

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/227693912>

Hidden Frequency Estimation with Data Tapers

Article *in* Journal of Time Series Analysis · February 2000

Impact Factor: 0.78 · DOI: 10.1111/1467-9892.00177

CITATIONS

14

READS

15

3 authors, including:



[Ka Ho Wu](#)

The Chinese University of Hong Kong

15 PUBLICATIONS 57 CITATIONS

[SEE PROFILE](#)

HIDDEN FREQUENCY ESTIMATION WITH DATA TAPERS

Zhao-Guo Chen¹, Ka Ho Wu² and Rainer Dahlhaus³

Statistics Canada, The Chinese University of Hong Kong and Universität Heidelberg

SUMMARY

Detecting and estimating hidden frequencies have long been recognized as an important problem in time series. This paper studies the asymptotic theory for two methods of high-precision estimation of hidden frequencies (secondary analysis method and maximum periodogram method) under the premise of using a data taper. In ordinary situations, a data taper may reduce the estimation precision slightly. However, when there are high peaks in the spectral density of the noise or other strong hidden periodicities with frequencies close to the hidden frequency of interest, the procedures of detection of the existence and the estimation for the hidden frequency of interest fail if data are non-tapered whereas they may work well if the data are tapered. The theoretical results are verified by some simulated examples.

¹Time Series Centre, 3H Coats Building, Statistics Canada, Ottawa, Ontario, Canada K1A 0T6

²Department of Statistics, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong

³Institut für Angewandte Mathematik, Universität Heidelberg, Im Neuenheimer Feld 294, D-6900, Heidelberg, Germany

Key words and phrases: Central limit theorem; Frequency leakage; Fourier transformation; Law of the iterated logarithm; Periodogram; Secondary analysis.

1 Introduction

The model under consideration is

$$y(n) = s(n) + x(n). \quad (1.1)$$

where the observations $y(n)$ consists of two real components $s(n)$ and $x(n)$. Here $x(n)$ is the stochastic component called “noise” and $s(n)$ is the deterministic component called “signal” defined as

$$s(n) = \sum_{k=-K}^K A_k \exp(i\omega_k n) = \sum_{k=1}^K 2|A_k| \cos(\omega_k n + \varphi_k), \quad 0 < \omega_k < \pi, \quad (1.2)$$

where $A_k = |A_k|e^{i\varphi_k}$, $-\pi < \varphi_k \leq \pi$, $\omega_0 = 0$ and $A_0 = 0$. Each term in (1.2) is called a hidden periodicity, and ω_k is called a hidden frequency. This model finds wide applications in the fields of science, engineering and economics.

The ordinary assumption for $\{x(n)\}$ in the literature is that it is a linear process of the form

$$x(n) = \sum_{j=0}^{\infty} \theta_j \varepsilon(n-j), \quad \sum_{j=0}^{\infty} |\theta_j| < \infty, \quad \theta_0 = 1; \quad (1.3)$$

where $\varepsilon(n)$ are independent, or more generally, martingale differences satisfying

$$E(\varepsilon(n)|\mathcal{F}_{n-1}) = 0, \quad E(\varepsilon(n)^2|\mathcal{F}_{n-1}) = \sigma^2, \quad (1.4)$$

where \mathcal{F}_n are the σ -algebras of events determined by $\varepsilon(t)$, $t \leq n$.

In the literature, it has long been recognized that the key step for modelling data as (1.1) and (1.2) consists of two steps: detection and estimation of the ω_k . The history of this research topic may go back as early as Fisher (1929), and others like Hartley (1949), Grenander and Rosenblatt (1957). All these authors assumed $\{x(n)\}$ be white noise, i.e. $x(n) = \varepsilon(n)$. The methodology they used for detection is the following.

Define the finite Fourier transform and periodogram of $\{y(n)\}$ by

$$d_y(\omega) = (2\pi N)^{-1/2} \sum_{n=1}^N y(n) \exp(-i\omega n) \quad (1.5)$$

and

$$I_y(\omega) = |d_y(\omega)|^2 = (2\pi N)^{-1} \left| \sum_{n=1}^N y(n) \exp(-i\omega n) \right|^2 \quad (1.6)$$

respectively, where N is the number of observations. Similar notations are $d_x(\omega)$, $I_x(\omega)$ for $\{x(n)\}$. If $\{s(n)\}$ is absent, then $I_y(\omega) = I_x(\omega)$, and the periodogram ordinates $I_y(2\pi j/N)$, $j = 1, 2, \dots, [N/2]$, are independently distributed as $(\sigma^2/4\pi)\chi_2^2$, if $\{x(n)\}$ is Gaussian white noise. Therefore, the statistic $g = \{\max_j I_y(2\pi j/N)\} / \sum_l I_y(2\pi l/N)$ can be used to detect ω_k in which a value of g , larger than a preassigned critical value, indicates the existence of an ω_k around the frequency where $I_y(2\pi j/N)$ is the largest. Whittle (1952), Hannan (1961) and other authors extended the methodology to $\{x(n)\}$ being a linear series. See Priestley (1981) for details.

Unfortunately, there are two major drawbacks of the above method. Firstly, when an ω_k falls around the center between two successive $2\pi j/N$, the power becomes particularly

low. Secondly, when there are more than one ω_k , large contributions of other ω_j to the denominator of the statistic g may result in an unacceptably low power of the test. For redesigning a procedure detecting hidden frequencies, the background theory should not restrict $I_y(\omega)$ to $\omega = 2\pi j/N$, and the testing statistics should not depend on the existence of other hidden frequencies which keep away from the hidden frequency of interest.

Let $f(\omega)$ be the spectral density of $\{x(n)\}$ and $f(\omega) > 0$ for all ω . Under the conditions (1.3) with *iid* $\epsilon(t)$, Turkman and Walker (1984) derived an asymptotic probability about $\max_{\omega} I_x(\omega)/f(\omega)$ which may be used to develop some testing procedures with higher power than above reviewed methods. Another fundamental result given by An, Chen and Hannan (1983) is that, under conditions (1.3) and (1.4),

$$\limsup_{N \rightarrow \infty} \max_{\omega} I_x(\omega) / \{f(\omega) \log N\} \leq 1, \quad a.s. \quad (1.7)$$

This result motivated the following approach. Let

$$\hat{z}_j = I_y\left(\frac{\pi j}{N}\right) / \left\{ \hat{f}_N\left(\frac{\pi j}{N}\right) \log N \right\} \quad (1.8)$$

where $\hat{f}_N(\lambda)$ is a nonparametric estimate of $f(\lambda)$. We may conclude that there exists a hidden frequency ω_k around $\bar{\omega}_k = \pi j/N$ if \hat{z}_j exceeds the threshold $1 + \varepsilon$ for some small $\varepsilon > 0$. The problem is to find an adequate estimate \hat{f}_N which is not influenced by large values of $I_y(\lambda)$ at the hidden frequencies. Chen (1988a, 1988b) had proposed a method which may be called “three-leave-out”-estimator. Von Sachs (1993) used a more general peak-insensitive estimator for $f(\lambda)$. Furthermore, he already used data tapers. If one regards this procedure as a testing procedure a heuristic rule for the selection of ε in dependence on the significance level α is given in von Sachs (1993, Lemma 2). In general

it is known that the so obtained initial estimate fulfills

$$|\bar{\omega}_k - \omega_k| < \pi/N \quad a.s. \quad (1.9)$$

The second step consists of improving the initial estimate $\bar{\omega}_k$. One straightforward method is the value $\hat{\omega}_k$ maximizing the periodogram $I_y(\omega)$ in some neighbourhood of $\bar{\omega}_k$ which we simply shall call MP. The other method is the so called secondary analysis (SA). SA is quite an old method which is restated in Priestley (1981, p.413). Chen (1988a) pointed out that there was a hidden flaw in the procedure and made a revision.

Both methods offer estimates $\hat{\omega}_k$ of ω_k with precision $O(N^{-3/2})$ which is a considerable improvement of $\bar{\omega}_k$. For $\hat{\omega}_k$ obtained by MP, Hannan (1973) proved the following central limit theorem (CLT):

$$N^{3/2}(\hat{\omega}_k - \omega_k) \xrightarrow{d} \mathcal{N}(0, 6R_k^{-2}). \quad (1.10)$$

Hannan and Mackisack (1986) also proved the strong consistency, more precisely, the following law of the iterated logarithm (LIL):

$$\limsup_{N \rightarrow \infty} (N^3 / \log \log N)^{1/2} |\hat{\omega}_k - \omega_k| \leq 12^{1/2} R_k^{-1}, \quad a.s. \quad (1.11)$$

where

$$R_k = |A_k| / (2\pi f(\omega_k))^{1/2} \quad (1.12)$$

is called the *local signal to noise ratio at ω_k* . Notice that in this paper, the amplitude of a sinusoid is $2|A_k|$ [see (1.2)], while in above mentioned papers, 2 is dropped; so in (1.10), the coefficient of R_k^{-2} is 4 times larger and in (1.11), the coefficient of R_k^{-1} is 2 times larger than those in the above mentioned papers. For $\hat{\omega}_k$ obtained by SA, Chen

(1988a) proved a LIL:

$$\limsup_{N \rightarrow \infty} (N^3 / \log \log N)^{1/2} |\hat{\omega}_k - \omega_k| \leq 2M^{3/2} C_M R_k^{-1} \quad a.s. \quad (1.13)$$

where $M \geq 2$ is an integer and C_M is given by (2.25) in the next section. However, the CLT for this estimate has not been proved so far. This paper fills the gap.

Theoretically, as $N \rightarrow \infty$, the above classical results seem to be satisfactory. However, for finite N , the periodogram may be heavily biased due to strong peaks in the spectrum of the stochastic component or due to more than one hidden frequency in the periodic component. This effect is called leakage effect. As we will show in Section 3 this leakage may affect the estimates of the hidden frequencies discussed above.

To guard against such effects we suggest in this paper the use of data tapers for the estimation of hidden frequencies. In ordinary spectral estimation, data-tapers are known to be an effective tool in reducing the bias due to frequency leakage (cf. Dahlhaus, 1988, 1990).

Define a taper function $h_N(u)$ of order (l, κ) as in Dahlhaus (1988, Definition 5.1.). Loosely speaking, l is the degree of smoothness (in particular at the edges) while $\kappa > 0$ means that the portion of tapered data goes to zero with a certain rate. When $\kappa = 0$, we simply put $h_N(u) = h(u)$. Here, we list three important and well-known taper functions. In all cases, $h_N(u) = 0$, if $u \notin (0, 1]$.

(a) $l = 0, \kappa = 0$:

$$h_N(u) = h(u) = 1, \quad 0 < u \leq 1; \quad (1.14)$$

(b) $l = 1, 0 \leq \kappa < 1/2$ ($\rho_N = \rho N^{-\kappa}$):

$$h_N(u) = \begin{cases} 2u/\rho_N, & 0 \leq u \leq \rho_N/2, \\ 1, & \rho_N/2 \leq u \leq 1/2, \\ h_N(1-u), & 1/2 \leq u \leq 1; \end{cases} \quad (1.15)$$

(c) $l = 2, 0 \leq \kappa < 1/2$ ($\rho_N = \rho N^{-\kappa/3}$):

$$h_N(u) = \begin{cases} [1 - \cos(2\pi u/\rho_N)]/2, & 0 \leq u \leq \rho_N/2, \\ 1, & \rho_N/2 \leq u \leq 1/2, \\ h_N(1-u), & 1/2 \leq u \leq 1; \end{cases} \quad (1.16)$$

where ρ_N is the fraction of the data that are tapered. Letting ρ_N tend to zero at some slow rate seems to be realistic in most situations — in particular if one regards tapering as smoothing the break from “data” to “no data” at the edges. Asymptotically, data tapers with $\rho_N \rightarrow 0$ very often lead to fully efficient procedures (compare for example our Theorems 2.1 and 2.3 below) while on the other side the advantages of tapers can be retained (cf. the discussion in Section 3).

In Section 2, we prove some CLT and LIL for the SA estimate and the MP estimate. All discussions are under the premise of using data taper. The coefficient on the right hand side of (1.13) is reduced from 2 to 1. The performance of the SA and MP procedures, and the effect of data taper are compared.

When the performance of a procedure for detecting or estimating hidden frequencies with data taper is judged, there is a fact that the bias due to leakage disappears asymptotically. But this advantage of data taper cannot be reflected by a classical CLT or LIL. For this reason, we adopt in Section 3 the special asymptotic approach of Dahlhaus (1988, 1990) for spectra with strong peaks: The peaks are assumed to increase with the

sample size leading to different asymptotic results which seem to be more realistic for finite sample situations with strong peaks. By this approach, we show that data tapers diminish the leakage effect and that this advantage overcomes the increase of estimation variance derived from the classical CLT and LIL in Section 2 which may be a misleading in the situations of “strong peak” or “more than one hidden frequency”.

In Section 4, we provide some simulation results which verify our theoretical propositions and Section 5 offers some concluding remarks.

2 Asymptotic properties of estimates

For a taper function $h_N(u)$ we use the following notations:

$$H_r^{(N)}(\omega) = \sum_{n=1}^N \left\{ h_N\left(\frac{n}{N}\right) \right\}^r \exp(-i\omega n), \quad H_r^{(N)} = H_r^{(N)}(0), \quad h_r^{(N)} = \int_0^1 \{h_N(u)\}^r du. \quad (2.1)$$

We assume that the taper function is symmetric about 0 with $h_N(0) = h_N(1) = 0$. The taper (1.14) which belongs to the non-tapered case does not fulfill this. However, all results proved below also hold for this case since the contribution resulting from the observation $y(N)$, [$u = 1$ in (1.14)] can be neglected asymptotically.

Furthermore, we always assume in this section that the taper is of the order $(l, 0)$. Setting $\kappa = 0$ is not a substantial restriction because if $\kappa > 0$, the situation is closer and closer to the non-tapered case as $N \rightarrow \infty$, i.e. $h_N(u)$ converges to (1.14). However, the case $\kappa = 0$ is easier to treat theoretically.

If $\kappa = 0$, $h_N(u)$ and $h_r^{(N)}$ are independent of N . We simply denote them by $h(u)$ and

h_r respectively. Obviously, for large N ,

$$H_r^{(N)} \doteq Nh_r. \quad (2.2)$$

2.1 How to use data tapers to SA

For the sake of simplicity, we concentrate our discussion on the hidden frequency ω_1 , and $\bar{\omega}_1$ is an initial estimate of ω_1 , satisfying (1.9). The main steps of SA are (see Chen 1988a):

(1) Choose a small integer $M \geq 2$ and divide the data of size N into M equal segments. As M is small, we may throw away the last few data points until N is divisible by M . On the other hand, note that $h_N(1) = 0$, so $y(N)$ is always diminished to zero and hence sometimes we may add $\{y(N)\}$ of any values to make N being divisible by M . In either case, we always can assume that M divides N . Let

$$\xi(s) = \xi_1(s) - i\xi_2(s) = \sum_{n=(s-1)N/M+1}^{sN/M} y(n) \exp(-i\bar{\omega}_1 n), \quad s = 1, 2, \dots, M \quad (2.3)$$

(2) Denote

$$z(s) = \arg(\xi(s)) \quad (2.4)$$

which takes values in, say, $[0, 2\pi)$ or $[-\pi, \pi)$. Form a linear regression model

$$z(s) = \alpha + \beta s + e(s), \quad s = 1, 2, \dots, M. \quad (2.5)$$

(3) The LSE of β is

$$\hat{\beta} = \frac{\sum_{s=1}^M z(s) \left(s - \frac{M+1}{2}\right)}{\sum_{s=1}^M \left(s - \frac{M+1}{2}\right)^2}, \quad (2.6)$$

and the final estimate $\hat{\omega}_1$ of ω_1 is

$$\hat{\omega}_1 = \bar{\omega}_1 + \hat{\beta}M/N, \quad \text{i.e.} \quad \hat{\beta} = (\hat{\omega}_1 - \bar{\omega}_1)N/M. \quad (2.7)$$

Chen (1988a) pointed out that step (2) should be revised as the result depends on the choice of an interval of length 2π . Using the “mean direction” $\bar{\beta}$, he puts $\xi'(s) = \xi(s) \exp(-i\bar{\beta}) = \xi'_1(s) + i\xi'_2(s)$ and replaces $z(s)$ in (2.4) by

$$z(s) = \arg(\xi'(s)), \quad -\pi < z(s) < \pi. \quad (2.4')$$

It can be shown that, for large N , values of $z(s)$ are within or around the interval $[-\pi/2, \pi/2]$. (See the end of Section 3, Chen 1988a). In this interval, 0 corresponds to the “main direction”.

We now introduce data taper to every segment of data. Suppose $h(u)$ is an ordinary taper function, put

$$h(u, s) = \begin{cases} h(Mu - (s - 1)), & (s - 1)/M < u \leq s/M, \\ 0, & \text{otherwise,} \end{cases} \quad (2.8)$$

$s = 1, 2, \dots, M$. Then instead of (2.3), we have

$$\xi(s) = \sum_{n=1}^N h\left(\frac{n}{N}, s\right) y(n) \exp(-i\bar{\omega}_1 n), \quad s = 1, 2, \dots, M. \quad (2.3')$$

Using (2.3'), (2.4'), (2.6) and (2.7), we obtain $\hat{\omega}_1$.

For taper (2.8), corresponding to (2.1), we may define $H_{r,s}^{(N)}(\omega)$ from $h(n/N, s)$ and

$$H_{r,s}^{(N)} = H_{r,s}^{(N)}(0) = \sum_{n=1}^N \left\{ h\left(\frac{n}{N}, s\right) \right\}^r = \sum_{n=1}^{N/M} \left\{ h\left(\frac{Mn}{N}\right) \right\}^r = H_{r,1}^{(N)} \doteq \frac{N}{M} h_r. \quad (2.9)$$

In view of (1.1), (1.2) and (2.3'), and notice that $h(u)$ is symmetric about $1/2$,

$$\begin{aligned} \xi(s) &= A_1 \exp(is(\omega_1 - \bar{\omega}_1)N/M) \sum_{n=1}^{N/M} h(nM/N) \exp(-i(\omega_1 - \bar{\omega}_1)n) \\ &\quad + \sum_{k \neq 1} A_k \sum_{n=1}^N h\left(\frac{n}{N}, s\right) \exp(i(\omega_k - \bar{\omega}_1)n) + \sum_{n=1}^N h\left(\frac{n}{N}, s\right) x(n) \exp(-i\bar{\omega}_1 n) \end{aligned}$$

which can be rewritten as

$$\xi(s) = A_1 G \exp(i\beta s) + \sum_{k \neq 1} A_k H_{1,s}^{(N)}(\omega_k - \bar{\omega}_1) + \xi^x(s) \quad (2.10)$$

where

$$\beta = (\omega_1 - \bar{\omega}_1)N/M, \quad (2.11)$$

$$G = H_{1,1}^{(N)}(\omega_1 - \bar{\omega}_1) = \sum_{n=1}^{N/M} h(nM/N) \exp(-i(\omega_1 - \bar{\omega}_1)n), \quad (2.12)$$

and the definition of $\xi^x(s)$ is self-clear from (2.10) by referring to (2.3') with $x(n)$ replacing $y(n)$. It is worth noting that $\bar{\omega}_1$, β , G all depend on N .

Lemma 2.1 *For any $\bar{\omega}_1$ satisfying (1.9), as $N \rightarrow \infty$,*

$$\frac{N}{M} h' \leq |G| \leq H_{1,1}^{(N)} \doteq \frac{N}{M} h_1 \quad a.s. \quad (2.13)$$

where

$$h' \doteq \inf_{0 \leq \psi \leq \pi/M} \left\{ \left(\int_0^1 h(u) (\cos \psi u) du \right)^2 + \left(\int_0^1 h(u) (\sin \psi u) du \right)^2 \right\}^{1/2} > 0.$$

If $|\omega_1 - \bar{\omega}_1| = o(N^{-1})$ a.s., then $|G| \doteq (N/M)h_1$ a.s..

Proof: Suppose $|\omega_1 - \bar{\omega}_1| = \psi/N$ with $|\psi| < \pi$. For $n = 1, \dots, N/M$, $0 \leq |\omega_1 - \bar{\omega}_1|n = (\psi/N)n < \pi/M$ always holds. Hence the first inequality of (2.13) follows from (2.12) immediately. The second inequality and the approximation in (2.13) are obvious.

2.2 CLT for SA

Lemma 2.2 *Suppose $x(n)$ satisfies*

$$x(n) = \sum_{j=-\infty}^{\infty} \theta_j \varepsilon(n-j), \quad \sum_{j=-\infty}^{\infty} \theta_j^2 < \infty, \quad (2.14)$$

where $\varepsilon(n)$ are i.i.d. with $E\varepsilon(n) = 0$ and $E\varepsilon(n)^2 = \sigma^2$, $f(\omega) = (2\pi)^{-1}\sigma^2 |\sum \theta_j e^{-i\omega j}|^2$ is piecewise continuous and continuous at ω_1 . Put $\xi^x(s) = \xi_1^x(s) - i\xi_2^x(s)$. If $\bar{\omega}_1 \rightarrow \omega_1$ a.s., then

$$(2\pi H_{2,1}^{(N)})^{-1/2} (\xi_1^x(1), -\xi_2^x(1), \dots, \xi_1^x(M), -\xi_2^x(M))^\tau \xrightarrow{d} \mathcal{N}_2(0, (f(\omega_1)/2)\mathbf{I}_{2M}) \quad (2.15)$$

where \mathbf{I}_{2M} is the unit matrix of order $2M$ and τ denotes transposition.

Proof: The proof can be viewed as an application of Theorems 8 and 10 of Chapter IV in Hannan (1970) by taking (referring to examples therein)

$$y_{2s-1}^{(N)}(n) = h\left(\frac{n}{N}, s\right) \cos \bar{\omega}_1 n, \quad y_{2s}^{(N)}(n) = h\left(\frac{n}{N}, s\right) \sin \bar{\omega}_1 n.$$

We omit the details.

Remark: Notice that (2.14) is a weaker condition than (1.3). Alternative conditions for (2.15) to hold are given in Hannan (1970, p.226) by assuming $\{x(n)\}$ being uniform mixing and some others.

Denote $\alpha = \arg(A_1 G)$,

$$\eta(s) = \left\{ \sum_{k \neq 1} A_k H_{1,s}^{(N)}(\omega_k - \bar{\omega}_1) + \xi^x(s) \right\} / |A_1 G|, \quad (2.16)$$

and

$$\eta'(s) = \exp(-i(\alpha + \beta s)) \eta(s). \quad (2.16')$$

In view of (2.10), we have

$$\xi(s)/|A_1G| = \exp(i(\alpha + \beta s))\{1 + \eta'(s)\}, \quad (2.17)$$

$$\xi'(s)/|A_1G| = \exp(-i\bar{\beta})\{\xi(s)/|A_1G|\} = \exp(i(\alpha' + \beta s))\{1 + \eta'(s)\} \quad (2.17')$$

where $\alpha' = \alpha - \bar{\beta}$. Denote $\eta'(s) = \eta'_1(s) + i\eta'_2(s)$. Since all ω_k ($k \neq 1$) stay away from ω_1 (and hence from $\bar{\omega}_1$), $H_{1,s}^{(N)}(\omega_k - \bar{\omega}_1) = O(1)$. In view of Lemmas 2.1 and 2.2, and (2.16') [i.e. the transformation from $(\eta_1(s), \eta_2(s))$ to $(\eta'_1(s), \eta'_2(s))$ is orthogonal], asymptotically, we have

$$\frac{|A_1G|}{(2\pi H_{2,1}^{(N)})^{1/2}}(\eta'_1(1), \eta'_2(1), \dots, \eta'_1(M), \eta'_2(M))^\tau \xrightarrow{d} \mathcal{N}_2(0, \frac{f(\omega_1)}{2}\mathbf{I}_{2M}). \quad (2.18)$$

The scaling factor on the left hand side is $O(N^{1/2})$, that means the mean of both $\eta'_1(s)$ and $\eta'_2(s)$ is $o(N^{-1/2})$ and the standard deviations of both $\eta'_1(s)$ and $\eta'_2(s)$ are of order $O(N^{-1/2})$. With some elemental discussion and inequalities in the probability theory, it is easy to show the following property: If a random sequence, x_N , has mean $o(N^{-1/2})$ and standard deviation $O(N^{-1/2})$, then $N^{1/2}x_N^2 \xrightarrow{p} 0$.

Now, the right hand side of (2.17') can be viewed as $|1 + \eta'(s)| \exp\{i(\alpha' + \beta s + e(s))\}$, where

$$e(s) = \arg(1 + \eta'(s)), \quad (2.19)$$

which is the expression of $e(s)$ in (2.5). By Taylor expansion and the above mentioned property, one sees that $N^{1/2}(e(s) - \eta'_2(s)) \xrightarrow{p} 0$. Hence,

$$\frac{|A_1G|}{(2\pi H_{2,1}^{(N)})^{1/2}}(e(1), \dots, e(M))^\tau \xrightarrow{d} \mathcal{N}(0, \frac{f(\omega_1)}{2}\mathbf{I}_M). \quad (2.20)$$

In view of (2.9), Lemma 2.1 and notation (1.12), if $|\omega_1 - \bar{\omega}_1| = o(N^{-1})$ a.s., then

$$(N/M)(2\pi H_{2,1}^{(N)} / |A_1 G|^2) f(\omega_1) / 2 \xrightarrow{P} (h_2/h_1^2) R_1^{-2} / 2. \quad (2.21)$$

If it only holds $|\omega_1 - \bar{\omega}_1| < \pi/N$ a.s., h_1 may be replaced by a stochastic number between h_1 and h' defined in Lemma 2.1.

From (2.6) and (2.5), it is easy to show that

$$\hat{\beta} - \beta = \sum_1^M (s - \frac{M+1}{2}) e(s) \bigg/ \sum_1^M (s - \frac{M+1}{2})^2. \quad (2.22)$$

Therefore, $\hat{\beta} - \beta$ is asymptotically normal with $\text{Var}(\hat{\beta} - \beta) = 12 \text{Var}\{e(1)\} / \{(M-1)M(M+1)\}$. By (2.7) and (2.11), $\hat{\omega}_1 - \omega_1 = (\hat{\beta} - \beta)M/N$. Combining these results, we have the following theorem.

Theorem 2.1 *For model (1.1) and (1.2), where $\{x(n)\}$ satisfies conditions in Lemma 2.2, if the initial estimate $\bar{\omega}_1$ of ω_1 satisfies $|\omega_1 - \bar{\omega}_1| = o(N^{-1})$ a.s., then as $N \rightarrow \infty$, the estimate $\hat{\omega}_1$ obtained by SA satisfies*

$$N^{3/2}(\hat{\omega}_1 - \omega_1) \xrightarrow{d} \mathcal{N}(0, \{6M^2/(M^2 - 1)\}(h_2/h_1^2)R_1^{-2}). \quad (2.23)$$

In the case where $\bar{\omega}_1$ only satisfies (1.9), then asymptotically, the mean of $\hat{\omega}_1 - \omega_1$ is $o(N^{-3/2})$ and its asymptotic variance is between $N^{-3}\{6M^2/(M^2 - 1)\}(h_2/h_1^2)R_1^{-2}$ and $N^{-3}\{6M^2/(M^2 - 1)\}(h_2/h'^2)R_1^{-2}$.

Remark: (i) Usually, the first $\bar{\omega}_1$ obtained in the stage of detecting the hidden frequency with special form of $\pi j/N$, (j is an integer in $[1, N]$) only satisfies (1.9); however, as the theorem does not require $\bar{\omega}_1$ to be of any special form, we may use the resulted $\bar{\omega}_1$ from

the preliminary use of SA which satisfies $|\bar{\omega}_1 - \omega_1| = O(N^{-3/2}) = o(N^{-1})$ *a.s.* (see next subsection); i.e. when the SA procedure is iterated twice, (2.23) is assured.

(ii) The data taper introduces the extra factor h_2/h_1^2 into the asymptotic variance (and its square root into the upper bound of the LIL — see Theorem 2.2 below). By using Cauchy-Schwarz inequality it can easily be shown that $h_2/h_1^2 \geq 1$ with equality if and only if $h(u) = 1$ for $0 < u \leq 1$ *a.s.* which is the non-tapered case. Equality also holds for an asymptotically vanishing taper with $\rho_N \rightarrow 0$. Thus, we have an asymptotic loss of efficiency by using a non-vanishing data taper. Nevertheless, tapering may be very beneficial in certain situations as we will point out in Sections 3 and 4.

2.3 LIL for SA

Theorem 2.2 *For model (1.1) and (1.2), where $\{x(n)\}$ is defined by (1.3) and (1.4) with some further conditions as stated in Theorem 3.1 of Chen (1988a), if $|dh(u)/du|$ is uniformly bounded in $(0, 1)$ and the initial estimate $\bar{\omega}_1$ satisfies $|\bar{\omega}_1 - \omega_1| = o(N^{-1})$; then as $N \rightarrow \infty$, the estimate $\hat{\omega}_1$ obtained by SA satisfies*

$$\limsup_{N \rightarrow \infty} (N^3 / \log \log N)^{1/2} |\hat{\omega}_1 - \omega_1| \leq M^{3/2} C_M (h_2^{1/2} / h_1) R_1^{-1}, \quad a.s. \quad (2.24)$$

where

$$C_M = \sum_{i=1}^M \left| s - \frac{M+1}{2} \right| / \sum_{j=1}^M \left(s - \frac{M+1}{2} \right)^2. \quad (2.25)$$

In the case that $\bar{\omega}_1$ only satisfies (1.8), h_1 in the right hand side of (2.24) should be replaced by h' defined in Lemma 2.1.

Proof: It is trivial to introduce the factor $h_2^{1/2}/h_1$ into the formula. For that, all we have to do is replacing “ $h(u, s) = 1$, if $(s-1)/M < u < s/M$; or 0, otherwise” (that

gives $h_1 = h_2 = 1$) by a general taper function satisfying the conditions mentioned in the theorem and carry on the proof in Chen (1988a). The major task is then to reduce coefficient 2 in (1.13) to 1.

Consider [see (3.30) in Chen, 1988a]

$$\sum_{n=1}^N h\left(\frac{n}{N}, s\right)x(n) \exp(-i\bar{\omega}_1 n) = \sum_{n=1}^N h_N^{(*)}\left(\frac{n}{N}, s\right)x(n) \exp(-i\omega_1 n), \quad (2.26)$$

where

$$h_N^{(*)}(u, s) = h(u, s) \cos N\tilde{\omega}_1 u - ih(u, s) \sin N\tilde{\omega}_1 u \quad (2.27)$$

and $\tilde{\omega}_1 = \bar{\omega}_1 - \omega_1$. As $|\tilde{\omega}_1| < \pi/N$, so, $|N\tilde{\omega}_1| < \pi$ and hence $|dh_N^{(*)}(u, s)/du|$ are uniformly bounded in u and in N .

Lemma 3.1 in Chen (1988a) may be extended to the complex situation: $\varphi_N = \varphi_N^{(1)} - i\varphi_N^{(2)}$, $\psi_N = \psi_N^{(1)} - i\psi_N^{(2)}$, with

$$P(\limsup_{N \rightarrow \infty} |\psi_N(U_N)| \leq \limsup_{N \rightarrow \infty} \sup_{x \in K} |\varphi_N(z)|) = 1, \quad (2.28)$$

where $\varphi_N^{(j)}$ and $\psi_N^{(j)}$ are sequences of linear functional in $C[0,1]$. Similar to (3.21), (3.22) in Chen (1988a), put

$$\begin{aligned} \varphi_N(z) &= \int_0^1 \dot{z}(u) h_N^{(*)}(u, s) du, \\ \psi_N(z) &= \int_0^1 \dot{z}(u) \tilde{h}_N^{(*)}(u, s) du = \sum_{n=1}^N \left\{ z\left(\frac{n}{N}\right) - z\left(\frac{n-1}{N}\right) \right\} h_N^{(*)}\left(\frac{n}{N}, s\right), \end{aligned}$$

then $\|\psi_N\|$ is bounded in N , $\limsup_{N \rightarrow \infty} \sup_{z \in K} |\psi_N(z) - \varphi_N(z)| = 0$ where K is a compact subset of $C[0,1]$ with its elements $z(u)$ satisfying $z(0) = 0$ and $\int_0^1 \dot{z}(u) du = 1$ where $\dot{z}(u)$ denotes $dz(u)/du$. Hence the conditions of the lemma are all satisfied. Thus (2.28) gives

$$\limsup_{N \rightarrow \infty} \left| \sum_{n=1}^N \left\{ U_N\left(\frac{n}{N}\right) - U_N\left(\frac{n-1}{N}\right) \right\} h_N^{(*)}\left(\frac{n}{N}, s\right) \right|$$

$$\begin{aligned}
&\leq \limsup_{N \rightarrow \infty} \sup_{z \in K} \left| \int_0^1 \dot{z}(u) h_N^{(*)} \left(\frac{n}{N}, s \right) du \right| \\
&\leq \limsup_{N \rightarrow \infty} \sup_{z \in K} \left(\int_0^1 \dot{z}(u)^2 du \right)^{1/2} \left(\int_0^1 |h_N^{(*)}(u, s)|^2 du \right)^{1/2} \\
&= \limsup_{N \rightarrow \infty} \left(\int_0^1 h(u, s)^2 \{ (\cos N\tilde{\omega}_1 u)^2 + (\sin N\tilde{\omega}_1 u)^2 \} du \right)^{1/2} \\
&= (h_2/M)^{1/2} \quad a.s.
\end{aligned} \tag{2.29}$$

Therefore except introducing some factors of h_1 and h_2 , the coefficient $2 + \epsilon$ in (3.31) of Chen (1988a) may reduce to $1 + \epsilon$. The rest of the proof is exactly the same. There is a slip in Chen (1988a) of ignoring the condition $|\tilde{\omega}_1 - \omega_1| = o(N^{-1})$ for (2.24) to be true which has been taken into consideration in Lemma 2.1 of this paper. As it was pointed out in the remark right after Theorem 2.1, iterating the SA procedure once more is suggested.

2.4 CLT and LIL for MP

Apart from the notations in (2.1) and (2.2), we introduce the following notations:

$$\left. \begin{aligned}
\dot{h}_r &= \int_0^1 u \{h(u)\}^r du & \ddot{h}_r &= \int_0^1 u^2 \{h(u)\}^r du \\
\dot{H}_r^{(N)}(\omega) &= \sum_{n=1}^N \left(\frac{n}{N}\right) \{h(\frac{n}{N})\}^r \exp(-i\omega n), & \ddot{H}_r^{(N)}(\omega) &= \sum_{n=1}^N \left(\frac{n}{N}\right)^2 \{h(\frac{n}{N})\}^r \exp(-i\omega n), \\
\dot{H}_r^{(N)} &= \dot{H}_r^{(N)}(0) \doteq N\dot{h}_r, & \ddot{H}_r^{(N)} &= \ddot{H}_r^{(N)}(0) \doteq N\ddot{h}_r.
\end{aligned} \right\} \tag{2.30}$$

We do not go into the details of proving theorems, as Hannan (1973) and Hannan and Mackisack (1986) have already got the results for the non-tapered case. We just show, under the conditions given in this paper, what happens if a data taper is introduced.

Imitating Hannan (1973), put

$$q_N(\omega) = N^{-2} \left| \sum_{n=1}^N y(n) h\left(\frac{n}{N}\right) \exp(-i\omega n) \right|^2. \quad (2.31)$$

By expressing $y(n)$ by (1.1) and (1.2) and focusing our attention to ω_1 , under conditions of either Theorem 2.3 or Theorem 2.4 below, we may show that:

$$\begin{aligned} N^{-1/2} \frac{dq_N(\omega_1)}{d\omega} &= N^{-5/2} i \{ (A_1 N H_1^{(N)} \sum_1^N h\left(\frac{n}{N}\right) \frac{n}{N} x(n) \exp(i\omega_1 n) - \text{the conjugate}) \\ &\quad - (A_1 N \dot{H}_1^{(N)} \sum_1^N h\left(\frac{n}{N}\right) x(n) \exp(i\omega_1 n) - \text{the conjugate}) \} \\ &\quad + O(N^{-1/2}), \end{aligned} \quad (2.32)$$

$$N^{-2} \frac{d^2 q_N(\omega'_1)}{d\omega^2} = -2|A_1|^2 (h_1 \ddot{h}_1 - \dot{h}_1^2) + O(N^{-1/2}), \quad \text{if } |\omega'_1 - \omega_1| = o(N^{-1}); \quad (2.33)$$

where the meaning of $O(N^{-1/2})$ in (2.32) and (2.33) is “ $O(N^{-1/2})$ a.s.”, and $dq(\omega_1)/d\omega$ etc means $\{dq(\omega)/d\omega\}_{\omega=\omega_1}$ etc. By Taylor expansion,

$$N^{-1/2} \frac{dq_N(\omega_1)}{d\omega} = -N^{3/2} (\hat{\omega}_1 - \omega_1) N^{-2} \frac{d^2 q_N(\omega'_1)}{d\omega^2}, \quad (2.34)$$

where $\hat{\omega}_1$ in the MP estimate satisfying $dq_N(\hat{\omega}_1)/d\omega = 0$ and ω'_1 is a value between ω_1 and $\hat{\omega}_1$. Similarly to Lemma 2.2 with $h_N(u, s)$ replaced by $h_N(u)$ or $uh_N(u)$, we can show that the right hand side of (2.32) is asymptotically normal with mean 0 and variance $2|A_1|^2 (h_1^2 \ddot{h}_2 + \dot{h}_1^2 h_2 - 2h_1 \dot{h}_1 \dot{h}_2) (2\pi f(\omega_1))$. Combining (2.32) through (3.34), we have the following theorem.

Theorem 2.3 *For model (2.1) and (2.2), where $\{x(n)\}$ satisfies the conditions in Lemma 2.2, the solution of (2.34), $\hat{\omega}_1$, exists (in probability) in a neighbourhood of ω_1 , and*

$$N^{3/2} (\hat{\omega}_1 - \omega_1) \xrightarrow{d} \mathcal{N}(0, (\mu_1/2\mu_2^2) R_1^{-2}), \quad (2.35)$$

where

$$\left. \begin{aligned} \mu_1 &= h_1^2 \ddot{h}_2 + \dot{h}_1^2 h_2 - 2h_1 \dot{h}_1 \dot{h}_2, \\ \mu_2 &= h_1 \ddot{h}_1 - \dot{h}_1^2. \end{aligned} \right\} \quad (2.36)$$

The proof of the existence of the solution of (2.34) [i.e. there is an $\hat{\omega}_1$ satisfying (2.34)] is outlined as follows. Consider all ω in a neighbourhood of ω_1 , say $\{\omega : |\omega - \omega_1| < N^{-9/8}\}$, and $\delta > 0$ (say, $1/8$), in view of (2.33) and (2.36), put (2.34) in the form of

$$N^{3/2-\delta}(\omega - \omega_1) = \{N^{-(1/2+\delta)} dq_N(\omega_1)/d\omega\} / \{2|A_1|^2 \mu_2 + O(N^{-1/2})\}. \quad (2.37)$$

Notice that $d^2 q_N(\omega'_1)/d\omega^2$ is a function of ω (since ω'_1 depends on ω), we may put (2.37) as $D_1(N)/D_2(N, \omega)$. Then $P(|D_1(N)| \leq c) \geq 1 - \epsilon$ ($c > 0$ is a constant) for any $\epsilon > 0$ and all sufficiently large N [since $N^\delta D_1(N)$ is asymptotically normal with finite variance]. For fixed large N , with $|D_1(N)| \leq c$, since $D_2(N, \omega)$ is a continuous function of ω and takes values around $2|A_1|^2 \mu_2$, $D_1(N)/D_2(N, \omega)$ is a continuous bounded function of ω . But since $N^{3/2-\delta}(\omega - \omega_1)$ is a linear function and reaches $\pm N^{1/4}$ in $\hat{\omega}_1 \in (\omega_1 - N^{-9/8}, \omega_1 + N^{-9/8})$, there must be a solution $\hat{\omega}_1$ in this neighbourhood. Similarly, $\hat{\omega}_1$ exists a.s. in the following theorem.

Theorem 2.4 *For model (1.1) and (1.2), under the conditions of Theorem 2.2, the solution of (2.34), $\hat{\omega}_1$, exists almost surely in a neighbourhood of ω_1 and*

$$\limsup_{N \rightarrow \infty} N^{3/2} (\log \log N)^{-1/2} |\hat{\omega}_1 - \omega_1| \leq (\mu_1^{1/2} / \mu_2) R_1^{-1}, \quad a.s. \quad (2.38)$$

Proof: From (2.32), we have

$$|dq_N(\omega_1)/d\omega| \doteq 2|A_1| \left| \cos \varphi_1 \sum_1^N g\left(\frac{n}{N}\right) x(n) \cos \omega_1 n - \sin \varphi_1 \sum_1^N g\left(\frac{n}{N}\right) x(n) \sin \omega_1 n \right|$$

$$\begin{aligned}
&= 2|A_1| \left| \operatorname{Re} \left\{ e^{i\varphi_1} \sum_1^N g\left(\frac{n}{N}\right) x(n) \exp(i\omega_1 n) \right\} \right| \\
&\leq 2|A_1| \left| \sum_1^N g\left(\frac{n}{N}\right) x(n) \exp(-i\omega_1 n) \right|, \tag{2.39}
\end{aligned}$$

where $A_1 = |A_1| \exp(i\varphi_1)$, $g\left(\frac{n}{N}\right) = h_1 h\left(\frac{n}{N}\right) \frac{n}{N} - \dot{h}_1 h\left(\frac{n}{N}\right)$. We have $\int_0^1 g(u)^2 du = \mu_1$. However, (see Theorem 3.1 in Chen, 1988a)

$$\begin{aligned}
&\limsup_{N \rightarrow \infty} (2\pi N \log \log N)^{-1/2} \left| \sum_1^N g\left(\frac{n}{N}\right) x(n) e^{-i\omega_1 n} \right| \\
&\leq \left\{ f(\omega) \int_0^1 g(u)^2 du \right\}^{1/2} \\
&= f(\omega)^{1/2} \mu_1^{1/2} \quad a.s., \tag{2.40}
\end{aligned}$$

therefore (2.38) follows from (2.40) and (2.32) through (2.34).

Remark: Data taper introduces the extra factor $\mu_1/(12\mu_2^2)$ into the asymptotic variance of the CLT and its square root into the upper bound of the LIL [cf. (1.10) and (1.11)]. Below, we prove that this factor is larger than or equal to 1 with equality in the non-tapered case which leads to the same situation as for SA estimates [see Remark (ii) after Theorem 2.1].

To prove the inequality $\mu_1/(12\mu_2^2) \geq 1$, we first note that due to the symmetry of the taper about 1/2 we have $\dot{h}_1/h_1 = 1/2$. We then get with the Cauchy-Schwarz inequality

$$\begin{aligned}
\mu_2^2 &= h_1 \ddot{h}_1 - \dot{h}_1^2 = \frac{1}{h_1^2} \left[\int_0^1 (uh_1 - \dot{h}_1)^2 h(u) du \right]^2 \\
&\leq \left[\int_0^1 (uh_1 - \dot{h}_1)^2 h(u)^2 du \right] \frac{1}{h_1^2} \int_0^1 (uh_1 - \dot{h}_1)^2 du \\
&= \mu_1 \left[\int_0^1 \left(u - \frac{\dot{h}_1}{h_1} \right)^2 du \right] = \frac{\mu_1}{12}.
\end{aligned}$$

Theorems 2.3 and 2.4 give (1.10) and (1.11) respectively, where equality holds in (2.38) for the non-tapered case. In fact, under some other conditions (Theorem 3 in Chen, 1990),

it can be shown that the equality in (2.38) holds for general taper functions. For SA we do not know whether equality in (2.24) holds under some conditions.

2.5 Comparison

For SA, the effect of a data taper produces a factor h_2/h_1^2 in the asymptotic variance and a factor $h_2^{1/2}/h_1$ in the *a.s.* upper bound; while for MP, correspondingly, $\mu_1/(12\mu_2^2)$ and $\mu_1^{1/2}/(12^{1/2}\mu_2)$. For the non-tapered case (1.14), and the cases of using tapers (1.15) and (1.16) with $\rho_N = \rho = 1$ (full size tapers), we list these values in Table 2.1. For tapers with smaller ρ , the efficiency loss is smaller.

Taper	SA			MP		
	h_2	h_1	$h_2^{1/2}/h_1$	μ_1	μ_2	$\mu_1^{1/2}/(12^{1/2}\mu_2)$
(1.14)	1	1	1	0.0833	0.0833	1
(1.15)	1/3	1/2	1.15	0.00208	0.01042	1.27
(1.16)	3/8	1/2	1.22	0.00148	0.00820	1.36

Intuitively, in SA, we may regard $\xi(s)$ as a vector in the complex plane of length about $|A_1G| \doteq h_1N/M$ turning an angle of β every time for $s = 1, \dots, M$ with a disturbance which has mean $O(1)$ and standard deviation in each dimension about $(2\pi N_{2,1}^{(N)} f(\omega_1)/2)^{1/2} \doteq (h_2\pi N f(\omega_1)/M)^{1/2}$ [see (2.10), (2.15) and (2.9)]. When a data taper is used, the disturbances become smaller (smaller h_2) but the vector becomes shorter (smaller h_1). As the result of balancing, $h_2^{1/2}/h_1$ increases and hence the stochastic error of the actual turning angle from β increases by the data taper. That leads to a

larger estimation error of β and a larger estimation error of ω_1 .

For MP, μ_1 may represent the effect of tapers on the scale of disturbances, and μ_2 represents the effect on the sharpness of the peak generated by ω_1 (a sharper peak leads to a more accurate estimate). As the result of balancing, $\mu_1^{1/2}/\mu_2$ increases, and so tapering also increases the stochastic error. In ordinary situations [no high peaks in the spectrum of $\{x(n)\}$, no other close and strong hidden frequencies], “no taper” or “light taper” (say, $\rho = 0.1$) are recommended (cf. Section 4).

Now, the performance of SA and MP are compared for the non-tapered case. Table 2.2 lists the asymptotic standard deviation (A.S.D.) for CLT and the almost sure boundary (A.S.B.) for LIL where $M = 2, 3, 5, 7$. Since R_1^{-1} is common in all the formulae, only the coefficients are listed. Theoretically, M may take any integer value greater than 1.

Table 2.2 A.S.D. and A.S.B. in the non-tapered case

		SA				MP
		M	2	3	5	7
A.S.D.	$6^{1/2}M/(M^2 - 1)^{1/2}$	2.83	2.60	2.50	2.48	$\mu_1^{1/2}/(2^{1/2}\mu_2) = 2.45$
A.S.B.	$M^{3/2}C_M$	5.7	5.2	6.7	7.9	$\mu_1^{1/2}/\mu_2 = 3.46$
notes	C_M	2	1	3/5	3/7	$\mu_1 = \mu_2 = 1/12$

If we take A.S.D. as the measure of accuracy, we observe that a large value of M is preferred for the SA method which converges to the A.S.D. of the MP method (i.e. as M increases, $6^{1/2}M/(M^2 - 1)^{1/2} \rightarrow \mu_1^{1/2}/(2^{1/2}\mu_2) = 6^{1/2} \doteq 2.45$). However, all differences are insignificant. If we take A.S.B. as the measure of accuracy, the minimum for SA is attained at $M = 3$ and increases rapidly as M increases. The values are much larger than

the A.S.B. via MP. However, for SA, we are not sure whether this bound can be achieved or not. It could be a conservative bound. Anyway, a large M is not recommended for SA. From Table 2.2, we see MP is better than SA in ordinary situations.

In practice, the solution via MP is obtained by a grid search (say, in a finer lattice than $\pi j/N$) rather than by solving $dI_y(\omega)/d\omega = 0$ [$I_y(\omega)$ is given by (1.6) with tapers]. SA obtains the solution by a simple analytic formula which is computationally far more efficient. In our experience, for $N = 150$, SA only takes about 1/50 of the time used by MP.

3 The advantages of using a data taper

3.1 A model coping with leakage effect

Theorems 2.1 and 2.3 only show a disadvantage of using data taper, namely the increase of the asymptotic variance of the SA estimates and the MP estimates. This is typical when data tapers are used. A similar increase of the asymptotic variance can be observed in several other situations, for example for tapered kernel spectral estimates (cf. Dahlhaus, 1990) or for tapered Whittle estimates and tapered Yule-Walker estimates (cf. Dahlhaus, 1988). Nevertheless, tapering may lead to dramatic improvements of the bias of the estimates. Simulations which show these improvements for kernel spectral estimates and tapered Yule-Walker estimates can be found in Dahlhaus (1988, 1990). The bias of non-tapered estimates usually results from spectral leakage from strong peaks in the spectrum, or - from the periodic components. The leakage effect is a finite sample effect

and the classical asymptotic analysis therefore is not suitable for describing this effect.

For this reason Dahlhaus (1988) had introduced a different type of asymptotic analysis. In his approach, the magnitudes of peaks in the spectrum were allowed to increase with the sample size. With this (different) asymptotic theory, the leakage effect does not disappear asymptotically and the benefit of data tapers can be seen. Thus, the use of this theory reflects some problems for a fixed sample size in a much better way.

We basically follow the notation of Dahlhaus (1990) where he introduced a class of stationary processes $\mathcal{X}(N, s_1, s_2, \delta_0, c_0)$. An element, $\{x(n)\}$, in $\mathcal{X}(N, s_1, s_2, \delta_0, c_0)$ can be represented in the form

$$x(n) = \sum_{j=-\infty}^{\infty} \theta_j^{(N)} \nu(n-j) \quad (3.1)$$

where, $\{\nu(n)\}$ is an ordinary stationary series with spectral density satisfying $c_0^{-1} \leq f_\nu(\omega) \leq c_0$. Dahlhaus also requires stationarity up to 4th order and that the 4th order spectral density is bounded, but these are not necessary in our study. The transfer function $\Theta^{(N)}(\omega) = \sum_{j=-\infty}^{\infty} \theta_j^{(N)} e^{-i\omega j}$ can be written as

$$\Theta^{(N)}(\omega) = \prod_{j=1}^{r_1} \{\Theta_{1j}^{(N)}(\omega - \lambda_{1j})\}^{s_{1j}} \bigg/ \prod_{j=1}^{r_2} \{\Theta_{2j}^{(N)}(\omega - \lambda_{2j})\}^{s_{2j}} \quad (3.2)$$

with $s_{ij} \leq s_i$ ($i = 1, 2$), $|\lambda_{i_1 j_1} - \lambda_{i_2 j_2}| > 2\delta_0 \pmod{2\pi}$ for $(i_1, j_1) \neq (i_2, j_2)$. If we denote $|\Theta_{ij}^{(N)}(\lambda)|^2 = g_{ij}^{(N)}(\lambda)$, then

$$c_0^{-1} \{L_{N_{ij}}(\lambda)\}^2 \leq \{g_{ij}^{(N)}(\lambda)\}^{-1} \leq c_0 \{L_{N_{ij}}(\lambda)\}^2, \quad (3.3)$$

where $N_{ij} \leq N$ and $L_N(\lambda)$ is a function defined as

$$L_N(\lambda) = \begin{cases} N, & |\lambda| \leq 1/N, \\ 1/|\lambda|, & 1/N < |\lambda| \leq \pi. \end{cases} \quad (3.4)$$

This means that the spectral density may have peaks of order $O(N^{2s_2})$ and troughs of order $O(N^{-2s_1})$.

We do not try to convince readers that there must be realistic series like that, but we use this model as a tool to describe the situations that for a given N , the heights of peaks or the depths of troughs of the spectral density are competitive with N^{2s_2j} or N^{-2s_1j} and embed such situations into the model for theoretical discussion. We now use this model to describe the advantages of using a taper.

3.2 A property of data tapers

We may introduce data tapers in (1.5) and (1.6) to define $d_y(\omega)$ and $I_y(\omega)$. Similarly,

$$I_x(\omega) = |d_x(\omega)|^2 = (2\pi H_2^{(N)})^{-1} \left| \sum_{n=1}^N x(n) h_N\left(\frac{n}{N}\right) \exp(-i\omega n) \right|^2. \quad (3.5)$$

Here, we temporarily assume that $\{x(n)\}$ is observable. From (2.1), it is easy to show that

$$EI_x(\omega) = \int_{-\pi}^{\pi} f(\omega - \lambda) (2\pi H_2^{(N)})^{-1} |H_1^{(N)}(\lambda)|^2 d\lambda. \quad (3.6)$$

The classical result is that as $N \rightarrow \infty$, $EI_x(\omega) \rightarrow f(\omega)$. However this is not necessarily true for $\{x(n)\} \in \mathcal{X}(N, s_1, s_2, \delta_0, c_0)$.

Put $\lambda = \psi/N$, and denote the Fourier transform of $h_N(u)$ by

$$H_N(\psi) = (2\pi)^{-1} \int_{-\infty}^{\infty} h_N(u) \exp(-i\psi u) du \doteq (2\pi N)^{-1} H_1^{(N)}\left(\frac{\psi}{N}\right), \quad (3.7)$$

then we have

$$EI_x(\omega) \doteq \frac{2\pi}{h_2} \int_{-N\pi}^{N\pi} f\left(\omega - \frac{\psi}{N}\right) |H_N(\psi)|^2 d\psi. \quad (3.8)$$

Using integration by parts, if $h_N(u)$ is of order (l, κ) , and $d^j h_N(u)/du^j$ is a piecewise continuous function plus Dirac- δ function, then for $j \leq l + 1$,

$$H_N(\psi) = (i\psi)^{-j} \int_{0-}^{1+} (d^j h_N(u)/du^j) \exp(-iu\psi) du / (2\pi). \quad (3.9)$$

For (1.14), (1.15) and (1.16), we obtain respectively:

$$H_N(\psi) = H(\psi) = (i\psi)^{-1} \{1 - \exp(-i\psi)\} / 2\pi, \quad (3.10a)$$

$$H_N(\psi) = (i\psi)^{-2} \{1 - \exp(-i\psi)\} \{1 - \exp(i\rho_N \psi/2)\} / (\pi\rho_N), \quad (3.10b)$$

$$\begin{aligned} H_N(\psi) = & (i\psi)^{-3} \{ (1 - \exp(-i\psi))(1 + \exp(i\rho_N \psi/2)) \\ & - \int_0^{\rho_N/2} \sin \frac{2\pi u}{\rho_N} \exp(-iu\psi) du \left(\frac{2\pi}{\rho_N}\right) \\ & + \int_{1-\rho_N/2}^1 \sin \frac{2\pi(1-u)}{\rho_N} \exp(-iu\psi) du \left(\frac{2\pi}{\rho_N}\right) \} \frac{\pi}{\rho_N^2}. \end{aligned} \quad (3.10c)$$

In general, we may write formulae of (3.10) as

$$H_N(\psi) = \psi^{-(l+1)} \rho_N^{-l} \phi_l(\psi, \rho_N), \quad (3.11)$$

where $|\phi_l(\psi, \rho_N)|$ is bounded for all ψ and ρ_N . Unlike Dahlhaus (1988, 1990), we simply define the order of a taper by (l, ζ) , if $H_N(\psi)$ has an expression like (3.11) and $\rho_N = \rho N^{-\zeta}$, where $\zeta \geq 0$, usually, is very small. Now (3.11) may be written as

$$H_N(\psi) = \psi^{-(l+1)} N^{\zeta l} \tilde{\phi}_l(\psi, N^{-\zeta}), \quad |\tilde{\phi}_l(\psi, N^{-\zeta})| < c, \quad (3.12)$$

where, $c > 0$ is a constant. From (3.12), we see that larger l gives smaller $|H_N(\psi)|$ (for $|\psi| > 1$), and hence (3.8) is closer to $f(\omega)$. In particular (3.8) shows that $EI_x(\omega)$ is less affected from peaks of $f(\lambda)$ at frequencies λ different from ω .

3.3 Eliminating leakage

In the following, we denote the spectral density of $\{x(n)\} \in \mathcal{X}(N, s_1, s_2, \delta_0, c_0)$ by $f_N(\omega)$.

A simulation (Dahlhaus, 1990) has shown that if $\{x(n)\} \in \mathcal{X}(N, s_1, s_2, \delta_0, c_0)$, with non-tapered data, there is no indication in $I_x(\omega)$ to reflect small peaks of $f_N(\omega)$. But with data tapers, the small peaks of $f_N(\omega)$ can be discovered by $I_x(\omega)$. We now show more detailed why leakage from large peaks of $f_N(\omega)$ can be eliminated by data tapers.

In view of (3.2) through (3.4) and $c_0^{-1} \leq f_\nu(\omega) \leq c_0$, for $f_N(\omega) = |\Theta^{(N)}(\omega)|^2 f_\nu(\omega)$, we have

$$\left\{ \begin{array}{ll} c_2 L_N(\omega - \lambda_{2j})^{2s_{2j}} \leq f_N(\omega) \leq c_1 L_N(\omega - \lambda_{2j})^{2s_{2j}}, & \omega \in [\lambda_{2j} - \delta_0, \lambda_{2j} + \delta_0]; \\ c_2 L_N(\omega - \lambda_{1j})^{2s_{1j}} \leq f_N(\omega)^{-1} \leq c_1 L_N(\omega - \lambda_{1j})^{2s_{1j}}, & \omega \in [\lambda_{1j} - \delta_0, \lambda_{1j} + \delta_0]; \\ c_2 \leq f_N(\omega) \leq c_1, & \text{otherwise;} \end{array} \right. \quad (3.13)$$

where $c_1, c_2 > 0$ are constants.

For simplicity of notation, suppose at $\lambda_{2j} = \lambda_2$, $f_N(\lambda)$ has a highest peak with $s_{2j} = s_2 \geq 1$. Fix $\omega \notin (\lambda_2 - \delta_0, \lambda_2 + \delta_0)$. The cumulated leakage from $f_N(\lambda)$, $\lambda \in (\lambda_2 - \delta_0, \lambda_2 + \delta_0)$, to $EI_x(\omega)$ is then the integral (3.6) but with lower integration bound $\omega - \lambda_2 - \delta_0$ and upper integration bound $\omega - \lambda_2 + \delta_0$, or in (3.8), the bounds are $N(\omega - \lambda_2 - \delta_0)$ and $N(\omega - \lambda_2 + \delta_0)$. In view of (3.11) and (3.13), for large N ,

$$\begin{aligned} & \int_{N(\omega - \lambda_2 - \delta_0)}^{N(\omega - \lambda_2 + \delta_0)} f_N(\omega - \frac{\psi}{N}) |H_N(\psi)|^2 d\psi \\ & \leq c g N^{-2(l+1)} \rho_N^{-2l} \int_{N(\omega - \lambda_2 - \delta_0)}^{N(\omega - \lambda_2 + \delta_0)} c_1 L_N(\omega - \lambda_2 - \frac{\psi}{N})^{2s_2} d\psi \\ & = c g N^{-2(l+1)+2l\zeta} \int_{-N\delta_0}^{N\delta_0} L_N(-\frac{\psi}{N})^{2s_2} d\psi = c g N^{-2l-1+2l\zeta} \int_{-\delta_0}^{\delta_0} L_N(\theta)^{2s_2} d\theta \\ & = c g N^{-2l-1+2l\zeta} \left(\frac{N^{2s_2}}{N} + \int_{\frac{1}{N}}^{\delta_0} \theta^{-2s_2} d\theta \right) = c g N^{2s_2-2(l+1)+2l\zeta}, \end{aligned} \quad (3.14)$$

where $g = \max\{|\omega - \lambda_2 + \delta_0|^{-2(l+1)}, |\omega - \lambda_2 - \delta_0|^{-2(l+1)}\}$ and $c > 0$ is some constant

that may be different in different formulae. Note that (3.14) is true only for $s_2 > 0$. If $s_2 = 0$, we obtain $cgN^{-2l-1+2l\zeta}$ for the last expression; it could be $O(N^{-1})$ even for $l = 0$ (non-tapered) by choosing $\zeta = 0$ — this is not the situation of our concern.

On the other hand, again by (3.11) and (3.13), for large N (so $N^{-1} \leq \delta_0$),

$$\begin{aligned}
& \int_{N(\omega-\lambda_2-\delta_0)}^{N(\omega-\lambda_2+\delta_0)} f_N(\omega - \frac{\psi}{N}) |H_N(\psi)|^2 d\psi \\
\geq & \rho_N^{-2l} \int_{N(\omega-\lambda_2-N^{-1})}^{N(\omega-\lambda_2+N^{-1})} c_2 L_N(\omega - \lambda_2 - \frac{\psi}{N})^{2s_2} |\psi|^{-2(l+1)} |\phi_l(\psi, \rho_N)|^2 d\psi \\
\doteq & c_2 |\omega - \lambda_2|^{-2(l+1)} \rho_N^{-2l} N^{2s_2-2(l+1)} \int_{-1}^1 |\phi_l(N(\omega - \lambda_2) + \theta, \rho_N)|^2 d\theta. \quad (3.15)
\end{aligned}$$

If data are non-tapered ($l = 0$), we obtain from (3.10a), with $\psi = N(\omega - \lambda_2)$,

$$\int_{-1}^1 |\phi_0(\psi + \theta, \rho_N)|^2 d\theta = \int_{-1}^1 |1 - \exp(-i\psi) \exp(-i\theta)|^2 (2\pi)^{-2} d\theta \geq c; \quad (3.16)$$

then, from (3.15), (3.16) and noticing $\rho_N^{-2l} \geq c$, we see that the leakage from this highest peak reaches the order $O(N^{2(s_2-1)})$. Therefore the periodogram can not be asymptotically unbiased if $s_2 \geq 1$.

In general, the leakage is dominated by (3.14), that is $O(N^{2(s_2-l+l\zeta-1)})$. For $l > 0$, from (3.10), we may also show that $\int_{-1}^1 |\phi_l(\psi + \theta, \rho_N)|^2 d\theta$ is bounded away from 0, so that (3.15) may reach the order of $O(N^{2(s_2-l+l\zeta-1)})$, it means that a data taper with an adequate l and ζ eliminates the leakage.

Not only high peaks of $f_N(\lambda)$ may produce large frequency leakage, strong hidden periodicities can also produce leakage. The main contribution of a hidden frequency, ω_1 , to $d_y(\omega)$ is like a term in the middle summation of (2.10). Referring to (2.1), (3.7) and (3.11), we have

$$A_k (2\pi N)^{-1/2} H_1^{(N)}(\omega - \omega_k)$$

$$\begin{aligned}
&= A_k(2\pi N)^{1/2}(2\pi N)^{-1} \sum_{n=1}^N h_N\left(\frac{n}{N}\right) \exp\left\{-iN(\omega - \omega_k)\frac{n}{N}\right\} \\
&\doteq A_k(2\pi N)^{1/2} H_N(N(\omega - \omega_k)) \\
&= A_k(2\pi)^{1/2} N^{-l-1/2} \rho_N^{-l} (\omega - \omega_k)^{-(l+1)} \phi_l(N(\omega - \omega_k), \rho_N). \tag{3.17}
\end{aligned}$$

If $|A_k|$ is a constant, (3.17) is the order of $O(N^{-l-1/2+l\zeta})$. The main contribution of ω_k to $I_y(\omega) = |d_y(\omega)|^2$ is the squared value of (3.17), which is $O(N^{-1})$ in the non-tapered situation for $\omega \notin (\omega_k - \delta_o, \omega_k + \delta_o)$. However, the coefficient in the $O(N^{-1})$ - term depends on $|A_k|^2$. Let us put

$$|A_k| = O(N^p), \quad p \geq 0 \quad (p = 0 \text{ is the ordinary case}). \tag{3.18}$$

Here again this model is used to describe such critical situations although it does not necessarily exist in the real world. The leakage from ω_k to $I_y(\omega)$ now is $O(N^{2(p-l+l\zeta-1/2)})$ (which also fades away as ω goes away from ω_k), and the use of a data taper ($l > 0$) again helps to eliminate the leakage.

In the following subsections, we discuss how data taper may improve the methods for detecting and estimating hidden frequencies discussed in this paper.

3.4 SA and MP may not work without data taper

Let ω_1 be a hidden frequency and $\bar{\omega}_1$ be an initial estimate, for them, (1.9) holds. Using the definition of $\xi(s)$ in (2.3'), we have

$$(2\pi N)^{-1/2} \xi^x(s) = d_x(\bar{\omega}_1) \quad \text{with taper function } h_N(u, s), \quad s = 1, \dots, M. \tag{3.19}$$

Suppose that $f_N(\omega)$ has a peak of order s_2 at λ_2 , $\omega_1 \notin (\lambda_2 - \delta_0, \lambda_2 + \delta_0)$. Since the asymptotic variance of $d_x(\bar{\omega}_1)$ is $E I_x(\bar{\omega}_1)$, in view of (3.14) and (3.15), the cumu-

lated leakage from $f_N(\lambda)$, $\lambda \in (\lambda_2 - \delta_0, \lambda_2 + \delta_0)$, to $EI_x(\bar{\omega}_1)$ may reach the order of $O(N^{2(s_2-l+l\zeta-1)})$; or this leakage to the asymptotic standard deviation of $\xi^x(s)$ may be as large as $O(N^{s_2-l-1/2+l\zeta})$.

For SA to work, the error of regression model (2.5), $e(s)$, must be of order at most $o(1)$. Since $\eta(s)$, $\eta'(s)$ and $e(s)$ [for the definition, see (2.16), (2.16') and (2.19)] are of the same order, so $\eta(s)$ must be of the order of at most $o(1)$. Now, $|A_1G| \doteq cN$ (Lemma 2.1), so, the deviation of $\xi^x(s)$ is allowed to be at most $o(N)$. Hence, the requirement for suppressing the leakage is

$$s_2 - l - 1/2 + l\zeta < 1. \quad (3.20)$$

For the leakage does not affect the asymptotic properties of the estimate (Theorems 2.1 and 2.2) which hold under the conditions that the deviation of $\xi^x(\omega)/|A_1G|$ is $O(N^{-1/2})$ the other term in (2.16) is of smaller order, we require that

$$s_2 - l - 1/2 + l\zeta < 1/2. \quad (3.21)$$

If we take $l = 0$, then for $s_2 = 1$, the left hand side of (3.21) is $1/2$. SA works but the asymptotic properties are jeopardized, while $s_2 = 2$, SA fails to work. Now, we take $l = 2$ and a small value for ζ (simply, $\zeta = 0$), the left hand side of (3.21) may be even less than 0, SA works as good as in the ordinary situations.

For MP, consider the effect of the cumulated leakage from $f_N(\lambda)$, $\lambda \in (\lambda_2 - \delta_0, \lambda_2 + \delta_0)$, to the standard deviation of $N^{1/2} \sum_1^N h_N(\frac{n}{N}) \frac{n}{N} x(n) \exp(i\omega_1 n)$ and $N^{1/2} \sum_1^N h_N(\frac{n}{N}) x(n) \exp(i\omega_1 n)$ in $N^{1/2} dq_N(\omega_1)/d\omega$ of (2.32), where notice that $N^{-5/2}(NH_1^{(N)}) = O(N^{-1/2})$, $h(u)$ and $\dot{h}(u) = h(u)u$ have the same order l [though $\dot{h}(u)$ is not a usual taper function as it is not symmetric about $1/2$]. This cumulated leakage may reach the order of $O(N^{s_2-l-1/2+l\zeta})$.

But the variance of the left side of (2.34) is $O(1)$ [similar to $EI_x(w)$]; so, only if the contribution of the leakage is of $o(1)$, then its effect on the estimate of ω_1 can be ignored. It means that $s_2 - l - 1/2 + l\zeta < 0$ is required, and we have to choose some $l > 0$ if $s_2 > 0$.

Data tapers also help if A_k in the model are of the form (3.18), i.e. $A_k = O(N^{p_k})$, $p_k \geq 0$. Applying (3.17) to $H_{1,s}^{(N)}(\omega - \omega_1)$, we have

$$A_k H_{1,s}^{(N)}(\omega_k - \bar{\omega}_1) = O(N^{p_k - l + l\zeta}). \quad (3.22)$$

Suppose that $\{x(n)\}$ is an ordinary series (no sharp high peaks in its spectrum) and $p_1 = 0$. Consider the estimation of ω_1 . Due to (3.19), a term in the summation of the right hand side of (2.10) is $O(N^{p_k - l + l\zeta})$. In the ordinary situation where all $p_k = 0$ and $l = 0$ (no taper) we have $O(N^{p_k - l + l\zeta}) = O(1)$. Now, some $p_k > 0$. For SA may work as well as in the ordinary situation, it requires $p_k - l + l\zeta \leq 0$. Note that the standard deviation of $\xi^x(s)$ is $O(N^{1/2})$. If $p_k - l + l\zeta < 1/2$ is not fulfilled then SA does not work. Obviously, these inequalities may be fulfilled for $p_k > 0$ if l is chosen large enough. A similar discussion can be carried out for the MP method.

3.5 A data taper helps detecting ω_1

The basic statistic for detecting ω_1 in model (1.1) and (1.2) is $\hat{z}(\omega) = I_y(\omega) / \{\hat{f}_N(\omega) \log N\}$,

where

$$\begin{aligned} I_y(\omega) &= (2\pi N)^{-1} \sum_{j,k=-K}^K A_k A_j^* H_1^{(N)}(\omega - \omega_k) H_1^{(N)}(\omega - \omega_j)^* \\ &\quad + (2\pi N)^{-1/2} \left(\sum_{k=-K}^K A_k H_1^{(N)}(\omega - \omega_k) \right) d_x(\omega)^* \end{aligned}$$

$$+(2\pi N)^{-1/2} \left(\sum_{k=-K}^K A_k H_1^{(N)}(\omega - \omega_k) \right)^* d_x(\omega) + I_x(\omega). \quad (3.23)$$

and $\hat{f}_N(\omega)$ is a nonparametric estimate of $f(\lambda)$ — usually an average of neighbouring periodogram values where $I_y(\omega)$ is left out, or some peak insensitive estimator for $f(\lambda)$ as in von Sachs (1993).

Consider ω in a neighbourhood of ω_1 . In ordinary situations, ω_1 contributes a peak to $I_y(\omega)$ at ω_1 through the first summation of (3.23), with the height of $(2\pi N)^{-1} |A_1 H_1^{(N)}|^2 = O(N)$, while the estimator $\hat{f}_N(\omega)$ should not be influenced too much by the peak of $I_y(\omega)$ at ω_1 [it is one of the major problems to design the estimator $\hat{f}_N(\omega)$ in such a way - however, this is not discussed in the present paper]. Ideally, $\hat{f}_N(\omega_1)$ should be of order $O(1)$ leading to a value for $\hat{z}(\omega)$ of magnitude $O(N/\log N)$. The test suggested in Section 1 then detects the existence of a hidden frequency.

When heavy leakage is present, these ideal properties of $I_y(\omega)$ and $\hat{f}_N(\omega)$ are corrupted. In particular $\hat{f}_N(\omega_1)$ may increase substantially which leads to a much lower value of $\hat{z}(\omega)$ around ω_1 . A similar situation may happen if there are other hidden periodicities of strong magnitude. However, leakage effect can be eliminated or reduced by data taper.

4 Simulation

4.1 An ordinary situation

The following specification is used in our simulation to verify the theorems in Section 2. Let $N = 150$, $\{x(n)\}$ in (1.1) be a Gaussian white noise with variance one, and three

hidden periodicities with $|A_k| = 1$ at frequencies $(10k + 4.5)\pi/150$ for $k = 0, 4, 5$, i.e.

$$\omega_0 = 0.0942, \quad \omega_4 = 0.9320, \quad \omega_5 = 1.1410. \quad (4.1)$$

Obviously either $(10k + 4)\pi/150$ or $(10k + 5)\pi/150$ can be used or regarded as the initial estimate of ω_k as ω_k is almost the average of these two points. The error of the initial estimate is always about ± 0.0105 . This setting provides an easy way to demonstrate the improvement of the final estimate and leads to a fair comparison for different procedures and different k . Notice that, for SA, the final estimate heavily depends on the initial estimate, while for MP, the initial estimate only serves as a reference of the location.

According to the above model, samples of 150 observations $y(n)$ with 1000 replications are generated with φ_k , $k = 0, 4, 5$, being independent uniform random numbers in $(-\pi, \pi]$ in each replication.

First, we use the procedure discussed in Section 1 [around (1.8)] with $\varepsilon = 0$ to detect ω_k . The choice of $\varepsilon = 0$ may lead to the detection of spurious hidden periodicities. However, we took this value since our concern was to study the effect that hidden frequencies which could not be discovered due to leakage. In our simulation study, all the ω_k , $k = 0, 4, 5$ are always detected throughout the 1000 replications with the initial estimate, $\bar{\omega}_k$ being either $(10k + 4)\pi/150$ or $(10k + 5)\pi/150$ no matter whether the data are tapered or not.

Let $\bar{\omega}_k^{(p)}$ be an initial estimate in the p th replication and $\hat{\omega}_k^{(p)}$ be the corresponding final estimate $\bar{\omega}_k^{(p)}$ obtained by either SA or MP. The quantities

$$MEAN(k) = \frac{1}{1000} \sum_{p=1}^{1000} \hat{\omega}_k^{(p)} \quad (4.2)$$

and

$$RMSE(k) = \left\{ \frac{1}{1000} \sum_{p=1}^{1000} (\hat{\omega}_k^{(p)} - \omega_k)^2 \right\}^{1/2}, \quad (4.3)$$

indicate the bias and the accuracy of the estimate respectively. Here RMSE stands for the “Root of Mean Square Error”. $RMSE(k) < 0.0105$ indicates that the estimate has been improved from its initial value.

Table 4.1 lists the $RMSE$ of both SA and MP procedures for all the combinations of $\rho = 0$ (non-taper), 0.4 (medium-size taper), 1.0 (full-size taper) and $M = 2, 3, 5, 7$. Since $MEAN(k)$ is very close to the true value ω_k in all cases, its values are not shown.

The numbers with brackets and without brackets under the SA part of Table 4.1 are the results of using SA once and twice respectively. It is evident that using SA twice reduces the error significantly in many cases. Among all the M used in SA, $M = 3$ and 5 provide better results ($M = 4$ should also be good).

For MP, we do not use the avenue of solving (2.34), since we believe that the remainder terms in (2.32) and (2.33) are too complicated. We simply calculate $I_y(j\pi/\{150 \times 100\})$, where j runs over all integers between $(10k + 3) \times 100$ and $(10k + 6) \times 100$; i.e. the values of the periodogram $I_y(\omega)$ in the range $[(10k + 3)\pi/150, (10k + 6)\pi/150]$ and on a lattice 100 times finer than $\pi/150$ are calculated. The final estimate of ω_k is then the value maximizing $I_y(\omega)$.

To compare the SA and MP procedures, we observe that we can often find a value in SA part which is better than, or at least, competitive to the corresponding value given by MP in each case. However, we see that, MP is more stable than SA.

For $\omega_0 = 0.0942$, the results confirm that data taper increases the estimation error which is theoretically demonstrated in Table 2.1. However, the results for ω_4 and ω_5 do not support the theoretical conclusion as data taper does not make $RMSE$ larger, but contrarily, often smaller.

In fact, this is the advantage of data taper mentioned in Section 3. This example makes

us aware that, even for such small $|A_4| = |A_5| = 1$, the effect of frequency leakage exists because ω_4 and ω_5 are close; but data taper may reduce this mutual effect. To illustrate this, an experiment was carried out as follows. Using the same parameter as before to generate $y(n)$ except for putting $|A_5| = 0$, the *RMSEs* of MP estimation procedure for ω_4 are shown in the brackets of MP column and the rows of ω_4 in Table 4.1. These numbers follow the same pattern as those in the rows of ω_0 , because now the estimate of ω_4 is no longer affected by the leakage from ω_5 . The numbers in the brackets of MP column and the rows of ω_5 are obtained by the similar way (putting $|A_4| = 0$).

4.2 Some critical situations

We now study the estimation of the same hidden frequencies $\omega_0, \omega_4, \omega_5$ under the situation where either additional strong peaks at frequencies λ_1, λ_2 or additional hidden frequencies ω_1, ω_2 with strong amplitudes are present. Here $\omega_k = \lambda_k = (10k + 4.5)\pi/150$, i.e. the strong peaks or hidden frequencies are at frequencies

$$\omega_1 = \lambda_1 = 0.3040; \quad \omega_2 = \lambda_2 = 0.5130. \quad (4.4)$$

Again, let

$$|A_0| = |A_4| = |A_5| = 1. \quad (4.5)$$

The following four specified situations are employed to demonstrate the advantage of data taper in detecting and estimating ω_0, ω_4 and ω_5 .

(A) $\{x(n)\} \in \mathcal{X}(N, 0, 2, \delta_0, c_0)$ is the *AR*(4)-model given by

$$\{1 - 2r_N(\cos\lambda_2)B + r_N^2 B^2\}^2 x(n) = \nu(n), \quad (4.6)$$

where B is the backward shift operator and $\{\nu(n)\}$ is white noise series with $N(0, 1)$ distribution. By choosing $r_N = 1 - 1.1/N$, $N = 150$ gives $r_N = 0.9927$. As $\{x(n)\}$ has

spectral density $f_N(\omega) = (2\pi)|1 - r_N e^{-i(\omega - \lambda_2)}|^{-4}$, we see that the peak of $f_N(\omega)$ at λ_2 is $(150/1.1)^4/2\pi$. Furthermore, $|A_1| = 0$ and $|A_2| = 0$.

(B) $\{x(n)\} \in \mathcal{X}(N, 0, 1, \delta_0, c_0)$ is given by

$$\{1 - 2r_N(\cos \lambda_1)B + r_N^2 B^2\}\{1 - 2q_N(\cos \lambda_2)B + q_N^2 B^2\}x(n) = \nu(n); \quad (4.7)$$

where $1 - q_N = 5(1 - r_N)$ gives $q_N = 0.9633$ for $N = 150$, $f_N(\omega)$ has peaks at λ_1 with height $(150/1.1)^2/2\pi$ and at λ_2 with height $(30/1.1)^2/2\pi$. $|A_1| = 0$ and $|A_2| = 0$.

(C) $\{x(n) = \nu(n)\}$ is a white noise series with $N(0, 1)$ distribution

$$|A_1| = 45, \quad |A_2| = 0 \quad \text{and (4.5) holds.} \quad (4.8)$$

(D) The same as in (C) except

$$|A_1| = 30, \quad |A_2| = 15. \quad (4.9)$$

In the simulation, all initial estimation errors are supposed to be $(\pi/150)/2 = 0.0105$. If the *RMSE* of the final estimate, given either by the SA method or the MP method, is larger than 0.0105 in a case, then the method is regarded as offering no further improvement and we simply say that the method “fails to work” in this case.

Tables 4.2 and 4.3 list the MEAN and the RMSE of MP estimates for 1000 replications. The column “*No.*” represents the number of times (among 1000 replications) that ω_k is detected for each ρ . Although ω_k may not be detected in many cases, the MP estimation procedure is carried out throughout all 1000 replications.

The results for ω_4 and ω_5 in these tables show that the medium-size and full-size taper may make the detection procedure and MP estimation procedure work very well while they fail to work or work badly for non-tapered data. The further ω_k stays away from those high peaks, the better estimation results we have. The results of ω_0 show the

limitation of data tapers since ω_0 is too close to the highest peak of the spectrum or the hidden frequency with largest $|A_k|$.

Here, the improvement of the estimates are in both MEAN and *RMSE*. For ω_4 and ω_5 , perhaps, medium-size taper is slightly more preferable. However for ω_0 , the most difficult situation, full-size taper makes MP work for (C) and (D), while medium-size taper fails at all.

These two tables also show that light taper does not make a difference for situation (A), but has more or less help for situations (B), (C) and (D).

Table 4.4 shows some *RMSE* of the SA estimation procedure. The results are the average of using both $(10k + 4)\pi/150$ and $(10k + 5)\pi/150$ as the first initial estimate of ω_k . Similar to the MP procedure, basically, SA fails to work for ω_0 , and $\rho = 0.1$ does not make much difference with $\rho = 0$. We only list the results for ω_4 and ω_5 with $\rho = 0.4$ and $\rho = 1.0$. The numbers with brackets and without brackets are again the results of using the SA once and twice respectively. In all these situations, different from the ordinary situation, $M = 2$ is always better than $M = 3$ [with one exception in situation (B) and $\rho = 1.0$]. SA almost “fails to work” for higher M (those results are not listed). The reason for this is that in critical situations with strong peaks or strong periodic components as much data as possible are needed for a reasonably good Fourier transform.

For ease of comparison of SA with MP, we put the corresponding numbers of MP (has been given in Tables 4.2 and 4.3) again in this table (columns MP). We see that for situation (A), SA is better than MP; for (B), both methods are competitive; for (C) and (D), MP is better.

Table 4.1 RMSEs of SA and MP estimates (white noise case)

		SA					
ω_k	ρ	$M = 2$	$M = 3$	$M = 5$	$M = 7$	MP	
0.0942 (ω_0)	0	0.0037 (0.0059)	0.0017 (0.0018)	0.0019 (0.0020)	0.0021 (0.0021)	0.0016	
	0.4	0.0036 (0.0059)	0.0018 (0.0019)	0.0041 (0.0041)	0.0025 (0.0026)	0.0018	
	1.0	0.0036 (0.0058)	0.0024 (0.0025)	0.0063 (0.0069)	0.0032 (0.0033)	0.0022	
0.9320 (ω_4)	0	0.0044 (0.0063)	0.0025 (0.0025)	0.0014 (0.0014)	0.0024 (0.0025)	0.0026 (0.0014)	
	0.4	0.0037 (0.0055)	0.0027 (0.0026)	0.0015 (0.0017)	0.0034 (0.0038)	0.0019 (0.0016)	
	1.0	0.0036 (0.0061)	0.0021 (0.0049)	0.0020 (0.0032)	0.0048 (0.0066)	0.0020 (0.0020)	
1.1410 (ω_5)	0	0.0044 (0.0063)	0.0025 (0.0025)	0.0014 (0.0014)	0.0023 (0.0024)	0.0022 (0.0014)	
	0.4	0.0037 (0.0054)	0.0026 (0.0024)	0.0015 (0.0017)	0.0033 (0.0040)	0.0019 (0.0017)	
	1.0	0.0036 (0.0066)	0.0021 (0.0067)	0.0020 (0.0035)	0.0048 (0.0068)	0.0021 (0.0021)	

Table 4.2 Detected numbers and MP estimates (cases A and B)

ω_k	ρ	Situation A			Situation B		
		<i>No.</i>	<i>Mean</i>	<i>RMSE</i>	<i>No.</i>	<i>Mean</i>	<i>RMSE</i>
0.0942	0	1	0.1187	0.0385	4	0.1109	0.0373
(ω_0)	0.1	0	0.1228	0.0400	3	0.01145	0.0385
	0.4	0	0.1254	0.0394	5	0.1140	0.0382
	1.0	1	0.1153	0.0422	6	0.1088	0.0403
0.9320	0	0	0.9082	0.0335	57	0.9237	0.0220
(ω_4)	0.1	0	0.9043	0.0337	156	0.9237	0.0219
	0.4	905	0.9305	0.0117	843	0.9312	0.0065
	1.0	872	0.9297	0.0101	705	0.9310	0.0077
1.1410	0	0	1.1233	0.0305	175	1.1388	0.0186
(ω_5)	0.1	9	1.1256	0.0281	729	1.1404	0.0191
	0.4	999	1.1401	0.0066	1000	1.1408	0.0027
	1.0	996	1.1391	0.0080	1000	1.1405	0.0038

Table 4.3 Detected numbers and MP estimates (cases C and D)

ω_k	ρ	Situation C			Situation D		
		<i>No.</i>	<i>Mean</i>	<i>RMSE</i>	<i>No.</i>	<i>Mean</i>	<i>RMSE</i>
0.0942	0	0	0.1089	0.0279	0	0.1008	0.0228
(ω_0)	0.1	0	0.1405	0.0484	0	0.1274	0.0415
	0.4	0	0.1158	0.0274	0	0.1065	0.0227
	1.0	0	0.0955	0.0080	149	0.0949	0.0058
0.9320	0	92	0.9311	0.0165	419	0.9308	0.0142
(ω_4)	0.1	551	0.9280	0.0140	644	0.9281	0.0138
	0.4	1000	0.9321	0.0022	1000	0.9321	0.0024
	1.0	1000	0.9320	0.0020	1000	0.9320	0.0020
1.1410	0	405	1.1411	0.0150	687	1.1416	0.0126
(ω_5)	0.1	1000	1.1411	0.0070	1000	1.1410	0.0068
	0.4	1000	1.1411	0.0020	1000	1.1410	0.0020
	1.0	1000	1.1410	0.0021	1000	1.1410	0.0021

Table 4.4 Some RMSEs of SA estimates (cases A,B,C,D)

		$\omega_4 = 0.9320$			$\omega_5 = 1.1410$		
		SA		MP	SA		MP
Situation	ρ	$M = 2$	$M = 3$		$M = 2$	$M = 3$	
A	0.4	0.0069	0.0196	0.0117	0.0076	0.0252	0.0066
		(0.0069)	(0.0190)		(0.0070)	(0.0235)	
	1.0	0.0066	0.0099	0.0101	0.0031	0.0057	0.0080
		(0.0066)	(0.0087)		(0.0053)	(0.0061)	
B	0.4	0.0075	0.0189	0.0065	0.0032	0.0067	0.0027
		(0.0070)	(0.0176)		(0.0054)	(0.0070)	
	1.0	0.0052	0.0070	0.0077	0.0031	0.0025	0.0038
		(0.0060)	(0.0070)		(0.0054)	(0.0026)	
C	0.4	0.0069	0.0197	0.0022	0.0078	0.0248	0.0020
		(0.0069)	(0.0191)		(0.0070)	(0.0235)	
	1.0	0.0066	0.0100	0.0020	0.0031	0.0057	0.0021
		(0.0066)	(0.0087)		(0.0053)	(0.0062)	
D	0.4	0.0070	0.0200	0.0024	0.0077	0.0251	0.0020
		(0.0069)	(0.0192)		(0.0070)	(0.0235)	
	1.0	0.0066	0.0100	0.0020	0.0031	0.0056	0.0021
		(0.0066)	(0.0087)		(0.0053)	(0.0062)	

5 Conclusion

In this article, we study the asymptotic properties of (the central limit theorem and the law of the iterated logarithm) of the hidden frequency estimates by using secondary analysis (SA) and the maximum periodogram method (MP). The general model is given by (1.1) and (1.2). We demonstrated how data taper affects these asymptotic properties. We concluded that data taper makes the estimation error larger in ordinary situations, i.e. the noise $\{x(n)\}$ follows a linear process with a flat spectrum and there is no strong hidden periodicity with its frequency close the hidden frequency of interest. MP is slightly better, but computationally less efficient than SA.

We also show that, when there are high peaks in the spectral density of $\{x(n)\}$, or there are strong hidden periodicities, with their frequencies close to the hidden frequency of interest, the nature of the estimates are no longer well described by these asymptotic theorems. In such situations, the contribution of such hidden frequency to the periodogram or Fourier transformation is blurred by frequency leakage from those high peaks or strong periodicities. However, leakage effect may be eliminated by data taper which makes these estimation procedures work again. A data taper may also make a procedure of detecting hidden frequencies work very well while the procedure fails with non-tapered data.

The theorems are verified by simulation. Our simulation results shows that, even in some ordinary situations, if two hidden frequencies are close, data taper may improve the estimate.

In ordinary situation, using SA for estimating ω_k with $M = 3, 4$ or 5 can possibly provide a better estimate than MP estimate but the result strongly depends on M . SA is not as stable as MP. In critical situation, $M = 2$ is the best choice for SA estimate and higher M are not recommended. Moreover, SA is recommended when the noise has high peaks in its spectral density, while for the situations where some very strong periodicities exist, MP estimates are much better.

If the spectrum of the series is flat and the hidden frequencies are well separated we

recommend a small taper with, say, $\rho = 0.1$. This has only a minor efficiency loss but protects a bit against leakage problems. If the spectrum really has strong peaks or hidden frequencies close to each other we recommend a medium- or full-size taper. In any case the periodogram with different tapers should be plotted prior to the analysis.

ACKNOWLEDGEMENT

This research is partially supported by the Max-Planck Fellowship Foundation of Germany and the Summer Research Grant of the Chinese University of Hong Kong. The authors are the most grateful to the referee for many very interesting and helpful comments.

REFERENCES

- An, H.Z., Chen, Z.G. and Hannan E.J. (1983). The maximum of periodogram. *J. Multi. Anal.* **13**, 383-400.
- Chen, Z.G. (1988a). Consistent estimates for hidden frequencies in a linear processes. *Adv. Appl. Prob.* **20**, 295-314.
- Chen, Z.G. (1988b). An alternative consistent procedure for detecting hidden frequencies. *J. Time Series Anal.* **9**, 301-317.
- Chen, Z.G. (1990). An extension of Lai and Wei's law of the iterated logarithm with applications to time series analysis and regression. *J. Multi. Anal.* **32**, 55-69.
- Dahlhaus, R. (1988). Small sample effects in time series analysis: a new asymptotic theory and new estimate *Ann. Statist.* **16**, 808-841.
- Dahlhaus, R. (1990). Nonparametric high resolution spectral estimation. *Prob. Theory Related Fields* **85**, 147-180.
- Fisher, A.R. (1929). Test significance in harmonic analysis. *Proc. R. Soc. London* **A 125**, 54-59.

- Grenander, U. and Rosenblatt, M. (1957) *Statistical Analysis of Stationary Time Series*. Wiley, New York.
- Hannan, E.J. (1961). Testing for a jump in the spectral function. *J. R. Statist. Soc. B* **23**, 394-404.
- Hannan, E.J. (1970). *Multiple Time Series*. John Wiley and Sons, Inc.
- Hannan, E.J. (1973). The estimation of frequency. *J. Appl. Prob.* **10**, 510-519.
- Hannan, E.J. and Mackisack, M. (1986). A law of the iterated logarithm for an estimate of frequency. *Stoch. Proc. Appl.* **22**, 103-109.
- Hartley, H.O. (1949). Tests of significance in harmonic analysis. *Biometrika* **36** 194-201.
- Priestley, M.B. (1981). *Spectral Analysis and Time Series*. Vol. 1. Academic Press, New York.
- Turkman, K.F. and Walker, A.M. (1984). On the asymptotic distribution of maxima of trigonometric polynomials with random coefficients. *Adv. Appl. Prob.* **16**, 819-842.
- von Sachs, R. (1993) Detecting periodic components in a stationary time series by an improved nonparametric procedure. Proceedings of the International Conference on Applications of Time Series Analysis in Astronomy and Meteorology, Univ. de Padova, Italy, Sept 6 - 10, 1993, pp. 115-118.
- Whittle, P. (1952). The simultaneous estimation of a time series harmonic components and covariance structure. *Trab. Estadist.* **3**, 43-57.