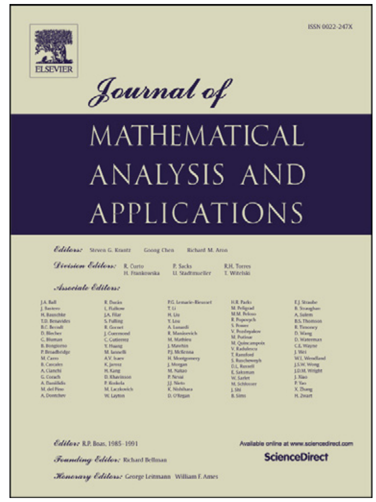
Accepted Manuscript



Almost automorphy and various extensions for stochastic processes

Fazia Bedouhene, Nouredine Challali, Omar Mellah,
Paul Raynaud de Fitte, Mannal Smaali

PII: S0022-247X(15)00334-0
DOI: http://dx.doi.org/10.1016/j.jmaa.2015.04.014
Reference: YJMAA 19387

To appear in: *Journal of Mathematical Analysis and Applications*

Received date: 13 November 2014

Please cite this article in press as: F. Bedouhene et al., Almost automorphy and various extensions for stochastic processes, *J. Math. Anal. Appl.* (2015), http://dx.doi.org/10.1016/j.jmaa.2015.04.014

Almost automorphy and various extensions for stochastic processes

Fazia Bedouhene[*], Nouredine Challali[†], Omar Mellah[‡], Paul Raynaud de Fitte[§], and Mannal Smaali[¶]

April 7, 2015

Abstract

We compare different modes of pseudo almost automorphy and variants for stochastic processes: in probability, in quadratic mean, or in distribution in various senses. We show by a counterexample that square-mean (pseudo) almost automorphy is a property which is too strong for stochastic differential equations (SDEs). Finally, we consider two semilinear SDEs, one with almost automorphic coefficients and the second one with pseudo almost automorphic coefficients, and we prove the existence and uniqueness of a mild solution which is almost automorphic in distribution in the first case, and pseudo almost automorphic in distribution in the second case.

Keywords. Weighted pseudo almost automorphic; square-mean almost automorphic; pseudo almost automorphic in quadratic mean; pseudo almost automorphic in distribution; Stepanov; Weyl; Besicovitch; Ornstein-Uhlenbeck; semilinear stochastic differential equation; stochastic evolution equation

2010 Mathematical Subject Classification. 60G05 (60B05 60B11 60H10 34C27)

1 Introduction

Almost automorphic functions, introduced by Bochner, are an important generalization of almost periodic functions. Almost automorphy is a property of regularity and recurrence of functions, which has been studied in the context of

[*]Mouloud Mammeri University of Tizi-Ouzou, Laboratoire de Mathématiques Pures et Appliquées, Tizi-Ouzou, Algeria E-Mail: fbedouhene@yahoo.fr

[†]Mouloud Mammeri University of Tizi-Ouzou, Laboratoire de Mathématiques Pures et Appliquées, Tizi-Ouzou, Algeria E-Mail: challalin@yahoo.fr

[‡]Mouloud Mammeri University of Tizi-Ouzou, Laboratoire de Mathématiques Pures et Appliquées, Tizi-Ouzou, Algeria E-Mail: omellah@yahoo.fr

[§]Normandie Univ, Laboratoire Raphaël Salem, UMR CNRS 6085, Rouen, France E-Mail: prf@univ-rouen.fr

[¶]Mouloud Mammeri University of Tizi-Ouzou, Laboratoire de Mathématiques Pures et Appliquées, Tizi-Ouzou, Algeria E-Mail: smaali_manel@yahoo.fr

differential equations and dynamical systems, and in other contexts. The question of studying the concept of almost automorphic stochastic processes arises naturally in connection with stochastic differential equations.

As can be seen in Tudor's survey [49], almost periodicity forks into many different notions when applied to stochastic processes: almost periodicity in probability, in p-mean, in one-dimensional distributions, in finite dimensional distributions, in distribution, almost periodicity of moments, etc. These notions are not all comparable: for example, almost periodicity in distribution does not imply almost periodicity in probability, and the converse implication is false too [3]. The situation for almost automorphy is no different.

Furthermore, things become even more complicated if one takes into account different generalizations of almost automorphy (which have their analogue in almost periodicity): on the one hand, changing the mode of convergence in the definition leads to the notions of Stepanov-like, Weyl-like and Besicovitch-like almost automorphy. On the other hand, for the study of asymptotic properties of functions, it is natural to consider functions which are the sum of an almost automorphic function and of a function vanishing in some sense at infinity. In this way, one gets the notions of asymptotically almost automorphic functions and of pseudo almost automorphic functions and their weighted variants.

Each of these notions can be interpreted in different manners for stochastic processes, exactly in the same way as for "plain" almost periodicity and almost automorphy. It is an objective of this paper to clarify the hierarchy of those various concepts. We did not try to list all possible variants, one can imagine many other extensions and combinations, as the reader will probably do. This is rather a preliminary groundwork, in which we investigate some notions we think particularly useful. For example, when we describe the different modes of pseudo almost automorphy in distribution, we concentrate on a stronger notion, which is not purely "distributional" but seems to be the relevant one for stochastic differential equations.

As we have in view applications in spaces of probability measures, which do not have a vector space structure and whose topology can be described by different non uniformly equivalent metrics, we are especially interested in the properties of almost automorphy and pseudo almost automorphy which are purely topological, i.e. which do not depend on a vector structure or on a particular metric, but only on the topology of the underlying space.

A natural application of these concepts is the study of stochastic differential equations with almost automorphic or more general coefficients. We provide two examples of stochastic semilinear evolution equations, with almost automorphic coefficients for the first one, and with pseudo almost automorphic coefficients for the second one, whose unique bounded solution is almost automorphic (respectively pseudo almost automorphic) in distribution. It is another objective of this paper to point out a common error in many papers which claim the existence of nontrivial solutions which are almost automorphic in quadratic mean. We show by a counterexample borrowed from [38] that this claim is false, even for several extensions of almost automorphy.

Historical comments These comments are not intended to provide a full historical account, only to highlight some steps in the history of almost automorphic stochastic processes and their generalizations.

The study of almost periodic random functions has a relatively long story, starting from Slutsky [47] in 1938, who focused on conditions for weakly stationary random processes to have almost periodic trajectories in Besicovitch's sense.

The investigation of almost periodicity in probability was initiated later by the Romanian school [44, 16, 45].

It is only in the late eighties that almost periodicity in distribution was considered, again by the Romanian school (mainly by Constantin Tudor), in connection with the study of stochastic diferential equations with almost periodic coefficients [29, 30, 39, 48, 2, 17].

Starting from 2007, many papers appeared, claiming the existence of square-mean almost periodic solutions to almost periodic semilinear stochastic evolution equations, using a fixed point method. Despite the counterexamples given in [37, 38], new papers in this vein continue to be published.

The story of almost automorphy and its generalizations is much shorter. Almost automorphic functions were invented by Bochner since 1955 [13, 11, 12] (the terminology stems from the fact that they were first encountered in [13] in the context of differential geometry on real or complex manifolds). Almost automorphic stochastic processes and their generalizations seem to have been investigated only since 2010, starting with [28], which was followed by many other papers. Most of these papers claim almost automorphy (or one of its generalizations) in square mean for solutions to stochastic equations. There are only few papers we are aware of [26, 27, 35] which investigate almost automorphy in a distributional sense.

Recently, the notion of almost automorphic random functions in probability has been introduced by Ding, Deng and N'Guérékata [23].

Organization of the paper In Section 2, we present the concept of almost automorphy and some of its generalizations: μ-pseudo almost automorphy, Stepanov-like, Weyl-like and Besicovitch-like μ-pseudo almost automorphy. Our setting is that of functions of a real variable with values in a metrizable space. Metrizability seems a sufficiently general frame to investigate almost automorphy in many useful spaces of probability measures, while avoiding complications. An extension to uniformizable spaces would be useful for applications in locally convex vector spaces, this could be done using projective limits of metrizable spaces as in [3]. We show that almost automorphy and a slightly generalized notion of pseudo almost automorphy can be defined in a topological way, without any reference to a metric nor to a vector structure.

In Section 3, we investigate several notions of almost automorphy and pseudo almost automorphy for stochastic processes. First, we investigate almost automorphy and its variants in pth mean: the stochastic processes are seen as almost automorphic (or, more generally, μ-pseudo almost automorphic) functions from

$\mathbb{R}$ to L^p, $p \geq 0$ ($p = 0$ corresponds to almost automorphy and its variants in probability). We show with the simple counterexample of Ornstein-Uhlenbeck process that even a one-dimensional linear equation with constant coefficients has no nontrivial solution which is almost automorphic (in any of the variants considered) in pth mean.

Then we move to almost automorphy in distribution and its variants. There are at least three kinds of almost automorphy in distribution: in one-dimensional distributions, in finite dimensional distributions, and in distribution of the whole process. In the deterministic case, the first two notions are equivalent to almost automorphy, the third one is equivalent to compact automorphy, a stronger notion. For $p > 0$, we introduce also the notion of almost automorphy in p-distribution, which is obtained by adding to the preceding notions a condition of p-uniform integrability. For μ-pseudo almost automorphy, the situation becomes even more complicated, because there are several ways to take into account the ergodic part. We introduce the notion of processes which are μ-pseudo almost automorphic in p-distribution, which are the sum of a process which is almost automorphic in p-distribution and a process which is μ-ergodic in pth mean. We use this notion in the next section. We do not address the notions of Stepanov-like, Weyl-like or Besicovitch-like (pseudo) almost automorphy for stochastic processes, these notions would probably have to be linked to a particular choice of a metric on a space of probability measures.

We study the superposition operator (also called Nemytskii operator) between spaces of processes which are almost periodic (compact almost automorphic, and μ-pseudo compact almost automorphic) in distribution.

We also carry out a comparison of the main notions of (generalized) almost automorphy for stochastic processes: in probability, in pth mean, and in p-distribution.

Finally, in Section 4, we consider two semilinear stochastic evolution equations in a Hilbert space. The first one has almost automorphic coefficients, and the second one has μ-pseudo almost automorphic coefficients. We show that each equation has a unique mild solution which is almost automorphic in 2-distribution in the first case, and μ-pseudo almost automorphic in 2-distribution in the second case.

2 Weighted pseudo almost automorphy in Banach spaces and in metric spaces

2.1 Notations and definitions

In the sequel, $\mathbb{X}$ and $\mathbb{Y}$ are metrizable topological spaces. When no confusion may arise, we denote by $\mathfrak{d}$ a distance on $\mathbb{X}$ (respectively on $\mathbb{Y}$) which generates the topology of $\mathbb{X}$ (respectively $\mathbb{Y}$). Most of our results depend only on the topology of those spaces, not on the choice of particular metrics. When $\mathbb{X}$ and $\mathbb{Y}$ are Banach spaces, their norms are indistinctly denoted by $\|.\|$, and $\mathfrak{d}$ is assumed to result from $\|.\|$.

We denote by $\mathrm{C}(\mathbb{X},\mathbb{Y})$ the space of continuous functions from $\mathbb{X}$ to $\mathbb{Y}$. When this space is endowed with the topology of uniform convergence on compact subsets of $\mathbb{X}$, it is denoted by $\mathrm{C}_k(\mathbb{X},\mathbb{Y})$.

For a continuous function $f:\mathbb{R}\to\mathbb{X}$, we define its *translation mapping*

$$\tilde{f}:\left\{\begin{array}{lcl}\mathbb{R} & \to & \mathrm{C}(\mathbb{R},\mathbb{X})\\ t & \to & f(t+.).\end{array}\right.$$

2.2 Almost periodicity and almost automorphy

Almost periodicity We say that a continuous function $f:\mathbb{R}\to\mathbb{X}$ is *almost periodic* if, for any $\varepsilon>0$, there exists $l(\varepsilon)>0$ such that any interval of length $l(\varepsilon)$ contains at least an ε*-almost period*, that is, a number τ for which

$$\mathfrak{d}\left(f(t+\tau),f(t)\right)\leq\varepsilon,\ \text{ for all } t\in\mathbb{R}.$$

We denote by $\mathrm{AP}(\mathbb{R},\mathbb{X})$ the space of $\mathbb{X}$-valued almost periodic functions.

By a result of Bochner [12], $f:\mathbb{R}\to\mathbb{X}$ is almost periodic if, and only if, the set $\left\{\tilde{f}(t),\,t\in\mathbb{R}\right\}=\{f(t+.),\,t\in\mathbb{R}\}$ is totally bounded in the space $\mathrm{C}(\mathbb{R},\mathbb{X})$ endowed with the norm $\|.\|_\infty$ of uniform convergence.

Another very useful characterization is *Bochner's double sequence criterion* [12]: f is almost periodic if, and only if, it is continuous and, for every pair of sequences (t_n') and (s_n') in $\mathbb{R}$, there are subsequences (t_n) of (t_n') and (s_n) of (s_n') respectively, with same indexes, such that, for every $t\in\mathbb{R}$, the limits

$$\lim_{n\to\infty}\lim_{m\to\infty}f(t+t_n+s_m)\text{ and }\lim_{n\to\infty}f(t+t_n+s_n),\tag{2.1}$$

exist and are equal. This very useful criterion shows that the set $\mathrm{AP}(\mathbb{R},\mathbb{X})$ depends only on the topology of $\mathbb{X}$, i.e. it does not depend on any uniform structure on $\mathbb{X}$, in particular it does not depend on the choice of any norm (if $\mathbb{X}$ is a vector space) or any distance on $\mathbb{X}$.

Almost automorphy Almost automorphic functions were introduced by Bochner [12] and studied in depth by Veech [51], see also the monographs [46, 40, 42] for applications to differential equations. A continuous mapping $f:\mathbb{R}\to\mathbb{X}$ is said to be *almost automorphic* if, for every sequence (t_n') in $\mathbb{R}$, there exists a subsequence (t_n) such that, for every $t\in\mathbb{R}$, the limit

$$g(t)=\lim_{n\to\infty}f(t+t_n)\tag{2.2}$$

exists and

$$\lim_{n\to\infty}g(t-t_n)=f(t).\tag{2.3}$$

The range R_f of f is then relatively compact, because we can extract from every sequence $(f(t_n))$ in R_f a convergent subsequence.

Clearly, the space of almost automorphic $\mathbb{X}$-valued functions depends only on the topology of $\mathbb{X}$.

Almost automorphic functions generalize almost periodic functions in the sense that f is almost periodic if, and only if, the above limits are uniform with respect to t.

Note that, in (2.2) and (2.3), the limit function g is not necessarily continuous. Let us consider the following property:

(**C**) For any choice of (t'_n) and (t_n), the function g of (2.2) and (2.3) is continuous.

Functions satisfying (**C**) are called continuous almost automorphic functions in [51]. They were re-introduced by Fink [25] under the name of *compact almost automorphic* functions. This terminology is now generally adopted, so we stick to it.

It has been shown by Veech in [51, Lemma 4.1.1] (see also [41, Theorem 2.6] and [43, 36]) that, if f satifies (**C**), f is uniformly continuous. The proof of Veech is given in the case when $\mathbb{X}$ is the field of complex numbers, but it extends to any metric space. Furthermore, f satifies (**C**) if, and only if, the convergence in (2.2) and (2.3) is uniform on the compact intervals. We denote by $\mathrm{AA}_c(\mathbb{R}, \mathbb{X})$ the subspace of functions satisfying (**C**).

We have the inclusions

$$\mathrm{AP}(\mathbb{R}, \mathbb{X}) \subset \mathrm{AA}_c(\mathbb{R}, \mathbb{X}) \subset \mathrm{AA}(\mathbb{R}, \mathbb{X}).$$

All these spaces depend only on the topological structure of $\mathbb{X}$ and not on its metric.

Almost automorphic functions depending on a parameter Following [33], we say that a function $f : \mathbb{R} \times \mathbb{Y} \to \mathbb{X}$ is *almost automorphic with respect to the first variable, uniformly with respect to the second variable in bounded subsets of* $\mathbb{Y}$ (respectively *in compact subsets of* $\mathbb{Y}$) if, for every sequence (t'_n) in $\mathbb{R}$, there exists a subsequence (t_n) such that, for every $t \in \mathbb{R}$ and every $y \in \mathbb{Y}$, the limit

$$g(t, y) = \lim_{n\to\infty} f(t + t_n, y)$$

exists and, for every bounded (respectively compact) subset B of $\mathbb{Y}$, the convergence is uniform with respect to $y \in B$, and if the convergence

$$\lim_{n\to\infty} g(t - t_n, y) = f(t, y)$$

holds uniformly with respect to $y \in B$. We denote by $\mathrm{AAU}_b(\mathbb{R} \times \mathbb{Y}, \mathbb{X})$ and $\mathrm{AAU}_c(\mathbb{R} \times \mathbb{Y}, \mathbb{X})$ respectively the spaces of such functions.

Similarly, one can define the spaces of functions $f : \mathbb{R} \times \mathbb{Y} \to \mathbb{X}$ which are *compact almost automorphic with respect to the first variable, uniformly with respect to the second variable in bounded* (or *in compact*) *subsets of* $\mathbb{Y}$. We denote these spaces by $\mathrm{AA}_c\mathrm{U}_c(\mathbb{R} \times \mathbb{Y}, \mathbb{X})$ and $\mathrm{AA}_c\mathrm{U}_b(\mathbb{R} \times \mathbb{Y}, \mathbb{X})$ respectively.

These notions are different from the notion of functions almost automorphic uniformly in y defined in [9, 7].

Proposition 2.1 *Let $f \in AA_cU_c(\mathbb{R}\times\mathbb{Y},\mathbb{X})$. Assume that f is continuous with respect to the second variable. Then f is continous on $\mathbb{R}\times\mathbb{Y}$, and, for every compact subset K of $\mathbb{Y}$, f is uniformly continuous on $\mathbb{R}\times K$.*

Proof For simplicity, we use the same notation $\mathfrak{d}$ for distances on $\mathbb{Y}$ and $\mathbb{X}$ which generate the topologies of $\mathbb{Y}$ and $\mathbb{X}$ respectively.

First step Let us show that f is jointly continuous. Let $(t,x)\in\mathbb{R}\times\mathbb{Y}$, and let (t_n,x_n) be a sequence in $\mathbb{R}\times\mathbb{Y}$ which converges to (t,x). Let $\varepsilon>0$. The set $K=\{x_n;\, n\in\mathbb{N}\}\cup\{x\}$ is compact, thus there exists $N_1\in\mathbb{N}$ such that, for any $y\in K$,

$$n,m\geq N_1 \Rightarrow \mathfrak{d}(f(t_n,y),f(t_m,y))<\varepsilon/3.$$

Now, there exists $N_2\in\mathbb{N}$ such that

$$n\geq N_2 \Rightarrow \mathfrak{d}(f(t_{N_1},x_n),f(t_{N_1},x))<\varepsilon/3.$$

We deduce, for $n\geq (N_1\vee N_2)$,

$$\begin{aligned}\mathfrak{d}(f(t_n,x_n),f(t,x)) &\leq \mathfrak{d}(f(t_n,x_n),f(t_{N_1},x_n))\\ &\quad+\mathfrak{d}(f(t_{N_1},x_n),f(t_{N_1},x))+\mathfrak{d}(f(t_{N_1},x),f(t,x))<\varepsilon/3+\varepsilon/3+\varepsilon/3=\varepsilon,\end{aligned}$$

which proves the continuity of f.

Second step Let (t'_n) be a sequence in $\mathbb{R}$. Let (t_n) be a subsequence of (t'_n) such that, for every $y\in\mathbb{Y}$, and for every $t\in\mathbb{R}$, the limit

$$g(t,y)=\lim_{n\to\infty} f(t+t_n,y)$$

exists, uniformly with respect to y in compact subsets of $\mathbb{Y}$, and

$$\lim_{n\to\infty} g(t-t_n,y)=f(t,y).$$

By our hypothesis, for each $y\in\mathbb{Y}$, the function $g(.,y)$ is continuous. A similar reasoning to that of the first step shows that g is continuous on $\mathbb{R}\times\mathbb{Y}$. Indeed, let $(t,x)\in\mathbb{R}\times\mathbb{Y}$, and let (s_k,x_k) be a sequence in $\mathbb{R}\times\mathbb{Y}$ such that $(t+s_k,x_k)$ converges to (t,x). Let $\varepsilon>0$. The set $K=\{x_n;\, n\in\mathbb{N}\}\cup\{x\}$ is compact, thus there exists an integer N such that, for every $y\in K$,

$$n\geq N \Rightarrow \mathfrak{d}\Big(g(t,y),f(t+t_n,y)\Big)<\varepsilon/3.$$

By continuity of f at the point $(t+t_N,x)$, there exists $N'\in\mathbb{N}$ such that

$$k\geq N' \Rightarrow \mathfrak{d}\Big(f(t+s_k+t_N,x_k),f(t+t_N,x)\Big)<\varepsilon/3.$$

We have thus, for $k\geq N'$,

$$\mathfrak{d}\Big(g(t+s_k,x_k),g(t,x)\Big)\leq \mathfrak{d}\Big(g(t+s_k,x_k),f(t+s_k+t_N,x_k)\Big)$$

$$
\begin{aligned}
&+ \mathfrak{d}\Big(f(t+s_k+t_N,x_k), f(t+t_N,x)\Big) \\
&+ \mathfrak{d}\Big(f(t+t_N,x), g(t,x)\Big) \\
&< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon.
\end{aligned}
$$

Third step Let K be a compact subset of $\mathbb{Y}$. Assume that f is not uniformly continuous on $\mathbb{R}\times K$. We can find two sequences (s_n, x_n) and (t_n, y_n) in $\mathbb{R}\times K$ such that $(s_n - t_n) + \mathfrak{d}(x_n, y_n)$ converges to 0 and $\mathfrak{d}\big(f(s_n, x_n), f(t_n, y_n)\big) > 2\delta$ for some $\delta > 0$ and for all $n \in \mathbb{N}$. By compactnes of K, and extracting if necessary a subsequence, we can assume that (x_n) and (y_n) converge to a common limit $x \in K$. We have thus

$$
\begin{aligned}
\liminf_{n\to\infty} \mathfrak{d}\Big(f(s_n,x_n), f(s_n,x)\Big) &+ \liminf_{n\to\infty} \mathfrak{d}\Big(f(t_n,y_n), f(t_n,x)\Big) \\
&\geq \liminf_{n\to\infty} \mathfrak{d}\Big(f(s_n,x_n), f(t_n,y_n)\Big) > 2\delta,
\end{aligned}
$$

which implies that at least one term in the left hand side is greater than δ. So, we can assume, without loss of generality, that

$$
\liminf_{n\to\infty} \mathfrak{d}\Big(f(t_n,y_n), f(t_n,x)\Big) > \delta. \tag{2.4}
$$

Extracting if necessary a further subsequence, we can assume also that there exists a function $g : \mathbb{R}\times\mathbb{Y} \to \mathbb{X}$ such that

$$
\lim_{n\to\infty} f(t_n, y) = g(0, y)
$$

uniformly with respect to $y \in K$. We have proved in the second step that g is continuous. But, then, we have

$$
\begin{aligned}
&\limsup_{n\to\infty} \mathfrak{d}\Big(f(t_n,y_n), f(t_n,x)\Big) \\
&\quad \leq \limsup_{n\to\infty} \bigg(\mathfrak{d}\Big(f(t_n,y_n), g(0,y_n)\Big) \\
&\qquad + \mathfrak{d}\Big(g(0,y_n), g(0,x)\Big) + \mathfrak{d}\Big(g(0,x), f(t_n,x)\Big)\bigg) = 0,
\end{aligned}
$$

which contradicts (2.4). ∎

2.3 (Weighted) pseudo almost automorphy

Pseudo almost periodic functions were invented by Zhang [54, 55, 56, 57]. The generalization of this concept to pseudo almost automorphic functions was investigated in [33]. To define pseudo almost automorphy, we need another class of functions. Assume for the moment that $\mathbb{X}$ is a Banach space. Let

$$
\mathcal{E}(\mathbb{R},\mathbb{X}) = \left\{ f \in \mathrm{BC}(\mathbb{R},\mathbb{X});\ \lim_{r\to\infty} \frac{1}{2r}\int_{[-r,r]} \|f(t)\|\, dt = 0 \right\},
$$

where $\mathrm{BC}(\mathbb{R},\mathbb{X})$ denotes the space of bounded continuous functions from $\mathbb{R}$ to $\mathbb{X}$. We say that a continuous function $f:\mathbb{R}\to\mathbb{X}$ is *pseudo almost automorphic* if it has the form

$$f = g + \Phi, \quad g \in \mathrm{AA}(\mathbb{R},\mathbb{X}), \quad \Phi \in \mathcal{E}(\mathbb{R},\mathbb{X}). \tag{2.5}$$

The space of $\mathbb{X}$-valued pseudo almost automorphic functions is denoted by $\mathrm{PAA}(\mathbb{R},\mathbb{X})$.

Weighted pseudo almost automorphic functions were introduced by Blot et al. in [6] and later generalized in [7]. They generalize the weighted pseudo almost periodic functions introduced by Diagana [18, 19, 20], see also [8]. Let μ be a Borel measure on $\mathbb{R}$ such that

$$\mu(\mathbb{R}) = \infty \text{ and } \mu(I) < \infty \text{ for every bounded interval } I. \tag{2.6}$$

We define the space $\mathcal{E}(\mathbb{R},\mathbb{X},\mu)$ of μ-*ergodic* $\mathbb{X}$-valued functions by

$$\mathcal{E}(\mathbb{R},\mathbb{X},\mu) = \left\{ f \in \mathrm{BC}(\mathbb{R},\mathbb{X});\ \lim_{r\to\infty} \frac{1}{\mu([-r,r])} \int_{[-r,r]} \|f(t)\|\, d\mu(t) = 0 \right\}.$$

The space $\mathrm{PAA}(\mathbb{R},\mathbb{X},\mu)$ of μ-*pseudo almost automorphic functions* with values in $\mathbb{X}$ is the space of continuous functions $f:\mathbb{R}\to\mathbb{X}$ of the form

$$f = g + \Phi, \quad g \in \mathrm{AA}(\mathbb{R},\mathbb{X}), \quad \Phi \in \mathcal{E}(\mathbb{R},\mathbb{X},\mu). \tag{2.7}$$

The space $\mathrm{PAA}(\mathbb{R},\mathbb{X},\mu)$ contains the *asymptotically almost automorphic* functions, that is, the functions of the form

$$f = g + \Phi, \quad g \in \mathrm{AA}(\mathbb{R},\mathbb{X}), \quad \lim_{|t|\to\infty} \|\Phi(t)\| = 0,$$

see [7, Corollary 2.16].

Note that, contrarily to (2.5), the decomposition (2.7) is not necessarily unique [7, Remark 4.4 and Theorem 4.7], even in the case of weighted almost periodic functions [32, 8]. A sufficient condition of uniqueness of the decomposition is that $\mathcal{E}(\mathbb{R},\mathbb{X},\mu)$ be translation invariant. This is the case in particular if Condition (**H**) of [7] is satisfied:

(**H**) For every $\tau \in \mathbb{R}$, there exist $\beta > 0$ and a bounded interval I such that $\mu(A+\tau) \le \beta\mu(A)$ whenever A is a Borel subset of $\mathbb{R}$ such that $A \cap I = \emptyset$.

The following elementary lemma will prove useful.

Lemma 2.2 *Let $f \in \mathcal{E}(\mathbb{R},\mathbb{X},\mu)$, with μ satisfying* (2.6). *There exists a sequence (t_n) in $\mathbb{R}$ such that $(|t_n|)$ converges to ∞ and $(f(t_n))$ converges to 0. If furthermore*

$$\liminf_{r\to\infty} \frac{\mu([0,r])}{\mu([-r,r])} > 0, \tag{2.8}$$

one can choose (t_n) converging to $+\infty$.

Proof Observe first that, for every bounded interval I, the function f satisfies

$$\lim_{r\to\infty} \frac{1}{\mu([-r,r]\setminus I)} \int_{[-r,r]\setminus I} \|f(t)\|\, d\mu(t) = 0$$

(see [7, Theorem 2.14] for a stronger result). Assume that the first part of the lemma is false. There exist $\varepsilon > 0$ and $R > 0$ such that, for $|t| \geq R$, $\|f(t)\| \geq \varepsilon$. Then we have

$$\begin{aligned} 0 = \lim_{r\to\infty} \frac{1}{\mu([-r,r]\setminus[-R,R])} \int_{[-r,r]\setminus[-R,R]} \|f(t)\|\, d\mu(t) \\ \geq \lim_{r\to\infty} \varepsilon \frac{\mu([-r,r]\setminus[-R,R])}{\mu([-r,r]\setminus[-R,R])} = \varepsilon, \end{aligned}$$

a contradiction.

In the case when (2.8) is satisfied, we have

$$\begin{aligned} &\limsup_{r\to\infty} \frac{1}{\mu([0,r]\setminus[0,R])} \int_{[0,r]\setminus[0,R]} \|f(t)\|\, d\mu(t) \\ &\leq \limsup_{r\to\infty} \frac{\mu([-r,r]\setminus[-R,R])}{\mu([0,r]\setminus[0,R])} \frac{1}{\mu([-r,r]\setminus[-R,R])} \int_{[-r,r]\setminus[-R,R]} \|f(t)\|\, d\mu(t) \\ &= \limsup_{r\to\infty} \frac{\mu([-r,r])}{\mu([0,r])} \frac{1}{\mu([-r,r]\setminus[-R,R])} \int_{[-r,r]\setminus[-R,R]} \|f(t)\|\, d\mu(t) \leq 0. \end{aligned}$$

Then we only need to reproduce the reasoning of the first part of the lemma, replacing $[-r,r]$ by $[0,r]$. ■

Pseudo almost automorphic functions depending on a parameter Let μ be a Borel measure on $\mathbb{R}$ satisfying (2.6). We say that a continuous function $f : \mathbb{R}\times\mathbb{Y} \to \mathbb{X}$ is *μ-ergodic with respect to the first variable, uniformly with respect to the second variable in bounded subsets of* $\mathbb{Y}$ (respectively *in compact subsets of* $\mathbb{Y}$) if, for every $x \in \mathbb{Y}$, $f(.,x)$ is μ-ergodic, and the convergence of $1/\mu([-r,r]) \int_{-r}^{r} \|f(t,x)\|\, d\mu(t)$ is uniform with respect to x in bounded subsets of $\mathbb{Y}$ (respectively compact subsets of $\mathbb{Y}$). The space of such functions is denoted by $\mathcal{E}\mathrm{U}_b(\mathbb{R}\times\mathbb{Y},\mathbb{X},\mu)$ (respectively $\mathcal{E}\mathrm{U}_c(\mathbb{R}\times\mathbb{Y},\mathbb{X},\mu)$).

Remark 2.3 If for each $x \in \mathbb{Y}$, $f(.,x) \in \mathcal{E}(\mathbb{R},\mathbb{X},\mu)$ and $f(t,x)$ is continuous with respect to x, uniformly with respect to t, then $f \in \mathcal{E}\mathrm{U}_c(\mathbb{R}\times\mathbb{Y},\mathbb{X},\mu)$.

Indeed, let K be a compact subset of $\mathbb{Y}$, and let $\varepsilon > 0$. Let $\mathfrak{d}$ be any distance on $\mathbb{Y}$ which generates the topology of $\mathbb{Y}$. There exists $\eta > 0$ such that, for all $x, y \in K$ satisfying $\mathfrak{d}\,(x,y) < \eta$, we have $\|f(t,x) - f(t,y)\| < \varepsilon$ for every $t \in \mathbb{R}$. Let $x_1,\dots,x_m$ be a finite sequence in K such that $K \subset \cup_{i=1}^{m} B(x_i,\eta)$. We have, for every $r > 0$,

$$\sup_{x\in K} \frac{1}{\mu([-r,r])} \int_{-r}^{r} \|f(t,x)\|\, d\mu(t)$$

$$\begin{aligned}
&\leq \max_{1\leq i\leq m} \sup_{x\in B(x_i,\eta)} \left(\frac{1}{\mu([-r,r])} \int_{-r}^{r} \|f(t,x)-f(t,x_i)\| \, d\mu(t) \right. \\
&\qquad\qquad\qquad \left. + \frac{1}{\mu([-r,r])} \int_{-r}^{r} \|f(t,x_i)\| \, d\mu(t) \right) \\
&\leq \varepsilon + \max_{1\leq i\leq m} \frac{1}{\mu([-r,r])} \int_{-r}^{r} \|f(t,x_i)\| \, d\mu(t),
\end{aligned}$$

which shows that $f \in \mathcal{E}\mathrm{U}_c(\mathbb{R}\times\mathbb{Y},\mathbb{X},\mu)$.

We say that a continuous function $f : \mathbb{R}\times\mathbb{Y} \to \mathbb{X}$ is *μ-pseudo almost automorphic with respect to the first variable, uniformly with respect to the second variable in bounded subsets of* $\mathbb{Y}$ (respectively *in compact subsets of* $\mathbb{Y}$) if it has the form

$$\begin{aligned}
f = g + \Phi, \quad & g \in \mathrm{AAU}_b(\mathbb{R}\times\mathbb{Y},\mathbb{X}), \quad \Phi \in \mathcal{E}\mathrm{U}_b(\mathbb{R}\times\mathbb{Y},\mathbb{X},\mu) \\
\text{(respectively } f = g + \Phi, \quad & g \in \mathrm{AAU}_c(\mathbb{R}\times\mathbb{Y},\mathbb{X}), \quad \Phi \in \mathcal{E}\mathrm{U}_c(\mathbb{R}\times\mathbb{Y},\mathbb{X},\mu)).
\end{aligned}$$

The space of such functions is denoted by $\mathrm{PAAU}_b(\mathbb{R}\times\mathbb{Y},\mathbb{X},\mu)$ (respectively $\mathrm{PAAU}_c(\mathbb{R}\times\mathbb{Y},\mathbb{X},\mu)$).

2.4 Stepanov, Weyl and Besicovitch-like pseudo almost automorphy

Stepanov-like pseudo almost automorphy and variants The notion of Stepanov-like almost automorphy was proposed by Casarino in [15]. Then Stepanov-like pseudo almost automorphy was first studied by Diagana [21]. Stepanov-like weighted pseudo almost automorphy seems to have been investigated first and simultaneously in [53] and [58].

Let $p > 0$. We say that a locally p-integrable function $f : \mathbb{R} \to \mathbb{X}$ is *$\mathbb{S}^p$-almost automorphic*, or *Stepanov-like almost automorphic* if, for every sequence (t_n) in $\mathbb{R}$, there exists a subsequence (t'_n) and a locally p-integrable function $g : \mathbb{R} \to \mathbb{R}$ such that

$$\lim_{n\to\infty} \|f(t+t'_n) - g(t)\|_{\mathbb{S}^p}$$

and

$$\lim_{n\to\infty} \|g(t-t'_n) - f(t)\|_{\mathbb{S}^p},$$

where, for any locally p-integrable function $h : \mathbb{R} \to \mathbb{X}$,

$$\|h\|_{\mathbb{S}^p} = \sup_{x\in\mathbb{R}} \left(\int_x^{x+1} \|h(t)\|^p \, dt \right)^{1/p}.$$

The space of $\mathbb{S}^p$-almost automorphic $\mathbb{X}$-valued functions is denoted by $\mathrm{AA}_{\mathbb{S}^p}(\mathbb{R},\mathbb{X})$.

The *Bochner transform*[1] of a function $f : \mathbb{R} \to \mathbb{X}$ is the function

$$f^b : \left\{ \begin{array}{lcl} \mathbb{R} & \to & \mathbb{X}^{[0,1]} \\ t & \mapsto & f(t+.). \end{array} \right.$$

[1] The terminology is due to the fact that Bochner was the first to use this transform, for Stepanov almost periodicity, in [10].

We have
$$\mathrm{AA}_{\mathbb{S}^p}(\mathbb{R},\mathbb{X}) = \left\{f;\ f^b \in \mathrm{AA}(\mathbb{R}, \mathrm{L}^p([0,1], dt, \mathbb{X}))\right\}.$$
We define the Stepanov-like μ-ergodic functions in a similar way:
$$\mathcal{E}_{\mathbb{S}^p}(\mathbb{R},\mathbb{X},\mu) = \left\{f;\ f^b \in \mathcal{E}(\mathbb{R}, \mathrm{L}^p([0,1], dt, \mathbb{X}), \mu)\right\}.$$

Let μ be a Borel measure on $\mathbb{R}$ satisfying (2.6). We say that $f : \mathbb{R} \to \mathbb{X}$ is $\mathbb{S}^p$*-pseudo almost automorphic*, or *Stepanov-like weighted pseudo almost automorphic* if f has the form
$$f = g + \Phi, \quad g \in \mathrm{AA}_{\mathbb{S}^p}(\mathbb{R},\mathbb{X}), \quad \Phi \in \mathcal{E}_{\mathbb{S}^p}(\mathbb{R},\mathbb{X},\mu).$$

A further extension has been imagined by Diagana [22]: it consists in adding a weight in the Stepanov norm $\|.\|_{\mathbb{S}^p}$. Let p and μ as before, and let ν be a Borel measure on the interval $[0,1]$ such that
$$0 < \nu([0,1]) < +\infty. \tag{2.9}$$
Set
$$\|h\|_{\mathbb{S}^p_\nu} = \sup_{x\in\mathbb{R}} \left(\int_0^1 \|h(x+t)\|^p \, d\nu(t) \right)^{1/p}.$$
Then, by replacing $\|.\|_{\mathbb{S}^p}$ by $\|.\|_{\mathbb{S}^p_\nu}$, one defines in the obvious way the space $\mathrm{AA}_{\mathbb{S}^p_\nu}(\mathbb{R},\mathbb{X})$ of $\mathbb{S}^p_\nu$*-almost automorphic* $\mathbb{X}$-valued functions, the space $\mathcal{E}_{\mathbb{S}^p_\nu}(\mathbb{R}, \mathrm{L}^p(\Omega, \mathrm{P}, \mathbb{R}), \mu)$ of $\mathbb{S}^p_\nu$*-μ-ergodic functions*, and the space $\mathrm{PAA}_{\mathbb{S}^p_\nu}(\mathbb{R},\mathbb{X})$ of $\mathbb{S}^p_\nu$*-pseudo almost automorphic* functions.

Weyl-like and Besicovitch-like pseudo almost automorphy The concept of Weyl-like pseudo almost automorphy has been recently explored by Abbas [1]. The definition is similar to that of Stepanov-like pseudo almost automorphy, replacing $\|.\|_{\mathbb{S}^p}$ by the weaker seminorm $\|.\|_{\mathbb{W}^p}$, defined by
$$\|h\|_{\mathbb{W}^p} = \lim_{r\to+\infty} \sup_{x\in\mathbb{R}} \left(\frac{1}{2r} \int_{x-r}^{x+r} \|h(t)\|^p \, dt \right)^{1/p}.$$
A further weakening leads to the Besicovitch seminorm, which does not seem to have been investigated in the context of almost automorphy:
$$\|h\|_{\mathbb{B}^p} = \limsup_{r\to+\infty} \left(\frac{1}{2r} \int_{-r}^{r} \|h(t)\|^p \, dt \right)^{1/p}.$$
We shall briefly consider this seminorm in Example 3.1.

2.5 Weighted pseudo almost automorphy in topological spaces

We have seen that, to define the space of almost automorphic functions, as well as that of almost periodic functions, on a space $\mathbb{X}$, there is no need to assume that $\mathbb{X}$ is a vector space, nor a metric space, these spaces depend only on the topological structure of $\mathbb{X}$. We prove in this section that the definition of $\mathrm{PAA}(\mathbb{R}, \mathbb{X}, \mu)$ is metric (independent of the vector structure of $\mathbb{X}$ but dependent on the metric), and that it can be made purely topological if one allows for a slight change of the definition. This leads to two topological concepts of μ-pseudo almost automorphy: in Tudor and Tudor's sense, and in the wide sense.

In this section, unless otherwise stated, $\mathbb{X}$ is only assumed to be a topological space, not necessarily metrizable.

Remark 2.4 Assume that $\mathbb{X}$ is a Banach space. By definition, the space $\mathcal{E}(\mathbb{R}, \mathbb{X}, \mu)$ consists of continuous functions $f : \mathbb{R} \to \mathbb{X}$ satisfying

(i) f is bounded,

(ii) $\displaystyle\lim_{r\to\infty} \frac{1}{\mu([-r,r])} \int_{[-r,r]} \|f(t)\| \, d\mu(t) = 0.$

Condition (i) is of metric nature. By [7, Theorem 2.14], Condition (ii) implies Condition (2.10) below, and the converse implication is true if we assume (i) :

$$\text{For any } \varepsilon > 0, \; \lim_{r\to\infty} \frac{\mu\{t \in [-r,r];\, \|f(t)\| > \varepsilon\}}{\mu([-r,r])} = 0. \tag{2.10}$$

Thus f satisfies (2.10) if, and only if,

$$\lim_{r\to\infty} \frac{1}{\mu([-r,r])} \int_{[-r,r]} \Big(\|f(t)\| \wedge 1 \Big) \, d\mu(t) = 0.$$

Condition (2.10) can be reformulated as

$$\text{For any neighbourhood } U \text{ of } 0, \; \lim_{r\to\infty} \frac{\mu\{t \in [-r,r];\, f(t) \notin U\}}{\mu([-r,r])} = 0. \tag{2.11}$$

The vector structure of $\mathbb{X}$ is still involved in (2.11) through the vector 0. But we use $\mathcal{E}(\mathbb{R}, \mathbb{X}, \mu)$ only in order to ensure that a function is close in a certain sense to $\mathrm{AA}(\mathbb{R}, \mathbb{X})$. To that end, as $\mathrm{AA}(\mathbb{R}, \mathbb{X})$ contains the constant functions, we can allow a generalization of $\mathcal{E}(\mathbb{R}, \mathbb{X}, \mu)$ by replacing 0 in (2.11) by any other fixed point x_0 of $\mathbb{X}$.[2]

An elegant metric definition of μ-pseudo almost periodic functions has been proposed by Constantin and Maria Tudor in [50]. If we adapt their definition, a continuous function $f : \mathbb{R} \to \mathbb{X}$ is *μ-pseudo almost automorphic in Tudor and Tudor's sense* if f has relatively compact range and there exists $g \in \mathrm{AA}(\mathbb{R}, \mathbb{X})$ such that the function $t \mapsto \mathfrak{d}(f(t), g(t))$ is in $\mathcal{E}(\mathbb{R}, \mathbb{R}, \mu)$. This definition is more restrictive than the standard one because the metric condition (i) is replaced by

[2] Actually, the terminology "ergodic" is misleading but it has the advantage of shortness. It would be more appropriate to follow Zhang's terminology (e.g. [55, 57]), and call *μ-ergodic perturbations* the elements of $\mathcal{E}(\mathbb{R}, \mathbb{X}, \mu)$, and call *ergodic* the functions f such that $1/\mu([-r,r]) \int_{-r}^{r} f(t)\, d\mu(t)$ converges to some limit, not necessarily 0.

the stronger topological condition that f have relatively compact range (note that a subset A of $\mathbb{X}$ is relatively compact if, and only if, it is bounded for *every* metric which generates the topology of $\mathbb{X}$, see [24, Problem 4.3.E.(c)]).

Let us say that a continuous function $f : \mathbb{R} \to \mathbb{X}$ is *μ-pseudo almost automorphic in the wide sense* if there exists $g \in \mathrm{AA}(\mathbb{R}, \mathbb{X})$ such that the function $t \mapsto \mathfrak{d}(f(t), g(t)) \wedge 1$ is in $\mathcal{E}(\mathbb{R}, \mathbb{R}, \mu)$. We can get rid of the distance $\mathfrak{d}$ in this definition by using the fact that the closure of the range of any function in $\mathrm{AA}(\mathbb{R}, \mathbb{X})$ is compact. On a compact space K, there is one and only one uniform structure, a basis of entourages of which consists of all sets of the form $V = \cup_{i=1}^{m} (U_i \times U_i)$, where $U_1, \dots, U_m$ is a finite open cover of K, see e.g. [14] or [24]. In this way, we obtain (ii) below, which shows that the space of functions which are μ-pseudo almost automorphic in the wide sense depends only on the topology of $\mathbb{X}$.

We have thus two possible topological definitions of μ-pseudo almost automorphy: in Tudor and Tudor's sense, and in the wide sense. The former is stronger than (2.7), while the latter is weaker.

Proposition 2.5 *(Topological characterization of μ-pseudo almost automorphy in the wide sense) Let $f : \mathbb{R} \to \mathbb{X}$ be continuous. Let μ be a Borel measure on $\mathbb{R}$ satisfying* (2.6)*. If $\mathbb{X}$ is a metric space, the following propositions are equivalent.*

(i) f is μ-pseudo almost automorphic in the wide sense, i.e. there exists a function $g \in AA(\mathbb{R}, \mathbb{X})$ such that

$$\lim_{r\to\infty} \frac{1}{\mu([-r,r])} \int_{[-r,r]} \Big(\mathfrak{d}\left(f(t), g(t)\right) \wedge 1 \Big)\, d\mu(t) = 0.$$

(ii) There exists a function $g \in AA(\mathbb{R}, \mathbb{X})$ such that, for any finite open cover $U_1, \dots, U_m$ of the closure K of $\{g(t);\, t \in \mathbb{R}\}$,

$$\lim_{r\to\infty} \frac{\mu\{t \in [-r,r];\ (f(t), g(t)) \notin V\}}{\mu([-r,r])} = 0,$$

where $V = \cup_{i=1}^{m} (U_i \times U_i)$.

Proof

(i) $\Rightarrow$ (ii). Recall that K is compact because $g \in \mathrm{AA}(\mathbb{R}, \mathbb{X})$. Let $V = \cup_{i=1}^{m} (U_i \times U_i)$ be as in (ii). Then V is an open neighborhood in $\mathbb{X} \times \mathbb{X}$ of the diagonal $\Delta = \{(x, x);\, x \in K\}$. Define the distance $\mathfrak{d}_2$ on $\mathbb{X} \times \mathbb{X}$ by

$$\mathfrak{d}_2\Big((x, y), (x', y')\Big) = \mathfrak{d}(x, x') + \mathfrak{d}(y, y').$$

As Δ is compact, the distance $\varepsilon = \mathfrak{d}_2(\Delta, \mathbb{X} \times \mathbb{X} \setminus V)$ is positive. Let

$$B_\varepsilon := \{(y, z) \in \mathbb{X} \times \mathbb{X};\, \mathfrak{d}_2\Big((y, z), \Delta\Big) < \varepsilon\} \subset V.$$

For each $(y, z) \in B_\varepsilon$, there exists $x \in K$ such that $\mathfrak{d}_2\Big((y, z), (x, x)\Big) < \varepsilon$, thus

$$\mathfrak{d}(y, z) \le \mathfrak{d}(y, x) + \mathfrak{d}(x, z) < \varepsilon.$$

On the other hand, if $z \in K$ and $\mathfrak{d}(y,z) < \varepsilon$, we have

$$\mathfrak{d}_2\Big((y,z),(z,z)\Big) = \mathfrak{d}(y,z) < \varepsilon,$$

thus $(y,z) \in B_\varepsilon$. Applying this to $(f(t), g(t))$, we get, for every $r > 0$,

$$\frac{\mu\{t \in [-r,r];\ (f(t),g(t)) \notin V\}}{\mu([-r,r])} \le \frac{\mu\{t \in [-r,r];\ \mathfrak{d}\,(f(t),g(t)) \ge \varepsilon\}}{\mu([-r,r])}.$$

But (i) means that the function $t \mapsto \mathfrak{d}(f(t),g(t)) \wedge 1$ is in $\mathcal{E}(\mathbb{R},\mathbb{R},\mu)$, thus, by [7, Theorem 2.14], the latter term goes to 0 when r goes to ∞.

(ii) $\Rightarrow$ (i). Let $\varepsilon > 0$. Let $U_1,\dots,U_m$ be a finite open cover of K such that $\mathrm{Diam}(U_i) > \varepsilon$, $i = 1,\dots,m$, and let $V = \cup_{i=1}^m (U_i \times U_i)$. We have

$$\frac{\mu\{t \in [-r,r];\ \mathfrak{d}\,(f(t),g(t)) > \varepsilon\}}{\mu([-r,r])} \le \frac{\mu\{t \in [-r,r];\ (f(t),g(t)) \notin V\}}{\mu([-r,r])},$$

where the latter term goes to 0 when r goes to ∞. The conclusion follows from [7, Theorem 2.14]. ∎

Remark 2.6 The reasoning of Proposition 2.5 can be applied without change to give a topological characterization of μ-pseudo almost periodic functions in the wide sense. More generally, $\mathrm{AA}(\mathbb{R},\mathbb{X})$ can be replaced in this reasoning by any class of functions which have relatively compact range.

Theorem 2.7 *(Uniqueness of the decompostion of pseudo almost automorphic functions in the wide sense) Let μ be a Borel measure on $\mathbb{R}$ satisfying* (2.6) *and Condition (**H**). Let $f : \mathbb{R} \to \mathbb{X}$ satisfying Condition* (ii) *of Proposition 2.5. Then the function $g \in AA(\mathbb{R},\mathbb{X})$ given by Condition* (ii) *is unique and satisfies*

$$\{g(t);\, t \in \mathbb{R}\} \subset \overline{\{f(t);\, t \in \mathbb{R}\}}.$$

Proof Let g and K as in Condition (ii). Let $\varphi : \mathbb{X} \to \mathbb{R}$ be a continuous function. Then $\varphi \circ g \in \mathrm{AA}(\mathbb{R},\mathbb{R})$. Let $\hat{U}_1,\dots,\hat{U}_m$ be a finite open cover of the closure $\varphi(K)$ of $\{\varphi \circ g(t);\, t \in \mathbb{R}\}$, and let $\hat{V} = \cup_{i=1}^m \left(\hat{U}_i \times \hat{U}_i\right)$. Then $U_1 = \varphi^{-1}(\hat{U}_1),\dots,U_m = \varphi^{-1}(\hat{U}_m)$ form a finite open cover of K. Let $V = \cup_{i=1}^m (U_i \times U_i)$. We have

$$\frac{\mu\{t \in [-r,r];\ (\varphi \circ f(t), \varphi \circ g(t)) \notin \hat{V}\}}{\mu([-r,r])} \le \frac{\mu\{t \in [-r,r];\ (f(t),g(t)) \notin V\}}{\mu([-r,r])}.$$

Thus, by Proposition 2.5, $\varphi \circ f$ is in $\mathrm{PAA}(\mathbb{R},\mathbb{R},\mu)$ and the function $\varphi \circ f - \varphi \circ g$ is in $\mathcal{E}(\mathbb{R},\mathbb{R},\mu)$. By [7, Theorem 4.1], we have

$$\{\varphi \circ g(t);\, t \in \mathbb{R}\} \subset \overline{\{\varphi \circ f(t);\, t \in \mathbb{R}\}},$$

and, by the uniqueness of the decomposition of vector-valued μ-pseudo almost automorphic functions [7, Theorem 4.7] if $g' \in \mathrm{AA}(\mathbb{R},\mathbb{X})$ satisfies the same condition as g, we have $\varphi \circ g' = \varphi \circ g$. As φ is arbitrary, we deduce $g' = g$. ∎

3 Weighted pseudo almost automorphy for stochastic processes

From now on, $\mathbb{X}$ and $\mathbb{Y}$ are assumed to Polish spaces, i.e. separable metrizable topological spaces whose topology is generated by a complete metric.

3.1 Weighted pseudo almost automorphy in pth mean

We assume here that $\mathbb{X}$ is a Banach space. Let $X = (X_t)_{t\in\mathbb{R}}$ be a continuous stochastic process with values in $\mathbb{X}$, defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$. Let μ be a Borel measure on $\mathbb{R}$ satisfying (2.6).

Let $p > 0$. We say that X is *almost automorphic in pth mean* (respectively *μ-pseudo almost automorphic in pth mean*) if the mapping $t \mapsto X(t)$ is in $\mathrm{AA}(\mathbb{R}, \mathrm{L}^p(\Omega, \mathrm{P}, \mathbb{X}))$ (respectively in $\mathrm{PAA}(\mathbb{R}, \mathrm{L}^p(\Omega, \mathrm{P}, \mathbb{X}), \mu)$, i.e., if it has the form $X = Y + Z$, where $Y \in \mathrm{AA}(\mathbb{R}, \mathrm{L}^p(\Omega, \mathrm{P}, \mathbb{X}))$ and $Z \in \mathcal{E}(\mathbb{R}, \mathrm{L}^p(\Omega, \mathrm{P}, \mathbb{X}), \mu)$). When $p = 2$, we say that X is *square-mean almost automorphic* (respectively *square-mean μ-pseudo almost automorphic*).

The process X is said to be *almost automorphic in probability* if the mapping $X : t \to \mathrm{L}^0(\Omega, \mathrm{P}, \mathbb{X})$ is almost automorphic, where $\mathrm{L}^0(\Omega, \mathrm{P}, \mathbb{X})$ is the space of measurable mappings from Ω to $\mathbb{X}$, endowed with the topology of convergence in probability. Recall that the topology of $\mathrm{L}^0(\Omega, \mathrm{P}, \mathbb{X})$ is induced by e.g. the distance

$$\mathfrak{d}_{\mathrm{Prob}}(U, V) = \mathrm{E}\left(\|U - V\| \wedge 1\right),$$

which is complete.

The process X is said to be *μ-pseudo almost automorphic in probability*, and we write $X \in \mathrm{PAA}(\mathbb{R}, \mathrm{L}^0(\Omega, \mathrm{P}, \mathbb{X}), \mu)$, if the mapping $t \mapsto X(t)$, $\mathbb{R} \to \mathrm{L}^0(\Omega, \mathrm{P}, \mathbb{X})$ is μ-pseudo almost automorphic in the wide sense (or, equivalently, if it is μ-pseudo almost automorphic when $\mathrm{L}^0(\Omega, \mathrm{P}, \mathbb{X})$ is endowed with $\mathfrak{d}_{\mathrm{Prob}}$), i.e. if it has the form $X = Y + Z$ where $Y \in \mathrm{AA}(\mathbb{R}, \mathrm{L}^0(\Omega, \mathrm{P}, \mathbb{X}))$ and Z satisfies

$$\lim_{r\to\infty} \frac{1}{\mu([-r, r])} \int_{[-r,r]} \mathrm{E}\left(\|Z(t)\| \wedge 1\right) d\mu(t) = 0. \tag{3.1}$$

We denote by $\mathcal{E}(\mathbb{R}, \mathrm{L}^0(\Omega, \mathrm{P}, \mathbb{X}), \mu)$ the set of stochastic processes Z satisfying (3.1).

Note that, for $p \geq 0$, the previous decompositions of $X \in \mathrm{PAA}(\mathbb{R}, \mathrm{L}^p(\Omega, \mathrm{P}, \mathbb{X}), \mu)$ are unique under the condition **(H)**.

Clearly, for $0 \leq p \leq q$, we have

$$\mathrm{AA}(\mathbb{R}, \mathrm{L}^q(\Omega, \mathrm{P}, \mathbb{X})) \subset \mathrm{AA}(\mathbb{R}, \mathrm{L}^p(\Omega, \mathrm{P}, \mathbb{X}))$$

and

$$\mathrm{PAA}(\mathbb{R}, \mathrm{L}^q(\Omega, \mathrm{P}, \mathbb{X}), \mu) \subset \mathrm{PAA}(\mathbb{R}, \mathrm{L}^p(\Omega, \mathrm{P}, \mathbb{X}), \mu).$$

Conversely, if the set $\{\|X(t)\|^q \,;\, t \in \mathbb{R}\}$ is uniformly integrable, we have the implications

$$\Big(X \in \mathrm{AA}(\mathbb{R}, \mathrm{L}^p(\Omega, \mathrm{P}, \mathbb{X}))\Big) \Rightarrow \Big(X \in \mathrm{AA}(\mathbb{R}, \mathrm{L}^q(\Omega, \mathrm{P}, \mathbb{X}))\Big),$$

$$\Big(X \in \mathrm{PAA}(\mathbb{R}, \mathrm{L}^p(\Omega, \mathrm{P}, \mathbb{X}), \mu) \Big) \Rightarrow \Big(X \in \mathrm{PAA}(\mathbb{R}, \mathrm{L}^q(\Omega, \mathrm{P}, \mathbb{X}), \mu) \Big).$$

A process X is *Stepanov-like almost automorphic in pth mean* if X is in $\mathrm{AA}_{\mathbb{S}^p}(\mathbb{R}, \mathrm{L}^p(\Omega, \mathrm{P}, \mathbb{X}))$. We define in the same way the processes which are *Stepanov-like μ-pseudo almost automorphic in pth mean*, *Weyl-like (μ-pseudo) almost automorphic in pth mean*, *Besicovitch-like (μ-pseudo) almost automorphic in pth mean*.

Explicit counterexample to square-mean pseudo almost automorphy
At the time of the submission of this paper, there are at least 24 papers listed in Mathematical Reviews related to almost automorphy of solutions to stochastic differential equations, which all have been published in 2010 or later. To our knowledge, except for [26, 27, 35], all other papers claim the existence of square-mean pseudo almost automorphic solutions to stochastic differential equations with coefficients having similar properties.

We show that a very simple counterexample from [37, 38] contradicts these claims. The other counterexamples given in [37, 38] also contradict these claims.

Example 3.1 (Stationary Ornstein-Uhlenbeck process) Let $W = (W(t))_{t\in\mathbb{R}}$ be a standard Brownian motion on the real line. Let $\alpha, \sigma > 0$, and let X be the stationary Ornstein-Uhlenbeck process (see [34]) defined by

$$X(t) = \sqrt{2\alpha\sigma} \int_{-\infty}^{t} e^{-\alpha(t-s)} dW(s). \tag{3.2}$$

Then X is the only L^2-bounded solution of the following SDE, which is a particular case of Equation (3.1) in [4]:

$$dX(t) = -\alpha X(t)\, dt + \sqrt{2\alpha}\sigma\, dW(t).$$

The process X is Gaussian with mean 0, and we have, for all $t \in \mathbb{R}$ and $\tau \geq 0$,

$$\mathrm{Cov}(X(t), X(t+\tau)) = \sigma^2 e^{-\alpha\tau}.$$

Assume that X is square-mean μ-pseudo almost automorphic, for some Borel measure μ on $\mathbb{R}$ satisfying (2.6) and (2.8). Then we can decompose X as

$$X = Y + Z, \quad Y \in \mathrm{AA}(\mathbb{R}, \mathrm{L}^2(\Omega, \mathbb{R})), \quad Z \in \mathcal{E}(\mathbb{R}, \mathrm{L}^2(\Omega, \mathbb{R}), \mu).$$

By Lemma 2.2, we can find an increasing sequence (t_n) of real numbers which converges to $+\infty$ such that $(Z(t_n))$ converges to 0 in $\mathrm{L}^2(\Omega, \mathbb{R})$. Then, we can extract a sequence (still denoted by (t_n) for simplicity) such that $(Y(t_n))$ converges in L^2 to a random variable $\widehat{Y}$. Thus $(X(t_n))$ converges in L^2 to $\widehat{Y}$. Necessarily $\widehat{Y}$ is Gaussian with law $\mathcal{N}(0, 2\alpha\sigma^2)$, and $\widehat{Y}$ is $\mathcal{G}$-measurable, where $\mathcal{G} = \sigma(X_{t_n};\, n \geq 0)$. Moreover $(X(t_n), \widehat{Y})$ is Gaussian for every n, and we have, for any integer n,

$$\mathrm{Cov}(X(t_n), \widehat{Y}) = \lim_{m\to\infty} \mathrm{Cov}(X(t_n), X(t_{n+m})) = 0,$$

because $(X^2(t))_{t\in\mathbb{R}}$ is uniformly integrable. This proves that $\widehat{Y}$ is independent of $X(t_n)$ for every n, thus $\widehat{Y}$ is independent of $\mathcal{G}$. Thus $\widehat{Y}$ is constant, a contradiction. Thus (3.2) has no square-mean μ-pseudo almost automorphic solution.

Let us show that X is not Weyl-like nor Besicovitch-like square-mean pseudo almost automorphic. It is enough to disprove the Besicovitch sense. Assume that X is Besicovitch-like square-mean pseudo almost automorphic. As before, using Lemma 2.2, we can find a sequence (t_n) converging to $+\infty$ and a process $\widehat{Y}$ such that

$$\lim_{n\to\infty}\left\|X(t_n)-\widehat{Y}\right\|_{\mathbb{B}^2}=0.$$

In particular, $(X(t_n))$ is Cauchy for $\|.\|_{\mathbb{B}^2}$, thus, for every $\varepsilon>0$, there exists $N(\varepsilon)\in\mathbb{N}$, such that, for all $n,m\in\mathbb{N}$,

$$(n\geq N(\varepsilon))\Rightarrow\left\|X(t_n)-X(t_{n+m})\right\|_{\mathbb{B}^2}\leq\varepsilon. \tag{3.3}$$

But we have

$$\begin{aligned}
&\left\|X(t_n)-X(t_{n+m})\right\|_{\mathbb{B}^2}^2\\
=&\limsup_{r\to+\infty}\frac{1}{2r}\int_{-r}^{r}\mathrm{E}\left|X(t_n+s)-X(t_{n+m}+s)\right|^2ds\\
=&\limsup_{r\to+\infty}\frac{1}{2r}\int_{-r}^{r}\mathrm{E}\big(X^2(t_n+s)+X^2(t_{n+m}+s)-2X(t_n+s)X(t_{n+m}+s)\big)ds\\
=&\limsup_{r\to+\infty}\frac{1}{2r}\int_{-r}^{r}\big(\sigma^2+\sigma^2-2\sigma^2e^{-\alpha(t_{n+m}-t_n)}\big)ds\\
=&2\sigma^2\big(1-e^{-\alpha(t_{n+m}-t_n)}\big).
\end{aligned}$$

For m large, the last term is arbitrarily close to $2\sigma^2$, which contradicts (3.3) for $\varepsilon<2\sigma^2$.

A similar calculation shows that, for any Borel measure μ on $\mathbb{R}$ satisfying (2.6) and (2.8), and for any Borel measure ν on $[0,1]$ satisfying (2.9), the process X is not square-mean $\mathbb{S}^2_\nu$-μ-pseudo almost automorphic.

3.2 Weighted pseudo almost automorphy in distribution

We denote by $\operatorname{law}(X)$ the law (or distribution) of a random variable X. For any topological space $\mathbb{X}$, we denote by $\mathcal{M}^{1,+}(\mathbb{X})$ the set of Borel probability measures on $\mathbb{X}$, endowed with the topology of narrow (or weak) convergence, i.e. the coarsest topology such that the mappings $\mu\mapsto\mu(\varphi)$, $\mathcal{M}^{1,+}(\mathbb{X})\to\mathbb{R}$ are continuous for all bounded continuous $\varphi:\mathbb{X}\to\mathbb{R}$.

If $\tau:\mathbb{X}\to\mathbb{Y}$ is a Borel measurable mapping and μ is a Borel measure on $\mathbb{X}$, we denote by $\tau_\sharp\,\mu$ the Borel measure on $\mathbb{Y}$ defined by

$$\tau_\sharp\,\mu(B)=\mu(\tau^{-1}(B))$$

for every Borel set of $\mathbb{Y}$.

Let $\mathrm{BC}(\mathbb{X},\mathbb{R})$ denote the space of bounded continuous functions from $\mathbb{X}$ to $\mathbb{R}$, which we endow with the norm

$$\|\varphi\|_\infty = \sup_{x\in\mathbb{X}} |\varphi(x)|.$$

For a given distance $\mathfrak{d}$ on $\mathbb{X}$, and for $\varphi \in \mathrm{BC}(\mathbb{X},\mathbb{R})$ we define

$$\|\varphi\|_{\mathrm{L}} = \sup\Big\{\frac{\varphi(x)-\varphi(y)}{\mathfrak{d}(x,y)};\, x\neq y\Big\}$$

$$\|\varphi\|_{\mathrm{BL}} = \max\{\|\varphi\|_\infty, \|\varphi\|_{\mathrm{L}}\}.$$

We denote

$$\mathrm{BL}(\mathbb{X},\mathbb{R}) = \big\{\varphi \in \mathrm{BC}(\mathbb{X},\mathbb{R}); \|\varphi\|_{\mathrm{BL}} < \infty\big\}.$$

The *bounded Lipschitz distance* $\mathfrak{d}_{\mathrm{BL}}$ associated with $\mathfrak{d}$ on $\mathcal{M}^{1,+}(\mathbb{X})$ is defined by

$$\mathfrak{d}_{\mathrm{BL}}(\mu,\nu) = \sup_{\substack{\varphi\in\mathrm{BL}(\mathbb{X},\mathbb{R})\\ \|\varphi\|_{\mathrm{BL}}\le 1}} \int_{\mathbb{X}} \varphi\, d(\mu-\nu).$$

This metric generates the narrow (or weak) topology on $\mathcal{M}^{1,+}(\mathbb{X})$.

Let $p \geq 0$, and let (X_n) be a sequence in $\mathrm{L}^p(\Omega,\mathrm{P},\mathbb{X})$. We say that (X_n) *converges in p-distribution* (or simply *converges in distribution* if $p=0$) to a random vector X if

(i) the sequence $(\mathrm{law}(X_n))$ converges to $\mathrm{law}(X)$ for the narrow topology on $\mathcal{M}^{1,+}(\mathbb{X})$,

(ii) if $p>0$, the sequence $(\|X_n\|^p)$ is uniformly integrable.

Almost automorphy in distribution If X is continuous with values in $\mathbb{X}$, we denote by $\tilde{X}(t)$ the random variable $X(t+.)$ with values in $\mathrm{C}(\mathbb{R},\mathbb{X})$.

We say that X is *almost automorphic in one-dimensional distributions* if the mapping $t\mapsto \mathrm{law}(X(t))$, $\mathbb{R}\mapsto\mathcal{M}^{1,+}(\mathbb{X})$ is almost automorphic.

Remark 3.2 One-dimensional distributions of a process reflect poorly its properties. For example, let X be the Ornstein-Uhlenbeck process of Example 3.1, and set $Y(t) = X(0)$, $t\in\mathbb{R}$. The processes X and Y have the same one-dimensional distributions with completely different trajectories and behaviors. The trajectories of Y are constant, whereas the covariance $\mathrm{Cov}(X(t+\tau),X(t))$ converges to 0 when τ goes to ∞.

We say that X is *almost automorphic in finite dimensional distributions* if, for every finite sequence $t_1,\dots,t_m\in\mathbb{R}$, the mapping $t\mapsto \mathrm{law}(X(t+t_1),\dots,X(t+t_m))$, $\mathbb{R}\mapsto\mathcal{M}^{1,+}(\mathbb{X}^m)$ is almost automorphic.

We say that X is *almost automorphic in distribution* if the mapping $t\mapsto \mathrm{law}\left(\tilde{X}(t)\right)$, $\mathbb{R}\mapsto\mathcal{M}^{1,+}(\mathrm{C}_k(\mathbb{R},\mathbb{X}))$ is almost automorphic, where $\mathrm{C}_k(\mathbb{R},\mathbb{X})$ denotes the space $\mathrm{C}(\mathbb{R},\mathbb{X})$ endowed with the topology of uniform convergence on compact subsets. The Ornstein-Uhlenbeck process X of Example 3.1 is almost automorphic in distribution because the mapping $t\mapsto\mathrm{law}\left(\tilde{X}(t)\right)$ is constant.

Remark 3.3 If X is a deterministic process $\mathbb{R} \to \mathbb{X}$, we have the equivalences

$$\begin{aligned} X \in \mathrm{AA}(\mathbb{R}, \mathbb{X}) \Leftrightarrow & X \text{ is almost automorphic in one-dimensional distributions} \\ \Leftrightarrow & X \text{ is almost automorphic in finite dimensional distributions} \end{aligned}$$

and

$$X \in \mathrm{AA}_c(\mathbb{R}, \mathbb{X}) \Leftrightarrow X \text{ is almost automorphic in distribution.}$$

Actually, as Remark 3.3 suggests, the definition of almost automorphy in distribution implies a stronger property for the mapping $t \mapsto \operatorname{law}\left(\tilde{X}(t)\right)$. We need first some notations. For simplicity, we set $\mathrm{C}_k = \mathrm{C}_k(\mathbb{R}, \mathbb{X})$. For every $t \in \mathbb{R}$, we define a continuous operator on C_k:

$$\tau_t : \begin{cases} \mathrm{C}_k & \to & \mathrm{C}_k \\ x & \mapsto & x(t + .) = \tilde{x}(t). \end{cases}$$

Proposition 3.4 *If X is almost automorphic in distribution, the mapping $t \mapsto \operatorname{law}\left(\tilde{X}(t)\right)$ is in $AA_c(\mathbb{R}, \mathcal{M}^{1,+}(\mathrm{C}_k(\mathbb{R}, \mathbb{X})))$. Furthermore, for any sequence (t_n) in $\mathbb{R}$ such that, for every $t \in \mathbb{R}$, $(\operatorname{law}\left(\tilde{X}(t + t_n)\right))$ converges to a limit $g(t) \in \mathcal{M}^{1,+}(\mathrm{C}_k)$, the function g satisfies, for every $t \in \mathbb{R}$, the consistency relation*

$$g(t) = (\tau_t)_\sharp\, g(0). \tag{3.4}$$

Proof Let us denote $f(t) = \operatorname{law}\left(\tilde{X}(t)\right)$. We have, for every $t \in \mathbb{R}$,

$$f(t) = \operatorname{law}(\tau_t \circ X) = (\tau_t)_\sharp\, f(0).$$

Let (t_n) be a sequence in $\mathbb{R}$ such that, for each $t \in \mathbb{R}$, the sequence $(f(t + t_n))$ converges to some $g(t) \in \mathcal{M}^{1,+}(\mathrm{C}_k)$. Then, for every $t \in \mathbb{R}$, (3.4) is satisfied by continuity of the operator $(\tau_t)_\sharp$. To prove the continuity of g, let us endow C_k with the distance

$$\underline{\mathfrak{d}}(x, y) = \sum_{k \geq 1} 2^{-k} \sup_{-k \leq t \leq k} \Big(\mathfrak{d}\big(x(t), y(t)\big) \wedge 1\Big), \tag{3.5}$$

and let $\underline{\mathfrak{d}}_{\mathrm{BL}}$ be the associated bounded Lipschitz distance on $\mathcal{M}^{1,+}(\mathrm{C}_k)$. Let us show that the convergence of $(f(. + t_n))$ is uniform for $\underline{\mathfrak{d}}_{\mathrm{BL}}$ on compact intervals. Let $r \geq 1$ be an integer. For every $t \in [-r, r]$, and for all $x, y \in \mathrm{C}_k$, we have

$$\begin{aligned} \underline{\mathfrak{d}}(\tau_t(x), \tau_t(y)) &= \sum_{k \geq 1} 2^{-k} \sup_{-k \leq s \leq k} \Big(\mathfrak{d}\big(x(s+t), y(s+t)\big) \wedge 1\Big) \\ &\leq \sum_{k \geq 1} 2^{-k} \sup_{-k-r \leq s \leq k+r} \Big(\mathfrak{d}\big(x(s), y(s)\big) \wedge 1\Big) \\ &\leq 2^r \underline{\mathfrak{d}}(x, y). \end{aligned}$$

Thus, for any 1-Lipschitz mapping $\varphi : \mathrm{C}_k \to \mathbb{R}$, the mapping $\varphi \circ \tau_t$ is 2^r-Lipschitz. We deduce that, if $\|\varphi\|_{\mathrm{BL}} \le 1$, we have $\|\varphi \circ \tau_t\|_{\mathrm{BL}} \le 1 + 2^r$. We have thus

$$\begin{aligned}\underline{\mathfrak{d}}_{\mathrm{BL}}\big(f(t+t_n), f(t+t_{n+m})\big) =& \sup_{\|\varphi\|_{\mathrm{BL}}\le 1} \mathrm{E}\Big(\varphi \circ \tau_t \circ \tilde{X}(t_n) - \varphi \circ \tau_t \circ \tilde{X}(t_{n+m})\Big)\\ \le& (1+2^r) \sup_{\|\varphi\|_{\mathrm{BL}}\le 1} \mathrm{E}\Big(\varphi \circ \tilde{X}(t_n) - \varphi \circ \tilde{X}(t_{n+m})\Big)\\ =& (1+2^r)\underline{\mathfrak{d}}_{\mathrm{BL}}\big(f(t_n), f(t_{n+m})\big),\end{aligned}$$

which shows that $(f(.+t_n))$ is uniformly Cauchy on $[-r, r]$. Thus g is continuous, and $f \in \mathrm{AA}_c(\mathbb{R}, \mathcal{M}^{1,+}(\mathrm{C}_k(\mathbb{R}, \mathbb{X})))$. ■

We denote by

- $\mathrm{AAD}_1(\mathbb{R}, \mathbb{X})$ the set of $\mathbb{X}$-valued processes which are almost automorphic in one-dimensional distributions,
- $\mathrm{AAD}_f(\mathbb{R}, \mathbb{X})$ the set of $\mathbb{X}$-valued processes which are almost automorphic in finite dimensional distributions,
- $\mathrm{AAD}(\mathbb{R}, \mathbb{X})$ the set of $\mathbb{X}$-valued processes which are almost automorphic in distribution.

(In these notations, we omit the probability space $(\Omega, \mathcal{F}, \mathrm{P})$, as there is no ambiguity here.) We have the inclusions

$$\mathrm{AAD}(\mathbb{R}, \mathbb{X}) \subset \mathrm{AAD}_f(\mathbb{R}, \mathbb{X}) \subset \mathrm{AAD}_1(\mathbb{R}, \mathbb{X}).$$

The following result is in the line of [3, Theorem 2.3].

Theorem 3.5 *Let X be an $\mathbb{X}$-valued stochastic process, and let $\mathfrak{d}$ be a distance on $\mathbb{X}$ which generates the topology of $\mathbb{X}$. Asume that X satisfies the tightness condition*

$$\forall [a,b] \subset \mathbb{R},\ \forall \varepsilon > 0,\ \forall \eta > 0,\ \exists \delta > 0,\ \forall r \in \mathbb{R},\\ \mathrm{P}\Big\{ \sup_{\substack{|t-s|<\delta\\ t,s\in[a,b]}} \mathfrak{d}\left(X(r+t), X(r+s)\right) > \eta \Big\} < \varepsilon. \quad (3.6)$$

Then the following properties are equivalent:

(a) $X \in AAD_f(\mathbb{R}, \mathbb{X})$.

(b) $X \in AAD(\mathbb{R}, \mathbb{X})$.

Proof Clearly (b)⇒(a). Assume that $X \in \mathrm{AAD}_f(\mathbb{R}, \mathbb{X})$. Let (γ'_n) be a sequence in $\mathbb{R}$, and, for $t_1, t_2, \ldots, t_k, t \in \mathbb{R}$ define (using notations of [49])

$$\mu_t^{t_1,\ldots,t_k} := \mathrm{law}\left(X(t_1+t), \ldots, X(t_k+t)\right).$$

By a diagonal procedure we can find a subsequence (γ_n) of (γ'_n) such that, for every $k \geq 1$, for all $q_1, q_2, \ldots, q_k \in \mathbb{Q} \bigcap \mathbb{R}$ (where $\mathbb{Q}$ is the set of rational numbers), and for every $t \in \mathbb{R}$,

$$\lim_n \lim_m \mu^{q_1,\ldots,q_k}_{t+\gamma_n-\gamma_m} = \mu^{q_1,\ldots,q_k}_t.$$

Let $\mathfrak{d}_k$ be the distance on $\mathbb{X}^k$ defined by

$$\mathfrak{d}_k\big((x_1,\ldots,x_k),(y_1,\ldots,y_k)\big) = \max_{1\leq i\leq k} \mathfrak{d}(x_i,y_i),$$

and let $\mathfrak{d}_{\mathrm{BL}}$ the associated bounded Lipschitz distance on $\mathcal{M}^{1,+}\left(\mathbb{X}^k\right)$. We have, for all $t_1, t_2, \ldots, t_k, t \in \mathbb{R}$, for all $q_1, q_2, \ldots, q_k \in \mathbb{Q} \bigcap \mathbb{R}$, and for all $n, m \in \mathbb{N}$,

$$\begin{aligned}\mathfrak{d}_{\mathrm{BL}}\Big(\mu^{q_1,\ldots,q_k}_{t+\gamma_n-\gamma_m}, \mu^{t_1,\ldots,t_k}_{t+\gamma_n-\gamma_m}\Big) &= \sup_{\|f\|_{\mathrm{BL}}\leq 1} \int_{\mathbb{X}^k} f\,d\big(\mu^{q_1,\ldots,q_k}_{t+\gamma_n-\gamma_m} - \mu^{t_1,\ldots,t_k}_{t+\gamma_n-\gamma_m}\big)\\ &\leq \max_{1\leq i\leq k} \int_\Omega \mathfrak{d}\Big(X(q_i+t+\gamma_n-\gamma_m), X(t_i+t+\gamma_n-\gamma_m)\Big)\,d\,\mathrm{P},\end{aligned}$$

so that, by (3.6), if $(q_1,\ldots,q_k) \to (t_1,\ldots,t_k)$, then

$$\mathfrak{d}_{\mathrm{BL}}\Big(\mu^{q_1,\ldots,q_k}_{t+\gamma_n-\gamma_m}, \mu^{t_1,\ldots,t_k}_{t+\gamma_n-\gamma_m}\Big) \to 0$$

uniformly with respect to $t \in \mathbb{R}$ and $n, m \in \mathbb{N}$. By a classical result on inversion of limits, we deduce that, for all $k \geq 1$ and $t_1, \ldots, t_k, t \in \mathbb{R}$,

$$\lim_n \lim_m \mu^{t_1,\ldots,t_k}_{t+\gamma_n-\gamma_m} = \mu^{t_1,\ldots,t_k}_t.$$

Therefore, to show that

$$\lim_n \lim_m \mathrm{law}\left(\tilde{X}(t+\gamma_n-\gamma_m)\right) = \mathrm{law}\left(\tilde{X}(t)\right),$$

it is enough to prove that $(\tilde{X}(t))_{t\in\mathbb{R}}$ is tight in $\mathrm{C}_k(\mathbb{R},\mathbb{X})$. Since $X \in \mathrm{AAD}_f(\mathbb{R},\mathbb{X})$, the family $(X(t))_{t\in\mathbb{R}} = (\tilde{X}(t)(0))_{t\in\mathbb{R}}$ is tight, by Prokhorov's theorem for relatively compact sets of probability measures on Polish spaces. By (3.6) and the Arzelà-Ascoli-type characterization of tight subsets of $\mathcal{M}^{1,+}(\mathbb{X})$ (see e.g. the proof of [5, Theorem 7.3] or [52, Theorem 4]), we conclude that $(\tilde{X}(t))_{t\in\mathbb{R}}$ is tight in $\mathrm{C}_k(\mathbb{R},\mathbb{X})$, which proves our claim. ■

Remark 3.6 Assume that $\mathbb{X}$ is a vector space. The spaces $\mathrm{AAD}(\mathbb{R},\mathbb{X})$, $\mathrm{AAD}_f(\mathbb{R},\mathbb{X})$, and $\mathrm{AAD}_1(\mathbb{R},\mathbb{X})$ are not vector spaces. Indeed, let X be the Ornstein-Uhlenbeck process of Example 3.1. For each $t \in \mathbb{R}$, let $Y(t) = X(0)$. The processes X and Y are stationary in the strong sense, thus they are in $\mathrm{AAD}(\mathbb{R},\mathbb{X})$. For each $t \in \mathbb{R}$, the variable $Z(t) = X(t) + Y(t)$ is Gaussian centered with variance

$$\mathrm{Var}\,Z(t) = \mathrm{E}\left(X^2(t)\right) + \mathrm{E}\left(Y^2(t)\right) + 2\,\mathrm{Cov}\big(X(t),Y(t)\big)$$

$$
\begin{aligned}
&= 2\sigma^2 + 2\sigma^2 \exp\left(-\alpha\,|t|\right) \\
&\to 2\sigma^2 \text{ when } |t| \to \infty.
\end{aligned}
$$

Thus $\operatorname{law}(Z(t))$ is the Gaussian distribution $\mathcal{N}(0, 2\sigma^2(1+\exp(-\alpha\,|t|)))$, which converges when $|t| \to \infty$ to $\mathcal{N}(0, 2\sigma^2)$. Set $\mathfrak{m}(t) = \mathcal{N}(0, 2\sigma^2)$, $t \in \mathbb{R}$. For each $t \in \mathbb{R}$, we have

$$
\begin{aligned}
\lim_{n\to\infty} \operatorname{law}\left(Z(t+n)\right) &= \mathfrak{m}(t) \\
\lim_{n\to\infty} \mathfrak{m}(t-n) = \mathfrak{m}(t) &\neq \operatorname{law}\left(Z(t)\right).
\end{aligned}
$$

Thus $Z \notin \mathrm{AAD}_1(\mathbb{R}, \mathbb{X})$.

This contradicts [26, Lemma 2.3].

Almost automorphy in p-distribution A useful variant of almost automorphy in distribution takes into account integrability of order p. Let $p \geq 0$. We say that a continuous $\mathbb{X}$-valued stochastic process is *almost automorph in p-distribution* if

(i) $X \in \mathrm{AAD}(\mathbb{R}, \mathbb{X})$,

(ii) if $p > 0$, the family $(\|X(t)\|^p)_{t\in\mathbb{R}}$ is uniformly integrable.

These conditions imply that the mapping $t \mapsto X(t)$, $\mathbb{R} \to \mathrm{L}^p(\Omega, \mathrm{P}, \mathbb{X})$, is continuous.

We denote by $\mathrm{AAD}^p(\mathbb{R}, \mathbb{X})$ the set of $\mathbb{X}$-valued processes which are almost automorphic in p-distribution, in particular we have $\mathrm{AAD}^0(\mathbb{R}, \mathbb{X}) = \mathrm{AAD}(\mathbb{R}, \mathbb{X})$. Similarly, for $p \geq 0$, one defines the sets $\mathrm{AAD}^p_f(\mathbb{R}, \mathbb{X})$ and $\mathrm{AAD}^p_1(\mathbb{R}, \mathbb{X})$ of processes which are respectively *almost automorphic in one-dimensional p-distributions* and *almost automorphic in finite dimensional p-distributions*.

Weighted pseudo almost automorphy in distribution and variants As usual, we assume that μ is a Borel measure on $\mathbb{R}$ satisfying (2.6).

Tudor and Tudor proposed in [50] a very natural and elegant notion of pseudo almost periodicity in (one-dimensional) distribution that can easily be extended to weighted μ-pseudo almost automorphy: X is *μ-pseudo almost periodic in one-dimensional distributions in Tudor and Tudor's sense* if it satisfies

(TT_1) The mapping $t \mapsto \operatorname{law}(X(t))$ is continuous with relatively compact range in $\mathcal{M}^{1,+}(\mathbb{X})$, and there exists an almost automorphic function $\mathfrak{m} : \mathbb{R} \to \mathcal{M}^{1,+}(\mathbb{X})$ such that

$$
\lim_{r\to\infty} \frac{1}{\mu([-r,r])} \int_{-r}^{r} \mathfrak{d}_{\mathrm{BL}}(\operatorname{law}(X(t)), \mathfrak{m}(t))\, d\mu(t) = 0. \tag{3.7}
$$

Similar definitions are easy to write for μ-pseudo almost automorphy in finite distributions or in distribution.

Recall that $\mathfrak{d}_{\text{BL}}$ is bounded. If we remove the condition of relatively compact range (as in Proposition 2.5), we get three distributional notions of *μ-pseudo almost automorphy in the wide sense: in one-dimensional distributions, in finite dimensional distributions, and in distribution.*

We propose three stronger notions of μ-pseudo almost automorphy in a distributional sense, that seem to be particularly useful for stochastic equations.

Assume that $\mathbb{X}$ is a vector space. Let $p \geq 0$. We say that X is *μ-pseudo almost automorphic in p-distribution* if X can be written

$$X = Y + Z, \text{ where } Y \in \text{AAD}^p(\mathbb{R}, \mathbb{X}) \text{ and } Z \in \mathcal{E}(\mathbb{R}, \text{L}^p(\Omega, \text{P}, \mathbb{X}), \mu).$$

The set of $\mathbb{X}$-valued processes which are μ-pseudo almost automorphic in p-distribution is denoted by $\text{PAAD}^p(\mathbb{R}, \mathbb{X})$. Similar definitions hold for the spaces $\text{PAAD}^p_1(\mathbb{R}, \mathbb{X}, \mu)$ and $\text{PAAD}^p_f(\mathbb{R}, \mathbb{X}, \mu)$ of processes which are *μ-pseudo almost automorphic in one-dimensional p-distributions* and *in finite dimensional p-distributions* respectively.

Remark 3.7 The definitions we propose for μ-pseudo almost automorphy in distribution, or in finite dimensional distributions, or in one-dimensional distributions, are in a way stronger and less natural than those in the wide distributional sense, because they involve (at least, apparently) not only the distribution of the process, but the probability space $(\Omega, \mathcal{F}, \text{P})$. For example, a random process X is in $\text{PAAD}^0_1(\mathbb{R}, \mathbb{X}, \mu)$ if, and only if, there exists a process Y defined on the same probability space such that $\mathfrak{m}(.) := \text{law}\,(Y(.))$ is in $\text{AA}(\mathbb{R}, \mathcal{M}^{1,+}(\mathbb{X}))$ and satisfies (3.7). Note also that, in our definition, the ergodic part is ergodic in pth mean. Our definitions are thus intermediate between μ-pseudo almost automorphy in a purely distributional sense and μ-pseudo almost automorphy in pth mean.

However, our definitions seem to be convenient for calculations, and Theorem 4.4 shows that, for some stochastic differential equations with μ-pseudo almost automorphic coefficients, the process Y appears naturally: it is the solution of the corresponding SDE where the coefficients are the almost automorphic parts of the coefficients of the original SDE.

Almost automorphy in distribution and some of its variants enjoy some stability properties, as shows the following superposition lemma. Similar results could be proved by the same method for one-dimensional or for finite dimensional distributions.

Theorem 3.8 *(Superposition lemma) Let X be a continuous $\mathbb{X}$-valued stochastic process, and let $f : \mathbb{R} \times \mathbb{X} \to \mathbb{Y}$ be a continuous mapping.*

1. *If X is almost periodic in distribution and f is almost periodic with respect to the first variable, uniformly with respect to the second variable in compact subsets of $\mathbb{X}$, then $f(., X(.))$ is almost periodic in distribution.*

2. *If X is almost automorphic in distribution and f is compact almost automorphic with respect to the first variable, uniformly with respect to the*

second variable in compact subsets of $\mathbb{X}$, then $f(.,X(.))$ is almost automorphic in distribution.

3. *Let μ be a Borel measure on $\mathbb{R}$ satisfying (2.6), and let $p \geq 0$. Assume that f is μ-pseudo compact almost automorphic with respect to the first variable, uniformly with respect to the second variable in compact subsets of $\mathbb{X}$, i.e.*

$$f = g + h, \text{ with } g \in AA_c U_c(\mathbb{R} \times \mathbb{X}, \mathbb{Y}) \text{ and } h \in \mathcal{E} U_c(\mathbb{R} \times \mathbb{X}, \mathbb{Y}, \mu).$$

Assume that g is continuous with respect to the second variable. Assume furthermore that f is uniformly continuous in the second variable in compact subsets of $\mathbb{X}$, uniformly with respect to the first variable, that is,

$$\begin{aligned} &\text{for every } \varepsilon > 0, \text{ and for every compact subset } K \text{ of } \mathbb{X},\\ &\text{there exists } \eta > 0 \text{ such that, for all } x, y \in K,\\ &\qquad\qquad \|x - y\| \leq \eta \Rightarrow \sup_{t \in \mathbb{R}} \|f(t,x) - f(t,y)\| \leq \varepsilon. \end{aligned} \tag{3.8}$$

If $p > 0$, assume also that f and g satisfy the growth condition

$$\|f(t,x)\| + \|g(t,x)\| \leq C(1 + \|x\|) \tag{3.9}$$

for all $(t,x) \in \mathbb{R} \times \mathbb{X}$ and for some constant C, and that f is Lipschitz with respect to the second variable, uniformly with respect to the first one. If X is μ-pseudo almost automorphic in p-distribution, then $f(.,X(.))$ is μ-pseudo almost automorphic in p-distribution.

Proof We only prove the second and third items, the first one can be proved in the same way as 2, using, for example, Bochner's double sequence criterion.

2. For each $t \in \mathbb{R}$, for each $x \in \mathrm{C}_k(\mathbb{R}, \mathbb{X})$, and for each $s \in \mathbb{R}$, let us denote

$$\tilde{f}(t,x)(s) = f(s + t, x(s)).$$

By continuity of f on $\mathbb{R} \times \mathbb{X}$, $\tilde{f}(t,.)$ maps $\mathrm{C}_k(\mathbb{R}, \mathbb{X})$ to $\mathrm{C}_k(\mathbb{R}, \mathbb{Y})$.

Let $\mathcal{K}$ be a compact subset of $\mathrm{C}_k(\mathbb{R}, \mathbb{X})$. By the Arzelà-Ascoli Theorem (see e.g. [24, Theorems 8.2.10 and 8.2.11]), this means that $\mathcal{K}$ is closed in $\mathrm{C}_k(\mathbb{R}, \mathbb{X})$ and equicontinuous, and that, for every compact interval I of $\mathbb{R}$, the set $\{x(t);\ x \in \mathcal{K},\ t \in I\}$ has compact closure in $\mathbb{X}$. Let $t \in \mathbb{R}$, and let (t_n) be a sequence in $\mathbb{R}$ converging to t. Let I be a compact interval of $\mathbb{R}$, and let K be the closure of $\{x(s);\ x \in \mathcal{K},\ s \in I\}$. We have, for any $y \in K$, and for any $s \in \mathbb{R}$,

$$\lim_n f(t_n + s, y) = f(t + s, y)$$

where the convergence is uniform with respect to $y \in K$ and $s \in I$, because f is compact almost automorphic uniformly with respect to the second variable in compact subsets of $\mathbb{X}$. In particular we have, uniformly with respect to $x \in \mathcal{K}$ and $s \in I$,

$$\lim_n \tilde{f}(t_n, x)(s) = \lim_n f(t_n + s, x(s)) = f(t + s, x(s)) = \tilde{f}(t,x)(s),$$

which proves that the mapping $\tilde{f}(.,x) : \mathbb{R} \to \mathrm{C}_k(\mathbb{R},\mathbb{X})$, is continuous, uniformly with respect to x in compact subsets of $\mathrm{C}_k(\mathbb{R},\mathbb{X})$.

Let us check that $\tilde{f} : \mathbb{R} \times \mathrm{C}_k(\mathbb{R},\mathbb{X}) \to \mathrm{C}_k(\mathbb{R},\mathbb{X})$ is compact almost automorphic with respect to the first variable, uniformly with respect to the second variable in compact subsets of $\mathrm{C}_k(\mathbb{R},\mathbb{X})$. Let (t'_n) be a sequence in $\mathbb{R}$. There exists a subsequence (t_n) such that, for every $t \in \mathbb{R}$ and every $y \in \mathbb{X}$,

$$\lim_n \lim_m f(t + t_n - t_m, y) = f(t,y),$$

where the convergence is uniform with respect to y in compact subsets of $\mathbb{X}$ and t in compact intervals of $\mathbb{R}$. For each $t \in \mathbb{R}$, and for each $s \in \mathbb{R}$, we have

$$\begin{aligned}\lim_n \lim_m \tilde{f}(t + t_n - t_m, x)(s) =& \lim_n \lim_m f(s + t + t_n - t_m, x(s))\\ =& f(s+t, x(s)) = \tilde{f}(t,x)(s),\end{aligned}$$

and these convergences are uniform with respect to t and s in compact intervals, and with respect to $x \in \mathcal{K}$, which proves our claim.

Let $\varphi : \mathrm{C}_k(\mathbb{R},\mathbb{Y}) \to \mathbb{R}$ be bounded Lipschitz, with $\|\varphi\|_{\mathrm{BL}} \le 1$, where $\|.\|_{\mathrm{BL}}$ is taken relatively to a distance $\underline{\mathfrak{d}}$ which generates the topology of $\mathrm{C}_k(\mathbb{R},\mathbb{Y})$, for example the distance defined by (3.5). Let (t'_n) be a sequence in $\mathbb{R}$. Let (t_n) be a subsequence such that, for every $t \in \mathbb{R}$ and every $y \in \mathbb{X}$,

$$\begin{aligned}\lim_n \lim_m f(t + t_n - t_m, y) &= f(t,y),\\ \lim_n \lim_m \mathrm{law}\left(\tilde{X}(t + t_n - t_m)\right) &= \mathrm{law}\left(\tilde{X}(t)\right).\end{aligned}$$

Let $\varepsilon > 0$. We can find a compact subset $\mathcal{K}_\varepsilon$ of $\mathrm{C}_k(\mathbb{R},\mathbb{X})$ such that, for every $t \in \mathbb{R}$,

$$\mathrm{P}\left\{\tilde{X}(t) \in \mathcal{K}_\varepsilon\right\} \ge 1 - \varepsilon.$$

Let $t \in \mathbb{R}$ be fixed, and, for all n, m, let $\Omega_{\varepsilon,n,m}$ be the measurable subset of Ω on which $\tilde{X}(t + t_n - t_m) \in \mathcal{K}_\varepsilon$. We have

$$\begin{aligned}&\left|\mathrm{E}\left(\varphi \circ \tilde{f}(t + t_n - t_m, \tilde{X}(t + t_n - t_m)) - \varphi \circ \tilde{f}(t, \tilde{X}(t))\right)\right|\\ \le&\left|\mathrm{E}\left(\varphi \circ \tilde{f}(t + t_n - t_m, \tilde{X}(t + t_n - t_m)) - \varphi \circ \tilde{f}(t, \tilde{X}(t + t_n - t_m))\right)\right|\\ &+\left|\mathrm{E}\left(\varphi \circ \tilde{f}(t, \tilde{X}(t + t_n - t_m)) - \varphi \circ \tilde{f}(t, \tilde{X}(t))\right)\right|\\ \le&\, \mathrm{E}\left(\mathbb{1}_{\Omega_{\varepsilon,n,m}} \underline{\mathfrak{d}}\Big(\tilde{f}(t + t_n - t_m, \tilde{X}(t + t_n - t_m)), \tilde{f}(t, \tilde{X}(t + t_n - t_m))\Big)\right)\\ &+ \mathrm{E}\left(\mathbb{1}_{\Omega^c_{\varepsilon,n,m}} \left|\varphi \circ \tilde{f}(t + t_n - t_m, \tilde{X}(t + t_n - t_m)) - \varphi \circ \tilde{f}(t, \tilde{X}(t + t_n - t_m)))\right|\right)\\ &+\left|\mathrm{E}\left(\varphi \circ \tilde{f}(t, \tilde{X}(t + t_n - t_m)) - \varphi \circ \tilde{f}(t, \tilde{X}(t))\right)\right|\\ =&A_{n,m} + B_{n,m} + C_{n,m}.\end{aligned}$$

We have

$$A_{n,m} \le \mathrm{E}\left(\mathbb{1}_{\Omega_{\varepsilon,n,m}} \sup_{x\in\mathcal{K}_\varepsilon} \mathfrak{d}\Big(\tilde f(t{+}t_n{-}t_m,x),\tilde f(t,x)\Big)\right) \le \sup_{x\in\mathcal{K}_\varepsilon} \mathfrak{d}\Big(\tilde f(t{+}t_n{-}t_m,x),\tilde f(t,x)\Big)$$

thus by the almost automorphy property of $\tilde f$, we have $\lim_n \lim_m A_{n,m} = 0$. Furthermore, $B_{n,m} \le 2\,\mathrm{P}(\Omega^c_{\varepsilon,n,m}) \le 2\varepsilon$ because $\|\varphi\|_{\mathrm{BL}} \le 1$. Finally, $\lim_n \lim_m C_{n,m} = 0$ by boundedness and continuity of $\tilde f(t,.) : \mathrm{C}_k(\mathbb{R},\mathbb{X}) \to \mathrm{C}_k(\mathbb{R},\mathbb{X})$ and the convergence in distribution of $\tilde X(t+t_n-t_m)$ to $\tilde X(t)$ for each $t\in\mathbb{R}$. As ε and φ are arbitrary, we have proved that

$$\lim_n \lim_m \mathrm{law}\left(\tilde f(t+t_n-t_m,\tilde X(t+t_n-t_m))\right) = \mathrm{law}\left(\tilde f(t,\tilde X(t))\right),$$

thus the mapping $t \mapsto \mathrm{law}\left(\tilde f(t,\tilde X(t))\right)$ is almost automorphic in distribution.

3. We use ideas of the proof of [7, Theorem 5.7]. Let (Y,Z) be a decomposition of X, namely,

$$X = Y+Z,\ Y\in \mathrm{AAD}^p(\mathbb{R},\mathbb{X}),\ Z\in\mathcal{E}(\mathbb{R},\mathrm{L}^p(\Omega,\mathrm{P},\mathbb{X}),\mu).$$

The function $f(.,X(.))$ can be decomposed as

$$f(t,X(t)) = g(t,Y(t)) + f(t,X(t)) - f(t,Y(t)) + h(t,Y(t)).$$

Let $G(t) = g(t,Y(t))$ and $H(t) = f(t,X(t)) - f(t,Y(t)) + h(t,Y(t))$. By using 2. and (3.9), we see that $t\mapsto G(t)$ is in $\mathrm{AAD}^p(\mathbb{R},\mathbb{Y})$. Furthermore, by (3.9), the continuity of f and the μ-pseudo almost automorphy in p-distribution of X, we have, using Vitali's theorem, that $f(.,X(.))$ is a continuous $\mathrm{L}^p(\Omega,\mathrm{P},\mathbb{Y})$-valued function. Indeed, if $t_n\to t$, then the sequence $(\|X(t_n)\|^p)$ is uniformly integrable, by continuity of the mapping $t\mapsto X(t)$, $\mathbb{R}\to\mathrm{L}^p(\Omega,\mathrm{P},\mathbb{X})$, and this entails that $(f(t_n,X(t_n)))$ is uniformly integrable. To show that $f(.,X(.))$ is in $\mathrm{PAAD}^p(\mathbb{R},\mathbb{Y})$, it is enough to prove that $H\in\mathcal{E}(\mathbb{R},\mathrm{L}^p(\Omega,\mathrm{P},\mathbb{Y}),\mu)$.

Clearly H is in $\mathrm{C}(\mathbb{R},\mathrm{L}^p(\Omega,\mathrm{P},\mathbb{Y}))$, and bounded in $\mathrm{L}^p(\Omega,\mathrm{P},\mathbb{X})$ (when $p>0$), by (3.9). As Y is in $\mathrm{AAD}(\mathbb{R},\mathbb{X})$, the family $(\tilde Y(t))_{t\in\mathbb{R}} = (Y(t{+}.))_{t\in\mathbb{R}}$ is uniformly tight in $\mathrm{C}_k(\mathbb{R},\mathbb{X})$. For each $\varepsilon>0$, there exists a compact subset $\mathcal{K}_\varepsilon$ of $\mathrm{C}_k(\mathbb{R},\mathbb{X})$ such that, for every $t\in\mathbb{R}$,

$$\mathrm{P}\left\{\tilde Y(t)\in\mathcal{K}_\varepsilon\right\} \ge 1-\varepsilon.$$

By the Arzelà-Ascoli Theorem (see e.g. [24, Theorems 8.2.10 and 8.2.11]), this implies that, for every $\varepsilon>0$, and for every compact interval I of $\mathbb{R}$, there exists a compact subset $K_{\varepsilon,I}$ such that, for every $t\in\mathbb{R}$;

$$\mathrm{P}\left\{(\forall s\in I)\ Y(t+s)\in K_{\varepsilon,I}\right\} \ge 1-\varepsilon.$$

In particular, the family $(Y(t))_{t\in\mathbb{R}}$ is tight, i.e., denoting $K_\varepsilon = K_{\varepsilon,\{0\}}$, we have, for every $t\in\mathbb{R}$,

$$\mathrm{P}\left\{Y(t)\in K_\varepsilon\right\} \ge 1-\varepsilon.$$

Let $\Omega_{\varepsilon,t}$ be the measurable subset of Ω on which $Y(t) \in K_\varepsilon$. The function g is uniformly continuous on $\mathbb{R} \times K_\varepsilon$ by Proposition 2.1. We deduce by (3.8) that there exists $\eta(\varepsilon) > 0$ such that, for all $y_1, y_2 \in K_\varepsilon$,

$$\|y_1 - y_2\| \leq \eta(\varepsilon) \Rightarrow \sup_{t\in\mathbb{R}}\big(\|h(t,y_1) - h(t,y_2)\|\big) \leq \varepsilon.$$

We can find a finite sequence $(y_i)_{1\leq i\leq m}$ in K_ε such that

$$K_\varepsilon \subset \bigcup_{i=1}^{m} B(y_i, \eta(\varepsilon)).$$

We have, for every $t \in \mathbb{R}$,

$$\begin{aligned}
&\mathrm{E}\big(\|h(t,Y(t))\| \wedge 1\big)\\
&\qquad \leq \mathrm{E}\left(\min_{1\leq i\leq m}\big(\mathbb{1}_{\Omega_{\varepsilon,t}}\|h(t,Y(t)) - h(t,y_i)\| \wedge 1\big)\right) + \max_{1\leq i\leq m}\|h(t,y_i)\|\\
&\qquad\quad + \mathrm{E}\big(\mathbb{1}_{\Omega^c_{\varepsilon,t}}\|h(t,Y(t))\| \wedge 1\big)\\
&\qquad \leq \varepsilon + \max_{1\leq i\leq m}\|h(t,y_i)\| \wedge 1 + \mathrm{P}\left(\Omega^c_{\varepsilon,t}\right)\\
&\qquad \leq \max_{1\leq i\leq m}\|h(t,y_i)\| + 2\varepsilon.
\end{aligned}$$

Since for all $i \in \{1, \dots, m\}$, the function $t \to h(t,y_i)$ satisfies

$$\lim_{r\to\infty}\frac{1}{\mu([-r,r])}\int_{[-r,r]}\|h(t,y_i)\|\,d\mu(t) = 0,$$

we deduce that, for every $\varepsilon > 0$,

$$\limsup_{r\to\infty}\frac{1}{\mu([-r,r])}\int_{[-r,r]}\mathrm{E}\big(\|h(t,Y(t))\| \wedge 1\big)d\mu(t) \leq 2\varepsilon.$$

This shows that $t \to h(t,Y(t))$ is in $\mathcal{E}(\mathbb{R}, \mathrm{L}^0(\Omega,\mathrm{P},\mathbb{Y}),\mu)$.

For $p > 0$, let $\delta > 0$. From the uniform integrability of $\big(\|h(t,Y(t))\|^p\big)_{t\in\mathbb{R}}$ (thanks to (3.9) and the uniform integrability of $(\|Y(t)\|^p)_{t\in\mathbb{R}}$), we can choose ε small enough such that, for any measurable $A \subset \Omega$ such that $\mathrm{P}(A) < \varepsilon$,

$$\sup_{t\in\mathbb{R}}\mathrm{E}\,(\mathbb{1}_A\|h(t,Y(t))\|^p) < \delta.$$

Note also that, for $p < 1$, the mapping $U \mapsto (\mathrm{E}\,\|U\|^p)^{1/p}$, $\mathrm{L}^p \to \mathbb{R}$, does not satisfy the triangular inequality. However, the mapping $(U_1,U_2) \mapsto \mathrm{E}\,\|U_1 - U_2\|^p$ is a distance on L^p. We deduce that, for all $U_1, U_2, U_3 \in \mathrm{L}^p$,

$$(\mathrm{E}\,\|U_1 + U_2 + U_3\|^p)^{1/p} \leq 3^{1/p-1}\left((\mathrm{E}\,\|U_1\|^p)^{1/p} + (\mathrm{E}\,\|U_2\|^p)^{1/p} + (\mathrm{E}\,\|U_3\|^p)^{1/p}\right).$$

To cover simultaneously the cases $p < 1$ and $p \geq 1$, we set $\kappa = \max(1, 3^{1/p-1})$. Using the same method as in the case when $p = 0$, we get

$$
\begin{aligned}
&\big(\mathrm{E}\,\|h(t,Y(t))\|^p\big)^{1/p}\\
&\qquad\le\kappa\bigg(\mathrm{E}\big(\min_{1\le i\le m}\mathbb{1}_{\Omega_{\varepsilon,t}}\,\|h(t,Y(t))-h(t,y_i)\|^p\big)\bigg)^{1/p}+\kappa\max_{1\le i\le m}\|h(t,y_i)\|\\
&\qquad\quad+\kappa\bigg(\mathrm{E}\big(\mathbb{1}_{\Omega^c_{\varepsilon,t}}\,\|h(t,Y(t))\|^p\big)\bigg)^{1/p}\\
&\qquad\le\kappa\Big(\max_{1\le i\le m}\|h(t,y_i)\|+\varepsilon+\delta\Big).
\end{aligned}
$$

We conclude, using the ergodicity of $h(t,y_i)$ for all $i\in\{1,\dots,m\}$, that $t\to h(t,Y(t))$ is in $\mathcal{E}(\mathbb{R},\mathrm{L}^p(\Omega,\mathrm{P},\mathbb{Y}),\mu)$.

Secondly, we show that $F(.):=f(.,X(.))-f(.,Y(.))$ is in $\mathcal{E}(\mathbb{R},\mathrm{L}^0(\Omega,\mathrm{P},\mathbb{Y}),\mu)$. Let

$$
\Phi:\left\{\begin{array}{ccc}\mathbb{X}&\to&\mathrm{C}(\mathbb{R},\mathbb{Y})\\ x&\mapsto&f(.,x).\end{array}\right.
$$

Let us endow $\mathrm{C}(\mathbb{R},\mathbb{Y})$ with the distance $\mathfrak{d}_\infty(\varphi,\psi)=\sup_{t\in\mathbb{R}}\|\varphi(t)-\psi(t)\|$. By the uniform continuity assumption, Φ is continuous. By [7, Lemma 5.6], for each $\varepsilon>0$, there exists $\eta>0$ such that, for all $x,y\in\mathbb{X}$,

$$
\Big(x\in K_\varepsilon\text{ and }\mathfrak{d}(x,y)\le\eta\Big)\ \Rightarrow\mathfrak{d}_\infty\Big(\Phi(x),\Phi(y)\Big)\le\varepsilon. \tag{3.10}
$$

For each $t\in\mathbb{R}$, let $\Omega_{\varepsilon,t}$ be the subset of Ω on which $Y(t)\in K_\varepsilon$. Since $Z(t)=X(t)-Y(t)$, we obtain, the following inequalities, with the help of (3.10) and Chebyshev's inequality:

$$
\frac{\mu\left\{t\in[-r,r];\ \mathrm{E}\big(\|F(t)\|\wedge 1\big)>3\varepsilon\right\}}{\mu([-r,r])}
$$

$$
\begin{aligned}
&\le \frac{\mu\left\{t\in[-r,r];\ \mathrm{E}\big(\mathbb{1}_{\Omega_{\varepsilon,t}}\mathbb{1}_{\{\|Z(t)\|>\eta\}}(\|F(t)\|\wedge 1)\big)>\varepsilon\right\}}{\mu([-r,r])}\\
&\quad+\frac{\mu\left\{t\in[-r,r];\ \mathrm{E}\big(\mathbb{1}_{\Omega_{\varepsilon,t}}\mathbb{1}_{\{\|Z(t)\|\le\eta\}}(\|F(t)\|\wedge 1)\big)>\varepsilon\right\}}{\mu([-r,r])}\\
&\quad+\frac{\mu\left\{t\in[-r,r];\ \mathrm{E}\big(\mathbb{1}_{\Omega^c_{\varepsilon,t}}(\|F(t)\|\wedge 1)\big)>\varepsilon\right\}}{\mu([-r,r])}\\
&= \frac{\mu\left\{t\in[-r,r];\ \mathrm{E}\big(\mathbb{1}_{\Omega_{\varepsilon,t}}\mathbb{1}_{\{\|Z(t)\|>\eta\}}(\|F(t)\|\wedge 1)\big)>\varepsilon\right\}}{\mu([-r,r])}\\
&\quad+\frac{\mu\left\{t\in[-r,r];\ \mathrm{E}\big(\mathbb{1}_{\Omega^c_{\varepsilon,t}}(\|F(t)\|\wedge 1)\big)>\varepsilon\right\}}{\mu([-r,r])}\\
&\le \frac{\mu\left\{t\in[-r,r];\ \mathrm{P}\{\|Z(t)\|>\eta\}>\varepsilon\right\}}{\mu([-r,r])}\\
&\quad+\frac{\mu\left\{t\in[-r,r];\ \mathrm{P}(\Omega^c_{\varepsilon,t})>\varepsilon\right\}}{\mu([-r,r])}\\
&\le \frac{\mu\left\{t\in[-r,r];\ \frac{1}{\eta}\mathrm{E}\big(\|Z(t)\|\big)>\varepsilon\right\}}{\mu([-r,r])}.
\end{aligned}
$$

Since Z is in $\mathcal{E}(\mathbb{R},\mathrm{L}^0(\Omega,\mathrm{P},\mathbb{X}),\mu)$, we have, for the above ε,

$$
\lim_{r\to\infty}\frac{\mu\left\{t\in[-r,r];\ \mathrm{E}\big(\|Z(t)\|\big)>\varepsilon\eta\right\}}{\mu([-r,r])}=0,
$$

which implies, using [7, Theorem 2.14] (see Remark 2.4),

$$
\limsup_{r\to\infty}\frac{1}{\mu([-r,r])}\int_{[-r,r]}\mathrm{E}\big(\|f(t,X(t))-f(t,Y(t))\|\wedge 1\big)d\mu(t)=0.
$$

Therefore $t\to f(t,X(t))-f(t,Y(t))$ is in $\mathcal{E}(\mathbb{R},\mathrm{L}^0(\Omega,\mathrm{P},\mathbb{Y}),\mu)$.

Assume now that $p>0$. If f is Lipschitz with respect to the second variable, uniformly with respect to the first one, then $\|F(t)\|\le K\,\|Z(t)\|$ for some constant K, thus, trivially, $F\in\mathcal{E}(\mathbb{R},\mathrm{L}^p(\Omega,\mathrm{P},\mathbb{Y}),\mu)$.

■

3.3 Pseudo-almost automorphy in p-mean vs in p-distribution

Let $X=(X_t)_{t\in\mathbb{R}}$ be a continuous stochastic process with values in $\mathbb{X}$, defined on a probability space $(\Omega,\mathcal{F},\mathrm{P})$. Let μ be a Borel measure on $\mathbb{R}$ satisfying (2.6). Clearly, we have for all $p\ge 0$,

$$
\Big(X\in\mathrm{AA}(\mathbb{R},\mathrm{L}^p(\Omega,\mathrm{P},\mathbb{X}))\Big)\ \Rightarrow\ \Big(X\in\mathrm{AAD}^p_f(\mathbb{R},\mathbb{X})\Big).
$$

Using Theorem 3.5, we can get more: if X satisfies (3.6), we deduce, for every $p \geq 0$,

$$\Big(X \in \mathrm{AA}(\mathbb{R}, \mathrm{L}^p(\Omega, \mathrm{P}, \mathbb{X}))\Big) \Rightarrow \Big(X \in \mathrm{AAD}^p(\mathbb{R}, \mathbb{X})\Big),$$
$$\Big(X \in \mathrm{PAA}(\mathbb{R}, \mathrm{L}^p(\Omega, \mathrm{P}, \mathbb{X}), \mu)\Big) \Rightarrow \Big(X \in \mathrm{PAAD}^p(\mathbb{R}, \mathbb{X}, \mu)\Big).$$

The converse implications are false. Indeed, Example 3.1 shows that a process which is almost automorphic in distribution is not necessarily almost automorphic in probability or in p-mean, see also [3, Counterexample 2.16][3].

The same counterexample also shows that a process which is μ-pseudo almost automorphic in p-distribution is not necessarily μ-pseudo almost automorphic in probability or in p-mean.

4 Pseudo almost automorphic solutions to stochastic differential equations

In the sequel, if $\mathbb{X}$ and $\mathbb{Y}$ are metric spaces, we denote $\mathrm{CUB}(\mathbb{X}, \mathbb{Y})$ the space of bounded uniformly continuous functions from $\mathbb{X}$ to $\mathbb{Y}$.

We are given two separable Hilbert spaces $\mathbb{H}_1$ and $\mathbb{H}_2$, and we consider the semilinear stochastic differential equation,

$$dX_t = AX(t)\,dt + f(t, X(t))\,dt + g(t, X(t))\,dW(t), \ t \in \mathbb{R} \tag{4.1}$$

where $A : \mathrm{Dom}(A) \subset \mathbb{H}_2 \to \mathbb{H}_2$ is a densely defined closed (possibly unbounded) linear operator, and $f : \mathbb{R} \times \mathbb{H}_2 \to \mathbb{H}_2$, and $g : \mathbb{R} \times \mathbb{H}_2 \to L(\mathbb{H}_1, \mathbb{H}_2)$ are continuous functions. In this section, we assume that:

(i) $W(t)$ is an $\mathbb{H}_1$-valued Wiener process with nuclear covariance operator Q (we denote by $\mathrm{tr}\, Q$ the trace of Q), defined on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}}, \mathrm{P})$.

(ii) $A : \mathrm{Dom}(A) \to \mathbb{H}_2$ is the infinitesimal generator of a C_0-semigroup $(S(t))_{t \geq 0}$ such that there exists a constant $\delta > 0$ with

$$\|S(t)\|_{L(\mathbb{H}_2)} \leq e^{-\delta t}, t \geq 0.$$

(iii) There exists a constant K such that the mappings $f : \mathbb{R} \times \mathbb{H}_2 \to \mathbb{H}_2$ and $g : \mathbb{R} \times \mathbb{H}_2 \to L(\mathbb{H}_1, \mathbb{H}_2)$ satisfy

$$\|f(t, x)\|_{\mathbb{H}_2} + \|g(t, x)\|_{L(\mathbb{H}_1, \mathbb{H}_2)} \leq K(1 + \|x\|_{\mathbb{H}_2}).$$

[3] Let us here point out an infortunate error in [38]: it is mistakenly said at the end of Section 1 of [38] that almost periodicity in square mean implies almost periodicity in distribution, and that the converse is true under a tightness condition. The first claim is true under the tightness condition (3.6), whereas Example 3.1, which is also Example 2.1 of [38], disproves the second claim.

(iv) The functions f and g are Lipschitz, more precisely there exists a constant K such that

$$\|f(t,x) - f(t,y)\|_{\mathbb{H}_2} + \|g(t,x) - g(t,y)\|_{L(\mathbb{H}_1,\mathbb{H}_2)} \leq K\|x-y\|_{\mathbb{H}_2}$$

for all $t \in \mathbb{R}$ and $x, y \in \mathbb{H}_2$.

(v) $f \in \mathrm{PAAU}_b(\mathbb{R}\times\mathbb{H}_2, \mathbb{H}_2, \mu)$ and $g \in \mathrm{PAAU}_b(\mathbb{R}\times\mathbb{H}_2, L(\mathbb{H}_1,\mathbb{H}_2), \mu)$ for some given Borel measure μ on $\mathbb{R}$ which satisfies (2.6) and Condition (**H**).

By [7, Theorem 3.5], Condition (**H**) implies that $\mathcal{E}(\mathbb{R}, \mathbb{X}, \mu)$ and $\mathrm{PAA}(\mathbb{R}, \mathbb{X}, \mu)$ are translation invariant.

In order to study the weighted pseudo almost automorphy property of solutions of SDEs, we need a result on almost automorphy.

Theorem 4.1 (Almost automorphic solution of an equation with almost automorphic coefficients) *Let the assumptions (i) - (iv) be fulfilled, and assume furthermore the following condition, which is stronger than (v) :*

(v') $f \in AAU_b(\mathbb{R}\times\mathbb{H}_2, \mathbb{H}_2)$ and $g \in AAU_b(\mathbb{R}\times\mathbb{H}_2, L(\mathbb{H}_1,\mathbb{H}_2))$.

Assume further that $\theta := \dfrac{K^2}{\delta}\left(\dfrac{1}{2\delta} + \operatorname{tr} Q\right) < 1$. Then there exists a unique mild solution X to (4.1) in the space $\mathrm{CUB}\big(\mathbb{R}, \mathrm{L}^2(\mathrm{P}, \mathbb{H}_2)\big)$ of bounded uniformly continuous mappings from $\mathbb{R}$ to $\mathrm{L}^2(\mathrm{P}, \mathbb{H}_2)$. Furthermore, X has a.e. continuous trajectories, and $X(t)$ satisfies, for each $t \in \mathbb{R}$:

$$X(t) = \int_{-\infty}^{t} S(t-s) f\big(s, X(s)\big) ds + \int_{-\infty}^{t} S(t-s) g\big(s, X(s)\big) dW(s). \tag{4.2}$$

If furthermore $\theta' := \dfrac{4K^2}{\delta}\left(\dfrac{1}{\delta} + \operatorname{tr} Q\right) < 1$, then X is almost automorphic in 2-distribution.

The proof of this theorem is very similar to that of [31, Theorem 3.1], which is the analogous result for SDEs with almost periodic coefficients. Only the almost automorphy part needs to be adapted. Such an adaptation is provided in [35], for SDEs driven by Lévy processes, but only for one-dimensional almost automorphy. We give the proof of this part for the convenience of the reader.

Let us first recall the following result, which is given in a more general form in [17]:

Proposition 4.2 *([17, Proposition 3.1-(c)]) Let $\tau \in \mathbb{R}$. Let $(\xi_n)_{0\leq n\leq\infty}$ be a sequence of square integrable $\mathbb{H}_2$-valued random variables. Let $(f_n)_{0\leq n\leq\infty}$ and $(g_n)_{0\leq n\leq\infty}$ be sequences of mappings from $\mathbb{R}\times\mathbb{H}_2$ to $\mathbb{H}_2$ and $L(\mathbb{H}_1,\mathbb{H}_2)$ respectively, satisfying* (iii) *and* (iv) *(replacing f and g by f_n and g_n respectively, and the constant K being independent of n). For each n, let X_n denote the solution to*

$$X_n(t) = S(t-\tau)\xi_n + \int_\tau^t S(t-s) f_n\big(s, X_n(s)\big)\,ds + \int_\tau^t S(t-s) g_n\big(s, X_n(s)\big)\,dW(s).$$

Assume that, for every $(t,x) \in \mathbb{R} \times \mathbb{H}_2$,

$$\lim_{n\to\infty} f_n(t,x) = f_\infty(t,x), \quad \lim_{n\to\infty} g_n(t,x) = g_\infty(t,x),$$
$$\lim_{n\to\infty} \mathfrak{d}_{\mathrm{BL}}(\mathrm{law}\,(\xi_n, W)\,, \mathrm{law}\,(\xi_\infty, W)) = 0,$$

(the last equality takes place in $\mathcal{M}^{1,+}\left(\mathbb{H}_2 \times \mathrm{C}(\mathbb{R}, \mathbb{H}_1)\right)$). Then we have in $\mathrm{C}([\tau, T]; \mathbb{H}_2)$, for any $T > \tau$,

$$\lim_{n\to\infty} \mathfrak{d}_{\mathrm{BL}}(\mathrm{law}\,(X_n)\,, \mathrm{law}\,(X_\infty)) = 0.$$

We need also a variant of Gronwall's lemma.

Lemma 4.3 *([31, Lemma 3.3]) Let $g : \mathbb{R} \to \mathbb{R}$ be a continuous function such that, for every $t \in \mathbb{R}$,*

$$0 \le g(t) \le \alpha(t) + \beta_1 \int_{-\infty}^t e^{-\delta_1(t-s)} g(s)\,ds + \cdots + \beta_n \int_{-\infty}^t e^{-\delta_n(t-s)} g(s)\,ds, \quad (4.3)$$

for some locally integrable function $\alpha : \mathbb{R} \to \mathbb{R}$, and for some constants $\beta_1, \ldots, \beta_n \ge 0$, and some constants $\delta_1, \ldots, \delta_n > \beta$, where $\beta := \sum_{i=1}^n \beta_i$. We assume that the integrals in the right hand side of (4.3) are convergent. Let $\delta = \min_{1\le i\le n} \delta_i$. Then, for every $\gamma \in]0, \delta - \beta]$ such that $\int_{-\infty}^0 e^{\gamma s}\alpha(s)\,ds$ converges, we have, for every $t \in \mathbb{R}$,

$$g(t) \le \alpha(t) + \beta \int_{-\infty}^t e^{-\gamma(t-s)} \alpha(s)\,ds.$$

In particular, if α is constant, we have

$$g(t) \le \alpha \, \frac{\delta}{\delta - \beta}.$$

Proof of Theorem 4.1 The proof of the existence and uniqueness of a mild solution to (4.1) in $\mathrm{CUB}\big(\mathbb{R}, \mathrm{L}^2(\mathrm{P}, \mathbb{H}_2)\big)$ is the same as that of Theorem 3.1 in [31] or Theorem 3.3.1 in [37].

For the almost automorphy part, let (γ'_n) be a sequence in $\mathbb{R}$. Since f and g are almost automorphic, there exists a subsequence (γ_n) and functions $\widehat{f} : \mathbb{R} \times \mathbb{H}_2 \to \mathbb{H}_2$ and $\widehat{g} : \mathbb{R} \times \mathbb{H}_2 \to L(\mathbb{H}_1, \mathbb{H}_2)$ such that

$$\lim_{n\to\infty} f(t+\gamma_n, x) = \widehat{f}(t,x), \quad \lim_{n\to\infty} \widehat{f}(t-\gamma_n, x) = f(t,x) \quad (4.4)$$
$$\lim_{n\to\infty} g(t+\gamma_n, x) = \widehat{g}(t,x), \quad \lim_{n\to\infty} \widehat{g}(t-\gamma_n, x) = g(t,x). \quad (4.5)$$

These limits are taken uniformly with respect to x in bounded subsets of $\mathbb{H}_2$.

For each fixed integer n, we consider

$$X_n(t) = \int_{-\infty}^{t} S(t-s)f(s+\gamma_n, X_n(s))\, ds + \int_{-\infty}^{t} S(t-s)g(s+\gamma_n, X_n(s))\, dW(s)$$

the mild solution to

$$dX_n(t) = AX_n(t)dt + f(t+\gamma_n, X_n(t))\, dt + g(t+\gamma_n, X_n(t))\, dW(t)$$

and

$$\widehat{X}(t) = \int_{-\infty}^{t} S(t-s)\widehat{f}(s, \widehat{X}(s))\, ds + \int_{-\infty}^{t} S(t-s)\widehat{g}(s, \widehat{X}(s))\, dW(s)$$

the mild solution to

$$d\widehat{X}(t) = A(t)\widehat{X}(t)dt + \widehat{f}(t, \widehat{X}(t))\, dt + \widehat{g}(t, \widehat{X}(t))\, dW(t).$$

Make the change of variable $\sigma + \gamma_n = s$, the process

$$X(t+\gamma_n) = \int_{-\infty}^{t+\gamma_n} S(t+\gamma_n - s)f(s, X(s))\, ds + \int_{-\infty}^{t+\gamma_n} S(t+\gamma_n - s)g(s, X(s))\, dW(s)$$

satisfies

$$X(t+\gamma_n) = \int_{-\infty}^{t} S(t-s)f(s+\gamma_n, X(s+\gamma_n))ds + \int_{-\infty}^{t} S(t-s)g(s+\gamma_n, X(s+\gamma_n))d\tilde{W}_n(s),$$

where $\tilde{W}_n(s) = W(s+\gamma_n) - W(\gamma_n)$ is a Brownian motion with the same distribution as $W(s)$. Thus the process $X(.+\gamma_n)$ has the same distribution as X_n.

Let us show that $X_n(t)$ converges in quadratic mean to $\widehat{X}(t)$ for each fixed $t \in \mathbb{R}$. We have

$$\mathrm{E}\|X_n(t) - \widehat{X}(t)\|^2$$

$$
\begin{aligned}
&= \mathrm{E}\Big\|\int_{-\infty}^{t} S(t-s)\Big(f(s+\gamma_n, X_n(s)) - \widehat{f}(s,\widehat{X}(s))\Big)\,ds \\
&\quad + \int_{-\infty}^{t} S(t-s)\Big(g(s+\gamma_n, X^n(s)) - \widehat{g}(s,\widehat{X}(s))\Big)\,dW(s)\Big\|^2 \\
&\leq 2\,\mathrm{E}\Big\|\int_{-\infty}^{t} S(t-s)\Big(f(s+\gamma_n, X_n(s)) - \widehat{f}(s,\widehat{X}(s))\Big)\,ds\Big\|^2 \\
&\quad + 2\,\mathrm{E}\int_{-\infty}^{t} S(t-s)\Big(g(s+\gamma_n, X_n(s)) - \widehat{g}(s,\widehat{X}(s))\Big)\,dW(s)\Big\|^2 \\
&\leq 4\,\mathrm{E}\Big\|\int_{-\infty}^{t} S(t-s)\Big(f(s+\gamma_n, X^n(s)) - f(s+\gamma_n,\widehat{X}(s))\Big)\,ds\Big\|^2 \\
&\quad + 4\,\mathrm{E}\|\int_{-\infty}^{t} S(t-s)\Big(f(s+\gamma_n, \widehat{X}(s)) - \widehat{f}(s,\widehat{X}(s))\Big)\,ds\Big\|^2 \\
&\quad + 4\,\mathrm{E}\|\int_{-\infty}^{t} S(t-s)\Big(g(s+\gamma_n, X^n(s)) - g(s+\gamma_n,\widehat{X}(s))\Big)\,dW(s)\Big\|^2 \\
&\quad + 4\,\mathrm{E}\|\int_{-\infty}^{t} S(t-s)\Big(g(s+\gamma_n, \widehat{X}(s)) - \widehat{g}(s,\widehat{X}(s))\Big)\,dW(s)\Big\|^2 \\
&\leq I_1 + I_2 + I_3 + I_4.
\end{aligned}
$$

Now, using (ii), (iv) and the Cauchy-Schwartz inequality, we obtain

$$
\begin{aligned}
I_1 &= 4\,\mathrm{E}\Big\|\int_{-\infty}^{t} S(t-s)\Big(f(s+\gamma_n, X_n(s)) - f(s+\gamma_n,\widehat{X}(s))\Big)\,ds\Big\|^2 \\
&\leq 4\,\mathrm{E}\Big(\int_{-\infty}^{t} \|S(t-s)\|\|f(s+\gamma_n, X_n(s)) - f(s+\gamma_n,\widehat{X}(s))\|\,ds\Big)^2 \\
&\leq 4\,\mathrm{E}\Big(\int_{-\infty}^{t} e^{-\delta(t-s)}\|f(s+\gamma_n, X_n(s)) - f(s+\gamma_n,\widehat{X}(s))\|ds\Big)^2 \\
&\leq 4\Big(\int_{-\infty}^{t} e^{-\delta(t-s)}\,ds\Big)\Big(\int_{-\infty}^{t} e^{-\delta(t-s)}\,\mathrm{E}\|f(s+\gamma_n, X_n(s)) - f(s+\gamma_n,\widehat{X}(s))\|^2 ds\Big) \\
&\leq \frac{4K^2}{\delta}\int_{-\infty}^{t} e^{-\delta(t-s)}\,\mathrm{E}\|X_n(s) - \widehat{X}(s)\|^2 ds.
\end{aligned}
$$

Then we have

$$
\begin{aligned}
I_2 &= 4\,\mathrm{E}\Big\|\int_{-\infty}^{t} S(t-s)[f(s+\gamma_n,\widehat{X}(s)) - \widehat{f}(s,\widehat{X}(s))]ds\Big\|^2 \\
&\leq 4\,\mathrm{E}\Big(\int_{-\infty}^{t} e^{-\delta(t-s)}\|f(s+\gamma_n,\widehat{X}(s)) - \widehat{f}(s,\widehat{X}(s))\|ds\Big)^2 \\
&\leq 4\,\mathrm{E}\Big(\int_{-\infty}^{t} e^{-\delta(t-s)}ds\Big)\Big(\int_{-\infty}^{t} e^{-\delta(t-s)}\|f(s+\gamma_n,\widehat{X}(s)) - \widehat{f}(s,\widehat{X}(s))\|^2 ds\Big)
\end{aligned}
$$

$$\leq 4\Big(\int_{-\infty}^{t} e^{-\delta(t-s)}ds\Big)^2 \sup_s \mathrm{E}\|f(s+\gamma_n,\widehat{X}(s)) - \widehat{f}(s,\widehat{X}(s))\|^2$$
$$\leq \frac{4}{\delta^2}\sup_s \mathrm{E}\|f(s+\gamma_n,\widehat{X}(s)) - \widehat{f}(s,\widehat{X}(s))\|^2,$$

which converges to 0 as $n \to \infty$ because $\sup_{t\in\mathbb{R}} \mathrm{E}\|\widehat{X}(t)\|^2 < \infty$ which implies that $(\widehat{X}(t))_t$ is tight relatively to bounded sets.

Applying Itô's isometry, we get

$$\begin{aligned}
I_3 &= 4\,\mathrm{E}\Big\|\int_{-\infty}^{t} S(t-s)\Big(g(s+\gamma_n,X_n(s)) - g(s+\gamma_n,\widehat{X}(s))\Big)\,dW(s)\Big\|^2\\
&\leq 4\,\mathrm{tr}\,Q\,\mathrm{E}\int_{-\infty}^{t}\|S(t-s)\|^2\,\|g(s+\gamma_n,X_n(s)) - g(s+\gamma_n,\widehat{X}(s))\|^2 ds\\
&\leq 4\,\mathrm{tr}\,Q\int_{-\infty}^{t} e^{-2\delta(t-s)}\,\mathrm{E}\|g(s+\gamma_n,X_n(s)) - g(s+\gamma_n,\widehat{X}(s))\|^2 ds\\
&\leq 4K^2\,\mathrm{tr}\,Q\int_{-\infty}^{t} e^{-2\delta(t-s)}\,\mathrm{E}\|X_n(s)-\widehat{X}(s)\|^2 ds,
\end{aligned}$$

and

$$\begin{aligned}
I_4 &= 4\,\mathrm{E}\Big\|\int_{-\infty}^{t} S(t-s)\Big(g(s+\gamma_n,\widehat{X}(s)) - \widehat{g}(s,\widehat{X}(s))\Big)\,dW(s)\Big\|^2\\
&\leq 4\,\mathrm{tr}\,Q\,\mathrm{E}\Big(\int_{-\infty}^{t}\|S(t-s)\|^2\|g(s+\gamma_n,\widehat{X}(s)) - \widehat{g}(s,\widehat{X}(s))\|^2 ds\Big)\\
&\leq 4\,\mathrm{tr}\,Q\Big(\int_{-\infty}^{t} e^{-2\delta(t-s)}ds\Big)\sup_{s\in\mathbb{R}}\mathrm{E}\|g(s+\gamma_n,\widehat{X}(s)) - \widehat{g}(s,\widehat{X}(s))\|^2\\
&\leq \frac{2\,\mathrm{tr}\,Q}{\delta}\sup_{s\in\mathbb{R}}\mathrm{E}\|g(s+\gamma_n,\widehat{X}(s)) - \widehat{g}(s,\widehat{X}(s))\|^2.
\end{aligned}$$

For the same reason as for I_2, the right hand term goes to 0 as $n \to \infty$.

We thus have

$$\begin{aligned}
\mathrm{E}\|X_n(t)-\widehat{X}(t)\|^2 \leq \alpha_n &+ \frac{4K^2}{\delta}\int_{-\infty}^{t} e^{-\delta(t-s)}\,\mathrm{E}\|X_n(s)-\widehat{X}(s)\|^2\,ds\\
&+ 4K^2\,\mathrm{tr}\,Q\int_{-\infty}^{t} e^{-2\delta(t-s)}\,\mathrm{E}\|X_n(s)-\widehat{X}(s)\|^2\,ds
\end{aligned}$$

for a sequence (α_n) such that $\lim_{n\to\infty}\alpha_n = 0$. Furthermore, $\beta := \frac{4K^2}{\delta} + 4K^2\,\mathrm{tr}\,Q < \delta$. We conclude by Lemma 4.3 that

$$\lim_{n\to\infty}\mathrm{E}\|X_n(t)-\widehat{X}(t)\|^2 = 0,$$

hence $X_n(t)$ converges in distribution to $\widehat{X}(t)$. But, since the distribution of $X_n(t)$ is the same as that of $X(t+\gamma_n)$, we deduce that $X(t+\gamma_n)$ converges in distribution to $\widehat{X}(t)$. By analogy and using (4.4), (4.5) we can easily prove that $\widehat{X}(t-\gamma_n)$ converges in distribution to $X(t)$.

Note that the sequence $(\|X_n(t)\|^2)$ is uniformly integrable, thus $(\|X(t+\gamma_n)\|^2)$ is uniformly integrable too. As (γ'_n) is arbitrary, this implies that the family $(\|X(t)\|^2)_{t\in\mathbb{R}}$ is uniformly integrable, because, if not, there would exist a sequence (γ'_n) and $t\in\mathbb{R}$ such that no subsequence of $(\|X(t+\gamma'_n)\|^2)$ is uniformly integrable.

We have thus proved that X has almost automorphic one-dimensional 2-distributions. To prove that X is almost automorphic in 2-distribution, we apply Proposition 4.2: for fixed $\tau\in\mathbb{R}$, let $\xi_n = X(\tau+\gamma_n)$, $f_n(t,x) = f(t+\gamma_n,x)$, $g_n(t,x) = g(t+\gamma_n,x)$. By the foregoing, (ξ_n) converges in distribution to some variable $Y(\tau)$. We deduce that (ξ_n) is tight, and thus (ξ_n, W) is tight also. We can thus choose a subsequence (still noted (γ_n) for simplicity) such that (ξ_n, W) converges in distribution to $(Y(\tau), W)$. Then, by Proposition 4.2, for every $T\geq\tau$, $X(.+\gamma_n)$ converges in distribution on $\mathrm{C}([\tau,T];\mathbb{H}_2)$ to the (unique in distribution) solution to

$$Y(t) = S(t-\tau)Y(\tau) + \int_\tau^t S(t-s)f\big(s,Y(s)\big)\,ds + \int_\tau^t S(t-s)g\big(s,Y(s)\big)\,dW(s).$$

Note that Y does not depend on the chosen interval $[\tau,T]$, thus the convergence takes place on $\mathrm{C}(\mathbb{R};\mathbb{H}_2)$. Similarly, $Y_n := Y(.-\gamma_n)$ converges in distribution on $\mathrm{C}(\mathbb{R};\mathbb{H}_2)$ to X. Thus X is almost automorphic in 2-distribution. ∎

We are now ready to prove our main result.

Theorem 4.4 (Weighted pseudo almost automorphic solution of an equation with weighted pseudo almost automorphic coefficients) *Let the assumptions (i) - (v) be fulfilled. Let (f_1,g_1) and (f_2,g_2) be respectively the decompositions of f and g, namely,*

$$\begin{gathered} f = f_1 + f_2, \quad g = g_1 + g_2,\\ f_1 \in AAU_b(\mathbb{R}\times\mathbb{H}_2,\mathbb{H}_2),\;\; f_2\in\mathcal{E}\,U_b(\mathbb{R}\times\mathbb{H}_2,\mathbb{H}_2,\mu),\\ g_1\in AAU_b(\mathbb{R}\times\mathbb{H}_2, L(\mathbb{H}_1,\mathbb{H}_2)),\;\; g_2\in\mathcal{E}\,U_b(\mathbb{R}\times\mathbb{H}_2, L(\mathbb{H}_1,\mathbb{H}_2),\mu). \end{gathered}$$

Assume that f_1 and g_1 satisfy the same growth and Lipschitz conditions (iii) - (v) as f and g respectively, with same coefficient K. Assume furthermore that

$$\theta' := \frac{4K^2}{\delta}\left(\frac{1}{\delta} + \operatorname{tr} Q\right) < 1.$$

Then there exists a unique mild solution X to (4.1) *in the space* $\mathrm{CUB}\big(\mathbb{R},\mathrm{L}^2(\mathrm{P},\mathbb{H}_2)\big)$ *of bounded uniformly continuous mappings from $\mathbb{R}$ to $\mathrm{L}^2(\mathrm{P},\mathbb{H}_2)$, X has a.e. continuous trajectories, and X satisfies* (4.2) *for every $t\in\mathbb{R}$. Furthermore, X is μ-pseudo almost automorphic in 2-distribution. More precisely, let $Y\in$* $\mathrm{CUB}\big(\mathbb{R},\mathrm{L}^2(\mathrm{P},\mathbb{H}_2)\big)$ *be the unique almost automorphic in distribution mild solution to*

$$dY(t) = AY(t)\,dt + f_1(t,Y(t))\,dt + g_1(t,Y(t))\,dW(t),\;\; t\in\mathbb{R}. \tag{4.6}$$

Then X has the decomposition

$$X = Y + Z, \quad Z \in \mathcal{E}\big(\mathbb{R}, \mathrm{L}^2(\mathrm{P}, \mathbb{H}_2), \mu\big).$$

The following technical lemma will be used several times.

Lemma 4.5 *Let $\mathfrak{h} \in \mathcal{E}(\mathbb{R}, \mathbb{R}, \mu)$. Then the function*

$$t \mapsto \left(\int_{-\infty}^{t} e^{-2\delta(t-s)} \mathfrak{h}^2(s)\, ds \right)^{1/2}$$

is also in $\mathcal{E}(\mathbb{R}, \mathbb{R}, \mu)$.

Proof by Condition (**H**) and [7, Theorem 3.9], we have, for every $u \in \mathbb{R}$,

$$\lim_{r \to +\infty} \frac{1}{\mu([-r,r])} \int_{[-r,r]} |\mathfrak{h}(t-u)|\; d\mu(t) = 0.$$

We deduce, by Lebesgue's dominated convergence theorem,

$$\begin{aligned}
&\frac{1}{\mu([-r,r])} \int_{[-r,r]} \left(\int_{-\infty}^{t} e^{-2\delta(t-s)} \mathfrak{h}^2(s)\, ds \right)^{1/2} d\mu(t) \\
&\qquad \leq \frac{1}{\big(\mu([-r,r])\big)^{1/2}} \left(\int_{[-r,r]} \int_{-\infty}^{t} e^{-2\delta(t-s)} \mathfrak{h}^2(s)\, ds\, d\mu(t) \right)^{1/2} \\
&\qquad = \frac{1}{\big(\mu([-r,r])\big)^{1/2}} \left(\int_{[-r,r]} \int_{0}^{+\infty} e^{-2\delta u} \mathfrak{h}^2(t-u)\, du\, d\mu(t) \right)^{1/2} \\
&\qquad = \frac{1}{\big(\mu([-r,r])\big)^{1/2}} \left(\int_{0}^{+\infty} e^{-2\delta u} \int_{[-r,r]} \mathfrak{h}^2(t-u)\, d\mu(t)\, du \right)^{1/2} \\
&\qquad \leq \left(\int_{0}^{+\infty} e^{-2\delta u} \left\|\mathfrak{h}\right\|_\infty \frac{\int_{[-r,r]} |\mathfrak{h}(t-u)|\, d\mu(t)}{\mu([-r,r])}\, du \right)^{1/2} \\
&\qquad \to 0 \text{ when } r \to +\infty.
\end{aligned}$$

■

Proof of Theorem 4.4 The existence and the properties of Y are guaranteed by Theorem 4.1.

As in Theorem 4.1, the existence and uniqueness of the mild solution X to (4.1) are proved as in [31, Theorem 3.1], using the classical method of the fixed point theorem for the contractive operator L on $\mathrm{CUB}\big(\mathbb{R}, \mathrm{L}^2(\mathrm{P}, \mathbb{H}_2)\big)$ defined by

$$LX(t) = \int_{-\infty}^{t} S(t-s) f\big(s, X(s)\big) ds + \int_{-\infty}^{t} S(t-s) g\big(s, X(s)\big) dW(s).$$

The solution X defined by (4.2) is thus the limit in $\mathrm{CUB}\big(\mathbb{R}, \mathrm{L}^2(\mathrm{P}, \mathbb{H}_2)\big)$ of a sequence (X_n) with arbitrary X_0 and, for every n, $X_{n+1} = L(X_n)$. To prove that X is μ-pseudo almost automorphic in 2-distribution we choose a special sequence. Set

$$X_0 = Y,\ X_{n+1} = L(X_n),\ Z_n = X_n - Y,\ n \in \mathbb{N}.$$

Let us prove that each Z_n is in $\mathcal{E}\big(\mathbb{R}, \mathrm{L}^2(\mathrm{P}, \mathbb{H}_2), \mu\big)$. We use some arguments of the proof of [7, Theorem 5.7]. We have, for every $n \in \mathbb{N}$ and every $t \in \mathbb{R}$,

$$\begin{aligned} Z_{n+1}(t) &= LX_n(t) - Y(t) \\ &= \int_{-\infty}^{t} S(t-s)\big(f(s, X_n(s)) - f(s, Y(s))\big)\, ds \\ &\quad + \int_{-\infty}^{t} S(t-s)\big(g(s, X_n(s)) - g(s, Y(s))\big)\, dW(s) \\ &\quad + \int_{-\infty}^{t} S(t-s)\big(f(s, Y(s)) - f_1(s, Y(s))\big)\, ds \\ &\quad + \int_{-\infty}^{t} S(t-s)\big(g(s, Y(s)) - g_1(s, Y(s))\big)\, dW(s) \\ &= \int_{-\infty}^{t} S(t-s)\big(f(s, X_n(s)) - f(s, Y(s))\big)\, ds \\ &\quad + \int_{-\infty}^{t} S(t-s)\big(g(s, X_n(s)) - g(s, Y(s))\big)\, dW(s) \\ &\quad + \int_{-\infty}^{t} S(t-s) f_2(s, Y(s))\, ds + \int_{-\infty}^{t} S(t-s) g_2(s, Y(s))\, dW(s). \end{aligned}$$

Assume that $Z_n \in \mathcal{E}\big(\mathbb{R}, \mathrm{L}^2(\mathrm{P}, \mathbb{H}_2), \mu\big)$. By the Lipschitz condition (iv),

$$\big(\mathrm{E}\, \|f(t, X_n(t)) - f(t, Y(t))\|^2\big)^{1/2} \le K \big(\mathrm{E}\, \|Z_n(t)\|^2\big)^{1/2}$$

thus the mapping

$$\mathfrak{f} : t \mapsto \big(\mathrm{E}\, \|f(t, X_n(t)) - f(t, Y(t))\|^2\big)^{1/2}$$

is in $\mathcal{E}(\mathbb{R}, \mathbb{R}, \mu)$. The same conclusion holds for

$$\mathfrak{g} : t \mapsto \big(\mathrm{E}\, \|g(t, X_n(t)) - g(t, Y(t))\|^2\big)^{1/2}.$$

We get, using Lemma 4.5,

$$\begin{aligned} \frac{1}{\mu([-r, r])} \int_{[-r,r]} \left(\mathrm{E} \left\| \int_{-\infty}^{t} S(t-s)\big(f(s, X_n(s)) - f(s, Y(s))\big)\, ds \right\|^2 \right)^{1/2} d\mu(t) \\ \le \frac{1}{\mu([-r, r])} \int_{[-r,r]} \left(\int_{-\infty}^{t} e^{-2\delta(t-s)} \mathfrak{f}^2(s)\, ds \right)^{1/2} d\mu(t) \\ \to 0 \text{ when } r \to +\infty, \end{aligned}$$

and

$$\frac{1}{\mu([-r,r])}\int_{[-r,r]}\left(\mathrm{E}\left\|\int_{-\infty}^{t}S(t-s)\big(g(s,X_n(s))-g(s,Y(s))\big)\,dW(s)\right\|^2\right)^{1/2}d\mu(t)$$
$$\leq (\operatorname{tr}Q)^{1/2}\frac{1}{\mu([-r,r])}\int_{[-r,r]}\left(\int_{-\infty}^{t}e^{-2\delta(t-s)}\mathfrak{g}^2(s)\,ds\right)^{1/2}d\mu(t)$$
$$\to 0 \text{ when } r\to+\infty.$$

To prove that Z_{n+1} is in $\mathcal{E}\big(\mathbb{R},\mathrm{L}^2(\mathrm{P},\mathbb{H}_2),\mu\big)$, there only remains to show that the process $\int_{-\infty}^{t}S(t-s)f_2(s,Y(s))\,ds+\int_{-\infty}^{t}S(t-s)g_2(s,Y(s))\,dW(s)$ belongs to $\mathcal{E}\big(\mathbb{R},\mathrm{L}^2(\mathrm{P},\mathbb{H}_2),\mu\big)$. As Y is almost automorphic in distribution, the family $(\tilde{Y}(t))=(Y(t+.))_{t\in\mathbb{R}}$ is uniformly tight in $\mathrm{C}_k(\mathbb{R},\mathbb{H}_2)$. In particular, for each $\varepsilon>0$ there exists a compact subset $\mathcal{K}_\varepsilon$ of $\mathrm{C}_k(\mathbb{R},\mathbb{H}_2)$ such that, for every $t\in\mathbb{R}$,

$$\mathrm{P}\left\{\tilde{Y}(t)\in\mathcal{K}_\varepsilon\right\}\geq 1-\varepsilon.$$

By the Arzelà-Ascoli Theorem (e.g. [24, Theorems 8.2.10 and 8.2.11]), this implies that, for every $\varepsilon>0$, and for every compact interval I of $\mathbb{R}$, there exists a compact subset $K_{\varepsilon,I}$ of $\mathbb{H}_2$ such that, for every $t\in\mathbb{R}$,

$$\mathrm{P}\left\{(\forall s\in I)\ Y(t+s)\in K_{\varepsilon,I}\right\}\geq 1-\varepsilon.$$

In particular, the family $(Y(t))_{t\in\mathbb{R}}$ is tight, i.e., denoting $K_\varepsilon=K_{\varepsilon,\{0\}}$, we have, for every $t\in\mathbb{R}$,

$$\mathrm{P}\left\{Y(t)\in K_\varepsilon\right\}\geq 1-\varepsilon.$$

By the uniform continuity property of f_2 and g_2 on K_ε, there exists $\eta(\varepsilon)>0$ such that, for all $x,y\in K_\varepsilon$,

$$\|x-y\|\leq\eta(\varepsilon)\Rightarrow\sup_{t\in\mathbb{R}}\max\big(\|f_2(t,x)-f_2(t,y)\|,\|g_2(t,x)-g_2(t,y)\|\big)\leq\varepsilon.$$

We can find a finite sequence $y_1,\dots,y_m$ such that

$$K_\varepsilon\subset\bigcup_{i=1}^{m}B(y_i,\eta(\varepsilon)).$$

By [31, Remark 3.6]), the condition $\theta'<1$ ensures that Y is bounded in $\mathrm{L}^p(\mathrm{P},\mathbb{H}_2)$ for some $p>2$ (the same result holds for X, but we do not need it). Note that $f_2=f-f_1$ and $g_2=g-g_1$ satisfy a condition similar to (iii), which implies that $f_2(.,Y(.))$ and $g_2(.,Y(.))$ are bounded in $\mathrm{L}^p(\mathrm{P},\mathbb{H}_2)$ and $\mathrm{L}^p\big(\mathrm{P},L(\mathbb{H}_1,\mathbb{H}_2)\big)$ respectively. Let

$$\mathfrak{M}_p=\sup_{t\in\mathbb{R}}\max\big(\mathrm{E}\|f_2(.,Y(.))\|^p,\mathrm{E}\|g_2(.,Y(.))\|^p\big)^{2/p}.$$

Let $q = p/(p-2)$. Let $t \in \mathbb{R}$, and let $\Omega_{\varepsilon,t}$ be the measurable subset of Ω on which $Y(t) \in K_\varepsilon$. We have

$$\begin{aligned}
&\left(\mathrm{E}\left\|f_2(t,Y(t))\right\|^2\right)^{1/2} \\
&\qquad \le \min_{1\le i\le m}\left(\mathrm{E}\big(\mathbb{1}_{\Omega_{\varepsilon,t}}\left\|f_2(t,Y(t))-f(t,y_i)\right\|^2\big)\right)^{1/2}+\max_{1\le i\le m}\left\|f_2(t,y_i)\right\| \\
&\qquad\quad +\left(\mathrm{E}\big(\mathbb{1}_{\Omega^c_{\varepsilon,t}}\left\|f_2(t,Y(t))\right\|^2\big)\right)^{1/2} \\
&\qquad \le \varepsilon+\max_{1\le i\le m}\left\|f_2(t,y_i)\right\|+\left(\mathrm{P}\left(\Omega^c_{\varepsilon,t}\right)\right)^{1/q}\left(\mathrm{E}\left\|f_2(t,Y(t))\right\|^p\right)^{2/p} \\
&\qquad \le \varepsilon+\max_{1\le i\le m}\left\|f_2(t,y_i)\right\|+\varepsilon^{1/q}\mathfrak{M}_p.
\end{aligned}$$

A similar result holds for $\mathrm{E}\left\|g_2(t,Y(t))\right\|^2$. Let us denote

$$\begin{aligned}
\mathfrak{E}(\varepsilon) &= \varepsilon+\varepsilon^{1/q}\mathfrak{M}_p, \\
\mathfrak{f}_\varepsilon(t) &= \max_{1\le i\le m}\left\|f_2(s,y_i)\right\|, \\
\mathfrak{g}_\varepsilon(t) &= \max_{1\le i\le m}\left\|g_2(s,y_i)\right\|.
\end{aligned}$$

Thanks to the ergodicity of f_2 and g_2, the functions $\mathfrak{f}_\varepsilon$ and $\mathfrak{g}_\varepsilon$ are in $\mathcal{E}(\mathbb{R},\mathbb{R},\mu)$. We have

$$\begin{aligned}
\frac{1}{\mu([-r,r])}\int_{[-r,r]}\Bigg(\mathrm{E}\bigg\|&\int_{-\infty}^t S(t-s)f_2(s,Y(s))\,ds \\
&+\int_{-\infty}^t S(t-s)g_2(s,Y(s))\,dW(s)\bigg\|^2\Bigg)^{1/2}d\mu(t)
\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{\mu([-r,r])} \int_{[-r,r]} \Bigg(\bigg(\Big(\int_{-\infty}^{t} e^{-2\delta(t-s)} \, \mathrm{E} \, \|f_2(s,Y(s))\|^2 \, ds \Big)^{1/2} \\
&\quad + (\operatorname{tr} Q)^{1/2} \Big(\int_{-\infty}^{t} e^{-2\delta(t-s)} \, \mathrm{E} \, \|g_2(s,Y(s))\|^2 \, ds \Big)^{1/2} \bigg) d\mu(t) \\
&\leq \frac{1}{\mu([-r,r])} \int_{[-r,r]} \Big(\int_{-\infty}^{t} e^{-2\delta(t-s)} \left(\mathfrak{f}_\varepsilon(s) + \mathfrak{E}(\varepsilon) \right)^2 \, ds \Big)^{1/2} d\mu(t) \\
&\quad + (\operatorname{tr} Q)^{1/2} \frac{1}{\mu([-r,r])} \int_{[-r,r]} \Big(\int_{-\infty}^{t} e^{-2\delta(t-s)} \left(\mathfrak{g}_\varepsilon(s) + \mathfrak{E}(\varepsilon) \right)^2 \, ds \Big)^{1/2} d\mu(t) \\
&\leq \frac{1}{\mu([-r,r])} \int_{[-r,r]} \Big(\int_{-\infty}^{t} e^{-2\delta(t-s)} \mathfrak{f}_\varepsilon^2(s) \, ds \Big)^{1/2} d\mu(t) \\
&\quad + (\operatorname{tr} Q)^{1/2} \frac{1}{\mu([-r,r])} \int_{[-r,r]} \Big(\int_{-\infty}^{t} e^{-2\delta(t-s)} \mathfrak{g}_\varepsilon^2(s) \, ds \Big)^{1/2} d\mu(t) \\
&\quad + \frac{1 + (\operatorname{tr} Q)^{1/2}}{\mu([-r,r])} \mathfrak{E}(\varepsilon) \int_{[-r,r]} \Big(\int_{-\infty}^{t} e^{-2\delta(t-s)} \, ds \Big)^{1/2} d\mu(t).
\end{aligned}
$$

In the right hand side of the last inequality, the last term is arbitrarily small and both other terms converge to 0 when r goes to $+\infty$, thanks to Lemma 4.5. We have thus proved that Z_{n+1} is in $\mathcal{E}\big(\mathbb{R}, \mathrm{L}^2(\mathrm{P}, \mathbb{H}_2), \mu\big)$. We deduce by induction that the sequence (Z_n) lies in $\mathcal{E}\big(\mathbb{R}, \mathrm{L}^2(\mathrm{P}, \mathbb{H}_2), \mu\big)$.

Now, the sequence (X_n) converges to X in $\mathrm{CUB}\big(\mathbb{R}, \mathrm{L}^2(\mathrm{P}, \mathbb{H}_2)\big)$, thus (Z_n) converges to $Z := X - Y$ in $\mathrm{CUB}\big(\mathbb{R}, \mathrm{L}^2(\mathrm{P}, \mathbb{H}_2)\big)$. Let $\varepsilon > 0$, and let n such that

$$
\sup_{t \in \mathbb{R}} \big(\mathrm{E} \, \|Z(t) - Z_n(t)\|^2 \big)^{1/2} \leq \varepsilon.
$$

We have

$$
\begin{aligned}
&\frac{1}{\mu([-r,r])} \int_{[-r,r]} \big(\mathrm{E} \, \|Z(t)\|^2 \big)^{1/2} d\mu(t) \\
&\qquad \leq \frac{1}{\mu([-r,r])} \int_{[-r,r]} \big(\mathrm{E} \, \|Z(t) - Z_n(t)\|^2 \big)^{1/2} d\mu(t) \\
&\qquad\quad + \frac{1}{\mu([-r,r])} \int_{[-r,r]} \big(\mathrm{E} \, \|Z_n(t)\|^2 \big)^{1/2} d\mu(t) \\
&\qquad \leq \varepsilon + \frac{1}{\mu([-r,r])} \int_{[-r,r]} \big(\mathrm{E} \, \|Z_n(t)\|^2 \big)^{1/2} d\mu(t).
\end{aligned}
$$

As ε is arbitrary, this proves that $Z \in \mathcal{E}\big(\mathbb{R}, \mathrm{L}^2(\mathrm{P}, \mathbb{H}_2), \mu\big)$. ∎

Remark 4.6 We did not use in the proof of Theorem 4.4 the hypothesis that $f_2 \in \mathcal{E}\mathrm{U}_b(\mathbb{R} \times \mathbb{H}_2, \mathbb{H}_2, \mu)$ and $g_2 \in \mathcal{E}\mathrm{U}_b(\mathbb{R} \times \mathbb{H}_2, L(\mathbb{H}_1, \mathbb{H}_2), \mu)$. Actually, we needed only to assume that, for each $x \in \mathbb{H}_2$, $f_2(.,x) \in \mathcal{E}(\mathbb{R}, \mathbb{H}_2, \mu)$ and $g_2(.,x) \in$

$\mathcal{E}(\mathbb{R}, L(\mathbb{H}_1, \mathbb{H}_2), \mu)$. By Remark 2.3 and the Lipschitz condition, this is equivalent to assume that $f_2 \in \mathcal{E}\mathrm{U}_c(\mathbb{R}\times\mathbb{H}_2, \mathbb{H}_2, \mu)$ and $g_2 \in \mathcal{E}\mathrm{U}_c(\mathbb{R}\times\mathbb{H}_2, L(\mathbb{H}_1, \mathbb{H}_2), \mu)$.

Theorem 4.7 (Weighted pseudo almost periodic solution of an equation with weighted pseudo almost periodic coefficients) *Assume the same hypothesis as in Theorem 4.4, and that f_1 and g_1 are almost periodic with respect to the first variable, uniformly with respect to the second variable in bounded sets. Then (with obvious definitions) the process Y of Theorem 4.4 is almost periodic in 2-distribution, thus the process X is μ-pseudo almost periodic in 2-distribution.*

Proof The proof is exactly the same as that of Theorem 4.4, replacing Theorem 4.1 by [31, Theorem 3.1]. ∎

Acknowledgements We thank an anonymous referee for pointing out an error in a previous version.

References

[1] Syed Abbas. A note on Weyl pseudo almost automorphic functions and their properties. *Math. Sci. (Springer)*, 6:Art. 29, 5, 2012.

[2] Ludwig Arnold and Constantin Tudor. Stationary and almost periodic solutions of almost periodic affine stochastic differential equations. *Stochastics Stochastics Rep.*, 64(3-4):177–193, 1998.

[3] Fazia Bedouhene, Omar Mellah, and Paul Raynaud de Fitte. Bochner-almost periodicity for stochastic processes. *Stoch. Anal. Appl.*, 30(2):322–342, 2012.

[4] Paul H. Bezandry and Toka Diagana. Existence of almost periodic solutions to some stochastic differential equations. *Appl. Anal.*, 86(7):819–827, 2007.

[5] Patrick Billingsley. *Convergence of probability measures.* Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[6] J. Blot, G. M. Mophou, G. M. N'Guérékata, and D. Pennequin. Weighted pseudo almost automorphic functions and applications to abstract differential equations. *Nonlinear Anal.*, 71(3-4):903–909, 2009.

[7] Joël Blot, Philippe Cieutat, and Khalil Ezzinbi. Measure theory and pseudo almost automorphic functions: new developments and applications. *Nonlinear Anal.*, 75(4):2426–2447, 2012.

[8] Joël Blot, Philippe Cieutat, and Khalil Ezzinbi. New approach for weighted pseudo-almost periodic functions under the light of measure theory, basic results and applications. *Appl. Anal.*, 92(3):493–526, 2013.

[9] Joël Blot, Philippe Cieutat, Gaston M. N'Guérékata, and Denis Pennequin. Superposition operators between various almost periodic function spaces and applications. *Commun. Math. Anal.*, 6(1):42–70, 2009.

[10] S. Bochner. Abstrakte Fastperiodische Funktionen. *Acta Math.*, 61(1):149–184, 1933.

[11] S. Bochner. Uniform convergence of monotone sequences of functions. *Proc. Nat. Acad. Sci. U.S.A.*, 47:582–585, 1961.

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

[13] Salomon Bochner. Curvature and Betti numbers in real and complex vector bundles. *Univ. e Politec. Torino. Rend. Sem. Mat.*, 15:225–253, 1955–56.

[14] N. Bourbaki. *Éléments de mathématique. Topologie générale. Chapitres 1 à 4.* Hermann, Paris, 1971.

[15] Valentina Casarino. Almost automorphic groups and semigroups. *Rend. Accad. Naz. Sci. XL Mem. Mat. Appl. (5)*, 24:219–235, 2000.

[16] Gh. Cenuşă and I. Săcuiu. Some properties of random functions almost periodic in probability. *Rev. Roumaine Math. Pures Appl.*, 25(9):1317–1325, 1980.

[17] G. Da Prato and C. Tudor. Periodic and almost periodic solutions for semilinear stochastic equations. *Stochastic Anal. Appl.*, 13(1):13–33, 1995.

[18] Toka Diagana. Weighted pseudo almost periodic functions and applications. *C. R. Math. Acad. Sci. Paris*, 343(10):643–646, 2006.

[19] Toka Diagana. *Pseudo almost periodic functions in Banach spaces.* Nova Science Publishers, Inc., New York, 2007.

[20] Toka Diagana. Weighted pseudo-almost periodic solutions to some differential equations. *Nonlinear Anal.*, 68(8):2250–2260, 2008.

[21] Toka Diagana. Existence of pseudo-almost automorphic solutions to some abstract differential equations with $\mathbf{S}^p$-pseudo-almost automorphic coefficients. *Nonlinear Anal.*, 70(11):3781–3790, 2009.

[22] Toka Diagana. Evolution equations in generalized Stepanov-like pseudo almost automorphic spaces. *Electron. J. Differential Equations*, pages No. 49, 19, 2012.

[23] Hui-Sheng Ding, Chao Deng, and Gaston M. N'Guérékata. Almost automorphic random functions in probability. *Abstr. Appl. Anal.*, pages Art. ID 243748, 6, 2014.

[24] Ryszard Engelking. *General topology*, volume 6 of *Sigma Series in Pure Mathematics*. Heldermann Verlag, Berlin, second edition, 1989. Translated from the Polish by the author.

[25] A. M. Fink. Almost automorphic and almost periodic solutions which minimize functionals. *Tôhoku Math. J. (2)*, 20:323–332, 1968.

[26] Miaomiao Fu. Almost automorphic solutions for nonautonomous stochastic differential equations. *J. Math. Anal. Appl.*, 393(1):231–238, 2012.

[27] Miaomiao Fu and Feng Chen. Almost automorphic solutions for some stochastic differential equations. *Nonlinear Anal.*, 80:66–75, 2013.

[28] Miaomiao Fu and Zhenxin Liu. Square-mean almost automorphic solutions for some stochastic differential equations. *Proc. Amer. Math. Soc.*, 138(10):3689–3701, 2010.

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

[30] A. Halanay, C. Tudor, and T. Morozan. Tracking almost periodic signals under white noise perturbations. *Stochastics*, 21(4):287–301, 1987.

[31] Mikhail Kamenski, Omar Mellah, and Paul Raynaud de Fitte. Weak Averaging of Semilinear Stochastic Differential Equations with Almost Periodic Coefficients. To appear in *J. Math. Anal. Appl.*, 2015.

[32] Jin Liang, Ti-Jun Xiao, and Jun Zhang. Decomposition of weighted pseudo-almost periodic functions. *Nonlinear Anal.*, 73(10):3456–3461, 2010.

[33] Jin Liang, Jun Zhang, and Ti-Jun Xiao. Composition of pseudo almost automorphic and asymptotically almost automorphic functions. *J. Math. Anal. Appl.*, 340(2):1493–1499, 2008.

[34] Georg Lindgren. Lectures on stationary stochastic processes. Technical report, Lund University, Lund, Sweden, 2006. A course for PHD students in Mathematical Statistics and other fields available at http://www.maths.lth.se/matstat/staff/georg/Publications/lecture2006.pdf.

[35] Zhenxin Liu and Kai Sun. Almost automorphic solutions for stochastic differential equations driven by Lévy noise. *J. Funct. Anal.*, 266(3):1115–1149, 2014.

[36] Carlos Lizama and Gaston M. N'Guérékata. Bounded mild solutions for semilinear integro differential equations in Banach spaces. *Integral Equations Operator Theory*, 68(2):207–227, 2010.

[37] Omar Mellah. *Étude des solutions presque périodiques des équations différentielles stochastiques*. PhD thesis, University Mouloud Mammeri, Tizi Ouzou, Algeria, and University of Rouen, France, 2012.

[38] Omar Mellah and Paul Raynaud de Fitte. Counterexamples to mean square almost periodicity of the solutions of some SDEs with almost periodic coefficients. *Electron. J. Differential Equations*, pages No. 91, 7, 2013.

[39] T. Morozan and C. Tudor. Almost periodic solutions of affine Itô equations. *Stochastic Anal. Appl.*, 7(4):451–474, 1989.

[40] Gaston M. N'Guerekata. *Almost automorphic and almost periodic functions in abstract spaces*. Kluwer Academic/Plenum Publishers, New York, 2001.

[41] Gaston M. N'guérékata. Comments on almost automorphic and almost periodic functions in Banach spaces. *Far East J. Math. Sci. (FJMS)*, 17(3):337–344, 2005.

[42] Gaston M. N'Guérékata. *Topics in almost automorphy*. Springer-Verlag, New York, 2005.

[43] Gaston M. N'Guérékata and Alexander Pankov. Stepanov-like almost automorphic functions and monotone evolution equations. *Nonlinear Anal.*, 68(9):2658–2667, 2008.

[44] Octav Onicescu and Vasile I. Istrătescu. Approximation theorems for random functions. *Rend. Mat. (6)*, 8:65–81, 1975. Collection of articles dedicated to Mauro Picone on the occasion of his ninetieth birthday.

[45] A. M. Precupanu. On the almost periodic functions in probability. *Rend. Mat. (7)*, 2(3):613–626, 1982.

[46] Wenxian Shen and Yingfei Yi. Almost automorphic and almost periodic dynamics in skew-product semiflows. *Mem. Amer. Math. Soc.*, 136(647):x+93, 1998.

[47] E. Slutsky. Sur les fonctions aléatoires presque périodiques et sur la décomposition des fonctions aléatoires stationnaires en composantes. Actual. sci. industr. 738, 33-55. (Confér. internat. Sci. math. Univ. Genève. Théorie des probabilités. V: Les fonctions aléatoires.) (1938)., 1938.

[48] C. Tudor. Almost periodic solutions of affine stochastic evolution equations. *Stochastics Stochastics Rep.*, 38(4):251–266, 1992.

[49] C. Tudor. Almost periodic stochastic processes. In *Qualitative problems for differential equations and control theory*, pages 289–300. World Sci. Publ., River Edge, NJ, 1995.

[50] C. A. Tudor and M. Tudor. Pseudo almost periodic solutions of some stochastic differential equations. *Math. Rep. (Bucur.)*, 1(51)(2):305–314, 1999.

[51] W. A. Veech. Almost automorphic functions on groups. *Amer. J. Math.*, 87:719–751, 1965.

[52] Ward Whitt. Weak convergence of probability measures on the function space $C[0, \infty)$. *Ann. Math. Statist.*, 41:939–944, 1970.

[53] Zhinan Xia and Meng Fan. Weighted Stepanov-like pseudo almost automorphy and applications. *Nonlinear Anal.*, 75(4):2378–2397, 2012.

[54] Chuan Yi Zhang. Integration of vector-valued pseudo-almost periodic functions. *Proc. Amer. Math. Soc.*, 121(1):167–174, 1994.

[55] Chuan Yi Zhang. Pseudo-almost-periodic solutions of some differential equations. *J. Math. Anal. Appl.*, 181(1):62–76, 1994.

[56] Chuan Yi Zhang. Pseudo almost periodic solutions of some differential equations. II. *J. Math. Anal. Appl.*, 192(2):543–561, 1995.

[57] Chuanyi Zhang. *Almost periodic type functions and ergodicity.* Science Press Beijing, Beijing; Kluwer Academic Publishers, Dordrecht, 2003.

[58] Rui Zhang, Yong-Kui Chang, and G. M. N'Guérékata. New composition theorems of Stepanov-like weighted pseudo almost automorphic functions and applications to nonautonomous evolution equations. *Nonlinear Anal. Real World Appl.*, 13(6):2866–2879, 2012.