DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS SERIES B
Volume **24**, Number **11**, November **2019**

doi:10.3934/dcdsb.2019113

pp. **5927–5944**

ALMOST PERIODIC SOLUTIONS AND STABLE SOLUTIONS FOR STOCHASTIC DIFFERENTIAL EQUATIONS

Yong Li

School of Mathematics, Jilin University
Changchun 130012, China
Center for Mathematics and Interdisciplinary Sciences, and School of
Mathematics and Statistics, Northeast Normal University
Changchun 130024, China

Zhenxin Liu

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

Wenhe Wang*

School of Basic Sciences
Changchun University of Technology
Changchun 130012, China

(Communicated by Xiaoying Han)

Abstract. In this paper, we discuss the relationships between stability and almost periodicity for solutions of stochastic differential equations. Our essential idea is to get stability of solutions or systems by some inherited properties of Lyapunov functions. Under suitable conditions besides Lyapunov functions, we obtain the existence of almost periodic solutions in distribution.

1. **Introduction.** In 1924–1926, Bohr founded the theory of almost periodic functions [8, 9, 10]. Roughly speaking, an almost periodic function means that it is periodic up to any desired level of accuracy. Since many differential equations arising from physics and other fields admit almost periodic solutions, almost periodicity becomes an important property of dynamical systems and is extensively studied in the area of differential equations and dynamical systems. We refer the reader to the books, e.g. Amerio and Prouse [1], Fink [15], Levitan and Zhikov [23], Yoshizawa [37] etc, for an exposition.

For deterministic differential equations, the existence of almost periodic solutions was studied under various stability assumptions. Markov [27] defined a kind of stability which implies almost periodicity. Deysach and Sell [14] assumed that there

2010 *Mathematics Subject Classification.* Primary: 60H10, 34C27; Secondary: 37B25.

Key words and phrases. Stochastic differential equation, almost periodic solution, asymptotically almost periodic solution, stable in distribution, Lyapunov function.

The first author is supported by National Research Program of China Grant 2013CB834100, Jilin DRC Grant 2017C028-1, NSFC Grants 11571065 and 11171132; the second author is supported by NSFC Grants 11522104 and 11871132, and the startup and Xinghai Jieqing funds from Dalian University of Technology.

* Corresponding author.

exists one bounded uniformly stable solution. Miller [28] assumed the existence of one bounded totally stable solution. Seifert [31] proposed the so-called Σ-stability, while Sell [32, 33] proposed the stability under disturbance from the hull; actually, these two concepts of stability are equivalent. Coppel [12] sharpened Miller's result without the uniqueness of solutions by using the properties of asymptotically almost periodic functions; Yoshizawa [36] developed the idea of Coppel and improved all the results mentioned above. On the other hand, the Lyapunov's second method was employed to investigate the existence of almost periodic solutions: Hale [17] and Yoshizawa [35] assumed the existence of Lyapunov functions for pairs of solutions to conclude the uniform asymptotic stability in the large of the bounded solution.

However, the various stability assumptions mentioned above are not easily verified directly in practice. It is known that some stabilities, such as uniform stability and uniform asymptotic stability, can be characterized by Lyapunov functions. So it seems that it is a good idea to give some explicit conditions on Lyapunov functions to study the existence of almost periodic solutions, as Hale and Yoshizawa did in [17, 35]. This is exactly what we do in the present paper for stochastic differential equations (SDE).

The study of stability of solutions for SDEs or equations with random coefficients dates back to the late 1950s. Bertram and Sarachik [5] considered mean and mean square stability. Kac and Krasovskii's work [18] is "extremely suggestive" and "has stimulated considerable further research" in Khasminskii's words [20]. Kozin [21] investigated almost sure stability. Kushner [22] and Khasminskii [19, 20] considered stability in probability by Lyapunov function method. Khasminskii [20] and Basak and Bhattachaya [3, 4] investigated the asymptotic stability in the sense of distribution, i.e. the transition probability of the diffusion converges weakly to the stationary distribution as t goes to infinity. Mao [25, 26] developed almost sure and moment exponential stability. In this paper, we introduce a new kind of stability in distribution to study almost periodic solutions for SDEs.

For the stochastically perturbed semilinear equation, almost periodic solutions were studied by assuming that the linear part of the equation satisfies the property of exponential dichotomy; see Halanay [16], Morozan and Tudor [29], Da Prato and Tudor [13], and Arnold and Tudor [2], among others. For general SDEs, Vârsana [34] studied asymptotically almost periodic (weaker than almost periodic) solutions by assuming that the stochastic system is totally stable. Very recently, Liu and Wang [24] investigated the almost periodic solutions to SDEs by the separation method.

This paper is organized as follows. Section 2 is a preliminary section. Section 3 contains main results of this paper, in which we study almost periodic solutions for SDEs by mainly the Lyapunov function method. In Section 4, we illustrate our results by some applications.

2. Preliminaries.

Assume that (M, d) is a complete metric space. The definition of M-valued (uniformly) almost periodic functions in the sense of Bohr needs the concept of relative density:

Definition 2.1. A set $A \subset \mathbb{R}$ is called *relatively dense* in $\mathbb{R}$ if there exists a constant $l > 0$ such that $(a, a + l) \cap A \neq \emptyset$ for every $a \in \mathbb{R}$.

And the definition of almost periodic functions is as follows:

Definition 2.2. (i) Assume that the function $\varphi : \mathbb{R} \to M$ is continuous. If for any $\epsilon > 0$ there is a relatively dense set $T(\epsilon, \varphi)$ such that for any $\tau \in T(\epsilon, \varphi)$ we have

$$\sup_{t\in\mathbb{R}} d(\varphi(t+\tau), \varphi(t)) < \epsilon,$$

then we say that the function φ is *almost periodic*.

(ii) A continuous function $f : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ is *almost periodic in t uniformly on compact sets* if for any $\epsilon > 0$ and compact set $S \subset \mathbb{R}^d$ there exists a relatively dense set $T(\epsilon, f, S)$ such that for every $\tau \in T(\epsilon, f, S)$ we have:

$$\sup_{(t,x)\in\mathbb{R}\times S} |f(t+\tau, x) - f(t, x)| < \epsilon.$$

We also say such f is *uniformly almost periodic* for short.

Remark 1. (i) It is well-known that if φ is almost periodic, then the range of φ is pre-compact, i.e. its closure is compact.

(ii) Bochner [6, 7] gave an equivalent condition to Bohr's almost periodicity in Definition 2.2. The above definition of uniform almost periodicity can be found in Yoshizawa's book [37]; Seifert and Fink made another definition of uniform almost periodicity (see Definition 2.1 in [15]).

Notations. Denote $\mathbb{R}_+ := [0, +\infty)$ and $\mathbb{R}_- := (-\infty, 0]$. In what follows, we follow Bochner (see [7] or [15, pp. 2-3] for details) on the notations of sequences and translation limits which are widely used by people. The sequence $\{\alpha_n\}$ is denoted by α. If $\beta = \{\beta_n\}$, then $\beta \subset \alpha$ means that β is a subsequence of α and $\alpha+\beta$ means the sequence $\{\alpha_n + \beta_n\}$. For a sequence $\alpha = \{\alpha_n\}$ and a function φ of t, we denote the translation operator by T_α: the notation $T_\alpha\varphi = \psi$ means that

$$\psi(t) = \lim_{n\to+\infty} \varphi(t + \alpha_n)$$

and is written only when the limit on the right-hand side exists. The mode of convergence, e.g. pointwise, uniformly on an interval, on $\mathbb{R}$ or on $\mathbb{R}_+$ etc., will be specified at each use of the notation, so this will not (hopefully) confuse the reader. Similarly, for a function f of (t, x) we denote by $T_\alpha f = g$ to mean that $g(t, x) = \lim_{n\to+\infty} f(t+\alpha_n, x)$ and only write so when the limit exists; the convergence mode will be pointed out explicitly when this notation is used.

For a given uniformly almost periodic function f, we denote by

$$\begin{aligned} H(f) := \{ g :\ & \text{there is a sequence } \alpha \text{ such that } T_\alpha f = g \\ & \text{uniformly on } \mathbb{R}\times S \text{ for each compact set } S \subset \mathbb{R}^d \ \} \end{aligned}$$

the *hull of f*. The hull has the following properties:

Proposition 1. *Let f be uniformly almost periodic. Then:*

(i) *any $g \in H(f)$ is also uniformly almost periodic and $H(g) = H(f)$;*
(ii) *for any $g \in H(f)$, there exists a sequence α with $\alpha_n \to +\infty$ (or $\alpha_n \to -\infty$) such that $T_\alpha f = g$ uniformly on $\mathbb{R} \times S$ for any compact set $S \subset \mathbb{R}^d$;*
(iii) *for any sequence α', there exists a subsequence $\alpha \subset \alpha'$ such that $T_\alpha f$ exists uniformly on $\mathbb{R} \times S$ for any compact set $S \subset \mathbb{R}^d$.*

Now we recall the definition of asymptotically almost periodic functions.

Definition 2.3. If the function $f : \mathbb{R}_+ \to M$ is continuous and there exists an almost periodic function $\eta : \mathbb{R} \to M$, such that

$$\lim_{t\to+\infty} d(f(t), \eta(t)) = 0, \tag{1}$$

then we say that f is *asymptotically almost periodic* (a.a.p. in short) on $\mathbb{R}_+$. The function η in (1) is called the *almost periodic part of* f. The function f being a.a.p. on $\mathbb{R}_-$ can be defined similarly.

Remark 2. Note that if f is a.a.p. on $\mathbb{R}_+$ or $\mathbb{R}_-$, then its almost periodic part is unique; see e.g. [37, Theorem 3.1] for details.

Lemma 2.4. *The following statements are equivalent to f being asymptotically almost periodic on $\mathbb{R}_+$:*

(i) For any sequence $\alpha' = \{\alpha'_n\}$ with $\alpha'_n \to +\infty$, there exists a subsequence $\alpha \subset \alpha'$ such that $T_\alpha f$ exists uniformly on $\mathbb{R}_+$.

(ii) For any sequence $\alpha' = \{\alpha'_n\}$ with $\alpha'_n \to +\infty$, there exists a subsequence $\alpha \subset \alpha'$ and a constant $\sigma = \sigma(\alpha) > 0$ such that $T_\alpha f$ exists pointwise on $\mathbb{R}_+$ and if sequences $\delta > 0$ (i.e. $\delta_n > 0$ for all n), $\beta \subset \alpha, \gamma \subset \alpha$ are such that

$$T_{\delta+\beta} f = h_1 \quad \text{and} \quad T_{\delta+\gamma} f = h_2$$

exist pointwise on $\mathbb{R}_+$, then either $h_1 \equiv h_2$ or $\inf_{t\in\mathbb{R}_+} d(h_1(t), h_2(t)) \geq 2\sigma$.

Similar results hold when f is asymptotically almost periodic on $\mathbb{R}_-$.

In this paper, we study the SDE:

$$\mathrm{d}X(t) = f(t, X(t))\mathrm{d}t + g(t, X(t))\mathrm{d}W(t), \tag{2}$$

where $f : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ is a continuous function, $g : \mathbb{R} \times \mathbb{R}^d \to M^{d\times m}$ is a $(d \times m)$-matrix-valued continuous function, and W is a standard m-dimensional Brownian motion. Note that almost periodicity is defined on the whole $\mathbb{R}$, but the Brownian motions in SDEs usually defined on $\mathbb{R}_+$. So we need to introduce the two-sided Brownian motion: for two independent Brownian motions W_1, W_2 on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$, let

$$W(t) := \begin{cases} W_1(t), & \text{for } t \geq 0, \\ -W_2(-t), & \text{for } t \leq 0. \end{cases}$$

Then W is a two-sided Brownian motion defined on the filtered probability space $(\Omega, \mathcal{F}, \mathbf{P}, \mathcal{F}_t)$ with $\mathcal{F}_t := \sigma\{W(u) : u \leq t\}, t \in \mathbb{R}$.

Furthermore, we always assume that the coefficients of (2) satisfy the following condition:

(H) The functions f, g are uniformly almost periodic. Moreover, there exists a constant $K > 0$ such that for all $t \in \mathbb{R}$ and $x, y \in \mathbb{R}^d$

$$|f(t, x) - f(t, y)| \vee |g(t, x) - g(t, y)| \leq K|x - y|, \tag{3}$$

where $a \vee b = \max\{a, b\}$ for $a, b \in \mathbb{R}$.

For simplicity, we also denote the equation (2) by (f, g); so equation $(T_\alpha f, T_\alpha g)$ has obvious meaning. By the definition of uniformly almost periodic functions and Remark 1, if the coefficients of (2) satisfy the condition **(H)**, they must satisfy the global linear growth condition, i.e. there is a constant $\hat{K} > 0$ such that

$$|f(t, x)| \vee |g(t, x)| \leq \hat{K}(1 + |x|) \text{ for } (t, x) \in \mathbb{R} \times \mathbb{R}^d. \tag{4}$$

Given an $\mathbb{R}^d$-valued random variable X on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$, we denote by $\mathcal{L}(X)$ the distribution (or law) of X on $\mathbb{R}^d$. We denote by $\mathcal{P}(\mathbb{R}^d)$ the space of all Borel probability measures on $\mathbb{R}^d$. For an $\mathbb{R}^d$-valued random variable X or stochastic process $\{Y(t)\}_{t\in\mathbb{R}}$, we define the following norms:

$$\|X\|_2 := \left(\int_\Omega |X(\omega)|^2 \mathrm{d}\mathbf{P}(\omega)\right)^{\frac{1}{2}}, \quad \|\{Y(t)\}_{t\in\mathbb{R}}\|_\infty := \sup_{t\in\mathbb{R}} \|Y(t)\|_2.$$

In what follows, we denote for a given constant $r > 0$:

$$\begin{aligned}
&\mathbf{L}^2(\mathbf{P}, \mathbb{R}^d) := \{X : \|X\|_2 < \infty\}, \quad \mathcal{B}_r := \{X \in \mathbf{L}^2(\mathbf{P}, \mathbb{R}^d) : \|X\|_2 \le r\},\\
&\mathcal{D}_r := \{\mu \in \mathcal{P}(\mathbb{R}^d) : \exists X \in \mathcal{B}_r \text{ such that } \mathcal{L}(X) = \mu\},\\
&\mathcal{B}_r^{(2)} = \mathcal{B}_r^{(f,g)} := \{X(\cdot) : (X, W) \text{ weakly solves equation } (f, g) \text{ on } \mathbb{R}\\
&\qquad\qquad\qquad\text{on some filtered probability space for some } W \text{ and}\\
&\qquad\qquad\qquad \|X(\cdot)\|_\infty \le r\},\\
&\mathcal{D}_r^{(2)} = \mathcal{D}_r^{(f,g)} := \{\mu(\cdot) : \mu(\cdot) = \mathcal{L}(X(\cdot)) \text{ for some } X(\cdot) \in \mathcal{B}_r^{(f,g)}\},\\
&\mathcal{B}^{(2)} = \mathcal{B}^{(f,g)} := \cup_{r>0}\mathcal{B}_r^{(f,g)}, \ \mathcal{D}^{(2)} = \mathcal{D}^{(f,g)} := \cup_{r>0}\mathcal{D}_r^{(f,g)}.
\end{aligned}$$

It is known that $\mathcal{P}(\mathbb{R}^d)$ can be metrized with some distance (which we denote by $\rho(\cdot,\cdot)$), such that the convergence under distance $\rho(\cdot,\cdot)$ is equivalent to the convergence under the weak-* topology of $\mathcal{P}(\mathbb{R}^d)$, and $\mathcal{P}(\mathbb{R}^d)$ is a complete metric space under $\rho(\cdot,\cdot)$ (see [30, Theorem 2.6.2] for details).

Since we concern almost periodic solutions in distribution, we merely need the stability to be defined among elements of $\mathcal{D}_r^{(f,g)}$. Hence we define, in contrast with references mentioned in Introduction, uniform stability of solutions for SDEs as follows:

Definition 2.5. For a given $t_0 \in \mathbb{R}$, we say element $\mu(\cdot) \in \mathcal{D}_r^{(f,g)}$ is *uniformly stable on* $[t_0, +\infty)$ *within* $\mathcal{D}_r^{(f,g)}$ if for every $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$ such that for any $t_1 \ge t_0$ and any other element $\eta(\cdot) \in \mathcal{D}_r^{(f,g)}$ satisfying

$$\rho(\mu(t_1), \eta(t_1)) < \delta$$

we have

$$\sup_{t\in[t_1,+\infty)} \rho(\mu(t), \eta(t)) < \epsilon.$$

If $\mu(\cdot)$ is uniformly stable on $[t_0, +\infty)$ for every $t_0 \in \mathbb{R}$, we call it *uniformly stable* for short.

In what follows, we get the stability of solutions in distribution mainly by Lyapunov functions, which satisfy the following condition:

(L) Assume that $V : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}_+$ is a function C^2 in $t \in \mathbb{R}$, C^3 in $x \in \mathbb{R}^d$. Assume that the differentials $D^i V$ of V for $i = 0, 1, 2$ and the derivatives $V_{tx_ix_j}$, $V_{x_ix_jx_k}$ for $i, j, k = 1, 2, \cdots, d$ are bounded on $\mathbb{R} \times S$ for every compact set $S \subset \mathbb{R}^d$. Furthermore,

$$\inf_{t\in\mathbb{R}} V(t, x) > 0 \text{ for each } x \ne 0, \text{ and } V(t, 0) = 0 \text{ for all } t \in \mathbb{R}. \tag{5}$$

In this paper, we need the following results from [24] for further discussions:

Proposition 2 ([24], Theorem 3.1)**.** *Consider the following family of Itô stochastic equations on* $\mathbb{R}^d$

$$\mathrm{d}X = f_n(t,X)\mathrm{d}t + g_n(t,X)\mathrm{d}W, \quad n = 1,2,\cdots, \tag{6}$$

where f_n *are* $\mathbb{R}^d$*-valued,* g_n *are* $(d\times m)$*-matrix-valued, and* W *is a standard* m*-dimensional Brownian motion. Assume that* f_n, g_n *satisfy the Lipschitz condition* (3) *and linear growth condition* (4) *with common Lipschitz constant* K *and linear growth constant* $\hat{K}$*. Assume further that* $f_n \to f$*,* $g_n \to g$ *pointwise on* $\mathbb{R}\times\mathbb{R}^d$ *as* $n\to\infty$*, and that* $X_n(\cdot)\in\mathcal{B}_r^{(f_n,g_n)}$ *for a constant* $r>0$*, independent of* n*. Then there is a subsequence of* $\{X_n(\cdot)\}$ *which converges in distribution, uniformly on compact intervals, to a solution* $X(\cdot)\in\mathcal{B}_r^{(f,g)}$*.*

Proposition 3 ([24], Lemma 4.1)**.** *Consider the equation* (2) *with coefficients satisfying condition* **(H)***. If* (2) *admits an* $\mathbf{L}^2$*-bounded solution* $X(\cdot)$ *on* $\mathbb{R}$ *which is asymptotically almost periodic in distribution on* $\mathbb{R}_+$*, then* (2) *admits a solution* $Y(\cdot)$ *on* $\mathbb{R}$ *which is almost periodic in distribution such that*

$$\lim_{t\to+\infty}\rho(\mathcal{L}(X(t)),\mathcal{L}(Y(t)))=0 \quad \textit{and} \quad \|Y(\cdot)\|_\infty \le \|X(\cdot)\|_\infty.$$

In particular, $\mathcal{L}(Y(\cdot))$ *is the almost periodic part of* $\mathcal{L}(X(\cdot))$*. A similar result holds when* $X(\cdot)$ *is asymptotically almost periodic in distribution on* $\mathbb{R}_-$*.*

3. Main results.

Consider (2) and let $V:\mathbb{R}\times\mathbb{R}^d\to\mathbb{R}_+$ be a function being C^1 in $t\in\mathbb{R}$ and C^2 in $x\in\mathbb{R}^d$. For $t\in\mathbb{R}$ and $x,y\in\mathbb{R}^d$, we denote

$$\begin{aligned}\mathcal{L}V(t,x-y) :=& \frac{\partial V}{\partial t}(t,x-y) + \sum_{i=1}^d \frac{\partial V}{\partial x_i}(t,x-y)[f_i(t,x)-f_i(t,y)]\\ &+ \frac{1}{2}\sum_{l=1}^m\sum_{i,j=1}^d [g_{il}(t,x)-g_{il}(t,y)]\frac{\partial^2 V}{\partial x_i\partial x_j}(t,x-y)\\ &\cdot[g_{jl}(t,x)-g_{jl}(t,y))].\end{aligned}$$

Theorem 3.1. *Assume that the coefficients of* (2) *satisfy condition* **(H)** *and that there is a function* V *satisfying condition* **(L)***. Assume further that there is a constant* $b>0$ *such that for all* $(t,x)\in\mathbb{R}\times\mathbb{R}^d$*,*

$$V(t,x)\le b|x|^2, \tag{7}$$

$$\mathcal{L}V(t,x-y)\le 0. \tag{8}$$

Then if $\mathcal{D}_r^{(f,g)}\ne\emptyset$ *for a constant* $r>0$*, all the elements of it are uniformly stable within* $\mathcal{D}_r^{(f,g)}$*; if* $\mathcal{D}_r^{(f,g)}$ *is finite, all of these elements are almost periodic.*

Proof. **Step 1. Uniform stability.** If there is an element $\mu(\cdot)\in\mathcal{D}_r^{(f,g)}$ which is not uniformly stable on $[t_0,+\infty)$ within $\mathcal{D}_r^{(f,g)}$ for some $t_0\in\mathbb{R}$, then there is a sequence $\{\mu_n(\cdot)\}\subset\mathcal{D}_r^{(f,g)}$ such that $\rho(\mu_n(t_0),\mu(t_0))\to 0$ and there are $t_n\ge t_0$ such that

$$\inf_n \rho(\mu_n(t_n),\mu(t_n))\ge\epsilon_0. \tag{9}$$

By the Skorohod representation theorem, there exist random variables $\hat{X}_n$, $\hat{X}$ such that $\mathcal{L}(\hat{X}_n)=\mu_n(t_0)$, $\mathcal{L}(\hat{X})=\mu(t_0)$ and $\hat{X}_n\xrightarrow{a.s.}\hat{X}$, by possibly extending the probability space $(\Omega,\mathcal{F},\mathbf{P})$. By the global Lipschitz condition of coefficients, there exist strong solutions $X_n(\cdot), X(\cdot) \in \mathcal{B}_r^{(f,g)}$ for a given Brownian motion W such that $X_n(t_0) = \hat{X}_n$, $X(t_0) = \hat{X}$, and $\mathcal{L}(X(\cdot)) = \mu(\cdot)$, $\mathcal{L}(X_n(\cdot)) = \mu_n(\cdot)$.

We want to prove that $\rho(\mu_n(t_n), \mu(t_n)) \to 0$ and hence get contradiction to (9). It suffices to prove that $X_n(\cdot)$ uniformly converge to $X(\cdot)$ in probability on $[t_0, +\infty)$; that is, for every $\epsilon > 0$, when n is large enough we have

$$\mathbf{P}\left\{\sup_{t\geq t_0} |X_n(t) - X(t)| \geq \epsilon\right\} < \epsilon. \tag{10}$$

Firstly, we prove that $V(\cdot, X_n(\cdot) - X(\cdot))$ is a supermartingale on $[t_0, +\infty)$ for each n. For $t \geq t_0$, we have

$$\begin{aligned} X_n(t) - X(t) =& \hat{X}_n - \hat{X} + \int_{t_0}^t [f(s, X_n(s)) - f(s, X(s))]\,\mathrm{d}s \\ &+ \int_{t_0}^t [g(s, X_n(s)) - g(s, X(s))]\,\mathrm{d}W(s). \end{aligned}$$

For every $\epsilon > 0$, let

$$V_\epsilon := \inf_{|x|\geq\epsilon, t\geq t_0} V(t, x). \tag{11}$$

By (5) we can see $V_\epsilon > 0$. For $t_0 \leq s < t < +\infty$ and $k, n \in \mathbb{N}$, we define a stopping time

$$\tau_k^n := \inf\{t \geq s : |X_n(t)| \vee |X(t)| > k\}.$$

By Itô's formula, we have

$$\begin{aligned} &V(\tau_k^n \wedge t, X_n(\tau_k^n \wedge t) - X(\tau_k^n \wedge t)) \\ =&V(s, X_n(s) - X(s)) + \int_s^{\tau_k^n\wedge t} \mathcal{L}V(u, X_n(u) - X(u))\mathrm{d}u \\ &+ \int_s^{\tau_k^n\wedge t} \sum_{i=1}^m \sum_{j=1}^d [g_{ji}(u, X_n(u)) - g_{ji}(u, X(u))] \frac{\partial V}{\partial x_j}(u, X_n(u) - X(u))\,\mathrm{d}W_i(u). \end{aligned}$$

Since

$$\begin{aligned} &\mathbf{E}\bigg(\int_s^{\tau_k^n\wedge t} \sum_{i=1}^m \sum_{j=1}^d [g_{ji}(u, X_n(u)) - g_{ji}(u, X(u))] \\ &\qquad \cdot \frac{\partial V}{\partial x_j}(u, X_n(u) - X(u))\,\mathrm{d}W_i(u) \bigg| \mathcal{F}_s \bigg) \\ =&0 \text{ a.s.,} \end{aligned}$$

we have

$$\begin{aligned} &\mathbf{E}(V(\tau_k^n \wedge t, X_n(\tau_k^n \wedge t) - X(\tau_k^n \wedge t))|\mathcal{F}_s) \\ =&\mathbf{E}(V(s, X_n(s) - X(s))|\mathcal{F}_s) + \mathbf{E}\left(\int_s^{\tau_k^n\wedge t} \mathcal{L}V(u, X_n(u) - X(u))\,\mathrm{d}u \middle| \mathcal{F}_s \right). \end{aligned}$$

Since $\mathcal{L}V \leq 0$ and $V(s, X_n(s) - X(s))$ is $\mathcal{F}_s$-measurable, we get

$$\begin{aligned} \mathbf{E}(V(\tau_k^n \wedge t, X_n(\tau_k^n \wedge t) - X(\tau_k^n \wedge t))|\mathcal{F}_s) \leq& \mathbf{E}(V(s, X_n(s) - X(s))|\mathcal{F}_s) \\ =& V(s, X_n(s) - X(s)), \text{ a.s..} \end{aligned}$$

Note that $\tau_k^n \xrightarrow{a.s.} +\infty$ as $k \to +\infty$ for every n, by Fatou's lemma we have:

$$\begin{aligned}\mathbf{E}(V(t, X_n(t) - X(t))|\mathcal{F}_s) =& \mathbf{E}\left(\liminf_{k\to+\infty}(V(\tau_k^n \wedge t, X_n(\tau_k^n \wedge t) - X(\tau_k^n \wedge t))\middle|\mathcal{F}_s\right)\\ \le& \liminf_{k\to+\infty} \mathbf{E}V(\tau_k^n \wedge t, X_n(\tau_k^n \wedge t) - X(\tau_k^n \wedge t))|\mathcal{F}_s)\\ \le& V(s, X_n(s) - X(s)), \text{ a.s..}\end{aligned} \tag{12}$$

So $V(\cdot, X_n(\cdot) - X(\cdot))$ is a supermartingale on $[t_0, +\infty)$.

Now we want to prove that $\mathbf{E}\sqrt{V(t_0, \hat{X}_n - \hat{X})}$ is sufficiently small when n is large enough. By Jensen's inequality and (12) we have

$$\begin{aligned}\mathbf{E}\left(\sqrt{V(t, X_n(t) - X(t))}\middle|\mathcal{F}_s\right) \le& \sqrt{\mathbf{E}(V(t, X_n(t) - X(t))|\mathcal{F}_s)}\\ \le& \sqrt{V(s, X_n(s) - X(s))}, \text{ a.s..}\end{aligned} \tag{13}$$

That is, $\sqrt{V(\cdot, X_n(\cdot) - X(\cdot))}$ is a supermartingale. So by the martingale inequality we have

$$\begin{aligned}\mathbf{P}\left\{\sup_{t\in[t_0,+\infty)} |X_n(t) - X(t)| \ge \epsilon\right\} \le& \mathbf{P}\left\{\sup_{t\in[t_0,+\infty)} \sqrt{V(t, X_n(t) - X(t))} \ge \sqrt{V_\epsilon}\right\}\\ \le& \frac{\mathbf{E}\sqrt{V(t_0, \hat{X}_n - \hat{X})}}{\sqrt{V_\epsilon}}.\end{aligned}$$

Note that $\hat{X}_n \xrightarrow{a.s.} \hat{X}$ and $\sup_n \mathbf{E}|\hat{X}_n|^2 \le r^2$, we have (cf. [11, Theorems 4.5.2, 4.5.4]):

$$\mathbf{E}|\hat{X}_n| \to \mathbf{E}|\hat{X}| \quad \text{as } n \to +\infty$$

and

$$\lim_{n\to+\infty} \mathbf{E}|\hat{X}_n - \hat{X}| = 0.$$

By (7) we have

$$\frac{\mathbf{E}\sqrt{V(t_0, \hat{X}_n - \hat{X})}}{\sqrt{V_\epsilon}} \le \frac{\sqrt{b}\mathbf{E}|\hat{X}_n - \hat{X}|}{\sqrt{V_\epsilon}},$$

which implies that if n is large enough such that

$$\mathbf{E}|\hat{X}_n - \hat{X}| < \frac{\epsilon\sqrt{V_\epsilon}}{\sqrt{b}},$$

then (10) holds. Thus

$$\sup_{t\ge t_0} \rho(\mu_n(t), \mu(t)) \to 0,$$

which is contradictory to (9). Thus each element of $\mathcal{D}_r^{(f,g)}$ is uniformly stable within $\mathcal{D}_r^{(f,g)}$.

Step 2. Inherited property; a.a.p. and almost periodicity when $\mathcal{D}_r^{(f,g)}$ is finite. Firstly, we prove that the consequence of step 1 is also valid for all the hull equations when $\mathcal{D}_r^{(f,g)}$ is finite.

Let the sequence α' be such that $(T_{\alpha'}f, T_{\alpha'}g)$ exists uniformly on $\mathbb{R} \times S$ for any compact set $S \subset \mathbb{R}^d$. Since V, V_t, V_{x_i} are bounded on $\mathbb{R} \times S$, $V(\cdot + \alpha_n', \cdot)$ are uniformly bounded and equi-continuous on $I \times S$ for any compact interval $I \subset \mathbb{R}$. By Arzela-Ascoli's theorem, there is a subsequence $\alpha \subset \alpha'$ such that $T_\alpha V$ exists uniformly on $I \times S$. By the diagonalization argument, the sequence α could be chosen such that $T_\alpha V$ exists uniformly on any compact subset of $\mathbb{R} \times \mathbb{R}^d$.

Similarly we can extract a further subsequence from α, which we still denote by α, such that $T_\alpha V_t$, $T_\alpha V_{x_i}$, $T_\alpha V_{x_i x_j}$ exist uniformly on compact subsets of $\mathbb{R} \times \mathbb{R}^d$. More precisely, we have

$$\frac{\partial T_\alpha V}{\partial t} = T_\alpha V_t,\ \frac{\partial T_\alpha V}{\partial x_i} = T_\alpha V_{x_i},\ \frac{\partial^2 T_\alpha V}{\partial x_i \partial x_j} = T_\alpha V_{x_i x_j}, \text{ for } i, j = 1, \cdots, d, \text{ on } \mathbb{R} \times \mathbb{R}^d.$$

So we have

$$T_\alpha V(t, x) \le b|x|^2,$$

$$\begin{aligned} \mathcal{L} T_\alpha V(t, x-y) =& \frac{\partial T_\alpha V}{\partial t}(t, x-y) + \sum_{i=1}^d \frac{\partial T_\alpha V}{\partial x_i}(t, x-y)[T_\alpha f_i(t,x) - T_\alpha f_i(t,y)] \\ &+ \frac{1}{2} \sum_{l=1}^m \sum_{i,j=1}^d [T_\alpha g_{il}(t,x) - T_\alpha g_{il}(t,y)] \frac{\partial^2 T_\alpha V}{\partial x_i \partial x_j}(t, x-y) \\ &\cdot [T_\alpha g_{jl}(t,x) - T_\alpha g_{jl}(t,y)] \le 0 \end{aligned}$$

for all $(t, x, y) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d$. Repeating Step 1, we obtain that all the elements of $\mathcal{D}_r^{(T_\alpha f, T_\alpha g)}$ are uniformly stable within $\mathcal{D}_r^{(T_\alpha f, T_\alpha g)}$.

By the uniform stability and the finiteness of the set $\mathcal{D}_r^{(f,g)}$, there is a separating constant $d(f,g)$, depending only on (f,g) but independent of $\mu(\cdot) \in \mathcal{D}_r^{(f,g)}$, such that for any two different elements $\eta(\cdot), \mu(\cdot) \in \mathcal{D}_r^{(f,g)}$ we have

$$\inf_{t \in \mathbb{R}_-} \rho(\eta(t), \mu(t)) > d(f,g). \tag{14}$$

By Proposition 1-(ii), we may assume without loss of generality that the above sequence α satisfies $\lim_{n\to\infty} \alpha_n = -\infty$, so it follows from (14) that

$$\inf_{t \in \mathbb{R}_-} \rho(T_\alpha \eta(t), T_\alpha \mu(t)) \ge d(f,g).$$

On the other hand, it follows from Proposition 2 that $T_\alpha \mu(\cdot) \in \mathcal{D}_r^{(T_\alpha f, T_\alpha g)}$, so $\mathcal{D}_r^{(T_\alpha f, T_\alpha g)}$ has no less elements than $\mathcal{D}_r^{(f,g)}$ does.

Conversely, by Proposition 1-(i), $(T_\alpha f, T_\alpha g)$ is uniformly almost periodic and $(f,g) \in H(T_\alpha f, T_\alpha g)$. So, by the same symmetric argument as above, $\mathcal{D}_r^{(f,g)}$ also has no less elements than $\mathcal{D}_r^{(T_\alpha f, T_\alpha g)}$ does and the separating constant $d(T_\alpha f, T_\alpha g) \le d(f,g)$. That is, all the equations in the hull $H(f,g)$ share the same number of elements as $\mathcal{D}_r^{(f,g)}$ and the same separating constant $d(f,g)$.

Now we prove that all the elements of $\mathcal{D}_r^{(f,g)}$ are a.a.p.. For the above sequence α with $\alpha_n \to -\infty$ and a given sequence $\delta = \{\delta_n\}$ with $\delta_n < 0$, by Proposition 1-(iii) there exist subsequences of α and δ, which we still denote by α and δ, such that $(T_{\alpha+\delta} f, T_{\alpha+\delta} g)$ exists uniformly on $\mathbb{R} \times S$ for any compact set $S \subset \mathbb{R}^d$. By Arzela-Ascoli's theorem there are subsequences $\beta, \gamma \subset \alpha$ such that $T_{\beta+\delta}\mu(t)$, $T_{\gamma+\delta}\mu(t)$ exist uniformly on compact intervals (see the proof of [24, Theorem 3.1] for details). By Proposition 2, $T_{\beta+\delta}\mu(\cdot), T_{\gamma+\delta}\mu(\cdot) \in \mathcal{D}_r^{(T_{\alpha+\delta} f, T_{\alpha+\delta} g)}$, then by the separating property obtained above we have

$$T_{\beta+\delta}\mu(t) \equiv T_{\gamma+\delta}\mu(t) \text{ or } \inf_{t \in \mathbb{R}_-} \rho(T_{\beta+\delta}\mu(t), T_{\gamma+\delta}\mu(t)) \ge d(f,g).$$

Then it follows from Lemma 2.4 that all the elements of $\mathcal{D}_r^{(f,g)}$ are all a.a.p. on $\mathbb{R}_-$.

By Proposition 3, there is an element $\hat{\mu}(\cdot) \in \mathcal{D}_r^{(f,g)}$, which is almost periodic and satisfies

$$\lim_{t\to-\infty} \rho(\mu(t), \hat{\mu}(t)) = 0.$$

By the separating property, $\hat{\mu}(\cdot) = \mu(\cdot)$, which implies that each element of $\mathcal{D}_r^{(f,g)}$ is almost periodic. The proof is complete. $\square$

If the averaged derivative of V is negative definite, then $\mathcal{D}_r^{(f,g)}$ will be a singleton set, which may be more convenient for use in some applications.

Theorem 3.2. *Let the coefficients of (2) satisfy condition* **(H)**. *Assume that there is a function V satisfying condition* **(L)**, *and that there are constants $a, b > 0$ such that*

$$a|x|^2 \le V(t,x) \le b|x|^2 \quad \text{for all } (t,x) \in \mathbb{R}\times\mathbb{R}^d. \tag{15}$$

Assume further that there is a positive definite function $c : \mathbb{R}_+ \to \mathbb{R}_+$ which is convex, increasing on $\mathbb{R}_+$, and

$$\mathcal{L}V(t, x-y) \le -c(|x-y|^2) \quad \text{for } t\in\mathbb{R} \text{ and } x,y\in\mathbb{R}^d. \tag{16}$$

Then either $\mathcal{D}^{(f,g)} = \emptyset$ or it consists of a unique element which is almost periodic.

Proof. We prove the uniqueness by contradiction. If there are two elements $\mu(\cdot), \eta(\cdot) \in \mathcal{D}^{(f,g)}$, then there is a constant $r > 0$ such that $\mu(\cdot), \eta(\cdot) \in \mathcal{D}_r^{(f,g)}$. Assume that $X(\cdot)$, $Y(\cdot)$ are two strong $\mathbf{L}^2$-bounded solutions of (2) for a given Brownian motion W such that $\mathcal{L}(X(t)) = \mu(t)$, $\mathcal{L}(Y(t)) = \eta(t)$ for $t \in \mathbb{R}$.

For a given $\epsilon > 0$, let $T(\epsilon) = 2br^2/c(\epsilon) + 1$. Firstly, we prove that for every $t \in \mathbb{R}$ there is $t_1 \in [t, t+T(\epsilon)]$ such that

$$\mathbf{E}|X(t_1) - Y(t_1)|^2 < \epsilon. \tag{17}$$

If this is not true, then there is $\hat{t} \in \mathbb{R}$ and $\epsilon_0 > 0$ such that

$$\inf_{t\in[\hat{t},\hat{t}+T(\epsilon_0)]} \mathbf{E}|X(t) - Y(t)|^2 \ge \epsilon_0.$$

Similar to the proof of Theorem 3.1, for a given $s \in \mathbb{R}$, we define

$$\tau_k := \inf\{t \ge s : |Y(t)| \vee |X(t)| > k\}.$$

Then it follows from Ito's formula that for $t \ge s$

$$\begin{aligned}
&V(\tau_k\wedge t, X(\tau_k\wedge t) - Y(\tau_k\wedge t))\\
=&\mathbf{E}V(s, X(s) - Y(s)) + \int_s^{\tau_k\wedge t} \mathbf{E}\mathcal{L}V(u, X(u) - Y(u))\,\mathrm{d}s\\
&+ \int_s^{\tau_k^n\wedge t} \sum_{i=1}^m \sum_{j=1}^d [g_{ji}(u, X_n(u)) - g_{ji}(u, X(u))] \frac{\partial V}{\partial x_j}(u, X_n(u) - X(u))\,\mathrm{d}W_i(u).
\end{aligned}$$

Since

$$\sup_{t\in\mathbb{R}} \mathbf{E}|X(t) - Y(t)|^2 \le 2r^2,$$

by (15), (16) we have

$$\begin{aligned}
&\mathbf{E}V(\tau_k\wedge t, X(\tau_k\wedge t) - Y(\tau_k\wedge t))\\
\le& b\mathbf{E}|X(s) - Y(s)|^2 - \mathbf{E}\int_s^{\tau_k\wedge t} c(|X(u) - Y(u)|^2)\,\mathrm{d}s\\
\le& 2br^2 - \mathbf{E}\int_s^{\tau_k\wedge t} c(|X(u) - Y(u)|^2)\,\mathrm{d}s.
\end{aligned}$$

Because $c(\cdot)$ is convex and increasing on $\mathbb{R}_+$, by Jensen's inequality we have:

$$\mathbf{E}\int_s^{\tau_k\wedge t} c(|X(u)-Y(u)|^2)\,\mathrm{d}s \geq \int_s^{\tau_k\wedge t} c(\mathbf{E}|X(u)-Y(u)|^2)\,\mathrm{d}s \geq c(\epsilon_0)(\tau_k\wedge t-s).$$

So

$$\mathbf{E}V(\tau_k\wedge t, X(\tau_k\wedge t)-Y(\tau_k\wedge t)) \leq 2br^2 - c(\epsilon_0)(\tau_k\wedge t-s). \tag{18}$$

Note that $\tau_k \xrightarrow{a.s.} +\infty$ as $k\to+\infty$, by Fatou's lemma and (18) we have

$$\begin{aligned}\mathbf{E}V(t,X(t)-Y(t)) &= \mathbf{E}\left(\liminf_{k\to+\infty} V(\tau_k\wedge t, X(\tau_k\wedge t)-Y(\tau_k\wedge t))\right)\\ &\leq \liminf_{k\to+\infty}\mathbf{E}[2br^2-c(\epsilon_0)(\tau_k\wedge t-s)]\\ &\leq 2br^2-c(\epsilon_0)(t-s).\end{aligned}$$

Letting $s=\hat{t}$ and $t=\hat{t}+T(\epsilon_0)$, we have

$$0\leq \mathbf{E}V(\hat{t}+T(\epsilon_0), X(\hat{t}+T(\epsilon_0))-Y(\hat{t}+T(\epsilon_0)))\leq 2br^2-c(\epsilon_0)T(\epsilon_0)=-c(\epsilon_0)<0,$$

a contradiction. Thus there is $t_1\in[t,t+T(\epsilon)]$ such that (17) is valid.

For a given $s\in\mathbb{R}$, assume that $t_1\in[s,s+T(\epsilon)]$ fulfills (17). By (15)–(17), for any $t\geq t_1$ we have:

$$a\mathbf{E}|X(t)-Y(t)|^2\leq \mathbf{E}V(t,X(t)-Y(t))\leq \mathbf{E}V(t_1,X(t_1)-Y(t_1))\leq b\epsilon.$$

Note that $s\in\mathbb{R}$ is arbitrarily chosen and $T(\epsilon)$ is only determined by $\epsilon>0$, so we actually have proved

$$\mathbf{E}|X(t)-Y(t)|^2\leq\epsilon \qquad \text{for all } t\in\mathbb{R}.$$

Thus $X(t)=Y(t)$ for all $t\in\mathbb{R}$ almost surely, which implies that $\mu(t)=\eta(t)$ for all $t\in\mathbb{R}$. That is, $\mathcal{D}^{(f,g)}$ has a unique element if it is not empty. Finally, it follows from Theorem 3.1 that this unique element is almost periodic. The proof is complete. ☐

Now we give another result to ensure the existence of almost periodic elements in $\mathcal{D}^{(f,g)}$ without the information of the number of its elements.

Theorem 3.3. *Assume that the coefficients of* (2) *satisfy condition* **(H)**, *and that there exists a function V which is C^1 in t, C^2 in x and satisfies the positive definite condition* (5). *Assume further that there exists a constant $b>0$ such that* (7) *is valid on $\mathbb{R}_+\times\mathbb{R}^d$, and that for all $t\in\mathbb{R}_+$, $s_1,s_2\in\mathbb{R}_+$ and $x,y\in\mathbb{R}^d$*

$$\begin{aligned}\mathcal{L}_{s_1,s_2}V(t,x-y):=&\frac{\partial V}{\partial t}(t,x-y)+\sum_{i=1}^d\frac{\partial V}{\partial x_i}(t,x-y)[f_i(t+s_1,x)-f_i(t+s_2,y)]\\ &+\frac{1}{2}\sum_{l=1}^m\sum_{i,j=1}^d[g_{il}(t+s_1,x)-g_{il}(t+s_2,y)]\frac{\partial^2 V}{\partial x_i\partial x_j}(t,x-y)\\ &\cdot[g_{jl}(t+s_1,x)-g_{jl}(t+s_2,y)]\leq 0.\end{aligned}\tag{19}$$

Then if (2) *has $\mathbf{L}^2$-bounded solutions, the distributions of these solutions are a.a.p. on $\mathbb{R}_+$ and* (2) *admits at least one solution with almost periodic distribution.*

Proof. For a sequence $\alpha=\{\alpha_n\}$ with $\alpha_n\to+\infty$, assume that $(T_\alpha f,T_\alpha g)$ exists uniformly on $\mathbb{R}\times S$ for any compact set $S\subset\mathbb{R}^d$, and that $T_\alpha\mu(\cdot)$ exists uniformly on compact intervals (see, again, the proof of [24, Theorem 3.1] for details). For a given constant $r > 0$, we want to prove that $T_\alpha \mu(\cdot)$ exists uniformly on $\mathbb{R}_+$ for any $\mu(\cdot) \in \mathcal{D}_r^{(f,g)}$.

By the Skorohod representation theorem, there are suitable random variables $\hat{X}_n$, $\hat{X}$ such that $\hat{X}_n \xrightarrow{a.s.} \hat{X}$ as $n \to +\infty$ and $\mathcal{L}(\hat{X}_n) = \mu(\alpha_n)$, $\mathcal{L}(\hat{X}) = T_\alpha \mu(0)$, by possibly extending the probability space $(\Omega, \mathcal{F}, \mathbf{P})$. By the global Lipschitz condition of coefficients, for a given Brownian motion W, there are strong solutions $X_n(\cdot) \in \mathcal{B}_r^{(f_{\alpha_n}, g_{\alpha_n})}$ such that $X_n(0) = \hat{X}_n$, and $\mathcal{L}(X_n(\cdot)) = \mu(\cdot + \alpha_n)$, where $f_{\alpha_n}(\cdot,\cdot) := f(\cdot + \alpha_n, \cdot)$ and similarly for g_{α_n}. Moreover, for $n, p \in \mathbb{N}$, we have for $t \geq 0$

$$\begin{aligned}
&X_{n+p}(t) - X_n(t) \\
=&\hat{X}_{n+p} - \hat{X}_n + \int_0^t [f(u + \alpha_{n+p}, X_{n+p}(u)) - f(u + \alpha_n, X_n(u))]\mathrm{d}u \\
&+ \int_0^t [g(u + \alpha_{n+p}, X_{n+p}(u)) - g(u + \alpha_n, X_n(u))]\mathrm{d}W(u).
\end{aligned}$$

We now show that $V(\cdot, X_{n+p}(\cdot) - X_n(\cdot))$ is a supermartingale on $\mathbb{R}_+$ for given n and p. For $0 \leq s < t < +\infty$, we define stopping times

$$\tau_k^{n,p} := \inf\{t \geq s; |X_n(t)| \vee |X_{n+p}(t)| > k\} \quad \text{for } k \in \mathbb{N}.$$

By Itô's formula we have

$$\begin{aligned}
&V(\tau_k^{n,p} \wedge t, X_{n+p}(\tau_k^{n,p} \wedge t) - X_n(\tau_k^{n,p} \wedge t)) \\
=&V(s, X_{n+p}(s) - X_n(s)) + \int_s^{\tau_k^{n,p} \wedge t} \mathcal{L}_{\alpha_{n+p}, \alpha_n} V(u, X_{n+p}(u) - X_n(u))\,\mathrm{d}u \\
&+ \int_s^{\tau_k^{n,p} \wedge t} \sum_{i=1}^m \sum_{j=1}^d [g_{ji}(u + \alpha_{n+p}, X_{n+p}(u)) - g_{ji}(u + \alpha_n, X_n(u))] \\
&\cdot \frac{\partial V}{\partial x_j}(u, X_{n+p}(u) - X_n(u))\,\mathrm{d}W_i(u).
\end{aligned}$$

Since

$$\begin{aligned}
&\mathbf{E}\Bigg(\int_s^{\tau_k^{n,p} \wedge t} \sum_{i=1}^m \sum_{j=1}^d [g_{ji}(u + \alpha_{n+p}, X_{n+p}(u)) - g_{ji}(u + \alpha_n, X_n(u))] \\
&\qquad \cdot \frac{\partial V}{\partial x_j}(u, X_{n+p}(u) - X_n(u))\,\mathrm{d}W_i(u)\Bigg|\mathcal{F}_s\Bigg) \\
=&0 \text{ a.s.},
\end{aligned}$$

it follows from (19) that

$$\begin{aligned}
&\mathbf{E}(V(\tau_k^{n,p} \wedge t, X_{n+p}(\tau_k^{n,p} \wedge t) - X_n(\tau_k^{n,p} \wedge t))|\mathcal{F}_s) \\
\leq&\mathbf{E}(V(s, X_{n+p}(s) - X_n(s))|\mathcal{F}_s) \\
=&V(s, X_{n+p}(s) - X_n(s)), \text{ a.s.}.
\end{aligned}$$

Noting that $\tau_k^{n,p} \xrightarrow{a.s.} +\infty$ as $k \to +\infty$ for fixed n and p, we have by Fatou's lemma (similar to (12)):

$$\mathbf{E}(V(t, X_{n+p}(t) - X_n(t))|\mathcal{F}_s)$$

$$
\begin{aligned}
&= \mathbf{E}\left(\liminf_{k\to+\infty} V(\tau_k^{n,p}\wedge t, X_{n+p}(\tau_k^{n,p}\wedge t) - X_n(\tau_k^{n,p}\wedge t)\Big|\mathcal{F}_s\right)\\
&\le \liminf_{k\to+\infty} \mathbf{E}V(\tau_k^{n,p}\wedge t, X_{n+p}(\tau_k^{n,p}\wedge t) - X_n(\tau_k^{n,p}\wedge t)|\mathcal{F}_s)\\
&\le V(s, X_{n+p}(s) - X_n(s)), \text{ a.s..}
\end{aligned}
$$

That is, $V(\cdot, X_{n+p}(\cdot) - X_n(\cdot))$ is a supermartingale for given p and n. Similar to (13), by Jensen's inequality,

$$
\begin{aligned}
\mathbf{E}\left(\sqrt{V(t, X_{n+p}(t) - X_n(t))}\ \Big|\mathcal{F}_s\right) &\le \sqrt{\mathbf{E}(V(t, X_{n+p}(t) - X_n(t))|\mathcal{F}_s)}\\
&\le \sqrt{V(s, X_{n+p}(s) - X_n(s))}, \text{ a.s..}
\end{aligned}
$$

So $\sqrt{V(\cdot, X_{n+p}(\cdot) - X_n(\cdot))}$ is also a supermartingale.

For any $\epsilon > 0$, we define $V_\epsilon > 0$ as in (11). Then by the martingale inequality, we have

$$
\begin{aligned}
\mathbf{P}\left\{\sup_{t\in\mathbb{R}_+} |X_{n+p}(t) - X_n(t)| \ge \epsilon\right\} &\le \mathbf{P}\left\{\sup_{t\in\mathbb{R}_+} \sqrt{V(t, X_{n+p}(t) - X_n(t))} \ge \sqrt{V_\epsilon}\right\}\\
&\le \frac{\mathbf{E}\sqrt{V(0, \hat{X}_{n+p} - \hat{X}_n)}}{\sqrt{V_\epsilon}}.
\end{aligned}
\tag{20}
$$

Because $\mathbf{E}|\hat{X}_n|^2 \le r^2$ and $\hat{X}_n \xrightarrow{a.s.} \hat{X}$, we have (see [11, Theorems 4.5.2, 4.5.4])

$$
\mathbf{E}|\hat{X}_n| \to \mathbf{E}|\hat{X}| \quad \text{as } n \to +\infty
$$

and

$$
\lim_{n\to+\infty} \mathbf{E}|\hat{X}_n - \hat{X}| = 0.
$$

So

$$
\lim_{n\to+\infty} \sup_{p\in\mathbb{N}} \mathbf{E}|\hat{X}_{n+p} - \hat{X}_n| = 0.
$$

When n is large enough, by (7) we have

$$
\sup_{p\in\mathbb{N}} \mathbf{E}\sqrt{V(0, \hat{X}_{n+p} - \hat{X}_n)} \le \sqrt{b} \sup_{p\in\mathbb{N}} \mathbf{E}|\hat{X}_{n+p} - \hat{X}_n| < \epsilon\sqrt{V_\epsilon}.
$$

This together with (20) implies

$$
\sup_{p\in\mathbb{N}} \mathbf{P}\left\{\sup_{t\in\mathbb{R}_+} |X_{n+p}(t) - X_n(t)| \ge \epsilon\right\} < \epsilon.
$$

By Theorem 4.1.3 in [11], there exists a suitable stochastic process $\widetilde{X}(\cdot)$ such that $X_n(\cdot) \to \widetilde{X}(\cdot)$ in probability uniformly on $\mathbb{R}_+$. Thus $\mu(\cdot + \alpha_n)$ uniformly converges to $T_\alpha\mu(\cdot)$ on $\mathbb{R}_+$.

By Lemma 2.4, we can see that each $\mu(\cdot) \in \mathcal{D}_r^{(f,g)}$ is a.a.p. on $\mathbb{R}_+$. So the distribution of any $\mathbf{L}^2$-bounded solution is a.a.p. on $\mathbb{R}_+$. By Proposition 3, there exists at least one $\mathbf{L}^2$-bounded solution of (2) with almost periodic distribution. The proof is complete. □

To discuss the almost periodicity of solutions for SDEs, we need to find ways to obtain $\mathbf{L}^2$-bounded solutions on $\mathbb{R}$, which may reduce to finding $\mathbf{L}^2$-bounded solutions on $\mathbb{R}_+$:

Proposition 4 (cf. [24], Theorem 4.7)**.** *Assume that the coefficients of* (2) *satisfy condition* **(H)** *and that* (2) *admits a solution* φ *on* $[t_0, +\infty)$ *for some* $t_0 \in \mathbb{R}$ *with* $\sup_{t\geq t_0} ||\varphi(t)||_2 \leq M$ *for a constant* $M > 0$*, then* (2) *has a solution* $\widetilde{\varphi}$ *on* $\mathbb{R}$ *with* $||\widetilde{\varphi}(\cdot)||_\infty \leq M$*.*

We now conclude this section by giving a sufficient condition for the existence of $\mathbf{L}^2$-bounded solutions via Lyapunov functions:

Theorem 3.4. *Assume that the coefficients of* (2) *satisfy condition* **(H)***. Assume moreover that there is a function* V *which is* C^1 *in* t *and* C^2 *in* x *such that* $V(t,x) \leq \bar{M}$ *uniformly for* $t \in \mathbb{R}$ *whenever* $|x| \leq R$*, where* $\bar{M}, R > 0$ *are constants; when* $|x| \geq R$*,* V *satisfies*

$$a|x|^2 \leq V(t,x) \leq b(t)|x|^2 + c(t)$$

and

$$LV := \frac{\partial V}{\partial t} + \sum_{i=1}^{d} \frac{\partial V}{\partial x_i} f_i + \sum_{l=1}^{m} \sum_{i,j=1}^{d} g_{il} \frac{\partial^2 V}{\partial x_i \partial x_j} g_{jl} \leq 0$$

for $t \in \mathbb{R}$*, where* $a > 0$ *is a constant and* $b(\cdot)$*,* $c(\cdot)$ *are positive functions on* $\mathbb{R}$*. Then if* $X(\cdot)$ *is a solution of* (2) *with initial condition* $\mathbf{E}|X(t_0)|^2 < +\infty$*,* $X(\cdot)$ *is* $\mathbf{L}^2$*-bounded on* $[t_0, +\infty)$*.*

Proof. Suppose that $X(\cdot)$ is the solution of (2) with $\mathbf{L}^2$-bounded initial value at the moment t_0. Since the coefficients satisfy condition **(H)**, the solution $X(\cdot)$ exists on $[t_0, +\infty)$. We define a sequence of stopping times:

$$\tau_n^R := \inf\{t \geq t_0 : |X(t)| \geq n \text{ or } |X(t)| \leq R\}$$

and

$$\tau^R := \inf\{t \geq t_0 : |X(t)| \leq R\}.$$

Then $\tau_n^R \xrightarrow{a.s.} \tau^R$ as $n \to +\infty$.

Denote by B_R the closed ball $\{x \in \mathbb{R}^d : |x| \leq R\}$. When $X(t_0)$ is supported on $\mathbb{R}^d - B_R$, by Itô's formula we have for $t \geq t_0$

$$\begin{aligned} \mathbf{E}V(t \wedge \tau_n^R, X(t \wedge \tau_n^R)) =& \mathbf{E}V(t_0, X(t_0)) + \mathbf{E}\int_{t_0}^{t\wedge\tau_n^R} LV(u, X(u))\, \mathrm{d}u \\ \leq& \mathbf{E}V(t_0, X(t_0)) \leq c(t_0) + b(t_0)\mathbf{E}|X(t_0)|^2. \end{aligned}$$

Then Fatou's lemma implies that

$$\mathbf{E}V(t \wedge \tau^R, X(t \wedge \tau^R)) \leq \mathbf{E}V(t_0, X(t_0)) \leq c(t_0) + b(t_0)\mathbf{E}|X(t_0)|^2 \tag{21}$$

by letting $n \to +\infty$.

When $X(t_0)$ is supported on $\mathbb{R}^d$, by (21) we have for $t \geq t_0$

$$\begin{aligned} \mathbf{E}V(t, X(t)) \leq& \mathbf{P}(\tau^R \geq t) \cdot \int_{\{X(t_0)>R\}} V(t_0, X(t_0, \omega))\, \mathrm{d}\mathbf{P}(\omega) \\ &+ \mathbf{P}(\tau^R < t) \cdot \left[\bar{M} + \int_{\{X(t_0)>R\}} V(t_0, X(t_0, \omega))\, \mathrm{d}\mathbf{P}(\omega)\right] \\ \leq& c(t_0) + b(t_0)\mathbf{E}|X(t_0)|^2 + \bar{M}. \end{aligned} \tag{22}$$

Note that in (22) either $|X(t)| \leq R$ or $a|X(t)|^2 \leq V(t, X(t))$, so $X(\cdot)$ is $\mathbf{L}^2$-bounded on $[t_0, +\infty)$. ∎

4. **Applications.** In this section, we illustrate our theoretical results by several examples. Firstly we consider the simplest case of almost periodic SDEs.

Example 4.1. Consider the one-dimensional SDE

$$\mathrm{d}X(t) = f(t, X(t))\mathrm{d}t + g(t, X(t))\mathrm{d}W(t), \tag{23}$$

where f, g satisfy condition **(H)** and are C^1 in x. Assume that there is a constant $c > 0$ such that

$$\sup_{(t,x)\in\mathbb{R}\times\mathbb{R}} \left|\frac{\partial g}{\partial x}(t,x)\right|^2 \leq c \text{ and } \sup_{(t,x)\in\mathbb{R}\times\mathbb{R}} \frac{\partial f}{\partial x}(t,x) \leq -c. \tag{24}$$

Then for equation (23), either $\mathcal{D}^{(f,g)} = \emptyset$ or it has a unique element which is almost periodic.

Proof. Let $V(t,x) = |x|^2$. Then it is easy to see that V satisfies condition **(L)** and

$$\frac{\partial V}{\partial t}(t,x) = 0, \quad \frac{\partial V}{\partial x}(t,x) = 2x, \quad \frac{\partial^2 V}{\partial x^2}(t,x) = 2.$$

By (24) and mean value theorem, for $x, y \in \mathbb{R}$ and $t \in \mathbb{R}$, if $x \neq y$, there exist $\hat{\xi} = \hat{\xi}(t,x,y)$, $\xi = \xi(t,x,y)$ such that $\hat{\xi}, \xi \in (x \wedge y, x \vee y)$ (where $x \wedge y = \min\{x, y\}$) and

$$(f(t,x) - f(t,y))(x-y) = \frac{\partial f}{\partial x}(t,\xi)(x-y)^2 \leq -c(x-y)^2,$$

$$(g(t,x) - g(t,y))^2 = (x-y)^2 \left|\frac{\partial g}{\partial x}(t,\hat{\xi})\right|^2 \leq c(x-y)^2.$$

So

$$\begin{aligned}\mathcal{L}V(t, x-y) =& 2(f(t,x) - f(t,y))(x-y) + (g(t,x) - g(t,y))^2 \\ \leq& -c(x-y)^2 = -c|x-y|^2.\end{aligned}$$

By Theorem 3.2 we can easily get the required result. □

Now let us consider some two-dimensional applications.

Example 4.2. Consider the two-dimensional SDE:

$$\begin{cases} \mathrm{d}X_1(t) = [f_1(t, X_1(t)) + \sigma X_2(t)]\mathrm{d}t + [A_1(t)X_1(t) + g_1(t)]\mathrm{d}W_1(t), \\ \mathrm{d}X_2(t) = [f_2(t, X_2(t)) - \sigma X_1(t)]\mathrm{d}t + [A_2(t)X_2(t) + g_2(t)]\mathrm{d}W_2(t), \end{cases} \tag{25}$$

where $f_i(t,x)$ are C^1 in x and satisfy condition **(H)** for $i = 1, 2$, and $\sigma \neq 0$ is a constant. Assume that A_i, g_i are almost periodic and $f_i(t, 0) \equiv 0$, $i = 1, 2$. Let $a(t) := \max\{A_i^2(t), g_i^2(t) : i = 1, 2\}$. Assume further that for $t, x \in \mathbb{R}$,

$$\frac{\partial f_i}{\partial x}(t,x) \leq -2a(t) - 1, \quad i = 1, 2. \tag{26}$$

Then $\mathcal{D}^{(25)}$ has a unique element which is almost periodic.

Proof. Let $V : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$, $V(t,x) = |x|^2 = x_1^2 + x_2^2$. Then V satisfies condition **(L)**, and for $(t,x) \in \mathbb{R} \times \mathbb{R}^2$, $i, j = 1, 2$,

$$\frac{\partial V}{\partial x_i}(t,x) = 2x_i, \quad \frac{\partial^2 V}{\partial x_i^2}(t,x) = 2, \quad \frac{\partial^2 V}{\partial x_i \partial x_j}(t,x) = 0 \text{ whenever } i \neq j.$$

By (26) and mean value theorem, for $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R}^2$, there are $\xi_i = \xi_i(t, x_i, y_i) \in (x_i \wedge y_i, x_i \vee y_i)$, $i = 1, 2$, such that

$$\begin{aligned}\mathcal{L}V(t, x-y) =& 2\sum_{i=1,2}[f_i(t,x_i) - f_i(t,y_i)](x_i - y_i) + \sum_{i=1,2} A_i^2(t)(x_i - y_i)^2 \\ \leq& \sum_{i=1,2}[a(t) + 2\frac{\partial f}{\partial x}(t,\xi_i)(x_i - y_i)^2] \\ \leq& (-3a(t) - 2)(x_i - y_i)^2 \leq -2|x-y|^2.\end{aligned}$$

Since $f_i(t, 0) = 0$, for every x_i and t, there exist $\hat{\xi}_i = \hat{\xi}_i(t, x_i) \in (x_i \wedge 0, x_i \vee 0)$ such that

$$f_i(t, x_i)x_i = \frac{\partial f_i}{\partial x_i}(t, \hat{\xi}_i)x_i^2 \leq -(2a(t) + 1)x_i^2.$$

So

$$\begin{aligned}LV(t,x) =& 2\sum_{i=1,2} f_i(t,x_i)x_i + \sum_{i=1,2}[A_i(t)x_i + g_i(t)]^2 \\ \leq& \sum_{i=1,2}[2A_i^2(t)x_i^2 + 2g_i^2(t) + 2\frac{\partial f}{\partial x}(t,\hat{\xi}_i)x_i^2] \\ \leq& \sum_{i=1,2}[-(2a(t)+2)x_i^2 + 2a(t)].\end{aligned}$$

Obviously $LV(t, x) \leq 0$ when $|x| \geq \sqrt{2}$. By the global Lipschitz condition of the coefficients, we can see that (25) must have $\mathbf{L}^2$-bounded solutions from Proposition 4 and Theorem 3.4. By Theorem 3.2 we can get the result required. □

Example 4.3. Consider the two-dimensional SDE:

$$\begin{cases} \mathrm{d}X_1(t) = & [-(A_1^2(t) + A_2^2(t) + 1)X_1(t) + 2A_1^2(t)X_2(t)]\mathrm{d}t \\ & +A_1(t)(X_1(t) - X_2(t))\mathrm{d}W_1(t), \\ \mathrm{d}X_2(t) = & [-(A_2^2(t) + A_1^2(t) + 1)X_2(t) + 2A_2^2(t)X_1(t)]\mathrm{d}t \\ & +A_2(t)(X_1(t) - X_2(t))\mathrm{d}W_2(t). \end{cases} \tag{27}$$

If $A_i(\cdot)$ are almost periodic for $i = 1, 2$, then (27) has $\mathbf{L}^2$-bounded solutions and all the $\mathbf{L}^2$-bounded solutions of (27) share the same distribution which is almost periodic.

Proof. Similar to the proof of Example 4.2, let $V(t, x) = x_1^2 + x_2^2$. For $t \in \mathbb{R}$, $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R}^2$, we have

$$\begin{aligned}&\mathcal{L}V(t, x-y) \\ =& 2\sum_{i=1,2}\Big(-[A_1^2(t) + A_2^2(t) + 1](x_i - y_i)^2 + A_i^2(t)(x_1 - y_1)(x_2 - y_2)\Big) \\ &+ \sum_{i=1,2}[(A_1^2(t) + A_2^2(t))(x_i - y_i)^2] - 2[A_1^2(t) + A_2^2(t)](x_1 - y_1)(x_2 - y_2) \\ \leq& -2|x-y|^2,\end{aligned}$$

and

$$\begin{aligned}LV(t,x) =& 2\sum_{i=1,2}[-(A_1^2(t) + A_2^2(t) + 1)x_i^2 + 2A_i^2(t)x_1x_2] + \sum_{i=1,2} A_i^2(t)(x_1 - x_2)^2 \\ \leq& -2|x|^2 \leq 0.\end{aligned}$$

By Proposition 4 and Theorem 3.4, (27) has $\mathbf{L}^2$-bounded solutions. By Theorem 3.2, $\mathcal{D}^{(27)}$ has a unique element which is almost periodic. $\square$

REFERENCES

[1] L. Amerio and G. Prouse, *Almost-periodic Functions and Functional Equations*, Van Nostrand Reinhold Co., New York-Toronto, Ont.-Melbourne, 1971. viii+184 pp.

[2] L. Arnold and C. Tudor, Stationary and almost periodic solutions of almost periodic affine stochastic differential equations, *Stochastics Stochastics Rep.*, **64** (1998), 177–193.

[3] G. K. Basak, A class of limit theorems for singular diffusions, *J. Multivariate Anal.*, **39** (1991), 44–59.

[4] G. K. Basak and R. N. Bhattachaya, Stability in distribution for a class of singular diffusions, *Ann. Probab.*, **20** (1992), 312–321.

[5] J. E. Bertram and P. E. Sarachik, Stability of circuits with randomly time-varying parameters, *IRE Trans. Information Theory*, **6** (1959), 260–270.

[6] S. Bochner, Beiträge zur theorie der fastperiodischen funktionen, I. Funktionen einer Variablen, (German), *Math. Ann.*, **96** (1927), 119–147.

[7] S. Bochner, A new approach to almost periodicity, *Proc. Nat. Acad. Sci. U.S.A.*, **48** (1962), 2039–2043.

[8] H. Bohr, Zur theorie der fastperiodischen funktionen, (German) I., *Acta Math.*, **45** (1925), 29–127.

[9] H. Bohr, Zur theorie der fastperiodischen funktionen, (German) II., *Acta Math.*, **46** (1925), 101–214.

[10] H. Bohr, Zur theorie der fastperiodischen funktionen, (German) III., *Acta Math.*, **47** (1926), 237–281.

[11] K. L. Chung, *A Course in Probability Theory*, 3^{nd} edition, Academic press, Inc., San Diego, CA, 2001. xviii+419 pp.

[12] W. A. Coppel, Almost periodic properties of ordinary differential equations, *Ann. Mat. Pura Appl.(4)*, **76** (1967), 27–49.

[13] G. Da Prato and C. Tudor, Periodic and almost periodic solutions for semilinear stochastic equations, *Stochastic Anal. Appl.*, **13** (1995), 13–33.

[14] L. G. Deysach and G. R. Sell, On the existence of almost periodic motion, *Michgan Math. J.*, **12** (1965), 87–95.

[15] A. M. Fink, *Almost Periodic Differential Equations*, Lecture Notes in Mathematics, Vol. 377. Springer-Verlag, Berlin-New York, 1974. viii+336 pp.

[16] A. Halanay, Periodic and almost periodic solutions to affine stochastic systems, in *Proceedings of the Eleventh International Conference on Nonlinear Oscillations*, (Budapest, 1987), 94–101, *János Bolyai Math. Soc.*, Budapest, 1987.

[17] J. Hale, Periodic and almost periodic solutions of functional-differential equations, *Arch. Rational Mech. Anal.*, **15** (1964), 289–304.

[18] I. Ya. Kac and N. N. Krasovskii, On the stability of systems with random parameters, (Russian), *Prikl. Mat. Meh.*, **24** (1960), 809–823. English translation: *J. Appl. Math. Mech.*, **24** (1960), 1225–1246.

[19] R. Khasminskii, On the stability of the trajectory of Markov processes,, *Prikl. Mat. Meh.*, **26** (1962), 1025–1032 (Russian); English translation: *J. Appl. Math. Mech.*, **26** (1962), 1554–1565.

[20] R. Khasminskii, *Stochastic Stability of Differential Equations*, 1^{nd} edition, Sijthoff & Noordhoff, Alphen aan den Rijn-Germantown, Md., 1980. xvi+344 pp; 2^{nd} edition, Springer, Heidelberg, 2012. xviii+339 pp.

[21] F. Kozin, On almost sure stability of linear systems with random coefficients, *J. Math. Phys.*, **42** (1963), 59–67.

[22] H. J. Kushner, *Stochastic Stability and Control*, Mathematics in Science and Engineering, Vol. 33 Academic Press, New York-London, 1967. xiv+161 pp.

[23] B. M. Levitan and V. V. Zhikov, *Almost Periodic Functions and Differential Equations*, Translated from the Russian by L. W. Longdon, Cambridge University Press, Cambridge-New York, 1982. xi+211 pp.

[24] Z. Liu and W. Wang, Favard separation method for almost periodic stochastic differential equations, *J. Differential Equations*, **260** (2016), 8109–8136.

[25] X. Mao, *Exponential Stability of Stochastic Differential Equations*, Monographs and Textbooks in Pure and Applied Mathematics, 182. Marcel Dekker, Inc., New York, 1994. xii+307 pp.

[26] X. Mao, *Stochastic Differential Equations and Their Applications*, Horwood Publishing Series in Mathematics and Applications, Horwood Publishing Limited, Chichester, 1997. xii+366 pp.

[27] A. A. Markov, Stabilität in Liapvanoffschen Sinne und Fastperiodizität, *Math. Z.*, **36** (1933), 708–738.

[28] R. K. Miller, Almost periodic differential equations as dynamical systems with applications to the extence of a. p. solution, *J. Differential Equations*, **1** (1965), 337–345.

[29] T. Morozan and C. Tudor, Almost periodic solutions of affine Itô equations, *Stoch. Anal. Appl.*, **7** (1989), 451–474.

[30] K. R. Parthasarathy, *Probability Measures on Metric Spaces*, Probability and Mathematical Statistics, No. 3 Academic Press, Inc., New York-London, 1967. xi+276 pp.

[31] G. Seifert, Almost periodic solutions for almost periodic systems of ordinary differential equations, *J. Differential Equations*, **2** (1966), 305–319.

[32] G. R. Sell, Nonautonomous differential equations and topological dynamics. I. The basic theory, *Trans. Amer. Math. Soc.*, **127** (1967), 241–262.

[33] G. R. Sell, Nonautonomous differential equations and topological dynamics. II. Limiting equations, *Trans. Amer. Math. Soc.*, **127** (1967), 263–283.

[34] C. Vârsana, Asymptotic almost periodic solutions for stochastic differential equations, *Tôhoku Math. J.*, **41** (1989), 609–618.

[35] T. Yoshizawa, Extreme stability and almost periodic solutions of functional-differential equations, *Arch. Rational Mech. Anal.*, **17** (1964), 148–170.

[36] T. Yoshizawa, Asymptotically almost periodic solutions of an almost periodic system, *Funkcial. Ekvac.*, **12** (1969), 23–40.

[37] T. Yoshizawa, *Stability Theory and the Existence of Periodic Solutions and Almost Periodic Solutions*, Applied Mathematical Sciences, Vol. 14. Springer-Verlag, New York-Heidelberg, 1975. vii+233 pp.

Received for publication September 2018.

E-mail address: liyong@jlu.edu.cn
E-mail address: zxliu@dlut.edu.cn
E-mail address: die_jahre@163.com