Ergodic Properties for Multirate Linear Systems

Damián Marelli and Minyue Fu, *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. We show that these properties are preserved under a number of operations, namely, downsampling, upsampling, addition, and uniformly stable linear (time-variant) filtering. We also 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 white (independent) random processes with uniformly bounded second (fourth) moments, then every signal in the system is strongly ergodic in the mean (correlation) and therefore ergodic in the mean (correlation). An application of these results is also discussed.**

***Index Terms*—Filter bank design and theory, multirate processing and multiresolution methods, nonstationary statistical signal processing.**

I. Introduction

MULTIRATE signal processing techniques find a wide range of applications (see, e.g., [1]–[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] and [7]).

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 (i.e., the time averages of the mean/correlation converge with probability one to a given constant). 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 this fact, we look for suitable notions of ergodicity for multirate linear systems.

Another fundamental aspect in stochastic analysis is the study of stationarity properties, i.e., whether a random process is stationary, cyclo-stationary (also called block-stationary), asymptotically mean stationary (AMS), etc. In the case of multirate linear systems, this is done in [8] and [9].

As pointed out in [10, Sec. 6.14], there is a conflict of terminology regarding the notion of ergodicity. In engineering language, ergodicity with respect to a measurement (e.g., mean or correlation) means that the time average of the measurement, calculated on a single realization of the random process, converges with probability one to a given value [11]. On the other hand, in the language of probability theory, a random process is ergodic if the only sets that are invariant under the time shift map (induced in the underlying probability space) have probability either zero or one [12], [13]. In this paper, we will use the term ergodicity to refer to the engineering definition and use the term *metrically transitive* to refer to the probability theory definition, following [14].

To prove that a random process is ergodic is essentially to verify that the Strong Law of Large Numbers (SLLN) [15] is satisfied. The results providing conditions to guarantee this law can be classified into two categories.

C1) In the first category are those results requiring some sort of stationarity condition on the random process considered. A first step in this direction is Kolmogorov's SLLN [15, Theorem 5.4.2, p. 133]. It requires the random process to be a sequence of independent, identically distributed, random variables with finite first absolute moments. We will call such a sequence an independent, strictly stationary random process. A generalization of Kolmogorov's SLLN is Birkhoff's ergodic theorem [16], which is the cornerstone of ergodic theory [12], [13]. In this generalization, the independence of the random process is relaxed to metric transitivity. This theorem provides conditions for ergodicity with respect to virtually any measurement (including the mean and correlation). A further generalization of Birkhoff's ergodic theorem relaxes the strict stationarity condition to AMS [17], [18]. We will refer to this generalization as the AMS ergodic theorem. This theorem provides the foundation to guarantee ergodicity of many commonly used random processes, e.g., strictly stationary random processes [19], Markov processes [14, Ch. V, Sec. 6] and hidden Markov processes [20]. We point out that there exist, in the probability theory literature, conditions stronger than metric transitivity, namely, *totally*

Manuscript received May 4, 2005; revised March 30, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Franz Hlawatsch.

The authors are with the School of Electrical Engineering and Computer Science, University of Newcastle, N.S.W. 2308 Australia (e-mail: damian.marelli@newcastle.edu.au; eemf@ece-mail.newcastle.edu.au).

Digital Object Identifier 10.1109/TSP.2006.885687

ergodic [18, p. 159] and *weakly* and *strongly mixing* [18, p. 149].

C2) In the second category are those results without stationarity requirements. A first attempt along this direction is Rajchman's SLLN [15, Theorem 5.1.2, p. 108]. It just requires that the random process be a sequence of pairwise uncorrelated random variables with uniformly bounded second absolute moments. Several results generalize Rajchman's SLLN by replacing the requirement on the uncorrelation of random variables with a requirement on the decaying rate of the autocorrelation function [21]–[23]. An earlier result for the continuous time case, which, with trivial modifications, is also valid for the discrete time case, was given in [24, pp. 94–96]. These results provide sufficient conditions for ergodicity in the mean. Another approach is taken in [25, Theorem 2.3], where it is shown that if an independent random process is filtered by a uniformly stable time-invariant linear filter, then the resulting random process is ergodic in the correlation.

A random process is said to have ergodic properties with respect to a measurement if the time average of the measurement, converges with probability one to some random variable [18, p. 124]. If, in addition, it is desired that the random process is ergodic with respect to that measurement (i.e., all the realizations of the measurement converge to the same constant value with probability one), then the random process is also required to be metrically transitive [14, Ch. X, Theorem 2.1]. According to the ergodic theory for AMS random processes, the AMS condition is necessary and sufficient for a random process to have ergodic properties with respect to all bounded measurements (i.e., not including mean and correlation) [18, Corollary 7.2.2]. Also, it states that the ergodicity of a random process is preserved if it is modified via any transformation that preserves the metrically transitive and AMS properties. Such transforms are known as ergodic channels and AMS channels, respectively [26]. It turns out that the kind of transformations that take place in a multirate, time-variant, linear system do not satisfy these requirements. More precisely, as we show in Section VIII, the AMS condition can be destroyed by time-variant operations, and the resulting non-AMS random process (which does not have ergodic properties with respect to all bounded measurements) still has the ergodic properties desired in a signal processing application, namely, in the mean and correlation. This suggests that the AMS property is a too strong requirement for a general multirate, time-invariant, linear system in a signal processing context. Therefore, we will align our analysis with results from category C2), i.e., those not requiring stationarity.

More specifically, we introduce the notions of strong ergodicity in the mean and strong ergodicity in the correlation. We show that these properties are preserved under a number of commonly used linear multirate operations, namely downsampling, upsampling, addition, and filtering by uniformly stable linear filters. We also 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.

The results in the present paper can be considered as a step further in the results from category C2). In contrast with [25, Theorem 2.3], we consider both ergodicity in the mean and ergodicity in the correlation, and we provide conditions to allow, not only (uniformly stable) time-variant linear filtering, but also multirate operations and addition of signals. To this end, we provide a condition on the decaying rate of the autocorrelation function which, in contrast with [21]–[24], is invariant under the four transformations mentioned above.

For readers who are more familiar with the ergodicity theory, we point out a difference between the notions of strong ergodicity and the known notion of total ergodicity [18, p. 159]. A random process is called totally ergodic if the random processes obtained by downsampling it, using all possible downsampling factors, are all metrically transitive. It is easy to see that the total ergodicity property is invariant under nonlinear, cyclo-stationary mappings. It is also straightforward to verify that independent identically distributed (i.i.d.) processes are totally ergodic. Therefore, the conceptual structure formed by the triad {total ergodicity, cyclo-stationary mappings, i.i.d. processes} can be considered as a C1) counterpart of the C2)-based triad {strong ergodicity in the mean (correlation), the four invariant transformations, white (mutually independent) processes with uniformly bounded second (fourth) moments}. In comparison with total ergodicity, strong ergodicity requires fewer stationarity constraints, and for that reason, it is a more suitable property to consider when analyzing multirate, time-variant, linear systems.

As mentioned above, an example of a multirate linear system is the so-called subband system identification technique. 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 illustrate our results, we introduce an application, namely, the analysis of convergence properties of a subband identification system. More precisely, we analyzed strong convergence (i.e., whether or not the outcome of the identification, as the number of samples goes to infinity, depends on the realizations of the involved random processes) and optimum convergence (i.e., whether or not the outcome of the identification converges to its "best possible"). It turns out that, in order to guarantee the required convergence properties, the subband signals involved need to be ergodic and jointly ergodic in the correlation. This result is developed in detail in [6], where some preliminary results on strong ergodicity are used.

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 the 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 Ω with σ-algebra $\mathcal{A}$ and probability measure $\mathbb{P}$. A random variable is an $\mathcal{A}$-measurable map $v : \Omega \to \mathbb{C}$, 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 pth (absolute) moment (or pth norm in a functional analysis context) 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 $\{x_i : i = 1, 2, \ldots, n\}$ is said to be *independent* if, for any $B_i \in \mathcal{B}$, $i = 1, 2, \ldots, n$, 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

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

A random process x is said to have *uniformly bounded pth (absolute) moments* with bound M_x, if

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

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 pth 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 $\ell_\infty(\mathbb{Z})$.

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. In addition, 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

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

A time-variant linear filter is characterized by a family of impulse responses $\{h_t \in \mathbb{C}^{\mathbb{Z}} : t \in \mathbb{Z}\}$ ($\mathbb{C}^{\mathbb{Z}}$ is the space of complex sequences indexed by the integers), i.e., if $x \in \mathcal{P}_2(\Omega, \mathcal{A}, \mathbb{P})$ is the filter's input signal and y is its output, then

$$y(t) = \sum_{\tau=-\infty}^{\infty} h_t(\tau)x(t-\tau).$$

A time-variant linear filter with impulse response $\{h_t \in \mathbb{C}^{\mathbb{Z}} : t \in \mathbb{Z}\}$ is said to be uniformly stable if there exists $h \in \ell_1(\mathbb{Z})$ such that $|h_t(\tau)| \le h(\tau), \forall t, \tau \in \mathbb{Z}$.

C. Ergodicity

In [11], a mean stationary random process is said to be ergodic in the mean (or mean ergodic) if

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} x(t) \stackrel{\text{w.p.1}}{=\!=} \mathcal{E}\{x(k)\} \quad \text{for all } k \in \mathbb{Z}. \tag{1}$$

As we mentioned in Section I, we need to deal with nonstationary random processes. To this end, a standard way to generalize (1) is as follows:

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} x(t) \stackrel{\text{w.p.1}}{=\!=} \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathcal{E}\{x(t)\}. \tag{2}$$

However, the existence of the limit in the right-hand side of (2) still requires some kind of stationarity. Hence, in order to make the definition of ergodicity in the mean completely independent of any stationarity condition, we introduce the following definition.

Definition 1: A random process $x \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$ is said to be *ergodic in the mean* if

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T}[x(t) - \mathcal{E}\{x(t)\}] \overset{\text{w.p.1}}{=\!=} 0. \tag{3}$$

If x is such that the limit in the right-hand side of (2) does not exist, intuitively, (3) says that, for any given realization, the sequence $(1/T)\sum_{t=1}^{T} x(t)$ approaches the sequence $(1/T)\sum_{t=1}^{T} \mathcal{E}\{x(t)\}$, with probability one, as T goes to infinity. Notice that, in this case, and according to the definitions in [18], the random process x is not AMS and does not have ergodic properties with respect to the expectation.

Following a similar reasoning, we define ergodicity in the correlation as follows.

Definition 2: two random processes $x, y \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$ are said to be *jointly ergodic in the correlation* if, for every $\tau \in \mathbb{Z}$

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T}[\bar{x}(t)y(t+\tau) - \mathcal{E}\{\bar{x}(t)y(t+\tau)\}] \overset{\text{w.p.1}}{=\!=} 0. \tag{4}$$

Also, a random process is *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.

1) Example 1: Consider the probability space $([-1/2, 1/2], \mathcal{B}, \lambda)$, where $\mathcal{B}$ denotes the Borel σ-algebra on the set $[-1/2, 1/2]$ and λ denotes the Lebesgue measure. Define the random process $x \in \mathcal{P}([-1/2, 1/2], \mathcal{B}, \lambda)$ 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

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

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

$$\frac{1}{T}\sum_{t=1}^{T}[\bar{z}(t,\omega)z(t+\tau,\omega) - \mathcal{E}\{\bar{z}(t,\omega)z(t+\tau,\omega)\}]$$
$$= \frac{1}{2}\text{sign}(\omega) \nrightarrow 0, \ T\to\infty. \tag{5}$$

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.

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

$$h_t(\tau) = \begin{cases} \delta(\tau), & t \text{ is even} \\ \delta(\tau-1), & t \text{ is odd} \end{cases}$$
$$\text{where} \quad \delta(\tau) = \begin{cases} 1, & \tau = 0 \\ 0, & \tau \neq 0 \end{cases}.$$

Let z be obtained by filtering x through $h_t(\tau)$, i.e.,

$$z(t,\omega) = \sum_{\tau=-\infty}^{\infty} h_t(\tau)x(t-\tau,\omega) = \frac{1}{2}(1+\text{sign}(\omega)).$$

Then (5) 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 that are ergodic in the correlation.

3) Example 3: Consider the probability space $([-1/2, 1/2], \mathcal{B}, \lambda)$. Define the random processes $x, y \in \mathcal{P}([-1/2, 1/2], \mathcal{B}, \lambda)$ 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 $\tau \in \mathbb{Z}$, $T \in \mathbb{N}$

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

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, (5) holds, and therefore, z is not ergodic in the correlation.

IV. Random Processes With Weakly Bounded Autocorrelation

In this section, we introduce a class of random processes with autocorrelation obeying a certain decaying rate bound. This property is instrumental in the analysis of strong ergodicity. As mentioned in Introduction, conditions of the decaying rate of the autocorrelation function have been introduced in [24, pp. 94–96], [21], [22]. The key difference between these conditions and our condition (6) is that (6) is invariant under uniformly stable linear filtering, downsampling, upsampling, and additions.

Definition 3: Let $x \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$ be a random process. We define

$$\|x\|_{\mathcal{S}} = \sup_{T\in\mathbb{N}, d\in\mathbb{Z}} \left(\frac{1}{T} \sum_{t,s=1}^{T} |\langle x(s+d), x(t+d)\rangle|^2 \right)^{1/4}. \quad (6)$$

We say x has ℓ_2*-norm bounded autocorrelation* [or simply *bounded autocorrelation* (BAC)] if there exists $f \in \ell_2(\mathbb{Z})$ such that

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

We say x has *weakly bounded autocorrelation* (WBAC) if $\|x\|_S < \infty$. The space of random processes with WBAC will be denoted by

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

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

$$\|x\|_{\mathcal{S}} \leq U_X, \ \forall x \in X. \quad (7)$$

In this case, U_X is called a *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 $x \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$ has BAC, it has WBAC.

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

1) $\|cx\|_{\mathcal{S}} = |c|\|x\|_{\mathcal{S}}$;
2) $\|x+y\|_{\mathcal{S}} \leq \|x\|_{\mathcal{S}} + \|y\|_{\mathcal{S}}$;
3) $\|x\|_{\mathcal{S}} = 0 \Rightarrow x = 0$.

Random processes with WBAC enjoy the following important properties.

Proposition 3: If a random processes $x \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$ has UWBAC, then

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

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

$$Z = \{z_i = x_i + y_i : i \in \mathcal{I}\}$$

also has UWBAC.

Proposition 5: Let $X \subset \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$ be a set of random processes with UWBAC and $\{h_t \in \ell_1(\mathbb{Z}) : t \in \mathbb{Z}\}$ 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

$$y(t) = \sum_{k=-\infty}^{\infty} h_t(k) x_k(t-k), \ \ x_k \in X.$$

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

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

V. Strong Ergodicity in the Mean

In this section, we introduce the notion of strong ergodicity in the mean and characterize random processes that are strongly ergodic in the mean.

Definition 4: A random process $x \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$ is said to be *strongly ergodic in the mean* if $x \in \mathcal{P}_2(\Omega, \mathcal{A}, \mathbb{P})$ and the random process ξ_x defined by

$$\xi_x(t) = x(t) - \mathcal{E}\{x(t)\} \quad (8)$$

has WBAC.

A random process which is strongly ergodic in the mean enjoys the following properties:

Proposition 7: Let $x \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$ be a random process that is strongly ergodic in the mean. Then, the following holds.

SEM1) x is ergodic in the mean;
SEM2) if the random process $y \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$ 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.

The next result shows that white random processes are strongly ergodic in the mean.

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

Propositions 7 and 8 together imply the following theorem.

Theorem 1: Given a finite set of random processes $u_m \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$, $m = 1, \ldots, M$, 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 $\ell_\infty(\mathbb{Z})$ (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. Property SEM1) says that strong ergodicity in the mean implies ergodicity in the mean, and SEM2) says that the addition of two signals that are strongly ergodic in the mean yields another signal that is also strongly ergodic in the mean. In contrast, we show via in Example 4 below that the addition of two independent signals, which are not mutually independent, can yield a signal which is nonergodic in the correlation. Since the definition of strong ergodicity in the correlation is expected to imply ergodicity in the correlation, it is not possible to have a condition analogous to SEM2).

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

$$\omega = \sum_{t=1}^{\infty} \omega_t 2^{-t}.$$

Define the random process $x \in \mathcal{P}([0,1), \mathcal{B}, \lambda)$ by

$$x(t,\omega) = 2\omega_t - 1.$$

Define a new random process y by

$$y(t,\omega) = \begin{cases} x(t,\omega), & \omega \in \left[0, \frac{1}{2}\right) \\ -x(t,\omega), & \omega \in \left[\frac{1}{2}, 1\right) \end{cases}.$$

By construction, x and y are independent signals [15, Example 4, p. 60] taking values in $\{-1, 1\}$, although they are not mutually independent. Define a new random process z by

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

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

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

It follows that

$$\frac{1}{T} \sum_{t=1}^{T} [\bar{z}(t,\omega) z(t,\omega) - \mathcal{E}\{\bar{z}(t,\omega) z(t,\omega)\}] = \begin{cases} 2, & \omega \in \left[0, \frac{1}{2}\right) \\ -2, & \omega \in \left[\frac{1}{2}, 1\right) \end{cases}.$$

Therefore, z is not ergodic in the correlation.

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

Definition 5: Let $x, y \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$ be random processes. Define a countable set of random processes $\Psi_{(x,y)} \subset \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$ by

$$\Psi_{(x,y)} = \{\psi_{(x,y,\alpha,\beta,a,b)} : \alpha, \beta \in \mathbb{N}; a, b \in \mathbb{Z}\} \tag{9}$$

where

$$\psi_{(x,y,\alpha,\beta,a,b)}(t) = \bar{x}(\alpha t + a) y(\beta t + b) - \mathcal{E}\{\bar{x}(\alpha t + a) y(\beta t + b)\}. \tag{10}$$

If $x, y \in \mathcal{P}_2(\Omega, \mathcal{A}, \mathbb{P})$ and $\Psi_{(x,y)}$ has UWBAC, then x and y are said to be *jointly strongly ergodic in the correlation*. Also, a random process is called *strongly ergodic in the correlation* if it is jointly strongly ergodic in the correlation with itself.

Two random processes that are strongly ergodic in the correlation enjoy the following properties.

Proposition 9: Let $x, y \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$, be two random processes which are jointly strongly ergodic in the correlation. Then, the following holds.

- SEC1) x and y are jointly ergodic in the correlation;
- SEC2) If the random process $z \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$ 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 the 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.

As in the previous section, we provide a result showing the relationship between independence and strong ergodicity.

Proposition 10: Let $x, y \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$ 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.

Propositions 9 and 10 lead to the following result.

Theorem 2: Given a finite set of random processes $u_m \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$, $m = 1, \ldots, M$, 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 5, this random process x cannot be strongly ergodic in the correlation, which in turn implies that $\Psi_{(x,x)}$ as defined in (9) does not have UWBAC. This is verified below.

2) Example 5: Consider the random process x from Example 1. Let $\psi_{(x,x,\alpha,\beta,a,b)}$ be defined as in (10). Then

$$\psi_{(x,x,1,1,0,0)}(t) = \bar{x}(t)x(t) - \mathcal{E}\{\bar{x}(t)x(t)\}.$$

It follows that

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

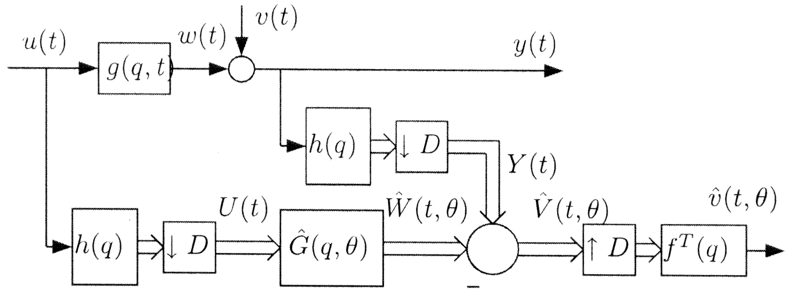


Fig. 1. Subband identification scheme.

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

$$\frac{1}{T}\sum_{t,s=1}^{T} |\langle \psi_{(x,x,1,1,0,0)}(t+d), \psi_{(x,x,1,1,0,0)}(s+d)\rangle|^2 = \frac{T}{16}.$$

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

VII. 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 Fig. 1. The idea is to identify the possibly time-variant linear system $g(q,t)$, by splitting both signals u and y (called fullband signals) into M subbands, using two identical analysis filter banks

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

Then, these subband signals are downsampled, by a factor of D. The results are denoted by two vector signals

$$U(t) = [U_1(t), \ldots, U_M(t)]^T$$
$$\text{and} \quad Y(t) = [Y_1(t), \ldots, Y_M(t)]^T$$

respectively. The subband parametric model is given by

$$\hat{G}(q,\theta) = \operatorname{diag}\{\hat{G}_m(q,\theta_m),\ m = 1,\ldots,M\}$$

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

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

and the subband prediction error is defined by

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

Each subband model $\hat{G}_m(q,\theta_m)$ is tuned to minimize the power of $\hat{V}_m(t,\theta_m)$. An upsampler and a synthesis filter bank

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

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

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

$$\lim_{N\to\infty} S_{\tilde{W}_m}(\theta_{m,N}) \overset{\text{w.p.1}}{=} \min_{\theta_m} S_{\tilde{W}_m}(\theta_m). \tag{11}$$

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

$$S_{\tilde{W}_m}(\theta_m) := \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathcal{E}\{|\hat{W}_m(t,\theta_m)|^2\}.$$

One of the technical conditions required to satisfy (11) is that the signals $U_m(t)$ and $Y_m(t)$ are individually ergodic and jointly ergodic in the correlation. It is very easy to guarantee that the time-variant system model $g(q,t)$ is uniformly stable. Therefore, in view of Theorem 2, the required 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. Note on Asymptotically Mean Stationary Random Processes

In the Introduction, we pointed out that Birkhoff's Ergodic Theorem was generalized to the so-called AMS random processes [18]. However, 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 $\mathcal{B}^{\mathbb{Z}}$ denote the Borel σ-algebra on $\mathbb{C}^{\mathbb{Z}}$. Given a random process $x \in \mathcal{P}(\Omega,\mathcal{A},\mathbb{P})$, we call any $E \in \mathcal{B}^{\mathbb{Z}}$ 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 6 [18]: We say that a random process $x \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$ is *asymptotically mean stationary* if

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{P}_x(q^{-t}E) \text{ exists for every event } E \in \mathcal{B}^{\mathbb{Z}}. \quad (12)$$

The following example shows that the AMS condition may not be preserved after uniformly stable (time-variant) linear filtering. Moreover, the produced non-AMS random process can still be ergodic in the mean and correlation.

1) Example 6: Consider an independent, stationary (and therefore AMS), Gaussian random process $x \in \mathcal{P}(\Omega, \mathcal{A}, \mathbb{P})$, such that, for all $t \in \mathbb{Z}$, the probability distribution $f_{x(t)}$ of $x(t)$ is given by

$$f_{x(t)}(\xi) = \frac{1}{\sqrt{2\pi}} e^{-(1/2)x^2}.$$

Also, consider the uniformly stable linear filter

$$h_t(\tau) = \begin{cases} 0, & t \le 0 \text{ or } \tau \le 0 \\ \frac{\sqrt{90}}{\pi^2\tau^4}, & t \in \cup_{n\in Z} \left[2^{4n}, 2^{4n+2}\right) \\ \frac{\sqrt{180}}{\pi^2\tau^4}, & t \in \cup_{n\in Z} \left[2^{4n+2}, 2^{5n}\right) \end{cases}.$$

Now, we form a new random process y by filtering x through $h_t(\tau)$. Then, for all $t \in \mathbb{Z}$, the probability distribution $f_{y(t)}$ of $y(t)$ is given by

$$f_{y(t)}(\xi) = \begin{cases} 0, & t \le 0 \\ \frac{1}{\sqrt{2\pi}} e^{-(1/2)x^2}, & t \in \cup_{n\in Z} \left[2^{4n}, 2^{4n+2}\right) \\ \frac{1}{2\sqrt{2\pi}} e^{-(1/2)(x/2)^2}, & t \in \cup_{n\in Z} \left[2^{4n+2}, 2^{5n}\right) \end{cases}.$$

Consider the event $F = \{\{c_k : k \in \mathbb{Z}\} : 1/2 \le c_0 < 1\} \in \mathcal{B}^{\mathbb{Z}}$, and let

$$q^{-t}F = \left\{ \{c_k : k \in \mathbb{Z}\} : \frac{1}{2} \le c_t < 1 \right\}.$$

Then, we get

$$\begin{aligned} &\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{P}_z(q^{-t}F) \\ &= \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} 0.6827 \times \chi_{\cup_{n\in Z}[2^{4n}, 2^{4n+2})}(t) \\ &\quad + 0.3829 \times \chi_{\cup_{n\in Z}[2^{4n+2}, 2^{5n})}(t) \end{aligned}$$

where $\chi_A(t)$ denotes the indicator function of the set A, i.e.,

$$\chi_A(t) = \begin{cases} 1, & t \in A \\ 0, & t \notin A \end{cases}.$$

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. However, according to Propositions 8 and 10, x is strongly ergodic in the mean and correlation (since it is independent and has uniformly bounded second and fourth moments). Also, h_t is clearly uniformly bounded, so, according to Propositions 7 and 9, y is strongly ergodic in the mean and correlation, and therefore, ergodic in the mean and correlation. In particular

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} y(t) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} [y(t) - \mathcal{E}\{y(t)\}] \overset{\text{w.p.1}}{=} 0.$$

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. The key properties of strong ergodicity in the mean (or correlation) are that it 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.

Appendix I
Proofs of Section IV

Proof of Proposition 1: Suppose x has BAC. Then

$$\begin{aligned} &\frac{1}{T} \sum_{t,s=1}^{T} |\langle x(s+d), x(t+d)\rangle|^2 \\ &\le \frac{1}{T} \sum_{t=1}^{T} \sum_{\tau=-(T-1)}^{T-1} |\langle x(t+d+\tau), x(t+d)\rangle|^2 \\ &\le \frac{1}{T} \sum_{t=1}^{T} \sum_{\tau=\infty}^{\infty} f^2(\tau) = \frac{1}{T} \sum_{t=1}^{T} \|f\|_2^2 = \|f\|_2^2. \end{aligned}$$

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

$$\begin{aligned} &\frac{1}{T} \sum_{t,s=1}^{T} |\langle y(s), x(t)\rangle|^2 \\ &\le \left(\frac{1}{T} \sum_{t,s=1}^{T} |\langle x(s), x(t)\rangle|^2\right)^{1/2} \left(\frac{1}{T} \sum_{t,s=1}^{T} |\langle y(s), y(t)\rangle|^2\right)^{1/2}. \end{aligned}$$

Proof: For every $t \in \mathbb{Z}$, the random variables $x(t)$ and $y(t)$ belong to $\mathcal{L}_2(\Omega, \mathcal{A}, \mathbb{P})$. Since $\mathcal{L}_2(\Omega, \mathcal{A}, \mathbb{P})$ is a Hilbert space, it has an orthonormal basis $\{e_n : n \in \mathcal{I} \subseteq N\}$. Then

$$\begin{aligned}
&\frac{1}{T}\sum_{t,s=1}^{T} |\langle y(s), x(t)\rangle|^2 \\
&\quad= \frac{1}{T}\sum_{t,s=1}^{T} \left|\sum_{n\in\mathcal{I}} \langle y(s), e_n\rangle \langle e_n, x(t)\rangle\right|^2 \\
&\quad= \frac{1}{T}\sum_{t,s=1}^{T} \sum_{n,m\in\mathcal{I}} \langle y(s), e_n\rangle \langle e_n, x(t)\rangle \langle e_m, y(s)\rangle \langle x(t), e_m\rangle \\
&\quad= \frac{1}{T}\sum_{n,m\in\mathcal{I}} \left(\sum_{t=1}^{T} \langle x(t), e_m\rangle \langle e_n, x(t)\rangle\right) \\
&\qquad\times \overline{\left(\sum_{s=1}^{T} \langle y(s), e_m\rangle \langle e_n, y(s)\rangle\right)} \\
&\quad\le \left(\frac{1}{T}\sum_{n,m\in\mathcal{I}} \left|\sum_{t=1}^{T} \langle x(t), e_m\rangle \langle e_n, x(t)\rangle\right|^2\right)^{1/2} \\
&\qquad\cdot \left(\frac{1}{T}\sum_{n,m\in\mathcal{I}} \left|\sum_{s=1}^{T} \langle y(s), e_m\rangle \langle e_n, y(s)\rangle\right|^2\right)^{1/2}
\end{aligned} \tag{13}$$

where the first equality follows from Parseval's identity [27, Theorem 4.13(e), p. 17] and the last inequality follows from Hölder's inequality [28, p. 26]. It follows that

$$\begin{aligned}
&\frac{1}{T}\sum_{n,m\in\mathcal{I}} \left|\sum_{t=1}^{T} \langle x(t), e_m\rangle \langle e_n, x(t)\rangle\right|^2 \\
&\quad= \frac{1}{T}\sum_{n,m\in\mathcal{I}} \sum_{t,s=1}^{T} \langle x(t), e_m\rangle \langle e_n, x(t)\rangle \langle e_m, x(s)\rangle \langle x(s), e_n\rangle \\
&\quad= \frac{1}{T}\sum_{t,s=1}^{T} \left(\sum_{n\in\mathcal{I}} \langle x(s), e_n\rangle \langle e_n, x(t)\rangle\right) \\
&\qquad\cdot \overline{\left(\sum_{m\in\mathcal{I}} \langle x(s), e_m\rangle \langle e_m, x(t)\rangle\right)} \\
&\quad= \frac{1}{T}\sum_{t,s=1}^{T} \left|\sum_{n\in\mathcal{I}} \langle x(s), e_n\rangle \langle e_n, x(t)\rangle\right|^2 \\
&\quad= \frac{1}{T}\sum_{t,s=1}^{T} |\langle x(s), x(t)\rangle|^2 .
\end{aligned}$$

A similar inequality above applies to the last term in (13). Substituting these inequalities into (13), we obtain the required result. ■

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

Proof: We have that

$$\begin{aligned}
\|x\|_{\mathcal{S}} &= \sup_{T\in\mathbb{N}, d\in\mathbb{Z}} \left(\frac{1}{T^2}\sum_{t,s=1}^{T} |\langle x(s+d), x(t+d)\rangle|^2\right)^{1/4} \\
&\ge \sup_{d\in\mathbb{Z}} \left(|\langle x(1+d), x(1+d)\rangle|^2\right)^{1/4} = \sup_{d\in\mathbb{Z}} \|x(d)\|_2 .
\end{aligned}$$

■

Proof of Proposition 2:

N1) It follows immediately from (6).

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

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

Let $z(t) = x(t) + y(t)$. Then

$$\begin{aligned}
\|z, z\|_{\#} &\le \|x, x\|_{\#} + \|y, y\|_{\#} + 2\|x, y\|_{\#} \\
&\le \|x, x\|_{\#} + \|y, y\|_{\#} + 2\|x, x\|_{\#}^{1/2} \|y, y\|_{\#}^{1/2} \\
&= \left(\|x, x\|_{\#}^{1/2} + \|y, y\|_{\#}^{1/2}\right)^2
\end{aligned}$$

where the first inequality follows from the triangular inequality for the 2-norm, and the second inequality follows from Lemma 1. Then, since $\|x\|_{\mathcal{S}} = \|x, x\|_{\#}^{1/2}$, it follows that

$$\|z\|_{\mathcal{S}} \le \|x\|_{\mathcal{S}} + \|y\|_{\mathcal{S}} .$$

N3) From Lemma 2, we have that $\|x\|_{\mathcal{S}} = 0$ 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 [15, Theorem 5.1.2, p. 108].

Lemma 3: Let x be a random process with

$$\sum_{T=1}^{\infty} \mathcal{E}\{|x(T)|^2\} < \infty \tag{14}$$

then

$$\lim_{T\to\infty} x(T) \overset{\text{w.p.1}}{=} 0. \tag{15}$$

Proof: From the Chebyshev inequality [15], we have

$$\mathbb{P}\{|x(T)| > \varepsilon\} \le \frac{1}{\varepsilon^2}\mathcal{E}\{|x(T)|^2\}, \ \forall \varepsilon > 0.$$

Then, (14) implies that

$$\sum_{T=1}^{\infty} \mathbb{P}\{|x(T)| > \varepsilon\} < \infty, \ \forall \varepsilon > 0. \tag{16}$$

Now, by the Borel–Cantelli Lemma [15], (16) implies that

$$\mathbb{P}\left\{\limsup_{T\to\infty} \{|x(T)| > \varepsilon\}\right\} = 0, \ \forall \varepsilon > 0$$

and (15) follows from Theorem 4.2.2 in [15]. ■

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

$$\|E_x(T)\|_2 \le \frac{K_x}{T^{1/4}}, \ \ \forall T \in \mathbb{N} \tag{17}$$

where

$$E_x(T) = \frac{1}{T}\sum_{t=1}^{T} x(t). \tag{18}$$

Then

$$\lim_{T\to\infty} E_x(T) \overset{\text{w.p.1}}{=\!=} 0. \tag{19}$$

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

$$C(L) = \frac{1}{L^4}|S(L^4)|$$
$$D(L) = \frac{1}{L^4}\max_{L^4 \le T < (L+1)^4} |S(T) - S(L^4)|.$$

We split the proof into three steps.

Step 1: We have

$$\mathcal{E}\{|C(L)|^2\} = \mathcal{E}\{|E_x(L^4)|^2\} = \left\|E_x(L^4)\right\|_2^2.$$

Then

$$\sum_{L=1}^{\infty}\mathcal{E}\{|C(L)|^2\} = \sum_{L=1}^{\infty}\left\|E_x(L^4)\right\|_2^2 \le \sum_{L=1}^{\infty}\frac{K_x^2}{L^2} < \infty.$$

Applying Lemma 3 on C, we get

$$\lim_{L\to\infty} C(L) \overset{\text{w.p.1}}{=\!=} 0. \tag{20}$$

Step 2: Let M_x be the bound of the second moments of $x(t)$. We have

$$\begin{aligned}
&\mathcal{E}\{|D(L)|^2\}\\
&\quad= \frac{1}{L^8}\mathcal{E}\left\{\max_{L^4\le T<(L+1)^4} |S(T) - S(L^4)|^2\right\}\\
&\quad= \frac{1}{L^8}\mathcal{E}\left\{\max_{L^4\le T<(L+1)^4} \sum_{t,s=L^4+1}^{T} \bar{x}(t)x(s)\right\}\\
&\quad\le \frac{1}{L^8}\mathcal{E}\left\{\max_{L^4\le T<(L+1)^4} \sum_{t,s=L^4+1}^{T} \frac{|x(t)|^2 + |x(s)|^2}{2}\right\}\\
&\quad\le \frac{1}{L^8}\max_{L^4\le T<(L+1)^4} \sum_{t,s=L^4+1}^{T} M_x^2\\
&\quad\le \frac{1}{L^8}((L+1)^4 - 1 - L^4 - 1)^2 M_x^2\\
&\quad= \frac{(4L^3 + 6L^2 + 4L - 1)^2}{L^8} M_x^2.
\end{aligned}$$

It follows that

$$\sum_{L=1}^{\infty}\mathcal{E}\{|D(L)|^2\} < \infty.$$

Applying Lemma 3 on D, we have

$$\lim_{L\to\infty} D(L) \overset{\text{w.p.1}}{=\!=} 0. \tag{21}$$

Step 3: Given any $T \in \mathbb{N}$, let $L \in \mathbb{N}$ be such that $L^4 \le T < (L+1)^4$. Then

$$\begin{aligned}
|E_x(T)| &\le \frac{1}{T}|S(L^4)| + \frac{1}{T}|S(T) - S(L^4)|\\
&\le \frac{1}{L^4}|S(L^4)| + \frac{1}{L^4}|S(T) - S(L^4)|\\
&\le C(L) + D(L).
\end{aligned}$$

Hence, (19) follows from (20) and (21). ■

Proof of Proposition 3: Let $E_x(T) = (1/T)\sum_{t=1}^{T} x(t)$. Then

$$\begin{aligned}
\|E_x(T)\|_2^2 &= \mathcal{E}\left\{\frac{1}{T^2}\sum_{t,s=1}^{T}\bar{x}(t)x(s)\right\} \le \frac{1}{T^2}\sum_{t,s=1}^{T}|\mathcal{E}\{\bar{x}(t)x(s)\}|\\
&\le \left(\frac{1}{T^2}\sum_{t,s=1}^{T}|\mathcal{E}\{\bar{x}(t)x(s)\}|^2\right)^{1/2} \le \frac{\|x\|_{\mathcal{S}}^2}{T^{1/2}}
\end{aligned}$$

where the second inequality follows from [28, 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 $i \in \mathcal{I}$

$$\|z_i\|_{\mathcal{S}} \le \|x_i\|_{\mathcal{S}} + \|y_i\|_{\mathcal{S}} \le U_X + U_Y$$

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 $\{h_t\}$ is uniformly stable, there exists $h \in \ell_1(\mathbb{Z})$ such that $|h_t(\tau)| \le h(\tau)$ for all $\tau \in \mathbb{Z}$. Then,

$$\begin{aligned}
|\langle y(s), y(t)\rangle| &= \left|\sum_{k,l=-\infty}^{\infty}\bar{h}_t(k)h_s(l)\langle x_l(s-l), x_k(t-k)\rangle\right|\\
&\le \sum_{k,l=-\infty}^{\infty} h(k)h(l)|\langle x_l(s-l), x_k(t-k)\rangle|.
\end{aligned}$$

Using the triangular inequality, it is straightforward to verify that (see the equation at the bottom of the next page). By Lemma 1, we get

$$\begin{aligned}
&\sup_{T\in\mathbb{N}, d\in\mathbb{Z}}\left(\frac{1}{T}\sum_{t,s=1}^{T}|\langle x_l(s-l+d), x_k(t-k+d)\rangle|^2\right)^{1/2}\\
&\quad\le \sup_{T\in\mathbb{N}, d\in\mathbb{Z}}\left(\frac{1}{T}\sum_{t,s=1}^{T}|\langle x_l(s-l+d), x_l(t-l+d)\rangle|^2\right)^{1/4}\\
&\quad\quad\cdot\sup_{T\in\mathbb{N}, d\in\mathbb{Z}}\left(\frac{1}{T}\sum_{t,s=1}^{T}|\langle x_k(s-k+d), x_k(t-k+d)\rangle|^2\right)^{1/4}\\
&\quad= \|x_l\|_{\mathcal{S}}\,\|x_k\|_{\mathcal{S}}.
\end{aligned}$$

Therefore

$$\begin{aligned}\|y\|_{\mathcal{S}}^2 &\leq \sum_{k,l=-\infty}^{\infty} h(k)h(l)\,\|x_l\|_{\mathcal{S}}\,\|x_k\|_{\mathcal{S}}\\ &\leq \left(\sum_{k,l=-\infty}^{\infty} h(k)h(l)\right) U_X^2 \leq \|h\|_1^2\, U_X^2\end{aligned}$$

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

$$\begin{aligned}\|y\|_{\mathcal{S}}^4 &= \sup_{T\in\mathbb{N},d\in\mathbb{Z}} \frac{1}{T}\sum_{t,s=1}^{T} |\langle x(D(s+d)), x(D(t+d))\rangle|^2\\ &\leq D \sup_{T\in\mathbb{N},d\in\mathbb{Z}} \frac{1}{DT}\sum_{t,s=1}^{DT} |\langle x(s+Dd), x(t+Dd)\rangle|^2\\ &\leq D\,\|x\|_{\mathcal{S}}^4 \leq DU_X^4.\end{aligned}$$

Upsampling: Let $y = \Upsilon_D x$. Since most of the of y are zero valued, we have that the sum $\sum_{t,s=1}^{T} |\langle y(s+d), y(t+d)\rangle|^2$ can be written as $\sum_{t,s=1}^{\hat{T}} \left|\left\langle x(s+\hat{d}), x(t+\hat{d})\right\rangle\right|^2$, where $\hat{T}$ and $\hat{d}$ are given by

$$\begin{aligned}\hat{T} &= \frac{(\lfloor T+d\rfloor_D - \lfloor d\rfloor_D)}{D}\\ \hat{d} &= \lfloor d\rfloor_D\end{aligned}$$

where $\lfloor \alpha\rfloor_D = D\lfloor \alpha/D\rfloor$ and $\lfloor \beta\rfloor$ denotes the largest integer smaller than or equal to β. Now, considering that $\hat{T} \leq T/D$, it follows that

$$\begin{aligned}\|y\|_{\mathcal{S}}^4 &= \sup_{T\in\mathbb{N},d\in\mathbb{Z}} \frac{1}{T}\sum_{t,s=1}^{T} |\langle y(s+d), y(t+d)\rangle|^2\\ &\leq \sup_{T\in\mathbb{N},d\in\mathbb{Z}} \frac{1}{D\hat{T}}\sum_{t,s=1}^{\hat{T}} \left|\left\langle x(s+\hat{d}), x(t+\hat{d})\right\rangle\right|^2\\ &= D^{-1} \sup_{\hat{T}\in\mathbb{N},\hat{d}\in\mathbb{Z}} \frac{1}{\hat{T}}\sum_{t,s=1}^{\hat{T}} \left|\left\langle x(s+\hat{d}), x(t+\hat{d})\right\rangle\right|^2\\ &\leq D^{-1}\|x\|_{\mathcal{S}}^4 \leq D^{-1}U_X^4.\end{aligned}$$

■

APPENDIX II
PROOFS OF SECTION V

Proof of Proposition 7: If $x \in \mathcal{P}_2(\Omega,\mathcal{A},\mathbb{P})$ is such that ξ_x has WBAC, then, Proposition 3 implies that (3) is satisfied, i.e., x is ergodic in the mean. Therefore, it suffices to show that the WBAC property is preserved under addition, uniformly stable linear filtering, downsampling and upsampling.

If $y \in \mathcal{P}_2(\Omega,\mathcal{A},\mathbb{P})$ is also such that ξ_y has WBAC, then, since $\xi_{x+y}(t) = \xi_x(t) + \xi_y(t)$, it follows from Proposition 4 that ξ_{x+y} also has WBAC.

Let $\{h_t \in \ell_1(\mathbb{Z}) : t \in \mathbb{Z}\}$ 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

$$\begin{aligned}\xi_y(t) &= \sum_{k=-\infty}^{\infty} h_t(k)x(t-k) - \mathcal{E}\left\{\sum_{k=-\infty}^{\infty} h_t(k)x(t-k)\right\}\\ &= \sum_{k=-\infty}^{\infty} h_t(k)(x(t-k) - \mathcal{E}\{x(t-k)\})\\ &= \sum_{k=-\infty}^{\infty} h_t(k)\xi_x(t-k)\end{aligned}$$

where the exchange of the expectation with the infinite sum in the second equality is justified by continuity of the functional $\mathcal{E} : \mathcal{L}_2(\Omega,\mathcal{A},\mathbb{P}) \to \mathbb{C}$ (note that $\sum_{k=-\infty}^{\infty} h_t(k)x(t-k) \in \mathcal{L}_2(\Omega,\mathcal{A},\mathbb{P})$ since $h_t \in \ell_1(\mathbb{Z})$ and x has uniformly bounded second moments). It follows from Proposition 5 that ξ_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 ξ_x be defined as in (8). We have that

$$|\langle \xi_x(t), \xi_x(t)\rangle| \leq M_x^2, \ \ \forall t \in \mathbb{N}. \tag{22}$$

Also, since x is white, if $\tau \neq 0$,

$$\langle \xi_x(t+\tau), \xi_x(t)\rangle = 0, \ \ \forall \tau \neq 0. \tag{23}$$

Equations (22) and (23) imply that the conditions for Proposition 1 hold for ξ_x. Hence, ξ_x has WBAC and therefore x is strongly ergodic in the mean. ■

$$\begin{aligned}\|y\|_{\mathcal{S}}^2 &= \sup_{T\in\mathbb{N},d\in\mathbb{Z}} \left(\frac{1}{T}\sum_{t,s=1}^{T} |\langle y(s+d), y(t+d)\rangle|^2\right)^{1/2}\\ &\leq \sum_{k,l=-\infty}^{\infty} h(k)h(l) \sup_{T\in\mathbb{N},d\in\mathbb{Z}} \left(\frac{1}{T}\sum_{t,s=1}^{T} |\langle x_l(s-l+d), x_k(t-k+d)\rangle|^2\right)^{1/2}.\end{aligned}$$

APPENDIX III
PROOFS OF SECTION VI

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

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

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 (9) is preserved under addition, uniformly bounded filtering, downsampling and upsampling.

The addition property follows from the fact

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

and Proposition 4.

To see the filtering property, we let $\{h_t \in \ell_1(\mathbb{Z}) : t \in \mathbb{Z}\}$ 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

$$\begin{aligned}
&\psi_{(x,z,\alpha,\beta,a,b)}(t)\\
&\quad= \bar{x}(\alpha t + a) \sum_{k=-\infty}^{\infty} h_{\beta t+b}(k) y(\beta t + b - k)\\
&\qquad - \mathcal{E}\left\{\bar{x}(\alpha t + a) \sum_{k=-\infty}^{\infty} h_{\beta t+b}(k) y(\beta t + b - k)\right\}\\
&\quad= \sum_{k=-\infty}^{\infty} h_{\beta t+b}(k) [\bar{x}(\alpha t + a) y(\beta t + b - k)\\
&\qquad - \mathcal{E}\{\bar{x}(\alpha t + a) y(\beta t + b - k)\}]\\
&\quad= \sum_{k=-\infty}^{\infty} h_{\beta t+b}(k) \psi_{(x,y,\alpha,\beta,a,b-k)}(t)\\
&\quad= \sum_{k=-\infty}^{\infty} h_{\beta t+b}(k) \psi_{(x,y,\alpha,\beta,a+k\alpha,b+k(\beta-1))}(t-k).
\end{aligned}$$

The exchanging of the expectation with the infinite sum in the second equality above is justified by continuity of the functional $\mathcal{E} : \mathcal{L}_2(\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{C}$ (note that $\sum_{k=-\infty}^{\infty} h_t(k) y(t-k) \in \mathcal{L}_2(\Omega, \mathcal{A}, \mathbb{P})$ since $h_t \in \ell_1(\mathbb{Z})$ and y has uniformly bounded second moments). Using Proposition 5, we know that the set of random processes $\Psi_{z,x}$ 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 α, β, 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 $z(t) = \Upsilon_D y(t)$. Let $\mathcal{T} = \{t_i \in \mathbb{Z}^2 : i \in \mathcal{I}\}$ and $\mathcal{K} = \{k_i \in \mathbb{Z}^2 : i \in \mathcal{I}\}$ be two sets of integers that satisfy

$$\beta t_i + b = k_i D. \tag{24}$$

We consider two cases. In the first case, there is no integer solution of (24), i.e., $\mathcal{I} = \emptyset$. It follows that

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

Therefore, $\psi_{(x,z,\alpha,\beta,a,b)}$ has WBAC. In the second case, (24) admits an integer solution. In this case, $\mathcal{I} = \mathbb{Z}$. Therefore, we can write

$$t_i = \gamma i + c \quad \text{and } k_i = \delta i + d$$

for some $\gamma, \delta \in \mathbb{N}$ and $c, d \in \mathbb{Z}$. It can be verified that

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

or, equivalently,

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

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

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

$$|\langle \psi_{(x,x,\alpha,\beta,a,b)}(t+\tau), \psi_{(x,x,\alpha,\beta,a,b)}(t) \rangle| \le 2M_x^4. \tag{25}$$

(The proof of (25) requires the use of Hölder's inequality [15, p. 50, eq. (18)] and Lyapunov's inequality [15, p. 50, eq. (21)].) Also, it is straightforward to verify that, for every $\alpha, \beta \in \mathbb{N}$ and $a, b \in \mathbb{Z}$, there exists $\hat{\tau} \in \mathbb{N}$ such that, if $|\tau| > \hat{\tau}$, then

$$\max\{|\alpha t + a|, |\beta t + b|\} < \min\{|\alpha(t+\tau) + a|, |\beta(t+\tau) + b|\}.$$

Using [15, p. 54, Theorem 3.3.2], the random variables $x(\alpha t + a)\bar{x}(\beta t + b)$ and $\bar{x}(\alpha(t+\tau) + a) x(\beta(t+\tau) + b)$ are independent when $|\tau| > \hat{\tau}$. It follows that

$$\langle \psi_{(x,x,\alpha,\beta,a,b)}(t+\tau), \psi_{(x,x,\alpha,\beta,a,b)}(t) \rangle = 0, \ \forall |\tau| > \hat{\tau}. \tag{26}$$

Since x has uniformly bounded fourth moments, $x \in \mathcal{P}_4(\Omega, \mathcal{A}, \mathbb{P}) \subset \mathcal{P}_2(\Omega, \mathcal{A}, \mathbb{P})$. Therefore, (25) and (26) imply that the conditions for Proposition 1 hold. Subsequently, $\Psi_{(x,x)}$ has UWBAC, and therefore x is strongly ergodic in the correlation. ■

Proof of Theorem 2: It follows from Proposition 10 that every member in $\{u_m : 1 \le m \le M\}$ is strongly ergodic in the correlation and every pair in $\{u_m : 1 \le m \le M\}$ jointly strongly ergodic in the correlation. We now claim that all the operations on the signals u_m, $m = 1, \ldots, M$, as mentioned in the theorem, can be decomposed into those in SEC2)–SEC5). For example, if w is downsampled 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 $y = u_n$ 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. ∎

ACKNOWLEDGMENT

The authors would like to thank Prof. M. Vetterli from the School of Computer and Communication Sciences, Ecole Polytechnique Fédérale de Lausanne, for suggesting the publication of the present material.

REFERENCES

[1] P. Vaidyanathan, *Multirate Systems and Filterbanks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[2] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1983.

[3] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[4] A. Gilloire and M. Vetterlli, "Adaptive filtering in subbands with critical sampling: Analysis, experiments, and application to acoustic echo cancellation," *IEEE Trans. Signal Process.*, vol. 40, no. 8, pp. 1862–1875, Aug. 1992.

[5] Y. Lu and J. Morris, "Gabor expansion for adaptive echo cancellation," *IEEE Signal Process. Mag.*, vol. 16, no. 2, pp. 68–80, Mar. 1999.

[6] D. Marelli and M. Fu, "Asymptotic properties of subband identification," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 142–154, Jan. 2004.

[7] ——, "Performance analysis for subband identification," *IEEE Trans. Signal Process.*, vol. 51, no. 12, pp. 3128–3142, Dec. 2003.

[8] V. Sathe and P. Vaidyanathan, "Effects of multirate systems on the statistical properties of random signals," *IEEE Trans. Signal Process.*, vol. 41, no. 1, pp. 131–146, Jan. 1993.

[9] S. Akkarakaran and P. Vaidyanathan, "Bifrequency and bispectrum maps: A new look at multirate systems with stochastic inputs," *IEEE Trans. Signal Process.*, vol. 48, no. 3, pp. 723–736, Mar. 2000.

[10] R. M. Gray and L. D. Davisson, *An Introduction to Statistical Signal Processing*. Cambridge, U.K.: Cambridge Univ. Press, 2004 [Online]. Available: http://www-ee.stanford.edu/~gray/sp.pdf

[11] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, ser. McGraw-Hill Series in Electrical Engineering. Communications and Information Theory, 2nd ed. New York: McGraw-Hill., 1984.

[12] P. Billingsley, *Ergodic Theory and Information*. New York: Wiley, 1965.

[13] P. Walters, *An Introduction to Ergodic Theory*. New York: Springer-Verlag, 1982, vol. 79, Graduate Texts in Mathematics.

[14] J. L. Doob, *Stochastic Processes*, ser. Wiley Classics Library. New York: Wiley, 1990.

[15] K. L. Chung, *A Course in Probability Theory*, 3rd ed. San Diego, CA: Academic Press, 2001.

[16] G. Birkhoff, "Proof of the ergodic theorem," *Proc. Nat. Acad. Sci. U.S.A.*, vol. 17, no. 12, pp. 656–660, Dec. 1931.

[17] R. M. Gray and J. C. Kieffer, "Asymptotically mean stationary measures," *Ann. Probab.*, vol. 8, no. 5, pp. 962–973, 1980.

[18] R. M. Gray, *Probability, Random Processes, and Ergodic Properties*. New York: Springer-Verlag, 1988 [Online]. Available: http://www-ee.stanford.edu/~gray/arp.pdf

[19] V. Anantharam and T. Konstantopoulos, "Stationary solutions of stochastic recursions describing discrete event systems," *Stoch. Process. Appl.*, vol. 68, no. 2, pp. 181–194, 1997.

[20] Y. Ephraim and N. Merhav, "Hidden Markov processes," *IEEE Trans. Inf. Theory*, vol. 48, pp. 1518–1569, 2002.

[21] B. Ninness, "Strong laws of large numbers under weak assumptions with application," *IEEE Trans. Autom. Control*, vol. 45, no. 11, pp. 2117–2122, 2000.

[22] H. Li and Q. Cheng, "Almost sure convergence analysis of mixed time averages and kth-order cyclic statistics," *IEEE Trans. Inf. Theory*, vol. 43, pp. 1265–1268, 1997.

[23] P. Révész, *The Laws of Large Numbers*, ser. Probability and Mathematical Statistics. New York: Academic, 1968, vol. 4.

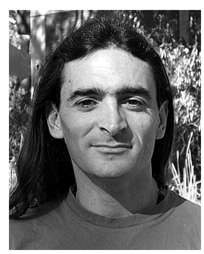
[24] H. Cramér and M. R. Leadbetter, *Stationary and Related Stochastic Processes. Sample Function Properties and their Applications*. New York: Wiley, 1967.

[25] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1999.

[26] R. M. Gray, *Entropy and Information Theory*. New York: Springer-Verlag, 1990.

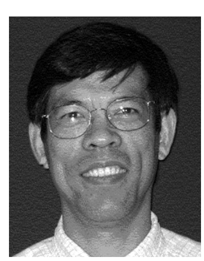
[27] J. B. Conway, *A Course in Functional Analysis*, ser. Graduate Texts in Mathematics, 2nd ed. New York: Springer-Verlag, 1990, vol. 96.

[28] G. Hardy, J. Littlewood, and G. Pólya, *Inequalities*, 2nd ed. London, U.K.: Cambridge Univ. Press, 1959.



Damián Marelli received the Bachelor's degree in electronics engineering from the Universidad Nacional de Rosario, Argentina, in 1995 and the Ph.D. degree in electrical engineering and a Bachelor's (Hon.) degree in mathematics from the University of Newcastle, Australia, in 2003.

He worked as a Software Engineer at Tesis Ingeniería Informática S.R.L., Rosario, Argentina, from 1995 to 1996 and at BLC S.A., Rosario, Argentina, in 1998. In 1997, he held a teaching and Research Assistant position at the Universidad Nacional de Rosario, Argentina. In 2003, he held a Research Associate position at the School of Electrical Engineering and Computer Science at the University of Newcastle, Australia. In 2004, he held a postdoctoral research fellowship at the Laboratoire d'Analyse Topologie et Probabilités, CNRS/Université de Provence, France, and during 2005, he held a combined teaching and research position at the Ecole Généraliste d'lngénieurs de Marseille, France. Since 2006, he has held a Research Academic position at the ARC Centre for Complex Dynamic Systems and Control at the University of Newcastle, Australia. His main research interests include multirate signal processing, time-frequency analysis, system identification, and statistical signal processing.



Minyue Fu (S'84–M'87–SM'94–F'04) received the Bachelor's degree from the University of Science and Technology of China, Hefei, in 1982 and the M.S. and Ph.D. degrees from the University of Wisconsin-Madison in 1983 and 1987, respectively, all in electrical engineering.

From 1983 to 1987, he held a teaching assistantship and a research assistantship at the University of Wisconsin–Madison. He worked as a Computer Engineering Consultant at Nicolet Instruments, Inc., Madison, WI, during 1987. From 1987 to 1989, he served as an Assistant Professor in the Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI. During summer 1989, he was employed by the Universite Catholoque de Louvain, Belgium, as a Maitre de Conferences Invited. He joined the Department of Electrical and Computer Engineering, the University of Newcastle, Australia, in 1989. He served as the Head of Department of Electrical and Computer Engineering from 1998 to 2002. Currently, he is a Chair Professor in Electrical Engineering and the Head of School of Electrical Engineering and Computer Science. In addition, he was a Visiting Associate Professor at the University of Iowa in 1995–1996 and a Senior Fellow/Visiting Professor at Nanyang Technological University, Singapore, in 2002. His main research interests include control systems, signal processing, and communications.

Dr. Fu has served as an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL for four years. Currently, he is an Area Editor for the *Journal of Optimization and Engineering* and Associate Editor for *Automatica*.