

Testing for Periodicity in Signals: An Application to Detect Partial Upper Airway Obstruction During Sleep

T. AITTOKALLIO^a, M. GYLLENBERG^{a*}, O. NEVALAINEN^a and O. POLO^b

^a*Department of Mathematical Sciences, University of Turku, and Turku Centre for Computer Science TUCS, FIN-20014, Turku, Finland;*

^b*Department of Physiology, University of Turku, Turku, Finland*

(Received 24 October 2000; Revised 23 April 2001; In final form 23 April 2001)

A simple method for detecting periodic components of unknown periodicity in a signal is presented. The method is based on spectral decomposition of signal using orthonormal functions. Traditionally, hypothesis testing together with harmonic functions is used, but we show that the same statistical properties are obtained for other systems of orthonormal functions as well. The appropriate behavior of the method is first demonstrated with simulation studies and then tested to identify visually determined clusters of high-frequency movements, which may repeat in synchrony with respiration during sleep. The good performance in the practical tests suggests that an automatic identification of these clusters could be based on Walsh functions.

Keywords: Fourier transform; Hypothesis testing; Periodogram analysis; Static charge-sensitive bed; Upper airway resistance syndrome; Walsh transform

1. INTRODUCTION

Relaxation of postural muscles is characteristic for the sleep state. In heavy snorers, the decreased postural muscles tone in the tongue and soft palate may result in marked narrowing (increased respiratory resistance, IRR) or complete obstruction (obstructive sleep apnea, OSA) of the upper airway. Repeated episodes of upper airway obstruction compromise pulmonary ventilation and cause arousals from sleep, sleep loss, excessive sleepiness and impaired cognitive function during daytime. These sleep-related breathing dis-

turbances could nowadays be effectively treated, if the condition is suspected and adequately diagnosed. Whereas a number of recording techniques are available to demonstrate episodes of OSA, only a few methods have been suggested to detect IRR.

The respiratory disturbances during sleep are conventionally diagnosed in a whole night sleep study with continuous monitoring of the respiratory effort and gas exchange. We have previously shown that by using a simple but sensitive movement sensor, the static charge-sensitive bed (SCSB), episodes of OSA as well as IRR can be monitored without

*Corresponding author. Tel.: +358-2-3336567. E-mail: mats.gyllenberg@utu.fi

attaching electrodes on the sleeping subject (Polo, 1992). The major advantage of this approach is minimal sleep disturbance during the monitoring session. Detection of the IRR with the SCSB is based on the appearance of the high-frequency respiratory movement (HFRM) patterns caused by increased respiratory efforts as response to partial upper airway obstruction (Kirjavainen, 1997). Since detection of IRR in large patient populations is a technical challenge, automatic detection and analysis of the HFRM events from the SCSB recordings would essentially facilitate and standardize the diagnostic procedure. We previously developed and evaluated the performance of a range of automated methods to detect all the HFRM events during a sleep recording (Aittokallio *et al.*, 2000). To improve the specificity to detect the clinically relevant episodes of IRR, which typically appear as clusters of periodic events rather than occasional events, we have developed a protocol to identify these clusters. The method presented here focuses on this problem and is proposed to replace the laborious visual scanning in clinical diagnostics and medical research. It could also facilitate locating periods of interests in all-night sleep recordings where more detailed analyses on the inspiratory flow shapes (Aittokallio *et al.*, 1999; 2001), and trends in ventilation or oxygenation are needed.

The problem of detecting unknown, or hidden, periodicities in a time series is well-known in statistics (Priestley, 1981). A test based on the Fourier periodogram referred to as Fisher's test, has been successfully applied in various real signals, e.g. control chart signal (Tatum, 1996). The popularity of Fisher's test results mainly from the simple and attractive test statistic for which Fisher (1929) also gave the exact null distribution. The test has been modified also for the case where the hidden periodicities are harmonics of some fundamental periodicity (Hinich, 1982). These types of test have been used in the case of discrete signals, or sequences, as well, e.g. nucleic acids sequence (Chechetkin and Turygin, 1995). Also certain specific algorithms for periodic sequences have been developed. The problem is that the algorithms for exact periodicity (Choffrut and Karhumäki, 1997) are not suitable for

many situations in practice, and the approximate algorithms are usually very slow for long sequences (Coward and Drabløs, 1998).

In this paper, a class of tests for periodicity is proposed that contains Fisher's test as a special case. The test is based on an orthonormal transform of the signal, and Fisher's test is obtained when the Fourier transform is used. We note that all such transforms can be implemented on the order of $N \log N$ operations, where N is the length of the signal (Andrews and Caspari, 1970). Moreover, we show that the statistical properties of Fisher's test can be derived also for the general test when N is large. In order to show the practical usefulness of the test, we consider in more detail the use of the Walsh periodogram in the problem of detecting periodic HFRM events in a binary sequence. In situations where signals have sharp edges, like in sequences, the Walsh periodogram can be more applicable than the Fourier periodogram (Beer, 1981) with the additional advantage that it can be produced more quickly since only $N \log N$ additions of real numbers are required in the fast Walsh transform (FWT).

The paper is structured as follows. Following a short overview of the theory of orthonormal transforms (Section 2), we give the general periodicity test procedure, including parametric and nonparametric distributions for significance of an observed test statistic value (Section 3). The operation of the test is investigated by means of simulation using both periodic and nonperiodic artificial data (Section 4). Finally, we consider the real data problem of identifying significant episodes of IRR (Section 5), and give some conclusions of the work (Section 6).

2. SYSTEMS OF ORTHONORMAL FUNCTIONS

In this section we recall some basic facts about orthonormal functions, the Fourier and Walsh transforms, and introduce notation to be used throughout the paper.

Let $S = (f_k)_{k=0}^{N-1}$ be a system of N real valued functions defined on a discrete domain of N points

labeled as $0, 1, \dots, N - 1$. Throughout this paper we assume that $N = 2^n$ with n is a positive integer. The system S is called orthonormal if

$$\sum_{t=0}^{N-1} f_i(t)f_j(t) = \delta_{ij}, \quad i, j = 0, 1, \dots, N - 1, \quad (1)$$

where δ_{ij} is Kronecker's delta function (i.e., $\delta_{ij} = 1$ if $i = j$, and zero otherwise). Let $\mathbf{U} = (f_k(t))_{k,t=0}^{N-1}$ be a matrix corresponding to an orthonormal system S . Then Eq. (1) implies that \mathbf{U} is an orthogonal matrix, that is,

$$\mathbf{U}^T \mathbf{U} = \mathbf{U} \mathbf{U}^T = \mathbf{I}, \quad (2)$$

where \mathbf{U}^T denotes the transpose of \mathbf{U} and \mathbf{I} is the identity matrix of order N . Consider a real valued vector, or signal, $\mathbf{x} = (x_t)_{t=0}^{N-1}$. The rows of \mathbf{U} form an orthonormal basis in \mathbb{R}^N , and therefore vector \mathbf{x} can be written uniquely as a linear combination of the elements of the basis, that is, $\mathbf{x} = \mathbf{U}^T \mathbf{c}$. Since $\mathbf{U}^{-1} = \mathbf{U}^T$ according to Eq. (2), the coordinate vector \mathbf{c} can be solved easily as

$$\mathbf{c} = \mathbf{U} \mathbf{x}. \quad (3)$$

We call the vector $\mathbf{c} = (c_k)_{k=0}^{N-1}$ in Eq. (3) the \mathbf{U} -transform of \mathbf{x} and the integer $k \in \{0, 1, \dots, N - 1\}$ is the *spectral index*. Note, that each \mathbf{U} -transform satisfies the Parseval relation

$$\mathbf{x}^T \mathbf{x} = \mathbf{c}^T \mathbf{c}, \quad (4)$$

which states that the transform conserves the energy of the signal. Next we recall two well-known orthonormal function systems which will be used in periodicity testing.

2.1 Fourier Transform

The Fourier transform applies to the orthogonal harmonic functions $h_j(x)$, which are defined as below on the interval $0 \leq x < 1$, and have zero value for all other values of x outside this interval. We note that the range of these functions is the interval $[-1, 1]$. The

definition is based on the cosine and sine functions:

$$\begin{aligned} h_{2j}(x) &= \cos(2\pi jx), \quad j = 0, 1, \dots, \\ h_{2j-1}(x) &= \sin(2\pi jx), \quad j = 1, 2, \dots, \end{aligned} \quad (5)$$

see Fig. 1 (left column). The normalized and time-discretized harmonic functions

$$f_\omega(t) = \begin{cases} h_0(t/N)/\sqrt{N} = 1/\sqrt{N} & \text{for } \omega = 0, \\ h_\omega(t/N)/\sqrt{N/2} & \text{for } \omega = 1, 2, \dots, N - 2, \\ h_N(t/N)/\sqrt{N} = (-1)^j/\sqrt{N} & \text{for } \omega = N - 1, \end{cases} \quad (6)$$

where $t = 0, 1, \dots, N - 1$, form an orthonormal system. In Fourier analysis, the integer ω is called the *frequency index* and the vector \mathbf{c} in Eq. (3) is known as the discrete Fourier transform (DFT) of the signal \mathbf{x} . The matrix \mathbf{U} induces a bijective linear mapping from time domain to frequency domain.

2.2 Walsh Transform

The Walsh transform uses the orthogonal square-wave functions $w_j(x)$, introduced by Walsh (1923), which have only two values ± 1 in the interval $0 \leq x < 1$ and the value zero elsewhere. The original definition of the Walsh functions was based on the following recursive equations:

$$\begin{aligned} w_{2j}(x) &= w_j(2x) + (-1)^j w_j(2x - 1), \\ j &= 1, 2, \dots, \\ w_{2j-1}(x) &= w_{j-1}(2x) - (-1)^{j-1} w_{j-1}(2x - 1), \\ j &= 1, 2, \dots \end{aligned} \quad (7)$$

with the initial condition $w_0(x) = 1$. Several other definitions of the Walsh functions exist in the literature which differ only in the ordering of the functions (Yuen, 1972). This ordering of the Walsh functions corresponds to the ordering of the harmonic functions, see Fig. 1 (right column). Walsh functions with even and odd orders are called the *cal* and *sal* functions, respectively, and they correspond to the cosine and sine functions in Fourier analysis. One should note, that in contrast to harmonic functions the

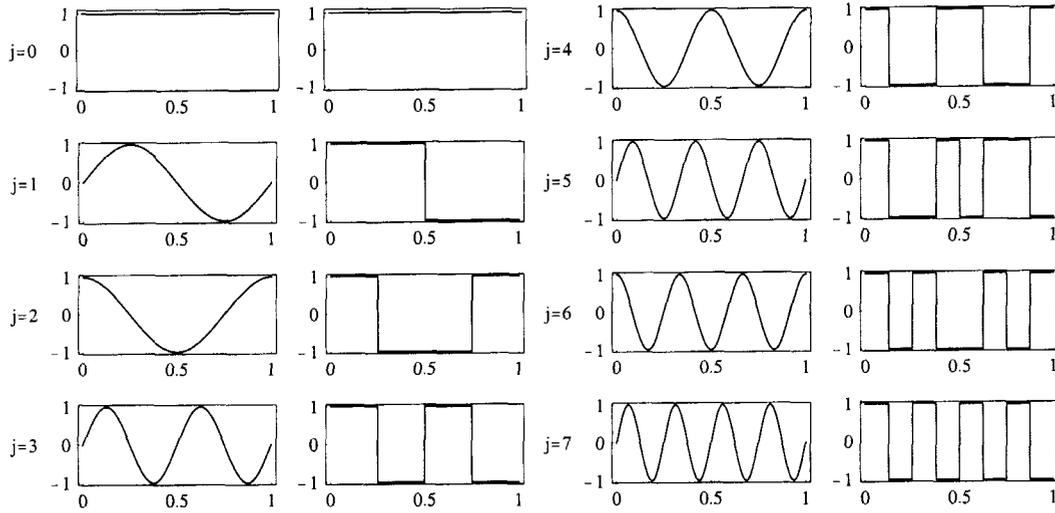


FIGURE 1 The first eight harmonic functions (Eq. (5) left column), and Walsh functions (Eq. (7) right column). The functions take the value zero when $x < 0$ or $x \geq 1$.

sign changes of the Walsh functions are not necessarily equidistant (e.g. Walsh functions of order 5 and 6 in Fig. 1). It is easy to see that the normalized and time-discretized Walsh functions

$$f_s(t) = \begin{cases} w_0(t/N)/\sqrt{N} = 1/\sqrt{N} & \text{for } s = 0, \\ w_s(t/N)/\sqrt{N} & \text{for } s = 1, 2, \dots, N-2 \\ w_N(t/N)/\sqrt{N} = (-1)^t/\sqrt{N} & \text{for } s = N-1, \end{cases} \quad (8)$$

where $t = 0, 1, \dots, N-1$, form an orthonormal system. In Walsh analysis, the integer s is called the *sequency index* and the vector \mathbf{c} in Eq. (3) is known as the discrete Walsh transform (DWT) of signal \mathbf{x} . The sequency s is used as a generalized frequency in communication theory (Harmuth, 1969). The matrix \mathbf{U} induces a bijective linear mapping from the time domain to the sequency domain. In this case, the entries of \mathbf{U} have only two possible values, namely $\pm 1/\sqrt{N}$. Moreover, since \mathbf{U} is symmetric, the inverse DWT is also generated by \mathbf{U} , that is, $\mathbf{x} = \mathbf{U}\mathbf{c}$.

3. TEST FOR PERIODICITY

Our tool in detection of hidden periodicities in a signal is the *U-periodogram* $\mathbf{p} = (p_k)_{k=0}^{N/2}$ of signal \mathbf{x}

defined as

$$p_k = \begin{cases} c_0^2 & \text{for } k = 0, \\ c_{2k-1}^2 + c_{2k}^2 & \text{for } k = 1, 2, \dots, N/2 - 1, \\ c_{N-1}^2 & \text{for } k = N/2. \end{cases} \quad (9)$$

The k th component of the \mathbf{U} -periodogram, when $k \in \{1, 2, \dots, N/2 - 1\}$, gives the amount of the signal energy distributed on the spectral indices $2k-1$ and $2k$ in the \mathbf{U} -transform. The energy of two successive indices is merged because in the systems we investigate (Fourier, Walsh), the successive functions (sin and cos, sal and cal) differ only in the respect of a time delay, that is, they have identical periodicity characteristics. In these systems, the first component represents the average of the signal since

$$\sqrt{p_0} = c_0 = \sum_{t=0}^{N-1} x_t/\sqrt{N} = \sqrt{N}\bar{x}, \quad (10)$$

where \bar{x} is the usual sample mean of signal \mathbf{x} . Thus,

$$m = \sqrt{p_0/N} \quad (11)$$

is an estimator of the signal mean. Moreover, the Parseval relation in Eq. (4) together with Eq. (10)

gives

$$\begin{aligned} \sum_{k=1}^{N/2} p_k &= \sum_{k=0}^{N-1} c_k^2 - c_0^2 = \sum_{t=0}^{N-1} x_t^2 - N\bar{x}^2 \\ &= \sum_{t=0}^{N-1} (x_t - \bar{x})^2 = (N-1)s^2, \end{aligned}$$

where s^2 is the usual sample variance of \mathbf{x} . Therefore,

$$v = \sum_{k=1}^{N/2} p_k / (N-1) \quad (12)$$

is an estimator of the signal variance. The last component in Eq. (9) describes the energy distributed on the spectral index corresponding to the function $(-1)^t/\sqrt{N}$, see Eqs. (6) and (8).

Heuristically we can state that if there exists periodicity in the signal, the periodogram components corresponding to the function pairs whose characteristics are close to the true periodicity will be large when compared with the other periodogram components. In Fisher's test, the whole range of $0, 1, \dots, N/2$ is investigated, but this is not necessary. Here we let $K \subseteq \{0, 1, \dots, N/2\}$ be the set of spectral indices among which one wishes to search for the periodicity. Because we want to recognize whether there exists any periodicity or not, a natural test statistic is therefore the maximum of the normalized U-periodogram components

$$T = \frac{\max_{k \in K} p_k}{\sum_{k \in K} p_k}. \quad (13)$$

Although we observe that the periodogram contains a peak we cannot conclude immediately that there is a genuine periodic component in a signal: even a non-periodic signal may produce peaks due merely to random fluctuations. Therefore, we must test whether the observed value of T is significantly larger than what would be likely to arise under the null hypothesis (H_0) that there are no genuine periodic components in the signal. Formally, the significance of the observed value t is defined by the p -value $p(t) = P(T > t | H_0)$. The smaller the p -value the more extreme the

observed value and the stronger the evidence against the null hypothesis. There exist many different techniques to test the significance of an observation, that is, to approximate the null distribution $P(T > t | H_0)$. In the parametric inference approach, an analytical formula is derived for the significance. On the other hand, in nonparametric methods (e.g. Monte Carlo) samples under the null hypothesis are generated and the significance is obtained with the help of a sample distribution function.

3.1 Parametric Approach

In this study, we set up the usual null hypothesis that the random signal \mathbf{x} under the null hypothesis originates from a white noise process. In other words, \mathbf{x} is a sequence of independently and identically distributed random variables with finite mean $E(x_t) = \mu$ and variance $\text{Var}(x_t) = \sigma^2$. Then it can be shown (see Appendix for details) that the variables p_k/σ^2 are asymptotically independent and exponentially distributed for all $k \in K \subseteq \{1, 2, \dots, N/2 - 1\}$, and the approximate parametric null distribution of T is

$$p_P(t) = \sum_{j=1}^{\lfloor 1/t \rfloor} (-1)^{j-1} \binom{l}{j} (1-jt)^{l-1}, \quad (14)$$

$$0 < t \leq 1,$$

where $\lfloor x \rfloor$ denotes the integer part of x and l is the number of indices in K . Also, $p_P(0) = 1$. In the special case, where $K = \{1, 2, \dots, N/2\}$, variables p_k/v have also asymptotically an exponential distribution since v in Eq. (12) under the white noise assumption is an unbiased and consistent estimator of σ^2 . Thus, one can derive a simpler asymptotic p -value for T in this case. Since $(N-1)T = \max p_k/v$, one gets as the maximum of $N/2$ independent exponential random variables

$$p_P(t) = 1 - [1 - \exp(-(N-1)t/2)]^{N/2}, \quad (15)$$

$$0 \leq t \leq 1.$$

The derivation of the parametric Eqs. (14) and (15) is

based on the central limit theory, and therefore, the approximation capability improves with increasing N .

3.2 Nonparametric Approach

Although the parametric approach benefits from the analytical formulae, which are easy to use, the nonparametric approach has one main advantage: the results are solely based on data of the particular experiment. In the Monte Carlo method, a sample of M_0 independent signals is generated under the null hypothesis. Let $\{t_1, t_2, \dots, t_{M_0}\}$ be the set of corresponding observed values of the test statistic T . The nonparametric approximation of the null distribution is then obtained by the sample distribution function

$$p_N(t) = |\{i|t_i > t\}|/M_0. \quad (16)$$

The accuracy of the approximation increases with increasing values of M_0 . One can generate signals under the null hypothesis by using the given null hypothesis model (like simulations of section 4), or by more sophisticated models, including randomization and bootstrap methods (Manly, 1997).

3.3 Test Procedure

Statistical hypothesis testing gives us the following simple rule to test H_0 against the alternative hypothesis H_1 that there exists genuine periodic components in the signal. First, choose a threshold value t . If the observed value of T exceeds t then the maximum of the periodogram components is large enough to justify the rejection of the null hypothesis, and we conclude that the signal contains periodicity.

When using the hypothesis testing procedure one can make two types of errors. The p -value is the probability of type I error, that is, the probability of rejecting H_0 when H_0 is true. The type II error is made if one accepts H_0 when H_0 is false. This occurs with a probability $P(T \leq t|H_1)$. Ideally, we should like both types of errors to have probability zero, but this is impossible. For example, in Eqs. (14)–(16), the p -value zero is obtained only at $t = \infty$. Therefore, it is

customary to fix only the probability of type I error at a preassigned (small) level α , called the *significance level* of the test (Rohatgi, 1984). Then one determines the threshold t_α for which $P(T > t_\alpha|H_0) = \alpha$ by using either Eqs. (14), (15), or Eq. (16). If the observed value of T is larger than t_α , the null hypothesis is rejected in favour of the alternative hypothesis at a level α . However, if we have signals generated under the alternative hypothesis, we are able to control the probability of the type II error as well. In the Monte Carlo method, the probability is obtained as

$$q_N(t) = |\{i|t_i \leq t\}|/M_1, \quad (17)$$

where $\{t_1, t_2, \dots, t_{M_1}\}$ are the observed values of the test statistic T computed from the M_1 independent signals generated under the alternative hypothesis.

4 SIMULATION STUDIES

In this section, we illustrate the use of the test procedures of the previous section by means of three examples. The theory of the previous section has been derived for all real signals. From now on, we concentrate only on binary sequences, $x_t \in \{0, 1\}$ for all $t = 0, 1, \dots, N - 1$. Here $x_t = 1$ means that a certain event takes place at time t . We are interested in recognizing if there exists any periodicity in the event sequence \mathbf{x} .

4.1 Nonperiodic Sequences

The first example is somewhat artificial describing the exact null hypothesis situation. Let us assume that the sequence \mathbf{x} is an outcome of $N = 8192$ independent Bernoulli trials having constant probability $\rho = 0.2$ of success, i.e. $P(x_t = 1) = 0.2$. The top row of Fig. 2 shows the normalized periodogram components $p_k/\sum_{k \in K} p_k$ for all $k \in \{1, 2, \dots, 4096\}$, using both Fourier and Walsh systems. Although the theoretical normalized periodogram of a white noise process is always flat having only values of $2/N = 0.00024$, a realization of such a process produces random fluctuations to a periodogram. This realization yields

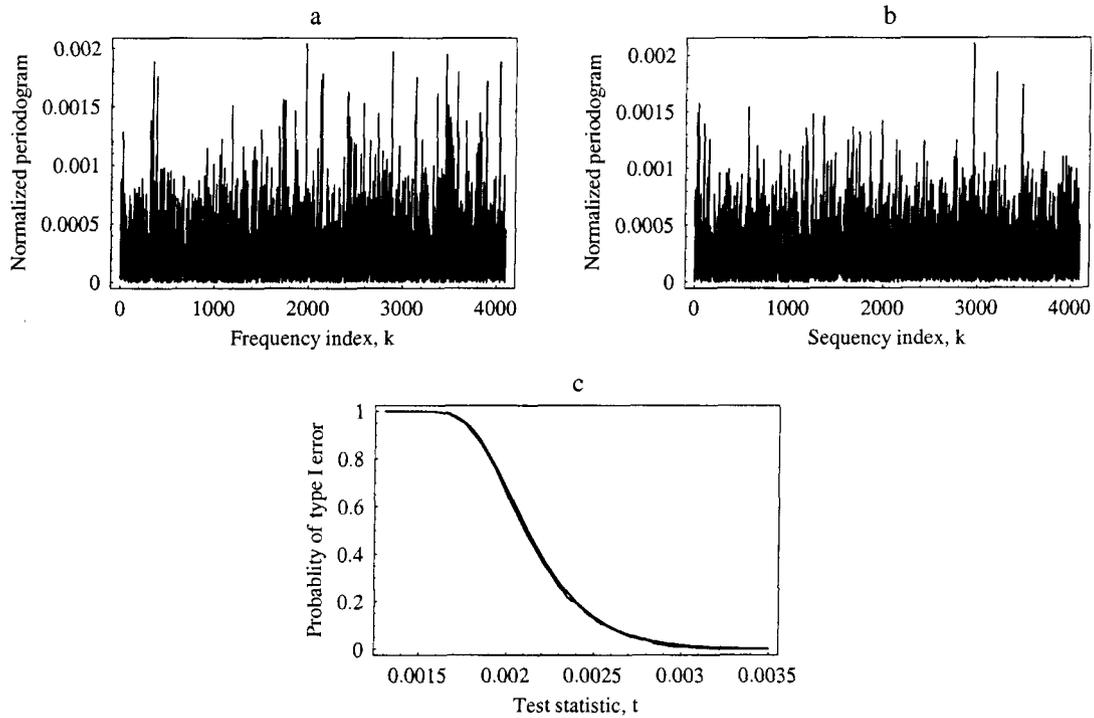


FIGURE 2 Simulation study of nonperiodic sequences. (a) Normalized Fourier periodogram of $N = 8192$ independent Bernoulli trials, and (b) normalized Walsh periodogram of the same sequence. (c) Comparison between parametric (Eq. (15)) and nonparametric (Eq. (16)) distributions.

the values $m = 0.206$ and $v = 0.164$ in Eqs. (11) and (12), respectively, as estimators of the signal mean and variance for both systems. The true mean is now $\mu = \rho = 0.2$ and the variance of the Bernoulli series is $\sigma^2 = \rho(1 - \rho) = 0.16$. The observed value of the test statistic computed from the Fourier periodogram is $t_F = 0.00204$ (Fig. 2(a)) and from the Walsh periodogram $t_W = 0.00210$ (Fig. 2(b)). Fig. 2 shows also the comparison of the parametric null distribution function (Eq. (15)) against the nonparametric distribution (Eq. (16)) computed using both systems with $M_0 = 1000$ independent Bernoulli series. As expected from the theory, the three distributions are very similar to each other (see Fig. 2(c)). The p -value pairs of the test statistics are $p_P(t_F) = 0.620$, $p_N(t_F) = 0.592$, and $p_P(t_W) = 0.523$, $p_N(t_W) = 0.505$. All these p -values give strong support in favor of the null hypothesis. If we choose, for instance, the significance level $\alpha = 0.05$, Eq. (15) gives the threshold value $t_{0.05} = 0.00276$ as the solution for $p_P(t_\alpha) = \alpha$.

Therefore, we accept the null hypothesis and conclude that the signal is nonperiodic. In fact, we can choose any level $\alpha < 0.505$.

The second example in Fig. 3 is already a more realistic one and it is closely related to the respiratory signal problem which will be described more accurately in the next section. Here the signal \mathbf{x} of length $N = 8192$ is constructed by placing a number of bursts of 1's in the signal. One burst contains 16 1's in the successive entries. We focus only on the indices in the neighborhood of $\lceil 8192/1000 \rceil = 8$ because 1000 points (when sampled with 250 Hz) corresponds to the average duration of an adult respiratory cycle. More precisely, we define that $K = \{1, 2, \dots, 13\}$. In the null hypothesis case the bursts are placed randomly. Although this process is not exactly white noise, the analytical null distribution can be used as a reference distribution, as can be seen from the comparison of the parametric Eq. (14) and nonparametric Eq. (16) null distributions in this case: some variations exists, but

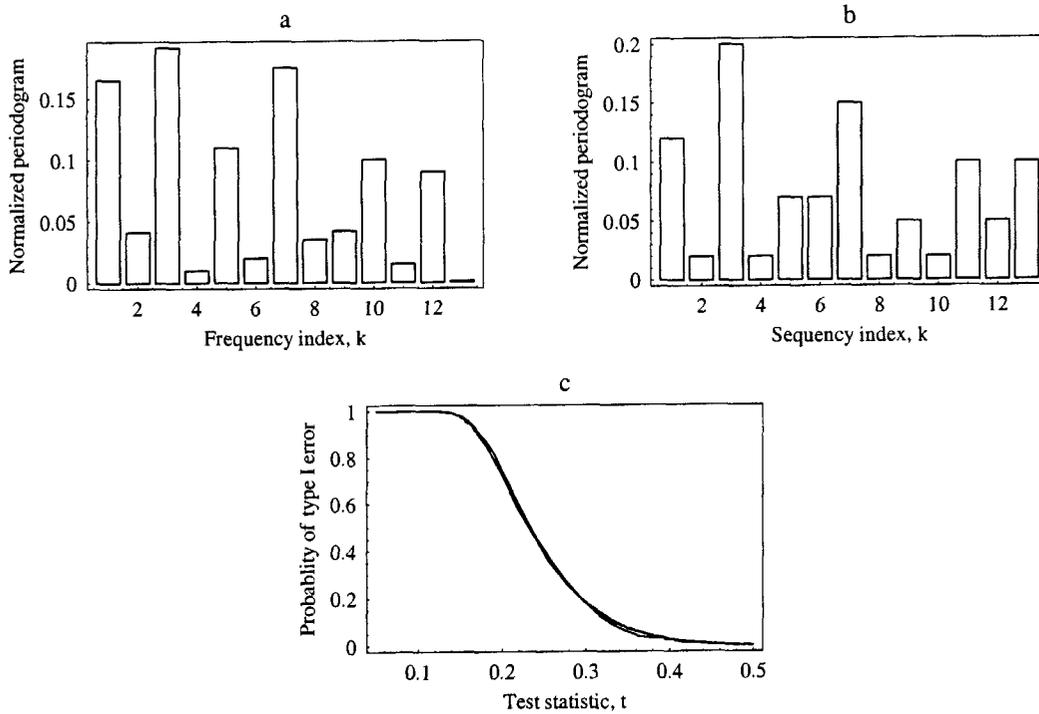


FIGURE 3 Simulation study of a random burst process with $K = \{1, \dots, 13\}$ (see text for details). (a) Normalized Fourier periodogram, and (b) normalized Walsh periodogram of the same sequence. (c) Comparison between parametric (Eq. (14)) and nonparametric (Eq. (16)) distributions.

the three distributions are again quite similar (see Fig. 3(c)). The normalized periodogram components of one realization of the process are shown in the top row of Fig. 3. The corresponding observed values of the test statistic are $t_F = 0.192$ (Fig. 3(a)) and $t_W = 0.200$ (Fig. 3(b)), which give the following probabilities of the type I error: $p_P(t_F) = 0.786$, $p_N(t_F) = 0.810$, and $p_P(t_W) = 0.728$, $p_N(t_W) = 0.723$. Thus, the maximum value of both periodograms is not significant at a level $\alpha = 0.05$ since the threshold value, computed now from the Eq. (14) with $l = 13$ is $t_{0.05} = 0.371$.

4.2 Periodic Sequences

As a final example we generate a set of time series under the alternative hypothesis H_1 . Here we let the distance between two bursts of 1's be a random variable which follows a normal distribution with mean 1000 and variance 100, 1000, or 10000. Fig. 4

shows the normalized periodograms of these processes for indices $k \in K$. It is seen that, when there are only moderate disturbances in the event occurrences (variance 100 and 1000), the maximum periodogram component is statistically significant at a level $\alpha = 0.05$ (Figs. 4(a)–(d)). However, the greater the variance the wider the set of indices the energy of the signal is distributed on. Here we can use the Eq. (17) to control the probability of type II error as well. If we use the threshold $t_{0.05} = 0.371$, we obtain the error probability $q_N(t_{0.05}) = 0$ for the cases where variance is 100 or 1000. This means that within $M_1 = 1000$ replicates the maximum periodogram in both systems is always larger than 0.371. In the case where the variance is 10000, however, the corresponding probabilities are 0.092 and 0.011 for the Fourier and Walsh systems, respectively. One such realization, where both periodograms are less than the threshold value, is shown in the bottom row of Fig. 4. If this signal describes a process under the

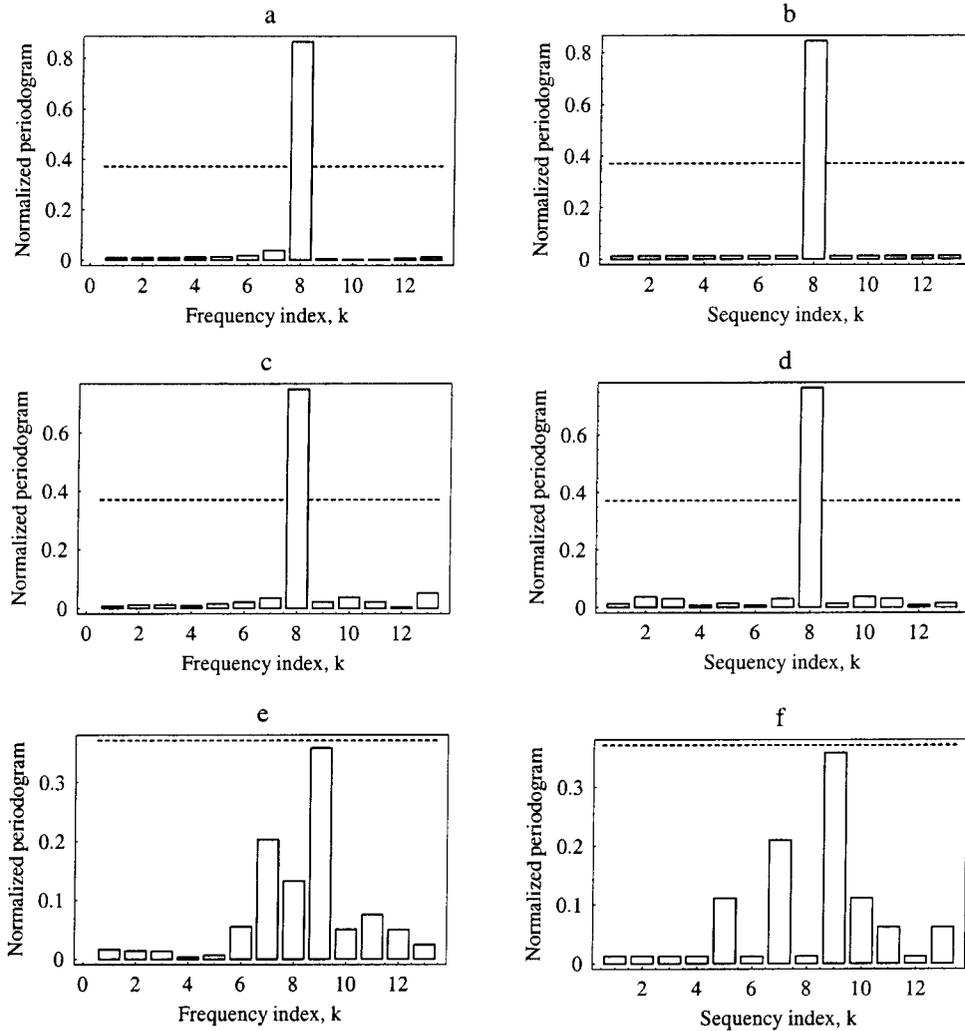


FIGURE 4 Simulation study of periodic signals. Each row represents a normalized Fourier (left column) and Walsh (right column) periodograms with an increasing variance (100, 1000, and 10000). Test statistic and its parametric significance are as follows: (a) $p_p(0.864) = 5.21 \times 10^{-10}$, (b) $p_p(0.844) = 2.70 \times 10^{-9}$, (c) $p_p(0.749) = 8.13 \times 10^{-7}$, (d) $p_p(0.764) = 3.88 \times 10^{-7}$, (e) $p_p(0.358) = 0.0637$, and (f) $p_p(0.358) = 0.0637$. Note the different scale in y-axis between the rows.

alternative hypothesis, then with the chosen threshold value, we have a positive probability for both types of errors.

These few examples should clarify that the main problem in the hypothesis testing procedure is the determination of the threshold value, which in turn determines the error probabilities. Whereas the type I error is easily controlled by the flexible analytical null distributions (Eqs. (14) and (15)), or by using the

Monte Carlo approximation (Eq. (16)), the control of probability of type II error is more complicated because it depends on the alternative hypothesis. Naturally we can no simulate all kinds of periodic signals in order to obtain the exact distribution $P(T \leq t | H_1)$ but with a large set of training examples one obtains a good approximation of the distribution by Eq. (17). This approach will be described in the next section. The simulations demonstrated also that both

the Fourier and Walsh periodogram analysis gave similar statistical decisions.

5. RESULTS ON SCSB SIGNAL

The SCSB is a cardiorespiratory monitoring method introduced by Alihanka *et al.* (1981). Because no electrodes need to be attached to the subject, it is well adapted for sleep recordings. The critical part of the bed is a movement sensor placed under a standard foam plastic mattress. The sensor operates like a capacitor, the charge of which is modified by static charge layers, which are activated by body movements. The potential differences between the capacitor plates produce the raw signal. The sleep study included continuous all-night monitoring of two SCSB signals and other conventional signals (e.g. EEG and ECG), which were all sampled at a frequency of 250 Hz. The present analysis focuses only on the high-frequency band (HFB) of the SCSB signal describing the frequency range of 6–16 Hz which is suitable for detection of the HFRM events. Fig. 5 shows some examples of parts of a HFB signal. During quiet breathing the HFB signal reflects the mechanical activity of the heart. The systolic waves are generated by the minor body movements forward due to rapid acceleration of blood in the ascending aorta (Fig. 5(a)). Increased respiratory efforts generate HFRM patterns that are superimposed on the cardiac activity (Fig. 5(c)). These events occur at regular intervals with distance of approximately 4 s, which is the average duration of an adult respiratory cycle.

Applying the Fourier periodogram analysis to the HFB signal is the first natural choice in order to detect the clusters of periodic HFRM events. In Fig. 5 we show the problems involved in this approach. As expected, there exists a peak in the Fourier periodogram at index $k = 8$ when the signal contains periodic HFRM patterns (Fig. 5(d)), but the other periodogram components are quite large as well. The problem is that the signal with no HFRM events may produce major peaks to the Fourier periodogram (Fig. 5(b)). In fact, in these example signals, the observed

value of the test statistic is larger in the nonperiodic case than in periodic case. Since the continuous HFB signal presents all kinds of periodic activities, the signal that contains either irregular HFRM events or no events at all is not a white noise signal in the Fourier sense. This kind of background signal is often called as colored noise. Moreover, some of these periodicities occur at the frequency band centered around $k = 8$, and therefore the procedure presented in section 3 is unsuitable for the detection problem in question. There have been several attempts to overcome this problem with so called *mixed-spectra procedures* (Priestley, 1981) which try to estimate simultaneously the spectrum of the background noise. However, the spectrum is usually approximated by smoothing the periodogram, and therefore these procedures are very uncertain in the situations where the periodogram of nonstationary background noise possesses large peaks.

To avoid the difficulty mentioned above, we took another, though a little bit more complicated approach to the problem. The first step was the preprocessing of the continuous HFB signal in order to identify each HFRM pattern (Aittokallio *et al.*, 2000). The preprocessing gives us a binary event sequence, where a burst of 1's indicates a HFRM pattern (Fig. 5(e)). Inspired by the simulation results of the previous section, we use the Walsh periodogram (Fig. 5(f)) in identifying the clinically relevant episodes of IRR during all night monitoring. More precisely, we compute the DWT of each window of length $N = 8192$ points in the binary sequence by the FWT routine of Beer (1981). Then the test statistic in Eq. (13) of the window is determined as the maximum of normalized Walsh periodogram components (Eq. (9)) within the set $K = \{1, 2, \dots, 13\}$. In order to examine the practical recognition performance of the method, we implemented these phases using the programming language C++. Although no particular attention was paid to the optimization of the algorithm, a normal PC (300 MHz Pentium) was sufficient and enabled real time cluster identification.

The rest of this section describes the selection of a good threshold value for the above algorithm (training), and the evaluation of the method (testing)

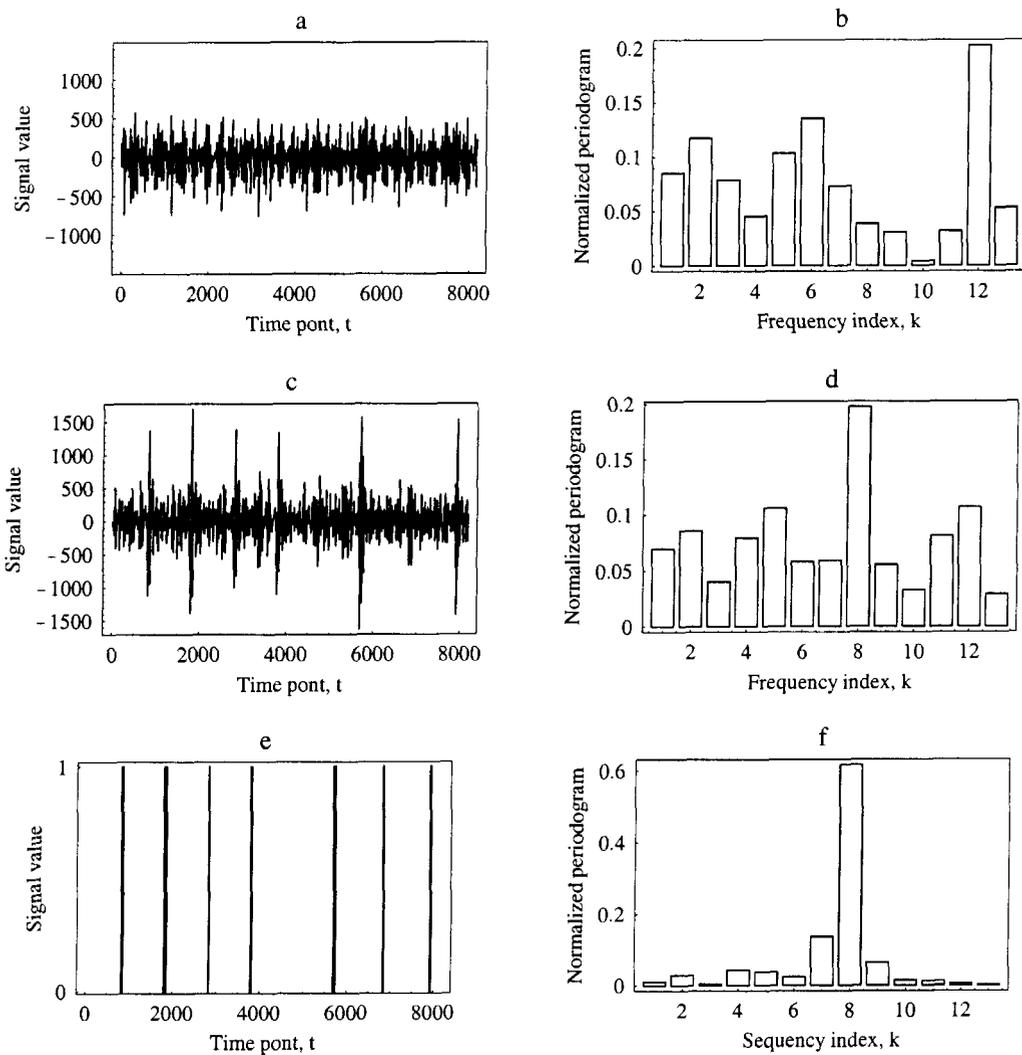


FIGURE 5 Examples of a real HFB signal. (a) Signal during quiet breathing, and (b) its normalized Fourier periodogram. Test statistic is $t_F = 0.202$, and its significance is $p_P(t_F) = 0.714$. (c) Signal with periodic HFRM patterns, and (d) its normalized Fourier periodogram. Test statistic is $t_F = 0.197$, and its significance is $p_P(t_F) = 0.749$. (e) Preprocessed binary sequence, and (f) its normalized Walsh periodogram. Test statistic is $t_W = 0.616$, and its significance is $p_P(t_W) = 0.000134$.

using independent test data. The training and test data were collected from three male patients with previous sleep apnea, which had been treated with surgery of the soft palate. After operation, all subjects improved in terms of sleep apnea frequency but continued to suffer from partial upper airway obstruction, which manifested as prolonged episodes of IRR in the SCSB recording. The persistence of partial obstruction after upper airway surgery is typical and has been described earlier by Polo *et al.* (1989). All subjects had the

clinical symptoms of snoring and excessive sleepiness. The patient characteristics with the conventional arterial oxyhemoglobin saturation (SaO_2) analyses are shown in Table 1. The SaO_2 was measured with a finger pulse oximeter. The frequency of sleep apnea episodes with significant ($\geq 4\%$ units) decreases in SaO_2 are presented as the oxyhemoglobin desaturation index (ODI_4). The patients represented a wide range of ODI_4 from 0 (only IRR) to 55.5 (IRR in conjunction of severe OSA).

5.1 Training

The threshold value of the algorithm was determined by utilizing the visual judgment of an expert physician, i.e., we chose an appropriate value so that the method identifies clusters that coincide with the clusters marked by the expert. The expert markings were made using a special-purpose graphical software. Two simple rules were employed when marking a cluster of HFRM events as periodic one. First, the length of the cluster must be more than $8192/250 \approx 30$ s, and secondly, the length between two successive periodic clusters must be more than 30 s as well. The goodness of a threshold value t was determined through the estimated error probabilities of the method based on the training data. Formally, let $\{t_1, t_2, \dots, t_{L-N+1}\}$ be the set of observed values of the test statistic Eq. (13) computed from the training sequence on a window-by-window basis. Here $L = 432 \times 60 \times 250 = 6480000$ is the length of the sequence. We can estimate the probability of type I error from the formula for $p_N(t)$ in Eq. (16), where $M_0 = 4592021$ is the number of points outside the periodic clusters marked by the expert. We use the number of the points i outside the clusters as the numerator which also satisfy the inequality $i - \max\{j|j \leq i, t_j > t\} < N$ to indicate that the periodicity of a window involves all points within it. (The numerator is sometimes known as the number of *false positive detections*.) Likewise, the probability of type II error can be estimated using the Eq. (17) for $q_N(t)$ with $M_1 = L - M_0 = 1887979$ and similar modification in the numerator (*false negative detections*). This type of Monte Carlo analysis produces very rough estimates, but the large number of sample signals under the both hypotheses justifies the approximation.

The nonparametric reference function $r_N(t)$ for the goodness of t is defined here as the average of the error probabilities:

$$r_N(t) = [p_N(t) + q_N(t)]/2. \quad (18)$$

The values $1-p$ and $1-q$ are called as the *specificity* and *sensitivity* of the method, respectively. Therefore, minimization of Eq. (18) corresponds to maximization of equally weighted specificity and sensitivity values. We perform this by a simple grid point method, see Fig. 6. Probability $p_N(t)$ is always a decreasing function of t , and $q_N(t)$ an increasing function of t (Fig. 6(a)). Note that $p_N(t)$ is again surprisingly close to the asymptotic distribution (Eq. (14)) (see Fig. 3(c)). The value $t_m = \operatorname{argmin}\{r_N(t)|t = 0, 0.02, \dots, 1\} = 0.34$ is selected as the threshold value of the method (Fig. 6(b)). This value yields the probabilities $p_N(t_m) = 0.0749$, $q_N(t_m) = 0.0645$, and $r_N(t_m) = 0.0697$. The reference function $r_N(t)$ remains relatively constant in the neighborhood of t_m from which we can conclude that the method is not very sensitive to the threshold value.

5.2 Testing

We treat the test data of the two subjects separately in order to study the inter-subject bias of the method as well. The recognition performance based on the the test subjects 1 and 2 is shown in Figs. 7 and 8, respectively. The length of the signal of test subject 1 was $L = 5670000$ which was divided to the points inside the periodic clusters marked by the expert ($M_1 = 3679602$) and to the points outside the clusters ($M_0 = 1990398$). The identifications made by the

TABLE I The characteristics of the training and test subjects with the conventional diagnostic measurements and frequencies of the visually scanned IRR episodes

Subject	Age (years)	BMI* (kg/m ²)	TIB† (min)	SaO ₂ mean(%)	ODI ₄ (events/h)	IRR‡ (min)
Training	51	28.7	432	95.4	0.0	29 (126)
Test 1	52	31.7	378	91.3	7.9	42 (245)
Test 2	57	35.7	474	89.6	55.5	49 (212)

* Body mass index.

† Time in bed.

‡ Number (duration) of IRR clusters.

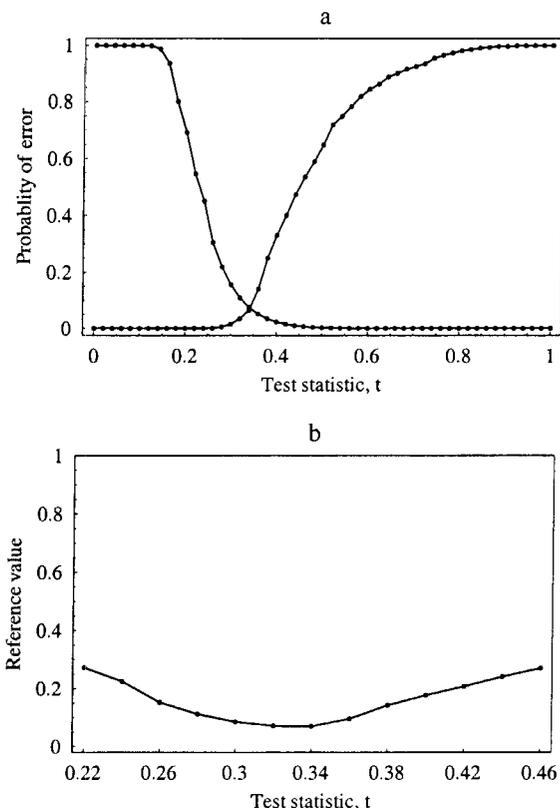


FIGURE 6 Training the method. (a) Probabilities of type I error (decreasing function) and type II error (increasing function). (b) Statistic value $t_m = 0.34$ is a local minimum of the reference function (Eq. (18)). The grid size is $\Delta t = 0.02$.

method correspond closely to those of the expert (Fig. 7). Here the detections of the method (i.e. $t_i > t_m = 0.34$) are extended to all points within the window as described in the previous section. The overall error probabilities, computed point-wise like in the training phase, were $p_N(t_m) = 0.138$ and $q_N(t_m) = 0.0712$. Thus, the overall goodness value of the method is $r_N(t_m) = 0.106$ when evaluated using the test subject 1. The recognition performance in the case of test subject 2 was consistent with the test subject 1 (Fig. 8). Here $L = 7110000$, $M_0 = 3930056$, and $M_1 = 3179944$. The corresponding probability values for the test subject 2 were $p_N(t_m) = 0.147$, $q_N(t_m) = 0.0722$, and $r_N(t_m) = 0.110$. The point-by-point analysis is naturally a very crude performance measure since

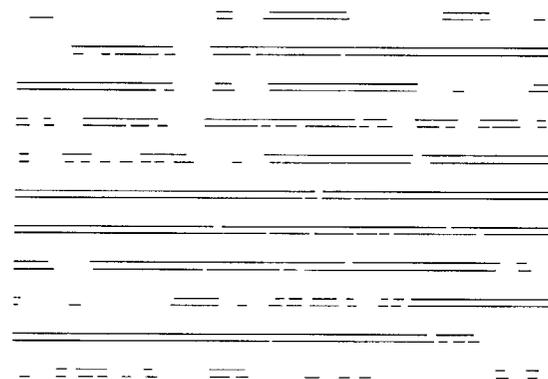


FIGURE 7 The markings of the expert (upper line) and of the method (lower line) for the test subject 1. Each row represents 500000 points, i.e. 2000 s.

it punishes for every erroneous instance in a 250^{-1} scale, but it is in its simplicity quite illustrative.

Although the probability of type I error (or the specificity) weakens from the training subject to the test subjects, the two test subjects produced almost identical error probabilities. The probability of type II error (sensitivity) is similar in all three subjects. In summary, the statistical method has a very good sensitivity to identify the same periodic clusters as marked by the expert physician, but it is inclined to identify some extra clusters as well. The test also suggests, that the inter-subject bias is negligible at least for the two test subjects.

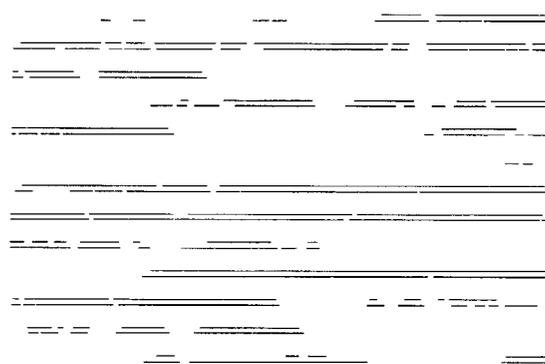


FIGURE 8 The markings for the test subject 2 (see the legend of Fig. 7). It is seen that the method usually identifies wider clusters than the expert which weakens the specificity value. We note that there occur only two predictions of the method in the row number five, but no markings by the expert.

6 DISCUSSION

We have described the statistical properties as well as the practical considerations of the problem of periodicity detection using an orthonormal transform together with hypothesis testing. The statistical inference included the derivation of the parametric null distribution for the test statistic and its comparison to a nonparametric one using Monte Carlo analysis in different situations (see Figs. 2 and 3). Although the DFT is the most common transform, the simulation studies demonstrated that the DWT has a similar detection power in the case of binary sequences (Fig. 4). The aim of the medical application was to analyze the SCSB signal with respect to the presence or absence of repetitive HFRM patterns which is a manifestation of an IRR period (Fig. 5). The practical implementation issues were the use of the procedure in a long SCSB signal by moving window, the selection of a good threshold value based on a set of training examples (Fig. 6), and the evaluation of the method using the test data (Figs. 7 and 8).

There are at least two aspects deserving still more attention in the basic periodogram analysis described in this work. First, the power of the hypothesis testing procedure decreases when the true periodicity falls midway between two consecutive spectral indices, and second if multiple periodicities are present the test suffers as well. These situations have received some interest in the statistical literature. To overcome the multiple peak problem, several alternatives for the test statistic in Eq. (13) have been suggested, including the use of r th largest (Hannan, 1970) or all large periodogram values (Siegel, 1980) in the numerator, and the use of winsorized (Bølviken, 1983a) or trimmed (Chiu, 1989) mean in the denominator. Also several works have proposed searching for the maximum of the periodogram among all possible periodicities, see e.g. the technique introduced by Chen (1988). Although all these investigations are based on the Fourier analysis, the results are extendable for other orthonormal function systems as well.

Although the number of training samples were large, the selection of the threshold value was based

on one subject only. One can increase the number of signals under the null hypothesis by the binary bootstrap method of Kim *et al.* (1993), but the problem is the generation of independent signals under the alternative hypothesis. Also, the evaluation of the method was performed using only two test subjects. This is because of the very tedious and time-consuming visual detection of periodic clusters. However, although the test recordings represented a wide range of variation in the manifestation of sleep-disordered breathing (see Table 1), the subject-to-subject variation in the recognition performance was found negligible. Therefore, the method based on the Walsh periodogram is considered good enough to warrant further testing in larger patient populations.

Acknowledgements

The authors thank PhD Jaakko Järvi for the software to perform the visual detection of sleep disturbances.

References

- Aittokallio, T., Nevalainen, O., Pursiheimo, U., Saaresranta, T. and Polo, O. (1999) "Classification of nasal inspiratory flow shapes by attributed finite automata", *Computers and Biomedical Research* **32**, 34–55.
- Aittokallio, T., Gyllenberg, M., Järvi, J., Nevalainen, O. and Polo, O. (2000) "Detection of high-frequency respiratory movements during sleep", *Computer Methods and Programs in Biomedicine* **61**, 171–185.
- Aittokallio, T., Saaresranta, T., Polo-Kantola, P., Nevalainen, O. and Polo, O. (2001) "Analysis of inspiratory flow shapes in patients with partial upper airway obstruction during sleep", *Chest* **119**, 37–44.
- Alihanka, J., Vaahtoranta, K. and Saarikivi, I. (1981) "A new method for long-term monitoring of the ballistocardiogram, heart-rate and respiration", *American Journal of Physiology* **240**, 384–392.
- Andrews, H.C. and Caspari, K.L. (1970) "A generalized techniques for spectral analysis", *IEEE Transactions on Computers* **19**, 16–25.
- Beer, T. (1981) "Walsh transforms", *American Journal of Physics* **49**, 466–472.
- Bølviken, E. (1983a) "New test of significance in periodogram analysis", *Scandinavian Journal of Statistics* **10**, 1–9.
- Bølviken, E. (1983b) "The distribution of certain rational functions of order statistics from exponential distributions", *Scandinavian Journal of Statistics* **10**, 117–123.
- Chechetkin, V.R. and Turygin, A.Y. (1995) "Search of hidden periodicities in DNA sequences", *Journal of Theoretical Biology* **175**, 477–494.

- Chen, Z.-G. (1988) "Consistent estimates for hidden frequencies in a linear process", *Advances in Applied Probability* **20**, 295–314.
- Chiu, S.-T. (1989) "Detecting periodic components in a white Gaussian time series", *Journal of the Royal Statistical Society B51*, 249–259.
- Choffrut, C. and Karhumäki, J. (1997) "Combinatorics of words", In: Rozenberg, G. and Salomaa, A., eds, *Handbook of Formal Languages* (Springer, Berlin) **1**, pp 329–438.
- Coward, E. and Drabløs, F. (1998) "Detecting periodic patterns in biological sequences", *Bioinformatics* **14**, 498–507.
- Feller, W. (1968) *An Introduction to Probability Theory and Its Applications*, Third ed. (John Wiley and Sons, New York) **1**.
- Fisher, R.A. (1929) "Tests of significance in harmonic analysis", *Proceedings of the Royal Society, London* **A125**, 54–59.
- Hannan, E.J. (1970) *Multiple Time Series* (John Wiley & Sons, New York).
- Harmuth, H.F. (1969) "Applications of Walsh functions in communications", *IEEE Spectrum* **6**, 82–91.
- Hinich, M.J. (1982) "Detecting a hidden periodic signal when its period is unknown", *IEEE Transactions on Acoustics, Speech, and Signal Processing* **30**, 747–750.
- Kim, Y.B., Haddock, J. and Willemain, T.R. (1993) "The binary bootstrap: inference with autocorrelated binary data", *Communications in Statistics* **22**, 205–216.
- Kirjavainen, T. (1997) "High-frequency respiratory movements during sleep", *Annales Universitatis Turkuensis* **D263**, 1–99.
- Manly, B.F.J. (1997) *Randomization, Bootstrap and Monte Carlo Methods in Biology*, Second ed. (Chapman and Hall, London).
- Polo, O. (1992) "Partial upper airway obstruction during sleep", *Acta Physiologica Scandinavica* **145**, 1–118.
- Polo, O., Brissaud, L., Fraga, J., Déjean, Y. and Billiard, M. (1989) "Partial upper airway obstruction in sleep after uvulopalatopharyngoplasty", *Arch Otolaryngol Head Neck Surg* **115**, 1350–1354.
- Priestley, M.B. (1981) *Spectral Analysis and Time Series* (Academic Press, London) **2**.
- Rohatgi, V.K. (1984) *Statistical Inference* (John Wiley and Sons, New York).
- Siegel, A.F. (1980) "Testing for periodicity in a time series", *Journal of the American Statistical Association* **75**, 345–348.
- Tatum, L.G. (1996) "Control charts for the detection of a periodic component", *Technometrics* **38**, 152–160.
- Walsh, J.L. (1923) "A closed set of normal orthogonal functions", *American Journal of Mathematics* **45**, 5–24.
- Yuen, C.-K. (1972) "Remarks on the ordering of Walsh functions", *IEEE Transactions on Computers* **21**, 1451.

APPENDIX A

This appendix gives the theoretical background of parametric inference approach of the section 3. The results follow from the orthonormality property (Eq.

(1)) and from the fact that the first row of the matrix \mathbf{U} consists entirely of constants $1/\sqrt{N}$ both in Fourier system (Eq. (6)) and in Walsh system (Eq. (8)). These two properties give

$$\sum_{t=0}^{N-1} f_k(t) = \sqrt{N} \delta_{k0}, \quad k = 0, 1, \dots, N-1. \quad (\text{A1})$$

We recall that the random signal \mathbf{x} under the null hypothesis is generated from a white noise process; \mathbf{x} is a sequence of independently and identically distributed random variables with finite mean $E(x_t) = \mu$ and variance $\text{Var}(x_t) = \sigma^2$. Since the \mathbf{U} -transform is given by Eq. (3), we get by using Eq. (A1) that $E(\mathbf{c}) = \mathbf{U}E(\mathbf{x}) = (\mu\sqrt{N}\delta_{k0})_{k=0}^{N-1}$, and further $\text{Cov}(\mathbf{c}) = \mathbf{U}\text{Cov}(\mathbf{x})\mathbf{U}^T = \mathbf{U}\sigma^2\mathbf{I}\mathbf{U}^T = \sigma^2\mathbf{I}$ because of Eq. (2). Moreover, the central limit theory for variable distributions (Feller, 1968) states that \mathbf{c} has asymptotically a N -dimensional normal distribution with mean vector $E(\mathbf{c})$ and covariance matrix $\text{Cov}(\mathbf{c})$.

We next consider the random variables c_k/σ , where $k = 0, 1, \dots, N-1$. By the above derivations, these are asymptotically independent and have asymptotic normal distributions with mean $\mu\sqrt{N}\delta_{k0}/\sigma$ and unity variance. According to Eq. (9), the random variable p_k/σ^2 has asymptotically a chi-square distribution with two degrees of freedom (i.e. exponential distribution with mean 2) for all $k = 1, 2, \dots, N/2 - 1$. Likewise, $p_{N/2}/\sigma^2$ has asymptotically a chi-square distribution with one degree of freedom, and p_0/σ^2 has asymptotically a non-central chi-square distribution with one degree of freedom and non-centrality parameter $\mu\sqrt{N}/\sigma^2$. For simplicity, we do not test the first and last periodogram components p_0 and $p_{N/2}$, and therefore the null distribution (Eq. (14)) follows from the results on rational functions of order statistics of independent random variables from exponential distributions derived by Bølviken (1983b).