}$ by setting

$$F_\phi^j(t) := \int_{\mathcal{U}}\phi d\mu_t^{x_{n_j}} \quad \text{for} \quad t\geqslant 0.$$

Clearly, F_ϕ^j is continuous and satisfies $F_\phi^j(0)=\phi(x_{n_j})$ and $|F_\phi^j(t)|\leqslant|\phi|_\infty$ for all $t\geqslant 0$. It follows from the definition of $\mu^{x_{n_j}}=(\mu_t^{x_{n_j}})_{t\in(0,\infty)}$ that

$$F_\phi^j(t) = F_\phi^j(0) + \lim_{r\to 0}\int_r^t\int_{\mathcal{U}}\mathcal{L}\phi d\mu_s^{x_{n_j}} ds, \quad \forall\,t\geqslant 0, \tag{2.9}$$

which, together with the assumption (H), gives rise to

$$|F_\phi^j(t) - F_\phi^j(r)| \leqslant (t-r)\times\max_{\operatorname{supp}(\phi)}(|a^{ij}\partial_{ij}^2\phi| + |V^i\partial_i\phi|), \quad \forall\,t\geqslant r\geqslant 0. \tag{2.10}$$

Thus, the family $\{F_\phi^j : j\in\mathbb{N}\}$ is uniformly bounded and equicontinuous, and hence, pre-compact under the topology of locally uniform convergence according to the Arzelà-Ascoli theorem and the standard diagonal argument.

Let $\{F_\phi^{j_k}\}_{k\in\mathbb{N}}$ be a subsequence of $\{F_\phi^j\}_{j\in\mathbb{N}}$ that locally uniformly converges to some function $F_\phi\in C([0,\infty))$. In particular, $F_\phi(0)=\lim_{k\to\infty}F_\phi^{j_k}(0)=\phi(x_0)$. For each $\eta\in C_0((0,\infty))$, we see from the dominated convergence theorem that

$$\lim_{k\to\infty}\int_0^\infty\eta(t)\int_{\mathcal{U}}\phi d\mu_t^{x_{n_{j_k}}} dt = \lim_{k\to\infty}\int_0^\infty\eta F_\phi^{j_k} dt = \int_0^\infty\eta F_\phi dt.$$

Moreover, as $\mu^{x_{n_{j_k}}}$ converges to μ under the weak*-topology as $k\to\infty$, it holds that

$$\lim_{k\to\infty}\int_0^\infty\eta(t)\int_{\mathcal{U}}\phi d\mu_t^{x_{n_{j_k}}} dt = \iint_{\mathcal{U}\times(0,\infty)}\eta\phi d\mu, \quad \forall\,\eta\in C_0((0,\infty)).$$

Hence,

$$\int_0^\infty\eta F_\phi dt = \iint_{\mathcal{U}\times(0,\infty)}\eta\phi d\mu, \quad \forall\,\eta\in C_0((0,\infty)). \tag{2.11}$$

Note that F_ϕ is the unique function in $C([0,\infty))$ that satisfies (2.11). This together with the pre-compactness of $\{F_\phi^j\}_{j\in\mathbb{N}}$ yields

$$\lim_{j\to\infty}F_\phi^j = F_\phi \quad \text{locally uniformly on} \quad [0,\infty). \tag{2.12}$$

Arguing as in the proof of [20, Lemma 4.2], we find a family of σ-finite Borel measures $\{\mu_t : t>0\}$ on $\mathcal{U}$ such that

$$F_\phi(t) = \int_{\mathcal{U}}\phi d\mu_t, \quad \forall\,\phi\in C_c^2(\mathcal{U}) \quad \text{and} \quad t>0.$$

It follows from (2.11) that

$$\int_0^\infty\eta(t)\int_{\mathcal{U}}\phi d\mu_t dt = \iint_{\mathcal{U}\times(0,\infty)}\eta\phi d\mu, \quad \forall\,\phi\in C_0^2(\mathcal{U}) \quad \text{and} \quad \eta\in C_0((0,\infty)).$$

Standard approximation arguments lead to

$$\int_0^\infty \int_{\mathcal{U}} \Phi(\cdot,t)d\mu_t dt = \iint_{\mathcal{U}\times(0,\infty)} \Phi d\mu, \quad \forall\, \Phi \in C_0(\mathcal{U}\times(0,\infty)),$$

i.e., $d\mu = d\mu_t dt$. This proves (2.7).

Note that (2.12) gives rise to

$$\lim_{j\to\infty} \int_{\mathcal{U}} \phi d\mu_t^{x_{n_j}} = \int_{\mathcal{U}} \phi d\mu_t, \quad \forall\, \phi \in C_0^2(\mathcal{U}) \quad \text{and} \quad t>0. \tag{2.13}$$

Since

$$\left|\int_{\mathcal{U}} \phi d\mu_{t_{n_j}}^{x_{n_j}} - \int_{\mathcal{U}} \phi d\mu_{t_*}\right| \leqslant \left|\int_{\mathcal{U}} \phi d\mu_{t_{n_j}}^{x_{n_j}} - \int_{\mathcal{U}} \phi d\mu_{t_*}^{x_{n_j}}\right| + \left|\int_{\mathcal{U}} \phi d\mu_{t_*}^{x_{n_j}} - \int_{\mathcal{U}} \phi d\mu_{t_*}\right|,$$

we conclude (2.8) from (2.10) and (2.13).

We further claim that

$$\mu_t = \mu_t^{x_*}, \quad \forall\, t>0. \tag{2.14}$$

Clearly, (2.13) and the Portmanteau theorem imply that

$$\mu_t(\mathcal{U}) \leqslant \liminf_{j\to\infty} \mu_t^{x_{n_j}}(\mathcal{U}) = 1, \quad \forall\, t>0. \tag{2.15}$$

Let us fix $\psi \in C_0(\mathcal{U})$ and let $\{\phi_n\}_{n\in\mathbb{N}} \subset C_0^2(\mathcal{U})$ converge uniformly to ψ. Note that

$$\begin{aligned}
&\left|\int_{\mathcal{U}} \psi d\mu_t^{x_{n_j}} - \int_{\mathcal{U}} \psi d\mu_t\right| \\
&\quad\leqslant \left|\int_{\mathcal{U}} \psi d\mu_t^{x_{n_j}} - \int_{\mathcal{U}} \phi_n d\mu_t^{x_{n_j}}\right| + \left|\int_{\mathcal{U}} \phi_n d\mu_t^{x_{n_j}} - \int_{\mathcal{U}} \phi_n d\mu_t\right| + \left|\int_{\mathcal{U}} \phi_n d\mu_t - \int_{\mathcal{U}} \psi d\mu_t\right| \\
&\quad\leqslant [\mu_t^{x_{n_j}}(\mathcal{U}) + \mu_t(\mathcal{U})] \times \max_{\mathcal{U}} |\phi_n - \psi| + \left|\int_{\mathcal{U}} \phi_n d\mu_t^{x_{n_j}} - \int_{\mathcal{U}} \phi_n d\mu_t\right| \\
&\quad\leqslant 2\max_{\mathcal{U}} |\phi_n - \psi| + \left|\int_{\mathcal{U}} \phi_n d\mu_t^{x_{n_j}} - \int_{\mathcal{U}} \phi_n d\mu_t\right|, \quad \forall\, t>0 \quad \text{and} \quad n\in\mathbb{N},
\end{aligned}$$

where we used $\mu_t^{x_{n_j}}(\mathcal{U}) = 1$ for all $t>0$ and (2.15) in the last inequality. It then follows from (2.13) that

$$\lim_{j\to\infty} \int_{\mathcal{U}} \psi d\mu_t^{x_{n_j}} = \int_{\mathcal{U}} \psi d\mu_t, \quad \forall\, t>0. \tag{2.16}$$

Fix $r,t \in (0,\infty)$ with $r<t$. It follows from (2.9) that

$$\begin{aligned}
\left|F_\phi^j(t) - F_\phi^j(0) - \int_r^t \int_{\mathcal{U}} \mathcal{L}\phi d\mu_s^{x_{n_j}} ds\right| &= \left|\lim_{r'\to 0} \int_{r'}^r \int_{\mathcal{U}} \mathcal{L}\phi d\mu_s^{x_{n_j}} ds\right| \\
&\leqslant r \times \max_{\mathrm{supp}(\phi)} |\mathcal{L}\phi|, \quad \forall\, \phi \in C_0^2(\mathcal{U}).
\end{aligned} \tag{2.17}$$

Since $\mathcal{L}\phi \in C_0(\mathcal{U})$ for each $\phi \in C_0^2(\mathcal{U})$, we find from (2.16) that

$$\lim_{j\to\infty} \int_r^t \int_{\mathcal{U}} \mathcal{L}\phi d\mu_s^{x_{n_j}} ds = \int_r^t \int_{\mathcal{U}} \mathcal{L}\phi d\mu_s ds.$$

Letting $j\to\infty$ in (2.17), we conclude from (2.12) that

$$\left|F_\phi(t) - F_\phi(0) - \int_r^t \int_{\mathcal{U}} \mathcal{L}\phi d\mu_s ds\right| \leqslant r \times \max_{\mathrm{supp}(\phi)} |\mathcal{L}\phi|, \quad \forall\, \phi \in C_0^2(\mathcal{U}),$$

which implies

$$\int_{\mathcal{U}} \phi d\mu_t = \phi(x_0) + \lim_{r\to 0} \int_r^t \int_{\mathcal{U}} \mathcal{L}\phi d\mu_s ds, \quad \forall\, \phi \in C_0^2(\mathcal{U}).$$

Moreover, in the presence of the unbounded strong Lyapunov function, [38, Theorem 2.7] ensures that $\mu_t(\mathcal{U})=1$ for all $t>0$. According to Definition 1.5, $\mu=(\mu_t)_{t>0}$ is the unique global probability solution of the Cauchy problems (1.3) and (1.5) with the initial value $\nu=\delta_{x_*}$, i.e., $\mu_t=\mu_t^{x_*}$ for all $t>0$. This proves (2.14).

Combining (2.8) and (2.14), we conclude in particular (2.6). This completes the proof. $\square$

Proof of Theorem 1.6. (1) $\Rightarrow$ (2). Let μ_0 be as in the statement. Then,

$$\mu_0(B)=\int_{\mathcal{U}}\delta_x(B)d\mu_0(x),\quad B\in\mathcal{B}(\mathcal{U}).$$

It follows from the global well-posedness of the Cauchy problems (1.3) and (1.5) and an approximation argument that

$$\int_{\mathcal{U}} f d\mu_t=\int_{\mathcal{U}}\left[\int_{\mathcal{U}} f d\mu_t^x\right]d\mu_0(x),\quad \forall\, t>0.$$

Fix $f\in C_0(\mathcal{U})$. By Lemma 2.3, the function $(x,t)\mapsto\int_{\mathcal{U}} f d\mu_t^x$ is measurable on $\mathcal{U}\times(0,\infty)$. We apply Fubini's theorem to find that

$$\begin{aligned}\frac{1}{t}\int_0^t\int_{\mathcal{U}} f d\mu_s ds&=\frac{1}{t}\int_0^t\int_{\mathcal{U}}\left[\int_{\mathcal{U}} f d\mu_s^x\right]d\mu_0(x)ds\\&=\frac{1}{t}\int_{\mathcal{U}}\int_0^t\left[\int_{\mathcal{U}} f d\mu_s^x\right]ds d\mu_0(x)\\&=\int_{\mathcal{U}}\left[\frac{1}{t}\int_0^t\int_{\mathcal{U}} f d\mu_s^x ds\right]d\mu_0(x),\quad \forall\, t>0.\end{aligned}$$

Passing to the limit $t\to\infty$, we conclude from (1) and the fact $\mu_0\ll\mu^*$ that

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t\int_{\mathcal{U}} f d\mu_s ds=\int_{\mathcal{U}} f d\mu^*,\quad \forall\, f\in C_0(\mathcal{U}).$$

(2) $\Rightarrow$ (3). It is easy to see that $\mu^*\in\mathcal{M}$. Suppose for contradiction that $\mu^*\notin\mathcal{M}^{\mathrm{ex}}$, i.e., there are $\mu^1,\mu^2\in\mathcal{M}$ with $\mu^1\neq\mu^2$ such that $\mu^*=p\mu^1+(1-p)\mu^2$ for some $p\in(0,1)$. Obviously, $\mu^1,\mu^2\ll\mu^*$. It follows from (2) that

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t\int_{\mathcal{U}} f d\mu_s^1 ds=\lim_{t\to\infty}\frac{1}{t}\int_0^t\int_{\mathcal{U}} f d\mu_s^2 ds,\quad \forall\, f\in C_0(\mathcal{U}),$$

where $(\mu_t^1)_{t\in(0,\infty)}$ and $(\mu_t^2)_{t\in(0,\infty)}$ are the unique global probability solutions of the Cauchy problems (1.3) and (1.5) with the initial conditions $\nu=\mu_0^1$ and $\nu=\mu_0^2$, respectively.

As $\mu_1,\mu_2\in\mathcal{M}$ and the Cauchy problems (1.3) and (1.5) are globally well-posed, it holds that $\mu_t^i=\mu^i$ for all $t\geqslant 0$ and $i=1,2$. As a result,

$$\frac{1}{t}\int_0^t\int_{\mathcal{U}} f d\mu_s^i ds=\int_{\mathcal{U}} f d\mu^i,\quad t>0,\quad i=1,2,$$

which implies that

$$\int_{\mathcal{U}} f d\mu^1=\int_{\mathcal{U}} f d\mu^2,\quad \forall\, f\in C_0(\mathcal{U}),$$

i.e., $\mu^1=\mu^2$. This gives rise to a contradiction. Hence, $\mu^*\in\mathcal{M}^{\mathrm{ex}}$. $\square$

3 Convergence to stationary measures

In this section, we study the weak ergodic theory of (1.2) and (1.3). Theorems 1.7 and 1.8 are respectively proven in Subsections 3.1 and 3.2.

3.1 Convergence in FPEs

We first recall from [21] the definition of continuous modifications of a global probability solution of the Cauchy problems (1.3) and (1.5).

Definition 3.1. Let $\mu = (\mu_t)_{t\in(0,\infty)}$ be a global probability solution of the Cauchy problems (1.3) and (1.5). A Borel measure $\tilde{\mu}$ on $\mathcal{U}\times(0,\infty)$ is called a *continuous modification* of μ if there exists a family of Borel measures $(\mu_t)_{t\in(0,\infty)}$ on $\mathcal{U}$ satisfying the following properties:

(1) $\tilde{\mu}_t = \mu_t$ for a.e. $t\in(0,\infty)$;

(2) the function $t\to\int_{\mathcal{U}}\phi d\tilde{\mu}_t$ is continuous on $(0,\infty)$ for each $\phi\in C_0^2(\mathcal{U})$; and

(3) it holds that the limit $\lim_{t\to 0}\int_{\mathcal{U}}\phi d\tilde{\mu}_t = \int_{\mathcal{U}}\phi d\nu$ for each $\phi\in C_0^2(\mathcal{U})$,

such that $d\tilde{\mu} = d\mu dt$. In this case, we write $\tilde{\mu} = (\tilde{\mu}_t)_{t\in(0,\infty)}$.

It is easy to see from the definition that there exists at most one continuous modification of a global probability solution of the Cauchy problems (1.3) and (1.5). The existence was proved in [21, Lemma 4.2] under the assumption (H). Hence, the following result holds.

Lemma 3.2. *Assume* (H). *Any global probability solution of the Cauchy problems* (1.3) *and* (1.5) *admits a unique continuous modification.*

Remark 3.3. The main advantage of using continuous modifications is as follows. If $\tilde{\mu} = (\tilde{\mu}_t)_{t\in(0,\infty)}$ is a continuous modification of μ, a global probability solution of the Cauchy problems (1.3) and (1.5), then Definition 1.5 implies that for each $\phi\in C_0^2(\mathcal{U})$ it holds that

$$\int_{\mathcal{U}}\phi d\tilde{\mu}_t = \int_{\mathcal{U}}\phi d\nu + \lim_{r\to 0}\int_r^t\int_{\mathcal{U}}\mathcal{L}\phi d\tilde{\mu}_t dt, \quad \forall\, t>0.$$

This allows us to get rid of the sets $\{J_\phi,\phi\in C_0^2(\mathcal{U})\}$ appearing in Definition 1.5.

Next, we deduce an estimate for global probability solutions of the Cauchy problems (1.3) and (1.5).

Lemma 3.4. *Assume* (H) *and $\mathcal{L}$ admits a strong Lyapunov function U with essential upper bound $\rho_M>0$. Let $\{\rho_n\}_{n\in\mathbb{N}}$ increase to ρ_M and $\mu=(\mu_t)_{t\geqslant 0}$ be a global probability solution of the Cauchy problems* (1.3) *and* (1.5). *Then there exists a $C>0$, independent of ν and μ, such that*

$$C_n\int_0^t\tilde{\mu}_\tau(\mathcal{U}\setminus\Omega_{\rho_n})d\tau \leqslant \int_{\mathcal{U}}Ud\nu + Ct, \quad \forall\, t>0 \quad \textit{and} \quad n\in\mathbb{N}, \tag{3.1}$$

where $d\tilde{\mu} = d\tilde{\mu}_t dt$ is the continuous modification of μ, and $C_n := -\sup_{\mathcal{U}\setminus\Omega_{\rho_n}}\mathcal{L}U > 0$.

Proof. We focus on the case $\rho_M<\infty$; the adaptation to the case $\rho_M=\infty$ is straightforward and hence left to the reader (see the proof of Lemma 2.1 for similar treatments).

By Lemma 3.2, μ admits a unique continuous modification, still denoted by $\mu=(\mu_t)_{t\geqslant 0}$. Then, for each $\phi\in C_0^2(\mathcal{U})$ it holds that

$$\int_{\mathcal{U}}\phi d\mu_t = \int_{\mathcal{U}}\phi d\nu + \lim_{r\to 0}\int_r^t\int_{\mathcal{U}}\mathcal{L}\phi d\mu_\tau d\tau, \quad \forall\, t>0. \tag{3.2}$$

Since U is a strong Lyapunov function with respect to $\mathcal{L}$, there is a $\rho_m>0$ such that $\mathcal{L}U\leqslant 0$ on $\mathcal{U}\setminus\overline{\Omega}_{\rho_m}$. Fix a $\rho_0>\rho_m$ and let $\{\zeta_\rho\}_{\rho\in(\rho_0,\rho_M)}$ be a family of smooth non-decreasing functions on $[0,\infty)$ satisfying

$$\zeta_\rho(t) = \begin{cases} 0, & t\in[0,\rho_m],\\ t, & t\in[\rho_0,\rho],\\ \dfrac{\rho+\rho_M}{4}, & t\in\left[\dfrac{\rho+\rho_M}{2},\rho_M\right), \end{cases} \quad \zeta_\rho(t)\leqslant t, \quad t\in[\rho_m,\rho_0] \quad \text{and} \quad \zeta_\rho''\leqslant 0 \quad \text{on} \ \left[\rho,\frac{\rho+\rho_M}{2}\right].$$

In addition, let the functions $\{\zeta_\rho\}_{\rho\in(\rho_0,\rho_M)}$ coincide on $[0,\rho_0]$. Clearly, for each $\rho\in(\rho_0,\rho_M)$, it holds that $\zeta_\rho(t)\leqslant t$ for all $t\geqslant 0$.

Obviously, $\zeta_\rho(U)-\frac{\rho+\rho_M}{4}\in C_0^2(\mathcal{U})$. Setting $\phi=\zeta_\rho(U)-\frac{\rho+\rho_M}{4}$ in (3.2), we find

$$\begin{aligned}\int_{\mathcal{U}}\Big(\zeta_\rho(U)-\frac{\rho+\rho_M}{4}\Big)d\mu_t&=\int_{\mathcal{U}}\Big(\zeta_\rho(U)-\frac{\rho+\rho_M}{4}\Big)d\nu+\lim_{r\to0}\int_r^t\int_{\mathcal{U}}\mathcal{L}\Big(\zeta_\rho(U)-\frac{\rho+\rho_M}{4}\Big)d\mu_\tau d\tau\\&=\int_{\mathcal{U}}\Big(\zeta_\rho(U)-\frac{\rho+\rho_M}{4}\Big)d\nu+\lim_{r\to0}\int_r^t\int_{\mathcal{U}}\mathcal{L}\zeta_\rho(U)d\mu_\tau d\tau.\end{aligned}$$

Since $\mu_t(\mathcal{U})=1$ for a.e. $t\in(0,\infty)$ (a direct consequence of the definition of the continuous modification) and $\mathcal{L}(\zeta_\rho(U))=\zeta_\rho'(U)\mathcal{L}U+\zeta_\rho''(U)a^{ij}\partial_iU\partial_jU$, we find

$$\int_{\mathcal{U}}\zeta_\rho(U)d\mu_t=\int_{\mathcal{U}}\zeta_\rho(U)d\nu+\lim_{r\to0}\int_r^t\int_{\mathcal{U}}\zeta_\rho'(U)\mathcal{L}Ud\mu_\tau d\tau+\lim_{r\to0}\int_r^t\int_{\mathcal{U}}\zeta_\rho''(U)a^{ij}\partial_iU\partial_jUd\mu_\tau d\tau \tag{3.3}$$

holds for a.e. $t\in(0,\infty)$.

The limiting $\lim_{n\to\infty}\rho_n=\rho_M$ and the fact $\rho_0\in(\rho_m,\rho_M)$ yield an $n_0\in\mathbb{N}$ such that $\rho_n>\rho_0$ for all $n>n_0$. Since $\zeta_\rho'=0$ on $[0,\rho_m]$, $\zeta_\rho'=1$ on $[\rho_0,\rho]$ and $\zeta_\rho'\geqslant0$ otherwise, we see from $\mathcal{L}U\leqslant0$ in $\mathcal{U}\setminus\Omega_{\rho_m}$ that

$$\zeta_\rho'(U)\mathcal{L}U\leqslant\begin{cases}\sup\limits_{\mathcal{U}\setminus\Omega_{\rho_n}}\mathcal{L}U, & \text{in}\quad \Omega_\rho\setminus\Omega_{\rho_n},\\ 0, & \text{otherwise.}\end{cases}$$

Thus,

$$\begin{aligned}\lim_{r\to0}\int_r^t\int_{\mathcal{U}}\zeta_\rho'(U)\mathcal{L}Ud\mu_\tau d\tau&\leqslant\sup_{\mathcal{U}\setminus\Omega_{\rho_n}}\mathcal{L}U\times\lim_{r\to0}\int_r^t\mu_\tau(\Omega_\rho\setminus\mathcal{U}_n)d\tau\\&=-C_n\int_0^t\mu_\tau(\Omega_\rho\setminus\Omega_{\rho_n})d\tau,\quad \rho>n_0,\end{aligned}\tag{3.4}$$

where $C_n:=-\sup_{(\mathcal{U}\setminus\Omega_{\rho_n})}\mathcal{L}U>0$ and the monotone convergence theorem is used in the above equality.

Since $\zeta_\rho''\neq0$ on $[\rho_m,\rho_0]$, $\zeta''\leqslant0$ on $[\rho,\frac{\rho+\rho_M}{2}]$ and $\zeta''=0$ otherwise, we find from the non-negative definiteness of (a^{ij}) that

$$\zeta_\rho''(U)a^{ij}\partial_iU\partial_jU\leqslant\begin{cases}C_*\max\limits_{\overline{\Omega}_{\rho_0}}a^{ij}\partial_iU\partial_jU, & \text{in}\quad\Omega_{\rho_0}\setminus\Omega_{\rho_m},\\ 0, & \text{otherwise,}\end{cases}$$

where $C_*:=\max_{t\in[\rho_m,\rho_0]}\zeta_\rho''(t)$ is independent of ρ due to the construction of $\{\zeta_\rho\}_{\rho\in(\rho_0,\rho_M)}$. Hence,

$$\int_r^t\int_{\mathcal{U}}\zeta_\rho''(U)a^{ij}\partial_iU\partial_jUd\mu_\tau d\tau\leqslant C_*\Big(\max_{\overline{\Omega}_{\rho_0}}a^{ij}\partial_iU\partial_jU\Big)\times(t-r)=:C(t-r).\tag{3.5}$$

Substituting (3.4) and (3.5) into (3.3), we find

$$\int_{\mathcal{U}}\zeta_\rho(U)d\mu_t\leqslant\int_{\mathcal{U}}\zeta_\rho(U)d\nu-C_n\int_0^t\mu_\tau(\Omega_\rho\setminus\Omega_{\rho_n})d\tau+Ct,\quad\text{for a.e.}\quad t>0.$$

Since $0\leqslant\zeta_\rho(t)\leqslant t$ for $t\geqslant0$, it follows from the above inequality that

$$C_n\int_0^t\mu_\tau(\Omega_\rho\setminus\Omega_{\rho_n})d\tau\leqslant\int_{\mathcal{U}}Ud\nu+Ct,\quad\text{for a.e.}\quad t>0.$$

Consequently, we pass to the limit $\rho\to\infty$ to find

$$C_n\int_0^t\mu_\tau(\mathcal{U}\setminus\Omega_{\rho_n})d\tau\leqslant\int_{\mathcal{U}}Ud\nu+Ct,\quad\text{for a.e.}\quad t>0.$$

By the monotone convergence theorem, the function $t\mapsto\int_0^t\mu_\tau(\mathcal{U}\setminus\Omega_{\rho_n})d\tau$ is continuous on $(0,\infty)$. Hence, the above inequality holds for all $t>0$.

This completes the proof. □

Finally, we prove Theorem 1.7.

Proof of Theorem 1.7. According to Lemma 3.2, we may replace $\mu = (\mu_t)_{t\in(0,\infty)}$ by its continuous modification, still denoted by $\mu = (\mu_t)_{t\in(0,\infty)}$. Since $\int_{\mathcal{U}} U d\nu < \infty$, it follows from Lemma 3.4 that there exists a $C > 0$ such that

$$C_n \int_0^t \mu_\tau(\mathcal{U} \setminus \Omega_{\rho_n}) d\tau \leqslant \int_{\mathcal{U}} U d\nu + Ct, \quad t > 0, \tag{3.6}$$

where $\lim_{n\to\infty} \rho_n = \rho_M$, $\rho_M > 0$ is the essential upper bound of U and $C_n := -\sup_{\mathcal{U}\setminus\Omega_{\rho_n}} \mathcal{L}U$.

For each $n \in \mathbb{N}$, define $\hat{\mu}_t := \frac{1}{t}\int_0^t \mu_t dt$ for $t > 0$. Then,

$$\int_{\mathcal{U}} \phi d\hat{\mu}_t = \frac{1}{t}\int_0^t \int_{\mathcal{U}} \phi d\mu_\tau d\tau, \quad \forall\, \phi \in C_0(\mathcal{U}). \tag{3.7}$$

Thanks to (3.6), we find

$$\hat{\mu}_t(\mathcal{U} \setminus \Omega_{\rho_n}) = \frac{1}{t}\int_0^t \mu_\tau(\mathcal{U} \setminus \Omega_{\rho_n}) d\tau \leqslant \frac{1}{C_n}\bigg(\int_{\mathcal{U}} U d\nu + C\bigg), \quad t > 1,$$

which together with the limit $\lim_{n\to\infty} C_n = \infty$ ensures the tightness of the family of the probability measures $\{\hat{\mu}_t, t > 1\}$.

Let $\{t_j\}_{j\in\mathbb{N}} \subset (0,\infty)$ satisfy $\lim_{j\to\infty} t_j = \infty$. We apply Prokhorov's theorem to find a subsequence, still denoted by $\{t_j\}_{j\in\mathbb{N}}$, such that $\hat{\mu}_{t_j}$ converges to some Borel probability measure $\tilde{\mu}$ on $\mathcal{U}$ under the weak*-topology as $j \to \infty$.

We show that $\tilde{\mu}$ is a stationary measure of (1.3). Note that for each $\phi \in C_0^2(\mathcal{U})$ it holds that

$$\int_{\mathcal{U}} \phi d\mu_t - \int_{\mathcal{U}} \phi d\mu_s = \int_s^t \int_{\mathcal{U}} \mathcal{L}\phi d\mu_\tau d\tau, \quad \forall\, t > s > 0. \tag{3.8}$$

It follows that

$$\lim_{j\to\infty} \frac{1}{t_j}\int_s^{t_j}\int_{\mathcal{U}} \mathcal{L}\phi d\mu_\tau d\tau = \lim_{j\to\infty}\frac{1}{t_j}\bigg[\int_{\mathcal{U}} \phi d\mu_t - \int_{\mathcal{U}} \phi d\mu_s\bigg] = 0, \quad \forall\, s > 0.$$

This together with (3.7) implies that

$$\lim_{j\to\infty}\int_{\mathcal{U}} \mathcal{L}\phi d\hat{\mu}_{t_j} = \lim_{j\to\infty}\frac{1}{t_j}\bigg[\int_0^s\int_{\mathcal{U}} \mathcal{L}\phi d\mu_\tau d\tau + \int_s^{t_j}\int_{\mathcal{U}} \mathcal{L}\phi d\mu_\tau d\tau\bigg] = 0.$$

By (H), $\mathcal{L}\phi \in C_0(\mathcal{U})$ for all $\phi \in C_0(\mathcal{U})$. Hence,

$$\int_{\mathcal{U}} \mathcal{L}\phi d\tilde{\mu} = \lim_{j\to\infty}\int_{\mathcal{U}} \mathcal{L}\phi d\hat{\mu}^{t_j} = 0, \quad \forall\, \phi \in C_0(\mathcal{U}).$$

By Definition 1.1, $\tilde{\mu}$ is a stationary measure of (1.3).

The "In particular" part follows readily. □

3.2 Convergence in SDEs

Consider the following initial value problem associated to the SDE (1.2):

$$\begin{cases} dx = V(x)dt + G(x)dW_t, \quad x \in \mathcal{U}, \\ x_0 \sim \nu, \end{cases} \tag{3.9}$$

where ν is a given Borel probability measure on $\mathcal{U}$. We assume that V and G are continuous on $\mathcal{U}$.

Recall that a (*globally defined*) *weak solution* of (3.9) is a triple of a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathrm{P})$, an adapted Wiener process $(W_t)_{t\geqslant 0}$ and an adapted stochastic process $(X_t)_{t\geqslant 0}$ such that

$$X_0 \sim \nu, \quad X_t = X_0 + \int_0^t V(X_\tau)d\tau + \int_0^t G(X_\tau)dW_\tau, \quad \forall\, t > 0.$$

In the sequel, we simply call $(X_t)_{t\geqslant 0}$ a weak solution of (3.9) without mentioning the underlying probability space and Wiener process. Recall that $\mathcal{L} := a^{ij}\partial^2_{ij} + V^i\partial_i$.

Lemma 3.5. *Let $(X_t)_{t\geqslant 0}$ be a weak solution of* (3.9) *and μ_t be the distribution of X_t for $t \geqslant 0$. Then, $(\mu_t)_{t\in(0,\infty)}$ is a global probability solution of the Cauchy problems* (1.3) *and* (1.5).

Proof. It is well known [25] that under the current assumptions on the coefficients, $(X_t)_{t\geqslant 0}$ induces a solution of the associated martingale problem. Hence, for each $\phi \in C_0^2(\mathcal{U})$, it holds that

$$\mathrm{E}\phi(X_t) - \mathrm{E}\phi(X_0) - \int_0^t \mathrm{E}[\mathcal{L}\phi(X_\tau)]d\tau = 0, \quad \forall\, t > 0,$$

i.e.,

$$\int_{\mathcal{U}} \phi d\mu_t - \int_{\mathcal{U}} \phi d\nu - \int_0^t \int_{\mathcal{U}} \mathcal{L}\phi d\mu_\tau d\tau = 0, \quad \forall\, t > 0.$$

The lemma then follows from Definition 1.5. □

Proof of Theorem 1.8. By the continuity assumption on V and G, the assumption (H) is satisfied. The theorem then follows directly from Lemma 3.5 and Theorem 1.7. □

4 Applications

In this section, we apply Theorem 1.7 to study the dynamics of Fokker-Planck equations associated with stochastic damping Hamiltonian systems and stochastic slow-fast systems.

4.1 Stochastic damping Hamiltonian systems

We consider the following stochastic damping Hamiltonian system:

$$\begin{cases} dx = ydt, \\ dy = -[b(x,y)y + \nabla V(x)]dt + F(x,y)dt + \sigma(x,y)dW_t, \end{cases} \quad (x,y) \in \mathbb{R}^d \times \mathbb{R}^d, \tag{4.1}$$

where the damping $b = (b^{ij}) : \mathbb{R}^d \times \mathbb{R}^d \mapsto \mathbb{R}^{d\times d}$ is continuous, the potential $V : \mathbb{R}^d \mapsto (0,\infty)$ is twice continuously differentiable, the external forces $F = (F^i) : \mathbb{R}^d \times \mathbb{R}^d \mapsto \mathbb{R}^d$ are continuous, the noise intensity $\sigma : \mathbb{R}^d \times \mathbb{R}^d \mapsto \mathbb{R}^{d\times m}$ belongs to $C(\mathbb{R}^d \times \mathbb{R}^d)$, where $p > d+2$ and $m \geqslant d$ are fixed, and $(W_t)_{t\in\mathbb{R}}$ is the standard m-dimensional Wiener process.

The Fokker-Planck equation associated to (4.1) reads

$$\partial_t u = \partial^2_{y_i y_j}(a^{ij}u) - \partial_{x_i}(y_i u) + \partial_{y_i}((b^{ij}y_j + \partial_{x_i}V - F^i)u), \quad (x,y,t) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}, \tag{4.2}$$

where $(a^{ij}) := \frac{\sigma\sigma^\top}{2}$ is the diffusion matrix. Set $\mathcal{L}_H := \partial_t + a^{ij}\partial^2_{y_i y_j} + y_i\partial_{x_i} - (b^{ij}y_j + \partial_{x_i}V - F^i)\partial_{y_i}$.

We make the following assumptions on the coefficients.

(A1) There is $b_0 > 0$ such that $b^{ij}y_iy_j \geqslant b_0|y|^2$ for all $y \in \mathbb{R}^d$.

(A2) The functions σ and F are uniformly bounded on $\mathbb{R}^d \times \mathbb{R}^d$.

(A3) There exists a lower bounded function $\Phi \in C^2(\mathbb{R}^d)$ such that

$$\sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d} \sum_{i,j=1}^d \left| -b^{ji}(x,y)\frac{x_j}{|x|} + \partial_{x_i}\Phi(x)\right| < \infty.$$

(A4) $\nabla V \cdot \frac{x}{|x|} \to \infty$ as $|x| \to \infty$.

We remark that (A1) says that the system (4.1) is damped. When $b(x,y)$ is bounded, the function Φ in (A3) can be taken to be 0. In addition, (A4) implies $V(x) \to \infty$ as $|x| \to \infty$.

Theorem 4.1. *Assume* (A1)–(A4). *Let $\mu = (\mu_t)_{t\in(0,\infty)}$ be a global probability solution of the Cauchy problem associated with* (4.2) *with the initial condition $\mu_0 = \nu \in \mathcal{M}_p(\mathbb{R}^d \times \mathbb{R}^d)$ being compactly supported. Then for any sequence of positive integers $\{n_j\}_{j\in\mathbb{N}}$ with $\lim_{j\to\infty} n_j = \infty$, there exist a subsequence, still denoted by $\{n_j\}_{j\in\mathbb{N}}$, and a stationary measure $\tilde{\mu}$ of* (4.2) *such that*

$$\lim_{j\to\infty} \frac{1}{t_j}\int_0^{t_j} \int_{\mathcal{U}} \phi d\mu_\tau d\tau = \int_{\mathcal{U}} \phi d\tilde{\mu}, \quad \forall\, \phi \in C_b(\mathbb{R}^d \times \mathbb{R}^d).$$

Proof. The theorem follows from Theorem 1.7 if a strong Lyapunov function with respect to $\mathcal{L}_H$ is established. We follow [12, 51]. Define

$$E(x,y)=\frac{|y|^2}{2}+V(x),\quad (x,y)\in\mathbb{R}^d\times\mathbb{R}^d,$$
$$G(x,y)=\eta(|x|)\frac{x\cdot y}{|x|},\quad (x,y)\in\mathbb{R}^d\times\mathbb{R}^d,$$

where $\eta\in C^\infty([0,\infty))$ satisfies

$$\eta(t)=\begin{cases}0, & t\leqslant \frac{1}{2},\\ 1, & t>1.\end{cases}$$

Let α, $\beta>0$ (to be chosen) and define

$$U(x,y)=\exp\{\alpha E(x,y)+\beta(G(x,y)+\Phi(x))\},\quad (x,y)\in\mathbb{R}^d\times\mathbb{R}^d,$$

where Φ is as in (A3). Clearly, $U\in C^{2,1}(\mathbb{R}^d\times\mathbb{R}^d)$ is positive and satisfies $U(x,y)\to\infty$ as $|x|+|y|\to\infty$, where the lower boundedness of Φ is used.

We compute

$$\begin{aligned}\frac{\mathcal{L}_HU}{U}&=\alpha\mathcal{L}_HE+\beta\mathcal{L}_H(G+\Phi)+a^{ij}(\alpha\partial_{y_i}E+\beta\partial_{y_i}G)(\alpha\partial_{y_j}E+\beta\partial_{y_j}G)\\&=\alpha\mathcal{L}_HE+\beta\mathcal{L}_H(G+\Phi)+a^{ij}\left(\alpha y_i+\beta\frac{x_i}{|x|}\right)\left(\alpha y_j+\beta\frac{x_j}{|x|}\right)\\&=\alpha\mathcal{L}_HE+\beta\mathcal{L}_H(G+\Phi)+\alpha^2a^{ij}y_iy_j+2\alpha\beta a^{ij}\frac{x_iy_j}{|x|}+\beta^2a^{ij}\frac{x_ix_j}{|x|^2},\quad \forall\,|x|>1\quad\text{and}\quad y\in\mathbb{R}^d.\end{aligned}$$

Direct calculations show that

$$\begin{aligned}&\mathcal{L}_HE=-b^{ij}y_iy_j+F^iy_i+\sum_{i=1}^d a^{ii},\\&\mathcal{L}_H\Phi=y_i\partial_{x_i}\Phi,\\&\mathcal{L}_HG=\mathcal{L}_H\left(\frac{x\cdot y}{|x|}\right)=-(b^{ij}y_j+\partial_{x_i}V-F^i)\cdot\frac{x_i}{|x|}+\frac{|y|^2}{|x|}-\frac{x_ix_jy_iy_j}{|x|^3},\quad\forall\,|x|>1\quad\text{and}\quad y\in\mathbb{R}^d.\end{aligned}$$

As $\sum_{ij}x_ix_jy_iy_j=[\sum_i(x_iy_i)]^2\geqslant 0$, we see that

$$\mathcal{L}_HG\leqslant-(b^{ij}y_j+\partial_{x_i}V-F^i)\cdot\frac{x_i}{|x|}+|y|^2,\quad\forall\,|x|>1\quad\text{and}\quad y\in\mathbb{R}^d.$$

Thus,

$$\begin{aligned}\frac{\mathcal{L}_HU}{U}\leqslant{}&\alpha\left(-b^{ij}y_iy_j+F^iy_i+\sum_{i=1}^d a^{ii}\right)\\&+\beta\left[-(b^{ij}y_j+\partial_{x_i}V-F^i)\cdot\frac{x_i}{|x|}+|y|^2+y_i\partial_{x_i}\Phi\right]\\&+\alpha^2a^{ij}y_iy_j+2\alpha\beta a^{ij}\frac{x_iy_j}{|x|}+\beta^2a^{ij}\frac{x_ix_j}{|x|^2},\quad\forall\,|x|>1\quad\text{and}\quad y\in\mathbb{R}^d.\end{aligned}$$

Setting

$$\begin{aligned}&(\mathrm{I}):=-\alpha b^{ij}y_iy_j+\alpha F^iy_i+\beta\left(-b^{ji}\frac{x_j}{|x|}+\partial_{x_i}\Phi\right)y_i+\beta|y|^2+\alpha^2a^{ij}y_iy_j+2\alpha\beta a^{ij}\frac{x_iy_j}{|x|},\\&(\mathrm{II}):=\alpha\sum_{i=1}^d a^{ii}-\beta\partial_{x_i}V\frac{x_i}{|x|}+\beta F^i\frac{x_i}{|x|}+\beta^2a^{ij}\frac{x_ix_j}{|x|^2},\end{aligned}$$

we find $\frac{\mathcal{L}_H U}{U} \leqslant \text{(I)} + \text{(II)}$.

Set

$$M_1 := \sup_{\mathbb{R}^d \times \mathbb{R}^d} \frac{|\sigma\sigma^\top|}{2}, \quad M_2 := \sup_{\mathbb{R}^d \times \mathbb{R}^d} |F| \quad \text{and} \quad M_3 := \sup_{\mathbb{R}^d \times \mathbb{R}^d} \sum_{i,j=1}^{d} \left| - b^{ji}(x,y)\frac{x_j}{|x|} + \partial_{x_i}\Phi(x) \right|.$$

Due to (A2) and (A4), M_1, M_2 and M_3 are finite. For (I), we see from (A1) and the definitions of M_1, M_2 and M_3 that

$$\begin{aligned}\text{(I)} &\leqslant -\alpha b_0|y|^2 + \alpha M_2|y| + \beta M_3|y| + \beta|y|^2 + \alpha^2 M_1|y|^2 + 2\alpha\beta M_1|y| \\ &\leqslant (-\alpha b_0 + \alpha^2 M_1 + \beta)|y|^2 + (\alpha M_2 + \beta M_3 + 2\alpha\beta M_1)|y|.\end{aligned}$$

Let us fix $0 < \alpha < \frac{b_0}{M_1}$ and then choose $\beta > 0$ so small that $-\alpha b_0 + \alpha^2 M_1 + \beta < 0$. It is clear that there exists $\delta_1 > 0$ such that (I) $\leqslant -1$ for all $|y| \geqslant \delta_1$. Similarly, we find

$$\text{(II)} \leqslant \alpha\sqrt{d}M_1 + \beta M_2 + \beta^2 M_1 - \beta\partial_{x_i}V\frac{x_i}{|x|}.$$

It is easy to see from (A4) that there is $\delta_2 > 1$ such that (II) $\leqslant -1$ for $|x| \geqslant \delta_2$. Thus,

$$\frac{\mathcal{L}_H U}{U} \leqslant -2, \quad \forall\,(x,y) \in \{(x,y) : |x| \geqslant \delta_1,\ |y| \geqslant \delta_2\},$$

which together with the fact that $\lim_{|x|\to\infty} U(x) = \infty$ yields $\mathcal{L}_H U(x) \to -\infty$ as $|x| \to \infty$, i.e., U is a strong Lyapunov function with respect to $\mathcal{L}_H$. This completes the proof. □

We remark that the uniqueness of stationary measures of (4.2) is only known when the coefficients are smooth, in which case the theory of hypoellipticity applies. Under the current conditions on the coefficients, it remains an interesting open question.

4.2 Stochastic slow-fast systems

Consider the following SDE:

$$\begin{cases} \epsilon\dot{x} = f(x,y), \\ dy = g(x,y)dt + \sigma(x,y)dW_t, \end{cases} \qquad (x,y) \in \mathbb{R}^m \times \mathbb{R}^n, \tag{4.3}$$

where $0 < \epsilon \ll 1$, $f = (f^k) : \mathbb{R}^m \times \mathbb{R}^n \mapsto \mathbb{R}^m$, $g = (g^i) : \mathbb{R}^m \times \mathbb{R}^n \mapsto \mathbb{R}^n$, $\sigma = (\sigma^{ij}) : \mathbb{R}^m \times \mathbb{R}^n \mapsto \mathbb{R}^{n\times\ell}$ is the *noise coefficient matrix* with $\ell \geqslant n$, and $W = (W_t)_{t\in\mathbb{R}}$ is a standard ℓ-dimensional Wiener process.

As here we are only interested in the dynamics of (4.3) for each fixed $0 < \epsilon \ll 1$, we set $\epsilon = 1$ in (4.3) and consider the following system for clarity:

$$\begin{cases} \dot{x} = f(x,y), \\ dy = g(x,y)dt + \sigma(x,y)dW_t, \end{cases} \qquad (x,y) \in \mathbb{R}^m \times \mathbb{R}^n.$$

The associated FPE reads

$$\partial_t u = \partial^2_{y_i y_j}(a^{ij}u) - \partial_{x_k}(f^k u) - \partial_{y_j}(g^i u), \quad (x,y,t) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}, \tag{4.4}$$

where $A := (a^{ij}) = \frac{1}{2}\sigma\sigma^\top$. Denote $\mathcal{L}_{SF} := a^{ij}\partial^2_{y_i y_j} + f^k\partial_{x_k} + g^i\partial_{y_i}$ as the diffusion operator.

We make the following assumptions on the coefficients.

(B1) $A(x,y)$ is positive definite for each $(x,y) \in \mathbb{R}^m \times \mathbb{R}^n$, and $a^{ij} \in C(\mathbb{R}^m \times \mathbb{R}^n)$ and $g^i \in C(\mathbb{R}^m \times \mathbb{R}^n)$ for each $i,j = 1,\dots,n$. Moreover, for each $a > 0$, it holds that

$$\sup_{B_a}\left(\sum_{i,j}|a^{ij}| + \sum_i |g^i|\right) < \infty,$$

where $B_a := \{(x,y) \in \mathbb{R}^m \times \mathbb{R}^n : |y| \leqslant a\}$.

(B2) There exists some compact function $U \in C^1(\mathbb{R}^m)$ with the essential upper bound $\rho_U > 0$ such that

$$\sup_{|y|\leqslant a} (\mathcal{L}_{SF}U)(x,y) \to -\infty \quad \text{as} \quad |x| \to \infty$$

holds for each $a > 0$, and

$$\begin{aligned} \mathcal{L}_{SF}U = 0 \quad &\text{on} \quad \{0_m\} \times \mathbb{R}^n, \\ \mathcal{L}_{SF}U < 0 \quad &\text{on} \quad (\mathbb{R}^m \setminus \{0_m\}) \times \mathbb{R}^n, \end{aligned}$$

where 0_m denotes the origin in $\mathbb{R}^m$.

We recall from Definition 1.2 the definition of a compact function.

Definition 4.2. A compact function $V \in C^2(\mathbb{R}^n)$ is called

(1) a *semi-Lyapunov function* with respect to $\mathcal{L}_{SF}$ if there exist positive constants γ and a such that

$$\mathcal{L}_{SF}V \leqslant -\gamma \quad \text{in} \quad \mathbb{R}^m \times B_a^c, \tag{4.5}$$

where $B_a^c := \{y \in \mathbb{R}^n : |y| > a\}$;

(2) a *strong semi-Lyapunov function* with respect to $\mathcal{L}_{SF}$ if $\lim_{|y|\to\infty} \mathcal{L}_{SF}V(y) = -\infty$.

If, in addition, the essential upper bound of V is infinity, V is called *unbounded*.

Theorem 4.3. *Assume* (B1), (B2) *and that $\mathcal{L}_{SF}$ admits a semi-Lyapunov function. Then there exists a stationary measure μ of* (4.4) *that satisfies* $\operatorname{supp}(\mu) = \{0_m\} \times \mathbb{R}^n$.

If, in addition, the semi-Lyapunov function is unbounded, then μ is the unique stationary measure of (4.4).

Proof. We write $\mathcal{L}_{SF}$ as $\mathcal{L}$ for notational simplicity. Let V be a semi-Lyapunov function with respect to $\mathcal{L}$ with the essential upper bound $\rho_V > 0$ and $\gamma, a > 0$ be constants such that (4.5) holds. The proof is divided into three steps. To be specific, Steps 1–3 are devoted to the existence of a stationary measure of (4.4) with support $\{0_m\} \times \mathbb{R}^n$. The uniqueness is shown in Step 4 when V is unbounded.

Step 1. We show that (4.4) admits a stationary measure μ. Define

$$W(x,y) := U(x) + V(y), \quad (x,y) \in \mathbb{R}^m \times \mathbb{R}^n.$$

Obviously, W is non-negative and satisfies $\mathcal{L}W = \mathcal{L}U + \mathcal{L}V \leqslant -\gamma$ in $\mathbb{R}^m \times B_a^c$, where we recall that $B_a := \{y \in \mathbb{R}^n : |y| \leqslant a\}$ and $B_a^c := \mathbb{R}^n \setminus B_a$.

It follows from (B1) that $\mathcal{L}V$ is bounded on $\mathbb{R}^m \times B_a$ and from (B2) that $\lim_{|x|\to\infty} \sup_{B_a} \mathcal{L}U = -\infty$. Hence, there is a constant $b > 0$ such that $\mathcal{L}W = \mathcal{L}U + \mathcal{L}V \leqslant -\gamma$ on $\{(x,y) \in \mathbb{R}^m \times \mathbb{R}^n : |y| \leqslant a,\ |x| > b\}$. As a result, we find $\mathcal{L}W \leqslant -\gamma$ on $\{(x,y) \in \mathbb{R}^m \times \mathbb{R}^n : |y| > a \text{ or } |x| > b\}$, i.e., W is a Lyapunov function with respect to $\mathcal{L}$. Then, Theorem 1.4 ensures the existence of a stationary measure μ of (4.4).

Step 2. We show that μ is supported on $\{0_m\} \times \mathbb{R}^n$. Since μ is a stationary measure, we find

$$\iint_{\mathbb{R}^m\times\mathbb{R}^n} \mathcal{L}\phi d\mu = 0, \quad \forall\, \phi \in C_0^2(\mathbb{R}^m \times \mathbb{R}^n). \tag{4.6}$$

For each $\alpha > 1$, we define $W_\alpha(x,y) := \alpha U(x) + V(y)$, $(x,y) \in \mathbb{R}^m \times \mathbb{R}^n$. Recall that the essential upper bounds of U and V are respectively ρ_U and ρ_V. Then, for each $\alpha > 1$ it holds that $W_\alpha(x,y) \to \alpha\rho_U + \rho_V$ as $|x| + |y| \to \infty$.

Let us fix $\alpha > 1$ and set $M_\alpha := \alpha\rho_U + \rho_V$. Let $\{\zeta_\rho\}_{\rho\in(0,M_\alpha)}$ be a family of smooth and non-decreasing functions on $[0, M_\alpha)$ satisfying

$$\zeta_\rho(t) = \begin{cases} t, & t \in [0,\rho], \\ \dfrac{\rho + M_\alpha}{4}, & t \in \left[\dfrac{\rho + M_\alpha}{2}, M_\alpha\right), \end{cases} \quad \text{and} \quad \zeta_\rho'' \leqslant 0 \quad \text{on} \quad \left[\rho, \dfrac{\rho + M_\alpha}{2}\right).$$

Clearly, $\zeta_\rho(W_\alpha) - \frac{1}{4}(\rho + M_\alpha) \in C_0^2(\mathbb{R}^m \times \mathbb{R}^n)$ for each $\rho \in (0, M_\alpha)$. Setting $\phi = \zeta_\rho(W_\alpha) - \frac{1}{4}(\rho + M_\alpha)$ in (4.6), we find from

$$\begin{aligned}\mathcal{L}\zeta_\rho(W_\alpha) &= \zeta_\rho'(W_\alpha)\mathcal{L}W_\alpha + \zeta_\rho''(W_\alpha)a^{ij}\partial_{y_i}W_\alpha\partial_{y_j}W_\alpha \\ &= \zeta_\rho'(W_\alpha)(\alpha\mathcal{L}U + \mathcal{L}V) + \zeta_\rho''(W_\alpha)a^{ij}\partial_{y_i}V\partial_{y_j}V\end{aligned}$$

that

$$\begin{aligned}0 &= \iint_{\mathbb{R}^m\times\mathbb{R}^n} \mathcal{L}\zeta_\rho(W_\alpha)d\mu \\ &= \alpha\iint_{\mathbb{R}^m\times\mathbb{R}^n} \zeta_\rho'(W_\alpha)\mathcal{L}U d\mu + \iint_{\mathbb{R}^m\times\mathbb{R}^n} \zeta_\rho'(W_\alpha)\mathcal{L}V d\mu \\ &\quad + \iint_{\mathbb{R}^m\times\mathbb{R}^n} \zeta_\rho''(W_\alpha)a^{ij}\partial_{y_i}V\partial_{y_j}V d\mu.\end{aligned} \tag{4.7}$$

As $\zeta_\rho' \geqslant 0$ on $[0, \infty]$, (4.5) implies that

$$\zeta_\rho'(W_\alpha)\mathcal{L}V \leqslant \begin{cases} -\gamma\zeta_\rho'(W_\alpha), & (x,y) \in \mathbb{R}^m \times B_a^c, \\ \Big(\max\limits_{\mathbb{R}^m\times B_a} |\mathcal{L}V|\Big)\zeta_\rho'(W_\alpha), & (x,y) \in \mathbb{R}^m \times B_a. \end{cases} \tag{4.8}$$

Since $\zeta_\rho'' \leqslant 0$ on $[\rho, \frac{\rho+M_\alpha}{2}]$ and $\zeta_\rho'' = 0$ otherwise, the non-negative definiteness of (a^{ij}) yields

$$\zeta_\rho''(W_\alpha)a^{ij}\partial_{y_i}V\partial_{y_j}V \leqslant 0. \tag{4.9}$$

Substituting (4.8) and (4.9) into (4.7), we find

$$-\alpha\iint_{\mathbb{R}^m\times\mathbb{R}^n} \zeta_\rho'(W_\alpha)\mathcal{L}U d\mu + \gamma\iint_{\mathbb{R}^m\times B_a^c} \zeta_\rho'(W_\alpha)d\mu \leqslant \Big(\max_{\mathbb{R}^m\times B_a} |\mathcal{L}V|\Big)\iint_{\mathbb{R}^m\times B_a} \zeta_\rho'(W_\alpha)d\mu.$$

As $|\zeta_\rho'| \leqslant 1$, we arrive at

$$-\alpha\iint_{\mathbb{R}^m\times\mathbb{R}^n} \zeta_\rho'(W_\alpha)\mathcal{L}U d\mu \leqslant \Big(\max_{\mathbb{R}^m\times B_a} |\mathcal{L}V|\Big)\iint_{\mathbb{R}^m\times B_a} \zeta_\rho'(W_\alpha)d\mu \leqslant \max_{\mathbb{R}^m\times B_a} |\mathcal{L}V|. \tag{4.10}$$

Note that $\lim_{\rho\to M_\alpha} \zeta_\rho(t) = t$ for each $t \in (0, M_\alpha)$. As a result, $\lim_{\rho\to M_\alpha} \zeta_\rho' = 1$ on $(0, M_\alpha)$. Passing to the limit $\rho \to M_\alpha^-$ in (4.10), we deduce

$$-\alpha\iint_{\mathbb{R}^m\times\mathbb{R}^n} \mathcal{L}U d\mu \leqslant \max_{\mathbb{R}^m\times B_a} |\mathcal{L}V|.$$

To see $\operatorname{supp}(\mu) \subset \{0_m\} \times \mathbb{R}^n$, we suppose on the contrary that there exists a closed set $B \subset \mathbb{R}^m$ satisfying $0_m \notin B$ such that $\mu(B \times \mathbb{R}^n) > 0$. It follows from (B2) that $\sup_{B\times\mathbb{R}^n} \mathcal{L}U < 0$, which results in

$$-\alpha\Big(\sup_{B\times\mathbb{R}^n} \mathcal{L}U\Big)\mu(B \times \mathbb{R}^n) \leqslant \max_{\mathbb{R}^m\times B_a} |\mathcal{L}V|.$$

A contradiction is derived by letting $\alpha \to \infty$ in the above inequality.

Step 3. We show that $\operatorname{supp}(\mu) = \{0_m\} \times \mathbb{R}^n$. Define

$$\mu_*(B) := \mu(\{0_m\} \times B), \quad \forall\, B \in \mathcal{B}(\mathbb{R}^n),$$

where $\mathcal{B}(\mathbb{R}^n)$ is the Borel σ-algebra of $\mathbb{R}^n$. We further define $\mathcal{L}_0 := \alpha^{ij}\partial^2_{y_iy_j} + \beta^i\partial_{y_i}$, where $\alpha^{ij}(y) = a^{ij}(0_m, y)$ and $\beta^i(y) := g^i(0_m, y)$ for $y \in \mathbb{R}^n$ and $i, j = 1, \ldots, n$.

As μ is a stationary measure of (4.4) and supported on $\{0_m\} \times \mathbb{R}^n$, it holds that $\mu_*(\mathbb{R}^n) = 1$ and $\int_{\mathbb{R}^n} \mathcal{L}_0\phi d\mu_* = 0$ for all $\phi \in C_0^2(\mathbb{R}^n)$, i.e., μ_* is a stationary measure of the following FPE:

$$\partial_t u = \partial^2_{y_iy_j}(\alpha^{ij}u) - \partial_{y_i}(\beta^i u), \quad y \in \mathbb{R}^n. \tag{4.11}$$

By (B1), (α^{ij}) is pointwise positive definite on $\mathbb{R}^n$. It follows from [5] that μ_* admits a positive density on $\mathbb{R}^n$, which implies $\mathrm{supp}(\mu_*) = \mathbb{R}^n$. Equivalently, $\mathrm{supp}(\mu) = \{0_m\} \times \mathbb{R}^n$.

Step 4. As V is indeed an unbounded Lyapunov function with respect to $\mathcal{L}_0$, we can follow [10, Example 5.1] to argue that μ_* is the unique stationary measure of (4.11). As a result, μ is the unique stationary measure of (4.4). □

A convergence result can be established if $\mathcal{L}_{SF}$ admits an unbounded strong semi-Lyapunov function.

Theorem 4.4. *Assume* (B1), (B2) *and* $\mathcal{L}_{SF}$ *admits an unbounded strong semi-Lyapunov function. Then for any global probability solution* $\mu = (\mu_t)_{t>0}$ *of the Cauchy problem associated with* (4.4) *with the initial condition* $\mu_0 = \nu$, *where* $\nu \in \mathcal{P}(\mathbb{R}^m \times \mathbb{R}^n)$ *is compactly supported, it holds that*

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \iint_{\mathbb{R}^m\times\mathbb{R}^n} \phi d\mu_\tau d\tau = \iint_{\mathbb{R}^m\times\mathbb{R}^n} \phi d\tilde{\mu}, \quad \forall \phi \in C_0(\mathbb{R}^m \times \mathbb{R}^n),$$

where $\tilde{\mu}$ *is the unique stationary measure of* (4.4) *and satisfies* $\mathrm{supp}(\tilde{\mu}) = \{0_m\} \times \mathbb{R}^n$.

Proof. Let V be the unbounded strong semi-Lyapunov function with respect to $\mathcal{L}_{SF}$. Then it is easy to show that $W(x,y) := U(x) + V(y)$, $(x,y) \in \mathbb{R}^m \times \mathbb{R}^n$ is a strong Lyapunov function with respect to $\mathcal{L}_{SF}$, where U is given in (B2). The conclusion then follows from Theorems 1.7 and 4.3. □

Acknowledgements The first author was supported by China Scholarship Council. The second author was supported by University of Alberta, and Natural Sciences and Engineering Research Council of Canada (Grant Nos. RGPIN-2018-04371 and DGECR-2018-00353). The third author was supported by Pacific Institute for the Mathematical Sciences-Canadian Statistical Sciences Institute Postdoctoral Fellowship, Pacific Institute for the Mathematical Sciences-Collaborative Research Group Grant, National Natural Science Foundation of China (Grant Nos. 11771026 and 11471344) and the Pacific Institute for the Mathematical Sciences-University of Washington site through National Science Foundation of USA (Grant No. DMS-1712701). The fourth author was supported by Natural Sciences and Engineering Research Council of Canada Discovery (Grant No. 1257749), Pacific Institute for the Mathematical Sciences-Collaborative Research Group Grant, University of Alberta, and Jilin University. The authors express their sincere thanks to the anonymous referees for carefully reading the manuscript and providing invaluable suggestions, and their special thanks to one of the anonymous referees for pointing out a gap in the proof of Theorem 1.6.

References

1 Alves J F, Bonatti C, Viana M. SRB measures for partially hyperbolic systems whose central direction is mostly expanding. Invent Math, 2000, 140: 351–398

2 Arnold L. Random Dynamical Systems. Springer Monographs in Mathematics. Berlin: Springer-Verlag, 1998

3 Benedicks M, Young L-S. Sinai-Bowen-Ruelle measures for certain Hénon maps. Invent Math, 1993, 112: 541–576

4 Birkhoff G D. Proof of the ergodic theorem. Proc Natl Acad Sci USA, 1931, 17: 656–660

5 Bogachev V I, Krylov N V, Röckner M. On regularity of transition probabilities and invariant measures of singular diffusions under minimal conditions. Comm Partial Differential Equations, 2001, 26: 2037–2080

6 Bogachev V I, Krylov N V, Röckner M, et al. Fokker-Planck-Kolmogorov Equations. Mathematical Surveys and Monographs, vol. 207. Providence: Amer Math Soc, 2015

7 Bogachev V I, Rökner M. A generalization of Khasminskii's theorem on the existence of invariant measures for locally integrable drifts. Theory Probab Appl, 2002, 45: 363–378

8 Bogachev V I, Rökner M, Shaposhnikov S V. On uniqueness problems related to elliptic equations for measures. J Math Sci (NY), 2011, 176: 759–773

9 Bogachev V I, Rökner M, Shaposhnikov S V. On positive and probability solutions to the stationary Fokker-Planck-Kolmogorov equation. Dokl Math, 2012, 85: 350–354

10 Bogachev V I, Röckner M, Stannat V. Uniqueness of solutions of elliptic equations and uniqueness of invariant measures of diffusions. Sb Math, 2002, 193: 945–976

11 Bowen R. Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms. Lecture Notes in Mathematics, vol. 470. Berlin-Heidelberg: Springer, 1975

12 Cattiaux P, León J R, Pineda Centeno A A, et al. An overlook on statistical inference issues for stochastic damping hamiltonian systems under the fluctuation-dissipation condition. Statistics, 2017, 51: 11–29

13 Daley D J, Vere-Jones D. An Introduction to the Theory of Point Processes Volume I: Elementary Theory and Methods, 2nd ed. New York-Berlin-Heidelberg: Springer-Verlag, 2003

14 de Simoi J, Liverani C. Limit theorems for fast-slow partially hyperbolic systems. Invent Math, 2018, 213: 811–1016

15 de Simoi J, Liverani C, Poquet C, et al. Fast-slow partially hyperbolic systems versus Freidlin-Wentzell random systems. J Stat Phys, 2017, 166: 650–679

16 Ethier S N, Kurtz T G. Markov Processes: Characterization and Convergence. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. New York: John Wiley & Sons, 1986

17 Giannakis D, Majda A J. Data-driven methods for dynamical systems: Quantifying predictability and extracting spatiotemporal patterns. In: Mathematical and Computational Modeling. Pure and Applied Mathematics. Hoboken: Wiley, 2015, 137–191

18 Hofmanová M, Seidler J. On weak solutions of stochastic differential equations II. Stoch Anal Appl, 2013, 31: 663–670

19 Huang W, Ji M, Liu Z, et al. Steady states of Fokker-Planck equations: I. Existence. J Dynam Differential Equations, 2015, 27: 721–742

20 Ji M, Qi W, Shen Z, et al. Existence of periodic probability solutions to Fokker-Planck equations with applications. J Funct Anal, 2019, 277: 108281

21 Ji M, Qi W, Shen Z, et al. Convergence to periodic probability solutions in Fokker-Planck equations. https://sites.ualberta.ca/~zhongwei/manuscript-Ji-Qi-Shen-Yi-convergence.pdf, 2019

22 Ji M, Shen Z, Yi Y. Convergence to equilibrium in Fokker-Planck equations. J Dynam Differential Equations, 2019, 31: 1591–1615

23 Jost J. Riemannian Geometry and Geometric Analysis, 7th ed. Cham: Springer, 2017

24 Just W, Kantz H, Rödenbeck C, et al. Stochastic modelling: Replacing fast degrees of freedom by noise. J Phys A, 2001, 34: 3199–3213

25 Karatzas I, Shreve S E. Brownian Motion and Stochastic Calculus, 2nd ed. Graduate Texts in Mathematics, vol. 113. New York: Springer-Verlag, 1991

26 Katsoulakis M A, Vlachos D G. Coarse-grained stochastic processes and kinetic Monte Carlo simulators for the diffusion of interacting particles. J Chem Phys, 2003, 119: 9412–9427

27 Khasminskii R Z. Ergodic properties of recurrent diffusion processes and stabilization of the solution of the Cauchy problem for parabolic equations (in Russian). Teor Veroyatn Primen, 1960, 5: 196–214

28 Khasminskii R Z. Stochastic Stability of Differential Equations, 2nd ed. Stochastic Modelling and Applied Probability, vol. 66. Heidelberg: Springer, 2012

29 Kotelenez P. Stochastic Ordinary and Stochastic Partial Differential Equations: Transition from Microscopic to Macroscopic Equations. Stochastic Modelling and Applied Probability, vol. 58. New York: Springer, 2008

30 Kunze M. Stochastic differential equations. https://www.uni-ulm.de/fileadmin/website_uni_ulm/mawi.inst.020/kunze/SDE/sde_skript.pdf, 2012

31 Le Bris C, Lions P-L. Existence and uniqueness of solutions to Fokker-Planck type equations with irregular coefficients. Comm Partial Differential Equations, 2008, 33: 1272–1317

32 Ledrappier F, Young L-S. The metric entropy of diffeomorphisms, I: Characterization of measures satisfying Pesin's entropy formula; II: Relations between entropy, exponents and dimension. Ann of Math (2), 1985, 122: 509–539; 540–574

33 Ledrappier F, Young L-S. Entropy formula for random transformations. Probab Theory Related Fields, 1988, 80: 217–240

34 Liao S. Certain ergodic properties of a differential system on a compact differentiable manifold. Acta Sci Natur Univ Pekinensis, 1963, 9: 241–265, 309–326

35 Liao S. On characteristic exponents: Construction of a new Borel set for the multiplicative ergodic theorem for vector fields. Beijing Daxue Xuebao Ziran Kexue Ban, 1993, 29: 277–302

36 MacKay R S. Langevin equation for slow degrees of freedom of Hamiltonian systems. In: Nonlinear Dynamics and Chaos: Advances and perspectives. Understanding Complex Systems. Berlin: Springer, 2010, 89–102

37 Majda A J, Wang X. Non-Linear Dynamics and Statistical Theories for Basic Geophysical Flows. Cambridge: Cambridge University Press, 2006

38 Manita O A, Shaposhnikov V S. On the Cauchy problem for Fokker-Planck-Kolmogorov equations with potential terms on arbitrary domains. J Dynam Differential Equations, 2016, 28: 493–518

39 Nadler B, Lafon S, Coifman R R, et al. Diffusion maps, spectral clustering and reaction coordinates of dynamical systems. Appl Comput Harmon Anal, 2006, 21: 113–127

40 Oseledec V I. A multiplicative ergodic theorem: Lyapunov characteristic number for dynamical systems. Trans Moscow Math Soc, 1968, 19: 197–231

41 Pope S B. Simple models of turbulent flows. Phys Fluids, 2011, 23: 011301

42 Qian H, Ao P, Tu Y, et al. A framework towards understanding mesoscopic phenomena: Emergent unpredictability, symmetry breaking and dynamics across scales. Chem Phys Lett, 2016, 665: 153–161

43 Risken H. The Fokker-Planck Equation: Methods of Solution and Applications, 2nd ed. Springer Series in Synergetics, vol. 18. Berlin: Springer-Verlag, 1989

44 Ruelle D. A measure associated with Axiom-A attractors. Amer J Math, 1976, 98: 619–654

45 Sinai Y G. Gibbs measures in ergodic theory (in Russian). Uspekhi Mat Nauk, 1972, 27: 21–69

46 Sreenivasan K R, Antonia R A. The phenomenology of small-scale turbulence. Annu Rev Fluid Mech, 1997, 29: 435–472

47 Veretennikov A Y. On polynomial mixing bounds for stochastic differential equations. Stochastic Process Appl, 1997, 70: 115–127

48 von Neumann J. Proof of the quasi-ergodic hypothesis. Proc Natl Acad Sci USA, 1932, 18: 70–82

49 Walters P. An Introduction to Ergodic Theory. Graduate Texts in Mathematics, vol. 79. New York-Berlin: Springer-Verlag, 1982

50 Wang Q, Young L. Toward a theory of rank one attractors. Ann of Math (2), 2008, 167: 349–480

51 Wu L. Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. Stochastic Process Appl, 2001, 91: 205–238

52 Young L-S. What are SRB measures, and which dynamical systems have them? J Stat Phys, 2002, 108: 733–754