ELECTRONIC RESEARCH ARCHIVE
Volume **29**, Number **4**, September **2021**

doi:10.3934/era.2021014
pp. **2791–2817**

AVERAGING PRINCIPLE ON INFINITE INTERVALS FOR STOCHASTIC ORDINARY DIFFERENTIAL EQUATIONS

David Cheban

School of Mathematical Sciences, Dalian University of Technology
Dalian 116024, China
State University of Moldova, Faculty of Mathematics and Informatics
Department of Mathematics
A. Mateevich Street 60, MD–2009 Chişinău, Moldova

Zhenxin Liu*

School of Mathematical Sciences, Dalian University of Technology
Dalian 116024, China

Dedicated to the memory of Professor Russell A. Johnson

Abstract. In contrast to existing works on stochastic averaging on finite intervals, we establish an averaging principle on the whole real axis, i.e. the so-called second Bogolyubov theorem, for semilinear stochastic ordinary differential equations in Hilbert space with Poisson stable (in particular, periodic, quasi-periodic, almost periodic, almost automorphic etc) coefficients. Under some appropriate conditions we prove that there exists a unique recurrent solution to the original equation, which possesses the same recurrence property as the coefficients, in a small neighborhood of the stationary solution to the averaged equation, and this recurrent solution converges to the stationary solution of averaged equation uniformly on the whole real axis when the time scale approaches zero.

1. **Introduction.** Highly oscillating systems may be "averaged" under some suitable conditions, and the evolution of the averaged system can reflect in some sense the dynamics of the original system. This idea of averaging dates back to the perturbation theory developed by Clairaut, Laplace and Lagrange in the 18th century, and is made rigorous by Krylov, Bogolyubov, Mitropolsky [1, 2, 18] for nonlinear oscillations. There are vast amount of works on averaging for deterministic systems which we will not mention here. Meantime, there are also many works on averaging principle for stochastic differential equations so far, see e.g. [4, 5, 12, 14, 15, 29, 30, 31, 32] among others. But to our best knowledge, except for the centre manifold approach to averaging (see e.g. [3, 22]), almost all the existing works on stochastic averaging are concerned with the so-called first Bogolyubov theorem, i.e. the convergence

---

2020 *Mathematics Subject Classification.* 34C29, 60H10, 37B20, 34C27.

Key words and phrases. Averaging principle, stochastic differential equations, second Bogolyubov theorem, periodic solution, quasi-periodic solution, almost periodic solution, Poisson stable solution.

This work is partially supported by NSFC Grants 11522104, 11871132, 11925102, and Xinghai Jieqing and DUT19TD14 funds from Dalian University of Technology.

* Corresponding author: Zhenxin Liu.

of the solution of the original equation to that of the averaged equation on finite intervals.

In the present paper, we establish an averaging principle on the whole real axis, i.e. the so-called second Bogolyubov theorem, for stochastic differential equations: if there exists a stationary solution for the averaged equation, then there exists in a small neighborhood (in the super-norm topology) a solution of the original equation which is defined on the whole axis and has the same recurrence property (in distribution sense) as the coefficients of the original equation. Furthermore, this recurrent solution is more general than the classical second Bogolyubov theorem, which only treats the almost periodic case. Note that the work [16] studies the averaging principle for stochastic differential equations with almost periodic coefficients, but they only show the convergence on the finite interval, not the super-norm topology on the whole axis.

To be more precise, we investigate the semilinear stochastic ordinary differential equation with Poisson stable (in particular, periodic, quasi-periodic, Bohr almost periodic, almost automorphic, Birkhoff recurrent, Levitan almost periodic, almost recurrent, pseudo periodic, pseudo recurrent) in time coefficients. Under some suitable conditions, this equation has a unique L^2-bounded solution which has the same recurrent properties as the coefficients, see [9, 20] for details. In this paper, we show that this recurrent solution converges to the unique stationary solution of the averaged equation uniformly on the whole real axis when the time scale goes to zero.

The paper is organized as follows. In the second section we collect some known notions and facts. Namely we present the construction of shift dynamical systems, definitions and basic properties of Poisson stable functions, Shcherbakov's comparability method, and the existence of compatible solutions for stochastic differential equations. In the third and fourth sections, we investigate the averaging principle on infinite intervals for linear and semilinear stochastic differential equations respectively.

2. **Preliminaries.**

2.1. **Shift dynamical systems.**

Let $(\mathcal{X},\rho)$ be a complete metric space and $(\mathcal{X},\mathbb{R},\pi)$ be a *dynamical system (or flow)* on $\mathcal{X}$, i.e. the mapping $\pi:\mathbb{R}\times\mathcal{X}\to\mathcal{X}$ is continuous, $\pi(0,x)=x$ and $\pi(t+s,x)=\pi(t,\pi(s,x))$ for any $x\in\mathcal{X}$ and $t,s\in\mathbb{R}$. The mapping $t\mapsto\pi(t,x)$ is called the *motion through* x. Denote by $C(\mathbb{R},\mathcal{X})$ the space of all continuous functions $\varphi:\mathbb{R}\to\mathcal{X}$ equipped with the distance

$$d(\varphi_1,\varphi_2):=\sum_{k=1}^{\infty}\frac{1}{2^k}\frac{d_k(\varphi_1,\varphi_2)}{1+d_k(\varphi_1,\varphi_2)},$$

where

$$d_k(\varphi_1,\varphi_2):=\sup_{|t|\le k}\rho(\varphi_1(t),\varphi_2(t)),$$

which generates the compact-open topology on $C(\mathbb{R},\mathcal{X})$. The space $(C(\mathbb{R},\mathcal{X}),d)$ is a complete metric space (see, e.g. [25, 26, 27, 28]).

Remark 1. (i) Let $\varphi,\varphi_n\in C(\mathbb{R},\mathcal{X})$ $(n\in\mathbb{N})$. Then $\lim\limits_{n\to\infty}d(\varphi_n,\varphi)=0$ if and only if $\lim\limits_{n\to\infty}\max\limits_{|t|\le l}\rho(\varphi_n(t),\varphi(t))=0$ for any $l>0$.

(ii) If there exists a sequence $l_n \to +\infty$ such that $\lim\limits_{n\to\infty} \max\limits_{|t|\le l_n} \rho(\varphi_n(t), \varphi(t)) = 0$, then $\lim\limits_{n\to\infty} d(\varphi_n, \varphi) = 0$ and vice versa. See [28] for details.

Let us now introduce two examples of shift dynamical systems which we will use later in this paper.

Example 2.1. For given $\varphi \in C(\mathbb{R}, \mathcal{X})$, we denote by φ^τ the *τ-translation of* φ, i.e. $\varphi^\tau(t) = \varphi(\tau + t)$ for $t \in \mathbb{R}$. Let $\sigma : \mathbb{R} \times C(\mathbb{R}, \mathcal{X}) \to C(\mathbb{R}, \mathcal{X})$ be a mapping defined by equality $\sigma(\tau, \varphi) := \varphi^\tau$ for $(\tau, \varphi) \in \mathbb{R} \times C(\mathbb{R}, \mathcal{X})$. Clearly $\sigma(0, \varphi) = \varphi$ and $\sigma(\tau_1 + \tau_2, \varphi) = \sigma(\tau_2, \sigma(\tau_1, \varphi))$ for $\varphi \in C(\mathbb{R}, \mathcal{X})$ and $\tau_1, \tau_2 \in \mathbb{R}$. It is immediate to check (see, e.g. [7, 25, 26, 28]) that the mapping $\sigma : \mathbb{R} \times C(\mathbb{R}, \mathcal{X}) \to C(\mathbb{R}, \mathcal{X})$ is continuous, and consequently the triplet $(C(\mathbb{R}, \mathcal{X}), \mathbb{R}, \sigma)$ is a dynamical system which is called *shift dynamical system* or *Bebutov dynamical system*.

The *hull of* φ, denoted by $H(\varphi)$, is the set of all the limits of φ^{τ_n} in $C(\mathbb{R}, \mathcal{X})$, i.e.

$$H(\varphi) := \{\psi \in C(\mathbb{R}, \mathcal{X}) : \psi = \lim_{n\to\infty} \varphi^{\tau_n} \text{ for some sequence } \{\tau_n\} \subset \mathbb{R}\}.$$

Note that the set $H(\varphi)$ is a closed and translation invariant subset of $C(\mathbb{R}, \mathcal{X})$ and consequently it naturally defines on $H(\varphi)$ a shift dynamical system – $(H(\varphi), \mathbb{R}, \sigma)$.

Example 2.2. Like in [9], we denote by $BUC(\mathbb{R} \times \mathcal{X}, \mathcal{X})$ the space of all continuous functions $f : \mathbb{R} \times \mathcal{X} \to \mathcal{X}$ which are bounded on every bounded subset from $\mathbb{R} \times \mathcal{X}$ and continuous in $t \in \mathbb{R}$ uniformly with respect to x on each bounded subset Q of $\mathcal{X}$. We equip this space with the topology of uniform convergence on bounded subsets of $\mathbb{R} \times \mathcal{X}$, which can be generated by the following metric

$$d(f, g) := \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{d_k(f, g)}{1 + d_k(f, g)}, \tag{1}$$

where

$$d_k(f, g) := \sup_{|t|\le k,\ x\in Q_k} \rho(f(t, x), g(t, x))$$

with $Q_k \subset \mathcal{X}$ being bounded, $Q_k \subset Q_{k+1}$ and $\cup_{k\in\mathbb{N}} Q_k = \mathcal{X}$.

For given $f \in BUC(\mathbb{R} \times \mathcal{X}, \mathcal{X})$ and $\tau \in \mathbb{R}$, we denote by f^τ the *τ-translation of* f, i.e. $f^\tau(t, x) := f(t+\tau, x)$ for $(t, x) \in \mathbb{R} \times \mathcal{X}$. Note that the space $BUC(\mathbb{R} \times \mathcal{X}, \mathcal{X})$ endowed with the distance (1) is a complete metric space and invariant with respect to translations. Now we define a mapping $\sigma : \mathbb{R} \times BUC(\mathbb{R} \times \mathcal{X}, \mathcal{X}) \to BUC(\mathbb{R} \times \mathcal{X}, \mathcal{X})$, $(\tau, f) \mapsto f^\tau$. It is clear that $\sigma(0, f) = f$ and $\sigma(\tau_2, \sigma(\tau_1, f)) = \sigma(\tau_1 + \tau_2, f)$ for all $f \in BUC(\mathbb{R} \times \mathcal{X}, \mathcal{X})$ and $\tau_1, \tau_2 \in \mathbb{R}$. It is immediate to see (e.g. [7, ChI]) that the mapping σ is continuous and consequently the triplet $(BUC(\mathbb{R} \times \mathcal{X}, \mathcal{X}), \mathbb{R}, \sigma)$ is a dynamical system. Similar to Example 2.1, for given $f \in BUC(\mathbb{R} \times \mathcal{X}, \mathcal{X})$, the hull $H(f)$ is a closed and translation invariant subset of $BUC(\mathbb{R} \times \mathcal{X}, \mathcal{X})$ and consequently it naturally defines on $H(f)$ a shift dynamical system – $(H(f), \mathbb{R}, \sigma)$.

Denote by $BC(\mathcal{X}, \mathcal{X})$ the space of all continuous functions $f : \mathcal{X} \to \mathcal{X}$ which are bounded on every bounded subset of $\mathcal{X}$ and equipped with the distance

$$d(f, g) := \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{d_k(f, g)}{1 + d_k(f, g)},\ d_k(f, g) := \sup_{x\in Q_k} \rho(f(x), g(x))$$

where Q_k are the same as above. Note that $(BC(\mathcal{X}, \mathcal{X}), d)$ is a complete metric space. For given $F \in BUC(\mathbb{R} \times \mathcal{X}, \mathcal{X})$, define $\mathcal{F} : \mathbb{R} \to BC(\mathcal{X}, \mathcal{X}), t \mapsto \mathcal{F}(t)$ by letting $\mathcal{F}(t) := F(t, \cdot) : \mathcal{X} \to \mathcal{X}$. Clearly, $\mathcal{F} \in C(\mathbb{R}, BC(\mathcal{X}, \mathcal{X}))$.

Remark 2. (i) Define the mapping $h : BUC(\mathbb{R}\times\mathcal{X},\mathcal{X}) \to C(\mathbb{R}, BC(\mathcal{X},\mathcal{X}))$ by equality $h(F) := \mathcal{F}$. It is immediate to see that the mapping h is one-one and continuous, and its inverse is also continuous. So it establishes an isometry between $BUC(\mathbb{R}\times\mathcal{X},\mathcal{X})$ and $C(\mathbb{R}, BC(\mathcal{X},\mathcal{X}))$.

(ii) By the definition of h we have $h(F^\tau) = \mathcal{F}^\tau$ for any $\tau \in \mathbb{R}$ and $F \in BUC(\mathbb{R}\times\mathcal{X},\mathcal{X})$, i.e. the shift dynamical systems $(BUC(\mathbb{R}\times\mathcal{X},\mathcal{X}),\mathbb{R},\sigma)$ and $(C(\mathbb{R}, BC(\mathcal{X},\mathcal{X})),\mathbb{R},\sigma)$ are (dynamically) homeomorphic.

2.2. **Poisson stable functions.** Let us recall the types of Poisson stable functions to be used in this paper; we refer the reader to [25, 26, 27, 28] for further details and the relations among these types of functions.

Definition 2.3. A function $\varphi \in C(\mathbb{R},\mathcal{X})$ is called *stationary* (respectively, τ-*periodic*) if $\varphi(t) = \varphi(0)$ (respectively, $\varphi(t+\tau) = \varphi(t)$) for all $t \in \mathbb{R}$.

Definition 2.4. (i) Let $\varepsilon > 0$. A number $\tau \in \mathbb{R}$ is called ε-*almost period* of the function $\varphi : \mathbb{R} \to \mathcal{X}$ if $\rho(\varphi(t+\tau),\varphi(t)) < \varepsilon$ for all $t \in \mathbb{R}$. Denote by $\mathcal{T}(\varphi,\varepsilon)$ the set of ε-almost periods of φ.

(ii) A function $\varphi \in C(\mathbb{R},\mathcal{X})$ is said to be *Bohr almost periodic* if the set of ε-almost periods of φ is *relatively dense* for each $\varepsilon > 0$, i.e. for each $\varepsilon > 0$ there exists a constant $l = l(\varepsilon) > 0$ such that $\mathcal{T}(\varphi,\varepsilon) \cap [a, a+l] \neq \emptyset$ for all $a \in \mathbb{R}$.

(iii) A function $\varphi \in C(\mathbb{R},\mathcal{X})$ is said to be *pseudo-periodic in the positive* (respectively, *negative*) *direction* if for each $\varepsilon > 0$ and $l > 0$ there exists a ε-almost period $\tau > l$ (respectively, $\tau < -l$) of the function φ. The function φ is called pseudo-periodic if it is *pseudo-periodic* in both directions.

Remark 3. A function $\varphi \in C(\mathbb{R},\mathcal{X})$ is pseudo-periodic in the positive (respectively, negative) direction if and only if there is a sequence $t_n \to +\infty$ (respectively, $t_n \to -\infty$) such that φ^{t_n} converges to φ uniformly with respect to $t \in \mathbb{R}$ as $n \to \infty$.

Definition 2.5. (i) A number $\tau \in \mathbb{R}$ is said to be ε-*shift* of $\varphi \in C(\mathbb{R},\mathcal{X})$ if $d(\varphi^\tau,\varphi) < \varepsilon$; denote $\mathfrak{T}(\varphi,\varepsilon) := \{\tau : d(\varphi^\tau,\varphi) < \varepsilon\}$. A function $\varphi \in C(\mathbb{R},\mathcal{X})$ is called *almost recurrent (in the sense of Bebutov)* if for every $\varepsilon > 0$ the set $\mathfrak{T}(\varphi,\varepsilon)$ is relatively dense.

(ii) A function $\varphi \in C(\mathbb{R},\mathcal{X})$ is called *Lagrange stable* if $\{\varphi^\tau : \tau \in \mathbb{R}\}$ is a relatively compact subset of $C(\mathbb{R},\mathcal{X})$.

(iii) A function $\varphi \in C(\mathbb{R},\mathcal{X})$ is called *Birkhoff recurrent* if it is almost recurrent and Lagrange stable.

Definition 2.6. A function $\varphi \in C(\mathbb{R},\mathcal{X})$ is called *Poisson stable in the positive* (respectively, *negative*) *direction* if for every $\varepsilon > 0$ and $l > 0$ there exists $\tau > l$ (respectively, $\tau < -l$) such that $d(\varphi^\tau,\varphi) < \varepsilon$. The function φ is called *Poisson stable* if it is Poisson stable in both directions.

In what follows, we denote as well $\mathcal{Y}$ a complete metric space.

Definition 2.7. (i) A function $\varphi \in C(\mathbb{R},\mathcal{X})$ is called *Levitan almost periodic* if there exists a Bohr almost periodic function $\psi \in C(\mathbb{R},\mathcal{Y})$ such that for any $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that $\mathcal{T}(\psi,\delta) \subseteq \mathfrak{T}(\varphi,\varepsilon)$.

(ii) A function $\varphi \in C(\mathbb{R},\mathcal{X})$ is said to be *almost automorphic* if it is Levitan almost periodic and Lagrange stable.

Remark 4. Note that:

1. every Bohr almost periodic function is Levitan almost periodic;

2. the function $\varphi \in C(\mathbb{R},\mathbb{R})$ defined by equality
$$\varphi(t) = \frac{1}{2+\cos t + \cos\sqrt{2}t}$$
is Levitan almost periodic, but it is not Bohr almost periodic [19, ChIV].

Definition 2.8. A function $\varphi \in C(\mathbb{R},\mathcal{X})$ is called *quasi-periodic with the spectrum of frequencies* $\nu_1,\nu_2,\dots,\nu_k$ if the following conditions are fulfilled:

1. the numbers $\nu_1,\nu_2,\dots,\nu_k$ are rationally independent;
2. there exists a continuous function $\Phi : \mathbb{R}^k \to \mathcal{X}$ such that $\Phi(t_1+2\pi, t_2+2\pi,\dots,t_k+2\pi) = \Phi(t_1,t_2,\dots,t_k)$ for all $(t_1,t_2,\dots,t_k) \in \mathbb{R}^k$;
3. $\varphi(t) = \Phi(\nu_1 t, \nu_2 t,\dots,\nu_k t)$ for $t \in \mathbb{R}$.

Let $\varphi \in C(\mathbb{R},\mathcal{X})$. Denote by $\mathfrak{N}_\varphi$ (respectively, $\mathfrak{M}_\varphi$) the family of all sequences $\{t_n\} \subset \mathbb{R}$ such that $\varphi^{t_n} \to \varphi$ (respectively, $\{\varphi^{t_n}\}$ converges) in $C(\mathbb{R},\mathcal{X})$ as $n \to \infty$. We denote by $\mathfrak{N}^u_\varphi$ (respectively, $\mathfrak{M}^u_\varphi$) the family of sequences $\{t_n\} \in \mathfrak{N}_\varphi$ (respectively, $\{t_n\} \in \mathfrak{M}_\varphi$) such that φ^{t_n} converges to φ (respectively, φ^{t_n} converges) uniformly with respect to $t \in \mathbb{R}$ as $n \to \infty$.

Remark 5. (i) The function $\varphi \in C(\mathbb{R},\mathcal{X})$ is pseudo-periodic in the positive (respectively, negative) direction if and only if there is a sequence $\{t_n\} \in \mathfrak{N}^u_\varphi$ such that $t_n \to +\infty$ (respectively, $t_n \to -\infty$) as $n \to \infty$.

(ii) Let $\varphi \in C(\mathbb{R},\mathcal{X})$, $\psi \in C(\mathbb{R},\mathcal{Y})$ and $\mathfrak{N}^u_\psi \subseteq \mathfrak{N}^u_\varphi$. If the function ψ is pseudo-periodic in the positive (respectively, negative) direction, then so is φ.

Definition 2.9. ([23, 26, 27]) A function $\varphi \in C(\mathbb{R},\mathcal{X})$ is called *pseudo-recurrent* if for any $\varepsilon > 0$ and $l \in \mathbb{R}$ there exists $L \ge l$ such that for any $\tau_0 \in \mathbb{R}$ we can find a number $\tau \in [l, L]$ satisfying
$$\sup_{|t|\le 1/\varepsilon} \rho(\varphi(t+\tau_0+\tau), \varphi(t+\tau_0)) \le \varepsilon.$$

Remark 6. ([23, 26, 27, 28])

1. Every Birkhoff recurrent function is pseudo-recurrent, but the inverse statement is not true in general.
2. If the function $\varphi \in C(\mathbb{R},\mathcal{X})$ is pseudo-recurrent, then every function $\psi \in H(\varphi)$ is pseudo-recurrent.
3. If the function $\varphi \in C(\mathbb{R},\mathcal{X})$ is Lagrange stable and every function $\psi \in H(\varphi)$ is Poisson stable, then φ is pseudo-recurrent.

Finally, we remark that a Lagrange stable function is not Poisson stable in general, but all other types of functions introduced above are Poisson stable.

Definition 2.10. A function $F \in BUC(\mathbb{R}\times\mathcal{X},\mathcal{X})$ is said to possess the property A in $t \in \mathbb{R}$ uniformly with respect to x on every bounded subset Q of $\mathcal{X}$ if the motion $\sigma(\cdot, F)$ through F with respect to the Bebutov dynamical system $(BUC(\mathbb{R}\times\mathcal{X},\mathcal{X}),\mathbb{R},\sigma)$ possesses the property A. Here the property A may be stationary, periodic, Bohr/Levitan almost periodic etc.

Remark 7. Note that a function $\varphi \in C(\mathbb{R},\mathcal{X})$ possesses the property A if and only if the motion $\sigma(\cdot,\varphi) : \mathbb{R} \to C(\mathbb{R},\mathcal{X})$ through φ with respect to the Bebutov dynamical system $(C(\mathbb{R},\mathcal{X}),\mathbb{R},\sigma)$ possesses this property.

2.3. **Shcherbakov's comparability method by character of recurrence.**

Definition 2.11. A function $\varphi \in C(\mathbb{R}, \mathcal{X})$ is said to be *comparable* (respectively, *strongly comparable*) *by character of recurrence* with $\psi \in C(\mathbb{R}, \mathcal{Y})$ if $\mathfrak{N}_\psi \subseteq \mathfrak{N}_\varphi$ (respectively, $\mathfrak{M}_\psi \subseteq \mathfrak{M}_\varphi$).

Theorem 2.12. ([26, ChII], [24]) *The following statements hold.*

1. $\mathfrak{M}_\psi \subseteq \mathfrak{M}_\varphi$ *implies* $\mathfrak{N}_\psi \subseteq \mathfrak{N}_\varphi$*, and hence strong comparability implies comparability.*
2. *Let* $\varphi \in C(\mathbb{R}, \mathcal{X})$ *be comparable by character of recurrence with* $\psi \in C(\mathbb{R}, \mathcal{Y})$*. If the function* ψ *is stationary (respectively,* τ*-periodic, Levitan almost periodic, almost recurrent, Poisson stable), then so is* φ*.*
3. *Let* $\varphi \in C(\mathbb{R}, \mathcal{X})$ *be strongly comparable by character of recurrence with* $\psi \in C(\mathbb{R}, \mathcal{Y})$*. If the function* ψ *is quasi-periodic with the spectrum of frequencies* $\nu_1, \nu_2, \dots, \nu_k$ *(respectively, Bohr almost periodic, almost automorphic, Birkhoff recurrent, Lagrange stable), then so is* φ*.*
4. *Let* $\varphi \in C(\mathbb{R}, \mathcal{X})$ *be strongly comparable by character of recurrence with* $\psi \in C(\mathbb{R}, \mathcal{Y})$ *and* ψ *be Lagrange stable. If* ψ *is pseudo-periodic (respectively, pseudo-recurrent), then so is* φ*.*

Lemma 2.13. ([9]) *Let* $\varphi \in C(\mathbb{R}, \mathcal{X})$*,* $\psi \in C(\mathbb{R}, \mathcal{Y})$*. The following statements hold.*

1. *If* $\mathfrak{M}^u_\psi \subseteq \mathfrak{M}^u_\varphi$*, then*
   (a) $\mathfrak{N}^u_\psi \subseteq \mathfrak{N}^u_\varphi$*;*
   (b) if the function ψ *is Bohr almost periodic, then so is* φ*.*
2. *If* $\mathfrak{N}^u_\psi \subseteq \mathfrak{N}^u_\varphi$ *and* ψ *is pseudo periodic, then so is* φ*.*

2.4. **Compatible solutions of semilinear stochastic differential equations.** Let $\mathfrak{B}$ be a real separable Banach space with the norm $|\cdot|$, and $L(\mathfrak{B})$ be the Banach space of all bounded linear operators acting on the space $\mathfrak{B}$ equipped with operator norm $\|\cdot\|$. Consider the linear homogeneous equation

$$\dot{x} = \mathcal{A}(t)x \tag{2}$$

on the space $\mathfrak{B}$, where $\mathcal{A} \in C(\mathbb{R}, L(\mathfrak{B}))$. Denote by $U(t, \mathcal{A})$ the Cauchy operator (see, e.g. [10]) of equation (2).

Definition 2.14. Equation (2) is said to be *uniformly asymptotically stable* if there are positive constants $\mathcal{N}$ and ν such that

$$\|G_{\mathcal{A}}(t, \tau)\| \le \mathcal{N} e^{-\nu(t-\tau)} \text{ for any } t \ge \tau \ (t, \tau \in \mathbb{R}), \tag{3}$$

where $G_{\mathcal{A}}(t, \tau) := U(t, \mathcal{A})U^{-1}(\tau, \mathcal{A})$ for any $t, \tau \in \mathbb{R}$.

If $\mathcal{A} \in C(\mathbb{R}, L(\mathfrak{B}))$, then by $H(\mathcal{A})$ we denote the closure in the space $C(\mathbb{R}, L(\mathfrak{B}))$ of all translations $\{\mathcal{A}^h : h \in \mathbb{R}\}$, where $\mathcal{A}^h(t) := \mathcal{A}(t + h)$ for $t \in \mathbb{R}$. Denote by $C_b(\mathbb{R}, \mathfrak{B})$ the Banach space of all continuous and bounded mappings $\varphi : \mathbb{R} \to \mathfrak{B}$ equipped with the norm $\|\varphi\|_\infty := \sup\{|\varphi(t)| : t \in \mathbb{R}\}$. Note that if $f \in C_b(\mathbb{R}, \mathfrak{B})$ and $\tilde{f} \in H(f)$, then $\|\tilde{f}\|_\infty \le \|f\|_\infty$.

Lemma 2.15. [6, ChIII] *Suppose that equation (2) is uniformly asymptotically stable such that inequality (3) holds. Then*

$$\|G_{\tilde{\mathcal{A}}}(t, \tau)\| \le \mathcal{N} e^{-\nu(t-\tau)}$$

for any $t \ge \tau$ $(t, \tau \in \mathbb{R})$ *and* $\tilde{\mathcal{A}} \in H(\mathcal{A})$*.*

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and $L^2(\mathbb{P}, \mathfrak{B})$ be the space of $\mathfrak{B}$-valued random variables X such that

$$\mathbb{E}|X|^2 := \int_\Omega |X|^2 \mathrm{d}\mathbb{P} < \infty.$$

Then $L^2(\mathbb{P}, \mathfrak{B})$ is a Banach space equipped with the norm $\|X\|_2 := \left(\int_\Omega |X|^2 \mathrm{d}\mathbb{P}\right)^{1/2}$.

Let $\mathcal{P}(\mathfrak{B})$ be the space of all Borel probability measures on $\mathfrak{B}$ endowed with the β metric:

$$\beta(\mu, \nu) := \sup\left\{\left|\int f \mathrm{d}\mu - \int f \mathrm{d}\nu\right| : \|f\|_{BL} \le 1\right\}, \quad \text{for } \mu, \nu \in \mathcal{P}(\mathfrak{B}),$$

where f are bounded Lipschitz continuous real-valued functions on $\mathfrak{B}$ with the norm

$$\|f\|_{BL} := Lip(f) + \|f\|_\infty, \text{ with } Lip(f) := \sup_{x \ne y} \frac{|f(x) - f(y)|}{|x - y|}, \ \|f\|_\infty := \sup_{x \in \mathfrak{B}} |f(x)|.$$

Recall that a sequence $\{\mu_n\} \subset \mathcal{P}(\mathfrak{B})$ is said to *weakly converges* to μ if $\int f \mathrm{d}\mu_n \to \int f \mathrm{d}\mu$ for all $f \in C_b(\mathfrak{B})$, where $C_b(\mathfrak{B})$ is the space of all bounded continuous real-valued functions on $\mathfrak{B}$. It is well-known (see, e.g. [13, ChXI]) that $(\mathcal{P}(\mathfrak{B}), \beta)$ is a separable complete metric space and that a sequence $\{\mu_n\}$ weakly converges to μ if and only if $\beta(\mu_n, \mu) \to 0$ as $n \to \infty$.

Definition 2.16. A sequence of random variables $\{X_n\}$ is said to *converge in distribution* to the random variable X if the corresponding laws $\{\mu_n\}$ of $\{X_n\}$ weakly converge to the law μ of X, i.e. $\beta(\mu_n, \mu) \to 0$.

In the following, we assume that $\mathcal{H}$ is a real separable Hilbert space. We still denote the norm in $\mathcal{H}$ by $|\cdot|$ and the operator norm in $L(\mathcal{H})$ by $\|\cdot\|$ which will not cause confusion. Let us consider the stochastic differential equation

$$\mathrm{d}X(t) = (\mathcal{A}(t)X(t) + F(t, X(t)))\mathrm{d}t + G(t, X(t))\mathrm{d}W(t), \tag{4}$$

where $\mathcal{A} \in C(\mathbb{R}, L(\mathcal{H}))$ and $F, G \in C(\mathbb{R} \times \mathcal{H}, \mathcal{H})$. Here W is a two-sided standard one-dimensional Brownian motion defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$; that is,

$$W(t) = \begin{cases} W_1(t), & \text{for } t \ge 0, \\ -W_2(-t), & \text{for } t < 0, \end{cases}$$

where W_1 and W_2 are two independent one-sided standard one-dimensional Brownian motions. We set $\mathcal{F}_t := \sigma\{W(u) : u \le t\}$.

Definition 2.17. An $\mathcal{F}_t$-adapted process $\{X(t)\}_{t \in \mathbb{R}}$ is said to be a mild solution of equation (4) on $\mathbb{R}$ if it satisfies the following stochastic integral equation

$$X(t) = G_\mathcal{A}(t, s)X(s) + \int_s^t G_\mathcal{A}(t, \tau)F(\tau, X(\tau))\mathrm{d}\tau + \int_s^t G_\mathcal{A}(t, \tau)G(\tau, X(\tau))\mathrm{d}W(\tau)$$

for all $t \ge s$ and each $s \in \mathbb{R}$, recalling that $G_\mathcal{A}$ is defined in Definition 2.14.

Definition 2.18. We say that functions F and G satisfy the condition

(C1) if there exists a constant $M \ge 0$ such that $|F(t, 0)| \vee |G(t, 0)| \le M$ for $t \in \mathbb{R}$;

(C2) if there exists a constant $L \ge 0$ such that $Lip(F) \vee Lip(G) \le L$, where $Lip(F) := \sup_{t \in \mathbb{R}, x \ne y} \frac{|F(t,x) - F(t,y)|}{|x-y|}$ and similarly for $Lip(G)$;

(C3) if F and G are continuous in t uniformly with respect to x on each bounded subset $Q \subset \mathcal{H}$.

Remark 8. (i) If F and G satisfy the conditions (C1)–(C3), then $F, G \in BUC(\mathbb{R} \times \mathcal{H}, \mathcal{H})$.

(ii) If F and G satisfy (C1)–(C2) with the constants M and L, then every pair of functions $(\tilde{F}, \tilde{G})$ in $H(F,G) := \overline{\{(F^\tau, G^\tau) : \tau \in \mathbb{R}\}}$, the hull of (F,G), also possess the same property with the same constants.

Definition 2.19. Let $\{\varphi(t)\}_{t\in\mathbb{R}}$ be a mild solution of equation (4). Then φ is called *compatible* (respectively, *strongly compatible*) *in distribution* if $\mathfrak{N}_{(\mathcal{A},F,G)} \subseteq \tilde{\mathfrak{N}}_\varphi$ (respectively, $\mathfrak{M}_{(\mathcal{A},F,G)} \subseteq \tilde{\mathfrak{M}}_\varphi$), where $\tilde{\mathfrak{N}}_\varphi$ (respectively, $\tilde{\mathfrak{M}}_\varphi$) means the set of all sequences $\{t_n\} \subset \mathbb{R}$ such that the sequence $\{\varphi(\cdot + t_n)\}$ converges to $\varphi(\cdot)$ (respectively, $\{\varphi(\cdot + t_n)\}$ converges) in distribution uniformly on any compact interval.

Theorem 2.20. *Consider the equation* (4). *Suppose that the following conditions hold:*

(a) $\sup_{t\in\mathbb{R}} \|\mathcal{A}(t)\| < +\infty$;
(b) equation (2) is uniformly asymptotically stable such that (3) *holds;*
(c) the functions F and G satisfy the conditions (C1) and (C2).

Then the following statements hold:

1. *if* $L < \frac{\nu}{\mathcal{N}\sqrt{2+\nu}}$, *then equation* (4) *has a unique solution* $\xi \in C(\mathbb{R}, B[0,r])$ *which satisfies*
$$\xi(t) = \int_{-\infty}^t G_{\mathcal{A}}(t,\tau)F(\tau,\xi(\tau))\mathrm{d}\tau + \int_{-\infty}^t G_{\mathcal{A}}(t,\tau)F(\tau,\xi(\tau))\mathrm{d}W(\tau), \tag{5}$$
where
$$r = \frac{\mathcal{N}M\sqrt{2+\nu}}{\nu - \mathcal{N}L\sqrt{2+\nu}} \tag{6}$$
and
$$B[0,r] := \{x \in L^2(\mathbb{P},\mathcal{H}) : \ \|x\|_2 \le r\};$$
2. *if additionally F and G satisfy (C3) and* $L < \frac{\nu}{2\mathcal{N}\sqrt{1+\nu}}$, *then*
   (a) $\mathfrak{M}^u_{(\mathcal{A},F,G)} \subseteq \tilde{\mathfrak{M}}^u_\xi$, *where* $\tilde{\mathfrak{M}}^u_\xi$ *means the the set of all sequences* $\{t_n\} \subset \mathbb{R}$ *such that the sequence* $\{\xi(\cdot + t_n)\}$ *converges in distribution uniformly on* $\mathbb{R}$;
   (b) the solution ξ is strongly compatible in distribution.

Proof. The proof is analogous to Theorem 4.6 in [9]. □

Corollary 1. *Under the conditions of* Theorem 2.20 *the following statements hold.*

1. *If the functions $\mathcal{A} \in C(\mathbb{R}, L(\mathcal{H}))$ and $F, G \in C(\mathbb{R} \times \mathcal{H}, \mathcal{H})$ are jointly stationary (respectively, τ-periodic, quasi-periodic with the spectrum of frequencies $\nu_1, \dots, \nu_k$, Bohr almost periodic, almost automorphic, Birkhoff recurrent, Lagrange stable, Levitan almost periodic, almost recurrent, Poisson stable), then equation (4) has a unique solution $\varphi \in C_b(\mathbb{R}, L^2(\mathbb{P},\mathcal{H}))$ which is stationary (respectively, τ-periodic, quasi-periodic with the spectrum of frequencies $\nu_1, \dots, \nu_k$, Bohr almost periodic, almost automorphic, Birkhoff recurrent, Lagrange stable, Levitan almost periodic, almost recurrent, Poisson stable) in distribution.*
2. *If the functions $\mathcal{A} \in C(\mathbb{R}, L(\mathcal{H}))$ and $F, G \in C(\mathbb{R} \times \mathcal{H}, \mathcal{H})$ are Lagrange stable and jointly pseudo-periodic (respectively, pseudo-recurrent), then equation (4) has a unique solution $\varphi \in C_b(\mathbb{R}, L^2(\mathbb{P},\mathcal{H}))$ which is pseudo-periodic (respectively, pseudo-recurrent) in distribution.*

Proof. These statements follow from Theorems 2.12 and 2.20 (see also Remark 2). □

3. **Averaging for linear equations.** Let ε_0 be some fixed positive number. Consider the equation

$$\mathrm{d}X(t) = \varepsilon(\mathcal{A}(t)X(t) + f(t))\mathrm{d}t + \sqrt{\varepsilon}g(t)\mathrm{d}W(t), \tag{7}$$

where $\mathcal{A} \in C(\mathbb{R}, L(\mathcal{H}))$, $f, g \in C(\mathbb{R}, L^2(\mathbb{P}, \mathcal{H}))$, $0 < \varepsilon \le \varepsilon_0$ and W is a two-sided standard one-dimensional Brownian motion defined on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t)$, where $\mathcal{F}_t := \sigma\{W(u) : u \le t\}$.

Definition 3.1. Let $f : \mathbb{R}\times(0, \varepsilon_0] \to \mathfrak{B}$. Following [17] we say that $f(t; \varepsilon)$ *integrally converges to* 0 if for any $L > 0$ we have

$$\lim_{\varepsilon\to 0} \sup_{|t-s|\le L} \left| \int_s^t f(\tau; \varepsilon)\mathrm{d}\tau \right| = 0.$$

If additionally there exists a constant $m > 0$ such that

$$|f(t; \varepsilon)| \le m$$

for any $t \in \mathbb{R}$ and $0 < \varepsilon \le \varepsilon_0$, then we say that $f(t; \varepsilon)$ *correctly converges to* 0 as $\varepsilon \to 0$.

Remark 9. [17, ChIV] If $f \in C(\mathbb{R}, \mathfrak{B})$ and

$$\lim_{T\to+\infty} \frac{1}{T} \left| \int_t^{t+T} f(s)\mathrm{d}s \right| = 0$$

uniformly with respect to $t \in \mathbb{R}$, then $f(t; \varepsilon) := f(\frac{t}{\varepsilon})$ integrally converges to 0 as $\varepsilon \to 0$. If additionally the function f is bounded on $\mathbb{R}$, then $f(t; \varepsilon)$ correctly converges to 0 as $\varepsilon \to 0$.

Let $\mathcal{A} \in L(\mathcal{H})$. Denote by $\sigma(\mathcal{A})$ the spectrum of $\mathcal{A}$. Below we will use the following conditions:

(**A1**) $\mathcal{A} \in C(\mathbb{R}, L(\mathcal{H}))$ and there exists $\bar{\mathcal{A}} \in L(\mathcal{H})$ such that

$$\lim_{T\to+\infty} \frac{1}{T} \int_t^{t+T} \mathcal{A}(s)\mathrm{d}s = \bar{\mathcal{A}} \tag{8}$$

uniformly with respect to $t \in \mathbb{R}$;

(**A2**) $f \in C(\mathbb{R}, L^2(\mathbb{P}, \mathcal{H}))$ and there exists $\bar{f} \in L^2(\mathbb{P}, \mathcal{H})$ such that

$$\lim_{T\to+\infty} \frac{1}{T} \left\| \int_t^{t+T} [f(s) - \bar{f}]\mathrm{d}s \right\|_2 = 0 \tag{9}$$

uniformly with respect to $t \in \mathbb{R}$;

(**A3**) $g \in C(\mathbb{R}, L^2(\mathbb{P}, \mathcal{H}))$ and there exists $\bar{g} \in L^2(\mathbb{P}, \mathcal{H})$ such that

$$\lim_{T\to+\infty} \frac{1}{T} \int_t^{t+T} \mathbb{E}|g(\tau) - \bar{g}|^2\mathrm{d}\tau = 0 \tag{10}$$

uniformly with respect to $t \in \mathbb{R}$.

Denote by Ψ the family of all decreasing, positive bounded functions $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ with $\lim_{t\to+\infty} \psi(t) = 0$.

Lemma 3.2. *Let $l > 0$ and $\psi \in \Psi$, then*

$$\lim_{\varepsilon \to 0} \sup_{0 \le \tau \le l} \tau \psi(\frac{\tau}{\varepsilon}) = 0.$$

Proof. Let ε and l be two arbitrary positive numbers, $\nu \in (0,1)$ and $\psi \in \Psi$, then we have

$$\sup_{0 \le \tau \le l} \tau \psi(\frac{\tau}{\varepsilon}) \le \sup_{0 \le \tau \le \varepsilon^{\nu}} \tau \psi(\frac{\tau}{\varepsilon}) + \sup_{\varepsilon^{\nu} \le \tau \le l} \tau \psi(\frac{\tau}{\varepsilon}) \le \varepsilon^{\nu} \psi(0) + l \psi(\varepsilon^{\nu - 1}).$$

Letting $\varepsilon \to 0$ we obtain the required result. $\square$

Remark 10. (i) By Lemma 2 in [8] equality (8) holds if and only if there exists a function $\omega \in \Psi$ satisfying

$$\left\| \frac{1}{T} \int_t^{t+T} \mathcal{A}(s) \mathrm{d}s - \bar{\mathcal{A}} \right\| \le \omega(T)$$

for any $T > 0$ and $t \in \mathbb{R}$.

(ii) Similarly equality (9) (respectively, equality (10)) holds if and only if there exists a function $\omega_1 \in \Psi$ (respectively, $\omega_2 \in \Psi$) satisfying

$$\frac{1}{T} \left\| \int_t^{t+T} [f(s) - \bar{f}] \mathrm{d}s \right\|_2 \le \omega_1(T)$$

(respectively, $\frac{1}{T} \int_t^{t+T} \mathbb{E}|g(\tau) - \bar{g}|^2 \mathrm{d}\tau \le \omega_2(T)$) for any $T > 0$ and $t \in \mathbb{R}$.

Theorem 3.3. [17, ChIV] *Suppose that $\mathcal{A} \in C_b(\mathbb{R}, L(\mathfrak{B}))$ and*

$$\lim_{T \to +\infty} \frac{1}{T} \int_t^{t+T} \mathcal{A}(s) \mathrm{d}s = \bar{\mathcal{A}}$$

uniformly with respect to $t \in \mathbb{R}$ and the operator $\bar{\mathcal{A}}$ is Hurwitz, i.e. $\mathcal{R}e\ \lambda < 0$ for any $\lambda \in \sigma(\bar{\mathcal{A}})$.

Then the following statements hold:

1. *there exists a positive constant $\alpha \le \varepsilon_0$ such that the equation*

$$\mathrm{d}x(\tau) = \mathcal{A}_{\varepsilon}(\tau) x(\tau) \mathrm{d}\tau,$$

where $\mathcal{A}_{\varepsilon}(\tau) := \mathcal{A}(\frac{\tau}{\varepsilon})$ for any $\tau \in \mathbb{R}$, is uniformly asymptotically stable for any $0 < \varepsilon \le \alpha$. Moreover there are constants $\mathcal{N} > 0$ and $\nu > 0$ such that

$$\|G_{\mathcal{A}_{\varepsilon}}(\tau, \tau_0)\| \le \mathcal{N} e^{-\nu(\tau - \tau_0)}$$

for any $\tau \ge \tau_0$ and $0 < \varepsilon \le \alpha$;

2. *there exists $\gamma_0 > 0$ such that*

$$\lim_{\varepsilon \to 0} \sup_{(\tau \ge \tau_0;\ \tau, \tau_0 \in \mathbb{R})} e^{\gamma_0(\tau - \tau_0)} \|G_{\mathcal{A}_{\varepsilon}}(\tau, \tau_0) - G_{\bar{\mathcal{A}}}(\tau, \tau_0)\| = 0\ . \tag{11}$$

Remark 11. (i) Note that Theorem 3.3 was proved for finite-dimensional almost periodic equations (this means that the matrix-function $\mathcal{A}(\cdot)$ is almost periodic). For the proof for infinite-dimensional almost periodic systems see [19, ChXI].

(ii) It is not difficult to show that Theorem 3.3 remains true in general case (see above) and can be proved with slight modifications of the reasoning from [17, ChIV].

(iii) Under the conditions of Theorem 3.3 there are positive constants $\alpha, \mathcal{N}$ and ν so that

$$\|G_{\mathcal{A}_{\varepsilon}}(t, \tau)\|, \|G_{\bar{\mathcal{A}}}(t, \tau)\| \le \mathcal{N} e^{-\nu(t - \tau)} \tag{12}$$

for any $0 < \varepsilon \le \alpha$ and $t \ge \tau$.

Lemma 3.4. *Let $f_\varepsilon \in C(\mathbb{R}, \mathfrak{B})$ for $\varepsilon \in (0, \alpha]$ be functions satisfying the following conditions:*

1. *there exists a positive constant A such that $|f_\varepsilon(t)| \le A$ for any $t \in \mathbb{R}$ and $\varepsilon \in (0, \alpha]$;*
2. *for any $l > 0$*
$$\lim_{\varepsilon \to 0} \sup_{|s| \le l,\ t \in \mathbb{R}} \left| \int_t^{t+s} f_\varepsilon(\sigma) \mathrm{d}\sigma \right| = 0. \tag{13}$$

Then for any $\nu > 0$ we have
$$\lim_{\varepsilon \to 0} \sup_{t \in \mathbb{R}} \left| \int_{-\infty}^t e^{-\nu(t-\tau)} f_\varepsilon(\tau) \mathrm{d}\tau \right| = 0.$$

Proof. To estimate the integral
$$I(t, \varepsilon) := \left| \int_{-\infty}^t e^{-\nu(t-\tau)} f_\varepsilon(\tau) \mathrm{d}\tau \right|,$$
we make the change $\tau - t = s$, then
$$\int_{-\infty}^t e^{-\nu(t-\tau)} f_\varepsilon(\tau) \mathrm{d}\tau = \int_{-\infty}^0 e^{\nu s} f_\varepsilon(t+s) \mathrm{d}s = \int_{-\infty}^0 e^{\nu s} \frac{\mathrm{d}}{\mathrm{d}s} \left(\int_t^{t+s} f_\varepsilon(\sigma) \mathrm{d}\sigma \right) \mathrm{d}s. \tag{14}$$

Since
$$\left| e^{\nu s} \int_t^{t+s} f_\varepsilon(\sigma) \mathrm{d}\sigma \right| \le A e^{\nu s} |s|$$
for any $s < 0$, we have
$$\lim_{s \to -\infty} e^{\nu s} \int_t^{t+s} f_\varepsilon(\sigma) \mathrm{d}\sigma = 0\ . \tag{15}$$

Integrating by parts and taking into consideration (15) we obtain
$$\begin{aligned}
&\int_{-\infty}^0 e^{\nu s} \frac{\mathrm{d}}{\mathrm{d}s} \left(\int_t^{t+s} f_\varepsilon(\sigma) \mathrm{d}\sigma \right) \mathrm{d}s \\
&= -\int_{-\infty}^0 \nu e^{\nu s} \int_t^{t+s} f_\varepsilon(\sigma) \mathrm{d}\sigma \mathrm{d}s \\
&= -\int_{-\infty}^{-l} \nu e^{\nu s} \int_t^{t+s} f_\varepsilon(\sigma) \mathrm{d}\sigma \mathrm{d}s - \int_{-l}^0 \nu e^{\nu s} \int_t^{t+s} f_\varepsilon(\sigma) \mathrm{d}\sigma \mathrm{d}s\ .
\end{aligned} \tag{16}$$

Note that
$$\left| -\int_{-\infty}^{-l} \nu e^{\nu s} \int_t^{t+s} f_\varepsilon(\sigma) \mathrm{d}\sigma \mathrm{d}s \right| \le A e^{-\nu l} \left(l + \frac{1}{\nu} \right) \tag{17}$$
and
$$\left| -\int_{-l}^0 \nu e^{\nu s} \int_t^{t+s} f_\varepsilon(\sigma) \mathrm{d}\sigma \mathrm{d}s \right| \le (1 - e^{-\nu l}) \sup_{|s| \le l,\ t \in \mathbb{R}} \left| \int_t^{t+s} f_\varepsilon(\sigma) \mathrm{d}\sigma \right|. \tag{18}$$

By (14)–(18) we get
$$\left| \int_{-\infty}^t e^{-\nu(t-\tau)} f_\varepsilon(\tau) \mathrm{d}\tau \right| \le A e^{-\nu l} \left(l + \frac{1}{\nu} \right) + (1 - e^{-\nu l}) \sup_{|s| \le l,\ t \in \mathbb{R}} \left| \int_t^{t+s} f_\varepsilon(\sigma) \mathrm{d}\sigma \right|.$$

Then

$$\sup_{t\in\mathbb{R}} I(t,\varepsilon) \le Ae^{-\nu l}\left(l+\frac{1}{\nu}\right) + (1-e^{-\nu l}) \sup_{|s|\le l,\ t\in\mathbb{R}} \left|\int_t^{t+s} f_\varepsilon(\sigma)\mathrm{d}\sigma\right|. \tag{19}$$

Since

$$\lim_{\varepsilon\to 0} \sup_{|s|\le l,\ t\in\mathbb{R}} \left|\int_t^{t+s} f_\varepsilon(\sigma)\mathrm{d}\sigma\right| = 0$$

for any $l>0$, letting $\varepsilon\to 0$ in (19) we have

$$\limsup_{\varepsilon\to 0} \sup_{t\in\mathbb{R}} I(t,\varepsilon) \le Ae^{-\nu l}\left(l+\frac{1}{\nu}\right).$$

Since l is arbitrary, it follows that

$$\lim_{\varepsilon\to 0} \sup_{t\in\mathbb{R}} I(t,\varepsilon) = 0.$$

The proof is complete. □

Remark 12. If the function $f\in C_b(\mathbb{R},\mathfrak{B})$ and $\bar{f}\in\mathfrak{B}$ are such that

$$\lim_{L\to+\infty} \frac{1}{L}\int_t^{t+L} [f(s)-\bar{f}]\mathrm{d}s = 0 \tag{20}$$

uniformly with respect to $t\in\mathbb{R}$, then the function $f_\varepsilon(\sigma) := f(\frac{\sigma}{\varepsilon}) - \bar{f}$ satisfies the conditions of Lemma 3.4. Indeed, note that

$$\int_t^{t+s} f_\varepsilon(\sigma)\mathrm{d}\sigma = \int_t^{t+s} [f(\frac{\sigma}{\varepsilon})-\bar{f}]\mathrm{d}\sigma = s\cdot\frac{\varepsilon}{s}\int_{t/\varepsilon}^{t/\varepsilon+s/\varepsilon} [f(\tilde{\sigma})-\bar{f}]\mathrm{d}\tilde{\sigma},$$

so the condition (13) of Lemma 3.4 holds by (20). Similarly, if the function g in (A3) is L^2-bounded, then the function $f_\varepsilon(\sigma) := \mathbb{E}|g(\frac{\sigma}{\varepsilon})-\bar{g}|^2$ satisfies as well the conditions of Lemma 3.4.

Let $W_\varepsilon(t) := \sqrt{\varepsilon}W(\frac{t}{\varepsilon})$ for $t\in\mathbb{R}$. Then W_ε is also a Brownian motion with the same distribution as W.

Theorem 3.5. *Suppose that $\mathcal{A}\in C_b(\mathbb{R},L(\mathcal{H}))$, $f,g\in C_b(\mathbb{R},L^2(\mathbb{P},\mathcal{H}))$ and conditions (A1)–(A3) are fulfilled. Suppose further that $\bar{\mathcal{A}}$ in (A1) is Hurwitz such that (11)–(12) holds. Then we have the following conclusions:*

1. *equation*

$$\mathrm{d}X_\varepsilon(t) = (\mathcal{A}_\varepsilon(t)X_\varepsilon(t)+f_\varepsilon(t))\mathrm{d}t + g_\varepsilon(t)\mathrm{d}W_\varepsilon(t) \tag{21}$$

has a unique bounded solution $\psi_\varepsilon\in C_b(\mathbb{R},L^2(\mathbb{P},\mathcal{H}))$ defined by equality

$$\psi_\varepsilon(t) = \int_{-\infty}^t G_{\mathcal{A}_\varepsilon}(t,\tau)f_\varepsilon(\tau)\mathrm{d}\tau + \int_{-\infty}^t G_{\mathcal{A}_\varepsilon}(t,\tau)g_\varepsilon(\tau)\mathrm{d}W_\varepsilon(\tau)$$

and it is strongly compatible in distribution (i.e. $\mathfrak{M}_{(\mathcal{A}_\varepsilon,f_\varepsilon,g_\varepsilon)}\subseteq\tilde{\mathfrak{M}}_{\psi_\varepsilon}$) and $\mathfrak{M}^u_{(\mathcal{A}_\varepsilon,f_\varepsilon,g_\varepsilon)}\subseteq\tilde{\mathfrak{M}}^u_{\psi_\varepsilon}$, where $\mathcal{A}_\varepsilon(t) := \mathcal{A}(\frac{t}{\varepsilon})$, $f_\varepsilon(t) := f(\frac{t}{\varepsilon})$ and $g_\varepsilon(t) := g(\frac{t}{\varepsilon})$ for $t\in\mathbb{R}$;

2. *equation*

$$\mathrm{d}X_\varepsilon(t) = (\mathcal{A}_\varepsilon(t)X_\varepsilon(t)+f_\varepsilon(t))\mathrm{d}t + g_\varepsilon(t)\mathrm{d}W(t)$$

has a unique bounded solution $\phi_\varepsilon\in C_b(\mathbb{R},L^2(\mathbb{P},\mathcal{H}))$ defined by equality

$$\phi_\varepsilon(t) = \int_{-\infty}^t G_{\mathcal{A}_\varepsilon}(t,\tau)f_\varepsilon(\tau)\mathrm{d}\tau + \int_{-\infty}^t G_{\mathcal{A}_\varepsilon}(t,\tau)g_\varepsilon(\tau)\mathrm{d}W(\tau); \tag{22}$$

3.

$$\lim_{\varepsilon\to 0}\sup_{t\in\mathbb{R}}\mathbb{E}|\phi_\varepsilon(t)-\bar\phi(t)|^2=0, \tag{23}$$

where $\bar\phi$ is the unique stationary solution of equation

$$\mathrm{d}X(t)=(\bar{\mathcal{A}}X(t)+\bar f)\mathrm{d}t+\bar g\mathrm{d}W(t); \tag{24}$$

4. for any $\varepsilon\in(0,\alpha]$ equation (7) has a unique solution $\varphi_\varepsilon\in C_b(\mathbb{R},L^2(\mathbb{P},\mathcal{H}))$ and it is strongly compatible in distribution (i.e. $\mathfrak{M}_{(\mathcal{A},f,g)}\subseteq\tilde{\mathfrak{M}}_{\varphi_\varepsilon}$) and $\mathfrak{M}^u_{(\mathcal{A},f,g)}\subseteq\tilde{\mathfrak{M}}^u_{\varphi_\varepsilon}$;

5.

$$\lim_{\varepsilon\to 0}\sup_{t\in\mathbb{R}}\beta(\mathcal{L}(\varphi_\varepsilon(\frac{t}{\varepsilon})),\mathcal{L}(\bar\phi(t)))=0,$$

where $\mathcal{L}(X)$ denotes the law of random variable X.

Proof. The first and second statements follow directly from Theorem 2.20.

We now verify the third statement, i.e. the uniform convergence of the unique bounded solution ϕ_ε to the unique stationary solution $\bar\phi$ of the averaged equation. By Theorem 2.20 equation (24) has a unique bounded and stationary solution $\bar\phi$, which is given by the formula

$$\bar\phi(t)=\int_{-\infty}^t G_{\bar{\mathcal{A}}}(t,\tau)\bar f\mathrm{d}\tau+\int_{-\infty}^t G_{\bar{\mathcal{A}}}(t,\tau)\bar g\mathrm{d}W(\tau), \tag{25}$$

where $G_{\bar{\mathcal{A}}}(t,\tau)=\exp\left\{\bar{\mathcal{A}}(t-\tau)\right\}$ for $t,\tau\in\mathbb{R}$. From (22) and (25) we get

$$\begin{aligned}
&\mathbb{E}|\phi_\varepsilon(t)-\bar\phi(t)|^2\\
&=\mathbb{E}\bigg|\int_{-\infty}^t G_{\mathcal{A}_\varepsilon}(t,\tau)f_\varepsilon(\tau)\mathrm{d}\tau+\int_{-\infty}^t G_{\mathcal{A}_\varepsilon}(t,\tau)g_\varepsilon(\tau)\mathrm{d}W(\tau)\\
&\qquad-\int_{-\infty}^t G_{\bar{\mathcal{A}}}(t,\tau)\bar f\mathrm{d}\tau-\int_{-\infty}^t G_{\bar{\mathcal{A}}}(t,\tau)\bar g\mathrm{d}W(\tau)\bigg|^2\\
&\le 2\bigg(\mathbb{E}\left|\int_{-\infty}^t[G_{\mathcal{A}_\varepsilon}(t,\tau)f_\varepsilon(\tau)-G_{\bar{\mathcal{A}}}(t,\tau)\bar f]\mathrm{d}\tau\right|^2\\
&\qquad+\mathbb{E}\left|\int_{-\infty}^t[G_{\mathcal{A}_\varepsilon}(t,\tau)g_\varepsilon(\tau)-G_{\bar{\mathcal{A}}}(t,\tau)\bar g]\,\mathrm{d}W(\tau)\right|^2\bigg)=:I_1(t,\varepsilon)+I_2(t,\varepsilon).
\end{aligned}$$

By equality (11) there exists a function $\mathcal{N}:(0,\alpha)\to\mathbb{R}_+$ such that $\mathcal{N}(\varepsilon)\to 0$ as $\varepsilon\to 0$ and

$$\|G_{\mathcal{A}_\varepsilon}(t,\tau)-G_{\bar{\mathcal{A}}}(t,\tau)\|\le\mathcal{N}(\varepsilon)e^{-\gamma_0(t-\tau)}$$

for any $t\ge\tau$ $(t,\tau\in\mathbb{R})$.

Note that

$$\begin{aligned}
&I_1(t,\varepsilon)\\
&:=2\mathbb{E}\left|\int_{-\infty}^t[G_{\mathcal{A}_\varepsilon}(t,\tau)f_\varepsilon(\tau)-G_{\bar{\mathcal{A}}}(t,\tau)\bar f]\mathrm{d}\tau\right|^2\\
&=2\mathbb{E}\left|\int_{-\infty}^t[G_{\mathcal{A}_\varepsilon}(t,\tau)f_\varepsilon(\tau)-G_{\mathcal{A}_\varepsilon}(t,\tau)\bar f]+[G_{\mathcal{A}_\varepsilon}(t,\tau)\bar f-G_{\bar{\mathcal{A}}}(t,\tau)\bar f]\mathrm{d}\tau\right|^2\\
&\le 4\mathbb{E}\left|\int_{-\infty}^t G_{\mathcal{A}_\varepsilon}(t,\tau)\left(f_\varepsilon(\tau)-\bar f\right)\mathrm{d}\tau\right|^2+4\mathbb{E}\left|\int_{-\infty}^t[G_{\mathcal{A}_\varepsilon}(t,\tau)-G_{\bar{\mathcal{A}}}(t,\tau)]\,\bar f\mathrm{d}\tau\right|^2
\end{aligned}$$

$$=: 4\left(I_{11}(t,\varepsilon)+I_{12}(t,\varepsilon)\right).$$

To estimate the integral

$$I_{11}(t,\varepsilon):=\mathbb{E}\left|\int_{-\infty}^{t} G_{\mathcal{A}_\varepsilon}(t,\tau)(f_\varepsilon(\tau)-\bar{f})\mathrm{d}\tau\right|^2,$$

making the change of variable $s:=\tau-t$ we obtain

$$\begin{aligned}\int_{-\infty}^{t} G_{\mathcal{A}_\varepsilon}(t,\tau)\left(f_\varepsilon(\tau)-\bar{f}\right)\mathrm{d}\tau &= \int_{-\infty}^{0} G_{\mathcal{A}_\varepsilon}(t,t+s)\left(f_\varepsilon(t+s)-\bar{f}\right)\mathrm{d}s \\ &= \int_{-\infty}^{0} G_{\mathcal{A}_\varepsilon}(t,t+s)\frac{\mathrm{d}}{\mathrm{d}s}\left(\int_{t}^{t+s}[f_\varepsilon(\sigma)-\bar{f}]\mathrm{d}\sigma\right)\mathrm{d}s.\end{aligned}\tag{26}$$

Since

$$\left\|G_{\mathcal{A}_\varepsilon}(t,t+s)\int_{t}^{t+s}[f_\varepsilon(\sigma)-\bar{f}]\mathrm{d}\sigma\right\|_2 \le 2\mathcal{N}\|f\|_\infty e^{\nu s}|s|$$

for any $s<0$, we have

$$\lim_{s\to-\infty} G_{\mathcal{A}_\varepsilon}(t,t+s)\int_{t}^{t+s}[f_\varepsilon(\sigma)-\bar{f}]\mathrm{d}\sigma=0.$$

Consequently, integrating by parts from (26) we get

$$\begin{aligned}&\int_{-\infty}^{0} G_{\mathcal{A}_\varepsilon}(t,t+s)\frac{\mathrm{d}}{\mathrm{d}s}\left(\int_{t}^{t+s}[f_\varepsilon(\sigma)-\bar{f}]\mathrm{d}\sigma\right)\mathrm{d}s \\ &= -\int_{-\infty}^{0}\frac{\partial G_{\mathcal{A}_\varepsilon}(t,t+s)}{\partial s}\left(\int_{t}^{t+s}[f_\varepsilon(\sigma)-\bar{f}]\mathrm{d}\sigma\right)\mathrm{d}s.\end{aligned}$$

Note that

$$\frac{\partial G_{\mathcal{A}}(t,\tau)}{\partial \tau} = -G_{\mathcal{A}}(t,\tau)\mathcal{A}(\tau),$$

so we have

$$\left\|\frac{\partial G_{\mathcal{A}_\varepsilon}(t,t+s)}{\partial s}\right\| \le \mathcal{N}\|\mathcal{A}\|_\infty e^{\nu s}$$

for any $t\in\mathbb{R}$ and $s<0$.

Let now l be an arbitrary positive number, then we have

$$\begin{aligned}&-\int_{-\infty}^{0}\frac{\partial G_{\mathcal{A}_\varepsilon}(t,t+s)}{\partial s}\left(\int_{t}^{t+s}[f_\varepsilon(\sigma)-\bar{f}]\mathrm{d}\sigma\right)\mathrm{d}s \\ &= -\int_{-\infty}^{-l}\frac{\partial G_{\mathcal{A}_\varepsilon}(t,t+s)}{\partial s}\left(\int_{t}^{t+s}[f_\varepsilon(\sigma)-\bar{f}]\mathrm{d}\sigma\right)\mathrm{d}s \\ &\quad -\int_{-l}^{0}\frac{\partial G_{\mathcal{A}_\varepsilon}(t,t+s)}{\partial s}\left(\int_{t}^{t+s}[f_\varepsilon(\sigma)-\bar{f}]\mathrm{d}\sigma\right)\mathrm{d}s\end{aligned}$$

and consequently

$$\begin{aligned}&\left\|-\int_{-\infty}^{0}\frac{\partial G_{\mathcal{A}_\varepsilon}(t,t+s)}{\partial s}\left(\int_{t}^{t+s}[f_\varepsilon(\sigma)-\bar{f}]\mathrm{d}\sigma\right)\mathrm{d}s\right\|_2 \\ &\le \left\|-\int_{-\infty}^{-l}\frac{\partial G_{\mathcal{A}_\varepsilon}(t,t+s)}{\partial s}\left(\int_{t}^{t+s}[f_\varepsilon(\sigma)-\bar{f}]\mathrm{d}\sigma\right)\mathrm{d}s\right\|_2 \\ &\quad +\left\|\int_{-l}^{0}\frac{\partial G_{\mathcal{A}_\varepsilon}(t,t+s)}{\partial s}\left(\int_{t}^{t+s}[f_\varepsilon(\sigma)-\bar{f}]\mathrm{d}\sigma\right)\mathrm{d}s\right\|_2\end{aligned}$$

$$\leq \mathcal{N}\|\mathcal{A}\|_\infty \left(2\|f\|_\infty \left|\int_{-\infty}^{-l} s e^{\nu s}\mathrm{d}s\right| + \left|\int_{-l}^{0} e^{\nu s}\mathrm{d}s\right| \sup_{|s|\leq l,\ t\in\mathbb{R}} \left\|\int_t^{t+s}[f_\varepsilon(\sigma)-\bar f]\mathrm{d}\sigma\right\|_2\right)$$
$$\leq \mathcal{N}\|\mathcal{A}\|_\infty \left(2\|f\|_\infty e^{-\nu l}\left(l+\frac{1}{\nu}\right) + \frac{1}{\nu}(1-e^{-\nu l}) \sup_{|s|\leq l,\ t\in\mathbb{R}} \left\|\int_t^{t+s}[f_\varepsilon(\sigma)-\bar f]\mathrm{d}\sigma\right\|_2\right).$$

Letting $\varepsilon \to 0$ in above inequality we get

$$\limsup_{\varepsilon\to 0}\sup_{t\in\mathbb{R}} \left\|-\int_{-\infty}^{0} \frac{\partial G_{\mathcal{A}_\varepsilon}(t,t+s)}{\partial s}\left(\int_t^{t+s}[f_\varepsilon(\sigma)-\bar f]\mathrm{d}\sigma\right)\mathrm{d}s\right\|_2$$
$$\leq 2\mathcal{N}\|\mathcal{A}\|_\infty\|f\|_\infty e^{-\nu l}\left(l+\frac{1}{\nu}\right).$$

Since l is arbitrary, we get by letting $l\to\infty$

$$\lim_{\varepsilon\to 0}\sup_{t\in\mathbb{R}} I_{11}(t,\varepsilon) = 0.$$

Note by Theorem 3.3–(ii) that

$$I_{12}(t,\varepsilon) := \mathbb{E}\left|\int_{-\infty}^{t}[G_{\mathcal{A}_\varepsilon}(t,\tau)-G_{\bar{\mathcal{A}}}(t,\tau)]\bar f\mathrm{d}\tau\right|^2 \leq \left(\frac{\|\bar f\|_2\mathcal{N}(\varepsilon)}{\gamma_0}\right)^2 \to 0$$

as $\varepsilon\to 0$. Consequently,

$$\lim_{\varepsilon\to 0}\sup_{t\in\mathbb{R}} I_1(t,\varepsilon) = 0.$$

Similarly we can show that

$$\lim_{\varepsilon\to 0}\sup_{t\in\mathbb{R}} I_2(t,\varepsilon) = 0. \tag{27}$$

In fact, using Itô's isometry property, the Cauchy-Schwartz inequality and reasoning as above we get

$$I_2(t,\varepsilon) = 2\mathbb{E}\left|\int_{-\infty}^{t}[G_{\mathcal{A}_\varepsilon}(t,\tau)g_\varepsilon(\tau)-G_{\bar{\mathcal{A}}}(t,\tau)\bar g]\mathrm{d}W(\tau)\right|^2 \tag{28}$$
$$= 2\mathbb{E}\int_{-\infty}^{t}\left|G_{\mathcal{A}_\varepsilon}(t,\tau)g_\varepsilon(\tau)-G_{\bar{\mathcal{A}}}(t,\tau)\bar g\right|^2\mathrm{d}\tau$$
$$\leq 4\Big(\mathbb{E}\int_{-\infty}^{t}\left|G_{\mathcal{A}_\varepsilon}(t,\tau)(g_\varepsilon(\tau)-\bar g)\right|^2\mathrm{d}\tau$$
$$+\,\mathbb{E}\int_{-\infty}^{t}\left|(G_{\mathcal{A}_\varepsilon}(t,\tau)-G_{\bar{\mathcal{A}}}(t,\tau))\,\bar g\right|^2\mathrm{d}\tau\Big)$$
$$\leq 4\left(\mathbb{E}\int_{-\infty}^{t}\mathcal{N}^2 e^{-2\nu(t-\tau)}|g_\varepsilon(\tau)-\bar g|^2\mathrm{d}\tau + \mathbb{E}\int_{-\infty}^{t}\mathcal{N}(\varepsilon)^2 e^{-2\gamma_0(t-\tau)}|\bar g|^2\mathrm{d}\tau\right)$$
$$= 4\left(\mathcal{N}^2\int_{-\infty}^{t}e^{-2\nu(t-\tau)}\mathbb{E}|g_\varepsilon(\tau)-\bar g|^2\mathrm{d}\tau + (\mathcal{N}(\varepsilon))^2\frac{\|\bar g\|_2^2}{2\gamma_0}\right).$$

By Lemma 3.4 the integral

$$\int_{-\infty}^{t}e^{-2\nu(t-\tau)}\mathbb{E}|g_\varepsilon(\tau)-\bar g|^2\mathrm{d}\tau \tag{29}$$

goes to 0 as $\varepsilon\to 0$ uniformly with respect to $t\in\mathbb{R}$.

Passing to the limit in (28) and taking into account (29) we obtain (27), and consequently $\lim\limits_{\varepsilon\to 0}\sup\limits_{t\in\mathbb{R}}\mathbb{E}|\phi_\varepsilon(t)-\bar\phi(t)|^2 = 0$.

To prove the fourth statement we note that the function $\varphi_\varepsilon(t) := \psi_\varepsilon(\varepsilon t)$ (for $t \in \mathbb{R}$) is a bounded solution of equation (7) if ψ_ε is a bounded solution of equation (21). The uniqueness follows from the fact if φ_i $(i = 1, 2)$ are two different bounded solutions of equation (7), then $\psi_i(t) := \varphi_i(\frac{t}{\varepsilon}), t \in \mathbb{R}$ $(i = 1, 2)$ are two different bounded solutions of equation (21), a contradiction to the first statement. It remains to show that $\mathfrak{M}_{(\mathcal{A},f,g)} \subseteq \tilde{\mathfrak{M}}_{\varphi_\varepsilon}$ and $\mathfrak{M}^u_{(\mathcal{A},f,g)} \subseteq \tilde{\mathfrak{M}}^u_{\varphi_\varepsilon}$. Let $\{t_n\} \in \mathfrak{M}_{(\mathcal{A},f,g)}$ (respectively, $\{t_n\} \in \mathfrak{M}^u_{(\mathcal{A},f,g)}$), then $\{\varepsilon t_n\} \in \mathfrak{M}_{(\mathcal{A}_\varepsilon,f_\varepsilon,g_\varepsilon)} \subseteq \tilde{\mathfrak{M}}_{\psi_\varepsilon}$ (respectively, $\{\varepsilon t_n\} \in \mathfrak{M}^u_{(\mathcal{A}_\varepsilon,f_\varepsilon,g_\varepsilon)} \subseteq \tilde{\mathfrak{M}}^u_{\psi_\varepsilon}$) by Theorem 2.20. By the relation between φ_ε and ψ_ε, we have $\{t_n\} \in \tilde{\mathfrak{M}}_{\varphi_\varepsilon}$ (respectively, $\{t_n\} \in \tilde{\mathfrak{M}}^u_{\varphi_\varepsilon}$).

Now we are in the position to prove the last statement. Since the L^2 convergence implies convergence in probability, it follows from (23) that

$$\lim_{\varepsilon\to 0} \sup_{t\in\mathbb{R}} \beta(\mathcal{L}(\phi_\varepsilon(t)), \mathcal{L}(\bar{\phi}(t))) = 0.$$

On the other hand taking into consideration that $\mathcal{L}(W) = \mathcal{L}(W_\varepsilon)$, we have $\mathcal{L}(\varphi_\varepsilon(\frac{t}{\varepsilon}))$ $= \mathcal{L}(\phi_\varepsilon(t))$ for any $t \in \mathbb{R}$, and consequently

$$\lim_{\varepsilon\to 0} \sup_{t\in\mathbb{R}} \beta(\mathcal{L}(\varphi_\varepsilon(\frac{t}{\varepsilon})), \mathcal{L}(\bar{\phi}(t))) = 0.$$

The proof is complete. □

Corollary 2. *Under the conditions of Theorem 3.5 the following statements hold:*

1. *if the functions $\mathcal{A} \in C(\mathbb{R}, L(\mathcal{H}))$ and $f, g \in C_b(\mathbb{R}, L^2(\mathbb{P}, \mathcal{H}))$ are jointly stationary (respectively, τ-periodic, quasi-periodic with the spectrum of frequencies $\nu_1, \dots, \nu_k$, Bohr almost periodic, almost automorphic, Birkhoff recurrent, Lagrange stable, Levitan almost periodic, almost recurrent, Poisson stable), then equation (7) has a unique solution $\varphi_\varepsilon \in C_b(\mathbb{R}, L^2(\mathbb{P}, \mathcal{H}))$ which is stationary (respectively, τ-periodic, quasi-periodic with the spectrum of frequencies $\nu_1, \dots, \nu_k$, Bohr almost periodic, almost automorphic, Birkhoff recurrent, Lagrange stable, Levitan almost periodic, almost recurrent, Poisson stable) in distribution;*
2. *if the functions $\mathcal{A} \in C(\mathbb{R}, L(\mathcal{H}))$ and $f, g \in C_b(\mathbb{R}, L^2(\mathbb{P}, \mathcal{H}))$ are Lagrange stable and jointly pseudo-periodic (respectively, pseudo-recurrent), then equation (7) has a unique solution $\varphi_\varepsilon \in C_b(\mathbb{R}, L^2(\mathbb{P}, \mathcal{H}))$ which is pseudo-periodic (respectively, pseudo-recurrent) in distribution;*
3. $$\lim_{\varepsilon\to 0} \sup_{t\in\mathbb{R}} \beta(\mathcal{L}(\varphi_\varepsilon(\frac{t}{\varepsilon}), \mathcal{L}(\bar{\phi}(t))) = 0\ .$$

Proof. These statements follow from Theorems 2.12 and 3.5. □

Remark 13. Note that the constant $\alpha > 0$ in Theorem 3.3, Remark 11, Lemma 3.4 and Theorem 3.5 can be chosen the same, e.g. we may choose the minimal one of them. But there is no restriction on ε_0 which we fix at the beginning of this section.

4. Averaging principle for semilinear stochastic differential equations.

As before, let ε_0 be some fixed positive constant. Consider the following stochastic differential equation

$$\mathrm{d}X(t) = \varepsilon\left(\mathcal{A}(t)X(t) + F(t, X(t))\right)\mathrm{d}t + \sqrt{\varepsilon}\, G(t, X(t))\mathrm{d}W(t), \tag{30}$$

where $\mathcal{A} \in C(\mathbb{R}, L(\mathcal{H}))$, $F, G \in C(\mathbb{R} \times \mathcal{H}, \mathcal{H})$, $0 < \varepsilon \le \varepsilon_0$ and W is a two-sided standard one-dimensional Brownian motion defined on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t)$, with $\mathcal{F}_t := \sigma\{W(u) : u \le t\}$. Below we will use the following conditions:

(G1) there exists a positive constant M such that

$$|F(t,0)| \vee |G(t,0)| \le M$$

for any $t \in \mathbb{R}$;

(G2) there exists a positive constant L such that

$$|F(t,x_1) - F(t,x_2)| \vee |G(t,x_1) - G(t,x_2)| \le L|x_1 - x_2|$$

for any $x_1, x_2 \in \mathcal{H}$ and $t \in \mathbb{R}$;

(G3) there exist functions $\omega_1 \in \Psi$ and $\bar{F} \in C(\mathcal{H}, \mathcal{H})$ such that

$$\frac{1}{T}\left|\int_t^{t+T} [F(s,x) - \bar{F}(x)]\mathrm{d}s\right| \le \omega_1(T)(1 + |x|)$$

for any $T > 0$, $x \in \mathcal{H}$ and $t \in \mathbb{R}$;

(G4) there exist functions $\omega_2 \in \Psi$ and $\bar{G} \in C(\mathcal{H}, \mathcal{H})$ such that

$$\frac{1}{T}\int_t^{t+T} \left|G(s,x) - \bar{G}(x)\right|^2 \mathrm{d}s \le \omega_2(T)(1 + |x|^2)$$

for any $T > 0$, $x \in \mathcal{H}$ and $t \in \mathbb{R}$;

(G5) $\mathcal{A} \in C(\mathbb{R}, L(\mathcal{H}))$ and there exists $\bar{\mathcal{A}} \in L(\mathcal{H})$ such that

$$\lim_{T \to +\infty} \frac{1}{T}\int_t^{t+T} \mathcal{A}(s)\mathrm{d}s = \bar{\mathcal{A}}$$

uniformly with respect to $t \in \mathbb{R}$.

Remark 14. Under the conditions $(\mathbf{G1}) - (\mathbf{G4})$ the functions $\bar{F}$ and $\bar{G}$ also possess the properties $(\mathbf{G1}) - (\mathbf{G2})$ with the same constants M and L.

We consider as well the following equations

$$\mathrm{d}X(t) = (\mathcal{A}_\varepsilon(t)X(t) + F_\varepsilon(t, X(t)))\mathrm{d}t + G_\varepsilon(t, X(t))\mathrm{d}W(t) \tag{31}$$

and

$$\mathrm{d}X(t) = (\mathcal{A}_\varepsilon(t)X(t) + F_\varepsilon(t, X(t)))\mathrm{d}t + G_\varepsilon(t, X(t))\mathrm{d}W_\varepsilon(t), \tag{32}$$

where $\mathcal{A}_\varepsilon(t) := \mathcal{A}(\frac{t}{\varepsilon})$, $F_\varepsilon(t,x) := F(\frac{t}{\varepsilon}, x)$ and $G_\varepsilon(t,x) := G(\frac{t}{\varepsilon}, x)$ for $t \in \mathbb{R}$, $x \in \mathcal{H}$ and $\varepsilon \in (0, \varepsilon_0]$. Here as before $W_\varepsilon(t) := \sqrt{\varepsilon}W(\frac{t}{\varepsilon})$ for $t \in \mathbb{R}$. Along with equations (31)–(32) we also consider the following averaged equation

$$\mathrm{d}X(t) = (\bar{\mathcal{A}}X(t) + \bar{F}(X(t)))\mathrm{d}t + \bar{G}(X(t))\mathrm{d}W(t). \tag{33}$$

Lemma 4.1. *Suppose $F, G \in C(\mathbb{R} \times \mathcal{H}, \mathcal{H})$ and that the conditions (G1)–(G2) hold. If φ is an L^2-bounded solution (i.e. $\|\varphi\|_\infty = \sup_{t \in \mathbb{R}} \mathbb{E}|\varphi(t)|^2 < +\infty$) of the equation*

$$\mathrm{d}X(t) = F(t, X(t))\mathrm{d}t + G(t, X(t))\mathrm{d}W(t).$$

then there exists a constant $C > 0$, depending only on $M, L, \|\varphi\|_\infty$, such that

$$\mathbb{E}\left|\varphi(t+h) - \varphi(t)\right|^2 \le Ch$$

and

$$\mathbb{E} \sup_{t \le s \le t+h} |\varphi(s)|^2 \le C(h^2 + 1)$$

for any $t \in \mathbb{R}$ and $h > 0$.

Proof. Since

$$\varphi(t+h) = \varphi(t) + \int_t^{t+h} F(\tau, \varphi(\tau))\mathrm{d}\tau + \int_t^{t+h} G(\tau, \varphi(\tau))\mathrm{d}W(\tau),$$

by Cauchy-Schwartz inequality and Itô's isometry property we have

$$\begin{aligned}
&\mathbb{E}|\varphi(t+h) - \varphi(t)|^2 \\
&\le 2\left(\mathbb{E}\left|\int_t^{t+h} F(\tau, \varphi(\tau))\mathrm{d}\tau\right|^2 + \mathbb{E}\left|\int_t^{t+h} G(\tau, \varphi(\tau))\mathrm{d}W(\tau)\right|^2\right) \\
&\le 2\left(\mathbb{E}\int_t^{t+h} 1^2\mathrm{d}\tau \cdot \mathbb{E}\int_t^{t+h} |F(\tau, \varphi(\tau))|^2\mathrm{d}\tau + \int_t^{t+h} \mathbb{E}|G(\tau, \varphi(\tau))|^2\mathrm{d}\tau\right) \\
&\le 2\left(h\int_t^{t+h} \mathbb{E}|F(\tau, \varphi(\tau))|^2\mathrm{d}\tau + \int_t^{t+h} \mathbb{E}|G(\tau, \varphi(\tau))|^2\mathrm{d}\tau\right) \\
&\le 4\left(h\int_t^{t+h} (M^2 + L^2\|\varphi\|_\infty^2)\mathrm{d}\tau + \int_t^{t+h} (M^2 + L^2\|\varphi\|_\infty^2)\mathrm{d}\tau\right) \\
&\le Ch.
\end{aligned}$$

Employing the BDG inequality (see, e.g. [11, Theorem 4.36] on page 114), we have

$$\begin{aligned}
&\mathbb{E}\sup_{t\le s\le t+h} |\varphi(s)|^2 \\
&\le 3\mathbb{E}|\varphi(t)|^2 + 3\mathbb{E}\sup_{t\le s\le t+h}\left|\int_t^s F(\tau, \varphi(\tau))\mathrm{d}\tau\right|^2 + 3\mathbb{E}\sup_{t\le s\le t+h}\left|\int_t^s G(\tau, \varphi(\tau))\mathrm{d}W(\tau)\right|^2 \\
&\le 3\|\varphi\|_\infty^2 + 3\mathbb{E}\sup_{t\le s\le t+h}\left|\int_t^s (M + L|\varphi(\tau)|)\mathrm{d}\tau\right|^2 + 3C\mathbb{E}\int_t^{t+h} |G(\tau, \varphi(\tau))|^2\mathrm{d}\tau \\
&\le 3\|\varphi\|_\infty^2 + 3h\int_t^{t+h} 2(M^2 + L^2\|\varphi\|_\infty^2)\mathrm{d}\tau + 3C\int_t^{t+h} 2(M^2 + L^2\|\varphi\|_\infty^2)\mathrm{d}\tau \\
&\le C(h^2+1),
\end{aligned}$$

where C denotes some positive constants which may change from line to line. □

Theorem 4.2. *Suppose that the following conditions hold:*

(a) $\sup_{t\in\mathbb{R}} \|\mathcal{A}(t)\| < +\infty$*;*

(b) the functions $\mathcal{A}$*,* F*,* G *satisfy the conditions* (G1)–(G5)*, and the operator* $\bar{\mathcal{A}}$ *in* (G5) *is Hurwitz, i.e.* $\mathcal{R}e\ \lambda < 0$ *for any* $\lambda \in \sigma(\bar{\mathcal{A}})$*;*

(c)

$$L < \frac{\nu}{\sqrt{3}\mathcal{N}\sqrt{2+\nu}},$$

where $\alpha, \mathcal{N}$ *and* ν *are the numbers figuring in Remark 11-(iii).*

Then there exists a positive constant $\varepsilon_1 \le \alpha$ *such that for any* $0 < \varepsilon \le \varepsilon_1$

1. *equation (30) has a unique solution* $\varphi_\varepsilon \in C_b(\mathbb{R}, L^2(\mathbb{P}, \mathcal{H}))$ *and* $\|\varphi_\varepsilon\|_\infty \le r$*, where* $r := \frac{\mathcal{N}M\sqrt{2+\nu}}{\nu - \mathcal{N}L\sqrt{2+\nu}}$*;*
2. *equation (31) has a unique solution* $\phi_\varepsilon \in C_b(\mathbb{R}, L^2(\mathbb{P}, \mathcal{H}))$ *and* $\|\phi_\varepsilon\|_\infty \le r$*;*

3. *if additionally F and G satisfy (C3) and $L < \frac{\nu}{2\mathcal{N}\sqrt{1+\nu}}$, then the solution φ_ε of equation (30) is strongly compatible in distribution (i.e. $\mathfrak{M}_{(\mathcal{A},F,G)} \subseteq \tilde{\mathfrak{M}}_{\varphi_\varepsilon}$) and $\mathfrak{M}^u_{(\mathcal{A},F,G)} \subseteq \tilde{\mathfrak{M}}^u_{\varphi_\varepsilon}$, recalling that $\tilde{\mathfrak{M}}^u_{\varphi_\varepsilon}$ means the set of all sequences $\{t_n\}$ such that $\varphi_\varepsilon(t+t_n)$ converges in distribution uniformly with respect to $t \in \mathbb{R}$;*
4. $$\lim_{\varepsilon\to 0}\sup_{t\in\mathbb{R}} \mathbb{E}|\phi_\varepsilon(t) - \bar{\phi}(t)|^2 = 0,$$
where $\bar{\phi}$ is the unique stationary solution of equation (33);
5. $$\lim_{\varepsilon\to 0}\sup_{t\in\mathbb{R}} \beta(\mathcal{L}(\varphi_\varepsilon(\frac{t}{\varepsilon})), \mathcal{L}(\bar{\phi}(t))) = 0.$$

Proof. By Theorem 3.3 (see also Remark 11-(iii)) there exist positive constants $\mathcal{N}, \nu$ and α such that equation
$$\mathrm{d}X(t) = \mathcal{A}_\varepsilon(t)X(t)\mathrm{d}t$$
is uniformly asymptotically stable for any $0 < \varepsilon \le \alpha$ and
$$\|G_{\mathcal{A}_\varepsilon}(t,\tau)\| \le \mathcal{N}e^{-\nu(t-\tau)}$$
for any $t \ge \tau$. By Theorem 3.3-(ii) there are $\gamma_0 > 0$ and $\mathcal{N} : (0,\alpha) \to \mathbb{R}_+$ such that $\mathcal{N}(\varepsilon) \to 0$ as $\varepsilon \to 0$ and
$$\|G_{\mathcal{A}_\varepsilon}(t,\tau) - G_{\bar{\mathcal{A}}}(t,\tau)\| \le \mathcal{N}(\varepsilon)e^{-\gamma_0(t-\tau)}$$
for any $t \ge \tau$.

Since $Lip(F_\varepsilon) = Lip(F) \le L$ and $Lip(G_\varepsilon) = Lip(G) \le L$, by Theorem 2.20 equation (30) (respectively, equation (31)) has a unique solution φ_ε (respectively, ϕ_ε) from $C_b(\mathbb{R}, L^2(\mathbb{P},\mathfrak{B}))$ with $\varphi_\varepsilon \in C(\mathbb{R}, B[0,r])$ (respectively, $\phi_\varepsilon \in C(\mathbb{R}, B[0,r])$), where
$$r := \frac{\mathcal{N}M\sqrt{2+\nu}}{\nu - \mathcal{N}L\sqrt{2+\nu}};$$
and the solution φ_ε is strongly compatible in distribution and $\mathfrak{M}^u_{(\mathcal{A},F,G)} \subseteq \tilde{\mathfrak{M}}^u_{\varphi_\varepsilon}$.

Let $\bar{\phi}$ be the unique stationary solution of equation (33). We now estimate $\mathbb{E}|\phi_\varepsilon(t) - \bar{\phi}(t)|^2$. To this end, we note that
$$\begin{aligned}
&\mathbb{E}|\phi_\varepsilon(t) - \bar{\phi}(t)|^2 \qquad (34)\\
&= \mathbb{E}\Big|\int_{-\infty}^t G_{\mathcal{A}_\varepsilon}(t,\tau)F_\varepsilon(\tau,\phi_\varepsilon(\tau))\mathrm{d}\tau + \int_{-\infty}^t G_{\mathcal{A}_\varepsilon}(t,\tau)G_\varepsilon(\tau,\phi_\varepsilon(\tau))\mathrm{d}W(\tau)\\
&\quad - \int_{-\infty}^t G_{\bar{\mathcal{A}}}(t,\tau)\bar{F}(\bar{\phi}(\tau))\mathrm{d}\tau - \int_{-\infty}^t G_{\bar{\mathcal{A}}}(t,\tau)\bar{G}(\bar{\phi}(\tau))\mathrm{d}W(\tau)\Big|^2\\
&\le 2\Big(\mathbb{E}\Big|\int_{-\infty}^t (G_{\mathcal{A}_\varepsilon}(t,\tau)F_\varepsilon(\tau,\phi_\varepsilon(\tau)) - G_{\bar{\mathcal{A}}}(t,\tau)\bar{F}(\bar{\phi}(\tau)))\mathrm{d}\tau\Big|^2\\
&\quad + \mathbb{E}\Big|\int_{-\infty}^t (G_{\mathcal{A}_\varepsilon}(t,\tau)G_\varepsilon(\tau,\phi_\varepsilon(\tau)) - G_{\bar{\mathcal{A}}}(t,\tau)\bar{G}(\bar{\phi}(\tau)))\mathrm{d}W(\tau)\Big|^2\Big)\\
&=: 2(I_1(t,\varepsilon) + I_2(t,\varepsilon)).
\end{aligned}$$
Since
$$I_1(t,\varepsilon) := \mathbb{E}\Big|\int_{-\infty}^t \big(G_{\mathcal{A}_\varepsilon}(t,\tau)F_\varepsilon(\tau,\phi_\varepsilon(\tau)) - G_{\bar{\mathcal{A}}}(t,\tau)\bar{F}(\bar{\phi}(\tau))\big)\,\mathrm{d}\tau\Big|^2$$

$$
\begin{aligned}
&\le 3\Big(\mathbb{E}\Big|\int_{-\infty}^{t} G_{\mathcal{A}_\varepsilon}(t,\tau)(F_\varepsilon(\tau,\phi_\varepsilon(\tau))-F_\varepsilon(\tau,\bar\phi(\tau)))\mathrm{d}\tau\Big|^2\\
&\quad+\mathbb{E}\Big|\int_{-\infty}^{t}(G_{\mathcal{A}_\varepsilon}(t,\tau)-G_{\bar{\mathcal{A}}}(t,\tau))F_\varepsilon(\tau,\bar\phi(\tau))\mathrm{d}\tau\Big|^2\\
&\quad+\mathbb{E}\Big|\int_{-\infty}^{t}G_{\bar{\mathcal{A}}}(t,\tau)[F_\varepsilon(\tau,\bar\phi(\tau))-\bar F(\bar\phi(\tau))]\mathrm{d}\tau\Big|^2\Big),
\end{aligned}
$$

using Cauchy-Schwartz inequality we get

$$
\begin{aligned}
I_1(t,\varepsilon) &\le 3\Big(\frac{\mathcal{N}^2L^2}{\nu}\int_{-\infty}^{t}e^{-\nu(t-\tau)}\mathbb{E}|\phi_\varepsilon(\tau)-\bar\phi(\tau)|^2\mathrm{d}\tau \qquad (35)\\
&\quad+\frac{2\mathcal{N}(\varepsilon)^2}{\gamma_0}\int_{-\infty}^{t}e^{-\gamma_0(t-\tau)}(M^2+L^2\|\bar\phi\|^2)\mathrm{d}\tau\\
&\quad+\mathbb{E}\Big|\int_{-\infty}^{t}G_{\bar{\mathcal{A}}}(t,\tau)[F_\varepsilon(\tau,\bar\phi(\tau))-\bar F(\bar\phi(\tau))]d\tau\Big|^2\Big)\\
&\le 3\Big(\frac{\mathcal{N}^2L^2}{\nu^2}\sup_{t\in\mathbb{R}}\mathbb{E}\left|\phi_\varepsilon(t)-\bar\phi(t)\right|^2+\frac{2\mathcal{N}(\varepsilon)^2}{\gamma_0^2}(M^2+L^2\|\bar\phi\|^2)\\
&\quad+\mathbb{E}\Big|\int_{-\infty}^{t}G_{\bar{\mathcal{A}}}(t,\tau)[F_\varepsilon(\tau,\bar\phi(\tau))-\bar F(\bar\phi(\tau))]d\tau\Big|^2\Big).
\end{aligned}
$$

We will show that

$$
\lim_{\varepsilon\to 0}\sup_{t\in\mathbb{R}}\mathbb{E}\Big|\int_{-\infty}^{t}G_{\bar{\mathcal{A}}}(t,\tau)[F_\varepsilon(\tau,\bar\phi(\tau))-\bar F(\bar\phi(\tau))]d\tau\Big|^2=0.
$$

To this end, making the change of variable $s=\tau-t$ and integrating by parts, we obtain for any $l>0$

$$
\begin{aligned}
&\mathbb{E}\Big|\int_{-\infty}^{t}G_{\bar{\mathcal{A}}}(t,\tau)[F_\varepsilon(\tau,\bar\phi(\tau))-\bar F(\bar\phi(\tau))]\mathrm{d}\tau\Big|^2 \qquad (36)\\
&=\mathbb{E}\Big|\int_{-\infty}^{0}G_{\bar{\mathcal{A}}}(t,t+s)[F_\varepsilon(t+s,\bar\phi(t+s))-\bar F(\bar\phi(t+s))]\mathrm{d}s\Big|^2\\
&=\mathbb{E}\Big|\int_{-\infty}^{0}G_{\bar{\mathcal{A}}}(t,t+s)\frac{\mathrm{d}}{\mathrm{d}s}\left(\int_{t}^{t+s}[F_\varepsilon(\sigma,\bar\phi(\sigma))-\bar F(\bar\phi(\sigma))]\mathrm{d}\sigma\right)\mathrm{d}s\Big|^2\\
&\le 2\mathbb{E}\Big|-\int_{-\infty}^{0}\frac{\partial G_{\bar{\mathcal{A}}}(t,t+s)}{\partial s}\left(\int_{t}^{t+s}[F_\varepsilon(\sigma,\bar\phi(\sigma))-\bar F(\bar\phi(\sigma))]\mathrm{d}\sigma\right)\mathrm{d}s\Big|^2\\
&\le 4\mathbb{E}\Big|-\int_{-\infty}^{-l}\frac{\partial G_{\bar{\mathcal{A}}}(t,t+s)}{\partial s}\left(\int_{t}^{t+s}[F_\varepsilon(\sigma,\bar\phi(\sigma))-\bar F(\bar\phi(\sigma))]\mathrm{d}\sigma\right)\mathrm{d}s\Big|^2\\
&\quad+4\mathbb{E}\Big|-\int_{-l}^{0}\frac{\partial G_{\bar{\mathcal{A}}}(t,t+s)}{\partial s}\left(\int_{t}^{t+s}[F_\varepsilon(\sigma,\bar\phi(\sigma))-\bar F(\bar\phi(\sigma))]\mathrm{d}\sigma\right)\mathrm{d}s\Big|^2\\
&\le 4\mathbb{E}\left(\int_{-\infty}^{-l}\Big|\int_{t}^{t+s}[F_\varepsilon(\sigma,\bar\phi(\sigma))-\bar F(\bar\phi(\sigma))]\mathrm{d}\sigma\Big|\mathcal{N}\|\bar{\mathcal{A}}\|e^{\nu s}\mathrm{d}s\right)^2\\
&\quad+4\mathbb{E}\sup_{-l\le s\le 0}\Big|\int_{t}^{t+s}\left[F_\varepsilon(\sigma,\bar\phi(\sigma))-\bar F(\bar\phi(\sigma))\right]\mathrm{d}\sigma\Big|^2\Big|\int_{-l}^{0}\mathcal{N}\|\bar{\mathcal{A}}\|e^{\nu s}\mathrm{d}s\Big|^2=:J_1+J_2.
\end{aligned}
$$

For J_1, we have

$$J_1 := 4\mathbb{E}\left(\int_{-\infty}^{-l}\left|\int_t^{t+s}[F_\varepsilon(\sigma,\bar{\phi}(\sigma)) - \bar{F}(\bar{\phi}(\sigma))]\mathrm{d}\sigma\right|\mathcal{N}\|\bar{A}\|e^{\nu s}\mathrm{d}s\right)^2 \tag{37}$$
$$\leq 4\mathcal{N}^2\|\bar{A}\|^2\int_{-\infty}^{-l}e^{\nu s}\mathrm{d}s\int_{-\infty}^{-l}\mathbb{E}\left|\int_t^{t+s}\left[F_\varepsilon(\sigma,\bar{\phi}(\sigma)) - \bar{F}(\bar{\phi}(\sigma))\right]\mathrm{d}\sigma\right|^2 e^{\nu s}\mathrm{d}s$$
$$\leq \frac{4\mathcal{N}^2\|\bar{A}\|^2}{\nu}e^{-\nu l}\int_{-\infty}^{-l}\mathbb{E}\left|\int_t^{t+s}\left[F_\varepsilon(\sigma,\bar{\phi}(\sigma)) - \bar{F}(\bar{\phi}(\sigma))\right]\mathrm{d}\sigma\right|^2 e^{\nu s}\mathrm{d}s$$
$$\leq \frac{4\mathcal{N}^2\|\bar{A}\|^2}{\nu}e^{-\nu l}\int_{-\infty}^{-l}s\int_t^{t+s}\mathbb{E}\left[F_\varepsilon(\sigma,\bar{\phi}(\sigma)) - \bar{F}(\bar{\phi}(\sigma))\right]^2\mathrm{d}\sigma e^{\nu s}\mathrm{d}s$$
$$\leq \frac{4\mathcal{N}^2\|\bar{A}\|^2}{\nu}e^{-\nu l}\int_{-\infty}^{-l}s\int_t^{t+s}8\left(M^2 + L^2\|\bar{\phi}\|_\infty^2\right)\mathrm{d}\sigma e^{\nu s}\mathrm{d}s$$
$$\leq \frac{32\mathcal{N}^2\|\bar{A}\|^2}{\nu}\left(M^2 + L^2\|\bar{\phi}\|_\infty^2\right)e^{-\nu l}\int_{-\infty}^{-l}s^2e^{\nu s}\mathrm{d}s$$
$$\leq \frac{32\mathcal{N}^2\|\bar{A}\|^2}{\nu}\left(M^2 + L^2\|\bar{\phi}\|_\infty^2\right)\left(\frac{l^2}{\nu} + \frac{2l}{\nu^2} + \frac{2}{\nu^3}\right)e^{-2\nu l}.$$

Divide $[0, l]$ into intervals of size δ, where $\delta > 0$ is a fixed constant depending on ε. Denote an adapted process $\tilde{\phi}$ such that $\tilde{\phi}(\sigma) = \bar{\phi}(t - k\delta)$ for any $\sigma \in (t - (k + 1)\delta, t - k\delta]$. By Lemma 4.1, we have

$$\mathbb{E}\sup_{-l\leq s\leq 0}\left|\int_t^{t+s}\left[F_\varepsilon(\sigma,\bar{\phi}(\sigma)) - \bar{F}(\bar{\phi}(\sigma))\right]\mathrm{d}\sigma\right|^2$$
$$= \mathbb{E}\sup_{-l\leq s\leq 0}\left|\int_t^{t+s}\left[F_\varepsilon(\sigma,\bar{\phi}(\sigma)) - F_\varepsilon(\sigma,\tilde{\phi}(\sigma)) + F_\varepsilon(\sigma,\tilde{\phi}(\sigma)) - \bar{F}(\tilde{\phi}(\sigma))\right.\right.$$
$$\left.\left. + \bar{F}(\tilde{\phi}(\sigma)) - \bar{F}(\bar{\phi}(\sigma))\right]\mathrm{d}\sigma\right|^2$$
$$\leq 6\mathbb{E}\sup_{-l\leq s\leq 0}\left|\int_t^{t+s}L|\bar{\phi}(\sigma) - \tilde{\phi}(\sigma)|\mathrm{d}\sigma\right|^2$$
$$+ 3\mathbb{E}\sup_{-l\leq s\leq 0}\left|\int_t^{t+s}\left[F_\varepsilon(\sigma,\tilde{\phi}(\sigma)) - \bar{F}(\tilde{\phi}(\sigma))\right]\mathrm{d}\sigma\right|^2$$
$$\leq 6\mathbb{E}\sup_{-l\leq s\leq 0}l\int_{t+s}^{t}L^2|\bar{\phi}(\sigma) - \tilde{\phi}(\sigma)|^2\mathrm{d}\sigma$$
$$+ 3\mathbb{E}\sup_{-l\leq s\leq 0}\left|\int_t^{t+s}\left[F_\varepsilon(\sigma,\tilde{\phi}(\sigma)) - \bar{F}(\tilde{\phi}(\sigma))\right]\mathrm{d}\sigma\right|^2$$
$$\leq 6L^2l^2C\delta + 3\mathbb{E}\sup_{-l\leq s\leq 0}\left|\int_t^{t+s}\left[F_\varepsilon(\sigma,\tilde{\phi}(\sigma)) - \bar{F}(\tilde{\phi}(\sigma))\right]\mathrm{d}\sigma\right|^2 =: 6L^2l^2C\delta + i_2.$$

For i_2, denote $s(\delta) := \left[\frac{|s|}{\delta}\right]$, we have

$$i_2 := 3\mathbb{E}\sup_{-l\leq s\leq 0}\left|\int_t^{t+s}\left(F_\varepsilon(\tau,\tilde{\phi}(\tau)) - \bar{F}(\tilde{\phi}(\tau))\right)\mathrm{d}\tau\right|^2 \tag{38}$$

$$
\begin{aligned}
&= 3\mathbb{E} \sup_{-l\le s\le 0} \Big| \sum_{k=0}^{s(\delta)-1} \int_{t-k\delta}^{t-(k+1)\delta} \left(F_\varepsilon(\tau,\bar\phi(t-k\delta)) - \bar F(\bar\phi(t-k\delta))\right) \mathrm{d}\tau \\
&\qquad + \int_{t-s(\delta)\cdot\delta}^{t+s} \left(F_\varepsilon(\tau,\bar\phi(t-s(\delta)\cdot\delta)) - \bar F(\bar\phi(t-s(\delta)\cdot\delta))\right) \mathrm{d}\tau \Big|^2 \\
&\le 6\left[\frac{l}{\delta}\right] \mathbb{E} \sup_{-l\le s\le 0} \sum_{k=0}^{s(\delta)-1} \left| \int_{t-k\delta}^{t-(k+1)\delta} \left(F_\varepsilon(\tau,\bar\phi(t-k\delta)) - \bar F(\bar\phi(t-k\delta))\right) \mathrm{d}\tau \right|^2 \\
&\quad + 6\mathbb{E} \sup_{-l\le s\le 0} \left| \int_{t-s(\delta)\cdot\delta}^{t+s} \left(F_\varepsilon(\tau,\bar\phi(t-s(\delta)\cdot\delta)) - \bar F(\bar\phi(t-s(\delta)\cdot\delta))\right) \mathrm{d}\tau \right|^2 \\
&=: i_2^1 + i_2^2.
\end{aligned}
$$

For i_2^1, by Lemma 4.1 we have

$$
\begin{aligned}
i_2^1 &:= 6\left[\frac{l}{\delta}\right] \mathbb{E} \sup_{-l\le s\le 0} \sum_{k=0}^{s(\delta)-1} \left| \int_{t-k\delta}^{t-(k+1)\delta} \left(F_\varepsilon(\tau,\bar\phi(t-k\delta)) - \bar F(\bar\phi(t-k\delta))\right) \mathrm{d}\tau \right|^2 \qquad (39)\\
&\le \frac{6l^2}{\delta^2} \mathbb{E} \sup_{-l\le s\le 0} \max_{0\le k\le s(\delta)-1} \left| \int_{\frac{t-k\delta}{\varepsilon}}^{\frac{t-(k+1)\delta}{\varepsilon}} \left(F(\tau,\bar\phi(t-k\delta)) - \bar F(\bar\phi(t-k\delta))\right) \varepsilon \mathrm{d}\tau \right|^2 \\
&\le \frac{12l^2}{\delta^2} \mathbb{E} \sup_{-l\le s\le 0} \max_{0\le k\le s(\delta)-1} \delta^2 \omega_1^2\left(\frac{\delta}{\varepsilon}\right) \left(1 + |\bar\phi(t-k\delta)|^2\right) \\
&\le 12 l^2 \left(C + Cl^2 + 1\right) \omega_1^2\left(\frac{\delta}{\varepsilon}\right).
\end{aligned}
$$

For i_2^2, we obtain

$$
\begin{aligned}
i_2^2 &:= 6\mathbb{E} \sup_{-l\le s\le 0} \left| \int_{t-s(\delta)\cdot\delta}^{t+s} \left(F_\varepsilon(\tau,\bar\phi(t-s(\delta)\cdot\delta)) - \bar F(\bar\phi(t-s(\delta)\cdot\delta))\right) \mathrm{d}\tau \right|^2 \qquad (40)\\
&\le 6\mathbb{E} \sup_{-l\le s\le 0} \delta \int_{t+s}^{t-s(\delta)\cdot\delta} \left(F_\varepsilon(\tau,\bar\phi(t-s(\delta)\cdot\delta)) - \bar F(\bar\phi(t-s(\delta)\cdot\delta))\right)^2 \mathrm{d}\tau \\
&\le 6\delta\mathbb{E} \sup_{-l\le s\le 0} \int_{t+s}^{t-s(\delta)\cdot\delta} 8\left(M^2 + L^2|\bar\phi(t-s(\delta)\cdot\delta)|^2\right) \mathrm{d}\tau \\
&\le 6\delta \int_{t-l}^{t} 8\left(M^2 + L^2\mathbb{E} \sup_{\sigma\in[t-l,t]} \|\bar\phi(\sigma)\|^2\right) \mathrm{d}\tau \\
&\le 48\left(M^2 + L^2C(l^2+1)\right) l\delta.
\end{aligned}
$$

Therefore, (38)–(40) imply

$$
i_2 \le 12\left(Cl^4 + (C+1)l^2\right)\omega_1^2\left(\frac{\delta}{\varepsilon}\right) + 48\left(M^2 + L^2C(l^2+1)\right) l\delta.
$$

Therefore, we have

$$
\begin{aligned}
J_2 \le \frac{4\mathcal{N}^2\|\bar{\mathcal{A}}\|^2}{\nu^2}\left(1 - e^{-\nu l}\right)^2 \Big[& 6L^2l^2C\delta + 12\left(Cl^4 + (C+1)l^2\right)\omega_1^2\left(\frac{\delta}{\varepsilon}\right) \qquad (41)\\
& + 48\left(M^2 + L^2C(l^2+1)\right) l\delta \Big].
\end{aligned}
$$

Combing (36), (37) and (41), we have

$$
\begin{aligned}
&\mathbb{E}\left|\int_{-\infty}^{t} G_{\bar{A}}(t,\tau)\left[F_\varepsilon(\tau,\bar{\phi}(\tau)) - \bar{F}(\bar{\phi}(\tau))\right]\mathrm{d}\tau\right|^2 \\
&\le \frac{32\mathcal{N}^2\|\bar{A}\|^2}{\nu}\left(M^2 + L^2\|\bar{\phi}\|_\infty^2\right)\left(\frac{l^2}{\nu} + \frac{2l}{\nu^2} + \frac{2}{\nu^3}\right)e^{-2\nu l} \\
&\quad + \frac{4\mathcal{N}^2\|\bar{A}\|^2}{\nu^2}\left(1 - e^{-\nu l}\right)^2\Big[6L^2 l^2 C\delta + 12\left(Cl^4 + (C+1)l^2\right)\omega_1^2\left(\frac{\delta}{\varepsilon}\right) \\
&\quad\quad + 48\left(M^2 + L^2 C(l^2+1)\right)l\delta\Big].
\end{aligned}
\tag{42}
$$

Taking $\delta = \sqrt{\varepsilon}$ and letting $\varepsilon \to 0$ in (42), we have

$$
\begin{aligned}
&\limsup_{\varepsilon\to 0}\sup_{t\in\mathbb{R}}\mathbb{E}\left|\int_{-\infty}^{t} G_{\bar{A}}(t,\tau)\left[F_\varepsilon(\tau,\bar{\phi}(\tau)) - \bar{F}(\bar{\phi}(\tau))\right]\mathrm{d}\tau\right|^2 \\
&\le \frac{32\mathcal{N}^2\|\bar{A}\|^2}{\nu}\left(M^2 + L^2\|\bar{\phi}\|_\infty^2\right)\left(\frac{l^2}{\nu} + \frac{2l}{\nu^2} + \frac{2}{\nu^3}\right)e^{-2\nu l}.
\end{aligned}
$$

Since l is arbitrary, by letting $l \to \infty$ we get

$$
\lim_{\varepsilon\to 0}\sup_{t\in\mathbb{R}}\mathbb{E}\left|\int_{-\infty}^{t} G_{\bar{A}}(t,\tau)\left[F_\varepsilon(\tau,\bar{\phi}(\tau)) - \bar{F}(\bar{\phi}(\tau))\right]\mathrm{d}\tau\right|^2 = 0. \tag{43}
$$

From (35) and (43) it follows that there exists a function $A : (0,\varepsilon_0) \to \mathbb{R}_+$ so that $A(\varepsilon) \to 0$ as $\varepsilon \to 0$ and

$$
I_1(t,\varepsilon) \le 3\frac{\mathcal{N}^2 L^2}{\nu^2}\sup_{t\in\mathbb{R}}\mathbb{E}|\phi_\varepsilon(t) - \bar{\phi}(t)|^2 + A(\varepsilon) \tag{44}
$$

for any $t \in \mathbb{R}$ and $\varepsilon \in (0,\varepsilon_0)$.

Now we will establish a similar estimation for $I_2(t,\varepsilon)$. Since

$$
\begin{aligned}
I_2(t,\varepsilon) &:= \mathbb{E}\left|\int_{-\infty}^{t}(G_{\mathcal{A}_\varepsilon}(t,\tau)G_\varepsilon(\tau,\phi_\varepsilon(\tau)) - G_{\bar{A}}(t,\tau)\bar{G}(\bar{\phi}(\tau)))\mathrm{d}W(\tau)\right|^2 \\
&\le 3\Bigg(\mathbb{E}\left|\int_{-\infty}^{t} G_{\mathcal{A}_\varepsilon}(t,\tau)\left(G_\varepsilon(\tau,\phi_\varepsilon(\tau)) - G_\varepsilon(\tau,\bar{\phi}(\tau))\right)\mathrm{d}W(\tau)\right|^2 \\
&\quad + \mathbb{E}\left|\int_{-\infty}^{t}\left(G_{\mathcal{A}_\varepsilon}(t,\tau) - G_{\bar{A}}(t,\tau)\right)G_\varepsilon(\tau,\bar{\phi}(\tau))\mathrm{d}W(\tau)\right|^2 \\
&\quad + \mathbb{E}\left|\int_{-\infty}^{t} G_{\bar{A}}(t,\tau)[G_\varepsilon(\tau,\bar{\phi}(\tau)) - \bar{G}(\bar{\phi}(\tau))]\mathrm{d}W(\tau)\right|^2\Bigg),
\end{aligned}
$$

using Itô's isometry property we have

$$
\begin{aligned}
I_2(t,\varepsilon) &\le 3\Bigg(\mathcal{N}^2 L^2\int_{-\infty}^{t} e^{-2\nu(t-\tau)}\mathbb{E}|\phi_\varepsilon(\tau) - \bar{\phi}(\tau)|^2\mathrm{d}\tau \\
&\quad + 2\mathcal{N}(\varepsilon)^2\int_{-\infty}^{t} e^{-2\gamma_0(t-\tau)}(M^2 + L^2\|\bar{\phi}\|_\infty^2)\mathrm{d}\tau \\
&\quad + \mathcal{N}^2\int_{-\infty}^{t} e^{-2\nu(t-\tau)}\mathbb{E}|G_\varepsilon(\tau,\bar{\phi}(\tau)) - \bar{G}(\bar{\phi}(\tau))|^2\mathrm{d}\tau\Bigg) \\
&\le 3\Bigg(\frac{\mathcal{N}^2 L^2}{2\nu}\sup_{t\in\mathbb{R}}\mathbb{E}|\phi_\varepsilon(t) - \bar{\phi}(\tau)|^2 + \frac{\mathcal{N}(\varepsilon)^2}{\gamma_0}(M^2 + L^2\|\bar{\phi}\|_\infty^2)
\end{aligned}
\tag{45}
$$

$$+\mathcal{N}^2\int_{-\infty}^{t}e^{-2\nu(t-\tau)}\mathbb{E}|G_\varepsilon(\tau,\bar{\phi}(\tau))-\bar{G}(\bar{\phi}(\tau))|^2d\tau\Big).$$

Now we prove that

$$\lim_{\varepsilon\to 0}\sup_{t\in\mathbb{R}}\left|\int_{-\infty}^{t}e^{-2\nu(t-\tau)}\mathbb{E}\left|G_\varepsilon(\tau,\bar{\phi}(\tau))-\bar{G}(\bar{\phi}(\tau))\right|^2d\tau\right|=0.$$

By Lemma 3.4, it suffices to show that

$$\lim_{\varepsilon\to 0}\sup_{|s|\le l,\ t\in\mathbb{R}}\left|\int_{t}^{t+s}\mathbb{E}\left|G_\varepsilon(\tau,\bar{\phi}(\tau))-\bar{G}(\bar{\phi}(\tau))\right|^2d\tau\right|=0.$$

To this end, define an adapted process $\hat{\phi}$ such that $\hat{\phi}(\sigma)=\bar{\phi}(t+k\delta)$ for any $\sigma\in[t+k\delta,t+(k+1)\delta)$. We can assume $s>0$ without loss of generality, then we have by Lemma 4.1

$$\begin{aligned}
&\int_{t}^{t+s}\mathbb{E}|G_\varepsilon(\tau,\bar{\phi}(\tau))-\bar{G}(\bar{\phi}(\tau))|^2d\tau\\
&\le\int_{t}^{t+s}\mathbb{E}\Big|G_\varepsilon(\tau,\bar{\phi}(\tau))-G_\varepsilon(\tau,\hat{\phi}(\tau))+G_\varepsilon(\tau,\hat{\phi}(\tau))-\bar{G}(\hat{\phi}(\tau))\\
&\qquad+\bar{G}(\hat{\phi}(\tau))-\bar{G}(\bar{\phi}(\tau))\Big|^2\mathrm{d}\tau\\
&\le 3\int_{t}^{t+s}\mathbb{E}\left|G_\varepsilon(\tau,\bar{\phi}(\tau))-G_\varepsilon(\tau,\hat{\phi}(\tau))\right|^2\mathrm{d}\tau+3\int_{t}^{t+s}\mathbb{E}\left|G_\varepsilon(\tau,\hat{\phi}(\tau))-\bar{G}(\hat{\phi}(\tau))\right|^2\mathrm{d}\tau\\
&\quad+3\int_{t}^{t+s}\mathbb{E}\left|\bar{G}(\hat{\phi}(\tau))-\bar{G}(\bar{\phi}(\tau))\right|^2\mathrm{d}\tau\\
&\le 6L^2lC\delta+3\int_{t}^{t+s}\mathbb{E}\left|G_\varepsilon(\tau,\hat{\phi}(\tau))-\bar{G}(\hat{\phi}(\tau))\right|^2\mathrm{d}\tau=:6L^2lC\delta+3J_3.
\end{aligned}$$

For J_3, we have

$$\begin{aligned}
J_3&:=\mathbb{E}\int_{t}^{t+s}\left|G_\varepsilon(\tau,\hat{\phi}(\tau))-\bar{G}(\hat{\phi}(\tau))\right|^2\mathrm{d}\tau\\
&\le\mathbb{E}\Big(\sum_{k=0}^{s(\delta)-1}\int_{t+k\delta}^{t+(k+1)\delta}\left|G_\varepsilon(\tau,\bar{\phi}(t+k\delta))-\bar{G}(\bar{\phi}(t+k\delta))\right|^2\mathrm{d}\tau\\
&\qquad+\int_{t+s(\delta)\cdot\delta}^{t+s}\left|G_\varepsilon(\tau,\bar{\phi}(t+s(\delta)\cdot\delta))-\bar{G}(\bar{\phi}(t+s(\delta)\cdot\delta))\right|^2\mathrm{d}\tau\Big)=:J_3^1+J_3^2.
\end{aligned}$$

Then

$$\begin{aligned}
J_3^1&:=\mathbb{E}\Big(\sum_{k=0}^{s(\delta)-1}\int_{t+k\delta}^{t+(k+1)\delta}\left|G_\varepsilon(\tau,\bar{\phi}(t+k\delta))-\bar{G}(\bar{\phi}(t+k\delta))\right|^2\mathrm{d}\tau\Big)\\
&\le\left[\frac{l}{\delta}\right]\max_{0\le k\le s(\delta)-1}\mathbb{E}\int_{t+k\delta}^{t+(k+1)\delta}\left|G_\varepsilon(\tau,\bar{\phi}(t+k\delta))-\bar{G}(\bar{\phi}(t+k\delta))\right|^2\mathrm{d}\tau\\
&=\left[\frac{l}{\delta}\right]\max_{0\le k\le s(\delta)-1}\mathbb{E}\int_{\frac{t+k\delta}{\varepsilon}}^{\frac{t+(k+1)\delta}{\varepsilon}}\left|G(\tau,\bar{\phi}(t+k\delta))-\bar{G}(\bar{\phi}(t+k\delta))\right|^2\varepsilon\mathrm{d}\tau\\
&\le l\omega_2\left(\frac{\delta}{\varepsilon}\right)\left(1+\|\bar{\phi}\|_\infty^2\right)
\end{aligned}$$

and

$$
\begin{aligned}
J_3^2 &:= \mathbb{E}\int_{t+s(\delta)\cdot\delta}^{t+s} \left|G_\varepsilon(\tau,\bar{\phi}(t+s(\delta)\cdot\delta)) - \bar{G}(\bar{\phi}(t+s(\delta)\cdot\delta))\right|^2 \mathrm{d}\tau \\
&\le 8\left(M^2 + L^2\|\bar{\phi}\|_\infty^2\right)\delta.
\end{aligned}
$$

Therefore we have

$$
\begin{aligned}
&\sup_{|s|\le l,\ t\in\mathbb{R}} \left|\int_t^{t+s} \mathbb{E}\left|G_\varepsilon(\tau,\bar{\phi}(\tau)) - \bar{G}(\bar{\phi}(\tau))\right|^2 \mathrm{d}\tau\right| \qquad (46)\\
&\le 6L^2 lC\delta + 24(M^2 + L^2\|\bar{\phi}\|_\infty^2)\delta + 3l\omega_2\left(\frac{\delta}{\varepsilon}\right)(1+\|\bar{\phi}\|_\infty^2).
\end{aligned}
$$

Taking $\delta = \sqrt{\varepsilon}$ and letting $\varepsilon \to 0$ in (46), we have

$$
\lim_{\varepsilon\to 0}\sup_{|s|\le l,\ t\in\mathbb{R}} \left|\int_t^{t+s} \mathbb{E}\left|G_\varepsilon(\tau,\bar{\phi}(\tau)) - \bar{G}(\bar{\phi}(\tau))\right|^2 \mathrm{d}\tau\right| = 0. \qquad (47)
$$

From (45) and (47) it follows that

$$
I_2(t,\varepsilon) \le 3(\mathcal{N}L)^2 \frac{1}{2\nu}\sup_{t\in\mathbb{R}} \mathbb{E}|\phi_\varepsilon(t) - \bar{\phi}(t)|^2 + B(\varepsilon), \qquad (48)
$$

where $B(\varepsilon)$ is some positive constant such that $B(\varepsilon) \to 0$ as $\varepsilon \to 0$.

Combing (34), (44) and (48), we have

$$
\left(1 - 3(\mathcal{N}L)^2\left(\frac{2}{\nu^2} + \frac{1}{\nu}\right)\right)\sup_{t\in\mathbb{R}} \mathbb{E}|\phi_\varepsilon(t) - \bar{\phi}(t)|^2 \le 2\left(A(\varepsilon) + B(\varepsilon)\right).
$$

Consequently

$$
\lim_{\varepsilon\to 0}\sup_{t\in\mathbb{R}} \mathbb{E}|\phi_\varepsilon(t) - \bar{\phi}(t)|^2 = 0
$$

because $1 - 3(\mathcal{N}L)^2\left(\frac{2}{\nu^2} + \frac{1}{\nu}\right) > 0$.

To finish the proof of the theorem we note that L^2-convergence implies convergence in distribution, so

$$
\lim_{\varepsilon\to 0}\sup_{t\in\mathbb{R}} \beta(\mathcal{L}(\phi_\varepsilon(t)), \mathcal{L}(\bar{\phi}(t))) = 0.
$$

Since $\mathcal{L}(\varphi_\varepsilon(\frac{t}{\varepsilon})) = \mathcal{L}(\phi_\varepsilon(t))$, we get

$$
\lim_{\varepsilon\to 0}\sup_{t\in\mathbb{R}} \beta(\mathcal{L}(\varphi_\varepsilon(\frac{t}{\varepsilon})), \mathcal{L}(\bar{\phi}(t))) = 0.
$$

The proof is complete. □

Corollary 3. *Under the conditions of Theorem 4.2 the following statements hold:*

1. *if the functions $\mathcal{A} \in C(\mathbb{R}, L(\mathcal{H}))$ and $F, G \in C(\mathbb{R}\times\mathcal{H}, \mathcal{H})$ are jointly stationary (respectively, τ-periodic, quasi-periodic with the spectrum of frequencies $\nu_1, \ldots, \nu_k$, Bohr almost periodic, almost automorphic, Birkhoff recurrent, Lagrange stable, Levitan almost periodic, almost recurrent, Poisson stable), then equation (30) has a unique solution $\varphi_\varepsilon \in C_b(\mathbb{R}, L^2(\mathbb{P}, \mathcal{H}))$ which is stationary (respectively, τ-periodic, quasi-periodic with the spectrum of frequencies $\nu_1, \ldots, \nu_k$, Bohr almost periodic, almost automorphic, Birkhoff recurrent, Lagrange stable, Levitan almost periodic, almost recurrent, Poisson stable) in distribution;*

2. *if the functions $\mathcal{A} \in C(\mathbb{R}, L(\mathcal{H}))$ and $F, G \in C(\mathbb{R} \times \mathcal{H}, \mathcal{H})$ are Lagrange stable and jointly pseudo-periodic (respectively, pseudo-recurrent), then equation (30) has a unique solution $\varphi_\varepsilon \in C_b(\mathbb{R}, L^2(\mathbb{P}, \mathcal{H}))$ which is pseudo-periodic (respectively, pseudo-recurrent) in distribution;*
3. $$\lim_{\varepsilon \to 0} \sup_{t \in \mathbb{R}} \beta(\mathcal{L}(\varphi_\varepsilon(\frac{t}{\varepsilon}), \mathcal{L}(\bar{\phi}(t))) = 0,$$
   with $\bar{\phi}$ being the unique stationary solution of the averaged equation (33).

Proof. This statement follows from Theorems 2.12 and 4.2 (see also Remark 2). □

Remark 15. To simplify the notations and highlight the idea, we consider only the one-dimensional noise. Indeed, the main results of this paper remain hold if we replace the one-dimensional Brownian motion W in Sections 3 and 4 by a Q-Wiener process, which brings no essential but just notational difference; see e.g. [20, 21] for details.

Remark 16. In the present paper, we only consider the second Bogolyubov theorem for semilinear stochastic *ordinary* differential equations, i.e. the linear part $\mathcal{A}(\cdot)$ is bounded operator valued. We will consider the case when $\mathcal{A}(\cdot)$ is an unbounded operator in future work, which can be applied to related stochastic partial differential equations.

Acknowledgments. We would like to thank Professor A. J. Roberts for drawing our attention to stochastic centre manifold approach for averaging and to references [3, 22]. We are grateful to the anonymous referees for their careful reading of our paper and valuable suggestions.

REFERENCES

[1] N. N. Bogolyubov, *On Some Statistical Methods in Mathematical Physics*, Akademiya Nauk Ukrainskoĭ SSR, 1945. (in Russian)

[2] N. N. Bogolyubov and Y. A. Mitropolsky, *Asymptotic Methods in the Theory of Non-Linear Oscillations*, Translated from the second revised Russian edition. International Monographs on Advanced Mathematics and Physics Hindustan Publishing Corp., Delhi, Gordon and Breach Science Publishers, New York, 1961.

[3] P. Boxler, A stochastic version of center manifold theory, *Probab. Theory Related Fields*, **83** (1989), 509–545.

[4] S. Cerrai and M. Freidlin, Averaging principle for a class of stochastic reaction-diffusion equations, *Probab. Theory Related Fields*, **144** (2009), 137–177.

[5] S. Cerrai and A. Lunardi, Averaging principle for nonautonomous slow-fast systems of stochastic reaction-diffusion equations: the almost periodic case, *SIAM J. Math. Anal.*, **49** (2017), 2843–2884.

[6] D. N. Cheban, *Asymptotically Almost Periodic Solutions of Differential Equations*, Hindawi Publishing Corporation, New York, 2009.

[7] D. N. Cheban, *Global Attractors of Nonautonomous Dynamical and Control Systems*, 2^{nd} edition, Interdisciplinary Mathematical Sciences, vol.18, River Edge, NJ: World Scientific, 2015.

[8] D. Cheban and J. Duan, Recurrent motions and global attractors of non-autonomous Lorenz systems, *Dyn. Syst.*, **19** (2004), 41–59.

[9] D. Cheban and Z. Liu, Periodic, quasi-periodic, almost periodic, almost automorphic, Birkhoff recurrent and Poisson stable solutions for stochastic differential equations, *J. Differential Equations*, **269** (2020), 3652–3685.

[10] Ju. L. Dalec'kiĭ and M. G. Kreĭn, *Stability of Solutions of Differential Equations in Banach Space*, Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 43. American Mathematical Society, Providence, R.I., 1974.

[11] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, 2^{nd} edition, Encyclopedia of Mathematics and its Applications, 152. Cambridge University Press, Cambridge, 2014.

[12] J. Duan and W. Wang, *Effective Dynamics of Stochastic Partial Differential Equations*, Elsevier Insights. Elsevier, Amsterdam, 2014.

[13] R. M. Dudley, *Real Analysis and Probability*, Revised reprint of the 1989 original. Cambridge Studies in Advanced Mathematics, 74. Cambridge University Press, Cambridge, 2002.

[14] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, Translated from the 1979 Russian original by Joseph Szücs. 3^{rd} edition. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 260. Springer, Heidelberg, 2012.

[15] R. Z. Has'minskiĭ, On the principle of averaging the itô's stochastic differential equations, *Kybernetika (Prague)*, **4** (1968), 260–279. (in Russian)

[16] M. Kamenskii, O. Mellah and P. Raynaud de Fitte, Weak averaging of semilinear stochastic differential equations with almost periodic coefficients, *J. Math. Anal. Appl.*, **427** (2015), 336–364.

[17] M. A. Krasnosel'skiĭ, V. Sh Burd and Yu. S. Kolesov, *Nonlinear Almost Periodic Oscillations*, Nauka, Moscow, 1970 (in Russian). [English translation: Nonlinear Almost Periodic Oscillations. A Halsted Press Book. New York etc.: John Wiley & Sons; Jerusalem- London: Israel Program for Scientific Translations. IX, 326 p., 1973]

[18] N. Krylov and N. Bogolyubov, *Introduction to Non-Linear Mechanics*, Annals of Mathematics Studies, no. 11. Princeton University Press, Princeton, N. J., 1943.

[19] B. M. Levitan and V. V. Zhikov, *Almost Periodic Functions and Differential Equations*, Moscow State University Press, Moscow, 1978 (in Russian). [English translation: Almost Periodic Functions and Differential Equations. Cambridge Univ. Press, Cambridge, 1982]

[20] X. Liu and Z. Liu, Poisson stable solutions for stochastic differential equations with Lévy noise, *Acta Math. Sin. (Engl. Ser.)*, In Press. (also available at `arXiv:2002.00395`)

[21] Z. Liu and K. Sun, Almost automorphic solutions for stochastic differential equations driven by Lévy noise, *J. Funct. Anal.*, **226** (2014), 1115–1149.

[22] A. J. Roberts, Normal form transforms separate slow and fast modes in stochastic dynamical systems, *Phys. A*, **387** (2008), 12–38.

[23] B. A. Ščerbakov, A certain class of Poisson stable solutions of differential equations, *Differencial'nye Uravnenija*, **4** (1968), 238–243. (in Russian)

[24] B. A. Ščerbakov, The comparability of the motions of dynamical systems with regard to the nature of their recurrence, *Differencial'nye Uravnenija*, **11** (1975), 1246–1255. (in Russian) [English translation: *Differential Equations* **11** (1975), 937–943].

[25] G. R. Sell, *Topological Dynamics and Ordinary Differential Equations*, Van Nostrand-Reinhold, 1971.

[26] B. A. Shcherbakov, *Topologic Dynamics and Poisson Stability of Solutions of Differential Equations*, Stiinca, Kishinev, 1972.

[27] B. A. Shcherbakov, *Poisson Stability of Motions of Dynamical Systems and Solutions of Differential Equations*, Ştiinţa, Chişinău, 1985. (in Russian)

[28] K. S. Sibirsky, *Introduction to Topological Dynamics*, Kishinev, RIA AN MSSR, 1970. (in Russian). [English translation: Introduction to Topological Dynamics. Noordhoff, Leyden, 1975]

[29] A. V. Skorokhod, *Asymptotic Methods in the Theory of Stochastic Differential Equations*, Translated from the Russian by H. H. McFaden. Translations of Mathematical Monographs, 78. American Mathematical Society, Providence, RI, 1989.

[30] A. Yu. Veretennikov, On large deviations in the averaging principle for SDEs with a "full dependence", *Ann. Probab.*, **27** (1999), 284–296.

[31] I. Vrkoč, Weak averaging of stochastic evolution equations, *Math. Bohem.*, **120** (1995), 91–111.

[32] W. Wang and A. J. Roberts, Average and deviation for slow-fast stochastic partial differential equations, *J. Differential Equations*, **253** (2012), 1265–1286.

Received August 2020; revised January 2021.

E-mail address: `cheban@usm.md`
E-mail address: `zxliu@dlut.edu.cn`