AVERAGING PRINCIPLE FOR STOCHASTIC DIFFERENTIAL EQUATIONS IN THE RANDOM PERIODIC REGIME

KENNETH UDA

Department of Mathematics, University of Edinburgh, Scotland, UK

ABSTRACT. We present the validity of stochastic averaging principle for non-autonomous slow-fast stochastic differential equations (SDEs) whose fast motions admit random periodic solutions. Our investigation is motivated by some problems arising from multi-scale stochastic dynamical systems, where configurations are time dependent due to nonlinearity of the underlying vector fields and the onset of time dependent random invariant sets. Averaging principle with respect to uniform ergodicity of the fast motion is no longer available in this scenario. Lyapunov second method together with synchronous coupling and strong Feller property of Markovian flows of SDEs are used to prove the ergodicity of time periodic measures of the fast motion on certain minimal Poincaré section and consequently identify the averaging limit.

Keywords: Periodic measures; PS-ergodicity; random periodic solutions; stochastic averaging; Hömander’s Lie bracket conditions; strong Feller property.

Contents

1. Introduction, notations and setup 1
   1.1. Introduction  1
   1.2. Notations     3
   1.3. The averaging setup  4
2. Random periodicity  7
   2.1. Random periodic solutions  8
   2.2. Existence of stable random periodic solutions  9
   2.3. Ergodic periodic measures on Poincaré sections  15
3. Stochastic averaging in the random periodic regime  30
   3.1. The averaged equation  30
   3.2. The Averaging limit  33
4. Conclusions and future work  39
References  39

1. Introduction, notations and setup

1.1. Introduction. Many nonlinear deterministic and stochastic dynamical models in science and engineering consist of motions at disparate spatiotemporal scales, giving rise to slow-fast systems. Multi-scale methods may be employed to analyse some of the desirable structures of these slow-fast motions in the long run. Multi-scale methods are well-established theoretical
and computational techniques in nonlinear stochastic dynamical systems and related areas, its applications range from signal processing, climate dynamics, material science, mathematical finance, molecular dynamics, etc. Averaging principle is an important multi-scale technique, it suggests that the slow subsystem of a slow-fast system may be approximated by a simpler system obtained by averaging over the fast motion. The active usage of various averaging could be traced back to 18th century as a well-known approximation procedure in celestial mechanics, for example, in the study of perturbation of planetary motion from the periodic variation in the orbital elements. Rigorous results justifying averaging principle go back to series of papers by Krylov, Bogolyubov and Mitropol’skii (e.g.,[6]). Generalisation of averaging principle to chaotic and stochastic dynamical systems with important applications in science and engineering are among the most investigated problems in nonlinear dynamical systems and related areas (see e.g., [28, 18, 31, 44]). In probabilistic language, averaging principle is a variant of weak law of large numbers.

The primary objective of the present paper is to investigate averaging principle in the random dynamical systems (RDS) framework (e.g., [3, 31, 32, 33, 34, 44, 50]) for fully coupled slow-fast SDEs with time dependent vector fields. Our investigation stems from some problems arising from long time behaviour of stochastic dynamical systems, where dynamical configurations are time dependent due to the existence of time dependent random invariant sets. Models whose evolution rule are time dependent slow-fast SDEs are particularly natural in numerous applications and are speedily becoming cradle for the development new mathematical ideas. Two applicable examples are the following:

- Climate-weather interactions when diurnal cycle and seasonal cycle are taken into consideration and one has to encode the time dependent statistics of the fast variables (the weather) when making predictions (e.g., [7, 11, 37, 38]).
- Neuronal learning models, where the strength of neurons connections evolve at a slower rate to account for synaptic events (synaptic plasticity), leading to a time dependent slow-fast SDE (e.g., [19]), etc.

The study of these class of systems are important and challenging problems in modern applied mathematics. One of the technical challenges is to identify the right notion of convergence of the slow motion to the averaged motion. There are no many theoretical results in this direction, however, significant progress have been made over the last decade in the study of the long time behaviour of time dependent RDS (e.g., [15, 16, 17, 47, 48, 53, 54, 55, 57]). Based on these results, one can justify averaging principle for time dependent RDS once a suitable notion of ergodicity is known.

The purpose of ergodicity is to study invariant measures and related problems, it is one of the well studied problems in dynamical systems, stochastic analysis, statistical physics and related areas. Intuitively, ergodic dynamical system is the one that the behaviour of observables averaged over long time is the same as the behaviour of observables averaged over phase space. In RDS, ergodicity is the cornerstone in the study of long time behaviour (e.g., part II and III of [1] and references therein) and in particular, in the investigation of averaging. Various caveats of ergodicity such as uniform/unique or non-uniform ergodicity exist in the literature, depending on the mode of convergence and or existence of unique or several extremal invariant measure(s). Averaging principle for SDEs when the fast subsystem is independent of the slow variable holds true whenever the fast motion is uniform or unique ergodic (e.g., [28, 44, 18]). In general, uniform
or unique ergodicity is not available in the fully coupled case and in particular, time dependent setting, so one has to resort to other form of ergodicity to identify the averaging limit. In the time dependent framework, some form of ergodicity like ergodicity on Poincaré sections (PS-ergodicity \cite{17, 49}) or semi-uniform ergodicity (e.g, \cite{46}) of the fast subsystem could be employed to identify the averaging limit. In this paper, we restrict ourselves to when the fast subsystem is SDE with time periodic forcing, then under some weak dissipative conditions, PS-ergodicity is proved and consequently, used to identify the averaging limit. Extra essential difficulty in our case is that we avoid some analytic assumption such as uniform ellipticity of the fast subsystem, at the expense of proving the existence of unique stable random periodic process and non-degeneracy of Malliavin covariance of solution flows under certain weak dissipative assumptions, restricted Hömander’s Lie bracket condition and Lyapunov second method.

Lyapunov second method is a powerful technique for the investigation of stability of solutions of nonlinear dynamical systems and control systems in finite and infinite dimensions. Its extension to random dynamical systems generated by SDEs is essentially due to Hasminskii (e.g., \cite{27}). This method was explored further by X. Mao (e.g., \cite{40}) and extended further by Schmalfuss in (\cite{45}) to the case of nontrivial stationary solutions and random attractors. The method involves the study of random invariant sets without explicit knowledge of solutions of the system. The framework of Lyapunov second method is simple in the sense that its verification is based explicitly on the underlying vector fields without explicit representation of flows of solutions.

Here, we employ Lyapunov second method together with synchronous coupling to prove the existence of unique stable random periodic solution and ergodicity of the corresponding periodic measures. These techniques are valid even when the drift of the SDE is only one-sided Lipschitz continuous. Importantly, it makes it easier to prove results in the narrow topology generated by the dual of the Markov evolutionary semigroup of the flows of SDEs.

The rest of the paper is organised as follows. In §1.2, we fix some relevent function spaces used in the paper and in §1.3, we formulate and describe the problem considered in the paper. The existence of random periodic solutions and ergodic periodic measures for fast subsystem are considered in §2. The validity of the averaging principle in the random periodic regime is proved in §3, and we conclude with a toy example arising from kinetic theories of turbulent flows.

1.2. Notations. Let \((M, d)\) be a complete separable metric space, we denote the set of bounded continuous real-valued functions by

\[ C_b(M) = \left\{ f : M \to \mathbb{R} \text{ continuous} : \|f\|_{\infty} < \infty \right\}, \quad \|f\|_{\infty} := \sup_{x \in M} |f(x)|. \]

The set of bounded Lipschitz continuous real-valued functions is denoted by

\[ \text{Lip}_b(M) = \left\{ f \in C_b(M) : \|f\|_{BL} < \infty \right\}, \]

\[ \|f\|_{BL} = \max\{\|f\|_{\infty}, \|f\|_L\}, \quad \text{and} \quad \|f\|_L = \sup \left\{ \frac{|f(y) - f(z)|}{d(y, z)} : y \neq z \right\}. \]

Let \((M, B(M), \mu)\) be a Borel measure space, we denote \(L^p(M)\), \(1 \leq p < \infty\) as the set of real-valued Lebesgue integrable functions \(f : M \to \mathbb{R}\) defined by

\[ L^p(M) := \left\{ f : M \to \mathbb{R} \text{ measurable} : \|f\|_p < \infty \right\}, \quad \|f\|_p = \left( \int_M |f|^p \, d\mu \right)^{1/p}. \]

Let \(\mathcal{M} \subseteq M\) be a non-empty Borel measurable set, we shall most at times restrict the set of functions above on \(\mathcal{M}\), for example, \(f \in C_b(\mathcal{M})\) means \(f\lvert_{\mathcal{M}} \in C_b(M)\), where \(f\lvert_{\mathcal{M}} : \mathcal{M} \cap \text{Dom} f \to \mathbb{R}\).
Definition 1.1 ([2, 36]). Let $l \in \mathbb{N}$ and $0 \leq \delta \leq 1$.

(a) Let $n, m \in \mathbb{N} \setminus \{0\}$, denote $C^{l, \delta}(\mathbb{R}^n; \mathbb{R}^m)$ as the Fréchet space of functions $f : \mathbb{R}^n \to \mathbb{R}^m$ which are continuously differentiable and $t$-th derivative is $\delta$-Hölder continuous, with the semi norms
\[
\|f\|_{l, 0, K} := \sum_{0 \leq |\alpha| \leq l} \sup_{x \in K} |D^\alpha f(x)|,
\]
\[
\|f\|_{l, \delta, K} := \|f\|_{l, 0, K} + \sum_{|\alpha| = l} \sup_{x, y \in K, x \neq y} \frac{|D^\alpha f(x) - D^\alpha f(y)|}{|x - y|^{\delta}}, \quad 0 < \delta \leq 1.
\]
Here $K$ is a compact convex subset of $\mathbb{R}^n$, $\alpha = (\alpha_1, \cdots, \alpha_n) \in \mathbb{N}^n$ and $|\alpha| = \alpha_1 + \cdots + \alpha_n$ and
\[
D^\alpha f(x) = \frac{\partial |\alpha| f}{(\partial x_1)^{\alpha_1} \cdots (\partial x_n)^{\alpha_n}}.
\]

(b) Let $C^l_0(\mathbb{R}^n; \mathbb{R}^m)$ be the Banach space of the functions $f : \mathbb{R}^n \to \mathbb{R}^m$ which are in the Fréchet space $C^{l, \delta}(\mathbb{R}^n; \mathbb{R}^e)$ with the norm
\[
\|f\|_{l} := \sup_{x \in \mathbb{R}^n} \frac{|f(x)|}{1 + |x|} + \sum_{1 \leq |\alpha| \leq l} \sup_{x \in \mathbb{R}^n} |D^\alpha f(x)|,
\]
\[
\|f\|_{l, \delta} := \|f\|_{l} + \sum_{|\alpha| = l} \sup_{x, y \in \mathbb{R}^n, x \neq y} \frac{|D^\alpha f(x) - D^\alpha f(y)|}{|x - y|^{\delta}} < \infty, \quad 0 < \delta \leq 1.
\]
We endow the space $C^l_0(\mathbb{R}^n; \mathbb{R}^m)$ with the norm
\[
\|f\|_l := \sup_{x \in \mathbb{R}^n} \frac{|f(x)|}{1 + |x|} + \sum_{1 \leq |\alpha| \leq l} \sup_{x \in \mathbb{R}^n} |D^\alpha f(x)|.
\]

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and $\mathcal{G} \subseteq \mathcal{F}$, we denote $L^p(\Omega, \mathcal{G}, \mathbb{P})$, $p \geq 1$ as the space of $\mathcal{G}$-measurable random variables $X : \Omega \to \mathbb{R}^m$ such that $\mathbb{E}|X|^p < \infty$ equipped with the $L^p$ norm $\|X\|_p = (\mathbb{E}|X|^p)^{1/p}$.

We shall fix the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ as the classical Wiener space, i.e., $\Omega = C_0(\mathbb{R}; \mathbb{R}^m)$, $m \in \mathbb{N}$, is a linear subspace of continuous functions that take zero at $t = 0$, endowed with compact open topology defined via
\[
d(\omega, \hat{\omega}) = \sum_{n=0}^{\infty} \frac{1}{2^n} \frac{\|\omega - \hat{\omega}\|_{n}}{1 + \|\omega - \hat{\omega}\|_{n}}, \quad \|\omega - \hat{\omega}\|_{n} = \sup_{t \in [-n, n]} |\omega(t) - \hat{\omega}(t)|.
\]
The sigma algebra $\mathcal{F}$ is the Borel sigma algebra generated by open subsets of $\Omega$ and $\mathbb{P}$ is the Wiener measure, i.e., that the law of the process $\omega \in \Omega$ with $\omega(0) = 0$.

1.3. The averaging setup. We are concerned with the following slow-fast SDE
\[
\begin{aligned}
&dX_t^\varepsilon = \varepsilon F(X_t^\varepsilon, Y_t^\varepsilon)dt, \quad X_0^\varepsilon = x, \\
&dY_t^\varepsilon = b(t, X_t^\varepsilon, Y_t^\varepsilon)dt + \sum_{k=1}^{N} \sigma_k(t, X_t^\varepsilon, Y_t^\varepsilon)dW_t^k, \quad Y_0^\varepsilon = y,
\end{aligned}
\tag{1.1}
\]
where $0 < \varepsilon \ll 1$ is the time scales separation between $X^\varepsilon$ and $Y^\varepsilon$. The vector fields $F : \mathbb{R}^d \times \mathbb{R}^N \to \mathbb{R}^d$, $b : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^N \to \mathbb{R}^N$ and $\sigma_k : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^N \to \mathbb{R}^N$ take $1 \leq k \leq N$, are such that $F \in C^{l-1, \alpha}_{b}(\mathbb{R}^d \times \mathbb{R}^N; \mathbb{R}^d)$ and $b, \sigma_k \in C^{l, \alpha}_b(\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^N, \mathbb{R}^N)$, $l \geq 2$. Given this regularity of the vector fields, the solutions of the SDE (1.1) exist and generate stochastic flow of $C^{l-1, \alpha}$ diffeomorphisms with $0 < \alpha \leq \delta$ (e.g., [14, 29, 35, 36]). We define the flow map $\Xi_t^\varepsilon(\omega) \equiv \Xi_t^\varepsilon(\omega) :
It is natural to think of the slow-fast SDE \((b, \sigma, \epsilon)x\) if \(R\) long ago by Hasminskii \([\ldots]\) following fast subsystem with the slow variable frozen topology.  

If \(\epsilon = 0\), the flow \(\Xi_t^0(\omega)\) reduces to \(\Xi_t^{0,x,y}(\omega) = (x, Y_t^{x,y}(\omega))\) where \(Y_t^{x,y}\) is the solution of the following fast subsystem with the slow variable frozen

\[
dY_t^{x,y} = b(t, x, Y_t^{x,y})dt + \sum_{k=1}^{N} \sigma_k(t, x, Y_t^{x,y})dW_t, \quad Y_0^{x,y} = y. \tag{1.2}
\]

It is natural to think of the slow-fast SDE (1.1) as a perturbation of the parametrised SDE (1.2). In this way, the flow map \(\Xi_t^{0,x,y}(\omega)\) describes idealised physical system where \(x = (x_1, \ldots, x_d)\) are fixed parameters, whereas the perturbed flow \(\Xi_t^{x,y}\) is describing real system with the evolution of the parameters taken into consideration.

In some previous works \(\text{e.g., [30, 5, 31, 32, 34, 50]}\), the validity of averaging principle for fully coupled ODEs and SDEs were investigated. The averaging principle for the fully coupled system is significantly different from the one when the fast motion is independent of the slow variables. The required assumptions in the fully coupled case depend on the specific class of systems and convergence of the slow motion to averaged motion is in a different sense \(\text{e.g., [18, 31, 42]}\).  

Consider the limit \(\bar{F}\) defined by

\[
\bar{F}(x) = \bar{F}_y(x) := \lim_{T \to \infty} \frac{1}{T} \int_0^T F(x,Y_t^{x,y})dt. \tag{1.3}
\]

For example, if \(\mu^x\) is an ergodic invariant measure of the fast motion \(Y^x = \{Y_t^{x,y} : (t, y) \in \mathbb{R}^+ \times \mathbb{R}^N\}\), then the limit (1.3) exists for \(\mu^x\)-almost all \(y\), due to ergodicity and it coincides with the following integral

\[
\bar{F}(x) = \bar{F}_\mu(x) := \int_{\mathbb{R}^N} F(x, y)\mu^x(dy). \tag{1.4}
\]

In the case, where the vector fields \(b, \sigma_k, 1 \leq k \leq N\) are independent of \((t, x)\), the fast flow defines a closed system, i.e., \(Y_t^x = Y_t\) and \(\mu^x = \mu\), in this case, one recovers the averaging principle proved long ago by Hasminskii \([28]\). So the averaged vector field \(\bar{F}\) inherit the regularities of slow vector field \(F\), thus, there exists a unique solution \(\bar{X}_t^{x,x}\) to the following averaged equation

\[
\frac{d\bar{X}_t^{x,x}}{dt} = \bar{F}(\bar{X}_t^{x,x}), \quad \bar{X}_0^{x,x} = x. \tag{1.5}
\]

In this case, the averaging principle states \(\text{e.g., [28, 33]}\), for \(\delta > 0\),

\[
\lim_{\epsilon \to 0} \mathbb{P} \left( \sup_{0 \leq t \leq T/\epsilon} |X_t^{x,y} - \bar{X}_t^{x,x}| > \delta \right) = 0, \quad \text{for all } (x,y) \in \Lambda, \tag{1.6}
\]

where \(\Lambda := \{(x,y) \in \mathbb{R}^d \times \mathbb{R}^N : \bar{F}_y(x) = \bar{F}_\mu(x)\}\).

If \(b, \sigma_k, 1 \leq k \leq N\), is autonomous but depend on \((x, y)\), the averaged vector field \(\bar{F}(x)\) need not be continuous in \(x\), let alone Lipschitz continuous; so the averaged equation (1.5) may admit infinitely many solutions or no solution at all. This is partly because in this case, the fast motions \(Y^x = \{Y_t^{x,y} : (t, y) \in \mathbb{R}^+ \times \mathbb{R}^N\}\) do not define a closed system, as \(Y_t^x\) may have different properties for different \(x\); hence, no well-defined family of invariant measures \(\{\mu^x : x \in \mathbb{R}^d\}\). Even in the rare case when \(\mu^x\) is the same for all \(x \in \mathbb{R}^d\), averaging in fully coupled system may not satisfy the limit (1.6), for example in the presence of resonance \(\text{e.g., [31, 42]}\). It is proved in \(\text{e.g.,[3, 30]}\).
that if the convergence in (1.3) is uniform in $x$ and $y$, that any limit points $\bar{X}_t^x$ in probability of $X_{t/\varepsilon}^{x,y}$ as $\varepsilon \to 0$, solves the averaged equation

$$\frac{dX_t^x}{dt} = F(X_t^x), \quad X_0^x = x.$$ \hfill (1.7)

Note that the limit (1.3) is uniform with respect to $y$ if and only if the fast flow $Y^x$ is uniquely ergodic, this occurs rather rarely and exclude some important models. In general, averaging in the fully coupled case requires different setup with stronger and specific assumptions (c.f. [31]). Precisely, it is noted in ([31, 33, 34]) that in the fully coupled case, it is useful to have necessary and sufficient conditions which ensure the following:

Given that $\mathbb{Q} \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^N)$ admit the disintegration $\mathbb{Q}(dx, dy) = \mu^x(dy)\nu(dx)$, with $\nu \in \mathcal{P}(\mathbb{R}^d)$ and $\mu^x \in \mathcal{P}(\mathbb{R}^N)$, then for $\mathbb{K} \subset \mathbb{R}^d$ and $\mathbb{C} \subset \mathbb{R}^N$ and any $\delta > 0$,

$$\lim_{\varepsilon \to 0} \int_{\mathbb{K}} \int_{\mathbb{C}} \mathbb{P} \left( \sup_{0 \leq t \leq T/\varepsilon} |X_t^{x,y} - X_t^{x}| > \delta \right) \mu^x(dy)\nu(dx) = 0. \hfill (1.8)$$

The averaging principle in the form (1.8) was proved long ago by Anosov in the case when the fast subsystem satisfy Liouville property (see chapter 2 of [42]). It was later generalised by Kifer to the case when the fast subsystem is axiom $A$ dynamical systems and to the case of Markovian SDEs ([31, 33, 34] and references therein).

In the present paper, the fast motion is time dependent, so ergodic invariant measures may not exist. However, if for example, the coefficients of the slow-fast SDE (1.1) are time periodic, quasi-periodic, uniform almost periodic or uniform almost automorphic, one may consider evolution system of probability measures $\{\mu_t^x : t \in \mathbb{R}\}$ (e.g., [10]) generated by the Markov evolution $\{\mathbb{P}_{s,t}^x : s \leq t\}$ of the fast motion, i.e.,

$$\int_{\mathbb{R}^N} \mathbb{P}_{s,t}^x f(y)\mu_s^x(dy) = \int_{\mathbb{R}^N} f(y)\mu_s^x(dy), \quad s \leq t, \quad f \in \mathcal{C}_0(\mathbb{R}^N). \hfill (1.9)$$

One can exploit the nice property of the above evolution system of probability measures with some dissipative conditions on the coefficients of the fast subsystem to identify the averaging limit. There are existing results in this time dependent case; for example, when the fast subsystem is time periodic and uniform elliptic was considered in (e.g., [51, 52]) and applied to investigate the long time behaviour of neuronal learning models in [19]. The case of uniform almost periodic fast subsystem was recently considered for stochastic reaction-diffusion equation in [9]. In this paper, we shall assume that the coefficients of the fast motion is periodic in time, i.e., exists $\tau > 0$ such that $b(t + \tau, x, y) = b(t, x, y)$, $\sigma(t + \tau, x, y) = \sigma(t, x, y)$. Under this time periodicity and regularity of the vector fields $b, \sigma_k$, $1 \leq k \leq N$, it have been rigorously established in (e.g., [15, 16, 17, 47, 48, 49, 57]) that the SDE (1.2) admit random periodic solution $S^x : \mathbb{R} \times \Omega \to \mathbb{R}^N$ (see §2.2). The law of the random periodic solution $\mu_t^x(A) := \mathbb{P}\{\omega : S^x(t, \omega) \in A\}$, $A \in \mathcal{B}(\mathbb{R}^N)$ generate evolution system of probability measures and, in particular, periodic measures (see §2.3). Under some dissipative conditions on the fast SDE (1.2), we establish ergodicity of the periodic measures $\mu_t^x, t \in [0, \tau)$ on certain minimal Poincaré section, we refer to this notion as PS-ergodicity of periodic measures as recently introduced in [17] (see §2.3).

\footnote{$\mathbb{K} \Subset \mathbb{M}$ means that $\mathbb{K}$ is a non-empty relative compact subset of $\mathbb{M}$.}
A useful way to visualise PS-ergodic periodic measures is to lift the stochastic flow to the cylinder \( S^1 \times \mathbb{R}^N \) \(^2\) and obtain an autonomous stochastic flow of the form \( \hat{Y}^\varepsilon,\mathcal{x},\mathcal{y} = (t + s \mod \tau, Y_{t+s}^\varepsilon,\mathcal{x},\mathcal{y}(\theta_{-s}\omega)) \) with the initial value \( \hat{y} = (s, y) \in S^1 \times \mathbb{R}^N \). The random periodic solution on \( S^1 \times \mathbb{R}^N \) then takes the form \( \hat{S}^\varepsilon(t, \omega) = (s \mod \tau, S^\varepsilon(t, \omega)) \). In this case, the slow-fast SDE on the extended phase space \( \mathbb{R}^d \times S^1 \times \mathbb{R}^N \), reads

\[
\begin{aligned}
&dX_t^\varepsilon,\mathcal{x},\mathcal{y} = \varepsilon F(X_t^\varepsilon,\mathcal{x},\mathcal{y}, Y_t^\varepsilon,\mathcal{x},\mathcal{y})dt,
&dY_t^\varepsilon,\mathcal{x},\mathcal{y} = \left( 1 \over b(s_t, x, Y_t^\varepsilon,\mathcal{x},\mathcal{y}) \right) dt + \left( 0 \over 0 \right) \sum_{k=1}^N \sigma_k(s_t, x, Y_t^\varepsilon,\mathcal{x},\mathcal{y}) \left( dW_t^0 \right) \left( dW_t^k \right) 
\end{aligned}
\]  

(1.10)

where \( W_t^0 \) is a one dimensional Brownian motion independent of the \( N \)-dimensional Brownian motion \( \{W_t^k : 1 \leq k \leq N, t \geq 0\} \). Apparently, this lifting does not reduce the complexity of the problem, but the SDE is now autonomous. In this way, under the regularity of the coefficients, the lifted slow-fast SDE (1.10) generate a Markovian RDS on the extended phase space \( \mathbb{R}^d \times S^1 \times \mathbb{R}^N \) (see [2], chapter 2 of [1] and references therein).

In addition to the PS-ergodicity of the periodic measures, it is important in this fully coupled case to prove that the branches of the periodic measures \( x \mapsto \tilde{\mu}^x_r, \, r \in [0, \tau) \) are Lipschitz continuous w.r.t. \( x \in \mathbb{R} \oplus \mathbb{R}^d \) in the narrow topology of \( \mathcal{P}(\mathcal{M}) \), where \( \mathcal{M} \subset S^1 \times \mathbb{R}^N \) is the union of minimal Poincaré sections (see §3.1). This Lipschitz property of periodic measures together with the regularity of the slow vector field \( F \), would guarantee the existence of \( \hat{C}_F > 0 \) such that

\[
\begin{aligned}
&\left| F(x) - F(z) \right| \leq \hat{C}_F |x - z|, \\
&\left| \hat{F}(x) \right| \leq \hat{C}_F,
\end{aligned}
\]  

(1.11)

where

\[
\hat{F}(x) := \frac{1}{\tau} \int_0^\tau \int_{\mathcal{M}} F(x, \pi_2(\tilde{y})) \tilde{\mu}^x_t(\tilde{y}) \, dr
\]  

(1.12)

and \( \pi_2 : S^1 \times \mathbb{R}^N \to \mathbb{R}^N \) with \( \pi_2(\tilde{y}) = y \). The local Lipschitz and boundedness property (1.11) would yield the existence of a unique solution \( \hat{X}^\varepsilon,\mathcal{x} \) to the averaged equation

\[
\frac{d\hat{X}^\varepsilon,\mathcal{x}_t}{dt} = \varepsilon \tilde{F}(\hat{X}^\varepsilon,\mathcal{x}_t), \quad \hat{X}^\varepsilon,\mathcal{x}_0 = x.
\]  

(1.13)

Finally, given the PS-ergodicity and Lipschitz dependence of the family \( \{\tilde{\mu}_t^x : r \in [0, \tau), \, x \in \mathbb{R} \oplus \mathbb{R}^d \} \), we are able to deduce averaging principle in the form of the limit (1.8) with the set \( C \) replaced by minimal Poincaré section \( \mathcal{M} \subset S^1 \times \mathbb{R}^N \) (see §3.2).

2. Random periodicity

In this section, we investigate the existence of unique stable random periodic solution and PS-ergodicity of periodic measures for Markovian RDS generated by some class of SDEs in finite dimensions. To facilitate our presentation, we first recall definitions of stochastic flows and RDS (see, e.g., [1, 2, 35, 36]) and random periodic solutions (see, e.g., [15, 16, 17, 57]).

\(^2\)i.e., \([0, \tau] \cong S^1\), with the homeomorphism \( t \mapsto e^{2\pi t/\tau} \).
2.1. Random periodic solutions.

**Definition 2.1** (Stochastic flow [35, 36]). Let \( \mathbb{M} \) be a smooth manifold and let \( \{ \Xi_{t,s}^\tau(\omega) : s, t \in T \subseteq \mathbb{R}, z \in \mathbb{M} \} \) be a random field defined on a complete probability space \((\Omega, \mathcal{F}, \mathbb{P})\). The map \( \Xi_{t,s}^\tau(\omega) \) is called a *stochastic flow of homeomorphism* if there exists a null set \( \mathcal{N} \subseteq \Omega \) such that for any \( \omega \notin \mathcal{N} \), the family of continuous maps \( \{ \Xi_{t,s}^\tau(\omega) : s, t \in T \} \) is a flow of homeomorphism in the following sense:

(i) \( \Xi_{t,s}^\tau(\omega) = \Xi_{t,u}^\tau(\omega) \circ \Xi_{u,s}^\tau(\omega) \) holds for any \( s, t, u \).

(ii) \( \Xi_{s,s}(\omega) = \text{Id}_{\mathbb{M}} \), for all \( s \).

(iii) the map \( \Xi_{t,s}(\omega) : \mathbb{M} \to \mathbb{M} \) is a homeomorphism for any \((s, t)\).

The map \( \Xi_{t,s}(\omega) \) is a *stochastic flow of \( C^k \)-diffeomorphism*, if it is a homeomorphism and \( \Xi_{t,s}^\tau(\omega) \) is \( k \)-times continuously differentiable with respect to \( z \) for all \( s, t \in T \subseteq \mathbb{R} \) and the derivatives are continuous in \((s, t, z)\).

Let \( \mathcal{F}_s^t \) be the smallest sub \( \sigma \)-field of \( \mathcal{F} \) containing \( \cap_{\varepsilon>0}\sigma(\Xi_{u,v} : s-\varepsilon \leq u, v \leq t+\varepsilon) \) and all null sets of \( \mathcal{F} \). Then the two parameter filtration \( \{ \mathcal{F}_s^t : s \leq t \} \) is the filtration generated by the stochastic flow \( \{ \Xi_{s,t} : s, t \in T \subseteq \mathbb{R} \} \).

**Definition 2.2** (RDS [1, 2]). Let \( \mathbb{M} \) be a measurable space and \( \theta \) be an ergodic flow of measurable transformations of a complete probability space \((\Omega, \mathcal{F}, \mathbb{P})\), a *random dynamical system* (RDS) over \( \theta \) is a mapping \( \Phi : T \times \Omega \times \mathbb{M} \to \mathbb{M} \) such that

(a) \((t, \omega, z) \mapsto \Phi(t, \omega, z) \) is measurable,

(b) \( \Phi(0, \omega, z) = \text{Id}_{\mathbb{M}} \) for all \( \omega \in \Omega \) and

(c) \( \Phi(t+s, \omega) = \Phi(t, \theta_s \omega) \circ \Phi(t, \omega) \), for all \( s, t \in T, \omega \in \Omega \) (cocycle property),

(d) if \( \mathbb{M} \) is a topological space, then \( \Phi \) is continuous if \((t, z) \mapsto \Phi(t, \omega, z) \) is continuous,

(e) if \( \mathbb{M} \) is a \( C^k \) manifold \( 1 \leq k \leq \infty \), then \( \Phi \) is smooth of class \( C^k \), if \( \Phi(t, \omega, z) \) \( k \)-times differentiable with respect to \( z \), and the derivatives are continuous with respect to \((t, z)\).

Let \( \theta : \mathbb{R} \times \Omega \to \Omega \), be a \((\mathcal{B}(\mathbb{R}) \otimes \mathcal{F}, \mathbb{P})\)-measurable flow defined by

\[
\theta_t(\omega(\cdot)) = \omega(t + \cdot) - \omega(t). 
\]  

(2.1)

It is well known that the measurable flow \((\theta_t)_{t \in \mathbb{R}}\) is ergodic and defines a filtered metric dynamical system \( \theta = (\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_s^t)_{s \leq t}, (\theta_t)_{t \in \mathbb{R}}) \) (e.g., [1, 2]). Furthermore, under regularity of the coefficients of autonomous SDEs together with appropriate adoption of two-sided stochastic calculus, it is well known that the solutions of autonomous SDEs on finite dimensions generate RDS over \( \theta \) (e.g., [1, 2, 14, 29, 36]).

**Definition 2.3** (Random periodic solution for stochastic flows [15, 16, 57]). A random periodic solution of period \( \tau > 0 \) of a stochastic flow \( \Xi : \Delta \times \Omega \times \mathbb{M} \to \mathbb{M} \), is an \( \mathcal{F} \)-measurable function \( S : T \times \Omega \to \mathbb{M} \) such that

\[
S(t + \tau, \omega) = S(t, \theta_\tau \omega) \quad \text{and} \quad \Xi_{s,t+s}^{S(t,\omega)}(\omega) = S(t + s, \omega), \quad \mathbb{P} \text{-a.s.,}
\]

for any \((t, s) \in \Delta \subseteq T \times T \) and \( \omega \in \Omega \), where \( \Delta := \{(t, s) \in T \times T : s \leq t\} \).

**Example 2.4.** Let \( \mathbb{M} = \mathbb{R} \), consider the following SDE

\[
\begin{cases}
    dZ_t = -\alpha(t)Z_t dt + \beta(t) dt + dW_t, \quad t \geq t_0 \\
    Z_{t_0} = z_0 \in \mathbb{R}.
\end{cases}
\]  

(2.2)
Assume that $\alpha, \beta : \mathbb{R} \to \mathbb{R}$ are continuous functions and there exists $\tau > 0$ such that $\alpha(t + \tau) = \alpha(t)$ and $\beta(t + \tau) = \beta(t)$ with
\[
\int_{-\infty}^{t} e^{-\int_{0}^{s} \alpha(u)du} \beta(s)ds + \int_{-\infty}^{t} e^{-2\int_{0}^{s} \alpha(u)du} ds < \infty, \quad \text{for} \quad 0 \leq t \leq \tau.
\]
The stochastic flow $\{\Xi_{t_0,t}(\omega) : t \geq t_0\}$ generated by the SDE (2.2) is defined by
\[
\Xi_{t_0,t}^\omega = z_0 e^{-\int_{0}^{t} \alpha(u)du} + \int_{0}^{t} e^{-\int_{s}^{t} \alpha(u)du} \beta(s)ds + \int_{0}^{t} e^{-\int_{s}^{t} \alpha(u)du} dW_s,
\]
and the process $S(t, \omega)$ defined by
\[
S(t, \omega) = \int_{-\infty}^{t} e^{-\int_{0}^{s} \alpha(u)du} \beta(s)ds + \int_{-\infty}^{t} e^{-\int_{s}^{t} \alpha(u)du} dW_s(\omega)
\]
is a random periodic solution of period $\tau$.

**Definition 2.5** (Random periodic solution for RDS [17, 57]). A random periodic path of period $\tau > 0$ of an RDS $\Phi : \mathbb{R}^+ \times \Omega \times \mathbb{M} \to \mathbb{M}$ is an $\mathcal{F}$-measurable function $S : \mathbb{R} \times \Omega \to \mathbb{M}$ such that for almost all $\omega \in \Omega$,
\[
S(t + \tau, \omega) = S(t, \theta_\tau \omega) \quad \text{and} \quad \Phi(t, \theta_\tau \omega, S(s, \omega)) = S(t + s, \omega)
\]
for $t \in \mathbb{R}^+$, $s \in \mathbb{R}$.

**Example 2.6** ([15, 16]). Let $b : \mathbb{R}^N \to \mathbb{R}^N$, $N \geq 2$ be a regular vector field and let $\{\Phi(t,) : t \in T \subseteq \mathbb{R}\}$ be a deterministic flow generated by the ODE
\[
\frac{dY_t}{dt} = b(Y_t).
\]
If $u : T \to \mathbb{R}^N$ is a periodic solution of the ODE (2.4) of period $\tau > 0$, i.e.,
\[
u(t + \tau) = u(t) \quad \text{and} \quad \Phi(t + s, u(s)) = u(t + s), \quad t, s \in T \subseteq \mathbb{R}.
\]
Consider the stochastic process $X_t(\omega) = u(t) + Z_t(\omega)$, where $Z_t$ solves the following time periodic forcing SDE
\[
dZ_t = \hat{b}(t, Z_t)dt + \hat{\sigma}(t, Z_t)dW_t,
\]
where $\hat{b}(t, z) := b(u(t) + z) - b(u(t))$ and $\hat{\sigma}(t, z) := \sigma(u(t) + z)$. If $v(t, \omega)$ is a random periodic solution of the time periodic forcing SDE (2.5), then $S(t, \omega) = u(t) + v(t, \omega)$ is a random periodic solution of the autonomous SDE
\[
\frac{dX_t}{dt} = b(X_t)dt + \sigma(X_t)dW_t.
\]

2.2. **Existence of stable random periodic solutions.** Here, we present a version of a random periodic results in ([47, 49]). This result will be used to prove PS-ergodicity and Lipschitz dependence of periodic measures on parameters. As pointed in the introduction, the technique here is a variant of Hasminskii’s method for stability of solutions of SDEs (see, e.g., [27, 40, 45]). The central idea is to investigate the infinitesimal separation of trajectories via the average growth of a Lyapunov function $V : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^+$ along the two-point motions $\{Y_{s,t}^z(\omega) - Y_{s,t}^y(\omega) : t \geq s, \ y, z \in \mathbb{R}^N, \ z \neq y\}$ which leads to the existence of (pre)-fixed point via Arzelà–Ascoli like argument.
First, we fix $s \in \mathbb{R}$ and recall that the transition probability function $P(s,y;t,.)$ induced by solutions of the SDE (1.2) is defined by
\[
P^x(s,y;t+s,A) = \mathbb{P}\{\omega \in \Omega : Y^{x,y}_{s,t+s}(\omega) \in A\}, \quad t \geq 0, \quad A \in \mathcal{B}(\mathbb{R}^N). \tag{2.7}
\]

The Markov evolution $P_{s,t+s}^x : \mathcal{C}_b(\mathbb{R}^N) \to \mathcal{C}_b(\mathbb{R}^N)$ is defined by
\[
P_{s,t+s}^x f(y) = \int_{\mathbb{R}^N} f(z) P^x(s,y;t+s,dz) = \mathbb{E}[f(Y_{s,t+s}(\omega)) | Y^x_{s,t} = y] \tag{2.8}
\]
and for any probability measure $\mu$ on $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$, the Markov dual $P_{s,t+s}^{x,\mu}$ is defined
\[
(P_{s,t+s}^{x,\mu})(A) = \int_{\mathbb{R}^N} P^x(s,y;t+s,A) \mu(dy), \quad t \geq 0, \quad A \in \mathcal{B}(\mathbb{R}^N). \tag{2.9}
\]

Next, the average growth of the function $V$ along the trajectory $Y^{x,y}_{s,t}(\omega)$ gives rise to the generator $\mathcal{L}_x$,
\[
\mathcal{L}_x V(t,y) = \lim_{h \to 0} \frac{\mathbb{E}[V(t+h,Y_{t,t+h}(\omega)) | Y^x_{t} = y] - V(t,y)}{h}. \tag{2.10}
\]

For $V \in C_b^{1,2}(\mathbb{R} \times \mathbb{R}^N)$, we use Itô’s formula to write $\mathcal{L}_x$ as
\[
\mathcal{L}_x V(t,y) = \frac{\partial V(t,y)}{\partial t} + \sum_{i=1}^N b^i(t,x,y) \frac{\partial V(t,y)}{\partial y^i} + \frac{1}{2} \sum_{i,j=1}^N \sum_{k=1}^N \sigma_{ik}(t,x,y) \sigma_{jk}(t,x,y) \frac{\partial^2 V(t,y)}{\partial y^i \partial y^j}.
\]

Now, consider the difference between two solutions of the fast subsystem starting from two different initial values $\{Y^{x,y}_{s,t+s}(\omega) - Y^{x,z}_{s,t+s}(\omega) : y, z \in \mathbb{R}^N, \ y \neq z, \ t \geq 0\}$. Here, we think of $Y^{x,y}_{s,t}(\omega) - Y^{x,z}_{s,t}(\omega)$ as a special case of coupling two copies of Itô process namely: $(y,z) \mapsto H(y,z) = y - z$, so that $H(Y^{x,y}_{s,t}(\omega),Y^{x,z}_{s,t}(\omega)) = Y^{x,y}_{s,t}(\omega) - Y^{x,z}_{s,t}(\omega)$. Then, by generalised Itô’s formula (Theorem 8.1 in [3]), the average growth of the function $V \in C_b^{1,2}(\mathbb{R} \times \mathbb{R}^N ; \mathbb{R}^+)$ along the trajectory $\{Y^{x,y}_{s,t+s}(\omega) - Y^{x,z}_{s,t+s}(\omega) : y \neq z, \ t \geq 0\}$, yields the second order differential operator $\mathcal{L}_x^{(2)}$, defined by
\[
\mathcal{L}_x^{(2)} V(t,y-z) = \partial_t V(t,y-z) + D_y V(t,y-z) (b(t,x,y) - b(t,x,z)) + \frac{1}{2} \text{trace}\left( (\sigma(t,x,y) - \sigma(t,x,z))^T D_y^2 V(t,y-z) (\sigma(t,x,y) - \sigma(t,x,z)) \right), \tag{2.11}
\]
where $\sigma(t,y) = (\sigma_{ik}(t,y))_{1 \leq i,k \leq N}$, $D_y V = \left( \frac{\partial V}{\partial y^i} \right)_{1 \leq i \leq N}$, $D_y^2 V = \left( \frac{\partial^2 V}{\partial y^i \partial y^j} \right)_{1 \leq i,j \leq N}$.

Finally, we return to the lifted fast subsystem on the cylinder $\mathbb{S}^1 \times \mathbb{R}^N$,
\[
dY^{x,\tilde{y}}_t = \tilde{b}(x,\tilde{Y}^{x,\tilde{y}}_t) dt + \tilde{\sigma}(x,\tilde{Y}^{x,\tilde{y}}_t) d\tilde{W}_t, \quad \tilde{Y}^{x,\tilde{y}}_0 = \tilde{y}. \tag{2.12}
\]

where $\tilde{y} = (s,y) \in \mathbb{S}^1 \times \mathbb{R}^N$, $\tilde{W}_t = (W^0_t,W^k_t)^T$, $1 \leq k \leq N$, the vector fields $\tilde{b}$, $\tilde{\sigma}$ are given by
\[
\tilde{b}(x,(s,y)) = (1,b(s,x,y))^T \quad \text{and} \quad \tilde{\sigma}(x,(s,y)) = \left( 0, \sum_{k=1}^N \sigma_k(s,x,y) \right).
\]

Throughout the remaining part of this section, we denote the stochastic flow generated by the lifted SDE (2.12) by
\[
\Phi^x(t,\omega,(s,y)) := \tilde{Y}^{x,(s,y)}_t(\omega) = \left( t + s \mod \tau, Y^{x,y}_{s,t+s}(\theta_{-s}\omega) \right).
\]

In some places, we may skip the dependence on $x$, when it does not cause confusion.
Next, recall that the transition probability function in space-time $\mathbb{R}^+ \times \mathbb{R}^N$ takes the form

$$
P(0, (s, y); t, \mathbb{T} \times B) := \mathbb{P} \left( \omega : (t + s, Y_{s+t+k}^\omega (\theta_{-s}\omega)) \in \mathbb{T} \times B \right)
= P(s, x; t + s, B) \mathbb{P}(t + s), \quad \mathbb{T} \in \mathcal{B}(\mathbb{R}^+), \; B \in \mathcal{B}(\mathbb{R}^N).
$$

(2.13)

In particular, the transition probability function $\tilde{P}(0, \tilde{y}; t, \cdot)$, $t \geq 0$ and the Markov evolution $(\tilde{P}_t)_{t \geq 0}$ (now Markov semigroup) on $S^1 \times \mathbb{R}^N$ is defined for $\tilde{y} \in S^1 \times \mathbb{R}^N$ as

$$
\tilde{P}(0, \tilde{y}; t, A) = \mathbb{P}(\omega : \Phi(t, \omega, \tilde{y}) \in A), \quad A \in \mathcal{B}(S^1) \otimes \mathcal{B}(\mathbb{R}^N),
$$

$$
\tilde{P}_t \tilde{h}(\tilde{y}) = \int_{S^1 \times \mathbb{R}^N} \tilde{h}(\tilde{z}) \tilde{P}(0, \tilde{y}; t, d\tilde{z}), \quad \tilde{h} \in C_b(S^1 \times \mathbb{R}^N).
$$

The Markov dual $(\tilde{P}_t^*)_{t \geq 0}$ on $S^1 \times \mathbb{R}^N$ is

$$
(\tilde{P}_t^* \tilde{\mu})(\tilde{A}) = \int_{S^1 \times \mathbb{R}^N} \tilde{P}(0, \tilde{y}; t, \tilde{A}) \tilde{\mu}(d\tilde{y}), \quad \tilde{\mu} \in \mathcal{P}(S^1 \times \mathbb{R}^N).
$$

With all the above notations in place, we present our result on the existence and uniqueness of stable random periodic solution on $S^1 \times \mathbb{R}^N$ under the following assumptions.

**Assumption 2.7.** In addition to the regularity and time periodicity of the coefficients of the fast subsystem, we suppose further that the following hold:

There exists a function $V \in C^{1,2}(\mathbb{R} \times \mathbb{R}^N; \mathbb{R}^+)$, $\tau$-periodic in the time component such that $V(t, 0) = 0$ and for all $t \in \mathbb{R}$, $y, z \in \mathbb{R}^N$, $x \in K \subseteq \mathbb{R}^d$,

$$
|y|^p \leq V(t, y) \leq C|y|^p \quad \text{and} \quad \mathcal{L}_x^{(2)}V(t, y - z) \leq \lambda(t, x)V(t, y - z)
$$

for $p \geq 1$, $C \geq 1$ with

$$
\lim_{t \to \infty} \sup_{u \in \mathbb{R}} \frac{1}{2t} \int_0^t \lambda(u)du := \beta < 0.
$$

(2.15)

where $\lambda(t) := \sup_{x \in \mathbb{R}} \lambda(t, x)$.

Moreover, let $\Phi^\tau(t, \omega, (s, y)) = (t + s \mod \tau, Y_{s+t+k}^\omega (\theta_{-s}\omega))$ be the lifted stochastic flow generated by the fast subsystem with the slow variable frozen such that

$$
\sup_{t \in [0, 1]} \log \mathbb{E} \left[ V(\Phi^\tau(t, \omega, (s, y)) - (s, y)) \right] < \infty,
$$

(2.16)

**Theorem 2.8** (Existence of stable random periodic path [47, 49]). Let $\{\gamma^\tau_{s,t}(\omega) : t \geq s\}$ be a stochastic flow generated by an SDE with the coefficients $b(t, \cdot), \sigma_k(t, \cdot) \in C_{b}^{1,2}(\mathbb{R}^N; \mathbb{R}^N)$, $1 \leq k \leq N$, $l \geq 2$, $\tau$-periodic and continuously differentiable in time component satisfying assumptions 2.7. Then, there exist a unique $F_{-\infty} \subseteq \bigwedge_{u \leq s} F_u$ measurable and exponential stable random periodic process $\tilde{S} : \mathbb{R} \times \Omega \to S^1 \times \mathbb{R}^N$.

**Proof.** Our presentation is in two steps, in the first step, we show that $\{\Phi(t, \omega, \Phi(k\tau, \theta_{-k\tau}\omega, \tilde{y}))\}_{k \in \mathbb{N}}$ is a Cauchy sequence in $C([0, \tau]; S^1 \times \mathbb{R}^N)$ and converges exponentially fast, the uniqueness of the limit follows by cocycle property of $\Phi$ on $S^1 \times \mathbb{R}^N$. The second step is to show via the cocycle property again and time periodicity of the coefficients of our SDE that the limit from the first step satisfies the definition of random periodic solution.
Step I: We investigate the separation of the trajectories $\Phi(t, \omega, (s, y))$ and $\Phi(t, \omega, \Phi(1, \theta_{-1}\omega, (s, y)))$. Let $(s, y) \neq (u, z)$ for some $(u, z) \in S^1 \times \mathbb{R}^d$, following same construction as in [45], we set

$$(s_1, X_1(\omega)) = \tilde{X}_1(\omega) := \begin{cases} (s, x), & \text{if} \quad (s, y) \neq (1 + s \mod \tau, X(s + 1, s, \theta_{-s-1}\omega, y)) \\ (u, z), & \text{if} \quad (s, y) = (1 + s \mod \tau, X(s + 1, s, \theta_{-s-1}\omega, y)). \end{cases}$$

Notice that $\tilde{X}_1$ is $\mathcal{F}_{-\infty}$-measurable and $V(\Phi(1, \theta_{-1}\omega, (s, x)) - \tilde{X}_1(s, \omega))$ satisfies the condition (2.16), since the random variable $V(\Phi(1, \theta_{-1}\omega, (s, y)) - (s, y))$ satisfies the same condition (2.16). Denote $\tilde{Y}(t, \omega) := \Phi(t, \omega, \tilde{X}_1(\omega))$ and $\tilde{Z}(t, \omega) := \Phi(t, \omega, \Phi(1, \theta_{-1}\omega, \tilde{y}))$, for $t \geq 0$. We apply Itô’s formula (e.g., [35, 45]) on the logarithmic process $\log V(\tilde{Y}(t, \omega) - \tilde{Z}(t, \omega))$, to obtain

$$\log V(\tilde{Y}(t, \omega) - \tilde{Z}(t, \omega)) = \log V(\tilde{X}_1(\omega) - \Phi(1, \theta_{-1}\omega, \tilde{y})) + \int_0^t \tilde{L}^{(2)} V(\tilde{Y}(r, \omega) - \tilde{Z}(r, \omega)) dr + N(t, \omega) - \frac{1}{2} q(t, \omega),$$

where

$$q(t, \omega) = \int_0^t \left( \mathcal{H} V(\tilde{Y}(r, \omega) - \tilde{Z}(r, \omega)) \right)^2 dr, \quad N(t, \omega) = \int_0^t \left( \mathcal{H} V(\tilde{Y}(r, \omega) - \tilde{Z}(r, \omega)) \right) dW_r,$$

$$\mathcal{H} V(\tilde{x}, \tilde{y}) = \frac{D_y V(\tilde{x} - \tilde{y})(\tilde{\sigma}(\tilde{x}) - \tilde{\sigma}(\tilde{y}))}{V(\tilde{x} - \tilde{y})}, \quad \text{and} \quad \tilde{L}^{(2)} = \mathcal{L}^{(2)} + \frac{\partial}{\partial s}. $$

Using a variant of Doob’s martingale inequality (c.f. Lemma 6.2 of [40], see also, [45]), we have

$$\mathbb{P} \left\{ \omega : \sup_{t \in [0, k]} \left( N(t, \omega) - \frac{1}{2} q(t, \omega) \right) > 2 \log k \right\} \leq \frac{1}{k^2}.$$

Moreover, by the $\mathbb{P}$-preserving property of $\theta$, we obtain

$$\mathbb{P} \left\{ \omega : \sup_{t \in [0, k]} \left( N(t, \theta_{-k+1}\omega) - \frac{1}{2} q(t, \theta_{-k+1}\omega) \right) > 2 \log k \right\} \leq \frac{1}{k^2}.$$

Next, we use conditions (2.14) and Borel Cantelli lemma, we have for $k > 1$,

$$\frac{1}{k - 1} \sup_{t \in [k^{-1}k]} \log |\tilde{Y}(t, \omega) - \tilde{Z}(t, \omega)| \leq \frac{1}{p(k - 1)} \sup_{t \in [k^{-1}k]} \log V(\tilde{Y}(t, \omega) - \tilde{Z}(t, \omega)) \leq \frac{1}{p(k - 1)} \sup_{t \in [k^{-1}k]} \log V(\tilde{X}_1(\omega) - \Phi(1, \theta_{-1}\omega, \tilde{y})) + \frac{1}{p(k - 1)} \sup_{t \in [k^{-1}k]} \int_0^t \lambda(r) dr + \frac{2 \log k}{p(k - 1)}. \quad (2.17)$$

Inequality (2.17) and conditions (2.14) together with the cocycle property of $\Phi$, lead to the following estimate

$$\frac{1}{(k - 1)} \sup_{t \in [k^{-1}k]} \log |\Phi(t, \theta_{-k+1}\omega, \tilde{X}_1(\theta_{-k+1}\omega)) - \Phi(t, \theta_{-k+1}\omega, \Phi(1, \theta_{-k}\omega, \tilde{y}))| \quad (2.18)$$

$$= \frac{1}{(k - 1)} \sup_{t \in [0, 1]} \log |\Phi(t, \omega, \Phi(k - 1, \theta_{-k+1}, \tilde{X}_1(\theta_{-k+1}\omega)) - \Phi(t, \omega, \Phi(k, \theta_{-k}\omega, \tilde{y}))| \leq \frac{1}{p(k - 1)} \log V(\tilde{X}_1(\theta_{-k+1}\omega) - \Phi(1, \theta_{-k+1}\omega, \tilde{y})) + \frac{1}{p(k - 1)} \sup_{t \in [k^{-1}k]} \int_0^t \lambda(r) dr + \frac{2 \log k}{p(k - 1)}$$

$$\leq \frac{1}{p(k - 1)} \log V(\tilde{X}_1(\theta_{-k+1}\omega) - \Phi(1, \theta_{-k+1}\omega, \tilde{x})) + \frac{\beta(k - 1)}{pk} + \frac{2 \log k}{p(k - 1)}. \quad (2.19)$$
As \( \log V(\Phi(1, \theta_{-1}\omega, \tilde{y}) - \bar{X}_1(\omega)) \) satisfies condition (2.16), there exist \( \hat{\Omega} \) with \( \mathbb{P}(\hat{\Omega}) = 1 \), a random variable \( \gamma(\omega) \) with \( \mathbb{E}[\gamma(\omega)] < \infty \) such that for any \( \omega \in \hat{\Omega} \) and \( 0 < \varepsilon < -\frac{\beta}{4p} \), there exists \( k(\omega) \in \mathbb{N} \) such that for \( k \geq k(\omega) \),

\[
\sup_{t \in [0,1]} |\Phi(t, \omega, \Phi(k - 1, \theta_{-k+1}, \bar{X}_1(\theta_{-k+1}\omega)) - \Phi(t, \omega, \Phi(k, \theta_{-k}\omega, \tilde{y}))| \\
\leq \gamma(\omega) \exp \left( \frac{1}{2p} \beta + \varepsilon \right) k, \quad \mathbb{P} \text{-a.s.} \tag{2.20}
\]

Finally, the cocycle property of \( \Phi \), leads to

\[
\sup_{t \in [0,\tau]} |\Phi(t, \omega, \Phi(k\tau - \tau, \theta_{-k+\tau}, \bar{X}_1(\theta_{-k+\tau}\omega)) - \Phi(t, \omega, \Phi(k\tau, \theta_{-k}\omega, \tilde{y}))| \\
\leq \hat{\gamma}(\omega) \exp \left( \frac{1}{2p} \beta + \varepsilon \right) k\tau, \quad \mathbb{P} \text{-a.s.} \tag{2.21}
\]

This implies that \( \{\Phi(\cdot, \omega, \Phi(k\tau, \theta_{-k}\omega, \tilde{y})\} \in \mathbb{N} \) is a Cauchy sequence in \( C([0, \tau]; S^1 \times \mathbb{R}^d) \). Letting \( \hat{S}(\cdot, \omega) \) be the limit of this sequence, then by the cocycle property of \( \Phi \), i.e., \( \Phi(t + k\tau, \theta_{-k}\omega, \tilde{y}) = \Phi(t, \omega, \Phi(k\tau, \theta_{-k}\omega, \tilde{y})) \), we have

\[
\lim_{k \to \infty} \Phi(r + k\tau, \theta_{-k}\omega, \tilde{y}) = \hat{S}(r, \omega), \quad \forall r \in [0, \tau]. \tag{2.22}
\]

Next, observe that the random variable \( \hat{S}(0, \omega) \) is \( \mathcal{F}^s_{-\infty} \)-measurable and given the \( \mathbb{P} \)-preserving property of \( \theta \), we can derive the same property by replacing \( \omega \) by \( \theta_t\omega \), \( t \in \mathbb{T} \subseteq \mathbb{R} \). Indeed, we have

\[
\Phi(t, \omega, \Phi(r, \omega, \Phi(k\tau, \theta_{-k}\omega, \tilde{y}))) = \Phi(r + t, \omega, \Phi(k\tau, \theta_{-k}\omega, \tilde{y})) \\
= \Phi(t + r + k\tau, \theta_{-k}\omega, \tilde{y}) \\
= \Phi(r, \theta_t\omega, \Phi(t + k\tau, \theta_{-t-k}\theta_t\omega, \tilde{y})),
\]

so that

\[
\Phi(m\tau, \theta_{m\tau}\omega, \hat{S}(r, \omega)) = \hat{S}(r, \theta_{m\tau}\omega), \quad \forall m \in \mathbb{Z}, \quad r \in [0, \tau]. \tag{2.23}
\]

The continuity of the flow map \( (t, \tilde{y}) \mapsto \Phi(t, \omega, \tilde{y}) \) for almost all \( \omega \) and the equality (2.23), then for any \( t, r \in \mathbb{T} \subseteq \mathbb{R} \) with \( t \geq r \), we have

\[
\Phi(t, \omega, \hat{S}(r, \omega)) = \Phi(t, \omega, \lim_{k \to \infty} \Phi(r + k\tau, \theta_{-k}\omega, \tilde{y})) \\
= \lim_{k \to \infty} \Phi(t, \omega, \Phi(r + k\tau, \theta_{-k}\omega, \tilde{y})) \\
= \lim_{k \to \infty} \Phi(t + r + k\tau, \theta_{-k}\omega, \tilde{y}) = \hat{S}(t + r, \omega). \tag{2.24}
\]

Finally, we show the exponential stability and the uniqueness of the limit \( \hat{S}(r, \omega), \ r \in [0, \tau] \). For this, let \( \tilde{Y}(\omega), \tilde{Z}(\omega) \) be any two \( \mathcal{F}^s_{-\infty} \)-measurable random variables, we deduce from (2.17) that

\[
\lim_{k \to \infty} \sup_{k \in [0,\tau]} |\Phi(t + k\tau, \omega, \tilde{Y}(\omega)) - \Phi(t + k\tau, \omega, \tilde{Z}(\omega))| = 0, \tag{2.25}
\]

exponentially fast, \( \mathbb{P} \)-a.s. In particular, take \( \tilde{Z}(\omega) = \hat{S}(r, \omega) \), we obtain the exponential stability and the uniqueness of \( \hat{S}(r, \omega) \).
**Step II:** Now, it only remains to show the random periodicity of $\tilde{S}(r, \omega)$. We start by using that the coefficients $b, \sigma$ of our SDE are $\tau$-periodic in time, so that for $t \geq s$,

$$
Y_{s,t}^y(\theta_\tau \omega) = y + \int_s^t b(u, Y_{s,u}^y(\theta_\tau \omega))du + \sum_{k=1}^N \int_s^t \sigma_k(u, Y_{s,u}^y(\theta_\tau \omega))dW_{u+\tau}^k(\omega)
$$

$$
= y + \int_{s+\tau}^{t+\tau} b(u, Y_{s-,u-\tau}^y(\theta_\tau \omega))du + \sum_{k=1}^N \int_{s+\tau}^{t+\tau} \sigma_k(u, Y_{s-,u-\tau}^y(\theta_\tau \omega))dW_u(\omega). \quad (2.26)
$$

On the other hand,

$$
Y_{s+t+s}^y(\omega) = y + \int_{s+\tau}^{t+\tau} b(u, Y_{s-u}^y(\omega))du + \sum_{k=1}^N \int_{s+\tau}^{t+\tau} \sigma(u, Y_{s-u}^y(\omega))dW_u(\omega). \quad (2.27)
$$

Then by uniqueness of solution of our SDE and the $\mathbb{P}$-preserving property of $\theta$, the equations (2.26) and (2.27) lead to

$$
Y_{s+t+s}^y(\omega) = Y_{s,t}^y(\theta_\tau \omega), \quad t \geq s, \quad \mathbb{P} - a.s. \quad (2.28)
$$

Now, we return to the lifted flow $\Phi(t, \omega, \cdot) : \mathbb{S}^1 \times \mathbb{R}^N \to \mathbb{S}^1 \times \mathbb{R}^N$,

$$
\Phi(t + k\tau, \theta_{-k\tau}, \omega, \tilde{y}) = (s + t + k\tau \mod \tau, Y_{s+t+k\tau+s}^y(\theta_{-k\tau-s}\omega)), \quad t \geq 0. \quad (2.29)
$$

Define a projection map $\Pi : \mathbb{S}^1 \times \mathbb{R}^N \to \mathbb{S}^1 \times \mathbb{R}^N$, $\tilde{S} \mapsto \Pi \tilde{S} \in \mathbb{R}^N$. Next, from the limit (2.22) and the random periodic property in (2.28) together with the continuity of the projection II, we have

$$
\Pi \tilde{S}(t, \theta_\tau \omega) = \lim_{k \to \infty} \Pi \Phi(t + k\tau, \theta_{-k\tau+s}\omega, \tilde{y})
$$

$$
= \lim_{k \to \infty} Y_{s+t+k\tau+s}^y(\theta_{-k\tau-s}\omega) = \lim_{k \to \infty} Y_{s+t+k\tau+s+t}^y(\theta_{-k\tau-s}\omega)
$$

$$
= \lim_{k \to \infty} \Pi\Big((s + t + k\tau \mod \tau, Y_{s+t+k\tau+s+t}^y(\theta_{-k\tau-s}\omega))\Big)
$$

$$
= \lim_{k \to \infty} \Pi \Phi(t + k\tau + \tau, \theta_{-k\tau+s}\omega, \tilde{y}) = \Pi \tilde{S}(t + \tau, \omega).
$$

The continuity of the flow map $(t, r, y) \mapsto Y_{s,t}^y(\omega)$ for almost all $\omega$ and the equality (2.24), yield

$$
Y_{s,t+s}^y(\omega) = \lim_{k \to \infty} Y_{s,t+k\tau+s}^y(\theta_{-k\tau-s}\omega) \quad (2.29)
$$

This implies that $\Pi \tilde{S}(r, \omega)$ is a random periodic solution of the $\tau$-periodic flow \{$Y_{r,t} : t, r \in \mathbb{T}$\} on $\mathbb{R}^N$ (see Definition 2.3). Set $\tilde{U}(r, \omega) := (r \mod \tau, \Pi \tilde{S}(r, \omega))$ for $r \in \mathbb{T} \subset \mathbb{R}$, it is relative easy to verify that

$$
U(r + \tau, \omega) = \tilde{U}(r, \theta_\tau \omega) \quad (2.31)
$$

and (2.30) leads to

$$
\Phi(t, \theta_\tau \omega, \tilde{U}(r, \omega)) = \Pi \tilde{S}(r, \omega) \quad \frac{r}{r+1} \quad \mathbb{P} - a.s. \quad (2.32)
$$
This implies that $\tilde{U}(r, \omega)$ is a random periodic solution of $\Phi$ on $\mathbb{S}^1 \times \mathbb{R}^N$ and the uniqueness of the limit (2.22), gives us $\tilde{S}(r, \omega) = \tilde{U}(r, \omega)$.

\textbf{Remark 2.9.} In practise, the construction of Lyapunov function is not simple and depends on a specific problem. However, one can construct a polynomial Lyapunov function growing at infinity like $|\hat{t}|^{2m}$, $m \in \mathbb{N} \setminus \{0\}$, for a class of SDEs with coefficients $b \in \mathcal{C}^1(\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^N)$, $\sigma_k \in C_b^2(\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^N; \mathbb{R}^N)$ satisfying the following dissipative conditions:

\begin{equation}
\begin{cases}
\langle b(t,x,y) - b(t,\hat{x},z),(x,y) - (\hat{x},z) \rangle & \leq -K_t \left( |y - z|^2 + |x - \hat{x}|^2 \right), \\
|\sigma_k(t,x,y) - \sigma_k(t,\hat{x},z)| & \leq L_t \left( |y - z| + |x - \hat{x}| \right), \quad 1 \leq k \leq N,
\end{cases}
\end{equation}

such that

\begin{equation}
\limsup_{t \to \infty} \frac{1}{2t} \int_0^t \lambda(s)ds < 0,
\end{equation}

where $\lambda(t) = -K_t + \frac{(p-1)}{2} N L_t^2$ for some $p > 1$. The function $K_t$ is defined by

$$K_t = \liminf_{R \to \infty} K_t(R),$$

where $K_t : \mathbb{R} \to \mathbb{R}$ is a Borel function defined by

$$K_t(R) = \inf \left\{ -\frac{\langle b(t,x,y) - b(t,\hat{x},z),(x,y) - (\hat{x},z) \rangle}{|y - z|^2 + |x - \hat{x}|^2} : |y - z| + |x - \hat{x}| = R \right\}.$$

There are many important class of SDEs in theory and applications satisfying the above dissipative conditons (2.33) – (2.34). The existence and uniqueness of random periodic solution for a class of SDEs satisfying related conditions was proved in [47, 49].

\section{2.3. Ergodic periodic measures on Poincaré sections.}

In this subsection, we discuss the ergodicity of the fast subsystem (2.12) in the random periodic regime. Many studies of invariant measures for Markovian RDS generated by SDEs are normally via stochastic analysis or dynamical systems perspectives (e.g., [1, 12]). From stochastic analysis perspective, invariant measures are investigated via Markov transition function, in this sense, ergodicity is based on the dynamics of Markov evolution. From dynamical description, one investigates random invariant probability measures whose conditional expectation with respect to a sub-algebra of $\mathcal{F}$ has one-to-one correspondence with the invariant measure of the Markov evolution. Here, we are interested in capturing the ergodicity of transition probability function of random periodic solutions (the law of random periodic solutions). Ergodicity on Poincaré sections (PS-ergodicity) is a form of ergodicity suitable in this case. PS-ergodicity of stochastic dynamical systems was recently developed by Feng and Zhao [17], it gives a new perspective and a generalised form of Poincaré-Bendixson theorem for stochastic dynamical systems. We would like to argue that the random periodic measure of the fast subsystem (1.2) is PS-ergodic under Assumption 2.7 and restricted Hömander Lie bracket conditions.

We start with preparatory definitions leading to the presentation of our ergodicity result (Theorem 2.19).

\textbf{Definition 2.10} (Periodic measure [17]). Let $\mathcal{M}$ be a Polish space, a measure $\mu : \mathbb{R} \to \mathcal{P}(\mathcal{M})$ is called a periodc measure of period $\tau > 0$, on the phase space $(\mathcal{M}, \mathcal{B}(\mathcal{M}))$ for the Markovian
stochastic flow \( \{Y_{t,t}(\omega) : t \geq r\} \), if for \( B \in \mathcal{B}(\mathbb{M}) \), the following hold

\[
\mu_{r+r} = \mu_r \quad \text{and} \quad \mu_{t+r}(B) = \int_{\mathbb{M}} P(r, x; t+r, B) \mu_r(dx) = (P^r_{t+r})_r(B), \quad r \in \mathbb{R}, \quad t \in \mathbb{R}^+. \tag{2.35}
\]

It is called a periodic measure with minimal period \( \tau \), if \( \tau > 0 \) is the smallest number such that (2.35) holds.

Let \( \tilde{S} : \mathbb{R} \times \Omega \to S^1 \times \mathbb{R}^N \) be a random periodic solution of the RDS \( \{\Phi(t, \omega) : t \geq 0\} \). Consider a probability measure \( \tilde{\mu}_t \in \mathcal{P}(S^1 \times \mathbb{R}^N) \) defined by

\[
\tilde{\mu}_r(\tilde{A}) := \left( \mathbb{P} \circ \tilde{S}^{-1}(r,) \right) (\tilde{A}) = \mathbb{P}(\{\omega : \tilde{S}(r, \omega) \in \tilde{A}\}), \quad \tilde{A} \in \mathcal{B}(S^1) \otimes \mathcal{B}(\mathbb{R}^N). \tag{2.36}
\]

Then the measure \( \mu_r \) is \( \tau \)-periodic as

\[
\mu_{r+r}(A) = \mathbb{P}\{\omega : \tilde{S}(r + \tau, \omega) \in \tilde{A}\} = \mathbb{P}\{\omega : \tilde{S}(r, \theta_r \omega) \in \tilde{A}\} = \mathbb{P}\{\omega : \tilde{S}(r, \omega) \in \tilde{A}\} = \mu_r(\tilde{A}), \tag{2.37}
\]

Moreover, as it was shown in [17], \( \tilde{\mu}_r \) satisfies (2.35). Thus, the law of random periodic solution satisfies Definition (2.10).

Next, we define a Poincaré section for the lifted Markov semigroup \( \tilde{\mathcal{P}} \) on \( S^1 \times \mathbb{R}^N \).

**Definition 2.11** (Poincaré section for lifted Markov semigroup [17]). The Borel subsets of \( \{M_r : r \in \mathbb{T} \subseteq \mathbb{R}\} \subset \mathcal{B}(S^1) \otimes \mathcal{B}(\mathbb{R}^N) \) are called Poincaré sections for the Markov semigroup \( \tilde{\mathcal{P}} \) on the cylinder \( S^1 \times \mathbb{R}^N \), if

\[
M_{r+r} = M_r, \quad \text{and} \quad \tilde{P}(0, \tilde{y}; t, M_{t+r}) = 1, \quad \forall \tilde{y} \in M_r, \quad t \geq 0. \tag{2.38}
\]

**Remark 2.12.** The choice of Poincaré section is not unique, for example \( M_r = S^1 \times \mathbb{R}^N \) and \( M_r = \text{supp}(\tilde{\mu}_r) \) satisfy the definition of Poincaré section. However, the family \( \{M_r = \text{supp}(\tilde{\mu}_r) : r \in \mathbb{R}\} \) is a minimal Poincaré section or \( k_r \)-irreducible Poincaré section [17]. To see this, fix \( r \in [0, \tau) \) and any open set \( A \subset \mathbb{R}^N \) such that \( \tilde{A}_r := \{r\} \times A \subset M_r = \text{supp}(\tilde{\mu}_r) \) with \( \mu_r(M_r \setminus \tilde{A}_r) > 0 \), we have for all \( \tilde{y} \in M_r \),

\[
\tilde{P}(0, \tilde{y}; k_r \tau + r, \tilde{A}_r) < 1.
\]

This implies that \( \tilde{A}_r \) is not a Poincaré section for the Markov semigroup \( \tilde{\mathcal{P}} \) on the cylinder \( S^1 \times \mathbb{R}^N \).

Let the periodic measure \( \{\tilde{\mu}_r\}_{r \in \mathbb{R}} \subset \mathcal{P}(S^1 \times \mathbb{R}^N) \) be given, we consider its time average \( \bar{\mu} \) over the interval \( [0, \tau) \) defined by

\[
\bar{\mu}(\tilde{A}) = \frac{1}{\tau} \int_0^\tau \tilde{\mu}_r(\tilde{A}) dr, \quad \tilde{A} \in \mathcal{B}(S^1) \otimes \mathcal{B}(\mathbb{R}^N). \tag{2.39}
\]

It is relatively easy to check that this time average measure \( \bar{\mu} \) is invariant under the Markov semigroup \( \tilde{\mathcal{P}} \) (see also [17]). Moreover, from the definition of random periodic solution
(Definition 2.3) and \( \mathbb{P} \)-preserving property of \( \theta \), we have for any \( \tilde{A} \in \mathcal{B}(\mathbb{S}^1 \times \mathbb{R}^N) \),

\[
\bar{\mu}(\tilde{A}) = \frac{1}{\tau} \int_0^\tau \bar{\mu}_r(\tilde{A})dr = \frac{1}{\tau} \int_0^\tau \mathbb{P}\{\omega : \tilde{S}(r, \omega) \in \tilde{A}\}dr
\]

\[
= \frac{1}{\tau} \mathbb{E}\left[ \int_0^\tau \mathbb{I}_\tilde{A}(\tilde{S}(r, \omega))dr \right]
\]

\[
= \mathbb{E}\left[ \frac{1}{\tau} m(\{r \in [0, \tau) : \tilde{S}(r, \omega) \in \tilde{A}\}) \right]
\]

\[
= \mathbb{E}\left[ \frac{1}{\tau} m(\{r \in [u, u + \tau) : \tilde{S}(r, \omega) \in \tilde{A}\}) \right].
\]

This implies that the expected time spent in a Borel set \( \tilde{A} \in \mathcal{B}(\mathbb{S}^1 \times \mathbb{R}^N) \) by the random periodic path \( r \mapsto \tilde{S}(r, \omega) \) over a time interval of exactly one period at any time is independent of the starting point. More precisely, this leads to the notion of PS-ergodicity of periodic measures.

**Definition 2.13 (PS-ergodicity [17]).** A family of \( \tau \)-periodic measures \( \{\bar{\mu}_r\}_{r \in \mathbb{R}} \) is said to be PS-ergodic if for each \( r \in [0, \tau) \), \( \bar{\mu}_r \) as an invariant measure of the Markov semigroup \( \{\tilde{P}^x_{r+k\tau}\}_{k \in \mathbb{N}} \) at the integral multiples of the period \( \tau \) on the Poincaré section \( \mathcal{M}_r \) is ergodic.

Equivalently, a family of \( \tau \)-periodic measures \( \{\bar{\mu}_r\}_{r \in \mathbb{R}} \) is PS-ergodic, if for any \( \tilde{A} \in \mathcal{B}(\mathbb{S}^1 \times \mathbb{R}^N) \) with \( \tilde{A} \subset \mathcal{M}_r \), the following holds

\[
\lim_{m \to \infty} \int_{\mathbb{S}^1 \times \mathbb{R}^N} \left| \int_0^\tau \left[ \frac{1}{m} \sum_{k=0}^{m-1} \tilde{P}(0, \tilde{y}; r + k\tau, \tilde{A}) - \bar{\mu}_r(\tilde{A}) \right]dr \right| \bar{\mu}(d\tilde{y}) = 0. \quad (2.40)
\]

The limit (2.40) is the Krylov-Bogolyubov scheme for periodic measures [17, 27]. Now, given the random periodic paths \( \{\tilde{S}^x(r, \omega) : r \in [0, \tau)\} \), we want to show that \( \{\bar{\mu}_r^x : r \in [0, \tau)\} \) is ergodic under the discrete Markov semigroup \( \{\tilde{P}_{r+k\tau}^x\}_{k \in \mathbb{N}} \) for fixed \( x \in \mathbb{K} \subset \mathbb{R}^d \) under the Assumption 2.7. Precisely, we want to prove the convergence of the Krylov–Bogolyubov scheme for periodic measures. To this end, we start with the following simple and useful Lemma.

**Lemma 2.14.** Let \( \{\bar{\mu}_r^x\}_{r \in [0, \tau)} \), \( x \in \mathbb{K} \) be \( \tau \)-periodic measures generated by the random periodic solutions of the SDE (2.12). If there exist \( 0 < K < \infty \) and \( \lambda \in L^1(\mathbb{R}) \) such that

\[
\lim_{t \to \infty} \sup_{t \in [0, 1]} \int_0^t \lambda(u)du < \beta < 0,
\]

and for all \( f \in C_b(\mathcal{M}_r) \), \( p \geq 1 \),

\[
\left| \tilde{P}_{r+k\tau}^x f - \int_{\mathcal{M}_r} f \bar{\mu}_r^x \right| \leq K ||f||_{\infty} \left[ \exp \left( p^{-1} \int_0^{r+k\tau} \lambda(u)du \right) + \exp \left( \frac{1}{2p} \beta k\tau \right) \right].
\]

Then,

\[
\lim_{m \to \infty} \int_{\mathbb{S}^1 \times \mathbb{R}^N} \left| \int_0^\tau \left[ \frac{1}{m} \sum_{k=0}^{m-1} \tilde{P}(0, \tilde{y}; r + k\tau, \tilde{A}) - \bar{\mu}_r(\tilde{A}) \right]dr \right| \bar{\mu}^x(\tilde{y}) = 0. \quad (2.43)
\]

**Proof.** The proof is relatively straightforward, the idea is to employ density argument. For this, let \( \tilde{A}_r \subset \mathcal{M}_r \subset \mathbb{S}^1 \times \mathbb{R}^N \) be a closed set, take \( f_r = \mathbb{I}_{\tilde{A}_r} \) and let the sequence of functions \( (f_n)_{n \in \mathbb{N}} \)
be defined by

\[ f_n(\tilde{y}) = \begin{cases} 
1, & \text{if } \tilde{y} \in \tilde{A}_r, \\
1 - 2^nd(\tilde{y}, \tilde{A}_r), & \text{if } d(\tilde{y}, \tilde{A}_r) \leq 2^{-n}, \\
0, & \text{if } d(\tilde{x}, \tilde{A}_r) \geq 2^{-n},
\end{cases} \]

where \( d(\tilde{y}, \tilde{A}_r) = \inf\{ |\tilde{y} - \tilde{z}| : \tilde{z} \in \tilde{A}_r \} \), \( \tilde{y} \in \mathcal{M}_r \). Then,

\[ f_n(\tilde{y}) \to f_r(\tilde{y}), \quad \text{as } n \to \infty, \quad \tilde{y} \in \mathcal{M}_r. \]

Consequently, for all \( r \in [0, \tau) \), we have

\[ P_{r+k\tau}^x f_n(\tilde{y}) \to P_{r+k\tau}^x f_r(\tilde{y}) = P_{r+k\tau}^x I_{\tilde{A}_r}(\tilde{y}). \]

This implies that \( P_{r+k\tau}^x f_n \in C_b(\mathcal{M}_r) \), so that by \((2.42)\) and as \( \tilde{\mu}_r^x \) is invariant w.r.t. \( P_{r+k\tau}^x \), we have

\[ |P_{r+k\tau}^x(0, \tilde{y}; r + k\tau, \tilde{A}_r) - \tilde{\mu}_r^x(\tilde{A}_r)| \leq \mu \left[ \exp \left( p^{-1} \int_0^{r+k\tau} \lambda(u)du \right) + \exp \left( \frac{1}{2p} \beta_k \right) \right]. \]

It then follows by covering lemma that for any \( \tilde{A} \in \mathcal{B}(\mathbb{S}^1 \times \mathbb{R}^N) \) with \( \tilde{A} \subset \mathcal{M}_r \),

\[ \int_0^\tau \left| P_{r+k\tau}^x(0, \tilde{y}; r + k\tau, \tilde{A}) - \tilde{\mu}_r^x(\tilde{A}) \right| dr \leq \mu \left[ \int_0^\tau \exp \left( p^{-1} \int_0^{r+k\tau} \lambda(u)du \right) dr + \exp \left( \frac{1}{2p} \beta_k \right) \right]. \]

By the condition \((2.41)\), there exist \( 0 < \tilde{K} < \infty, \quad 0 < \beta < 1 \), such that

\[ \int_0^\tau \left| P_{r+k\tau}^x(0, \tilde{y}; r + k\tau, \tilde{A}) - \tilde{\mu}_r^x(\tilde{A}) \right| dr \leq \tilde{K} \beta^{k\tau}, \]

leading to

\[ \int_{\mathbb{S}^1 \times \mathbb{R}^N} \left| \int_0^\tau \left[ \frac{1}{m} \sum_{k=0}^{m-1} P_{r+k\tau}^x(0, \tilde{y}; r + k\tau, \tilde{A}) - \tilde{\mu}_r^x(\tilde{A}) \right] dr \right| \tilde{\mu}(d\tilde{y}) \leq \tilde{K} \sum_{k=0}^{m-1} \beta^{k\tau}. \]

This gives the desired convergence. \( \Box \)

It is not very simple to verify the inequality \((2.42)\), some non-degenerate conditions would be required. In fact, the sufficient condition for inequality \((2.42)\) to hold is the strong Feller property of the Markov semigroup \( (P_{r+k\tau}^x)_{k \in \mathbb{N}} \). Recall, that a Markov semigroup \( (P_{r+k\tau}^x)_{k \in \mathbb{N}} \) has strong Feller property, if for any \( f \in B_b(\mathbb{S}^1 \times \mathbb{R}^N) \), we have \( P_{r+k\tau}^x f \in C_b(\mathbb{S}^1 \times \mathbb{R}^N) \). Equivalently, a Markov semigroup \( (P_{r+k\tau}^x)_{k \in \mathbb{N}} \) has strong Feller property, if and only if

(i) \((P_{r+k\tau}^x)_{k \in \mathbb{N}}\) is a Feller semigroup, i.e., \( P_{r+k\tau}^x : C_b(\mathbb{S}^1 \times \mathbb{R}^N) \to C_b(\mathbb{S}^1 \times \mathbb{R}^N) \), and
(ii) the family \( \{ P(0, \tilde{y}_m; t + k\tau, \cdot) : m \in \mathbb{N} \} \) is equicontinuous.

The first item follows from the existence of stochastic flows (e.g., \([36, 27]\)) and the second item require non-degeneracy and certain integrability of Malliavin covariance \( \hat{C}^x(\omega, \tilde{y}) \) of the flows of solutions of the fast motion \( Y^\tilde{z}_t \). Intuitively, the strong Feller property states that for any nearby initial data \( \tilde{y} \) and \( \tilde{z} \) and any realisation \( \omega \) of the past of the driving noise, it is possible to construct a coupling between two solutions \( \Phi(t, \omega, \tilde{y}) \) and \( \Phi(t, \omega, \tilde{z}) \) such that with a probability
close to 1 as \( \hat{y} \to \hat{z} \), one has \( \Phi(t, \omega, \hat{y}) = \Phi(t, \omega, \hat{z}) \) for \( t \geq 1 \) (e.g., [22, 24, 25]). As elucidated by Hairer, Mattingly and others (e.g., [22, 25, 26, 21]), one way of achieving such a coupling is via change of measure on driving process for one of the two solutions such that the noises \( W^\hat{y}_t \) and \( W^\hat{z}_t \) driving the solutions \( \Phi(t, \omega, \hat{y}) \) and \( \Phi(t, \omega, \hat{z}) \) respectively, are related by

\[
dW^\hat{z}_t = dW^\hat{y}_t + v^\hat{z}_t dt,
\]

where \( v \) is a control process that steers the solution \( \Phi(t, \omega, \hat{y}) \) towards the solution \( \Phi(t, \omega, \hat{z}) \) (cf., [26]). If one sets \( \hat{z} = \hat{y} + \varepsilon \eta \) and look for controls of the form \( v^{\hat{y}, \hat{z}}_t = \varepsilon v \), then as \( \varepsilon \to 0 \), the random variable \( v \) will induce a deformation onto the solution \( \Phi(t, \omega, \hat{y}) \) after time \( t \) in the form of Malliavin derivative \(^3\) of \( \Phi(t, \omega, \hat{y}) \) at \( \omega \) in the direction of \( v \), i.e.,

\[
D \Phi(t, \omega, \hat{y}) \cdot v = D_v \Phi(t, \omega, \hat{y}) =: A_t v.
\]

On the other hand, the effect of the perturbation of the initial condition by \( \eta \) is given as the directional derivative of the solution \( \Phi(t, \omega, \hat{y}) \) at \( \hat{y} \) in the direction of \( \eta \), i.e.,

\[
D_x \Phi(t, \omega, \hat{y}) \eta = J_{0,t} \eta.
\]

As extensively discussed in [22] for a more general and highly degenerate infinite-dimensional stochastic dynamical systems, in order to prove strong Feller property, the main task is to find a control \( v \) such that

\[
J_{0,t} A_t v = J_{0,t} \eta, \quad \text{or} \quad A_t v = \eta.
\]

(2.44)

To be able to construct the above control process satisfying (2.44) with appropriate integrable condition, we impose the following restricted Hömander’s Lie bracket condition.

**Assumption 2.15.** Let \( b^x \in C_b^{l-1}(\mathbb{R} \times \mathbb{R}^N; \mathbb{R}^N) \) and \( \sigma_k^x \in C_b^{l,\delta}(\mathbb{R} \times \mathbb{R}^N; \mathbb{R}^N), \quad 1 \leq k \leq N \), for \( l \geq 2 \) such that

\[
\dim(\text{span Lie} \{ \sigma_k^x : 1 \leq k \leq N \})(0, x) = N,
\]

(2.45)

or equivalently, there exists \( M(y) =: M \in \mathbb{N}, \) \( C_M(y) =: C_M > 0 \), such that for all \( \eta \in \mathbb{S}^{d-1} \) we have

\[
\sum_{\ell=0}^{M} \sum_{Z \in \Sigma^x_{\ell}} (\eta \cdot Z)^2(0, y) \geq C_M
\]

(2.46)

where \( \Sigma^x_{\ell} = \{ \sigma_k^x : 1 \leq k \leq N \}, \quad \Sigma^x_{\ell+1} = \{ [\sigma_k^x, Z] : 1 \leq k \leq N, \quad Z \in \Sigma^x_{\ell} \} \) and \([F, G]\) is the Lie bracket between the vector fields \( F \) and \( G \) defined by

\[
[F, G](t, y) := D_y G(t, y) F(t, y) - D_y F(t, y) G(t, y).
\]

The restricted Hömander brackets condition above generalises uniform ellipticity condition; this follows by noting that uniform ellipticity condition implies that for every point \((t, y)\) in the neighborhood of \((0, y)\), that the linear span of \( \sigma_k^x(t, y) ; 1 \leq k \leq N \), is the whole of \( \mathbb{R}^N \).

For convenience, we briefly recall derivation on the Wiener space \( \Omega \), we refer interested reader to (e.g., [20, 24, 25, 39, 41, 43, 56] ) for accessible treatment. Let \( \mathcal{H} = L^2([0, T]; \mathbb{R}^N) \) and let \( E \)

\(^3\) see equation (2.47) below.
be a real separable Hilbert space. An $E$-valued random variable $\Theta : \Omega \to E$ will be called smooth if it admit the following representation

$$\Theta = \sum_{j=1}^{M} \vartheta_j \langle W_{t_1}, \ldots, W_{t_m} \rangle v_j,$$

where $\vartheta_j \in \mathcal{C}_b^\infty(\mathbb{R}^{Nm})$, $t_1, \ldots, t_m \in [0, T]$ and $v_1, \ldots, v_M \in E$.

The derivative $D$ of a smooth random variable is a random variable taking values on the Hilbert space $\mathcal{H} \otimes E = L^2([0, T]; \mathbb{R}^N \otimes E)$ given by

$$D_i^t \Theta = \sum_{j=1}^{M} \sum_{k=1}^{m} \frac{\partial \vartheta_j}{\partial x_i^k} \langle W_{t_1}, \ldots, W_{t_m} \rangle \|_{[0, t]}(t) v_j,$$

for each $t \in [0, T]$ and $i = 1, \ldots, N$.

Observe that for any $h \in \mathcal{H}$, the random variable $\sum_{i=1}^{N} \int_0^T D_i^t \Theta h^i(t) dt$ can be interpreted as the directional derivative

$$\frac{d}{dz} \bigg|_{z=0} \Theta \left( \omega(\cdot) + \varepsilon \int_0^T h(s) ds \right),$$

where $W_t(\omega) = \omega(t)$.

Moreover, to establish strong Feller property, the following preparatory lemmas would be useful. We skip the dependence on $x$ in the proofs to simplify notation.

**Lemma 2.16.** Let $b^x, \sigma^x \in \mathcal{C}_b^1(\mathbb{R} \times \mathbb{R}^N; \mathbb{R}^N)$, $1 \leq k \leq N$, $l \geq 2$, and $\Phi(t, \omega, \tilde{y})$ be stochastic flow generated by the lifted SDE (2.12) with the Jacobian matrix $J_{0,t}(\omega, \tilde{y})$ at $\tilde{y} \in S^1 \times \mathbb{R}^N$. Then, for all $t > 0$ and $p \geq 1$, the following holds,

$$\begin{cases}
\int_0^t E \|D_i^t \Phi(u, \omega, \tilde{y})\|^{2p} \leq \left(\frac{N+1}{p L_{ba}}\right)^{2p} \exp(-p L_{ba} r)[\exp(p L_{ba} t) - 1],
\int_0^t E \|D_i^t J_{0,u}\|^{2p} du \leq \frac{L_{ba} (N+1)^2}{p L_{ba}} \exp(-p L_{ba} r)\left[\frac{1}{r} \exp(3p L_{ba} t) - \exp(p L_{ba} t)\right].
\end{cases}$$

(2.48)

where $L_{ba}, \tilde{L}_{ba}$ are constants depending on $b, \sigma$ and their derivatives up to second order.

Moreover, there exist constants $0 < \tilde{C}_j(N, p, L_{ba}) < \infty$, $j = 1, 2,$, depending on $N$, $p$, the coefficients $b, \sigma$ and their derivatives w.r.t $\tilde{y} = (s, y)$ up to second order such that

$$\mathbb{E}\|D_i^t \tilde{C}_u\|^{2p} \leq \exp(-2p L_{ba} r) \left\{ C_1(N, p, L_{ba}) \left[ \frac{1}{12p L_{ba}} \exp(12p L_{ba} u) - \frac{1}{4p L_{ba}} \exp(4p L_{ba} u) \right] \right\}^{1/2}
\leq \exp(8p L_{ba} u) - 1 \right\}^{1/2} + C_2(N, p, L_{ba}) \left[ \exp(4p L_{ba} u) - 1 \right]^{1/2} \left[ \exp(4p L_{ba} u) - 4p L_{ba} u - 1 \right]^{1/2},$$

(2.49)

where $\tilde{C}_u(\omega, \tilde{y}) := \langle D \Phi(u, \omega, \tilde{y}), D \Phi(\omega, \tilde{y}) \rangle$ is the Malliavin covariance of the flow map $\Phi(u, \omega, \tilde{y})$, also defined for $\xi \in S^1 \times \mathbb{R}^N$, as follows

$$A^T \xi = \dot{\sigma}^T (\Phi(u, \omega, \tilde{y}))(J_{0,u}^{-1})^T \xi$$

$$\tilde{C}_u \xi = A^T \xi = \int_0^u J_{0,t}^{-1} \dot{a}(\Phi(\ell, \omega, \tilde{y}))(J_{0,t}^{-1})^T \xi dt.$$

**Proof.** First, use that $b, \sigma \in C_b^1(\mathbb{R} \times \mathbb{R}^N; \mathbb{R}^N)$, $l \geq 2$, and Gronwall’s inequality, for all $u \in [0, t]$ and $p \geq 1$, to obtain

$$\begin{cases}
\mathbb{E}\|J_{0,u}^{-1}\|^{2p} \leq (N+1)^p \exp(p L_{b,\sigma} u),
\mathbb{E}\|J_{0,u}\|^{2p} \leq (N+1)^p \exp(p L_{b,\sigma} u)
\end{cases}$$

(2.50)
Let $u, r \in [0, t]$, then the Malliavin derivative of $\Phi(t, \omega, \tilde{x})$ for $u > r$ is
\[
D^i_r\Phi(u, \omega, \tilde{y}) = \int_r^u D\tilde{b}(\Phi(\ell, \omega, \tilde{y}))D^i_r\Phi(\ell, \omega, \tilde{y})d\ell + \sum_{k=1}^m \int_r^u D\tilde{\sigma}_k(\Phi(\ell, \omega, \tilde{y}))D^i_r\Phi(\ell, \omega, \tilde{y})dW^k_{\ell} + \tilde{\sigma}_i(\Phi(r, \omega, \tilde{y})),
\]
and for $u \leq r$, we have $D^i_r\Phi(u, \omega, \tilde{y}) = 0$. In fact, the Malliavin derivative is simply the composition of the derivative flows $J_{0,u} \circ J_{0,r}^{-1}\tilde{\sigma}_i(\Phi(r, \omega, \tilde{y})) = J_{r,u}\tilde{\sigma}_i(\Phi(r, \omega, \tilde{y}))$ starting from $\tilde{\sigma}_i(\Phi(r, \omega, \tilde{y}))$ at time $u = r$. So, by the regularity of the coefficients $b^x$, $\tilde{\sigma}^x$ and Gronwall’s inequality on $J_{r,u}$, one arrive at
\[
E\|D^i_r\Phi(u, \omega, \tilde{y})\|^{2p} = E\|J_{r,u}\tilde{\sigma}_i(\Phi(r, \omega, \tilde{y}))\|^{2p} \leq (N + 1)^{2p}\exp(pL_{ba}(u - r)).
\]
Integrating over $u \in [0, t]$, we have
\[
\int_0^t E\|D^i_r\Phi(u, \omega, \tilde{y})\|^{2p}du \leq \frac{(N + 1)^{2p}}{pL_{ba}} \exp(-pL_{ba}r) \left[\exp(pL_{ba}t) - 1\right].
\]
For the second assertion, we recall from (e.g., Proposition 5.2 in [23]) that the Malliavin derivative of the Jacobian matrix $J_{0,u}$, $u \in [0, t]$, solves the following variational equation
\[
D^i_rJ_{0,u} = \int_0^u J_{\ell,u}D^2\tilde{b}(\Phi(\ell, \omega, \tilde{y}))\left(J_{0,\ell}, D^i_r\Phi(\ell, \omega, \tilde{y})\right)d\ell + \int_0^u J_{\ell,u}D^2\tilde{\sigma}(\Phi(\ell, \omega, \tilde{y}))\left(J_{0,\ell}, D^i_r\Phi(\ell, \omega, \tilde{y})\right)dW_{\ell}.
\]
So that,
\[
E\|D^i_rJ_{0,u}\|^{2p} \leq 2^{2p-1}E \int_0^u \|J_{\ell,u}D^2\tilde{b}(\Phi(\ell, \omega, \tilde{y}))\left(J_{0,\ell}, D^i_r\Phi(\ell, \omega, \tilde{y})\right)\|^{2p}d\ell + 2^{2p-1}E \int_0^u \|J_{\ell,u}D^2\tilde{\sigma}(\Phi(\ell, \omega, \tilde{y}))\left(J_{0,\ell}, D^i_r\Phi(\ell, \omega, \tilde{y})\right)\|^{2p}d\ell
\]
\[
\leq 2^{2p-1}\tilde{L}_{ba}^{2p}E\left(\|J_{0,u}\|^{2p} \int_0^u \|D^i_r\Phi(\ell, \omega, \tilde{y})\|^{2p}d\ell\right)
\]
\[
\leq 2^{2p-1}\tilde{L}_{ba}^{2p}(N + 1)^{2p}\exp(pL_{ba}u) \int_0^u E\|D^i_r\Phi(\ell, \omega, \tilde{y})\|^{2p}d\ell
\]
\[
\leq \frac{2^{p-1}\tilde{L}_{ba}^{2p}(N + 1)^{3p}}{pL_{ba}} \exp(pL_{ba}(u - r)) \left[\exp(2pL_{ba}u) - 1\right].
\]
It then follows that
\[
\int_0^t E\|D^i_rJ_{0,u}\|^{2p}du \leq \frac{\tilde{L}_{ba}^{2p}(N + 1)^{2p}}{pL_{ba}} \exp(-pL_{ba}r) \left[\frac{1}{3} \exp(3pL_{ba}t) - \exp(pL_{ba}t)\right].
\]
For the final assertion, by product rule of differentiation, we have
\[
D^i_r\tilde{C}_u = \int_0^u D^i_r[J_{0,\ell}^{-1}\tilde{a}(\Phi(\ell, \omega, \tilde{y}))(J_{0,\ell}^{-1})^T]d\ell + \int_0^t J_{0,\ell}^{-1}D^i_r[\tilde{a}(\Phi(\ell, \omega, \tilde{y}))(J_{0,\ell}^{-1})^T]d\ell + \int_0^t J_{0,\ell}^{-1}\tilde{a}(\Phi(\ell, \omega, \tilde{y}))D^i_r[(J_{0,\ell}^{-1})^T]d\ell.
\]
So that
\[
E\|D^1_t\hat{C}_u\|^{2p} \leq 4^{2p-1} \int_0^u E\|J_{0,\ell}^{-1}D^1_t[J_{0,\ell}J_{0,\ell}^{-1}\hat{a}(\Phi(\ell,\omega,\tilde{y}))](J_{0,\ell}^{-1})^T\|^p d\ell
\]
\[+ 4^{2p-2} \int_0^u E\|J_{0,\ell}^{-1}\}2p\|D^1_t\Phi(\ell,\omega,\tilde{y}))\|^{4p} d\ell
\]
\[\leq 4^{2p-1} \int_0^u \left[ E\|J_{0,\ell}^{-1}\|^p\|D^1_t\Phi(\ell,\omega,\tilde{y}))\|^{4p}\right]^{1/2} \left[ E\|D^1_t\Phi(\ell,\omega,\tilde{y}))\|^{4p}\right]^{1/2} d\ell
\]
\[+ 4^{2p-2} \int_0^u \left[ \|J_{0,\ell}^{-1}\|^p\|D^1_t\Phi(\ell,\omega,\tilde{y}))\|^{4p}\right]^{1/2} \left[ E\|D^1_t\Phi(\ell,\omega,\tilde{y}))\|^{4p}\right]^{1/2} d\ell
\]
\[\leq 4^{2p-1} \left( \int_0^u E\|J_{0,\ell}^{-1}\|^p\|D^1_t\Phi(\ell,\omega,\tilde{y}))\|^{4p} d\ell \right)^{1/2} \left( \int_0^u \left[ E\|D^1_t\Phi(\ell,\omega,\tilde{y}))\|^{4p}\right]^{1/2} d\ell \right)^{1/2}
\]
\[+ 4^{2p-2} \left( \int_0^u \left[ E\|J_{0,\ell}^{-1}\|^p\|D^1_t\Phi(\ell,\omega,\tilde{y}))\|^{4p}\right]^{1/2} d\ell \right)^{1/2} \left( \int_0^u \left[ E\|D^1_t\Phi(\ell,\omega,\tilde{y}))\|^{4p}\right]^{1/2} d\ell \right)^{1/2} d\ell
\]

In a similar fashion as in the first and second assertions, there exist constants \( \hat{C}_j(N,p,L_{bs}) < \infty \), \( j = 1, 2, \) depending on \( N, p \), the coefficients \( b, \sigma \) and their derivatives up to second order, such that
\[
E\|D^1_t\hat{C}_u\|^{2p} \leq \exp(-2pL_{bs}r) \left\{ C_1(N,p,L_{bs}) \left[ \frac{1}{12pL_{bs}} \exp(12pL_{bs}u) - \frac{1}{4pL_{bs}} \exp(4pL_{bs}u) \right]^{1/2}
\]
\[\times [\exp(8pL_{bs}u) - 1]^{1/2} + C_2(N,p,L_{bs}) \left[ \exp(4pL_{bs}u) - 1 \right]^{1/2} \left[ \exp(4pL_{bs}u) - 4pL_{bs}u - 1 \right]^{1/2} \right\}
\]

\[\square\]

**Lemma 2.17.** Let \( b^x, \sigma^x \in C^1_{0}(\mathbb{R} \times \mathbb{R}^N; \mathbb{R}^N) \) \( 1 \leq k \leq m \) for \( l \geq 2 \), be \( \tau \)-periodic in time. If the assumptions 2.7 and 2.15 are satisfied. Then, there exists \( 0 < K_1, K_2 < \infty \) such that
\[
\left\{ \begin{array}{ll}
E \left( \int_0^u |h_\eta(u)|^2 du \right) & \leq K_1 |\eta|^2, \\
E \left( \int_0^u \text{tr}[D_u h_\eta(r)D_r h_\eta(u)] du dr \right) & \leq K_1 |\eta|^2.
\end{array} \right.
\]

(2.51)

for all \( \eta \in \mathbb{S}^1 \times \mathbb{R}^N \), where \( h_\eta(u) := A_u^{T}\hat{C}_u^{-1} J_{0,u}\eta \).

**Proof.** To derive the first estimate, we notice that
\[
|h_\eta(u)|^2 \leq \|A^T_u\|^2 \|\hat{C}_u^{-1}\|^2 \|J_{0,u}\|^2 |\eta|^2
\]
(2.52)

The restricted Hörmander condition 2.15, implies there exists \( M > 0 \) such that (e.g., [20, 43]),
\[
\sup_{u \in [0,l]} \|\hat{C}_u^{-1}\|^2 < M,
\]
(2.53)

and since \( \sigma^x \in C^1_{0}(\mathbb{R} \times \mathbb{R}^N; \mathbb{R}^N) \), \( 1 \leq k \leq N, l \geq 2 \), we have
\[
\|A^T_u\|^2 \leq L_0^2 \|J_{0,u}\|^2.
\]
(2.54)
Finally, substituting (2.53), (2.54) and (2.50) into (2.52), then for all $t > 0$, we have
\[
\mathbb{E} \left( \int_0^t \left| h_\eta(u) \right|^2 du \right) \leq (N + 1)^2 ML^2 \frac{1}{2L_b}\left( \exp(2L_b\sigma t) - 1 \right) |\eta|^2 \\
= K^1_t |\eta|^2, \quad \eta \in \mathcal{S}^1 \times \mathbb{R}^d. \tag{2.55}
\]
Now, we turn to the second estimate, we note that
\[
\mathbb{E} \left( \int_0^t \int_0^t \text{trace}[\mathcal{D}_u h_\eta(r) \mathcal{D}_r h_\eta(u)]drdu \right) = \mathbb{E} \left( \int_0^t \int_0^t \|\mathcal{D}_u h_\eta(u)\|^2_{\text{HS}} drdu \right) \\
= \mathbb{E} \left( \sum_{i=1}^m \int_0^t \int_0^t |\mathcal{D}_r^i h_\eta(u)|^2 du dr \right), \tag{2.56}
\]
and
\[
\mathbb{E} |\mathcal{D}_r^i h_\eta(u)|^2 \leq \mathbb{E} |\mathcal{D}_i^1 [A^T \hat{C}^{-1} J_{0,\eta}]|^2 \\
\leq 4\mathbb{E} |\mathcal{D}_i^1 [A^T \hat{C}^{-1} J_{0,\eta}]|^2 + 4\mathbb{E} |A^T \hat{C}^{-1} \mathcal{D}_i^1 [J_{0,\eta}]|^2 + 4\mathbb{E} |A^T \mathcal{D}_r^i [\hat{C}^{-1}] J_{0,\eta}]|^2 \\
=: 4\mathbb{E} |\mathcal{I}_1|^2 + 4\mathbb{E} |\mathcal{I}_2|^2 + 4\mathbb{E} |\mathcal{I}_3|^2. \tag{2.57}
\]
We derive bound on $\mathbb{E} |\mathcal{I}_1|^2$ as follows
\[
\mathbb{E} |\mathcal{I}_1|^2 \leq \mathbb{E} \left[ \|\mathcal{D}_u^4 A_t\|^2 \right] \|\hat{C}^{-1}\| \|J_{0,t}\|^2 |\eta|^2 \leq M\mathbb{E} \|\mathcal{D}_u^4 A_t\|^4 \|J_{0,t}\|^4 |\eta|^2 \\
\leq M(N + 1)^2 \exp(2L_b\sigma t) \mathbb{E} |\mathcal{D}_u^4 A_t|^4 |\eta|^2. \tag{2.58}
\]
It remains to find bound for $\mathbb{E} |\mathcal{D}_r^i A_t|^4$. To this end, we have,
\[
\mathbb{E} (\|\mathcal{D}_u^4 A_t\|^4) \leq 8L^4 \int_0^t \mathbb{E} |\mathcal{D}_r^i J_{0,t}^{-1}|^4 du + 8L^4 \int_0^t \mathbb{E} |\mathcal{D}_r^i \Phi(u, \omega, \tilde{x}) J_{0,t}^{-1}|^4 du \tag{2.59}
\]
Recall that the Fréchet derivative of a square matrix $A$ in the direction of $H$ is given by $D_H A^{-1} = -A^{-1} H A^{-1}$, so that by chain rule, we have
\[
\int_0^t \mathbb{E} |\mathcal{D}_r^i J_{0,t}^{-1}|^4 du \leq \int_0^t \mathbb{E} |J_{0,u}^{-1} D^i_r [J_{0,u}] J_{0,u}^{-1}|^4 du \\
\leq \int_0^t \left[ \mathbb{E} \|J_{0,u}^{-1}\|^8 \right]^{1/2} \left[ \mathbb{E} |\mathcal{D}_r^i J_{0,u}|^4 \right]^{1/2} du \\
\leq (N + 1)^2 \int_0^t \exp(2L_b\sigma u) \mathbb{E} |\mathcal{D}_r^i J_{0,u}|^4 \left[ \mathbb{E} \|J_{0,u}^{-1}\|^8 \right]^{1/2} du \\
\leq (N + 1)^2 \left( \int_0^t \exp(4L_b\sigma u) du \right)^{1/2} \left( \int_0^t \mathbb{E} |\mathcal{D}_r^i J_{0,u}|^4 du \right)^{1/2} \\
\leq \frac{(N + 1)^2}{2L_b^{1/2}} \left( \exp(4L_b\sigma t) - 1 \right)^{1/2} \left( \int_0^t \mathbb{E} |\mathcal{D}_r^i J_{0,u}|^4 du \right)^{1/2},
\]
and,
\[
\int_0^t \mathbb{E} |\mathcal{D}_r^i \Phi(u, \omega, \tilde{y}) J_{0,u}^{-1}|^4 du \leq \frac{(N + 1)^2}{2L_b^{1/2}} \left[ \exp(4L_b\sigma t) - 1 \right]^{1/2} \mathbb{E} \left( \int_0^t \mathbb{E} |\mathcal{D}_r^i \Phi(u, \omega, \tilde{y})|^8 du \right)^{1/2}.
\]
Now, we apply Lemma 2.16 to obtain

\[
\int_0^t \mathbb{E}\|D^i_r J_{0,u}^{-1}\|^4 du = \int_0^t \mathbb{E}\|D^i_r \Phi(u, \omega, \tilde{y}) J_{0,u}^{-1}\|^4 du \\
\leq \frac{\tilde{L}^4_{ba}(N + 1)^6}{2L_{ba}} \exp(-\frac{1}{2} L_{ba}r) \left[ \frac{1}{3} \exp(12L_{ba}t) - \exp(4L_{ba}t) \right]^{1/2} \left[ \exp(4L_{ba}t) - 1 \right]^{1/2} \tag{2.60}
\]

and,

\[
\int_0^t \mathbb{E}\|D^i_r \Phi(u, \omega, \tilde{y}) J_{0,u}^{-1}\|^4 du \leq \frac{(N + 1)^4}{2\sqrt{2}L_{b,\sigma}} \left[ \exp(4L_{ba}t) - 1 \right]^{1/2} \exp(-L_{ba}r) \left[ \exp(2L_{ba}t) - 1 \right]^{1/2} \tag{2.61}
\]

Substituting (2.60) and (2.61) into (2.59), we have

\[
\mathbb{E}\left(\|D^i_r A_i \|^4\right) \leq 4\tilde{L}^4_{\sigma} T_{ba}(N + 1)^6 \exp(-\frac{1}{2} L_{ba}r) \left[ \exp(12L_{ba}t) - \exp(4L_{ba}t) \right]^{1/2} \left[ \exp(4L_{ba}t) - 1 \right]^{1/2} + \frac{4(N + 1)^4}{\sqrt{2}L_{b,\sigma}} \exp(-\frac{1}{2} L_{ba}r) \left[ \exp(4L_{ba}t) - 1 \right]^{1/2} \left[ \exp(2L_{ba}t) - 1 \right]^{1/2} \tag{2.62}
\]

Put the inequality (2.62) in (2.58), we have

\[
\mathbb{E}|I|^2 \leq 2M(N + 1)^4 \exp(-\frac{1}{4} L_{ba}r) \left( (N + 1)^2 \tilde{L}^4_{\sigma} L_{ba} \left[ \exp(12L_{ba}t) - \exp(4L_{ba}t) \right] \right)^{1/2} + \left[ \exp(2L_{ba}t) - 1 \right]^{1/2} \frac{1}{\sqrt{2}L_{ba}} \left[ \exp(16L_{ba}t) - \exp(8L_{ba}t) \right]^{1/4} |\eta|^2 \tag{2.63}
\]

Next, we derive bound on $\mathbb{E}|I_2|^2$ as follows

\[
\mathbb{E}|I_2|^2 = \mathbb{E}\left|A^T \tilde{C}^{-1} D^i_r [J_{0,u}] J_{0,u} \eta \right|^2 \\
\leq \mathbb{E} \left[ \|A\|^2 \|\tilde{C}^{-1}\|^2 \|D^i_r [J_{0,u}] J_{0,u}\|^2 \|\eta\|^2 \right] \tag{2.64}
\]

\[
\leq M^2 L_{\sigma}^2 \mathbb{E} \left[ \|J_{0,u}\|^4 \|J_{0,u}\|^2 \|D^i_r [J_{0,u}] J_{0,u}\|^2 \|\eta\|^2 \right] \leq M^2 L_{\sigma}^2 \mathbb{E} \left[ \|J_{0,u}\|^4 \|J_{0,u}\|^4 \left[ \|D^i_r [J_{0,u}] J_{0,u}\|^4 \right]^\frac{1}{2} \left[ \|\eta\|^2 \right]^\frac{1}{2} \right] \leq 2M^2 (N + 1)^2 \frac{\tilde{L}^2_{ba}(N + 1)^3}{L_{ba}^{1/2}} \exp(L_{ba}(2u - r)) \left[ \exp(2L_{ba}t) - 1 \right]^{1/2} |\eta|^2 \tag{2.65}
\]

Finally, for $\mathbb{E}|I_3|^2$, again, we use that the Fréchet derivative of a square matrix $A$ in the direction of $H$ is given by $D_H A^{-1} = -A^{-1} H A^{-1}$, we have

\[
D^i_r [\tilde{C}^{-1}] = -[D^i_r \tilde{C}] \tilde{C}^{-1}
\]
We obtain the following estimate,

\[ \mathbb{E}|I_3|^2 = \mathbb{E} |A^\varepsilon | = \mathbb{E} \left| \dot{C}_u^\varepsilon \right|^2 J_{0,u} \eta^2 \]

Then, by Lemma 2.16, we have constants 0 < \tilde{C}_j(N, p, L_{b\sigma}) < \infty, j = 1, 2,\ldots, depending on N, p = 2, the coefficients b, \sigma and their derivatives up to second order such that

\[ \mathbb{E}|I_3|^2 \leq \exp(2L_{b\sigma}(2u - r)) \left\{ M^2 C_1(N, 2, L_{b\sigma}) \left( \frac{1}{24L_{b\sigma}} \exp(24L_{b\sigma} u) - \frac{1}{8L_{b\sigma}} \exp(8L_{b\sigma} u) \right) \right\}^{1/2} \left[ \exp(16L_{b\sigma} u) - 1 \right]^{1/2} + C_2(d, 2, L_{b\sigma}) \left[ \exp(8L_{b\sigma} u) - 1 \right]^{1/2} \left[ \exp(8L_{b\sigma} u) - 8L_{b\sigma} u - 1 \right]^{1/2} \right\} \frac{1}{\eta^2}. \]

If we put the estimates (2.63), (2.64) and (2.66) into the inequality (2.57), and then substitute into the double integral (2.56), we obtain the desired bound.

With the above preparatory estimates at hand, we derive the strong Feller property of the Markov semigroup \((\tilde{P}_t)_{t \geq 0}\); we shall skip the dependence on x to simplify notation.

**Proposition 2.18.** Let \( b^x, \sigma^x \in C^{l,\delta}_b(\mathbb{R} \times \mathbb{R}^N; \mathbb{R}^N) \) 1 \( \leq l \leq N \), for \( l \geq 2 \), be \( \tau \)-periodic in time. If the assumptions 2.7 and 2.15 are satisfied. Then, there exists 0 < \( K_t < \infty \) such that for \( \tilde{x}, \tilde{y} \in S^1 \times \mathbb{R}^d \) with \( t \in [0, \tau] \), and \( \varphi \in C_b(S^1 \times \mathbb{R}^N) \), we have

\[ |\tilde{P}_t \varphi(\tilde{y}) - \tilde{P}_t \varphi(\tilde{z})| \leq K_t ||\varphi||_{\infty} |\tilde{y} - \tilde{z}|. \]  

**Proof.** First, recall the Malliavin integration by part formula (e.g., [20, 24, 25, 39, 43, 4]) yield

\[ D(\tilde{P}_t \varphi)(\tilde{y}) \eta = \mathbb{E} \left( \varphi(\Phi(t, \omega, \tilde{y})) \int_0^t h_n(u) \ast d\tilde{W}_u \right), \quad \varphi \in C^1_b(S^1 \times \mathbb{R}^N), \]

where \( h_n(u) := \tilde{\sigma}^T J_{u,t} \tilde{C}_t J_{0,t} \eta, \) for \( \eta \in S^1 \times \mathbb{R}^d \), and \( J_{u,t} = J_{0,t} \circ J_{0,u}^{-1} \). Note that \( h_n \) need not be adapted to \( \sigma \{ W_t : t \geq 0 \} \), however, one can interprete the stochastic integral (2.68) in the sense of Skorokhod (cf. [20]).

Next, since \( C^1_b(S^1 \times \mathbb{R}^N) \) is dense in \( C_b(S^1 \times \mathbb{R}^N) \), we have \( (\varphi_n)_{n \in \mathbb{N}} \subset C^1_b(S^1 \times \mathbb{R}^N) \) such that

\[ \lim_{n \to \infty} \tilde{P}_t \varphi_n(\tilde{y}) = \tilde{P}_t \varphi(\tilde{y}), \]

\[ \lim_{n \to \infty} D(\tilde{P}_t \varphi_n)(\tilde{y}) \eta = \mathbb{E} \left( \varphi(\Phi(t, \omega, \tilde{y})) \int_0^t h_n(u) \ast d\tilde{W}_u \right), \]

for all \( \varphi \in C_b(S^1 \times \mathbb{R}^N) \), convergence being uniform (e.g., [13]). On the other hand, the regularity of the coefficients \( \tilde{b}, \tilde{\sigma} \) and assumption 2.15 ensure the existence of a function 0 < \( \tilde{\rho}_t \in C^1_b(S^1 \times \mathbb{R}^N) \),
\( \mathbb{R}^N \times \mathcal{C}_b^2(\mathbb{S}^1 \times \mathbb{R}^N) \), such that \( \tilde{P}(0, \tilde{y}, t, d\tilde{z}) = \tilde{\rho}_t(\tilde{y}, \tilde{z})d\tilde{z} \). This implies that
\[
\lim_{n \to \infty} D(\tilde{P}_t \varphi_n)(\tilde{y}) = \lim_{n \to \infty} \int_{\mathbb{S}^1 \times \mathbb{R}^N} \varphi_n(\tilde{y}) D\tilde{\rho}_t(\tilde{y}, \tilde{z})d\tilde{z} = \int_{\mathbb{S}^1 \times \mathbb{R}^N} \varphi(\tilde{y}) D\tilde{\rho}_t(\tilde{y}, \tilde{z})\eta d\tilde{z} = D(\tilde{P}_t \varphi)(\eta). \tag{2.70}
\]
The equalities (2.69) and (2.70), yield (2.68) for all \( \varphi \in \mathcal{C}_b(\mathbb{S}^1 \times \mathbb{R}^N) \). By Cauchy Schwartz inequality, we have
\[
|D(\tilde{P}_t \varphi)(\tilde{x})\eta| \leq \sqrt{(P_t \varphi^2)(\tilde{x})} \left( \mathbb{E} \left| \int_0^t h_\eta(u) \ast dW_u \right|^2 \right)^{1/2}, \quad \varphi \in \mathcal{C}_b(\mathbb{S}^1 \times \mathbb{R}^N) \tag{2.71}
\]
Then by generalised Itô isometry (cf. [20, 43]) and Lemma 2.17, we have
\[
\mathbb{E} \left| \int_0^t h_\eta(u) \ast dW_u \right|^2 = \mathbb{E} \left( \int_0^t |h_\eta(u)|^2 du \right) + \mathbb{E} \left( \int_0^t \int_0^t \text{trace}[D_u h_\eta(r)D_r h_\eta(u)]du dr \right) \leq K_t |\eta|^2
\]
It follows from (2.71), that
\[
|D(\tilde{P}_t \varphi)(\tilde{y})\eta| \leq K_t \|\varphi\|_{\infty} |\eta|, \quad \tilde{y}, \eta \in \mathbb{S}^1 \times \mathbb{R}^N.
\]
Finally, let \( \tilde{\xi}_\ell = \ell \tilde{y} + (1 - \ell)\tilde{x} \), \( \tilde{y}, \tilde{z} \in \mathbb{S}^1 \times \mathbb{R}^N \), \( \ell \in [0, 1] \) and \( \eta = \tilde{y} - \tilde{z} \), by the mean value theorem, we have
\[
|\tilde{P}_t \varphi(\tilde{y}) - \tilde{P}_t \varphi(\tilde{z})| \leq \int_0^1 |D(\tilde{P}_t \varphi)(\tilde{\xi}_\ell)\eta| d\ell \leq K_t \|\varphi\|_{\infty} |\tilde{y} - \tilde{z}|
\]
\[ \square \]
\textbf{Theorem 2.19.} Suppose that the coefficients of the SDE (1.2) are \( \tau \)-periodic in time and satisfy Assumptions 2.7 and 2.15. Suppose further that
\[
\sup_{t \in [0, 1]} \mathbb{E} \left[ V(\tilde{y} - \Phi(t, \omega, \tilde{y})) \right] < \infty, \tag{2.72}
\]
then, for fix \( x \in \mathbb{K} \), the family of periodic measures \( (\tilde{\mu}_r^x)_{r \in [0, \tau]} \) induced by the random periodic path \( \tilde{S}^x(r, \omega) \) is PS-ergodic.
\emph{Proof.} The idea of the proof is to use the construction carried out in Theorem 2.8 to verify the conditions of Lemma 2.14. To this end, we skip the dependence on \( x \) to simplify notation and divide the proof into three steps.

\textbf{Step I:} We show that for any \( \tilde{Y}, \tilde{Z} \in L^p(\Omega, \mathcal{F}_\tau, \mathbb{P}) \), \( p \geq 1 \) and for \( k \in \mathbb{Z} \), the following inequality holds
\[
\mathbb{E}|\Phi(r + k\tau, \omega, \tilde{Y}(\omega)) - \Phi(r + k\tau, \omega, \tilde{Z}(\omega))|^p \leq C \exp \left( \int_{-r}^{k\tau} \lambda(u + r)du \right) \mathbb{E}|\tilde{Y}(\omega) - \tilde{Z}(\omega)|^p. \tag{2.73}
\]
To derive this, set \( \alpha(t) = \exp \left( -\int_0^t \lambda(r)ds \right) \) and
\[
M(t, \omega, \tilde{y}) = M(t, \omega, (s, y)) = \int_0^t \begin{pmatrix} 0 & 0 \\ \sum_{k=1}^N \sigma_k(r + s, x, y) \\ 0 \end{pmatrix} \begin{pmatrix} dW_r^0 \\ dW_r^k \end{pmatrix},
\]
then, by Itô’s formula ([35, 36, 45]), we obtain
\[
d\left(\alpha(t)V(\Phi(t, \omega, \bar{Y}) - \Phi(t, \omega, \bar{Z}))\right) \\
= -\lambda(t)\alpha(t)V(\Phi(t, \omega, \bar{Y}) - \Phi(t, \omega, \bar{Z}))dt + \alpha(t)L^{(2)}V(\Phi(t, \omega, \bar{Y}) - \Phi(t, \omega, \bar{Z}))dt \\
+ \alpha(t)V(\Phi(t, \omega, \bar{Y}) - \Phi(t, \omega, \bar{Z}))d\left(M(\Phi(t, \omega, \bar{Y}) - M(\Phi(t, \omega, \bar{Z}))\right) \\
\leq -\lambda(t)\alpha(t)V(\Phi(t, \omega, \bar{Y}) - \Phi(t, \omega, \bar{Z}))dt + \lambda(t)\alpha(t)V(\Phi(t, \omega, \bar{Y}) - \bar{X}(t, \omega, \bar{Z}))dt \\
+ \alpha(t)V(\Phi(t, \omega, \bar{Y}) - \Phi(t, \omega, \bar{Z}))d\left(M(\Phi(t, \omega, \bar{Y}) - M(\Phi(t, \omega, \bar{Z}))\right).
\]

This implies for \(k \in \mathbb{N}\),
\[
\mathbb{E}\left(V(\Phi(r + k\tau, \omega, \bar{Y}(\omega)) - \Phi(r + k\tau, \omega, \bar{Z}(\omega))\right) \leq \exp\left(\int_{-r}^{r} \lambda(u + r)du\right)\mathbb{E}[V(\bar{Y}(\omega) - \bar{Z}(\omega))]
\]

We use that \(|y|^p \leq V(t, y)|, p \geq 1\), to obtain for \(k \in \mathbb{N}\),
\[
\mathbb{E}\left(|\Phi(r + k\tau, \omega, \bar{Y}(\omega)) - \Phi(r + k\tau, \omega, \bar{Z}(\omega))|^p\right) \leq \mathbb{E}\left(V(\Phi(r + k\tau, \omega, \bar{Y}(\omega)) - \Phi(r + k\tau, \omega, \bar{Y}(\omega))\right)
\]
\[
\leq C \exp\left(\int_{-r}^{r} \lambda(u + r)du\right)\mathbb{E}|\bar{Y}(\omega) - \bar{Z}(s, \omega)|^p.
\]

**Step II:** In this step, we show that there exists \(0 < \tilde{C} < \infty\) such that for any \(\bar{Y} \in L_p(\Omega, \mathcal{F}_{\tau}, \mathbb{P})\), the following inequality holds
\[
\mathbb{E}|\Phi(r + k\tau, \omega, \bar{Y}(\omega)) - \bar{S}(r, \theta_{k\tau}\omega)|^p \leq \tilde{C} H_r(k\tau)
\] (2.74)

where \(H_r(k\tau) = \eta_r(k\tau) + \exp\left(\frac{1}{2} \beta k\tau\right)\) and
\[
\eta_r(k\tau) = \exp\left(\int_{-r}^{r} \lambda(u + r)du\right) \sum_{j=1}^{k-1} \exp\left(\int_{0}^{j\tau} \lambda(u)du\right), \quad k \in \mathbb{N}.
\]

To see this, recall a construction done in Theorem 2.8, precisely, equation (2.22) reads
\[
\bar{S}(r, \omega) = \lim_{k \to \infty} \Phi(r + k\tau, \theta_{-k\tau}\omega, \bar{y}), \quad r \in [0, \tau].
\]

As \(\bar{S}(r + k\tau, \omega) = \bar{S}(r, \theta_{k\tau}\omega)\), then by equation (2.73) in step I, estimate (2.21) and triangle inequality, we have
\[
\mathbb{E}\left(|\Phi(r + k\tau, \omega, \bar{y}) - \bar{S}(r, \theta_{k\tau}\omega)|^p\right) \\
\leq \mathbb{E}\left(|\Phi(r + k\tau, \omega, \bar{y}) - \Phi(k\tau, \omega, \Phi(r + k\tau, \theta_{-k\tau}\omega, \bar{y}))|^p\right) + \tilde{C}(\omega) \exp\left(\frac{1}{2} \beta k\tau\right) \\
= \mathbb{E}\left(|\Phi(r + k\tau, \omega, \bar{y}) - \Phi(r + 2k\tau, \theta_{-k\tau}\omega, \bar{y})|^p\right) + \tilde{C}(\omega) \exp\left(\frac{1}{2} \beta k\tau\right) \\
= \mathbb{E}\left(|\Phi(r + k\tau, \omega, \bar{y}) - \Phi(r + k\tau, \omega, \Phi(k\tau, \theta_{-k\tau}\omega, \bar{y}))|^p\right) + \tilde{C}(\omega) \exp\left(\frac{1}{2} \beta k\tau\right) \\
\leq C \exp\left(\int_{-r}^{r} \lambda(u + r)du\right) \mathbb{E}|\bar{y} - \Phi(k\tau, \theta_{-k\tau}\omega, \bar{y})|^p + \tilde{C}(\omega) \exp\left(\frac{1}{2} \beta k\tau\right),
\]
where \( \hat{\gamma} \) and \( \beta \) are random variables appearing in (2.21). Now, set \( \beta_r(t) := \exp \left( \int_{-r}^{t} \lambda(u + r) \, du \right) \), by the cocycle property of \( \Phi \) and step I again, we have

\[
C\beta_r(k\tau)E|\tilde{y} - \Phi(k\tau, \theta_{-k\tau}, \tilde{y})|^p
\leq C\beta_r(k\tau) \left\{ \mathbb{E} \left[ |\tilde{y} - \Phi(1, \theta_{-1}, \tilde{y})|^p + |\Phi(1, \theta_{-1}, \tilde{y}) - \Phi(2, \theta_{-2}, \tilde{y})|^p \right]
+ \cdots + |\Phi([k\tau] - 1, \theta_{1-[k\tau]}, \tilde{y}) - \Phi([k\tau], \theta_{-1}, \tilde{y})| \right\} 
\]

\[
= C\beta_r(k\tau) \left\{ \mathbb{E} \left[ |\tilde{y} - \Phi(1, \theta_{-1}, \tilde{y})|^p + |\Phi(1, \theta_{-1}, \tilde{y}) - \Phi(1, \theta_{-1}, \Phi(1, \theta_{-1}, \tilde{y}))|^p \right]
+ \cdots + |\Phi([k\tau] - 1, \theta_{1-[k\tau]}, \tilde{y}) - \Phi([k\tau], \theta_{-1}, \Phi(1, \theta_{-1}, \tilde{y}))| \right\} 
\]

\[
\leq C\beta_r(k\tau) \left\{ C\beta_0(1) + \cdots + C\beta_0([k\tau] - 1) \right\} \mathbb{E}|\tilde{y} - \Phi(1, \theta_{-1}, \tilde{y})|^p 
= C^2\beta_r(k\tau) \sum_{j=1}^{[k\tau]-1} \beta_0(j) \mathbb{E}|\tilde{y} - \Phi(1, \theta_{-1}, \tilde{y})|^p. 
\]

Consequently, we have

\[
\mathbb{E} \left[ |\Phi(r + k\tau, \omega, \tilde{y}) - \hat{S}(r, \theta_r, \omega)|^p \right] \leq C^2\eta_r(k\tau) \mathbb{E}|\tilde{y} - \Phi(1, \theta_{-1}, \tilde{y})|^p + \hat{\gamma}(\omega) \exp \left( \frac{1}{2} \beta k\tau \right) 
\]

and using that \( |\tilde{y}|^p \leq V(\tilde{y}) \), we obtain

\[
\mathbb{E} \left[ |\Phi(k\tau, \omega, \tilde{y}) - \hat{S}(s, \theta_{k\tau}, \omega)|^p \right] \leq C^2\eta_r(k\tau) \left( \mathbb{E}V(\tilde{y} - \Phi(1, \theta_{-1}, \tilde{y})) \right) + \hat{\gamma}(\omega) \exp \left( \frac{1}{2} \beta k\tau \right). 
\]

Finally, by the condition (2.72) together with the \( \mathbb{P} \)-preserving property of \( \theta \), we arrive at the required estimate (2.74).

**Step III:** In this step, we show that the conditions of Lemma 2.14 are satisfied. For this purpose, we use step II and invariance of \( \tilde{\mu}_r \) w.r.t. the discrete Markov semigroup \( (\tilde{\mathcal{P}}_{r+k\tau})_{k \in \mathbb{N}} \) and for any \( h \in \text{Lip}_b(\mathcal{M}_r), \ r \in [0, \tau) \), to obtain

\[
\left| \tilde{\mathcal{P}}_{r+k\tau} h - \int_{\mathcal{M}_r} h d\tilde{\mu}_r \right| \leq \int_{S^1 \times \mathbb{R}_N} \mathbb{E}[|h(\Phi(r + k\tau, \omega, \cdot) - h(\tilde{S}(r + k\tau, \omega))|] d\tilde{\mu}_0 
\leq \text{Lip}(h) \int_{S^1 \times \mathbb{R}_N} \mathbb{E}[|\Phi(r + k\tau, \omega, \cdot) - \tilde{S}(r + k\tau, \omega)|] d\tilde{\mu} 
\leq \hat{C}\text{Lip}(h) H_r(k\tau), 
\]

i.e., there exists \( 0 < \hat{C} < \infty \) such that

\[
\left| \tilde{\mathcal{P}}_{k\tau+r} h(x) - \langle h, \tilde{\mu}_r \rangle \right| \leq \hat{C}\text{Lip}(h) \left[ \exp \left( p^{-1} \int_0^{k\tau+r} \lambda(u + r) \, du \right) + \exp \left( \frac{1}{2p} \beta k\tau \right) \right]. \tag{2.75} 
\]
Finally, let \( f \in C_b(\mathcal{M}_r) \) be given. Setting \( h = \tilde{P}_{r+\tau}f \) in (2.75), we have by Proposition 2.18 and by the invariance of \( \tilde{\mu}_t \) under the Markov semigroup \((\tilde{P}_{r+n\tau})_{k \in \mathbb{N}}\), we have
\[
\left| \tilde{P}_{r+\tau+k\tau}f(\tilde{y}) - \langle \tilde{P}_{r+\tau}f, \tilde{\mu}_r \rangle \right| = \left| \int_{\mathcal{M}_r} \left( \tilde{P}_{r+\tau+k\tau}f(\tilde{y}) - \tilde{P}_{r+k\tau+\tau}f(\tilde{y}) \right) \tilde{\mu}_r(d\tilde{y}) \right| \\
\leq C\text{Lip}(\tilde{P}_{r+\tau}f) \left[ \exp \left( p^{-1} \int_0^{r+k\tau} \lambda(u)du \right) + \exp \left( \frac{1}{2p} \beta k\tau \right) \right],
\]
where \( \tilde{K}_r = K_r\tilde{C} \). It then follows that for any \( f \in C_b(\mathcal{M}_r) \) and \( k > 1 \),
\[
\left| \tilde{P}_{r+k\tau}f(\tilde{y}) - \int_{\mathcal{M}_r} f d\tilde{\mu}_r \right| \leq \tilde{K}_r\|f\|_{\infty} \left[ \exp \left( p^{-1} \int_0^{r+k\tau-r} \lambda(u)du \right) + \exp \left( \frac{1}{2p} \beta k\tau \right) \right],
\]
this implies that there \( 0 < \tilde{K} < \infty \) such that (2.42) holds. Then, by Lemma 2.14, we have the convergence of Krylov–Bogolyubov scheme for periodic measures.

**Remark 2.20.** The approach adopted in this subsection relies on the construct periodic measure from the random periodic path \( \tilde{S}(r, \omega) \). It is also possible to construct random periodic paths given periodic measures. To see this, we proceed as in [17] and suppose that we are given the periodic measures \((\tilde{\mu}_r)_{r \in \mathbb{R}} \subset \mathcal{P}(\mathbb{S}^1 \times \mathbb{R}^N)\) of period \( \tau > 0 \), for a Markovian RDS \( \Phi \) on \( \mathbb{S}^1 \times \mathbb{R}^N \). Let the skew product flow \( \chi : \Omega \times \mathbb{S}^1 \times \mathbb{R}^N \rightarrow \Omega \times \mathbb{S}^1 \times \mathbb{R}^N \) be defined by
\[
\chi_t(\omega, \tilde{y}) = (\theta_t \omega, \Phi(t, \omega, \tilde{y})), \quad t \in \mathbb{R}. \tag{2.76}
\]
Note that the two sided skew product in (2.76) is defined by running two independent Brownian motions \( \{W^k_t : t \geq 0, 1 \leq k \leq N\} \) and \( \{W^k_t : t \geq 0, 1 \leq k \leq N\} \) (e.g., [1, 36]). Set
\[
\begin{align*}
\hat{\omega} &= (\omega, \tilde{y}), \quad \hat{\Omega} := \Omega \times \mathbb{S}^1 \times \mathbb{R}^N, \quad \hat{\mathcal{F}} := \mathcal{F} \otimes \mathcal{B}(\mathbb{S}^1 \times \mathbb{R}^N) \\
\tilde{\mu}(\hat{A}) := \frac{1}{\tau} \int_0^\tau \tilde{\mu}_r(\hat{A})dr, \quad \Psi(t, \hat{\omega}, \cdot) := \Phi(t, \omega, \cdot), \quad t \in \mathbb{R}.
\end{align*}
\]
It is relatively straightforward to check that \( \Psi \) defines an RDS over the ergodic base dynamical system \( \hat{\Omega} = (\hat{\Omega}, \hat{\mathcal{F}}, \tilde{\mu}, (\chi_t)_{t \in \mathbb{R}}) \). A random periodic path \( R : \mathbb{R}^+ \times \hat{\Omega} \rightarrow \mathbb{S}^1 \times \mathbb{R}^N \) for the enlarged RDS \( \Psi \) is constructed as follows: for any \( \hat{\omega}^* = (\omega^*, \tilde{y}^*(\theta_{-r}\omega^*)) \in \hat{\Omega}, \ r \in \mathbb{R} \)
\[
R(t, \omega^*) := \Phi(t + r, \theta_{-r}\omega^*, \tilde{y}^*(\theta_{-r}\omega^*)), \quad t \in \mathbb{R}^+ \tag{2.77}
\]
Finally, the transition probability of \( \Psi(t, \hat{\omega}, \tilde{y}) \) is the same as \( \tilde{P}(0, \tilde{y}; t, \cdot) \) and the law of \( R \) is
\[
\tilde{P} \circ R^{-1}(t, \cdot) = \tilde{\mu}_t, \quad \text{for all} \ t \in \mathbb{R}^+.
\]
Note the resemblance of the process \( R(t, \omega^*) \) and the limit \( \tilde{S}(t, \omega) \) of the Cauchy sequence in equation (2.22) from Theorem 2.8, the major difference is that the construction involve different ergodic base dynamical system. Related idea was previously employed in [48] to construct stable random periodic solutions for a Markovian RDS on a cylinder \( \mathbb{S}^1 \times \mathbb{R}^N \).
3. STOCHASTIC AVERAGING IN THE RANDOM PERIODIC REGIME

One of the difficulties in averaging for non-autonomous SDEs is to give a suitable notion of ergodic evolution system of probability measures. This issue has been resolved in subsection §2.3 for some class of SDEs with time periodic forcing. Having established the ergodicity of periodic measures for a class of systems satisfying the Assumptions 2.7 and 2.15, the proof of averaging principle will follow the same line of argument as in (e.g., [18, 28, 31, 42]) once we know the manner in which the PS-ergodic measures depend on the frozen slow variables $x \in K \subseteq \mathbb{R}^d$.

First, we deal with dependency of the PS-ergodic measures $\tilde{\mu}_x$ on the parameter $x \in K \subseteq \mathbb{R}^d$ in §3.1 and subsequently, identify the averaged equation. We prove an averaging principle in the random periodic regime in §3.2 and conclude with a toy example arising from kinetic theories of turbulent flows (e.g., [8]).

3.1. The averaged equation. A crucial step in averaging principle is to identify the averaged equation via the vector field of the slow motion and the ergodic measures of the fast subsystem. As the averaging framework considered in this paper is fully coupled, some degree of regularity of the invariant measures $x \mapsto \tilde{\mu}^x$ will be required. In fact, Lipschitz continuity with respect to the parameter $x \in K \subseteq \mathbb{R}^d$ is enough for our purpose. This Lipschitz property together with the regularity of the slow vector field $F$ would ensure the unique local solvability of the averaged equation. We employ similar argument as the one in §2.3 to prove Lipschitz dependence of periodic measures on parameters.

We recall the notion of Lipschitz continuity of probability measures in the sense of the narrow topology on $\mathcal{P}(S^1 \times \mathbb{R}^N)$. Consider a real-valued function $f : S^1 \times \mathbb{R}^N \to \mathbb{R}$ and the space of $1$-Lipschitz functions defined by

$$\text{Lip}(1, S^1 \times \mathbb{R}^N) = \{ f \in \text{Lip}_b(S^1 \times \mathbb{R}^N) : \| f \|_{BL} \leq 1 \}.$$  

The space of $1$-Lipschitz functions Lip$(1, S^1 \times \mathbb{R}^N)$ generate a narrow topology on $\mathcal{P}(S^1 \times \mathbb{R}^N)$ (e.g., [12]) with the metric

$$d_{BL}(\mu, \nu) = \sup \left\{ \int_{S^1 \times \mathbb{R}^N} f d(\mu - \nu) : f \in \text{Lip}(1, S^1 \times \mathbb{R}^N) \right\}. \quad (3.1)$$

**Definition 3.1** (Lipschitz continuity of probability measures). We say that a probability measure $(x_1, \cdots, x_d) = x \mapsto \tilde{\mu}^x \in \mathcal{P}(S^1 \times \mathbb{R}^N)$ is Lipschitz continuous in the sense of narrow topology, if there exists a constant $\tilde{C} > 0$ such that

$$d_{BL}(\tilde{\mu}^x, \tilde{\mu}^z) \leq \tilde{C}|x - z|.$$  

With the above notations in place, we have the following result on the Lipschitz dependence of PS-ergodic periodic measures on parameters.

**Proposition 3.2.** Let $K$ be a non-empty relative compact subset of $\mathbb{R}^d$, i.e., $K \subseteq \mathbb{R}^d$ and let $\{\tilde{\mu}_t^x : t \geq 0, x \in K\}$ be a family of periodic measures generated by the SDE (1.2) on $S^1 \times \mathbb{R}^N$ with coefficients satisfying Assumptions 2.7 and 2.15. Then, there exists $\tilde{C} > 0$ such that for all $x, z \in K$, we have

$$d_{BL}(\tilde{\mu}^x, \tilde{\mu}^z) \leq \tilde{C}|x - z|. \quad (3.2)$$
Proof. Recall from the definition of periodic measure that for \( f \in C_b(S^1 \times \mathbb{R}^N) \),
\[
\langle f, \bar{\mu}_i^\tau \rangle = \int_{S^1 \times \mathbb{R}^N} f d\bar{\mu}_i^\tau = \int_{S^1 \times \mathbb{R}^N} \mathbb{E}[f(\tilde{S}^x(t + k\tau, \omega))] d\mu_0, \quad k \in \mathbb{N}.
\]
In particular, for any \( f \in \text{Lip}(1, S^1 \times \mathbb{R}^N) \), we have
\[
\langle f, \bar{\mu}_i^\tau - \bar{\mu}_i \rangle = \int_{S^1 \times \mathbb{R}^N} \mathbb{E}[f(\tilde{S}^x(t + k\tau, \omega))] d\bar{\mu}_0 - \int_{S^1 \times \mathbb{R}^N} \mathbb{E}[f(\tilde{S}^x(t + k\tau, \omega))] d\bar{\mu}_0
\]
\[
\leq \int_{S^1 \times \mathbb{R}^N} \left| \mathbb{E}[f(\tilde{S}^x(t + k\tau, \omega))] - \mathbb{E}[f(\Phi^x(t + k\tau, \omega, .))] \right| d\bar{\mu}_0
\]
\[
+ \int_{S^1 \times \mathbb{R}^N} \left| \mathbb{E}[f(\tilde{S}^x(t + k\tau, \omega))] - \mathbb{E}[f(\Phi^x(t + k\tau, \omega, .))] \right| d\bar{\mu}_0
\]
\[
+ \int_{\mathbb{R}^d} \left| \mathbb{E}[f(\Phi^x(t + k\tau, \omega, .))] - \mathbb{E}[f(\Phi^x(t + k\tau, \omega, .))] \right| d\bar{\mu}_0
\]
\[
\leq \int_{S^1 \times \mathbb{R}^N} \mathbb{E}|\tilde{S}^x(t + k\tau, \omega) - \Phi^x(t + k\tau, \omega, .)| d\bar{\mu}_0
\]
\[
+ \int_{S^1 \times \mathbb{R}^N} \mathbb{E}|\tilde{S}^x(t + k\tau, \omega) - \Phi^x(t + k\tau, \omega, .)| d\bar{\mu}_0
\]
\[
+ \int_{S^1 \times \mathbb{R}^N} \mathbb{E}|\Phi^x(t + k\tau, \omega, .) - \Phi^x(t + k\tau, \omega, .)| d\bar{\mu}_0.
\]
Next, by Step I in Theorem 2.19, i.e., equation (2.73) and the stability of random periodic solution, we have for any \( t \in [0, \tau) \),
\[
\langle f, \bar{\mu}_i^\tau - \bar{\mu}_i \rangle \leq \limsup_{k \to \infty} C \exp \left( \frac{1}{2kt} \int_0^{r+k\tau} \lambda(u) du \right)^{2k\tau} \int_{S^1 \times \mathbb{R}^N} \mathbb{E}|(x, \tilde{y}) - (z, \tilde{y})| d\bar{\mu}_0(d\tilde{y}). \tag{3.3}
\]
Integrating both sides of (3.3) over \( t \in [0, \tau) \) and dividing by the length of the interval, we obtain
\[
\frac{1}{\tau} \int_0^\tau \langle f, \bar{\mu}_i^\tau - \bar{\mu}_i \rangle dt \leq \tilde{C} |x - z|. \tag{3.4}
\]
Finally, taking supremum of (3.4) over \( f \in \text{Lip}(1, \mathcal{M}) \), we arrive at the desired inequality
\[
d_{BL}(\bar{\mu}_i^\tau, \bar{\mu}^\tau) = \sup \left\{ \frac{1}{\tau} \int_0^\tau \langle f, \bar{\mu}_i^\tau - \bar{\mu}_i \rangle dt : f \in \text{Lip}(1, \mathcal{M}) \right\} \leq \tilde{C} |x - z|,
\]
\(\square\)

Now, given the ergodic property of periodic measures \( \{\mu_i^\tau : r \in [0, \tau), \ x \in \mathbb{K} \in \mathbb{R}^d \} \) and Lipschitz dependence on parameter \( x \in \mathbb{K} \), we identify the averaged vector field \( \bar{F} \) from the slow vector field \( F \). Importantly, we show that \( \bar{F} \) is regular enough to ensure local solvability of the corresponding averaged equation. The argument here is fairly straightforward, due to the regularity of the slow vector field and that of the periodic measures with respect to parameters \( x \in \mathbb{K} \). Consider a vector field \( \bar{F} : \mathbb{R}^d \to \mathbb{R}^d \) defined by
\[
\bar{F}(x) = \frac{1}{\tau} \int_0^\tau \int_{S^1 \times \mathbb{R}^N} F(x, \pi_2(\tilde{y})) \bar{\mu}_i^\tau(\tilde{y}) d\tilde{y} dr = \int_{S^1 \times \mathbb{R}^N} F(x, \pi_2(\tilde{y})) \bar{\mu}_i^\tau(\tilde{y}) \tag{3.5}
\]
where \( \pi_2 : S^1 \times \mathbb{R}^N \to \mathbb{R}^N \) is a projection map, i.e., \( \pi_2(\tilde{y}) = \pi_2(s, y) = y \).
Proposition 3.3. Let $F \in C^1_b(\mathbb{R}^d \times \mathbb{R}^N; \mathbb{R}^d)$, $l \geq 2$ and let $\{\tilde{\mu}_r^x : r \in [0, \tau], x \in \mathbb{K} \subset \mathbb{R}^d\}$ be a family of periodic measures generated by the fast SDE satisfying Assumptions 2.7 and 2.15. Then, the family of vector fields

$$\left\{ \int_{\mathcal{S}^1 \times \mathbb{R}^N} F(x, \pi_2(\bar{y})) \tilde{\mu}_r^x(d\bar{y}) : r \in [0, \tau), x \in \mathbb{K} \right\} \quad (3.6)$$

is $\tau$-periodic. Moreover, there exists $\tilde{C}_F > 0$ such that for all $x, z \in \mathbb{K}$

$$\begin{align*}
|\tilde{F}(x) - \tilde{F}(z)| &\leq \tilde{C}_F|x - z|, \\
|F(x)| &\leq \tilde{C}_F.
\end{align*} \quad (3.7)$$

Proof. The periodicity part follows from the $\tau$-periodicity of the periodic measure $\tilde{\mu}_r^x$ and the fact that $F$ is time independent. It only remains to prove the local Lipschitz and boundedness property (3.7). As $F \in C^1_b(\mathbb{R}^d \times \mathbb{R}^N; \mathbb{R}^d)$, $l \geq 2$, then by the mean value theorem and Lipschitz property of $x \mapsto \tilde{\mu}_r$, $r \in [0, \tau)$, we have the existence $L_F, \tilde{C} \geq 0$, such that for $x, z \in \mathbb{K} \subset \mathbb{R}^d$

$$|\tilde{F}(x) - \tilde{F}(z)| = \left| \frac{1}{\tau} \int_0^\tau \int_\mathcal{M} F(x, \pi_2(\bar{y})) \tilde{\mu}_r^x(d\bar{y})dr - \frac{1}{\tau} \int_0^\tau \int_\mathcal{M} F(z, \pi_2(\bar{y})) \tilde{\mu}_r^x(d\bar{y})dr \right|
\leq L_F|x - z| + L_F \frac{1}{\tau} \int_0^\tau \int_\mathcal{M} \mathbb{E}|\tilde{\Pi} \tilde{S}^x(r, \omega) - F(z, \pi_2(\bar{y}))|d\tilde{\mu}_0dr
\leq L_F|x - z| + L_F \tilde{C} |x - z| \leq \tilde{C}_F|x - z|.
$$

For the second assertion, since $F \in C^1_b(\mathbb{R}^d \times \mathbb{R}^N; \mathbb{R}^d)$, there exists $L_F > 0$ such that

$$|F(x, y)| \leq L_F(1 + |x| + |y|). \quad (3.8)$$

Now, consider the vector fields $F_n(x, y)$, $n \in \mathbb{N}$, defined by

$$F_n(x, \pi_2(\bar{y})) = \begin{cases} F(x, \pi_2(\bar{y})), & \text{if } |\bar{y}| \leq 2^n, \\ F\left(x, \frac{2^n \pi_2(\bar{y})}{\pi_2(\bar{y})}\right), & \text{if } |\bar{y}| > 2^n. \end{cases}$$

It follows that $F_n(., .) : \mathbb{R}^N \to \mathbb{R}^N$ is bounded Lipschitz continuous and for any $R > 0$, we have

$$\sup_{|x| \leq R} |F_n(x, .)| \leq \tilde{C}_F.$$ 

We use the inequality (3.8) and the moment bound on the periodic measures or random periodic solutions, i.e., equation (2.73), to obtain

$$\left| \int_{\{|\bar{y}| > 2^n\}} [F(x, \pi_2(\bar{y})) - F_n(x, \pi_2(\bar{y}))] \tilde{\mu}_r^x(d\bar{y}) \right| \leq L_F \int_{\{|\bar{y}| > 2^n\}} (1 + |x| + |\pi_2(y)|)\tilde{\mu}_r^x(d\bar{y}) \leq 2^{-n}K_F(1 + |x|^2).$$

This implies that, for any $\delta > 0$ and $R > 0$, we can choose $n = n(\delta, R) \in \mathbb{N} \setminus \{0\}$ large enough such that

$$\sup_{x \in B_R} \left| \int_{\{|\bar{y}| > 2^n\}} [F(x, \pi_2(\bar{y})) - F_n(x, \pi_2(\bar{y}))] \tilde{\mu}_r^x(d\bar{y}) \right| < \delta.$$
So, we obtain
\[ |\tilde{F}(x)| \leq \left| \int_{\mathcal{M}} F_n(x, \pi_2(\hat{y})) \tilde{F}_n(\hat{y}) \, d\hat{y} \right| + \left| \int_{\{\hat{y} : |\hat{y}| > 2^n\}} [F(x, \pi_2(\hat{y})) - F_n(x, \pi_2(\hat{y}))] \tilde{F}_n(\hat{y}) \, d\hat{y} \right| \]
\[ \leq \tilde{C}_F + \delta. \]

Finally, since \( \delta > 0 \) is arbitrary, we may take limit as \( \delta \to 0 \) to arrive at the required bound.

\( \square \)

3.2. The Averaging limit. In this subsection, we present an averaging result in the random periodic regime; we follow Hasminskii’s time discretisation scheme (e.g., [18, 28, 31]). The idea of the scheme is to decompose the time interval \([0, t]\) into subintervals of length \( \frac{1}{n(\varepsilon)} = \Delta(\varepsilon) \), such that in each subinterval the slow variable \( X^\varepsilon \) is almost everywhere constant and the fast variable \( \hat{Y}^\varepsilon_t \) after appropriate rescaling is well captured by the long time behaviour of the one-point motion \( \hat{Y}^\varepsilon_t \hat{x} \) for fixed \( x \in \mathbb{R} \subset \mathbb{R}^d \). First, we write the slow-fast SDE in the fast time scale \( t \mapsto t/\varepsilon \) as follows,
\[
\begin{align*}
\frac{dX^\varepsilon_t}{\varepsilon} &= F(X^\varepsilon_t, \Pi \hat{Y}^\varepsilon_t) \, dt \\
\frac{d\hat{Y}^\varepsilon_t}{\varepsilon} &= \frac{1}{\varepsilon} \tilde{b}(X^\varepsilon_t, \hat{Y}^\varepsilon_t) \, dt + \frac{1}{\varepsilon} \tilde{\sigma}(X^\varepsilon_t, \hat{Y}^\varepsilon_t) \, d\tilde{W}_t,
\end{align*}
\]
where \( \Pi : \mathbb{S}^1 \times \mathbb{R}^N \to \mathbb{R}^N \) is a projection map. Next, we write \([0, t]\) as a union of subintervals \([\frac{jL}{n}, \frac{(j+1)L}{n}] : j = 0, 1, \cdots, n - 1\) and set \( \hat{Y}^\varepsilon \) as an auxiliary process defined on each subinterval as follows (c.f. [18])
\[
\begin{align*}
\hat{Y}^\varepsilon_{jt/n} &= \hat{Y}^\varepsilon_{jt/n}, \\
\frac{d\hat{Y}^\varepsilon_u}{\varepsilon} &= \frac{1}{\varepsilon} \tilde{b}(X^\varepsilon_{jt/n}, \hat{Y}^\varepsilon_u) \, du + \frac{1}{\varepsilon} \tilde{\sigma}(X^\varepsilon_{jt/n}, \hat{Y}^\varepsilon_u) \, d\tilde{W}_u, \quad jt/n < u \leq (j + 1)t/n,
\end{align*}
\]
where \( \tilde{W}_t \) is the same Brownian path used in the definition of \( \hat{Y}^\varepsilon \) as a stochastic flow on \( \mathbb{S}^1 \times \mathbb{R}^N \).

Next, we give some preparatory lemmas on how to choose the integer \( n(\varepsilon) \) such that the interval spacing is not too small as PS-ergodicity need some time to occur.

Lemma 3.4 (c.f. [52]). Let \( F \in C^{l-1}_b(\mathbb{R}^d \times \mathbb{R}^N; \mathbb{R}^d), \ b, \sigma \in C^{l}_b(\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^N; \mathbb{R}^N), \ l \geq 2, \ 1 \leq k \leq N \). Then, there exists \( K > 0 \) such that
\[
\sup_{0 \leq u \leq t} \mathbb{E} |\hat{Y}^\varepsilon_u - \tilde{Y}^\varepsilon_u|^2 \leq K \left( \frac{1}{\varepsilon^2 n^2} + \frac{1}{\varepsilon n^2} \right) \exp \left( \frac{K}{\varepsilon^2 n^2} + \frac{K}{\varepsilon n} \right). \tag{3.9}
\]

Proof. Let \( u \in [0, t], \) then there exists \( j = j(u) \) such that for any \( jt/n \leq u \leq (j + 1)t/n \) we have
\[
\hat{Y}^\varepsilon_u - \tilde{Y}^\varepsilon_u \leq \frac{1}{\varepsilon} \int_{jt/n}^{u} \left[ \tilde{b}(X^\varepsilon_x, \pi_2(\tilde{y}), \hat{Y}^\varepsilon_u, \tilde{y}) - \tilde{b}(X^\varepsilon_{jt/n}, \hat{Y}^\varepsilon_u, \tilde{y}) \right] d\xi \\
+ \frac{1}{\varepsilon} \int_{jt/n}^{u} \left[ \tilde{\sigma}(X^\varepsilon_x, \pi_2(\tilde{y}), \hat{Y}^\varepsilon_u, \tilde{y}) - \tilde{\sigma}(X^\varepsilon_{jt/n}, \hat{Y}^\varepsilon_u, \tilde{y}) \right] d\tilde{W}_\xi.
\]
Next, we use Cauchy Schwartz inequality on the drift part and Itô isometry on the diffusion part to obtain
\[
\mathbb{E} \left| \hat{Y}^\varepsilon_u - \tilde{Y}^\varepsilon_u \right|^2 \leq \frac{1}{\varepsilon^2 n^2} \int_{jt/n}^{u} \mathbb{E} \left| \tilde{b}(X^\varepsilon_x, \pi_2(\tilde{y}), \hat{Y}^\varepsilon_u, \tilde{y}) - \tilde{b}(X^\varepsilon_{jt/n}, \hat{Y}^\varepsilon_u, \tilde{y}) \right|^2 d\xi \\
+ \frac{1}{\varepsilon} \int_{jt/n}^{u} \mathbb{E} \left| \tilde{\sigma}(X^\varepsilon_x, \pi_2(\tilde{y}), \hat{Y}^\varepsilon_u, \tilde{y}) - \tilde{\sigma}(X^\varepsilon_{jt/n}, \hat{Y}^\varepsilon_u, \tilde{y}) \right|^2 d\xi.
\]
By the regularity of \( b, \sigma \), then by the mean value theorem, there exist \( L_b, L_\sigma > 0 \) such that
\[
\mathbb{E} \left| \hat{Y}^{\varepsilon,x,y}_{u} - \hat{Y}^{\varepsilon,x,y}_{0} \right|^2 \leq \frac{L_b}{\varepsilon^2 n} \int_{jt/n}^{u} \left\{ \mathbb{E} \left| X^{\varepsilon,x,\pi_2(y)}_{\xi} - X^{\varepsilon,x,\pi_2(y)}_{jt/n} \right|^2 + \mathbb{E} \left| \hat{Y}^{\varepsilon,x,y}_{\xi} - \hat{Y}^{\varepsilon,x,y}_{jt/n} \right|^2 \right\} d\xi
\]
\[
+ \frac{L_\sigma}{\varepsilon} \int_{jt/n}^{u} \left\{ \mathbb{E} \left| X^{\varepsilon,x,\pi_2(y)}_{\xi} - X^{\varepsilon,x,\pi_2(y)}_{jt/n} \right|^2 + \mathbb{E} \left| \hat{Y}^{\varepsilon,x,y}_{\xi} - \hat{Y}^{\varepsilon,x,y}_{jt/n} \right|^2 \right\} d\xi.
\]
Next, by the regularity of \( F \), there exist \( L_F > 0 \) such that
\[
\mathbb{E} \left| X^{\varepsilon,x,\pi_2(y)}_{\xi} - X^{\varepsilon,x,\pi_2(y)}_{jt/n} \right|^2 \leq L_F |\xi - jt/n|,
\]
so that
\[
\mathbb{E} \left| \hat{Y}^{\varepsilon,x,y}_{u} - \hat{Y}^{\varepsilon,x,y}_{0} \right|^2 \leq L_F \left( \frac{L_b}{\varepsilon^2 n} + \frac{L_\sigma}{\varepsilon} \right) \int_{jt/n}^{u} |\xi - jt/n| d\xi
\]
\[
+ \left( \frac{L_b}{\varepsilon^2 n} + \frac{L_\sigma}{\varepsilon} \right) \int_{jt/n}^{u} \mathbb{E} \left| \hat{Y}^{\varepsilon,x,y}_{\xi} - \hat{Y}^{\varepsilon,x,y}_{jt/n} \right|^2 d\xi
\]
\[
\leq K \left( \frac{1}{\varepsilon^2 n^2} + \frac{1}{\varepsilon n^2} \right) + K \left( \frac{1}{\varepsilon^2 n} + \frac{1}{\varepsilon} \right) \int_{jt/n}^{u} \mathbb{E} \left| \hat{Y}^{\varepsilon,x,y}_{\xi} - \hat{Y}^{\varepsilon,x,y}_{jt/n} \right|^2 d\xi.
\]
Finally, by Gronwall lemma, we have the desired bound
\[
\sup_{0 \leq u \leq t} \mathbb{E} \left| \hat{Y}^{\varepsilon,x,y}_{u} - \hat{Y}^{\varepsilon,x,y}_{0} \right|^2 \leq K \left( \frac{1}{\varepsilon^2 n^2} + \frac{1}{\varepsilon n^2} \right) \exp \left( \frac{K}{\varepsilon^2 n^2} + \frac{K}{\varepsilon n} \right).
\]

We choose \( n(\varepsilon) = \frac{1}{\varepsilon \ln(\varepsilon^{-1})^l} \), thanks to the above Lemma 3.4 (see, also [18, 31]). The next lemma deals with the interval spacing. We take \( t_0 > 0 \) and consider the solution \( \hat{Y}^{\varepsilon,x,y}_{u} \) of the equation
\[
\begin{cases}
\frac{d\hat{Y}^{\varepsilon,x,y}_{u}}{du} = \frac{1}{\sqrt{\varepsilon}} \hat{b}(x, \hat{Y}^{\varepsilon,x,y}_{u}) du + \frac{1}{\sqrt{\varepsilon}} \sigma(x, \hat{Y}^{\varepsilon,x,y}_{u}) d\hat{W}_u, & u > t_0 \\
\hat{Y}^{\varepsilon,x,y}_{t_0} = \tilde{y} \in \mathbb{S}^1 \times \mathbb{R}^N.
\end{cases}
\]

**Lemma 3.5.** Let \( F \in C^{l-1}_{b}(\mathbb{R} \times \mathbb{R}^N; \mathbb{R}^d) \), \( b, \sigma_k \in C^l_b(\mathbb{R} \times \mathbb{R} \times \mathbb{R}^N; \mathbb{R}^d) \), \( l \geq 2 \), \( 1 \leq k \leq N \) and Assumptions 2.7 and 2.15 hold. Then, there exist constants \( C > 0 \) such for any \( T > 0 \), \( x \in \mathbb{R}^d \), \( \tilde{y} \in \mathbb{S}^1 \times \mathbb{R}^N \), we have
\[
\mathbb{E} \left| \frac{1}{T} \int_{t_0}^{t_0+T} F(x, \Pi^\varepsilon_u \hat{Y}^{\varepsilon,x,y}_{u}) du - \hat{F}(x) \right|^2 \leq C \left( 1 + |x|^2 + |	ilde{y}|^2 \right)^2 + \alpha(T, x)
\]
for some mapping \( \alpha : \mathbb{R}^+ \times \mathbb{R}^d \rightarrow \mathbb{R}^+ \) with
\[
\sup_{T \in \mathbb{R}^+} \alpha(T, x) \leq C(1 + |x|^2)
\]
and, for any \( K \in \mathbb{R}^d \)
\[
\lim_{T \to \infty} \sup_{x \in K} \alpha(T, x) = 0.
\]
Proof. We proceed as in the proof of Lemma 7.2 in [9] or as in the proof of Lemma 3.2 in [52], we have
\[
\mathbb{E}\left( \frac{1}{T} \int_{t_0}^{t_0+T} F(x, \Pi \tilde{Y}_{u,x}^\varepsilon) du - \bar{F}(x) \right)^2
\]
\[
= \frac{1}{T^2} \int_{t_0}^{t_0+T} \left( \mathbb{E} \left[ \left( F(x, \Pi \tilde{Y}_{u,x}^\varepsilon) - \bar{F}(x) \right)^2 \right] du \right)
\]
\[
= \frac{2}{T^2} \int_{t_0}^{t_0+T} \mathbb{E} \left[ \left( F(x, \Pi \tilde{Y}_{u,x}^\varepsilon) - \bar{F}(x) \right)^2 \right] du dr.
\]
Next, by Markov property of \( \Pi \tilde{Y}_{t,x}^\varepsilon \), we have for \( r \leq u, \)
\[
\mathbb{E} \left[ \left( F(x, \Pi \tilde{Y}_{r,x}^\varepsilon) - \bar{F}(x) \right)^2 \right] = \mathbb{E} \left[ \left( F(x, \Pi \tilde{Y}_{r,x}^\varepsilon) - \bar{F}(x) \right)^2 \right].
\]
By Cauchy-Schwartz inequality, we have
\[
\mathbb{E} \left( \frac{1}{T} \int_{t_0}^{t_0+T} F(x, \Pi \tilde{Y}_{u,x}^\varepsilon) du - \bar{F}(x) \right)^2
\]
\[
\leq \frac{2}{T^2} \int_{t_0}^{t_0+T} \left( \mathbb{E} \left( \left( F(x, \Pi \tilde{Y}_{r,x}^\varepsilon) - \bar{F}(x) \right)^2 \right) \right)^{1/2} dr.
\]
Next, we estimate the integrand using the regularity of vector fields \( F, \bar{F} \) and PS-ergodicity of the fast motion (see §2.3), these yield the existence of \( C > 0 \) such that
\[
\mathbb{E} |F(x, \Pi \tilde{Y}_{r,x}^\varepsilon) - \bar{F}(x)|^2 \leq C \left( 1 + |x|^2 + \mathbb{E} |\tilde{Y}_{r,x}^\varepsilon|^2 \right)
\]
\[
\leq C \left( 1 + |x|^2 + \exp \left( \int_{t_0}^{t_0+T} \lambda(\xi) d\xi \right) |\tilde{y}|^2 \right).
\]
Also, as \( F \in C_{l}^{l-1}(\mathbb{R}^d \times \mathbb{R}^N; \mathbb{R}^d), l \geq 2, \) there exists \( L_F > 0 \) such that
\[
|F(x,y) - F(x,z)| \leq L_F |y - z|.
\]
The inequality (3.12) together with PS-ergodicity of the fast variables yield
\[
\mathbb{E} |\tilde{P}_{u-r} F(x, \Pi \tilde{Y}_{u,x}^\varepsilon) - \bar{F}(x)|^2 \leq C \left( 1 + |x|^2 + |\tilde{y}|^2 \right) \exp \left( \int_{t_0}^{u-r} \lambda(\xi) d\xi \right).
\]
We substitute the above estimates (3.11) and (3.13) into the integral (3.10) to obtain
\[
\mathbb{E} \left( \frac{1}{T} \int_{t_0}^{t_0+T} F(x, \Pi \tilde{Y}_{u,x}^\varepsilon) du - \bar{F}(x) \right)^2
\]
\[
\leq \frac{2C}{T^2} \int_{t_0}^{t_0+T} \int_{t_0}^{t_0+T} (1 + |x|^2 + |\tilde{y}|^2) \exp \left( \int_{t_0}^{T} \lambda(\xi) d\xi \right) du dr
\]
\[
\leq \frac{2C}{T} (1 + |x|^2 + |\tilde{y}|^2)^2 \frac{1}{T^2} \int_{t_0}^{t_0+T} \int_{t_0}^{t_0+T} \exp \left( \int_{t_0}^{u-r} \lambda(\xi) d\xi \right) du dr
\]
\[
\leq \frac{C}{T} (1 + |x|^2 + |\tilde{y}|^2)^2.
\]
Finally, from Proposition 3.3, we see that the family of vector fields
\[
\left\{ \int_{S^1 \times \mathbb{R}^N} F(x, \pi_2(\tilde{y})) \mu^x_t(\tilde{y}) : r \in [0, \tau), \ x \in \mathbb{K} \right\}
\]
is \(\tau\)-periodic, then by PS-ergodicity of the fast motion, we know that
\[
\lim_{k \to \infty} \frac{1}{kT} \int_{t_0}^{t_0 + kT} F(x, \pi_2(\tilde{y})) \mu^x_t(\tilde{y}) dt = \tilde{F}(x).
\]
In view of this, we may define the mapping \(\alpha : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^+\) by
\[
\alpha(T, x) = 2 \left| \frac{1}{T} \int_{t_0}^{t_0 + T} \int_{\mathcal{M}} F(x, \pi_2(\tilde{y})) \mu^x_t(\tilde{y}) dt - \tilde{F}(x) \right|^2.
\]

With the above preparatory lemmas and notations in place, we are ready to state and prove a stochastic averaging principle in the random periodic regime.

**Theorem 3.6.** Assume that \(F \in \mathcal{C}^{l-1}(\mathbb{R}^d \times \mathbb{R}^N ; \mathbb{R}^d)\) and \(b, \sigma_k \in \mathcal{C}^{l}(\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^N ; \mathbb{R}^N), \ 1 \leq k \leq N, \ l \geq 2.\) If for fixed \(x \in \mathbb{K} \subset \mathbb{R}^d,\) the coefficients \(b, \sigma_k, \ 1 \leq k \leq N\) are time periodic and satisfy Assumptions 2.7 and 2.15. Moreover, if \(\{\mathcal{M}_r : r \in [0, \tau)\} = : \mathcal{M} \subset S^1 \times \mathbb{R}^N\) is such that each \(\mathcal{M}_r\) is a \(k, \tau\)-irreducible Poincaré section, then
\[
\lim_{\varepsilon \to 0} \int_{\mathbb{K}} \int_{\mathcal{M}} E \left( \sup_{0 \leq t \leq T/\varepsilon} \left| X_{t, \varepsilon, \pi_2(\tilde{y})} - X_{t, \varepsilon, x} \right| \right) \tilde{\mu}^x(\tilde{y}) \nu(dx) = 0,
\]
where \(\nu \in \mathcal{P}(\mathbb{K})\) and \(\tilde{\mu}^x \in \mathcal{P}(\mathcal{M}).\)

**Proof.**

We start by using the Lipschitz property of the averaged vector field \(\tilde{F}\) to obtain
\[
\left| X_{t, \varepsilon, \pi_2(\tilde{y})} - X_{t, \varepsilon, x} \right| = \varepsilon \left| \int_{0}^{t} \left[ F \left( X_{u, \varepsilon, \pi_2(\tilde{y})} , \Pi Y_{u, \varepsilon, \pi_2(\tilde{y})} \right) - \tilde{F}(X_{u, \varepsilon, x}) \right] du \right| \\
\leq \varepsilon \left| \int_{0}^{t} \left[ F \left( X_{u, \varepsilon, \pi_2(\tilde{y})} , \Pi Y_{u, \varepsilon, \pi_2(\tilde{y})} \right) - \tilde{F}(X_{u, \varepsilon, x}) \right] du \right| + \varepsilon \tilde{C}_F \int_{0}^{t} \left| X_{u, \varepsilon, \pi_2(\tilde{y})} - X_{u, \varepsilon, x} \right| du
\]
and by Gronwall’s lemma, we have
\[
\sup_{0 \leq t \leq T/\varepsilon} \left| X_{t, \varepsilon, \pi_2(\tilde{y})} - X_{t, \varepsilon, x} \right| \leq \varepsilon e^{T \tilde{C}_F} \sup_{0 \leq t \leq T/\varepsilon} \left| \int_{0}^{t} \left[ F \left( X_{u, \varepsilon, \pi_2(\tilde{y})} , \Pi Y_{u, \varepsilon, \pi_2(\tilde{y})} \right) - \tilde{F}(X_{u, \varepsilon, \pi_2(\tilde{y})}) \right] du \right|.
\]

Next, from Lemmas 3.4 and 3.5, we can choose \(n(\varepsilon) = \frac{1}{\varepsilon \sqrt{\ln(\varepsilon^{-1})}}\) and set \(t(\varepsilon) = \frac{T}{\varepsilon n(\varepsilon)}\) (see, also [18, 31]) and write the integral in RHS of the inequality (3.15) in the form
\[
\varepsilon \left| \int_{0}^{t(\varepsilon)} \left[ F \left( X_{j, \varepsilon, \pi_2(\tilde{y})} , \Pi Y_{j, \varepsilon, \pi_2(\tilde{y})} \right) - \tilde{F}(X_{j, \varepsilon, \pi_2(\tilde{y})}) \right] du \right| \\
= \varepsilon \left| \int_{0}^{t(\varepsilon)} \left\{ \left[ F \left( X_{j+u, \varepsilon, \pi_2(\tilde{y})} , \Pi Y_{j+u, \varepsilon, \pi_2(\tilde{y})} \right) - F \left( X_{j+u, \varepsilon, \pi_2(\tilde{y})} , \Pi Y_{j+u, \varepsilon, \pi_2(\tilde{y})} \right) \right] + \tilde{F}(X_{j(t)\varepsilon}) - \tilde{F}(X_{j(t)\varepsilon}) \right\} du \right| \\
=: \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3.
Now, by the regularity of the vector field $F$, we have the existence of $L_F > 0$, such that

$$I_1 = \varepsilon \left| \int_0^{t(\varepsilon)} \left[ F \left( X_{j(t)}^\varepsilon, x, \pi_2(y) \right) - F \left( X_{j(t)}^\varepsilon, x, \pi_2(y) \right) \right] du \right| \leq \varepsilon L_F \int_0^{t(\varepsilon)} \left| X_{j(t)}^\varepsilon - X_{j(t)}^\varepsilon \right| \leq \tilde{L}_F (\varepsilon t(\varepsilon))^2. \quad (3.16)$$

In a similar fashion, by the regularity of $F$ and that of averaged vector field $\tilde{F}$, we obtain

$$I_3 = \varepsilon \left| \int_0^{t(\varepsilon)} \left[ \tilde{F} \left( X_{j(t)}^\varepsilon, x, \pi_2(y) \right) - \tilde{F} \left( X_{j(t)}^\varepsilon, x, \pi_2(y) \right) \right] du \right| \leq [\tilde{L}_F + \tilde{C}_F L_F] (\varepsilon t(\varepsilon))^2. \quad (3.17)$$

Given $n(\varepsilon)$ with $t(\varepsilon) = \frac{T}{\varepsilon n(\varepsilon)}$, we use (3.16) and (3.17) to obtain the following inequality (see also [31]),

$$e^n \mathbb{E} \left( \sup_{0 \leq t \leq T/\varepsilon} \left| \int_0^t F \left( X_{j(t)}^\varepsilon, x, \pi_2(y) \right), \Pi_{\tilde{y}}^{\varepsilon, x, \tilde{y}} \right| du \right) \leq 2L_F e^n t(\varepsilon) + e^n \mathbb{E} \left( \sum_{j=0}^{n(\varepsilon)-1} \left| \int_0^{t(\varepsilon)} \left[ F \left( X_{j(t)}^\varepsilon, x, \pi_2(y) \right), \Pi_{\tilde{y}}^{\varepsilon, x, \tilde{y}} \right] du \right| \right)$$

$$= 2L_F e^n t(\varepsilon) + e^n \mathbb{E} \left( \sum_{j=0}^{n(\varepsilon)-1} \left| \int_0^{t(\varepsilon)} \left[ F \left( X_{j(t)}^\varepsilon, x, \pi_2(y) \right), \Pi_{\tilde{y}}^{\varepsilon, x, \tilde{y}} \right] du \right| \right)$$

$$\leq 2L_F e^n t(\varepsilon) + e^n n(\varepsilon) t(\varepsilon) [\tilde{L}_F + \tilde{C}_F L_F]$$

$$+ e^n t(\varepsilon) \left( \sum_{j=0}^{n(\varepsilon)-1} \mathbb{E} \left| \frac{1}{t(\varepsilon)} \int_0^{t(\varepsilon)} F \left( X_{j(t)}^\varepsilon, x, \pi_2(y) \right), \Pi_{\tilde{y}}^{\varepsilon, x, \tilde{y}} \right| du \right)$$

$$\leq 2L_F e^n t(\varepsilon) + e^n n(\varepsilon) t(\varepsilon) [\tilde{L}_F + \tilde{C}_F L_F]$$

$$+ e^n t(\varepsilon) \left( \sum_{j=0}^{n(\varepsilon)-1} \mathbb{E} \left| \frac{1}{t(\varepsilon)} \int_0^{t(\varepsilon)} F \left( X_{j(t)}^\varepsilon, x, \pi_2(y) \right), \Pi_{\tilde{y}}^{\varepsilon, x, \tilde{y}} \right| du \right)$$

Recalling from the discretisation scheme that the slow motion $X_{j(t)}^\varepsilon, y = 0, \ldots n(\varepsilon) - 1$, is $\mathbb{P}(d\omega)\tilde{\mu}^x(dy)\nu(dx)$ almost everywhere constant on each sub-interval of length $t(\varepsilon)$, then by Lemma 3.5, we have

$$\lim_{\varepsilon \to 0} \max_{0 \leq j \leq n(\varepsilon)-1} \int_{\mathcal{X}} \mathbb{E} |A_{x}(j, x, y)|^2 \tilde{\mu}^x(dy)\nu(dx) = 0 \quad (3.19)$$

where

$$2\mathbb{E} |A_{x}(j, x, y)|^2 = 2\mathbb{E} \left| \frac{1}{t(\varepsilon)} \int_0^{t(\varepsilon)} F \left( X_{j(t)}^\varepsilon, x, \pi_2(y) \right), \Pi_{\tilde{y}}^{\varepsilon, x, \tilde{y}} \right| du \right| F \left( X_{j(t)}^\varepsilon, x, \pi_2(y) \right) \right|^2$$

$$= \mathbb{E} \left[ \alpha \left( t(\varepsilon), X_{j(t)}^\varepsilon \right) \right]. \quad (3.20)$$
Integrating both sides of (3.15) with respect to \( \mathbb{P}(d\omega)\tilde{\mu}^x(dy)\nu(dx) \) and taking note of (3.18) and (3.20), we have
\[
\int_\mathcal{K} \int_\mathcal{M} \mathbb{E} \left( \sup_{0 \leq t \leq T/\varepsilon} \left| X^x_{t\varepsilon}(\tilde{y}) - X^x_{t\varepsilon}(\tilde{y}) \right| \right) \tilde{\mu}^x(dy)\nu(dx) \leq \varepsilon T \mathbb{E} \left[ \frac{2L_F T}{n(\varepsilon)} + \varepsilon T [\tilde{L}_F + \tilde{C}_F L_F] \right] + T \max_{0 \leq j \leq n(\varepsilon)-1} \int_\mathcal{K} \int_\mathcal{M} \left( \mathbb{E} |\Lambda_c(j, x, \tilde{y})|^2 \right)^{1/2} \tilde{\mu}^x(dy)\nu(dx) \tilde{\mu}^x(dy)\nu(dx).
\]
Taking limit as \( \varepsilon \to 0 \), and recalling that \( n(\varepsilon) \to \infty \) as \( \varepsilon \downarrow 0 \) and in view of (3.19), the proof is complete.

**Example 3.7.** We consider the following time periodic forcing SDE in two dimensions with \( d = 1 \) and \( N = 1 \).

\[
\begin{aligned}
    dX &= \varepsilon \left[ Y^2 + \alpha X + \vartheta X^3 \right] dt \\
    dY &= \left[ -\gamma(X)Y + \beta \cos(2\pi t) \right] dt + \sigma dW_t
\end{aligned}
\]

where \( 0 < \varepsilon \ll 1, \alpha, \vartheta \in \mathbb{R}, \beta \neq 0, \sigma \neq 0 \) and \( \gamma : \mathbb{R} \to \mathbb{R}^+ \setminus \{0\} \) is a polynomial.

**Proof.** Here, the solution of fast subsystem \( Y^{x,y}_t \) can be written explicitly as
\[
    Y^{x,y}_t = e^{-\gamma(x)t} y + \beta \int_0^t e^{-\gamma(x)(t-s)} \cos(2\pi s) ds + \sigma \int_0^t e^{-\gamma(x)(t-s)} dW_s,
\]
and the random periodic solution \( S^x(t, \omega) \) of period \( \tau = 1 \) is given by
\[
    S^x(t, \omega) = \beta \int_{-\infty}^t e^{-\gamma(x)(t-s)} \cos(2\pi s) ds + \sigma \int_{-\infty}^t e^{-\gamma(x)(t-s)} dW_s.
\]
The lifted random periodic solution \( \tilde{S}^x(t, \omega) \) on the cylinder \([0, 1) \times \mathbb{R}\) is
\[
    \tilde{S}^x(t, \omega) = (t \mod 1, S^x(t, \omega)).
\]
The corresponding periodic measure \( \tilde{\mu}^x_t \) is given by
\[
    \tilde{\mu}^x_t(dy) = \frac{1}{\sqrt{2\pi C(x)}} \exp \left( -\frac{(y - \beta v(t, x))^2}{2C(x)} \right) I_{[t \mod 1]}(s) dy ds
\]
where
\[
v(t, x) = \int_0^t e^{-\gamma(x)(t-s)} \cos(2\pi s) ds
\]
and \( C(x) \) solves the Lyapunov equation \( 2\gamma(x)C(x) = \sigma^2 \).

Given the periodic measure \( \tilde{\mu}^x_t \), we calculate the averaged vector field \( \bar{F}(x) \) as follows
\[
\bar{F}(x) = \int_0^1 \int_{[0, 1) \times \mathbb{R}} y^2 \tilde{\mu}^x_t(dy) ds dt + \alpha x + \vartheta x^3
\]
\[
= C(x) + \int_0^1 \beta^2 v^2(t, x) dt + \alpha x + \vartheta x^3.
\]

We compute the process \( v(t, x) \) as
\[
v(t, x) = \frac{\gamma(x) \cos 2\pi t + 2\pi \sin 2\pi t}{4\pi^2 + \gamma^2(x)}
\]
and the integral of \( v^2(t, x) \) over \([0, 1)\) is
\[
\int_0^1 v^2(t, x) dt = \int_0^1 \left( \frac{\gamma(x) \cos 2\pi t + 2\pi \sin 2\pi t}{\gamma^2(x) + 4\pi^2} \right)^2 dt = \frac{2}{\gamma^2(x) + 4\pi^2}.
\]
Finally, the rescaled slow motion $X^{x,y}_{t/\varepsilon}$ as $\varepsilon \downarrow 0$, can be approximated by the ODE
\[
\frac{d\tilde{X}_t}{dt} = \frac{\sigma^2}{2\gamma(X_t)} + \frac{2\beta^2}{\gamma^2(X_t) + 4\pi^2} + \alpha \tilde{X}_t + \vartheta \tilde{X}_t^3.
\] (3.22)

4. Conclusions and future work

We proved a stochastic averaging principle for fully coupled SDEs with time periodic forcing. The ergodicity of evolution system of periodic probability measures on an irreducible Poincaré section is a key step to the identification and validity of the averaging limit. This key step is natural in applications like climate change studies and neural networks, for example, only the statistics and not a particular path of the weather (fast variables) are important in climate evolution (e.g., [3]).

Averaging is one of the important steps for climatological applications, but it does not give the whole applicable recipes required in climate change studies. Some moderate and large deviation results for the slow motion from averaged motion are required for sensitive and rare events analysis. It is possible to use our framework to develop large deviation results and fluctuation dissipation theorems (FDTs) (e.g., [7, 37, 38]), particularly important in climate and neuronal models governed by multi-scale SDEs with time periodic forcing, where usual detailed balance condition (symmetry of stationary measures) fails. We shall investigate this aspect in the future publications and hope to carry out Hasselmann’s program (e.g., [3, 5, 32]), i.e., approximation of slow subsystem by a nonlinear SDE where random fluctuations are taken into account, providing description for hopping of the slow variables (climate) between local basins of attractors/metastable states (extreme climate events).

References

[1] L. Arnold, Random Dynamical Systems, Springer, 1998.
[2] L. Arnold and M. Scheutzow, Perfect cocycles through stochastic differential equations, Probability Theory and Related Fields Vol. 101 (1995), 65-88.
[3] L. Arnold, Hasselmann’s program revisited: the analysis of stochasticity in deterministic climate models, Stochastic Climate Models: Progress in Probability, Vol. 49 (2001) , 141 – 157, Birkhäuser, Basel.
[4] M. Arnaudon and A. Thalmaier, The differentiation of hypoelliptic diffusion semigroups, Illinois Journal of Mathematics, 54 (2010), 1285 – 1311.
[5] V. Bakhtin and Y. Kifer, Diffusion approximation for slow motion in fully coupled averaging, Probability Theory and Related Fields, Vol. 129 (2004), 157 – 181.
[6] N. N. Bogolyubov, Y. A. Mitropol’skii, Asymptotic methods in theory of nonlinear oscillations, Gordon and Breach, Delhi, 1961.
[7] M. Branicki and A. J. Majda, Quantifying uncertainty for predictions with model error in non-Gaussian systems with intermittency, Nonlinearity, Vol. 25 (2012) 2543 – 2578.
[8] F. Bouchet, T. Grafke, T. Tangarife and E. Vanden-Eijnden, Large deviations in fast-slow systems, Journal of Statistical Physics, Vol. 162 (2016), 793 – 812.
[9] S. Cerrai and A. Lunardi, Averaging principle for nonautonomous slow-fast systems of stochastic reaction-diffusion equations: The almost periodic case, SIAM Journal of Mathematical Analysis, Vol. 49 (2017), 2843 – 2884.
[10] X. Chen, J. Duan and M. Scheutzow, Evolution systems of measures for stochastic flows, Dynamical Systems, Vol. 26 (2011), 323 – 334.
[11] M. D. Chekroun, E. Simonnet and M. Ghil, Stochastic climate dynamics: Random attractors and time dependent invariant measures, Physica D, Vol. 240 (2011), 1685 – 1700.
[12] H. Crauel, Random Probability Measures on Polish spaces, Taylor and Francis, 2002.
[13] G. Da Prato, Introduction to Stochastic Analysis and Malliavin Calculus, Edizioni della Normale, 2008.
[14] K. D. Elworthy, *Stochastic dynamical systems and their flows*. In: Stochastic Analysis, ed. A. Friedman, M. Pinsky, London-New York press, (1978) 79 - 95.

[15] C. Feng, H. Zhao and B. Zhou, *Pathwise random periodic solutions of stochastic differential equations*, Journal of Differential Equations, 251(2011) 119 - 149.

[16] C. Feng and H. Zhao, *Random Periodic Solutions of SPDEs via Integral Equations and Wiener-Sobolev Compact Embedding*, Journal of Functional Analysis, Vol. 262 (2012) 4377 - 4422.

[17] C. Feng and H. Zhao, *Random periodic processes, periodic measures and ergodicity*, available at arXiv:1408.1897v4 [math.PR]

[18] M. I. Freidlin and A. D. Wentzell, *Random perturbations of Dynamical Systems*, Springer, 1998.

[19] M. Galtier and G. Wainrib, *Multiscale analysis of slow-fast neuronal learning models with noise*, The Journal of Mathematical Neuroscience, Vol. 2 (2012), 1 – 64.

[20] M. Hairer and A. J. Majda, *A simple framework to justify linear response theory*, Nonlinearity 23 (2010), 909 - 922.

[21] M. Hairer, J. C. Mattingly and M. Scheutzow, *Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equation*, Probab. Theory Relat. Fields, 149 (2011), 223 - 259.

[22] M. Hairer and J. C. Mattingly, *A theory of hypoellipticity and unique ergodicity for semilinear stochastic PDEs*, Electronic Journal of Probability, Vol. 16 (2011), 658 - 738.

[23] M. Hairer, *Advanced stochastic analysis*, Lecture Note Univ. Warwick. Available: http://www.hairer.org/Course.pdf.

[24] M. Hairer, *Convergence of Markov processes*, Lecture Note Univ. Warwick. Available: http://www.hairer.org/notes/Convergence.pdf.

[25] M. Hairer, *On Malliavin's proof of Hörmander's theorem*, Bull. Sci. Math. 135 (2011), 650 - 666.

[26] M. Hairer and N. S. Pillai, *Ergodicity of hypoelliptic SDEs driven by fractional Brownian motion*, Annales de l’Institut Henri Poincaré - Prob. Stat. Vol. 47 (2011), 601 - 625.

[27] R. Z. Hasminskii, *Stochastic stability of differential equations*, Sijthoff and Noordhoff, 1980 (Translated from Russia).

[28] R. Z. Hasminskii, *On the averaging principle for stochastic differential Itô equations*, Kibernetika 4, (1968), 260 - 279.

[29] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North Holland - Kodansha, Tokyo, 1981.

[30] Y. Kifer, *Averaging and climate models*, Stochastic Climate Models: Progress in Probability, Vol. 49 (2001), 171 – 188, Birkhäuser, Basel.

[31] Y. Kifer, *Large deviations and adiabatic transitions for dynamical systems and Markov processes in fully coupled averaging*, American Mathematical Society, 2009.

[32] Y. Kifer, *Averaging principle for fully coupled dynamical systems and large deviations*, Ergodic Theory and Dynamical Systems, Vol. 24 (2004), 847 – 871.

[33] Y. Kifer, *L² diffusion approximation for slow motion in averaging*, Stochastics and Dynamics, Vol. 3 (2003), 213 – 246.

[34] Y. Kifer, *Limit theorems in averaging for dynamical systems*, Ergodic Theory and Dynamical Systems, Vol. 15 (1995), 1143 – 1172.

[35] H. Kunita, *Stochastic differential equations and stochastic flow of diffeomorphisms*, École d’été de Probabilités de Saint-Flour 12, 1982 Lecture Notes in Mathematics, Vol. 1097 (1984), 143 - 303.

[36] H. Kunita, *Stochastic flows and stochastic differential equations*, Cambridge University press, Cambridge, 1990.

[37] A. J. Majda and X. Wang, *Linear response theory for statistical ensembles in complex systems with time-periodic forcing*, Communication of Mathematical Sciences, Vol. 8 (2010), 145 – 172.

[38] A. J. Majda, *Introduction to Turbulent Dynamical Systems in Complex Systems*, Springer, 2016.

[39] P. Malliavin, *Stochastic analysis*, Grundlehren der Mathematischen Wissenschaften 313, Springer, 1997.

[40] X. Mao, *Exponential stability of stochastic differential equations*, Marcel Dekker, New York, 1994.

[41] A. Millet, D. Nualart and M. Sanz, *Integration by parts and time reversal for diffusion processes*, The Annals of Probability, Vol. 17 (1989), 208 - 238.

[42] P. Lochak and C. Meunier, *Multiphase Averaging for Classical Systems with Application to Adiabatic Theorems*, Springer, 1988.

[43] D. Nualart, *The Malliavin calculus and related topics*, Springer, 2006.

[44] G. A. Pavliotis and A. M. Stuart, *Averaging and Homogenization*, Springer, 2008.

[45] B. Schmalfuss, *Lyapunov functions and non-trivial stationary solutions of stochastic differential equations*, Dynamical Systems, Vol. 16 No 8 (2001) 303 - 317.

[46] R. Sturman and J. Stark, *Semi-uniform ergodic theorems and applications to forced systems*, Nonlinearity, Vol. 13 (2000) 113 – 143.
[47] K. Uda, A qualitative approach to the existence of random periodic solutions, Loughborough University Doctoral Thesis, 2015.
[48] K. Uda, Existence of random invariant periodic curves via random semiuniform ergodic theorem, Stochastics and Dynamics, Vol. 17 (2017), 1750007.
[49] K. Uda and H. Zhao, Random periodic solutions and ergodicity for stochastic differential equations. Submitted, 2018. Available at arXiv:1811.05503 [math.PR].
[50] A. Y. Veretennikov, On large deviations in the averaging principle for SDEs with full dependence, The Annals of Probability, Vol. 27 (1999), 284 – 296.
[51] A. Y. Veretennikov, On large deviations in the averaging principle for stochastic differential equations with periodic coefficients II, Math. USSR Izvestiya, Vol. 39 (1992), 677.
[52] G. Wainrib, Double averaging for periodically forced slow-fast systems, Electronic Communications in Probability, Vol. 18 (2013), 1 –12.
[53] B. Wang, Periodic and almost periodic random inertial manifolds for non-autonomous stochastic equations, page 189 - 208, Continuous and Distributed Systems II: Theory and Applications, Edited by V. A. Sadovnichiy and M. Z. Zgurovsky, Springer, 2015.
[54] B. Wang, Stochastic bifurcation of pathwise almost random periodic and almost automorphic solutions for random dynamical systems, Discrete and Continuous Dynamical Systems, Vol. 36, issue 1, (2015), 3745 - 3769.
[55] B. Wang, Existence, stability and bifurcation of random complete and periodic solutions of stochastic parabolic equations, Nonlinear Analysis: Theory, Methods and Applications, Vol. 103 (2014), 9 - 25.
[56] S. Watanabe, Lectures on stochastic differential equations and Malliavin calculus, Springer, 1984.
[57] H. Zhao and Z. Zheng, Random Periodic Solutions of Random Dynamical systems, Journal of differential equations, Vol. 246, No: 5 (2009), 2020 - 2038.