Ergodic Properties for Multirate Linear Systems

Damian Marelli\textsuperscript{1} and Minyue Fu\textsuperscript{2} Fellow, IEEE

Abstract—Stochastic analysis of a multirate linear system typically requires the signals in the system to possess certain ergodic properties. Among them, ergodicity in the mean and ergodicity in the correlation are the most commonly used ones. We show that multirate operations and time-variant linear filtering can destroy these ergodic properties. Motivated by this fact, we introduce the notions of strong ergodicity in the mean and strong ergodicity in the correlation which are preserved under a number of operations, namely, downsampling, upsampling, addition and uniformly stable linear (time-variant) filtering. We provide conditions for random processes to be strongly ergodic. Using these conditions, we show that white random processes with uniformly bounded second moments are strongly ergodic in the mean and that mutually independent random processes with uniformly bounded fourth moments are jointly strongly ergodic in the correlation. The main implication of these results is that if a multirate linear system is driven by independent random processes with uniformly bounded second (or fourth) moments, then every signal in the system is strongly ergodic in the mean (or correlation) and therefore ergodic in the mean (or correlation). An application of these results is also discussed.

Index Terms—Stochastic processes, multirate systems, signal processing, digital filters, multirate digital filters.

I. INTRODUCTION

Multirate signal processing techniques find a wide range of applications; see, e.g., [1], [2], [3]. Downsampling and upsampling are two basic multirate operations. Using these operations and filterbanks, sophisticated signal processing can be carried out in subbands. An example of such applications is the so-called subband adaptive filtering [4], [5], where filtering is done in individual subbands to gain a number of numerical advantages. Another example is the so-called subband system identification; see, e.g., [6], [7]. In this technique, the input and output signals of a system to be identified are split into subbands by using filterbanks. A parametric model is then tuned in every subband. By doing so, better convergence and faster computation can be achieved in many cases.

In order to understand the statistical behavior of multirate systems, stochastic analysis is essential. To this end, the notion of ergodicity plays a fundamental role. More specifically, random signals in the system are required to be ergodic in the mean and/or correlation. When dealing with subband decompositions, it is desirable that various stochastic analysis results can be carried over to subbands. This implies that random signals in each subband are required to be ergodic, which is a property taken for granted in most multirate stochastic analysis. This raises the following fundamental question: What conditions are required on the fullband signals so that the subband signals are ergodic? We will show that the ergodicity of a random signal may not be preserved under a number of operations, including downsampling and time-variant (uniformly stable) linear filtering.

Motivated by the discussion above, we look for suitable notions of ergodicity for multirate linear systems. More specifically, we introduce the notions of strong ergodicity in the mean and strong ergodicity in the correlation, which are properties preserved under a number of commonly used linear multirate operations, namely downsampling, upsampling, addition and filtering by uniformly stable linear filters. We establish conditions for random processes to be strongly ergodic. Using these conditions, we show that white random processes with uniformly bounded second moments are strongly ergodic in the mean and that mutually independent random processes with uniformly bounded fourth moments are jointly strongly ergodic in the correlation.

These results mean that most commonly used random processes in multirate linear systems are strongly ergodic (in the mean and in the correlation). Hence, every signal in the system is strongly ergodic and therefore ergodic. In order to illustrate our results, we introduce an application. We refer to [6] and discuss the use of strong ergodicity in the analysis of some convergence properties in a subband identification system.

Although the results in this paper are simply stated and their importance in stochastic analysis of multirate systems is quite natural, their derivations turn out to be much more involved. Readers familiar with ergodic theory, which is a major field of mathematics, may wonder whether it is possible to derive our results from known results on ergodicity. We justify below that this is not the case.

To prove that a random process is ergodic is essentially to verify that the Strong Law of Large Numbers (SLLN) [8] is satisfied. There are two approaches to verify this law. One uses Rajchman’s SLLN [8, Th 5.1.2, p103] and the other uses Kolmogorov’s SLLN [8, Th 5.4.2, p126]. Both require the random process to be either uncorrelated or independent. The main difference is that Kolmogorov’s SLLN requires the random process to be strictly stationary while Rajchman’s SLLN does not. A generalization of Kolmogorov’s SLLN is Birkhoff’s ergodic theorem, which is the cornerstone of ergodic theory [9]. In this generalization, the random process is no longer required to be independent. The main advantage of this approach is that it provides conditions for ergodicity in a very general sense (including the mean and correlation). Also, these conditions are preserved if the random process is transformed via any nonlinear, measurable and stationary

\textsuperscript{1}Affiliation of author: Damian Marelli is with the Laboratoire d’Analyse, Topologie et Probabilités, Centre de Mathématiques et d’Informatique, Université de Provence, 13453 Marseille, France. E-mail: marelli@cmi.univ-mrs.fr.

\textsuperscript{2}Affiliation of author: Minyue Fu is with the School of Electrical Engineering and Computer Science, University of Newcastle, N.S.W. 2308 Australia. E-mail: eemf@ecemail.newcastle.edu.au.
transformation. However, this approach is not suitable for our purposes because it applies only to stationary random processes, and thus excludes time-variant operations such as upsampling and time-variant filtering. Another generalization of this approach is given in [10] where the requirement for strict stationarity is relaxed to the so-called asymptotic mean stationarity (AMS). It turns out that this generalization is not sufficient for our purposes because, as we will show, the AMS condition can be destroyed by time-variant operations. Therefore, we turn to Rajchman’s SLLN. This approach has been used in [11] to show that if a time-invariant linear filter is uniformly stable, then a filtered independent random process is ergodic in the correlation. However, this result does not generalize to other multirate operations. Therefore, it is necessary to introduce stronger notions of ergodicity, as we have mentioned earlier. A lot of care has been given in defining the notions of strong ergodicity so that commonly used random processes in multirate systems can be verified to be strongly ergodic.

The rest of this paper is organized as follows. Section II introduces some necessary background on random processes, including the definitions of ergodicity in the mean and in the correlation. Section III shows, through using examples, that ergodicity in the correlation can be destroyed by downsampling, time-variant linear filtering, and addition. Section IV introduces a technical concept needed to study strong ergodicity. Section V studies strong ergodicity in the mean and Section VI does the same for strong ergodicity in the correlation. Section VII discusses the application mentioned above. Section VIII shows that the AMS condition can be destroyed by time-variant operations. We conclude the paper in Section IX. For ease of readability, all proofs are contained in the Appendix.

II. PRELIMINARIES

In this section, we introduce the necessary notation and definitions for this paper.

A. Random Processes

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space defined on a set $\Omega$ with $\sigma$-algebra $\mathcal{A}$ and probability measure $\mathbb{P}$. A random variable is an $\mathcal{A}$-measurable map $\nu : \Omega \to \mathbb{C}$, where $\mathbb{C}$ denotes the set of complex numbers. We denote the set of all random variables so defined by

$$\mathcal{L}(\Omega, \mathcal{A}, \mathbb{P}) = \{v : \Omega \to \mathbb{C} : v \text{ is } \mathcal{A}\text{-measurable}\}$$

The $p$-th (absolute) moment (or $p$-th norm) of a random variable $v$ is defined by

$$\|v\|_p = \mathcal{E}\{|v|^p\}^{1/p}$$

where $\mathcal{E}\{\cdot\}$ denotes the expected value. We denote

$$\mathcal{L}_p(\Omega, \mathcal{A}, \mathbb{P}) = \{v \in \mathcal{L}(\Omega, \mathcal{A}, \mathbb{P}) : \|v\|_p < \infty\}$$

If $u, v \in \mathcal{L}_2(\Omega, \mathcal{A}, \mathbb{P})$, then their inner product is defined as

$$\langle u, v \rangle = \mathcal{E}\{\bar{u}v\}$$

Two random variables $x, y \in \mathcal{L}(\Omega, \mathcal{A}, \mathbb{P})$ are said to be uncorrelated if $\mathcal{E}\{xy\} = \mathcal{E}\{x\}\mathcal{E}\{y\}$. A set of random variables $\{x_i : i = 1, 2, \cdots, n\}$ is said to be independent if, for any $B_i \in \mathcal{B}$, $i = 1, 2, \cdots, n$, where $\mathcal{B}$ denotes the Borel algebra of complex numbers,

$$\mathbb{P}\left\{\bigcap_{i=1}^n \{\omega \in \Omega : x_i(\omega) \in B_i\}\right\} = \prod_{i=1}^n \mathbb{P}\{\omega \in \Omega : x_i(\omega) \in B_i\}$$

A (discrete-time) random process is a map $x : \mathbb{Z} \to \mathcal{L}(\Omega, \mathcal{A}, \mathbb{P})$, where $\mathbb{Z}$ is the set of integers. In order to simplify the notation, we write

$$x(t, \omega) = (x(t))(\omega), \quad t \in \mathbb{Z}, \omega \in \Omega$$

A random process $x$ is said to have uniformly bounded $p$-th (absolute) moments with bound $M_x$, if

$$\|x(t)\|_p \leq M_x, \ \forall t \in \mathbb{Z}$$

We denote

$$\mathcal{P}(\Omega, \mathcal{A}, \mathbb{P}) = \{x : \mathbb{Z} \to \mathcal{L}(\Omega, \mathcal{A}, \mathbb{P})\}$$

and

$$\mathcal{P}_p(\Omega, \mathcal{A}, \mathbb{P}) = \left\{x \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P}) : \sup_{t \in \mathbb{Z}} \|x(t)\|_p < \infty\right\}$$

i.e., $\mathcal{P}_p(\Omega, \mathcal{A}, \mathbb{P})$ is a set of random processes in $\mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$ with uniformly bounded $p$-th moments. Since a deterministic signal can also be interpreted as a random process, the set $\mathcal{P}_p(\Omega, \mathcal{A}, \mathbb{P})$ includes deterministic signals in $l_\infty(\mathbb{Z})$.

A random process is said to be white if its samples form a set of pairwise uncorrelated random variables, and it is said to be independent if its samples form a set of independent random variables. Also, a set of random processes is said to be mutually independent if their samples form a set of independent random variables.

B. Multirate Operations and Filtering

Let $x \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$ be a random process and $D \in \mathbb{N}$, where $\mathbb{N}$ is the set of positive integer numbers. The downsampled random process with downsampling rate $D$ is given by

$$y(t) = \Delta_D x(t) = x(Dt)$$

Similarly, the upsampled random process with upsampling rate $D$ is given by

$$y(t) = \Upsilon_D x(t) = \left\{\begin{array}{ll} x(t/D), & t/D \in \mathbb{Z} \\ 0, & t/D \notin \mathbb{Z} \end{array}\right.$$

A time-variant linear filter with impulse response $\{h_t \in l_1(\mathbb{Z}) : t \in \mathbb{Z}\}$ is said to be uniformly stable if there exists $h \in l_1(\mathbb{Z})$ such that $|h_t(\tau)| \leq h(\tau), \forall t, \tau \in \mathbb{Z}$. 
C. Ergodicity

A random process \( x \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P}) \) is said to be \textit{ergodic in the mean} if

\[
\lim_{T \to -\infty} \frac{1}{T} \sum_{t=1}^{T} [x(t) - \mathbb{E}\{x(t)\}] \xrightarrow{w.p.1} 0
\]

(1)

Similarly, two random processes \( x, y \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P}) \) are said to be \textit{jointly ergodic in the correlation} if, for every \( \tau \in \mathbb{Z} \),

\[
\lim_{T \to -\infty} \frac{1}{T} \sum_{t=1}^{T} [\bar{x}(t)y(t + \tau) - \mathbb{E}\{\bar{x}(t)y(t + \tau)\}] \xrightarrow{w.p.1} 0
\]

(2)

Also, a random process is \textit{ergodic in the correlation} if it is jointly ergodic in the correlation with itself.

III. Motivating Examples

In this section we point out that ergodicity in the correlation can be lost after some transformations involved in a multirate linear system. We do this by introducing examples.

The following example shows that ergodicity in the correlation can be lost after downsampling.

\textbf{Example 1:} Consider the probability space \( ([\frac{1}{2}, \frac{3}{2}], \mathcal{B}, \lambda) \), where \( \mathcal{B} \) denotes the Borel \( \sigma \)-algebra on the set \([\frac{1}{2}, \frac{3}{2}]\) and \( \lambda \) denotes the Lebesgue measure. Define the random process \( x \in \mathcal{P}([\frac{1}{2}, \frac{3}{2}], \mathcal{B}, \lambda) \) as follows:

\[
x(t, \omega) = \begin{cases} 
\frac{1}{2}(1 + \text{sign}(\omega)), & t \text{ is even} \\
\frac{1}{2}(1 - \text{sign}(\omega)), & t \text{ is odd} 
\end{cases}
\]

where \( \text{sign}(\omega) = |\omega|/\omega \). Then,

\[
\frac{1}{T} \sum_{t=1}^{T} [\bar{x}(t, \omega)x(t + \tau, \omega) - \mathbb{E}\{\bar{x}(t, \omega)x(t + \tau, \omega)\}] \\
= \begin{cases} 
\frac{1}{T} \sum_{t=1}^{T} \frac{1}{2}(-1)^{\tau}\text{sign}(\omega), & \tau \text{ is even} \\
0, & \tau \text{ is odd}
\end{cases}
\rightarrow 0, \quad T \to \infty
\]

Therefore, \( x \) is ergodic in the correlation. Let \( z \) be obtained by downsampling \( x \) with a factor of 2, i.e.,

\[
z(t, \omega) = x(2t, \omega) = \frac{1}{2}(1 + \text{sign}(\omega))
\]

Then,

\[
\frac{1}{T} \sum_{t=1}^{T} [\bar{z}(t, \omega)z(t+\tau, \omega) - \mathbb{E}\{\bar{z}(t, \omega)z(t+\tau, \omega)\}] = \frac{1}{2}\text{sign}(\omega) \rightarrow 0
\]

(3)

Hence, \( z \) is not ergodic in the correlation.

The next example shows that ergodicity in the correlation can be lost after uniformly stable linear filtering.

\textbf{Example 2:} Consider the random process \( x \) in Example 1 and the uniformly stable time-variant linear system with impulse response

\[
h_{\text{st}}(\tau) = \begin{cases} 
\delta(\tau), & t \text{ is even} \\
\delta(\tau - 1), & t \text{ is odd}
\end{cases}
\]

where \( \delta(\tau) = \begin{cases} 
1, & \tau = 0 \\
0, & \tau \neq 0
\end{cases} \)

Let \( z \) be obtained by filtering \( x \) through \( h_{\text{st}}(\tau) \), i.e.,

\[
z(t, \omega) = \sum_{\tau=-\infty}^{\infty} h_{\text{st}}(\tau)x(t - \tau, \omega) = \frac{1}{2} (1 + \text{sign}(\omega))
\]

Then (3) holds, and therefore, \( z \) is not ergodic in the correlation.

The following example shows that ergodicity in the correlation is not closed under addition, i.e., it can be lost by adding two random processes which are ergodic in the correlation.

\textbf{Example 3:} Consider the probability space \( ([\frac{1}{2}, \frac{3}{2}], \mathcal{B}, \lambda) \). Define the random processes \( x, y \in \mathcal{P}([\frac{1}{2}, \frac{3}{2}], \mathcal{B}, \lambda) \) as follows:

\[
x(t, \omega) = \frac{1}{2}\text{sign}(\omega)
\]

\[
y(t, \omega) = \frac{1}{2}
\]

It is easy to verify that, for all \( \tau \in \mathbb{Z} \), \( T \in \mathbb{N} \),

\[
\frac{1}{T} \sum_{t=1}^{T} [\bar{x}(t, \omega)x(t + \tau, \omega) - \mathbb{E}\{\bar{x}(t, \omega)x(t + \tau, \omega)\}] = 0
\]

\[
\frac{1}{T} \sum_{t=1}^{T} [\bar{y}(t, \omega)y(t + \tau, \omega) - \mathbb{E}\{\bar{y}(t, \omega)y(t + \tau, \omega)\}] = 0
\]

Therefore, \( x \) and \( y \) are ergodic in the correlation. We define a new random process \( z \) by

\[
z(t, \omega) = x(t, \omega) + y(t, \omega) = \frac{1}{2} (1 + \text{sign}(\omega))
\]

Then (3) holds, and therefore, \( z \) is not ergodic in the correlation.

IV. Random Processes with Weakly Bounded Auto-Correlation

In this section, we introduce a class of random processes with auto-correlation obeying a certain bound. This property is instrumental in the analysis of strong ergodicity.

\textbf{Definition 1:} Let \( x \in \mathcal{P}_2(\Omega, \mathcal{A}, \mathbb{P}) \) be a random process. We define

\[
\|x\|_{S} = \sup_{T \in \mathbb{N}, d \in \mathbb{Z}} \left( \frac{1}{T} \sum_{t,s=1}^{T} |\langle x(s + d), x(t + d) \rangle|^{1/4} \right)
\]

(4)

We say \( x \) has \( l_2 \)-norm bounded auto-correlation (or simply \textit{bounded auto-correlation} (BAC)) if there exists \( f \in l_2(\mathbb{Z}) \) such that\( \rightarrow \infty \)

\[
|\langle x(t + \tau), x(t) \rangle| \leq f(\tau), \quad \forall t \in \mathbb{Z}
\]

We say \( x \) has \textit{weakly bounded auto-correlation} (WBAC) if \( \|x\|_{S} < \infty \). The space of random processes with WBAC will be denoted by

\[
\mathcal{S}(\Omega, \mathcal{A}, \mathbb{P}) = \{ x \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P}) : \|x\|_{S} < \infty \}
\]

Further, a countable set of random processes \( X \subset \mathcal{P}_2(\Omega, \mathcal{A}, \mathbb{P}) \) is said to have \textit{uniformly WBAC} (UWBAC) if there exists a constant \( U_X > 0 \) such that

\[
\|x\|_{S} \leq U_X, \quad \forall x \in X
\]

(5)

In this case, \( U_X \) is called a \textit{uniform weak bound} of \( X \).
Note that the definition of WBAC implies that BAC is a stronger condition than WBAC. Also, the notation for the map \( \| \cdot \|_S \) implies that it is a norm operator. These two facts are stated in the two propositions below.

**Proposition 1:** If a random process \( x \in \mathcal{P}_2(\Omega, A, \mathbb{P}) \) has BAC, it has WBAC.

**Proposition 2:** The map \( \| \cdot \|_S : S(\Omega, A, \mathbb{P}) \rightarrow \mathbb{R}^+ \) defines a norm on \( S \), i.e., for all \( x, y \in S(\Omega, A, \mathbb{P}) \) and \( c \in \mathbb{C} \), the following conditions hold:

1. \( \| cx \|_S = |c| \| x \|_S \);
2. \( \| x + y \|_S \leq \| x \|_S + \| y \|_S \);
3. \( \| x \|_S = 0 \Rightarrow x = 0 \).

Random processes with WBAC enjoy the following important properties.

**Proposition 3:** If a countable set of random processes \( X \subset \mathcal{P}_2(\Omega, A, \mathbb{P}) \) has UWBAE, then, for all \( x \in X \),

\[
\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} x(t) \overset{w.p.}{=} 0
\]

**Proposition 4:** Consider two countable sets of random processes \( X, Y \subset \mathcal{P}_2(\Omega, A, \mathbb{P}) \) with UWBAE. Denoting the elements of \( X \) and \( Y \) by \( x_i \) and \( y_i \), respectively, with \( i \in \mathbb{Z} \). Then the countable set of random processes

\[
Z = \{ z_i = x_i + y_i : i \in \mathbb{Z} \}
\]

also has UWBAE.

**Proposition 5:** Let \( X \subset \mathcal{P}_2(\Omega, A, \mathbb{P}) \) be a countable set of random processes with UWBAE and \( \{ h_t \in l_1(\mathbb{Z}) : t \in \mathbb{Z} \} \) be the impulse response for a uniformly stable (time-variant) linear filter. Define a filtered set of random processes \( Y \) with elements \( y \) given by

\[
y(t) = \sum_{k=-\infty}^{\infty} h_t(k)x_k(t-k), \quad x_k \in X
\]

(Note that a set of random processes \( \{ x_k \} \) may be used to generate each \( y \).) Then, \( Y \) is a countable set of random processes with UWBAE.

**Proposition 6:** If \( X \subset \mathcal{P}_2(\Omega, A, \mathbb{P}) \) is a countable set of random processes with UWBAE and \( D \in \mathbb{N} \), then the countable set of random processes \( Y \) generated from \( X \) by downsampling or upsampling with rate \( D \) also has UWBAE.

V. STRONG ERGODICITY IN THE MEAN

In this section, we introduce the notion of strong ergodicity in the mean and study conditions under which a random process is strongly ergodic in the mean.

**Definition 2:** A random process \( x \in \mathcal{P}(\Omega, A, \mathbb{P}) \) is said to be strongly ergodic in the mean if the following conditions hold:

(SEM1) \( x \) is ergodic in the mean;

(SEM2) If the random process \( y \in \mathcal{P}(\Omega, A, \mathbb{P}) \) is also strongly ergodic in the mean, then \( x + y \) is strongly ergodic in the mean;

(SEM3) The filtering of \( x \) by a uniformly stable linear filter yields a random process that is strongly ergodic in the mean;

(SEM4) The downsampling of \( x \) by any factor \( D \in \mathbb{N} \) yields a random process that is strongly ergodic in the mean;

(SEM5) The upsampling of \( x \) by any factor \( D \in \mathbb{N} \) yields a random process that is strongly ergodic in the mean.

Next, we characterize random processes which are strongly ergodic in the mean. We first provide a sufficient condition for strong ergodicity in the mean (Proposition 7). This result is then used to show that all uncorrelated random processes are strongly ergodic in the mean (Proposition 8).

**Proposition 7:** Let \( x \in \mathcal{P}_2(\Omega, A, \mathbb{P}) \) be a random process. If the random process \( \xi_x \) defined by

\[
\xi_x(t) = x(t) - E\{x(t)\}
\]

has WBAC, then \( x \) is strongly ergodic in the mean.

**Proposition 8:** If a random process \( x \in \mathcal{P}(\Omega, A, \mathbb{P}) \) is white and has uniformly bounded second moments, then it is strongly ergodic in the mean.

Proposition 8 and Definition 2 together imply the following theorem, which we state without proof.

**Theorem 1:** Given a finite set of random processes \( u_m \in \mathcal{P}(\Omega, A, \mathbb{P}) \), \( m = 1, \ldots, M \), suppose each of them is white and has uniformly bounded second moments. Let \( v \) be a random process formed from \( \{ u_m \} \) by any finite combinations of additions, downsampling, upsampling and filtering by a uniformly stable linear filter. Then \( v \) is ergodic in the mean.

**Remark 1:** Note that Theorem 1 includes the possibility that some of the random processes \( u_m \) are deterministic signals in \( l_\infty(\mathbb{Z}) \) (see Section II-A).

VI. STRONG ERGODICITY IN THE CORRELATION

The aim of this section is to give a presentation of strong ergodicity in the correlation, in a way similar to that of Section V. However, this cannot be done straightforwardly. In Example 4 below, we show that the addition of two independent (and therefore uncorrelated) signals can yield a non-ergodic signal. In contrast, condition (SEM2) says that the addition of two signals which are strongly ergodic in the mean yields another signal which is also strongly ergodic in the mean. In view of Example 4, an analogous condition can not be imposed in the definition of strong ergodicity in the correlation, because otherwise it would prevent independent signals from being strongly ergodic in the correlation, which in turn means that this notion of ergodicity would be inappropriate.

**Example 4:** Consider the probability space \((0, 1), B, \lambda\). For every \( \omega \in (0, 1) \), consider its binary expansion

\[
\omega = \sum_{t=1}^{\infty} \omega_t 2^{-t}
\]

Define the random process \( x(\omega) \) by

\[
x(t, \omega) = 2\omega_t - 1
\]

Define a new random process \( y(\omega) \) by

\[
y(t, \omega) = \begin{cases} x(t, \omega), & \omega \in [0, \frac{1}{2}) \\ -x(t, \omega), & \omega \in \left[\frac{1}{2}, 1\right) \end{cases}
\]
By construction, $x$ and $y$ are independent signals ([8, Example 4, p.56]) taking values in $\{-1, 1\}$, although they are not mutually independent. Define a new random process $z$ by

$$z(t, \omega) = x(t, \omega) + y(t, \omega) = \begin{cases} 2x(t, \omega), & \omega \in [0, \frac{1}{2}) \\ 0, & \omega \in [\frac{1}{2}, 1) \end{cases}$$

Then, for all $t \in \mathbb{Z}$,

$$\bar{z}(t, \omega)z(t, \omega) = \begin{cases} 4, & \omega \in [0, \frac{1}{2}) \\ 0, & \omega \in [\frac{1}{2}, 1) \end{cases}$$

It follows that

$$\frac{1}{T} \sum_{t=1}^{T} |\bar{z}(t, \omega)z(t, \omega) - \mathbb{E}\{\bar{z}(t, \omega)z(t, \omega)\}| = \begin{cases} 2, & \omega \in [0, \frac{1}{2}) \\ -2, & \omega \in [\frac{1}{2}, 1) \end{cases}$$

Therefore, $z$ is not ergodic in the correlation.

The inconvenience above can be dealt with by adding an extra requirement, namely that the signals to be added should possess certain joint ergodicity conditions. Then, we define strong ergodicity in the correlation as a condition of two signals.

**Definition 3:** Given two random processes $x, y \in \mathcal{P}(\Omega, A, \mathbb{P})$, $x$ is said to be jointly strongly ergodic in the correlation with $y$ if the following conditions hold:

- (SEC0) $x$ and $y$ are jointly ergodic in the correlation;
- (SEC1) $y$ is jointly strongly ergodic in the correlation with $x$, i.e., the condition is symmetric;
- (SEC2) If the random process $z \in \mathcal{P}(\Omega, A, \mathbb{P})$ is jointly strongly ergodic in the correlation with both $x$ and $y$, then $z$ is also jointly strongly ergodic in the correlation with $x + y$;
- (SEC3) The filtering of $y$ by a uniformly stable linear filter yields a random process that is jointly strongly ergodic in correlation with $x$;
- (SEC4) The downsampling of $y$ by any factor yields a random process that is jointly strongly ergodic in the correlation with $x$;
- (SEC5) The upsampling of $y$ by any factor yields a random process that is jointly strongly ergodic in the correlation with $x$.

A random process is called strongly ergodic in the correlation if it is jointly strongly ergodic in the correlation with itself.

As in the previous section, we provide two results, one giving sufficient conditions for strong ergodicity, and another showing the relationship between independency and strong ergodicity.

**Proposition 9:** Let $x, y \in \mathcal{P}_2(\Omega, A, \mathbb{P})$ be random processes. Define a countable set of random processes $\Psi_{(x,y)} \subset \mathcal{P}(\Omega, A, \mathbb{P})$ by

$$\Psi_{(x,y)} = \{\psi_{(x,y,\alpha,\beta,a,b)} : \alpha, \beta \in \mathbb{N}; a, b \in \mathbb{Z}\}$$

where

$$\psi_{(x,y,\alpha,\beta,a,b)}(t) = x(\alpha t + a)y(\beta t + b) - \mathbb{E}\{x(\alpha t + a)y(\beta t + b)\}$$

Suppose $\Psi_{(x,y)}$ has UWBAC, then $x$ and $y$ are jointly strongly ergodic in the correlation.

**Proposition 10:** Let $x, y \in \mathcal{P}_2(\Omega, A, \mathbb{P})$ be random processes with uniformly bounded fourth moments. We have the following results:

(i) If $x$ is independent, then it is strongly ergodic in the correlation;
(ii) If $x$ and $y$ are mutually independent, then $x$ and $y$ are jointly strongly ergodic in the correlation.

**Theorem 2:** Given a finite set of random processes $u_m \in \mathcal{P}(\Omega, A, \mathbb{P})$, $m = 1, \ldots, M$, suppose each of them has uniformly bounded fourth moments and the set $\{u_m\}$ is mutually independent. Let $v$ be a random process formed from $\{u_m\}$ by any finite combinations of additions, downsampling, upsampling and filtering by a uniformly stable linear filter. Then $v$ is ergodic in the correlation.

**Remark 2:** Remark 1 also applies to strong ergodicity in the correlation.

**Remark 3:** Note that if a random process is Gaussian, then the uniform boundedness of the fourth moments is equivalent to that of the second moments.

Recall in Example 1 where we show that ergodicity in the correlation can be destroyed by downsampling and linear filtering. By Definition 3, this random process $x$ cannot be strongly ergodic in the correlation, which in turn implies that $\Psi_{(x,x)}$ as defined in (7) does not have UWBAC. This is verified below.

**Example 5:** Consider the random process $x$ from Example 1. Let $\psi_{(x,x,1,1,0,0)}$ be defined as in (8). Then,

$$\psi_{(x,x,1,1,0,0)}(t) = x(t)x(t) - \mathbb{E}\{x(t)x(t)\}$$

It follows that

$$|\langle \psi_{(x,x,1,1,0,0)}(t), \psi_{(x,x,1,1,0,0)}(s) \rangle| = \frac{1}{4}, \forall t, s \in \mathbb{Z}$$

Consequently, for any $T \in \mathbb{N}$ and $d \in \mathbb{Z}$,

$$\frac{1}{T} \sum_{t,s=1}^{T} |\langle \psi_{(x,x,1,1,0,0)}(t + d), \psi_{(x,x,1,1,0,0)}(s + d) \rangle|^2 = \frac{T}{8}$$

This means that $\|\psi_{(x,x,1,1,0,0)}\|_S$ is not bounded. Therefore, $\Psi_{(x,x)}$ does not have UWBAC.

**VII. An Application of Theorem 2**

An applications of Theorem 2 is discussed in this section, namely, the analysis of strong and optimal convergence in a subband identification system.

As mentioned in Introduction, subband identification is a system identification method that uses multirate signal processing techniques. The scheme is depicted in Figure 1. The idea is to split both signals $u$ and $y$ (called fullband signals) into $M$ subbands using two identical analysis filterbanks.

$$h(q) = [h_1(q), \ldots, h_M(q)]^T$$

Then, these subband signals are down-sampled, by a factor of $D$. The results are denoted by two vector signals

$$U(t) = [U_1(t), \ldots, U_M(t)]^T$$
and

\[ Y(t) = [Y_1(t), \cdots, Y_M(t)]^T \]

respectively. The subband parametric model is given by

\[ \hat{G}(q, \theta) = \text{diag}\{\hat{G}_m(q, \theta_m), m = 1, \ldots, M\} \]

where \( \hat{G}_m(q, \theta_m) \) are FIR models of tap size \( \hat{n} \), parameterized by a vector \( \theta_m \). Its output is denoted by

\[ \hat{W}(t, \theta) = [\hat{W}_1(t, \theta_1), \cdots, \hat{W}_M(t, \theta_M)]^T \]

and the subband prediction error is defined by

\[ \hat{V}(t, \theta) = [\hat{V}_1(t, \theta_1), \cdots, \hat{V}_M(t, \theta_M)]^T = Y(t) - \hat{W}(t, \theta) \]

Each subband model \( \hat{G}_m(q, \theta_m) \) is tuned to minimize the power of \( \hat{V}_m(t, \theta_m) \). An up-sampler and a synthesis filterbank

\[ f(q) = [f_1(q), \cdots, f_M(q)]^T \]

are used to reconstruct the fullband prediction error \( \hat{v}(t, \theta) \).

![Subband identification scheme.](image)

The subband identification scheme above has been investigated in details in \[6\], where technical conditions for strong and optimal convergence of \( \theta_m \) in each subband are provided. By strong and optimal convergence, it means that

\[ \lim_{N \to \infty} S_{\hat{W}_m}(\theta_{m,N}) \overset{w,p.1}{\rightarrow} \min \theta_m S_{\hat{W}_m}(\theta_m) \]  \hspace{1cm} (9)

In the above, \( \theta_{m,N} \) denotes the optimal parameter vector \( \theta_m \) up to time \( N \) in the \( m \)-th subband, calculated based on a given optimization criterion. The term \( S_{\hat{W}_m}(\theta_m) \) denotes the power of the signal \( \hat{W}_m(t, \theta_m) \), defined by

\[ S_{\hat{W}_m}(\theta_m) := \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\{|\hat{W}_m(t, \theta_m)|^2\} \]

One of the technical conditions required to satisfy (9) is that the signals \( U_m(t) \) and \( Y_m(t) \) are individually ergodic and jointly ergodic in the correlation. In view of Theorem 2, this condition is guaranteed if the input signals \( u(t) \) and \( v(t) \) are assumed to be generated from a set of mutually independent random processes (which includes the possibility of deterministic signals) by any combinations of uniformly stable linear filtering, downsampling, upsampling and addition.

VIII. A NOTE ON ASYMPTOTICALLY MEAN STATIONARY (AMS) RANDOM PROCESSES

In the Introduction, we pointed out that Birkhoff’s Ergodic Theorem was generalized to the so-called AMS random processes \[10\]. But we claimed that the AMS condition can be destroyed by time-variant operations. We justify this claim in this section.

We start by introducing some notation and the definition of AMS. Let \( \mathbb{C}^2 \) denote the space of two-sided sequences of complex numbers and let \( \mathbb{B}^2 \) denote the Borel \( \sigma \)-algebra on \( \mathbb{C}^2 \). Given a random process \( x \in \mathcal{P}(\Omega, A, \mathbb{P}) \), we call any \( E \in \mathbb{B}^2 \) an event and denote its probability (or measure) by \( \mathbb{P}_x(E) \), i.e.,

\[ \mathbb{P}_x(E) = \mathbb{P}\{\omega \in \Omega : \{x(t, \omega) : t \in \mathbb{Z}\} \in E\} \]

Let \( q \) denote the forward shift operator, i.e. \( q(x(t)) = x(t+1) \).

**Definition 4:** \[10\] We say that a random process \( x \in \mathcal{P}(\Omega, A, \mathbb{P}) \) is asymptotically mean stationary if

\[ \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{P}_x(q^{-t}E) \text{ exists for every event } E \in \mathbb{B}^2 \]  \hspace{1cm} (10)

The following example shows that the AMS condition may not be preserved after a combination of uniformly stable (time-variant) linear filtering and downsampling.

**Example 6:** Consider the probability space \( ([0,1], \mathcal{B}, \lambda) \), where \( \mathcal{B} \) denotes the Borel \( \sigma \)-algebra on the set \([0,1] \) and \( \lambda \) denotes the Lebesgue measure. Define the random process \( x \in \mathcal{P}([0,1], \mathcal{B}, \mathbb{P}) \) as follows:

\[ x(t, \omega) = \begin{cases} 1, & t \text{ is even} \\ 0, & t \text{ is odd} \end{cases} \]

which is a deterministic function. Since \( x \) is an alternating sequence, the limit (10) exists for any event \( E \in \mathbb{B}^2 \). More precisely,

\[ \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{P}_x(q^{-t}E) = \frac{1}{2} [\mathbb{P}_x(E) + \mathbb{P}_x(qE)] \]  \hspace{1cm} (11)

Now we form a new random process \( y \) (also deterministic) as follows: For each value of \( t \in \mathbb{Z} \), we consider the pair \( x(2t) \) and \( x(2t-1) \). We can either take these two values as they are or swap them (i.e., set \( y(2t) = x(2t) \) and \( y(2t-1) = x(2t-1) \)) or swap them (i.e., set \( y(2t) = x(2t-1) \) and \( y(2t-1) = x(2t) \)). The decision for swapping or not is made such that, for all \( \omega \in \Omega \),

\[ y(2t, \omega) = \begin{cases} 1, & 2^{2n} < t \leq 2^{2n+2} \text{ for some } n \in \mathbb{Z} \\ 0, & 2^{2n+2} < t \leq 2^{2n} \text{ for some } n \in \mathbb{Z} \end{cases} \]  \hspace{1cm} (12)

It is easy to see that \( y \) can be generated by filtering \( x \) with the following uniformly stable linear filter

\[ h_t(\tau) = \begin{cases} \delta(\tau), & t \in \mathcal{I} \\ \delta(\tau-1), & \tau \in J_0 \\ \delta(\tau+1), & \tau \in J_1 \end{cases} \]

where \( \mathcal{I} \) denotes the set of values of \( t \in \mathbb{Z} \) at which \( y(t) \) is generated without swapping, and \( J_0 \) (respectively, \( J_1 \)) denotes the set of \( t \in \mathbb{Z} \) at which \( y(t) \) is generated by forward (respectively, backward) shifting. Now let \( z(t) = \ldots, 2^{2n+2} < t \leq 2^{2n} \text{ for some } n \in \mathbb{Z} \end{cases} \]  \hspace{1cm} (12)

It is easy to see that \( y \) can be generated by filtering \( x \) with the following uniformly stable linear filter

\[ h_t(\tau) = \begin{cases} \delta(\tau), & t \in \mathcal{I} \\ \delta(\tau-1), & \tau \in J_0 \\ \delta(\tau+1), & \tau \in J_1 \end{cases} \]

where \( \mathcal{I} \) denotes the set of values of \( t \in \mathbb{Z} \) at which \( y(t) \) is generated without swapping, and \( J_0 \) (respectively, \( J_1 \)) denotes the set of \( t \in \mathbb{Z} \) at which \( y(t) \) is generated by forward (respectively, backward) shifting. Now let \( z(t) =
y(2t) (obtained by downsampling y with D = 2) and consider
the event \( F = \{ c_k : k \in \mathbb{Z} \} : c_0 = 1 \) \( \in \mathcal{B}^2 \). Using (12) and
\[
q^{-t}F = \{ c_k : k \in \mathbb{Z} : c_{-t} = 1 \}
\]
we get
\[
\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} P_z(q^{-t}F) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \left\{ \begin{array}{ll}
1, & 2^{4n} < t \leq 2^{4n+2}, \text{for some } n \in \mathbb{Z} \\
0, & 2^{4n+2} < t \leq 2^{4n+2}, \text{for some } n \in \mathbb{Z}
\end{array} \right.
\]
It is straightforward to verify that the limit above does not exist. This implies that the AMS property of \( x \) is lost after uniformly stable linear filtering followed by downsampling.

**IX. CONCLUSION**

In this paper we have introduced two new notions of ergodicity, namely strong ergodicity in the mean and strong ergodicity in the correlation. These notions are used to provide an adequate theoretical framework for the stochastic analysis of multirate linear systems, and are motivated by the fact that ergodicity in the mean (or correlation) can be destroyed by a number of transformations involved in this kind of system. By definition, strong ergodicity in the mean (or correlation) implies ergodicity in the mean (or correlation), and that it is invariant under the transformations involved in a multirate linear system, i.e., uniformly stable linear filtering, downsampling, upsampling and addition (in the case of strong ergodicity in the correlation, the signals to be added need also be jointly strongly ergodic in the correlation). We have shown that independent random processes, including deterministic bounded signals, are strongly ergodic in the mean (or correlation) and that mutually independent random processes are jointly strongly ergodic in the correlation. Therefore, all the signals generated from these signals by the transformations mentioned above are strongly ergodic in the mean (or correlation). As a consequence, most commonly used signals in multirate applications are strongly ergodic in the mean (or correlation) and thus ergodic in the mean (or correlation). It follows that a lot of stochastic analysis results for single-rate systems, requiring the ergodicity of signals in the system, can be readily applied to multirate systems.

As we mentioned in Introduction, there is another approach to ergodicity analysis which is based on the Kolmogorov’s SLLN. Using this approach, an elegant set of conditions can be given for ergodicity for every measurable function of the given random process, including mean and correlation as special cases. However, one of the required conditions is that the random process needs to be AMS which, as we have shown, is an inadequate condition for multirate signal processing. It remains a challenging problem to find a unified approach to general ergodicity analysis for random processes in multirate signal processing.

**APPENDIX**

**Proof of Proposition 1:** Suppose \( x \) has BAC. Then,
\[
\frac{1}{T} \sum_{t,s=1}^{T} |\langle x(s+d), x(t+d) \rangle|^2 \\
\leq \left( \frac{1}{T} \sum_{t,s=1}^{T} |\langle x(s), x(t) \rangle|^2 \right)^{1/2} \left( \frac{1}{T} \sum_{t,s=1}^{T} |\langle y(s), y(t) \rangle|^2 \right)^{1/2}
\]

Hence, \( x \) has WBAC.

To prove Proposition 2, we need Lemmas 1 and 2 below.

**Lemma 1:** If \( x, y \in \mathcal{S}(\Omega, \mathcal{A}, \mathbb{P}) \), then
\[
\frac{1}{T} \sum_{t,s=1}^{T} |\langle y(s), x(t) \rangle|^2 \\
= \left( \frac{1}{T} \sum_{t,s=1}^{T} \left| \langle y(s), e_n \rangle \langle e_n, x(t) \rangle \right|^2 \right)^{1/2} \\
= \frac{1}{T} \sum_{t,s=1}^{T} \left( \sum_{n \in \mathbb{N}} \langle y(s), e_n \rangle \langle e_n, x(t) \rangle \langle e_m, y(s) \rangle \langle x(t), e_m \rangle \right) \\
= \frac{1}{T} \sum_{n,m \in \mathbb{N}} \left( \sum_{t=1}^{T} \langle x(t), e_m \rangle \langle e_n, x(t) \rangle \right) \\
\cdot \left( \sum_{s=1}^{T} \langle y(s), e_m \rangle \langle e_n, y(s) \rangle \right) \\
\leq \left( \frac{1}{T} \sum_{n,m \in \mathbb{N}} \left| \sum_{t=1}^{T} \langle x(t), e_m \rangle \langle e_n, x(t) \rangle \right|^2 \right)^{1/2} \\
\cdot \left( \frac{1}{T} \sum_{n,m \in \mathbb{N}} \left| \sum_{s=1}^{T} \langle y(s), e_m \rangle \langle e_n, y(s) \rangle \right|^2 \right)^{1/2}
\]

where the first equality follows from Parseval’s identity [12, Theorem 4.13 (e), p.17] and the last inequality follows from
Hölder’s inequality [13, p. 26]. It follows that

\[
\frac{1}{T} \sum_{n,m \in \mathbb{N}} \left| \sum_{t=1}^{T} \langle x(t), e_m \rangle \langle e_n, x(t) \rangle \right|^2
\]

\[
= \frac{1}{T} \sum_{t,s=1}^{T} \sum_{n \in \mathbb{N}} \langle x(s), e_n \rangle \langle e_n, x(t) \rangle
\]

\[
= \frac{1}{T} \sum_{t,s=1}^{T} \left( \sum_{n \in \mathbb{N}} \langle x(s), e_n \rangle \langle e_n, x(t) \rangle \right)
\]

\[
\leq \frac{1}{T} \sum_{t,s=1}^{T} \left( \sum_{n \in \mathbb{N}} \langle x(s), e_n \rangle \langle e_n , x(t) \rangle \right)
\]

\[
= \frac{1}{T} \sum_{t,s=1}^{T} \left| \langle x(s), x(t) \rangle \right|^2
\]

A similar inequality above applies to the last term in (15). Substituting these inequalities into (15), we obtain the required result. ■

Lemma 2: Suppose \( x \in S(\Omega, A, \mathbb{P}) \). Then, \( \|x(t)\|_2 \leq \|x\|_S \) for all \( t \in \mathbb{Z} \). In particular, \( x \) has uniformly bounded second moments.

Proof: We have that

\[
\|x\|_S = \sup_{T \in \mathbb{N}, d \in \mathbb{Z}} \left( \frac{1}{T^2} \sum_{t,s=1}^{T} \|x(s + d), x(t + d)\|^2 \right)^{1/4}
\]

\[
\geq \sup_{d \in \mathbb{Z}} \left( \|x(1 + d), x(1 + d)\|^2 \right)^{1/4} = \sup_{d \in \mathbb{Z}} \|x(d)\|_2
\]

Proof of Proposition 2:

(N1): It follows immediately from (4).

(N2): For fixed \( d \in \mathbb{Z} \) and \( T \in \mathbb{N} \), we use the notation

\[
\|x, y\|_\# = \left( \frac{1}{T} \sum_{t,s=1}^{T} \|x(s + d), y(t + d)\|^2 \right)^{1/2}
\]

Let \( z(t) = x(t) + y(t) \). Then,

\[
\|z\|_\# \leq \|x, x\|_\# + \|y, y\|_\# + 2 \|x, y\|_\#
\]

\[
\leq \|x, x\|_\# + \|y, y\|_\# + 2 \|x, x\|^{1/2} \|y, y\|^{1/2}
\]

\[
= \left( \|x, x\|^{1/2} + \|y, y\|^{1/2} \right)^2
\]

where the first inequality follows from the triangular inequality for the 2-norm, and the second inequality follows from Lemma 1. Then, since \( \|x\|_S = \|x, x\|^{1/2} \), it follows that

\[
\|z\|_S \leq \|x\|_S + \|y\|_S
\]

(N3): From Lemma 2, we have that \( \|x\|_S = 0 \) implies that \( \|x(t)\|_2 = 0 \), for all \( t \in \mathbb{N} \), and therefore \( x = 0 \).

In order to prove Proposition 3, we need Lemmas 3 and 4, which are proved by following the proof of Rajchman’s SLLN [8, Theorem 5.1.2, p.103].

Lemma 3: Let \( x \) be a random process with

\[
\sum_{T = 1}^{\infty} \mathcal{E}\{|x(T)|^2\} < \infty
\]

then

\[
\lim_{T \to \infty} \mathbb{P}\{x(T) \neq 0\} = 1
\]

Proof: From the Chebyshev inequality [8], we have

\[
\mathbb{P}\{|x(T)| > \epsilon\} \leq \frac{1}{\epsilon^2} \mathbb{E}\{|x(T)|^2\}, \quad \forall \epsilon > 0
\]

Then, (16) implies that

\[
\sum_{T = 1}^{\infty} \mathbb{P}\{|x(T)| > \epsilon\} < \infty, \quad \forall \epsilon > 0
\]

Now, by the Borel-Cantelli Lemma [8], (18) implies that

\[
\mathbb{P}\left\{ \lim_{T \to \infty} \sup \{\|x(T)\| > \epsilon\} \right\} = 0, \quad \forall \epsilon > 0
\]

and (17) follows from Theorem 4.2.2 in [8]. ■

Lemma 4: Suppose a random process \( x \) has uniformly bounded second moments and there exists \( K_x > 0 \) such that

\[
\|E_x(T)\|_2 \leq \frac{K_x}{T^{1/4}}, \quad \forall T \in \mathbb{N}
\]

where

\[
E_x(T) = \frac{1}{T} \sum_{t=1}^{T} x(t)
\]

Then,

\[
\lim_{T \to \infty} E_x(T) \overset{w.p.1}{=} 0
\]

Proof: Let \( S(T) = \sum_{t=1}^{T} x(t) \) and define

\[
C(L) = \frac{1}{L^2} |S(L^4)|
\]

\[
D(L) = \frac{1}{L^2} \max_{L^4 \leq T < (L+1)^4} |S(T) - S(L^4)|
\]

We split the proof into three steps.

Step 1: We have

\[
\mathbb{E}\{|C(L)|^2\} = \mathbb{E}\{|E_x(L^4)|^2\} = \|E_x(L^4)\|_2^2
\]

Then,

\[
\sum_{L=1}^{\infty} \mathbb{E}\{|C(L)|^2\} = \sum_{L=1}^{\infty} \|E_x(L^4)\|_2^2 \leq \sum_{L=1}^{\infty} \frac{K_x^2}{L^2} < \infty
\]

Applying Lemma 3 on \( C \), we get

\[
\lim_{L \to \infty} C(L) \overset{w.p.1}{=} 0
\]
Step 2: Let $M_x$ be the bound of the second moments of $x(t)$. We have
\[
\mathcal{E}\{|D(L)|^2\} = \frac{1}{L^8} \mathcal{E}\left\{\max_{L^4 \leq T \leq (L+1)^4} |S(T) - S(L^4)|^2 \right\} \\
= \frac{1}{L^8} \mathcal{E}\left\{\max_{L^4 \leq T \leq (L+1)^4} \sum_{t,s=L^4+1}^T |\tilde{x}(t)x(s)| \right\} \\
\leq \frac{1}{L^8} \mathcal{E}\left\{\max_{L^4 \leq T \leq (L+1)^4} \sum_{t,s=L^4+1}^T |x(t)|^2 + |x(s)|^2 \right\} \\
\leq \frac{1}{L^8} \max_{L^4 \leq T \leq (L+1)^4} \sum_{t,s=L^4+1}^T M_x^2 \\
\leq \frac{1}{L^8} ((L+1)^4 - 1 - L^4 - 1)^2 M_x^2 \\
= \frac{(4L^3 + 6L^2 + 4L - 1)^2 M_x^2}{L^8}
\]
It follows that
\[
\sum_{L=1}^{\infty} \mathcal{E}\{|D(L)|^2\} < \infty
\]
Applying Lemma 3 on $D$, we have
\[
\lim_{L \to \infty} D(L) \overset{w}{\to} 0 \quad (23)
\]
Step 3: Given any $T \in \mathbb{N}$, let $L \in \mathbb{N}$ be such that $L^4 \leq T < (L + 1)^4$. Then,
\[
|E_x(T)| \leq \frac{1}{T} |S(L^4)| + \frac{1}{T} |S(T) - S(L^4)| \\
\leq \frac{1}{T^4} |S(L^4)| + \frac{1}{L^4} |S(T) - S(L^4)| \\
\leq C(L) + D(L)
\]
Hence, (21) follows from (22) and (23).

Proof of Proposition 3: Let $E_x(T) = \frac{1}{T} \sum_{t=1}^T x(t)$. Then,
\[
\|E_x(T)\|_2^2 = \mathcal{E}\left\{\frac{1}{T^2} \sum_{t,s=1}^T |\tilde{x}(t)x(s)| \right\} \\
\leq \left( \frac{1}{T^2} \sum_{t,s=1}^T |\mathcal{E}\{\tilde{x}(t)x(s)\}| \right)^{1/2} \leq \frac{|x|_S \|x\|_S}{T^{1/2}}
\]
where the second inequality follows from [13, p. 26, eq. (2.9.1)]. Also, from Lemma 2, $x$ has uniformly bounded second moments. Then, the result follows from Lemma 4.

Proof of Proposition 4: From Proposition 2, we have, for every $i \in \mathbb{Z}$,
\[
\|z_i\|_S \leq \|x_i\|_S + \|y_i\|_S \leq U_X + U_Y
\]
where $U_X$ and $U_Y$ are the uniform weak bounds of $X$ and $Y$, respectively. Hence, $Z$ has UWBAC.

Proof of Proposition 5: Since $\{h_k\}$ is uniformly stable, there exists $h \in l_1(\mathbb{Z})$ such that $|h_k(\tau)| \leq h(\tau)$ for all $\tau \in \mathbb{Z}$. Then,
\[
|\langle y(s), y(t) \rangle| = \left| \sum_{k,l=-\infty}^{\infty} h_k(k)h_s(l) \langle x_l(s-l), x_k(t-k) \rangle \right| \\
\leq \sum_{k,l=-\infty}^{\infty} h_k(k)h_s(l) |\langle x_l(s-l), x_k(t-k) \rangle|
\]
Using the triangular inequality, it is straightforward to verify that
\[
\|y\|_S^2 = \sup_{T \in \mathbb{N}, d \in \mathbb{Z}} \left( \frac{1}{T} \sum_{t,s=1}^T |\langle y(s+d), y(t+d) \rangle|^2 \right)^{1/2} \\
\leq \sum_{k,l=-\infty}^{\infty} h_k(k)h_s(l) \sup_{T \in \mathbb{N}, d \in \mathbb{Z}} \Pi(k, l, T, d)^{1/2}
\]
where
\[
\Pi(k, l, T, d) = \frac{1}{T} \sum_{t,s=1}^T |\langle x_l(s-l+d), x_k(t-k+d) \rangle|^2
\]
By Lemma 1, we get
\[
\sup_{T \in \mathbb{N}, d \in \mathbb{Z}} \left( \frac{1}{T} \sum_{t,s=1}^T |\langle x_l(s-l+d), x_k(t-k+d) \rangle|^2 \right)^{1/2} \\
\leq \sup_{T \in \mathbb{N}, d \in \mathbb{Z}} \left( \frac{1}{T} \sum_{t,s=1}^T |\langle x_l(s-l+d), x_l(t-l+d) \rangle|^2 \right)^{1/4} \\
\leq \sup_{T \in \mathbb{N}, d \in \mathbb{Z}} \left( \frac{1}{T} \sum_{t,s=1}^T |\langle x_k(s-k+d), x_k(t-k+d) \rangle|^2 \right)^{1/4}
\]
Therefore,
\[
\|y\|_S^2 \leq \sum_{k,l=-\infty}^{\infty} h_k(k)h_s(l) \|x_l\|_S \|x_k\|_S \\
\leq \left( \sum_{k,l=-\infty}^{\infty} h_k(k)h_s(l) \right) U_X^2 \leq \|h\|_1^2 U_X^2
\]
where $U_X$ is the uniform weak bound of $X$.

Proof of Proposition 6: Let $x \in X$ with uniform weak bound $U_X$. We analyze downsampling and upsampling separately.

Downsampling: Let $y = \Delta_D x$. We have
\[
\|y\|_S^2 = \sup_{T \in \mathbb{N}, d \in \mathbb{Z}} \left( \frac{1}{T} \sum_{t,s=1}^T |\langle x(D(s+d)), x(D(t+d)) \rangle|^2 \right) \\
\leq D \sup_{T \in \mathbb{N}, d \in \mathbb{Z}} \left( \frac{1}{DT} \sum_{t,s=1}^T |\langle x(s+Dd), x(t+Dd) \rangle|^2 \right) \\
\leq D \|x\|_S^2 \leq DU_X^2
\]
Upsampling: Let $y = \nabla_D x$. Let
\[
\hat{T} = ([T + d]_D - |d|_D)/D
\]
and \( \hat{d}(D,d) = \lfloor d \rceil_D \), where \( \lfloor x \rfloor_D = D \lfloor x/D \rfloor \) and \( \lfloor y \rfloor_D \) denotes the largest integer smaller than or equal to \( y \). Then,

\[
\| y \|_S^2 = \sup_{T \in \mathbb{N}, d \in \mathbb{Z}} \left\{ \frac{1}{T} \sum_{t,s=1}^{T} \| \langle y(s+d), y(t+d) \rangle \|^2 \right\}
\]

\[
= \sup_{T \in \mathbb{N}, d \in \mathbb{Z}} \left\{ \frac{1}{T} \sum_{t,s=1}^{T} \| \langle x(s+d, \hat{d}(D,d)), x(t+\hat{d}(D,d)) \rangle \|^2 \right\}
\]

\[
\leq D^{-1} \| x \|_S^2 \leq D^{-1} U_N^4
\]

In the last step above, we used the fact that \( \hat{T} \leq T/D \).

**Proof of Proposition 7:** Suppose \( x \in P_2(\Omega, A, \mathbb{P}) \) is such that \( \xi_x \) has WBAC. Then, Proposition 3 implies that (1) is satisfied, i.e., \( x \) is ergodic. Therefore, it suffices to show that the WBAC property is preserved under addition, uniformly stable linear filtering, downsampling and upsampling.

Suppose \( y \in P_2(\Omega, A, \mathbb{P}) \) is also such that \( \xi_y \) has WBAC. Since \( \xi_{x+y} = \xi_x(t) + \xi_y(t) \), it follows from Proposition 4 that \( \xi_{x+y} \) also has WBAC.

Let \( \{ h_t \in l_1(\mathbb{Z}) : t \in \mathbb{Z} \} \) be the impulse response of a uniformly stable linear filter and let \( y \) be given by

\[
y(t) = \sum_{\tau=-\infty}^{\infty} h_t(\tau)x(t-\tau)
\]

Then,

\[
\xi_y(t) = \sum_{k=-\infty}^{\infty} h_t(k)x(t-k) - \mathcal{E} \left\{ \sum_{k=-\infty}^{\infty} h_t(k)x(t-k) \right\}
\]

\[
= \sum_{k=-\infty}^{\infty} h_t(k)(x(t-k) - \mathcal{E} \{ x(t-k) \})
\]

\[
= \sum_{k=-\infty}^{\infty} h_t(k)\xi_x(t-k)
\]

where the exchange of the expectation with the infinite sum in the second equality is justified by continuity of the functional \( \mathcal{E} : L_2(\Omega, A, \mathbb{P}) \rightarrow \mathbb{C} \). It follows from Proposition 5 that \( \xi_y \) has WBAC.

The downsampling and upsampling properties follow from Proposition 6.

**Proof of Proposition 8:** Let \( M_x \) be the uniform bound of the second moments of \( x \) and let \( \xi_x \) be defined as in (6). We have that

\[
| \langle \xi_x(t), \xi_x(t) \rangle | \leq M_x^2, \quad \forall t \in \mathbb{N}
\]

(24)

Also, since \( x \) is white, if \( \tau \neq 0 \),

\[
\langle \xi_x(t+\tau), \xi_x(t) \rangle = 0, \quad \forall \tau \neq 0
\]

(25)

Equations (24) and (25) imply that the conditions for Proposition 1 hold for \( \xi_x \). Hence, \( \xi_x \) has WBAC. Using Proposition 7, it follows that \( x \) is strongly ergodic in the mean.

**Proof of Proposition 9:** Consider the set of random processes defined by

\[
\{ \psi_{(x,y,1,0,\tau)} : \tau \in \mathbb{Z} \} \subset \Psi_{(x,y)}
\]

It is obvious that this set has UWBAC. It follows from Proposition 3 that \( x \) and \( y \) are jointly ergodic. Thus, it suffices to show that the UWBAC property for \( \Psi_{(x,y)} \) in (7) is symmetric with respect to \( x \) and \( y \) and is preserved under addition, uniformly bounded filtering, downsampling and upsampling.

The symmetry property follows immediately from the fact

\[
\psi_{(x,y,\alpha,\beta,a,b)}(t) = \psi_{(y,x,\beta,\alpha,b,a)}(t)
\]

The addition property follows from the fact

\[
\psi_{(x,z+y,\alpha,\beta,a,b)}(t) = \psi_{(z,x,\alpha,\beta,a,b)}(t) + \psi_{(z,y,\alpha,\beta,a,b)}(t)
\]

and Proposition 4.

To see the filtering property, we let \( \{ h_t \in l_1(\mathbb{Z}) : t \in \mathbb{Z} \} \) be the impulse response of a uniformly stable linear filter. The filtered \( y \) is given by

\[
z(t) = \sum_{k=-\infty}^{\infty} h_t(k)y(t-k)
\]

It follows that

\[
\psi_{(x,z,\alpha,\beta,a,b)}(t) = \hat{x}(at + a) \sum_{k=-\infty}^{\infty} h_{\beta t + b}(k)y(\beta t + b - k) + \mathcal{E} \left\{ \hat{x}(at + a) \sum_{k=-\infty}^{\infty} h_{\beta t + b}(k)y(\beta t + b - k) \right\}
\]

\[
= \sum_{k=-\infty}^{\infty} h_{\beta t + b}(k)\psi_{(x,y,\alpha,\beta,a,b)}(t)
\]

\[
= \sum_{k=-\infty}^{\infty} h_{\beta t + b}(k)\psi_{(x,y,\alpha,\beta,a+b+k(\beta-1))}(t-k)
\]

Note that the exchanging of the expectation with the infinite sum in the second equality above is justified by continuity of the functional \( \mathcal{E} : L_2(\Omega, A, \mathbb{P}) \rightarrow \mathbb{C} \). Using Proposition 5, we know that the set of random processes \( \Psi_{x,z} \) has UWBAC.

The downsampling property holds because for any \( z(t) = \Delta_D y(t), D \in \mathbb{N} \), we have

\[
\psi_{(x,z,\alpha,\beta,a,b)}(t) = \psi_{(x,y,\alpha,D\beta a,Db)}(t)
\]

for all admissible \( \alpha, \beta, a \) and \( b \), which implies that \( \Psi_{x,z} \) has UWBAC.

Similarly, the upsampling property is confirmed as follows. Let \( \alpha, \beta \in \mathbb{N} \) and \( a, b \in \mathbb{Z}, D \in \mathbb{N} \) and \( z(t) = \Delta_D y(t) \). Let \( T = \{ t_i \in \mathbb{Z}^2 : i \in \mathcal{T} \} \) and \( K = \{ k_i \in \mathbb{Z}^2 : i \in \mathcal{T} \} \) be two sets of integers that satisfy

\[
\beta t_i + b = k_i D
\]

(26)

We consider two cases. In the first case, there is no integer solution of (26), i.e., \( \mathcal{T} = \emptyset \). It follows that

\[
\psi_{(x,z,\alpha,\beta,a,b)}(t) = 0 \quad \text{for all} \quad t \in \mathbb{Z}
\]

Therefore, \( \psi_{(x,z,\alpha,\beta,a,b)} \) has WBAC. In the second case, (26) admits an integer solution. In this case, \( \mathcal{T} = \mathbb{Z} \). Therefore, we can write

\[
t_i = \gamma i + c \quad \text{and} \quad k_i = \delta i + d
\]
for some $\gamma, \delta \in \mathbb{N}$ and $c, d \in \mathbb{Z}$. It can be verified that

$$
\psi_{(x,z,a,\alpha,\beta,b)}(t) = \begin{cases} 
\psi_{(x,y,\alpha\gamma,\delta,\alpha c+a,\beta t+b)}(\frac{t-x}{\gamma}), & \frac{t-x}{\gamma} \in \mathbb{Z} \\
0, & \frac{t-x}{\gamma} \notin \mathbb{Z}
\end{cases}
$$

or equivalently,

$$
\psi_{(x,z,a,\alpha,\beta,b)}(t) = q^{-e}Y_{\gamma}\psi_{(x,y,\alpha\gamma,\delta,\alpha c+a,\beta t+b)}(t)
$$

where $q$ is the forward shift operator. Now, the set of random processes $\{\psi_{(x,y,\alpha\gamma,\delta,\alpha c+a,\beta t+b)} : \alpha, \beta \in \mathbb{N}; a, b \in \mathbb{Z}\}$ has UWBC. Moreover, from Proposition 6, the operator $Y_{\gamma}$ preserves the UWBC property. Finally, the operator $q^{-e}$ is a special uniformly stable linear filter. By Proposition 5, $\Psi_{(x,z)}$ has UWBC.

Proof of Proposition 10: We only prove Part (i). Part (ii) can be shown using a similar but simpler argument. To show (i), let $M_x$ be a bound of the fourth moments of $x$. Then, for every $\tau \in \mathbb{Z}$, it can be verified that

$$
|\langle \psi_{(x,x,a,\alpha,\beta,b)}(t+\tau), \psi_{(x,x,a,\alpha,\beta,b)}(t) \rangle| \leq 2M_x^4
$$

(The proof of (27) requires the use of H"{o}lder’s inequality [8, p. 47, eq. (18)] and Lyapunov’s inequality [8, p. 47, eq. (21)]). Also, it is straightforward to verify that, for every $\alpha, \beta \in \mathbb{N}$ and $a, b \in \mathbb{Z}$, there exists $\tilde{\tau} \in \mathbb{N}$ such that, if $|\tau| > \tilde{\tau}$, then $\max\{\alpha t + a, |\beta t + b|\} \leq \min\{\alpha(t+\tau) + a, |\beta(t+\tau) + b|\}$. Using [8, p.51, Theorem 3.3.2], the random variables $x(\alpha t + a)\hat{x}(\beta t + b)$ and $\bar{x}(\alpha(t+\tau) + a)x(\beta(t+\tau) + b)$ are independent when $|\tau| > \tilde{\tau}$. It follows that

$$
\langle \psi_{(x,x,a,\alpha,\beta,b)}(t+\tau), \psi_{(x,x,a,\alpha,\beta,b)}(t) \rangle = 0, \ \forall|\tau| > \tilde{\tau}
$$

Since $x$ has uniformly bounded fourth moments, $x \in P_4(\mathbb{N}, \mathbb{A}, \mathbb{P}) \subset P_2(\mathbb{N}, \mathbb{A}, \mathbb{P})$. Therefore, equations (27) and (28) imply that the conditions for Proposition 1 hold. Subsequently, $\Psi_{(x,x)}$ has UWBC. Using Proposition 7, $x$ is strongly ergodic in the correlation.

Proof of Theorem 2: It follows from Proposition 10 that every member in $\{u_m : 1 \leq m \leq M\}$ is strongly ergodic in the correlation and every pair in $\{u_m : 1 \leq m \leq M\}$ jointly strongly ergodic in the correlation. We now claim that all the operations on the signals $u_m$, $m = 1, \cdots, M$, as mentioned in the theorem can be decomposed into those in (SEC2)-(SEC5). For example, if $w$ is downsamped from some $u_n$ with $1 \leq n \leq M$, then this operation can be decomposed into two applications of (SEC4). In the first application, we take $x = u_m$ with any $1 \leq m \leq M$ and $y = u_n$. It follows that $w$ is jointly strongly ergodic in the correlation with every $u_m$, $1 \leq m \leq M$. In the second application of (SEC4), we take $y = u_m$ and $x = w$, which yields that $w$ is strongly ergodic in the correlation. Thus, the new set $\{w, u_m : 1 \leq m \leq M\}$ has the same ergodicity properties as $\{u_m : 1 \leq m \leq M\}$ has. This procedure can continue until the required random process $v$ is generated. By induction, $v$ is also strongly ergodic in the correlation.

REFERENCES

LIST OF FIGURES

• Fig. 1. Subband identification scheme.