On Locally Dyadic Stationary Processes

Theodoros Moysiadis
Institute of Applied Biosciences, Centre for Research and Technology Hellas

and

Konstantinos Fokianos
Department of Mathematics & Statistics, University of Cyprus

April 24, 2015

Abstract
We introduce the concept of local dyadic stationarity, to account for non-stationary time series, within the framework of Walsh-Fourier analysis. We define and study the time varying dyadic ARMA models (tvDARMA). It is proven that the general tvDARMA process can be expressed both as a tvDMA and a tvDAR process. The local dyadic stationary behavior is established, under mild assumptions, for all three classes.

Keywords: dyadic stationarity, local stationarity, spectral density, stationarity, Walsh functions, Walsh-Fourier analysis.
1 Introduction

The concept of stationarity is crucial in the statistical theory of time series analysis and especially for the development of asymptotic theory. However, the assumption of stationarity is very often unrealistic in practice. For example, a time series can display significant changes through time and therefore stationarity is a questionable assumption. One of the most important consequences, is that attempts to develop asymptotic results are, in general, groundless, since future information of the process does not necessarily contain any information regarding the present of the process. In addition, there is no natural generalization of stationarity to non-stationarity, since non-stationary processes might exhibit trend or and periodicity and other types of non-standard behavior.

Priestley (1965) considered non-stationary processes whose characteristics are changing slowly over time and developed the theory of evolutionary spectra (see Priestley (1966, 1981, 1988)). However, such an approach makes it difficult to obtain asymptotic results, which are needed for developing estimation theory. In order to apply standard asymptotic theory for non-stationary processes, Dahlhaus, in a series of contributions, introduced an appropriate theoretical framework, based on the concept of local stationarity (see for example Dahlhaus (1996c, 1997, 2000)). The definition of local stationarity is based on the existence of a time varying spectral representation (see for example Dahlhaus (1996c)). Dahlhaus (2012) gives an excellent and detailed overview of the theory of locally stationary processes. A comparison between the methodology of Dahlhaus and Priestley is given in Dahlhaus (1996a). Some other works related to locally stationary time series include the works by Granger and Hatanaka (1964), Rao and Tong (1972), Tjøstheim (1976), Martin (1981), Melard and Schutter (1989), Neumann and Von Sachs (1997), Nason et al. (2000), Sakiyama and Taniguchi (2004) and Davis et al. (2006), among others.

The main goal of this contribution is to utilize the idea of local stationarity, in the sense of the above mentioned papers, to study the spectral behavior of time series, based on the system of Walsh functions. Walsh functions led to the development of Walsh-Fourier (square wave) analysis,
just like the sinusoidal functions led to Fourier (trigonometric) analysis. The motivation behind Walsh-Fourier analysis was the need to approximate stationary time series, which display square waveforms with abrupt switches (e.g. in communications and engineering). We introduce the concept of local stationarity within the orthogonal system of Walsh functions to account for such phenomena that exhibit, in addition, non-stationary behavior. We study important general classes of time series, similar in concept with the time varying ARMA (tvARMA) process- see Dahlhaus (2012).

The Walsh functions were introduced by Walsh (1923), take only two values, $+1$ and $-1$, and have similar behavior and properties with the trigonometric functions (although they are not periodic). The introduction of Walsh functions has been followed by a series of papers, related to their mathematical properties and generalizations (Paley (1932a,b), Fine (1949, 1950, 1957), Chrestenson (1955) and Selfridge (1955), among others), which provided the theoretical framework for various applications, see e.g. Harmuth (1969, 1977), Grove (1983), Maqusi (1981), Beauchamp (1984) Stoffer (1991) and Abbasi et al. (2012). Stoffer (1991) gives an excellent account of the history of Walsh functions and a comparison between Walsh and Fourier analysis.

The statistical analysis via Walsh functions has been based both on real and dyadic time. We refer to the real time in the typical sense for a time series. The dyadic time is based on the concept of dyadic addition (see Subsection 2.1) and for the time points $m, n$, the real time sum $m + n$ is now replaced by the dyadic sum $m \oplus n$. Morettin (1981) reviewed the work on Walsh spectral analysis in both considerations. Walsh-Fourier analysis of real time stationary processes has been studied by Kohn (1980a,b), Morettin (1983) and Stoffer (1985, 1987, 1990). The dyadic time stationarity is defined in respect to the real time stationarity as in Subsection 2.2 see also Morettin (1974a) and Nagai (1977), among others. Further references related to the Walsh-Fourier analysis of dyadic stationary processes are Morettin (1974b, 1978, 1981), Nagai (1980), Nagai and Taniguchi (1987) and Taniguchi et al. (1989). In particular, Morettin (1974b, 1978) studied the finite Walsh trans-
form, considered the Walsh periodogram as an estimator of the Walsh spectrum and studied its theoretical properties. Nagai (1977) proved that a dyadic stationary process has always unique spectral representation in terms of the system of Walsh functions and studied the dyadic linear process (see also Morettin (1974b)). Nagai (1980) also studied dyadic autoregressive and moving average processes and their relation.

In this article, we introduce the concept of local dyadic stationarity and discuss the advantages and the perspectives of such consideration in the framework of Walsh functions. In Section 2 of this article, we recall some definitions and review some fundamental results for dyadic stationary processes. In Section 3, we introduce the concept of local dyadic stationarity and study the time varying dyadic moving average process (tvDMA) process. In Section 4, we define the general class of time varying dyadic autoregressive moving average process (tvDARMA) processes and show that they exhibit locally dyadic stationarity. The article concludes with several remarks concerning further research in this topic.

2 Preliminaries

2.1 Dyadic addition

We recall the definition of dyadic addition and of a dyadic process following Kohn (1980a). Consider \( m \) and \( n \) to be non-negative integers that have the following dyadic expansions

\[
m = \sum_{k=0}^{f} m_k 2^k, \quad n = \sum_{k=0}^{f} n_k 2^k, \quad \text{where} \quad m_k, n_k \in \{0, 1\}.
\]

Then, the dyadic sum \( m \oplus n \) is defined as

\[
m \oplus n = \sum_{k=0}^{f} |m_k - n_k| 2^k.
\]
Consider now $x$ and $y$ to be real numbers that belong to the interval $I = [0, 1)$. We write

$$x = \sum_{k=1}^{\infty} x_k 2^{-k}, \quad y = \sum_{k=1}^{\infty} y_k 2^{-k}, \quad \text{where } x_k, y_k \in \{0, 1\}.$$  

Each of the above representations is not unique in general. We follow the convention that if, e.g. $x$, can be written both through a finite representation and an infinite one, we choose the finite, or equivalently the representation where $x_k = 0, \forall k > k_0$. Then, the dyadic sum $x \oplus y$ is defined as

$$x \oplus y = \sum_{k=1}^{\infty} |x_k - y_k| 2^{-k}.$$ 

Recall that the $j^{th}$ Rademacher function is $\phi_k(x) = (-1)^{x_{k+1}}$, $\forall x \in I, \forall k \geq 0$. Then the system of Walsh functions, $W(n, x)$, $n = 0, 1, 2, \ldots, x \in I$, is defined as follows. If $n = 0$, set $W(0, x) = 1, \forall x \in I$. If $n > 0$, expand $n$ dyadically, i.e. $n = 2^{n_{[i]}} + 2^{n_{[2]}} + \ldots + 2^{n_{[\nu]}}$, where $n_{[i]}, i = 1, 2, \ldots, \nu$ corresponds to the $i^{th}$ non-zero term of the expansion. Then

$$W(n, x) = \begin{cases} 
1, & n = 0, \\
\phi_{n_{[i]}}(x) \phi_{n_{[2]}}(x) \cdots \phi_{n_{[\nu]}}(x), & n > 0,
\end{cases} \quad \forall x \in I.$$ 

We mention briefly some characteristic properties of the Walsh functions.

(i) The system of Walsh functions is orthonormal in $I$, that is

$$\int_0^1 W(n, x) W(m, x) dx = \begin{cases} 
1 & \text{for } n = m, \\
0 & \text{for } n \neq m,
\end{cases}$$

and constitutes a complete set. Hence, for $x \in I$, if $f(x)$ is a square integrable function, it can be expanded in a Walsh-Fourier series $f(x) = \sum_{n=0}^{\infty} c_n W(n, x)$, with $c_n = \int_0^1 f(x) W(n, x) dx$.

(ii) $\forall n, m \in \mathbb{N}, x, y \in I$, $W(n, x) W(m, x) = W(n \oplus m, x)$ and $W(n, x) W(n, y) = W(n, x \oplus y)$. 

5
2.2 Dyadic stationarity

We call a stochastic process \( \{ X_t \}_{t \in \mathbb{N}} \) dyadic stationary if it has constant mean and finite second moments and its covariance function

\[
R(n, m) = \text{cov}(X_n, X_m) = E[(X_n - E[X_n])(X_m - E[X_m])], \quad n, m \in \mathbb{N},
\]

is invariant under the shift by the dyadic addition \( \oplus \). Equivalently \( \forall n, m, k \in \mathbb{N} \)

\[
R(n, m) = R(n \oplus k, m \oplus k) = R(n \oplus m).
\]

In the following assume that \( E[X_t] = 0, \ E[X_t^2] = 1, \ \forall t \in \mathbb{N}. \) We recall some important results about dyadic stationary processes.

A dyadic stationary process \( \{ X_t \}_{t \in \mathbb{N}} \) has a dyadic spectral representation given by (Nagai (1977))

\[
X_t = \int_0^1 W(t, x) dZ_X(x), \quad t \in \mathbb{N},
\]

where \( \{ Z_X(x) \}_{x \in [0,1)} \) is a real random process with orthogonal increments, such that

\[
E[(dZ_X(x))^2] = dG_X(x), \quad x \in I.
\]

The function \( G(x) \), defined on \( I \), is a unique distribution, which is called dyadic spectral distribution of the process \( \{ X_t \}_{t \in \mathbb{N}} \). In addition

\[
R(n, m) = \int_0^1 W(n \oplus m, x) dG_X(x).
\]

**Example 1.** A simple example of a dyadic stationary process is a sequence \( \{ \varepsilon_t \}_{t \in \mathbb{N}} \) of independent random variables with \( E(\varepsilon_t) = 0 \) and \( E(\varepsilon_t^2) = \sigma^2, \ \forall t \in \mathbb{N}. \) It is straightforward to show that its covariance function is

\[
R(n, n \oplus \tau) = R(\tau) = E(\varepsilon_n \varepsilon_{n \oplus \tau}) = \begin{cases} \sigma^2, & \text{if } \tau = 0, \\ 0, & \text{if } \tau \neq 0. \end{cases}
\]
Since the sequence \( \{\varepsilon_t\}_{t \in \mathbb{N}} \) is dyadic stationary, it has a dyadic spectral representation of the form
\[
\varepsilon_t = \int_0^1 W(t, x) dZ_\varepsilon(x), \quad t \in \mathbb{N},
\]
with
\[
E[(dZ_\varepsilon(x))^2] = dG_\varepsilon(x) = \sigma^2 dx, \quad x \in I.
\]

A stochastic process \( \{X_t\}_{t \in \mathbb{N}} \) is a linear dyadic process if it can be represented as (Morettin (1974b))
\[
X_t = \sum_{k=0}^{\infty} a_k \varepsilon_{t \oplus k},
\tag{1}
\]
where \( \{\varepsilon_t\}_{t \in \mathbb{N}} \) is the sequence of i.i.d. variables as in Example\[1\] and \( \{a_k\}_{k \in \mathbb{N}} \) are real numbers, which satisfy \( \sum_{k=0}^{\infty} a_k^2 < \infty \). This definition is similar in spirit to the definition of the general linear process model in Priestley (1981, Ch.6, p.415).

Nagai (1977) shows that a linear dyadic process of the form (1) is dyadic stationary (see also Morettin (1974b)), since its covariance function \( R(n, m) \) can be written
\[
R(n, m) = \int_0^1 W(n \oplus m, x) \left( \sigma \sum_{k=0}^{\infty} a_k W(k, x) \right)^2 dx.
\tag{2}
\]
In addition, it has an absolutely continuous dyadic spectral distribution function and its dyadic spectral density function has the form
\[
g(x) = \sigma^2 \left( \sum_{k=0}^{\infty} a_k W(k, x) \right)^2 \sigma^2 A^2(x),
\tag{3}
\]
where \( A(x) = \sum_{k=0}^{\infty} a_k W(k, x) \).

Setting in (1), \( a_q \neq 0 \) and \( a_k = 0, \forall k > q \), \( X_t \) is said to be a dyadic moving average process of order \( q \) (DMA\(q\)). In general, the process \( X_t \) defined by (1) is called a DMA(\( \infty \)) process.
3 Local Dyadic Stationarity

Suppose now that the function $A(\cdot)$ in (3) is time dependent, i.e. $A_t(\cdot)$. We rescale $A_t(\cdot)$ from the axis of the first $T$ non negative integers ($t = 1, 2, \ldots, T$) to the unit interval $I$, where $T$ denotes the sample size. The reason for this rescaling will be clear later on. The rescaled form of $A_t(\cdot)$ is denoted by $A_{t,T}(\cdot)$ and $X_t$ is now expressed as $X_{t,T}$. We give a general definition regarding local dyadic stationarity for a process $X_{t,T}$, in the spirit of Dahlhaus (e.g. Dahlhaus (1996c, 1997)).

**Definition 1.** A sequence of stochastic processes \{\(X_{t,T}, \ t = 1, 2, \ldots, T\)\} is called locally dyadic stationary with transfer function $A_{t,T}(\cdot)$ and trend function $\mu(\cdot)$ if there exists a representation

$$X_{t,T} = \mu\left(\frac{t}{T}\right) + \int_0^1 W(t, x)A_{t,T}(x)dU(x),$$

where the following hold:
(i) $U(x)$ is a stochastic real valued process on $I$ and

$$\text{cum}\{dU(x_1), \ldots, dU(x_k)\} = \eta\left(\sum_{j=1}^k x_j\right) g_k(x_1, \ldots, x_{k-1})dx_1, \ldots, dx_k,$$

where $\text{cum}\{\ldots\}$ denotes the cumulant of the $k^{th}$ order, $g_1 = 0$, $g_2(x_1) = 1$, $|g_k(x_1, \ldots, x_{k-1})| \leq C_k$, \(\forall k\), \(\eta(x) = \sum_{j=-\infty}^{\infty} \delta(x \oplus j)\) is the "periodic" extension of the Dirac delta function.
(ii) There exists a constant $K$ and a periodic function $A : [0, 1] \times \mathbb{R} \to \mathbb{R}$, with $A(u, -x) = A(u, x)$ and

$$\sup_{t,x} \left|A_{t,T}(x) - A\left(\frac{t}{T}, x\right)\right| \leq \frac{K}{T}, \ \forall T.$$  

The functions $A(u, x)$ and $\mu(u)$ are assumed to be continuous with respect to $u$.

Note that the continuity of $A(u, x)$ and $\mu(u)$ in $u$ is required for the process $X_{t,T}$ to exhibit locally dyadic stationary behavior.
Example 2. Suppose $Y_t$ is a dyadic stationary process with dyadic spectral representation

$$Y_t = \int_0^1 W(t, x) A(x) dZ(x).$$

Define $X_{t,T}$ by

$$X_{t,T} = \mu \left( \frac{t}{T} \right) + \sigma \left( \frac{t}{T} \right) Y_t.$$

Then

$$X_{t,T} = \mu \left( \frac{t}{T} \right) + \int_0^1 W(t, x) A_{t,T}(x) dZ(x),$$

where $A_{t,T}(x) = A(t/T, x) = \sigma(t/T) A(x)$. Hence $X_{t,T}$ is locally a dyadic stationary process.

Consider now the process $\{X_t\}_{t \in \mathbb{N}}$

$$X_t = \sum_{k=0}^{\infty} a_{k,t} \varepsilon_{t \oplus k},$$

(6)

where $\{\varepsilon_t\}_{t \in \mathbb{N}}$ is an i.i.d. sequence and $\{a_{k,t}\}_{k,t \in \mathbb{N}}$ is a time-dependent process of real numbers such that $\forall t, \sum_{k=0}^{\infty} a_{k,t}^2 < \infty$. We call this process a time varying dyadic moving average process of infinite order (tvDMA(\infty)). If we set in (6) $a_{q,t} \neq 0$ and $a_{k,t} = 0, \forall k > q$, then we call $X_t$ a time varying dyadic moving average process of order $q$ (tvDMA(q)). Rescale now the parameter curves $a_{k,t}$ to the unit interval $I$. The non-rescaled parameter curves $a_{k,t}$ are now written as $a_{k,t,T}$ and equation (6) becomes now

$$X_{t,T} = \sum_{k=0}^{\infty} a_{k,t,T} \varepsilon_{t \oplus k}.$$

(7)

We assume that there exist functions $a_k(\cdot) : [0,1] \rightarrow \mathbb{R}$, which satisfy some regularity conditions (see Remark 1). Roughly speaking, we assume that $a_{k,t,T} \approx a_k(t/T)$. The reasons for the rescaling are described in detail, e.g. in Dahlhaus (2012, Sec.2). Briefly, suppose that we choose $a_{k,t,T}$ to be polynomials of $t$. Then, as $t \rightarrow \infty$, $a_{k,t,T} \rightarrow \infty$ as well, which violates the condition $\sum_{k=0}^{\infty} a_{k,t}^2 < \infty$. In addition, this rescaling enables us to impose smoothing conditions through the continuity of the functions $a_k(t/T)$, ensuring that the process exhibits locally dyadic stationary...
behavior. Furthermore, the number of observations within the neighborhood of a fixed point \( u_o \in [0, 1] \) increases as \( T \to \infty \) and makes reasonable enough to develop and apply locally asymptotic results. Suppose that the process \( X_{t,T} \) defined by (7) is written as

\[
X_{t,T} = \sum_{k=0}^{\infty} a_k \left( \frac{t}{T} \right) \varepsilon_{t\oplus k}.
\]

We assume that \( a_k(u) = a_k(0) \) for \( u < 0 \) and \( a_k(u) = a_k(1) \) for \( u > 1 \) and that the functions \( a_k(\cdot) \) satisfy some smoothness conditions. Consider now a fixed point \( u_0 = t_0/T \) and its neighborhood \([u_0 - \epsilon/2, u_0 + \epsilon/2]\). If the length of this segment is reasonably small, the process \( X_{t,T} \) can be approximated by the process \( \tilde{X}_t(u_0) \), which is defined as

\[
\tilde{X}_t(u_0) = \sum_{k=0}^{\infty} a_k(u_0) \varepsilon_{t\oplus k},
\]

where \( a_k(u_0) \) are constants, with \( u_0 \) indicating their dependence from the fixed point \( u_0 \) (see also Dahlhaus (2012)). \( \tilde{X}_t(u_0) \) is dyadic stationary. Indeed, we can write

\[
\tilde{X}_t(u_0) = \int_0^1 W(t, x) \left( \sum_{k=0}^{\infty} a_k(u_0) W(k, x) \right) dZ_{\varepsilon}(x) = \int_0^1 W(t, x) A(u_0, x) dZ_{\varepsilon}(x), \quad (8)
\]

where \( A(u_0, x) = \sum_{k=0}^{\infty} a_k(u_0) W(k, x) \).

As \( T \to \infty \), the number of observations of \( X_{t,T} \) in the neighborhood of \( u_0 \) increases and we may consider locally the application for \( X_{t,T} \) of known asymptotic results for dyadic stationary processes. From equations (2) and (3) and since \( u_0 \) in \( a_k(u_0) \) just indicates the dependence from the fixed point \( u_0 \), we have that \( \tilde{X}_t(u_0) \) has covariance function given by

\[
R(u_0, n \oplus m) = \int_0^1 W(n \oplus m, x) \left( \sigma \sum_{k=0}^{\infty} a_k(u_0) W(k, x) \right)^2 dx,
\]

and a unique dyadic spectral density function given by

\[
g(u_0, x) = \left( \sigma \sum_{k=0}^{\infty} a_k(u_0) W(k, x) \right)^2 = \sigma^2 A^2 (u_0, x) . \quad (9)
\]
We can show that for \( \{u = t/T : |t/T - u_0| \leq \epsilon/2\} \) it holds that \( |X_{t,T} - \tilde{X}_t(u_0)| = O_p(1/T) \), see Theorem 1. Therefore, we can say that \( X_{t,T} \) has locally the same covariance and dyadic spectral density function as \( \tilde{X}_t(u_0) \) and therefore exhibits locally dyadic stationary behavior.

We show in Theorem 1 that the tvDMA(\( \infty \)) process \( X_{t,T} \) in (7) is locally dyadic stationary, since it has a time varying spectral representation as in (4). Indeed, we have

\[
X_{t,T} = \sum_{k=0}^{\infty} \left( a_{k,t,T} \int_{0}^{1} W(t \oplus k, x)dZ_\varepsilon(x) \right) = \int_{0}^{1} W(t, x)A_{t,T}(x)dZ_\varepsilon(x),
\]

(10)

where \( A_{t,T}(x) = \sum_{k=0}^{\infty} a_{k,t,T}W(k, x) \) is the time varying transfer function.

**Theorem 1.** For a sequence of stochastic processes \( \{X_{t,T}, t = 1, 2, \ldots, T\} \) which satisfy a representation of the form (4) where \( A_{t,T}(x) \) is the time varying transfer function and \( \mu \) is the trend and under the assumptions of Definition 1 it holds that

\[
|X_{t,T} - \tilde{X}_t(u_0)| = O_p(1/T),
\]

*Proof. Set \( \mu(t/T) = 0 \) for ease of presentation. From equations (8) and (10) we have that

\[
|X_{t,T} - \tilde{X}_t(u_0)| = \left| \int_{0}^{1} W(t, x)A_{t,T}(x)dZ_\varepsilon(x) \right| - \int_{0}^{1} W(t, x)A(u_0, x)dZ_\varepsilon(x) \right| 
\leq \int_{0}^{1} |W(t, x)| \cdot |A_{t,T}(x) - A(u_0, x)| dZ_\varepsilon(x) 
\]

\[
= \int_{0}^{1} |A_{t,T}(x) - A(u_0, x)| dZ_\varepsilon(x),
\]

(11)

since \( W(t, x) \in \{-1, 1\} \). In addition

\[
|A_{t,T}(x) - A(u_0, x)| \leq |A_{t,T}(x) - A\left(\frac{t}{T}, x\right)| + |A\left(\frac{t}{T}, x\right) - A(u_0, x)| 
\leq \frac{K}{T} + |A(u, x) - A(u_0, x)|,
\]

(12)
Figure 1: The spectral density function for the tvMA(1) process (left) and the tvDMA(1) process (right).

from (5) in assumption (ii) of Definition 1. However, the same assumption states that \( A(u, x) \) is continuous. Therefore, since \( \{u = t/T : |t/T - u_0| \leq \epsilon/2\} \) and for any \( \epsilon' > 0 \) we can choose \( \epsilon > 0 \) to be such that \( |A(u, x) - A(u_0, x)| < \epsilon' \),

(12) becomes

\[
|A_{t,T}(x) - A(u_0, x)| \leq \frac{K^*}{T} 
\]

for some positive constant \( K^* \). Finally, from (11) and (13), we obtain that

\[
E|X_{t,T} - \tilde{X}_t(u_0)| = O(1/T). 
\]

and hence we have the desired result.

Remark 1. Equation (5) implies a similar assumption for the \( \sup_{t,x} \left| a_{k,t,T} - a_k \left( \frac{t}{T} \right) \right| \) and the above discussion still holds.
Example 3. If we consider now the infinite time varying MA (tvMA(∞)) representation \( X_{t,T} = \sum_{k=0}^{\infty} a_{k,t,T} \varepsilon_{t-k} \), in the real time, then its spectral density function is given by

\[
 f(u, \lambda) = \left( \frac{\sigma^2}{2\pi} \right) \left( \sum_{k=0}^{\infty} a_k(u) \exp(-i\lambda k) \right)^2.
\]

We compare the behavior of functions \( g(u_0, x) \) (in (9)) and \( f(u, \lambda) \) for the same order of the respective processes and for the same representation of the time varying coefficients \( a_k(u) \) (set \( \sigma^2 = 1 \)). Figure [1] shows the spectral density function of the tvMA(1) and the tvDMA(1) processes. We set \( a_0(u) = -1.8 \cos(1.5 - \cos(4\pi u)), a_1(u) = 0.81 \). Figure [2] shows the spectral density function of the tvMA(2) and the tvDMA(2) processes. In this case we set \( a_0(u) = 1.2 \cos(2\pi u), a_1(u) = 2 \cos(1.5 - \cos(8\pi u)), a_2(u) = u \). Both figures reveal the differences between real and dyadic stationarity. The square waveform of Walsh functions allows a more oscillatory behavior of the dyadic spectral density function.

4 tvDARMA processes

It is well known that autoregressive, moving average, and ARMA models can be regarded as special cases of the general linear process. Nagai (1980) shows that a dyadic autoregressive process of finite order is always inverted into a dyadic moving average process of finite order, and vice versa. We obtain similar results, but within a time varying framework. We define the time varying, dyadic, autoregressive, moving average (tvDARMA) process as follows.

Definition 2. A stochastic process \( \{X_t, t = 1, 2, \ldots, T\} \) is called tvDARMA(\( p, r \)) if it can be represented by

\[
 \sum_{k=0}^{p} b_{k,t} X_{t\oplus k} = \sum_{n=0}^{r} a_{n,t} \varepsilon_{t\oplus n},
\]

13
Figure 2: The spectral density function for the tvMA(2) process (left) and the tvDMA(2) process (right).

where $p, r \in \mathbb{Z}^+$ with $p = 2^m - 1$, $r = 2^f - 1$, the sequences of parameters $\{b_{k,t}\}_{k=0,1,...,p}, \{a_{n,t}\}_{n=0,1,...,p}$ are real numbers with at least two non-zero parameters $b_{k_0,t}, a_{n_0,t}$ for $2^m-1 \leq k_0 \leq 2^m - 1$ and $2^f-1 \leq n_0 \leq 2^f - 1$. In addition, $\{\varepsilon_t\}_{t \in \mathbb{N}}$ is the i.i.d. sequence.

If we rescale, the parameters $b_{k,t}, a_{n,t}$ to the unit interval, then (14) can be written as

$$\sum_{k=0}^{p} b_{k,t,T} X_{t \oplus k,T} = \sum_{n=0}^{r} a_{n,t,T} \varepsilon_{t \oplus n},$$

(15)

Assume that $b_{0,t} = a_{0,t} = 1$. If we set in (14), $p = 0$, then the tvDMA process arises as in (6), but for a finite order $r$. In case we set in (14) $r = 0$, then (14) becomes

$$\sum_{k=0}^{p} b_{k,t} X_{t \oplus k} = \varepsilon_t.$$

(16)

We call $X_{t,T}$ in (16) a time varying, dyadic, autoregressive (tvDAR) process of order $p$ (tvDAR($p$)).
Rescaling the parameters \( b_{k,t} \) to the unit interval, equation (16) can be written as

\[
\sum_{k=0}^{p} b_{k,t} X_{k\otimes k,t} = \varepsilon_t,
\]

which is the analogue of the rescaled tvDMA process in (7).

We show that a tvDAR process, and even more generally, a tvDARMA process, can be expressed as a tvDMA process. This is an interesting result, since it enables transfer the results on local stationarity of tvDMA processes to the framework of tvDAR and tvDARMA processes, respectively. It turns out that a tvDAR\((p)\) is not dyadic stationary, since its autocovariance function \( R(n,m) \) does not depend only on \( n \oplus m \). We can show, however, that the tvDAR\((p)\) can be written as a tvDMA\((p)\) process and therefore it is locally dyadic stationary. But first, we need a couple of lemmas. We follow the spirit of Nagai and Taniguchi (1987) who study multiple dyadic stationary processes. Set

\[
\phi_t(x) = \sum_{j=0}^{p} b_{j,t} W(j, x), \quad x \in I,
\]

where \( p = 2^m - 1, m \in \mathbb{N} \) and \( b_{j,t} \) are real numbers. Denote by \( \Sigma_t \) the \((p+1) \times (p+1)\) matrix, which is given by

\[
\Sigma_t = \begin{pmatrix}
  b_0 \oplus 0, t & b_0 \oplus 1, t & \cdots & b_0 \oplus p, t \\
  b_1 \oplus 0, t & b_1 \oplus 1, t & \cdots & b_1 \oplus p, t \\
  \vdots & \vdots & \ddots & \vdots \\
  b_p \oplus 0, t & b_p \oplus 1, t & \cdots & b_p \oplus p, t
\end{pmatrix}.
\]

**Lemma 1.** The following equation holds

\[
\det[\Sigma_t] = \prod_{j=0}^{p} \phi_t(x_j),
\]

where \( x_j = j/(p+1), \quad j = 0, 1, \ldots, p \). Therefore the function \( \phi_t(x) \neq 0 \) if and only if \( \det[\Sigma_t] \neq 0 \).
Lemma 2. Assume that $\phi_t(x) \neq 0$ in (18). Then there exists a function $\eta_t(x)$, which is defined by

$$
\eta_t(x) = \sum_{m=0}^{p} d_{m,t} W(m, x), \ d_{m,t} \in \mathbb{R}, \ x \in I,
$$

and satisfies $\phi_t(x) \eta_t(x) = 1$. The coefficients $d_{m,t}$ are uniquely determined by

$$
\Sigma_t \left( \begin{array}{cccc}
d_{0,t} & d_{1,t} & \cdots & d_{p,t}
\end{array} \right)' = 1.
$$

Define $S_t$ to be the $(r + 1) \times (r + 1)$ matrix

$$
S_t = \left( \begin{array}{cccc}
a_{0\oplus 0,t} & a_{0\oplus 1,t} & \cdots & a_{0\oplus r,t} \\
a_{1\oplus 0,t} & a_{1\oplus 1,t} & \cdots & a_{1\oplus r,t} \\
\vdots & \vdots & \ddots & \vdots \\
a_{r\oplus 0,t} & a_{r\oplus 1,t} & \cdots & a_{r\oplus r,t}
\end{array} \right).
$$

The following theorem states that a tvDARMA process can be represented both as a tvDMA and a tvDAR process. We denote by $\Sigma_{t,T}$ and $S_{t,T}$ the rescaled matrices $\Sigma_t$ and $S_t$, respectively. Lemmas 1 and 2 still hold under the rescaling.

Theorem 2. Suppose that $\{X_{t,T}, t = 1, 2, \ldots, T\}$ is a tvDARMA$(p, r)$ as in (15). Set $\mu = \max(p, r)$. Then the following hold

(i) If $\det[\Sigma_{t,T}] \neq 0$, then $X_{t,T}$ can be represented as a tvDMA$(\mu)$ process.

(ii) If $\det[S_{t,T}] \neq 0$, then $X_{t,T}$ can be represented as a tvDAR$(\mu)$ process.

If we set in Theorem 2, $r = 0$, $a_{0,t,T} = 1$, then (15) becomes

$$
\sum_{k=0}^{p} b_{k,t,T} X_{t\oplus k,T} = \varepsilon_t,
$$

which is the definition of tvDAR$(p)$ process in (17). Then we have the following result.

Corollary 1. Suppose that $\{X_{t,T}, t = 1, 2, \ldots, T\}$ is a tvDAR$(p)$ as in (17). If $\det[\Sigma_{t,T}] \neq 0$, $X_{t,T}$ can be represented as a tvDMA$(p)$ process.
Hence, both a tvDAR and a tvDARMA process can be transformed in a tvDMA form and, therefore the results regarding local stationarity for the tvDMA process still hold in their case.

5 Conclusions

The above discussion constitutes the groundwork for future work in the field of applications. In this direction, the concepts of Walsh spectrum and Walsh transform have to be considered. The Walsh spectrum for a real-valued dyadic stationary process \( \{X_t\}_{t \in \mathbb{N}} \) is defined by

\[
f(x) = \sum_{\tau=0}^{\infty} R(\tau)W(\tau, x), \quad 0 \leq x < \infty,
\]

where the covariance function \( R(\cdot) \) satisfies \( \sum_{\tau=0}^{\infty} |R(\tau)| < \infty \), (see e.g. Morettin (1974b, 1978, 1981)). Inverting (19), the covariance is given by

\[
R(\tau) = \int_{0}^{1} W(\tau, x)f(x)dx.
\]

The finite Walsh transform is given by

\[
d^{(N)}(x) = \sum_{n=0}^{N-1} X_n W(n, x), \quad \text{with} \quad x \in I.
\]

To estimate the Walsh spectrum, Morettin (1981) defined the Walsh periodogram, by

\[
I^{(N)}(x) = N^{-1}[d^{(N)}(x)]^2,
\]

and showed that \( I^{(N)}(x) \) is asymptotically an unbiased, but inconsistent, estimator of \( f(x) \). He also considered the smooth Walsh periodogram and other classes of estimates. Dyadic stationarity is necessary to estimate the Walsh spectrum. Therefore, in the case of local dyadic stationarity, it would be reasonable to divide the rescaled interval \( I \) into subintervals and estimate the Walsh spectrum within each subinterval, where local dyadic stationarity is satisfied. The number of the
observations within each subinterval \( \{ u = t/T : |t/T - u_0| \leq \epsilon/2 \} \) increases as \( T \) tends to infinity and the above asymptotical results still hold. A similar method is applied by Dahlhaus and Giraitis (1998) for real-time stationary processes.

Kohn (1980a,b) studied the system of Walsh functions for real time stationary processes. He defined the \( j-th \) logical autocovariances \( \tau(j) \) and the Walsh-Fourier spectral density function \( F(x) \), respectively, instead of \( R(\tau) \) and \( f(x) \) above. He considered the finite Walsh-Fourier transform \( \omega_N(x) \) and studied its asymptotic properties. A class of estimators for \( F(x) \) was obtained, the average Walsh periodogram being a member of this class. The concept of local stationarity could also be applied in this real time setting and we conjecture that similar results could be obtained also in this case.
Appendix

Proof of Lemma 1 Recall that $p = 2^m - 1$. The Walsh-ordered Hadamard matrix $H_W(m)$ is a $(2^p \times 2^p)$ matrix with elements of the form $W(n, x_m)$, $x_m = m/2^p$, $m, n = 0, 1, 2, \ldots, 2^p - 1$, see also Stoffer (1991). Then the following relations hold:

$$
\Sigma_t H_W(m) = \begin{pmatrix}
  b_{0\oplus 0, t} & b_{0\oplus 1, t} & \cdots & b_{0\oplus p, t} \\
b_{1\oplus 0, t} & b_{1\oplus 1, t} & \cdots & b_{1\oplus p, t} \\
  \vdots & \vdots & \ddots & \vdots \\
b_{p\oplus 0, t} & b_{p\oplus 1, t} & \cdots & b_{p\oplus p, t}
\end{pmatrix}
\begin{pmatrix}
  W(0, x_0) & W(0, x_1) & \cdots & W(0, x_p) \\
  W(1, x_0) & W(1, x_1) & \cdots & W(1, x_p) \\
  \vdots & \vdots & \ddots & \vdots \\
  W(p, x_0) & W(p, x_1) & \cdots & W(p, x_p)
\end{pmatrix}
\begin{pmatrix}
  \phi_t(x_0) \\
  \phi_t(x_1) \\
  \vdots \\
  \phi_t(x_p)
\end{pmatrix}
\begin{pmatrix}
  \phi(x_0) \\
  \phi(x_1) \\
  \vdots \\
  \phi(x_p)
\end{pmatrix}
= \{ \sum_{l=0}^{p} b_{(i-1)\oplus l, t} W(l, x_{j-1}) \} (i, j) = \{ \phi_t(x_{j-1}) W(i - 1, x_{j-1}) \} (i, j)
= H_W(m) \cdot \text{diag}[\phi_t(x_0), \phi_t(x_1), \ldots, \phi_t(x_p)].
$$

(A-1)

But, since $\det[H_W(m)] = (p + 1)^{p+1}/2 \neq 0$, we get from (A-1) that

$$
\det[\Sigma_t] = \det[\text{diag}[\phi_t(x_0), \phi_t(x_1), \ldots, \phi_t(x_p)]] = \prod_{j=0}^{p} \det[\phi_t(x_j)].
$$

Proof of Lemma 2

$$
\phi_t(x) \eta_t(x) = \left( \sum_{l=0}^{p} b_{l, t} W(l, x) \right) \left( \sum_{m=0}^{p} d_{m, t} W(m, x) \right) = \sum_{l=0}^{p} \sum_{m=0}^{p} b_{l, t} d_{m, t} W(l \oplus m, x)
= \sum_{h=0}^{p} \left[ \sum_{j=0}^{p} b_{j\oplus h, t} d_{j, t} \right] W(h, x).
$$

(A-2)

In order for $\phi_t(x) \eta_t(x) = 1$ to hold, we have from equation (A-2) that

$$
\sum_{h=0}^{p} \left[ \sum_{j=0}^{p} b_{j\oplus h, t} d_{j, t} \right] W(h, x_j) = 1, \quad j = 0, 1, \ldots, p,
$$

19
which is equivalently written in matrix notation as

\[
H_W'(m) \begin{pmatrix}
\sum_{j=0}^{p} b_{j,t} d_{j,t} \\
\sum_{j=0}^{p} b_{j+1,t} d_{j,t} \\
\vdots \\
\sum_{j=0}^{p} b_{j+p,t} d_{j,t}
\end{pmatrix} = \begin{pmatrix}
1 \\
1 \\
\vdots \\
1
\end{pmatrix}.
\]  \hspace{1cm} (A-3)

But \(H_W(m)H'_W(m) = 2^m I_{2^m}\). Hence, since from assumption we have that \(\det [\Sigma_t] \neq 0\), equation (A-3) gives that

\[
2^m I_{2^m} \begin{pmatrix}
\sum_{j=0}^{p} b_{j,t} d_{j,t} \\
\sum_{j=0}^{p} b_{j+1,t} d_{j,t} \\
\vdots \\
\sum_{j=0}^{p} b_{j+p,t} d_{j,t}
\end{pmatrix} = H_W(m) \begin{pmatrix}
1 \\
1 \\
\vdots \\
1
\end{pmatrix} = 2^m \begin{pmatrix}
1 \\
0 \\
\vdots \\
0
\end{pmatrix} \implies \begin{pmatrix}
d_{0,t} \\
d_{1,t} \\
\vdots \\
d_{p,t}
\end{pmatrix} = \Sigma_t^{-1} \begin{pmatrix}
1 \\
0 \\
\vdots \\
0
\end{pmatrix}.
\]

**Proof of Theorem 2** Suppose that \(X_t\) and \(\varepsilon_t\) have Walsh spectral representations

\[
X_t = \int_0^1 W(t, x)dZ_X(x) \quad \text{and} \quad \varepsilon_t = \int_0^1 W(t, x)dZ_\varepsilon(x), \quad t \in \mathbb{N}.
\]

Then the LHS and RHS of equation (13) can be written as

\[
\sum_{k=0}^{p} b_{k,t,T} X_{t \oplus k; T} = \int_0^1 W(t, x) \left( \sum_{k=0}^{p} b_{k,t,T} W(k, x) \right) dZ_X(x) = \int_0^1 W(t, x) \phi_{1,t,T}(x)dZ_X(x)
\]

\[
\sum_{n=0}^{r} a_{n,t,T}\varepsilon_{t \oplus n} = \int_0^1 W(t, x) \left( \sum_{n=0}^{r} a_{n,t,T} W(n, x) \right) dZ_\varepsilon(x) = \int_0^1 W(t, x) \phi_{2,t,T}(x)dZ_\varepsilon(x).
\]

The system of Walsh functions is complete and equation (13) holds for \(t = 1, 2, \ldots, T\), hence, we have that \(\int_0^1 W(t, x) \phi_{1,t,T}(x)dZ_X(x) = \int_0^1 W(t, x) \phi_{2,t,T}(x)dZ_\varepsilon(x)\).

(i) Since \(\det[\Sigma_{t,T}] \neq 0\), then \(\phi_{1,t,T}(x) \neq 0\) and therefore there exists a function \(\eta_{1,t,T}(x) = \frac{1}{\phi_{1,t,T}(x)}\).
\[ \sum_{k=0}^{p} g_{k,t,T} W(k, x) \] such that

\[
dZ_X(x) = \eta_{1,t,T}(x) \phi_{2,t,T}(x) dZ_\varepsilon(x)
\]

\[
= \begin{cases} 
\sum_{j=0}^{\mu} \left( \sum_{l=0}^{p} g_{l,t,T} a_{l \oplus j,t,T} \right) W(j, x) dZ_\varepsilon(x), & p \leq r; \\
\sum_{j=0}^{\mu} \left( \sum_{l=0}^{r} g_{l,t,T} a_{l \oplus j,t,T} \right) W(j, x) dZ_\varepsilon(x), & p > r.
\end{cases}
\]

Hence,

\[
dZ_X(x) = \sum_{j=0}^{\mu} K_{j,t,T} W(j, x) dZ_\varepsilon(x),
\]

where

\[
K_{j,t,T} = \begin{cases} 
\sum_{s=0}^{p} g_{s,t,T} a_{s \oplus j,t,T}, & p \leq r, \\
\sum_{s=0}^{r} g_{s \oplus j,t,T} a_{s,t,T}, & p > r, 
\end{cases} \quad j = 0, 1, \ldots, p.
\]

Therefore,

\[
X_t = \sum_{j=0}^{\mu} K_{j,t,T} \int_{0}^{1} W(t \oplus j, x) dZ_\varepsilon(x) = \sum_{j=0}^{\mu} K_{j,t,T} \varepsilon_{t \oplus j}.
\]

(ii) Similarly with (i).

**Proof of Corollary**

Suppose that \( X_{t,T} = \int_{0}^{1} W(t, x) dZ_X(x) \). The LHS of (17) is

\[
\sum_{k=0}^{p} b_{k,t,T} X_{t \oplus k} = \sum_{k=0}^{p} b_{k,t,T} \int_{0}^{1} W(t \oplus k, x) dZ_X(x) = \int_{0}^{1} W(t, x) \phi_{t,T}(x) dZ_X(x),
\]

where \( \phi_{t,T}(x) = \sum_{k=0}^{p} b_{k,t,T} W(k, x) \). From Lemma[2], since by assumption \( \det[\Sigma_{t,T}] \neq 0 \), \( \exists \eta_{t,T}(x) = \sum_{m=0}^{p} d_{m,t,T} W(m, x) \), \( d_{m,t,T} \in \mathbb{R} \), such that \( \phi_{t,T}(x) \eta_{t,T}(x) = 1 \). Therefore, and since the system of Walsh functions is complete, we have that

\[
dZ_X(x) = \eta_{t,T}(x) dZ_\varepsilon(x) = \sum_{m=0}^{p} d_{m,t,T} W(m, x) dZ_\varepsilon(x).
\]
Hence,

\[ X_{t,T} = \int_0^1 W(t, x) \sum_{m=0}^p d_{m,t,T} W(m, x) dZ_\varepsilon(x) = \sum_{m=0}^p d_{m,t,T} \int_0^1 W(t \oplus m, x) dZ_\varepsilon(x) \]

\[ = \sum_{m=0}^p d_{m,t,T} \varepsilon_{t \oplus m}. \]
References

Abbasi, S. A., A. Alamoud, et al. (2012). FPGA based Walsh and inverse Walsh transforms for signal processing. *Elektronika ir Elektrotechnika* 18, 3–8.

Beauchamp, K. G. (1984). *Walsh functions and their applications*. Academic Press.

Chrestenson, H. (1955). A class of generalized Walsh functions. *Pacific J. Math* 5, 17–31.

Dahlhaus, R. (1996a). Asymptotic statistical inference for nonstationary processes with evolutionary spectra. In *Athens conference on applied probability and time series analysis*, pp. 145–159. Springer.

Dahlhaus, R. (1996b). Maximum likelihood estimation and model selection for locally stationary processes. *Journal of Nonparametric Statistics* 6, 171–191.

Dahlhaus, R. (1996c). On the Kullback-Leibler information divergence of locally stationary processes. *Stochastic Processes and their Applications* 62, 139–168.

Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *The Annals of Statistics* 25, 1–37.

Dahlhaus, R. (2000). A likelihood approximation for locally stationary processes. *Annals of Statistics* 28, 1762–1794.

Dahlhaus, R. (2012). Locally stationary processes. In T. S. Rao, S. S. Rao, and C. R. Rao (Eds.), *Handbook of Statistics: Time Series Analysis-Methods and Applications*, Volume 30, pp. 351–408. Amsterdam: Elsevier B. V.

Dahlhaus, R. and L. Giraitis (1998). On the optimal segment length for parameter estimates for locally stationary time series. *Journal of Time Series Analysis* 19, 629–655.
Davis, R. A., T. C. M. Lee, and G. A. Rodriguez-Yam (2006). Structural break estimation for nonstationary time series models. *Journal of the American Statistical Association* 101, 223–239.

Fine, N. (1950). The generalized Walsh functions. *Transactions of the American Mathematical Society*, 66–77.

Fine, N. (1957). Fourier-Stieltjes series of Walsh functions. *Transactions of the American Mathematical Society* 86, 246–255.

Fine, N. J. (1949). On the Walsh functions. *Transactions of the American Mathematical Society* 65, 372–414.

Granger, C. W. J. and M. Hatanaka (1964). *Spectral analysis of economic time series*. Princeton: Univ. Press.

Grove, A. (1983). An introduction to Walsh functions and their applications. *International Journal of Mathematical Education in Science and Technology* 14, 43–53.

Harmuth, H. F. (1969). Applications of Walsh functions in communications. *Spectrum, IEEE* 6, 82–91.

Harmuth, H. F. (1977). *Sequency theory: Foundations and applications*. New York: Academic Press.

Kohn, R. (1980a). On the spectral decomposition of stationary time series using Walsh functions. I. *Advances in Applied Probability*, 183–199.

Kohn, R. (1980b). On the spectral decomposition of stationary time series using Walsh functions. II. *Advances in Applied Probability*, 462–474.
Maquasi, M. (1981). *Applied Walsh Analysis*. Heyden and Sons Ltd., London.

Martin, W. (1981). Line tracking in nonstationary processes. *Signal Processing* 3, 147–155.

Melard, G. and A. H.-d. Schutter (1989). Contributions to evolutionary spectral theory. *Journal of Time Series Analysis* 10, 41–63.

Morettin, P. (1978). Estimation of the spectrum and of the covariance function of a dyadic-stationary series. *Bulletin of the Brazilian Mathematical Society* 9, 83–88.

Morettin, P. A. (1974a). Limit theorems for stationary and dyadic-stationary processes. *Bulletin of the Brazilian Mathematical Society* 5, 97–104.

Morettin, P. A. (1974b). Walsh-function analysis of a certain class of time series. *Stochastic processes and their applications* 2, 183–193.

Morettin, P. A. (1981). Walsh spectral analysis. *SIAM Review* 23, 279–291.

Morettin, P. A. (1983). A note on a central limit theorem for stationary processes. *Journal of Time Series Analysis* 4, 49–52.

Nagai, T. (1977). Dyadic stationary processes and their spectral representations. *Bulletin of Mathematical Statistics* 17, 65–73.

Nagai, T. (1980). On finite parametric linear models of dyadic stationary processes. *Bulletin of Mathematical Statistics* 19, 45–53.

Nagai, T. and M. Taniguchi (1987). Walsh spectral analysis of multiple dyadic stationary processes and its applications. *Stochastic Processes and their Applications* 24, 19–30.
Nason, G. P., R. Von Sachs, and G. Kroisandt (2000). Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 62, 271–292.

Neumann, M. H. and R. Von Sachs (1997). Wavelet thresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra. *The Annals of Statistics* 25, 38–76.

Paley, R. (1932a). A remarkable series of orthogonal functions (I). *Proceedings of the London Mathematical Society* 2, 241–264.

Paley, R. (1932b). A remarkable series of orthogonal functions (II). *Proceedings of the London Mathematical Society* 2, 265–279.

Priestley, M. (1966). Design relations for non-stationary processes. *Journal of the Royal Statistical Society. Series B (Methodological)* 28, 228–240.

Priestley, M. B. (1965). Evolutionary spectra and non-stationary processes. *Journal of the Royal Statistical Society. Series B (Methodological)* 27, 204–237.

Priestley, M. B. (1981). *Spectral analysis and time series*. Academic Press, London.

Priestley, M. B. (1988). *Non-linear and non-stationary time series analysis*. Academic Press, London.

Rao, T. S. and H. Tong (1972). A test for time-dependence of linear open-loop systems. *Journal of the Royal Statistical Society. Series B (Methodological)* 34, 235–250.

Sakiyama, K. and M. Taniguchi (2004). Discriminant analysis for locally stationary processes. *Journal of Multivariate Analysis* 90, 282–300.
Selfridge, R. (1955). Generalized Walsh transforms. *Pacific J. Math* 5, 451–480.

Stoffer, D. S. (1985). Central limit theorems for finite Walsh-Fourier transforms of weakly stationary time series. *Journal of Time Series Analysis* 6, 261–267.

Stoffer, D. S. (1987). Walsh-Fourier analysis of discrete-valued time series. *Journal of Time Series Analysis* 8, 449–467.

Stoffer, D. S. (1990). Multivariate Walsh-Fourier analysis. *Journal of Time Series Analysis* 11, 57–73.

Stoffer, D. S. (1991). Walsh-Fourier analysis and its statistical applications. *Journal of the American Statistical Association* 86, 461–479.

Taniguchi, M., L. Zhao, P. Krishnaiah, and Z. Bai (1989). Statistical analysis of dyadic stationary processes. *Annals of the Institute of Statistical Mathematics* 41, 205–225.

Tjøstheim, D. (1976). Spectral generating operators for non-stationary processes. *Advances in Applied Probability* 8, 831–846.

Walsh, J. (1923). A closed set of normal orthogonal functions. *American Journal of Mathematics* 45, 5–24.