Statistica Sinica **10**(2000), 1041-1067

ASYMPTOTIC DISTRIBUTION OF ESTIMATES FOR A TIME-VARYING PARAMETER IN A HARMONIC MODEL WITH MULTIPLE FUNDAMENTALS

Rafael A. Irizarry

Johns Hopkins University

Abstract: Window-based estimates for stochastic harmonic regression models are useful for cases where harmonic parameters appear to be time-varying. Least squares estimates for harmonic models with one fundamental have been studied and asymptotic variance expressions have been developed. This paper extends these results to weighted least squares for the multiple fundamental case, and presents an application in signal processing.

Key words and phrases: Asymptotic variance, harmonic regression, signal processing, sound analysis, time-varying parameters, weighted least squares estimates.

1. Introduction

Consider the signal plus noise model

$$y_t = s(t;\beta) + \epsilon_t,\ (t = 1,\ldots,T), \tag{1}$$

where the signal $s(t;\beta)$ is composed of J periodic components

$$s(t;\beta) = \sum_{j=1}^{J} s_j(t;\beta_j),\ \beta = (\beta_1,\ldots,\beta_J)', \tag{2}$$

and each component $s_j(t;\beta_j)$ is a sum of K_j sinusoidal components

$$s_j(t;\beta_j) = \sum_{k=1}^{K_j}\{A_{j,k}\cos(\omega_{j,k}t) + B_{j,k}\sin(\omega_{j,k}t)\}, \tag{3}$$

$$\beta_j = (A_{j,1}, B_{j,1},\ldots,A_{j,K_j}, B_{j,K_j}, \omega_{j,1},\ldots,\omega_{j,K_j})' \tag{4}$$

that are somehow related to each other, for example by a condition like (7) below.

Here the noise ϵ_t is a strictly stationary real valued random process, with autocovariance function $c_{\epsilon\epsilon}(u) = \mathrm{Cov}\,\{\epsilon_{t+u},\epsilon_t\}$ that satisfies Condition 1 below, and has power spectrum

$$f_{\epsilon\epsilon}(\lambda) = \frac{1}{2\pi}\sum_u c_{\epsilon\epsilon}\exp(-i\lambda u), \quad -\infty < \lambda < \infty. \tag{5}$$

Condition 1. The noise process $\{\epsilon_t\}$ is such that all its moments exist, with zero mean, and with $c_{\epsilon\ldots\epsilon}(u_1,\ldots,u_{L-1})$ the joint cumulant function of order L for $L=2,3,\ldots$. Furthermore, the

$$C_L = \sum_{u_1=-\infty}^{\infty} \cdots \sum_{u_{L-1}=-\infty}^{\infty} |c_{\epsilon\ldots\epsilon}(u_1,\ldots,u_{L-1})| \tag{6}$$

satisfy $\sum_k C_k z^k/k! < \infty$ for z in a neighborhood of 0.

We are interested in studying this model when a harmonic relation exists between the frequencies $\omega_{j,k}$ in each of the J components $s_j(t;\beta_j)$. We assume that there exist J different fundamental frequencies $0<\theta_j<2\pi$, $(j=1,\ldots,J)$ such that

$$\omega_{j,k} = k\theta_j,\ \ k=1,\ldots,K_j \text{ for each } j=1,\ldots,J. \tag{7}$$

We refer to K_j as the number of harmonics associated with the jth fundamental frequency.

We define the parameter $\beta^c = (\beta_1^c,\ldots,\beta_J^c)'$ for the model under constraint (7), where $\beta_j^c = (A_{j,1},B_{j,1},\ldots,A_{j,K_j},B_{j,K_j},\theta_j)'$ for each $j=1,\ldots,J$. The parameters of this model are identifiable provided the frequency parameters $\omega_{j,k}$ are different from each other.

Several authors have studied various forms of the model defined by (1) – (6) with $J=1$ and when no constraint, for example like (7), exists on the frequencies $\omega_{j,k}$. Notice that if no constraint exists, even if $J>1$, no relation exists between the $\omega_{j,k}, k=1,\ldots,K_j$ and there is no need for the j indexes. Then we can rewrite the model as

$$s(t;\beta) = \sum_{k=1}^{K} \{A_k\cos(\omega_k t) + B_k \sin(\omega_k t)\},\ \beta = (A_1,B_1,\ldots,A_K,B_K,\omega_1,\ldots,\omega_K)' \tag{8}$$

with $K=\sum_{j=1}^{J} K_j$.

For this model Walker (1971) establishes weak consistency and asymptotic normality for estimates that are asymptotically equivalent to least squares estimates under the assumption that the ϵ_t are distributed independently and identically with mean zero and finite variance. For functions $s(t;\beta)$ which do not necessarily have the form (2),(3) or (8), but which satisfy certain regularity conditions, Hannan (1971) finds estimates of β by minimizing

$$Q_T(\beta) = T^{-1}\sum_{j=0}^{T-1} \phi(\lambda_j) I_2^T(\lambda_j;\beta)$$

with $\lambda_j = 2\pi j/T$, $(j = 0,\ldots,T-1)$, $\phi(\lambda) > 0$ a frequency domain weight function, and

$$I_2^T(\lambda;\beta) = (2\pi T)^{-1}\left|\sum_{t=1}^{T}\{y_t - s(t;\beta)\}\exp(-i\lambda t)\right|^2,$$

the residual periodogram. Hannan (1971) proves strong consistency and asymptotic normality for these estimates under the condition that the error process $\{\epsilon_t\}$ is a purely non-deterministic linear process. Furthermore, Hannan (1971) shows that the asymptotic variance of the estimates is minimized by taking $\phi(\lambda) = f_{\epsilon\epsilon}^{-1}(\lambda)$. Hannan (1971, 1973) notices that for $s(t;\beta)$, as in (8), the regularity conditions presented in Hannan (1971, pp.768-769) are not satisfied, yet consistency and asymptotic normality, for the estimates obtained by minimizing $Q_T(\beta)$, hold. Furthermore, Hannan (1973) notices that for this case the asymptotic distribution is independent of $\phi(\lambda)$, and thus it is only necessary to consider the case $\phi(\lambda) = 1$, i.e. the least squares estimates. Hannan (1974) extends this result to the case where the constraint (7) holds and reports, without proof, an expression for the asymptotic variance of the estimates. Brown (1990) finds a minor mistake in this expression and derives a correct one by computing a first-order Taylor expansion of the gradient of $Q_T(\beta)$. Brown (1990) does not take advantage of the fact that the asymptotic variance does not depend on $\phi(\lambda)$; however, by computing the result for $\phi(\lambda) = f_{\epsilon\epsilon}^{-1}(\lambda)$, he gives the correct expression for the asymptotic variance.

The harmonic regression signal plus noise model is widely used, for example for the study of biological rhythm data (Greenhouse, Kass and Tsay (1987)) and sound analysis (Irizarry (1998)). In the latter example, models with multiple fundamental frequencies can be used to study reverberated sound signals. Because the harmonic parameters may be time-varying we present the asymptotic distribution of window-based estimates below.

2. Weighted Least Squares

In some applications, it may be useful to fit models in order to obtain estimates of parameter functions that depend on time. For these cases, it is only natural to consider window-based estimates.

Given that we have T observations for which we approximate the time-varying parameter with a constant β, the weighted least squares method consists of choosing $\hat{\beta}$ to minimize the criterion

$$S_T(\beta) = \sum_{t=1}^{T} w(\frac{t}{T})\,\{y_t - s(t;\beta)\}^2 \tag{9}$$

with $w(s)$ a weight function. Some of the results regarding the asymptotic behavior of these estimates require that the weight function satisfy the following condition.

Condition 2. The function $w(s)$ is non-negative, bounded, of bounded variation, has support $[0,1]$, $W_0 > 0$, and $W_1^2 - W_0 W_2 \neq 0$ where

$$W_n = \int_0^1 s^n w(s)\, ds. \tag{10}$$

We are interested in finding the asymptotic properties of the weighted least squares estimate of the parameter β^c of the model defined by (1) – (7). To do this we extend the result developed by Hannan (1973) for the asymptotic behavior of the weighted least squares estimates for this model, but without constraint (7) and under the assumption that the error process is a linear process. In this paper we extend this to the more general condition for the error process defined by (6).

As mentioned, the model defined by (1)–(6) can be rewritten as model (8). Because this is the model considered by Hannan (1973), for simplicity we consider model (8) for the remainder of this section.

We define weighted versions of the estimates defined by Hannan (1973): for $k = 1, \ldots, K$,

$$\tilde{A}_{k,T} = 2 \sum_{t=1}^{T} w(\frac{t}{T}) y_t \cos(\tilde{\omega}_{k,T}\, t) / \sum_{t=1}^{T} w(\frac{t}{T}), \tag{11}$$

$$\tilde{B}_{k,T} = 2 \sum_{t=1}^{T} w(\frac{t}{T}) y_t \sin(\tilde{\omega}_{k,T}\, t) / \sum_{t=1}^{T} w(\frac{t}{T}), \tag{12}$$

where, if we write $\omega = (\omega_1, \ldots, \omega_K)$ and $\tilde{\omega}_T = (\tilde{\omega}_{1,T}, \ldots, \tilde{\omega}_{K,T})$, $\tilde{\omega}_T$ is such that

$$q_T(\tilde{\omega}_T) = \max_{0 \leq \omega \leq \pi} q_T(\omega) \tag{13}$$

and q_T is defined by

$$q_T(\omega) = \sum_{k=1}^{K} \left| T^{-1} \sum_{t=1}^{T} w(\frac{t}{T}) y_t \exp(it\omega_k) \right|^2. \tag{14}$$

We notice that these estimates are asymptotically equivalent to the weighted least squares estimates $\hat{A}_{k,T}$, $\hat{B}_{k,T}$ and $\hat{\omega}_{k,T}$ for $k = 1, \ldots, K$, the values that minimize (9). This result is best understood by first considering the case of one sinusoidal component ($K = 1$)

$$s(t; \beta_0) = A_0 \cos(\omega_0 t) + B_0 \sin(\omega_0 t) \tag{15}$$

with $\beta_0 = (A_0, B_0, \omega_0)'$, and then generalizing to the case of several sinusoidal components.

As done in Walker (1971) for the unweighted case, we notice that if we define

$$R_T(\beta) = \sum_{t=1}^{T} w(\frac{t}{T})y_t^2 + \frac{1}{2}(A^2+B^2)\sum_{t=1}^{T} w(\frac{t}{T}) - 2\sum_{t=1}^{T} w(\frac{t}{T})y_t\{A\cos(\omega t)+B\sin(\omega t)\} \tag{16}$$

with $\beta = (A, B, \omega)$, then

$$S_T(\beta) - R_T(\beta) = \frac{1}{2}\sum_{t=1}^{T} w(\frac{t}{T})\{(A^2 - B^2)\cos(2\omega t) + 2AB\sin(2\omega t)\}. \tag{17}$$

Here $S_T(\beta)$ is the weighted residual sum of squares of (9). The difference in (17) is deterministic and, using Lemma 1 (Appendix C), we can show it is bounded as $T \to \infty$ if $0 < \omega < \pi$.

By taking derivatives and setting to 0, we see that the ω that maximizes the periodogram of the tapered data $w(t/T)y_t$ also minimizes $R_T(\beta)$. This and (17) may be used to show that the estimates in (11), (12), and (14) are asymptotically equivalent to the weighted least squares estimates.

For the case of more than one frequency we use the previous result as done in Walker (1971). The function corresponding to (16), whose minimization yields approximate weighted least squares estimators, becomes

$$\begin{aligned} R_T(\beta) = &\sum_{t=1}^{T} w(\frac{t}{T})y_t^2 + \frac{1}{2}\sum_{k=1}^{K}(A_k^2 + B_k^2)\sum_{t=1}^{T} w(\frac{t}{T}) \\ &-2\sum_{k=1}^{K}\sum_{t=1}^{T} w(\frac{t}{T})y_t\{A_k\cos(\omega_k t) + B_k\sin(\omega_k t)\}. \end{aligned} \tag{18}$$

In this case, to obtain (18) from (9) we need terms of the form

$$A_k A_l \sum_{t=1}^{T} w(\frac{t}{T})\cos(\omega_k t)\cos(\omega_l t) \text{ and } B_k B_l \sum_{t=1}^{T} w(\frac{t}{T})\sin(\omega_k t)\sin(\omega_l t)$$

to be bounded, since they are included in $S_T(\beta) - R_T(\beta)$. Some conditions need to be imposed to avoid having the ω_k's too close together, thus preventing the estimators of two or more frequencies from converging in probability to the same value. An appropriate condition is $\lim_{T\to\infty}\min_{k\neq l}(T|\hat{\omega}_k - \hat{\omega}_l|) = \infty$. Walker (1971) proposes maximizing $q_T(\omega)$ subject to

$$\min_{k\neq l}(|\omega_k - \omega_l|) = T^{-\frac{1}{2}}. \tag{19}$$

We redefine the estimates of the ω's as those that maximize (13) but under constraint (19).

We now can prove Theorem 1, in a way similar to Walker (1971) or Hannan (1973) by using estimates that are asymptotically equivalent to the weighted least squares estimates of interest. Appendix C has a sketch of the proof containing the key differences for the weighted case.

Theorem 1. *Let $\hat{\beta}_T$ be the weighted least squares estimates of β for the model defined in* (8) *obtained by minimizing equation* (9)*, with $w(s)$ satisfying Condition* 2*. Then $\hat{\beta}_T$ is a consistent estimate of β and for $k = 1,\ldots,K$, the vectors $\{T^{\frac{1}{2}}(\hat{A}_{k,T} - A_k), T^{\frac{1}{2}}(\hat{B}_{k,T} - B_k), T^{\frac{3}{2}}(\hat{\omega}_{k,T} - \omega_k)\}'$ converge in distribution to mutually independent normal vectors with zero mean and covariance matrices*

$$V_k = \frac{4\pi f_{\epsilon\epsilon}(\omega_k)}{A_k^2 + B_k^2} \begin{pmatrix} c_1 A_k^2 + c_2 B_k^2 & -c_3 A_k B_k & -c_4 B_k \\ -c_3 A_k B_k & c_2 A_k^2 + c_1 B_k^2 & c_4 A_k \\ -c_4 B_k & c_4 A_k & c_0 \end{pmatrix}. \tag{20}$$

Here $c_0,\ldots,c_4$ are constants depending on the weight function $w(s)$ and are defined at (A.1) *in Appendix* A.

Remark 1. If $w(s) = 1$ for all s, the constants above reduce to $c_1 = 1$, $c_2 = 4$, $c_3 = 3$, $c_4 = 6$ and $c_0 = 12$, and the covariance matrix reduces to that obtained in the equally weighted case by, for example, Walker (1971).

3. Harmonic Model with Multiple Fundamentals

In this section we present results describing the asymptotic properties of the weighted least squares estimates for the parameter β^c of the model defined in (1)–(7). As mentioned above, Brown (1990) finds the asymptotic distribution of least squares estimates, $w(s) = 1$, for the case $J = 1$, by computing a first-order Taylor expansion of the gradient $\nabla Q_T(\beta)$. This requires tedious computations, especially if we consider $J > 1$. Using Theorem 1, and a technique similar to the one used by Brillinger (1980) to estimate a bifrequency, a result for the weighted least squares when $J \geq 1$ is obtained in a simpler manner. Computations showing how this result is obtained are in Appendix D.

Corollary 1. *Let $\hat{\beta}^c_{j,T}$ be the weighted least squares estimate of β^c_j for the model defined by* $(1)-(7)$*. For each $j = 1,\ldots,J$, let $N_j(T)$ be a $(2K_j + 1) \times (2K_j + 1)$ diagonal matrix whose first $2K_j$ diagonal entries are $T^{1/2}$ and whose $(2K_j + 1)$st diagonal entry is $T^{3/2}$. Then for each $j = 1,\ldots,J$, if $w(s)$ satisfies Condition* 2*, $\hat{\beta}^c_{j,T}$ is a consistent estimate of β^c_j and the $N_j(T)(\hat{\beta}^c_{j,T} - \beta^c_j)$ converge in distribution to mutually independent multivariate normal vectors with zero mean and* *covariance matrix*

$$4\pi \left\{ \sum_{k=1}^{K_j} k^2 (A_{j,k}^2 + B_{j,k}^2)/f_{\epsilon\epsilon}(k\theta_j) \right\}^{-1} \begin{pmatrix} D_j + c_0^{-1} E_j E_j' & E_j \\ E_j' & c_0 \end{pmatrix}, \tag{21}$$

where the matrices D_j and E_j are defined at (*A*.3), (*A*.4), *and* (*A*.5) *in Appendix* B.

Remark 2. Corollary 1 provides a useful approximation of the variance of the estimates of the fundamental frequencies

$$\mathrm{Var}\,(\hat{\theta}_{j,T}) \approx 4\pi c_0\, T^{-3} \left\{ \sum_{k=1}^{K_j} k^2 (A_{j,k}^2 + B_{j,k}^2)/f_{\epsilon\epsilon}(k\theta_j) \right\}^{-1}, \tag{22}$$

where c_0 is as in (A.1). Notice that the denominator in (22) is a sum of weighted signal-to-noise ratios. This implies that the precision of the estimate increases with the total magnitude of the respective harmonic components.

Remark 3. In some instances it might be useful to find estimates for the amplitudes of the harmonic components defined by $\rho_{j,k} = (A_{j,k}^2 + B_{j,k}^2)^{1/2}$. Using Corollary 1, it is easy to verify that the amplitude estimates defined by $\hat{\rho}_{j,k,T} = (\hat{A}_{j,k,T}^2 + \hat{B}_{j,k,T}^2)^{1/2}$ are consistent estimates of the $\rho_{j,k}$'s, and are asymptotically mutually independent normal. We may approximate the variance with $\mathrm{Var}\,(\hat{\rho}_{j,k,T}) \approx 4\pi c_1 T^{-1} f(k\theta_j)$ where c_1 is as in (A.1).

4. Model Selection

When considering the model defined by (1)–(7), the number of fundamental frequencies J, as well as the number of harmonics for each fundamental frequency, $K_1, \ldots, K_J$, can be considered to be unknown parameters. In practice we must make a decision on how many to include in the model we fit to the data.

For models defined by (8), He (1984) suggests a simple intuitive procedure to estimate the number of periodic components K, and proves the procedure to be strongly consistent under some conditions. Quinn (1989) suggests an Akaike information criteria (AIC) type estimator for K and proves strong consistency when the noise process is a certain kind of white noise. Wang (1991) extends these results under assumptions like that of Condition 2. In this section we extend Wang's criterion in order to use it when weight estimates are being considered, then we extend it to the case of models with multiple fundamentals.

Consider the model defined by (8) with K sinusoidal components. We will consider K to be the true number of sinusoidal components. To estimate K when it is unknown consider the following scheme.

Let $\Omega_1 = (0,\pi)$ and let $\hat\omega_1$ be the maximum point of $q_T(\omega) = |T^{-1}\sum_{t=1}^{T} w(t/T)\ y_t \exp(it\omega)|^2$. For $l > 1$, given that Ω_{l-1} and $\hat\omega_{l-1}$ are defined, let $\Omega_l = \Omega_{l-1} \setminus (\hat\omega_{l-1} - u_T, \hat\omega_{l-1} + u_T)$ and $\hat\omega_l$ be the value of $\omega \in \Omega_l$ that maximizes $q_T(\omega)$.

One can repeat this procedure until $(0,\pi)$ is exhausted. The $\hat\omega_l$'s will be referred to as the maximum periodogram frequencies.

Define

$$A_T(\lambda) = 2\sum_{t=1}^{T} w(\frac{t}{T}) y_t \cos(\lambda\, t) / \sum_{t=1}^{T} w(\frac{t}{T}),$$

$$B_T(\lambda) = 2\sum_{t=1}^{T} w(\frac{t}{T}) y_t \sin(\lambda\, t) / \sum_{t=1}^{T} w(\frac{t}{T})$$

and, for any $k = 1, 2, \ldots$, let

$$\sigma_T^2(k) = \frac{1}{T}\sum_{t=1}^{T}\left[y_t - \sum_{l=1}^{k}\{A_T(\hat\omega_l)\cos(\hat\omega_l t) + B_T(\hat\omega_l)\sin(\hat\omega_l t)\}\right]^2.$$

Notice that σ_T^2 is asymptotically equivalent to the residual mean squares if we were fitting model (8). Notice also that $\sigma_T^2(k_1) \le \sigma_T^2(k_2)$ for all $k_1 > k_2$, thus σ_T^2 is not an appropriate criterion for estimating K.

As done by Wang (1991) we let $\mathrm{BIC}_T(k) = T\,\log\sigma_T^2(k) + b_T k$ (best information criterion) with the sequence $\{b_T\}$ satisfying $b_T/T \to 0$ as $T \to \infty$. An AIC type estimator for K can be defined by

$$\hat K = \min\{k : \mathrm{BIC}_T(k) \le \mathrm{BIC}_T(k+1)\}. \tag{23}$$

From Theorem 4.5.1 in Brillinger (1981) we have that for ϵ_t satisfying Condition 1,

$$\lim_{T\to\infty}\sup_\lambda |d_T(\lambda)|\,(T\log T)^{-1/2} \le 2\{2\pi U_0 \sup_\lambda f_{\epsilon\epsilon}(\lambda)\}^{1/2} \tag{24}$$

with probability 1, where

$$d_T(\lambda) = \sum_{t=1}^{T} w(\frac{t}{T})\epsilon_t \exp(it\lambda) \tag{25}$$

is the discrete Fourier transform of the tapered errors and U_0 is defined by (A.2) in Appendix A.

This implies that by choosing the sequences $\{b_T\}$ and $\{u_T\}$ so that

$$\liminf_{T\to\infty}\frac{b_T}{\log T} \le 4U_0\frac{\sup_\lambda f_{\epsilon\epsilon}(\lambda)}{(2\pi)^{-1}\int_{-\pi}^{\pi} f_{\epsilon\epsilon}(\lambda)\,d\lambda}$$

and $u_T \to 0$ with $(T\log T)^{1/2}u_T \to \infty$ (notice that if we use (19) this is satisfied), we have that the amount we minimize the BIC when $k > K$, $\mathrm{BIC}_T(K) - \mathrm{BIC}_T(k)$, is asymptotically corrected by the quantity $b_T(k-K)$. Similarly, for $k < K$ the penalty $b_T(K-k)$ is asymptotically negligible compared to the amount we minimize the BIC by adding $K-k$ parameters. In fact we can show that for large enough T,

$$\hat{K}_T = K \tag{26}$$

with probability 1. See Wang (1991) for the details.

Now we turn our attention to estimating the number of fundamentals, J, and their respective number of periodicities, $K_1, \ldots, K_J$, when considering the model defined by (1)–(7). To do this we use the result just described for model (8).

First we find an estimate $\hat{K}$ of the total number of sinusoidal components $K = \sum_{j=1}^{J} K_j$ in the model defined by (1)–(7). From (26) we know that for large enough T we have $\hat{K} = \sum_{j=1}^{J} K_j$ with probability 1. To estimate the number of fundamentals J, let $M_0 = \{\hat{\omega}_{(k)}, k = 1, \ldots, \hat{K}\}$ be the set of ordered maximum periodogram frequencies. Consider $\hat{\omega}_{1,1} = \hat{\omega}_{(1)}$ to be an estimate of what we consider to be the first fundamental. The frequencies $M_1 = \{\omega \in M_0 : |\omega - k\hat{\omega}_{1,1}| \le T^{-1/2}$ for some $k = 1, 2, \ldots\}$ are considered to be the set containing the harmonics related to $\hat{\omega}_{1,1}$. Given that we have defined fundamentals $1, \ldots, j-1$ and their respective harmonics, contained in the sets $M_1, \ldots, M_{j-1}$, define the jth fundamental $\hat{\omega}_{j,1}$ as the smallest frequency in $M_0 \setminus \cup_{l=1}^{j-1} M_l$ and $M_j = \{\omega \in M_0 \setminus \cup_{l=1}^{j-1} M_l : |\hat{\omega}_{j,1} - k\hat{\omega}_{j,1}| \le T^{-1/2}$ for some $k = 1, 2, \ldots\}$. Continue this process until all $\hat{K}$ maximum periodogram frequencies are exhausted. The number of fundamental frequencies found will be the estimate $\hat{J}$ of J and the number of elements in M_j will be the estimate $\hat{K}_j$ of K_j for each $j = 1, \ldots, \hat{J}$.

Theorem 1 implies that, since $\hat{K}_T = K$ with probability 1 for large enough T, $\hat{J}$ and $\hat{K}_j$ are consistent estimates of J and K_j for each $j = 1, \ldots, J$.

5. Estimating Time-Varying Parameters

In some applications the parameters of the harmonic structure may be time-varying. For example, in the case of signals studied in musical sound analysis, the performer generally changes the sound being produced by the instrument. Examples here are changes of note or pitch, vibrato, and tremolo, to mention a few. For this reason the model defined by (1) – (7) may not be appropriate. Instead a version with time-varying parameters needs to be considered.

In signal processing in general, it is common that the *sample rate* (observations taken per unit time) is large. For the case of sound signals the harmonic parameters appear to change slowly in time. This motivates estimation procedures where the fundamental frequency, and other harmonic parameters, are assumed fixed within segments of short duration. The asymptotic theory presented in

Theorem 1 and Corollary 1 may be used to give approximate distributions for estimates obtained this way. To use this asymptotic theory we need the sampling rate to be large so that the short segments considered for the estimation contain enough observations. Furthermore, for the asymptotics to make sense, we need the fundamental frequency, or frequencies, to be large enough so that within small segments the signal still contains harmonic-type behavior. For sound signals, the application of interest in this paper, we have both these properties. In this section we present these heuristic arguments in a theoretical framework.

Consider the following sequence of processes that are equivalent to the model defined by (1)–(7), but with time-varying parameters

$$y_{n,N} = s\left\{\frac{n}{N}, \beta^c_N(\frac{n}{N})\right\} + \epsilon_{n,N} \text{ for } n = 1, \ldots, N \times D, N \geq 1. \tag{27}$$

Here $s\{t;\beta^c_N(t)\} = \sum_{j=1}^J s_j\{t;\beta_{j,N}(t)\}, \beta^c_N(t) = \{\beta_{1,N}(t), \ldots, \beta_{J,N}(t)\}'$, $t \in [0, D]$, where each component $s_j\{t;\beta_{j,N}(t)\}$ is a sum of K_j sinusoidal components with time-varying parameters $s_j\{t;\beta_{j,N}(t)\} = \sum_{k=1}^{K_j}[A_k(t)\cos\{k\,\theta_{j,N}(t)\,t\} + B_k(t)\sin\{k\,\theta_{j,N}(t)\,t\}]$ for $j = 1, \ldots, J$. The duration of the signal is assumed without loss of generality to be $D = 1$, and $\beta_{j,N}(t) = \{A_{j,1}(t), B_{j,1}(t), \ldots, A_{j,K_j}(t), B_{j,K_j}(t), \theta_{j,N}(t)\}'$ whence, for each $k = 1, \ldots, K_j$, $j = 1, \ldots, J$, $A_{j,k}(t)$ and $B_{j,k}(t)$ are continuous bounded functions for $t \in [0, 1]$. We make sure that each one of the time-varying fundamental frequencies $\theta_{j,N}, j = 1, \ldots, J$ is large with respect to the sampling rate N by using the following assumption.

Condition 3. For each $j = 1, \ldots, J$, there exists a continuous function $\theta_j(t)$ with $0 < \theta_j(t) < 2\pi$ for $t \in [0, 1]$, such that the sequence of functions $\theta_{j,N}(t) - N\theta_j(t)$ converges uniformly to 0 for $t \in [0, 1]$.

The sequence of stochastic processes is defined by considering, for each $N \geq 1$, $\{\epsilon_{n,N}, n = 1, \ldots, N\}$ to be N observations of a stationary processes $\{\epsilon_n\}$ satisfying Condition 1.

We must notice that the signal $s\{t, \beta^c_N(t)\}$ is different for each N. In fact, if $\theta_N(t) = \theta_N$ is constant in time for each N, then the number of cycles per unit time $N\theta_N$ tends to infinity with N. Therefore, we must not interpret the asymptotics as having a fixed signal from which we can obtain better estimates as we increase the sample rate N.

A more reasonable interpretation of the asymptotics is the following: as N increases we observe signals for which the size (in units of time) of segments containing, say, H observations become smaller, thus the time-varying parameters are closer to constant. Condition 3 assures that instantaneous fundamental frequencies are large enough so that within such segments we have a model that approximates the harmonic model defined by (1)–(7).

To see this more clearly, for any $t_0 \in (0,1)$ consider a small enough *estimation segment* or *estimation window* size h_N so that we can act as if the functional parameters are constant in time, i.e., $\beta_N(t) \approx \beta_N(t_0)$ for t in the segment $(t_0 - h_N/2, t_0 + h_N/2)$.

Letting $H_N = \lfloor h_N N \rfloor$ we have that within the estimation window the signal $s\{t;\beta_N^c(t)\}$ is approximately

$$\sum_{j=1}^{J}\sum_{k=1}^{K_j} [A_{j,k}(t_0)\cos\{k\,\theta_{j,N}(t_0)\,(t_0 - h_N/2 + n/N)\}+ \\ B_{j,k}(t_0)\,\sin\{k\,\theta_{j,N}\,(t_0)(t_0 - h_N/2 + n/N)\}]\,, \quad n = 1,\dots H_N. \tag{28}$$

If N is large, then Condition 3 suggests that $\theta_{j,N}(t_0)/N \approx \theta_j(t_0)$ with $0 < \theta_j(t_0) < 2\pi$ for each $j = 1,\dots,J$. Letting $\theta_{j,0} = \theta_j(t_0)$, $A_{j,k,0} = C_A A_{j,k}(t_0)$, and $B_{j,k,0} = C_B B_{j,k}(t_0)$ for all j,k, an approximation for (28) is $\sum_{j=1}^{J}\sum_{k=1}^{K_j}$ $\{A_{j,k,0}\cos(k\theta_0 n) + B_{j,k,0}\sin(k\theta_0 n)\}$, $n = 1,\dots,H_N$ with constants C_A and C_B not depending on n that correct for the phase, and are easily obtained using trigonometric identities. If $H_N \to \infty$ as $N \to \infty$, these approximations imply that within the estimation segment we have a model that approximates the harmonic model defined in (1)-(7) with a large number of observations H_N. Therefore, we should be able to obtain reasonable estimates of $\beta_N^c(t_0)$ for large values of N. Now we will make this asymptotic theory precise, but first we need an assumption regarding the smoothness of the time-varying parameters $\beta_N^c(t)$.

Condition 4. There exists an M such that for each $k = 1,\dots,K_j, j = 1,\dots,J$, $\sup_{t\in[0,1]} |A'_{j,k}(t)|, \sup_{t\in[0,1]} |B'_{j,k}(t)|$, and $\sup_{t\in[0,1]} |\theta'_{j,N}(t)|$ are all bounded by M for all N.

Intuitively this assumption prevents the local behavior of the function $s\{t;\beta_N^c(t)\}$ from being too different from a sum of sinusoids and thereby preserving some sort of local harmonic structure known to be present in sound signals. We can now define an estimate of the time-varying parameters.

For each $t_0 \in (0,1)$ define the local weighted least square estimate using span h_N in the following way. Let $H_N = \lfloor h_N \times N \rfloor$, $n_0 = \lfloor t_0 \times N \rfloor$, and $l = n_0 - H_N/2$, $u = n_0 + H_N/2$, then

$$\hat{A}_{j,k,N}(t_0) = 2\sum_{n=l+1}^{u} w(\frac{n-l}{H_N})\, y_{n,N}\cos(\hat{\omega}_{j,k,N}\, n)/\sum_{n=1}^{H_N} w(\frac{n}{H_N}), \tag{29}$$

$$\hat{B}_{j,k,N}(t_0) = 2\sum_{n=l+1}^{u} w(\frac{n-l}{H_N})\, y_{n,N}\sin(\hat{\omega}_{j,k,N}\, n)/\sum_{n=1}^{H_N} w(\frac{n}{H_N}), \tag{30}$$

where, if we write $\omega = (\omega_{1,1}, \ldots, \omega_{K_J,J})$ and $\hat{\omega}_N = (\hat{\omega}_{1,1,N}, \ldots, \hat{\omega}_{K_J,J,N})$, $\hat{\omega}_N$ is such that

$$q_N(\hat{\omega}) = \max_{0 \leq \omega \leq \pi} q_N(\omega), \tag{31}$$

where $q_N(\omega)$ is now defined by:

$$q_N(\omega) = \sum_{j=1}^{J} \sum_{k=1}^{K_j} \left| (H_N)^{-1} \sum_{n=1}^{H_N} w(\frac{n}{H_N})\, y_{n+l,N} \exp(in\omega_{j,k}) \right|^2 . \tag{32}$$

We obtain estimates $\hat{\theta}_{j,N}(t_0)$ from the $\hat{\omega}_{j,k,N}$'s using the method presented in the proof of Corollary 1, namely

$$\hat{\lambda}_{j,N}(t_0) = \frac{\sum_{k=1}^{K_j} k\, \hat{\omega}_{j,k,N} \{\hat{A}^2_{j,k,N}(t_0) + \hat{B}^2_{j,k,N}(t_0)\}}{\sum_{k=1}^{K_j} k^2 \{\hat{A}^2_{j,k,N}(t_0) + \hat{B}^2_{j,k,N}(t_0)\}} \tag{33}$$

for each $j = 1, \ldots, J$.

Corollary 2. *For any $t_0 \in (0,1)$ let the sequence of segment sizes $\{h_N,\ N > 1\}$ be such that $h_N \downarrow 0$ and $H_N = \lfloor h_N \times N \rfloor \to \infty$ as $N \to \infty$. Then if Condition 4 holds we have for each $k = 1, \ldots, K_j$, $j = 1, \ldots, J$, $\hat{A}_{j,k,N}(t_0)$ and $\hat{B}_{j,k,N}(t_0)$ are consistent estimates of $A_{j,k}(t_0)$ and $B_{j,k}(t_0)$ respectively, and $\lim_{N\to\infty} H_N |\hat{\theta}_{j,N}(t_0) - \theta_j(t_0)| = 0$ in probability for each $j = 1, \ldots, J$, where the estimates are defined by equations (29)−(33). Furthermore, for each $j = 1, \ldots, J$, let $H_{j,N}$ be a $(2K_j + 1) \times (2K_j + 1)$ diagonal matrix whose first $2K_j$ diagonal entries are $H_N^{1/2}$ and whose $(2K_j+1)$st diagonal entry is $H_N^{3/2}$. Then the vectors $H_{j,N}\{\hat{\beta}_{j,N}(t_0) - \beta_j(t_0)\}, j = 1, \ldots, J$ converge in distribution to mutually independent multivariate normal vectors with zero mean and variance matrix as in (21) in Corollary 1, but with harmonic parameters the time-varying parameter functions evaluated at t_0.*

The proof of Corollary 2 is in Appendix E. As done for Theorem 1, the estimates considered in this Corollary may be shown to be equivalent to the weighted least squares estimates.

Remark 4. Notice that in Corollary 2 the weight function $w(s)$ may be defined to be equally weighted, $w(s) = 1, 0 \leq s \leq 1$. The asymptotics work since for the local estimates only H_N points are given positive weight regardless of the shape of the weight function $w(s)$. In practice, functions that give more weight to points near the middle of the estimation window are used since there is an a priori belief that there is more *information* about the time-varying parameter function evaluated at t_0, in points near t_0. Exploring how different window functions may provide more "efficient" estimates from a theoretical point of view is of interest, but will not be discussed further in this paper.

6. Applications in Sound Analysis

The study of musical sound has become a popular research field within signal processing. Stochastic harmonic regression models have been used to analyze sound waves produced by musical instruments (see for example, Rodet (1997)). Least squares estimation provides a way to obtain useful parametric representations of sound signals (Irizarry (1998)). Harmonic parameters in sound analysis models are considered to be time-varying (Rodet (1997)), thus it is useful to consider window-based estimates when performing estimation.

The sound studied in this example is a pipe organ playing two consecutive notes, $F\sharp$ (fundamental frequency of about 368 Hz.) and E (fundamental frequency of about 325 Hz.), for a total duration of two seconds. The room where the recording was made, Hertz Hall in U.C. Berkeley, is a concert hall characterized as having quite a bit of echo. When the second note is played, the first note can still be heard. This is called reverberation.

During the recording of the organ sound, 44100 observations of the air wave pressure were recorded per second. Figure 1a shows a time series plot of the sound. Notice that after 1.1 seconds or so there appears to be an abrupt change, due to the note change.

In Figure 1b we see a *spectrogram* of the data. To obtain the spectrogram, the data was divided into 300 overlapping segments, each with 2647 observations (segments of approximately 60 milliseconds duration). For each segment, the periodogram of the data is computed and plotted in an image plot, with darker shades of grey representing higher values. When the data has periodic components at certain frequencies, the periodogram will show peaks at these frequencies, thus the spectrogram of a sound wave will show dark horizontal lines at the frequencies corresponding to the fundamental frequency being played, and corresponding harmonics. In this spectrogram, we can see that after 1.1 seconds or so, the second note begins. The vertical line is at the note change. In this figure we can see the frequency component related to the main fundamental frequency change to a smaller value after 1.1 seconds, from about 368 Hz. to about 325 Hz. We also notice that frequency components of the first note remain during the playing of the second note. As well, there is a relatively dark horizontal line around a low frequency of 50 Hz.

The spectrogram seen in Figure 1b seems to suggest that fitting harmonic models to this data may be appropriate. By looking at Figure 1a, it is apparent that when looking at the entire signal the total amplitude is slowly varying. It seems appropriate to use the windowed estimation procedure described in Section 5.

An analysis like the ones typically found in this literature (Rodet (1997)) would consider segments of small duration (less than 20 milliseconds) and fit a model like (8) with large values of K to the data contained in such segments. This analysis would fail to identify different fundamental frequencies and their respective harmonic structure. Analyses that define a fundamental frequency usually define only one. If we consider 368 Hz. and 325 Hz. to be the fundamental frequencies for the first and second parts of the signal, the time-varying parameters within segments of 20 millisecond duration appear to be usefully constant. As an example, in Figure 1c we show a 20 millisecond segment around time $t_0 = 0.115$ seconds.

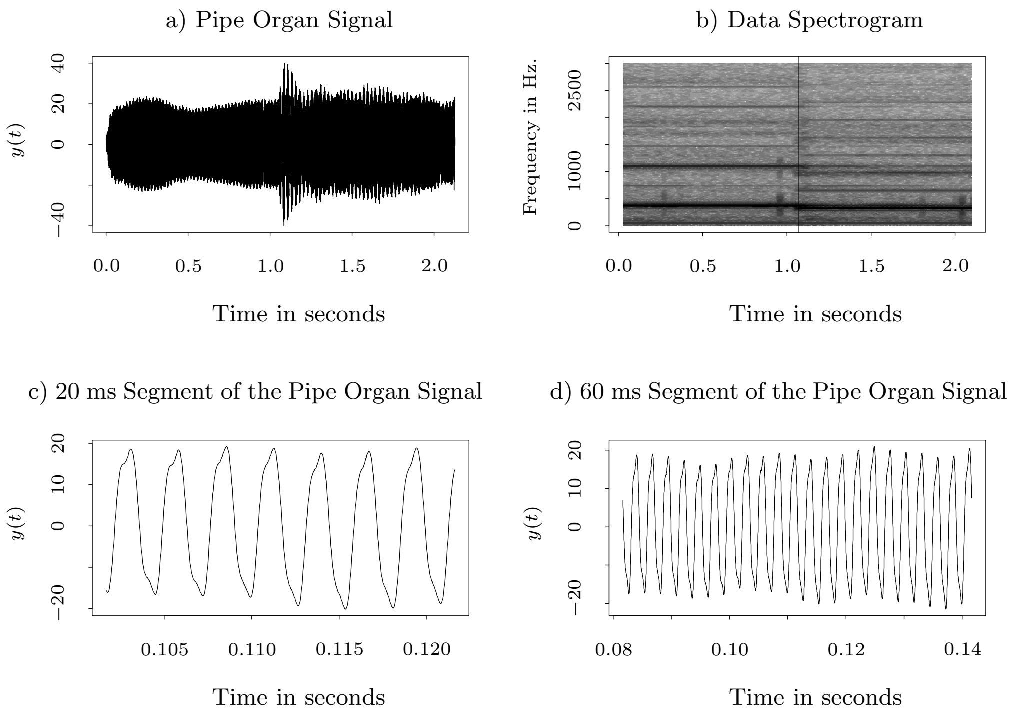

Figure 1. Time series plot of a sound with reverberation produced by a pipe organ playing two consecutive notes, F♯ followed by E and respective spectrogram (in the spectrogram the vertical line is at the note change); two segments around time $t_0 = 0.115$.

Figure 2a shows the estimates of the fundamental frequency when the model defined by (1)–(7), with $J = 1$ and $K_1 = 12$, is fitted to each 20 millisecond segment. The spectrogram of the residuals obtained from fitting this model can be seen in Figure 2b. Notice that in the part of the spectrogram corresponding to the part of the signal where the reverberation was occurring, the harmonic structure produced by the echo of the previous note can be seen as well as the low frequency component observed in the spectrogram of the original data. This model does not seem to provide an appropriate fit.

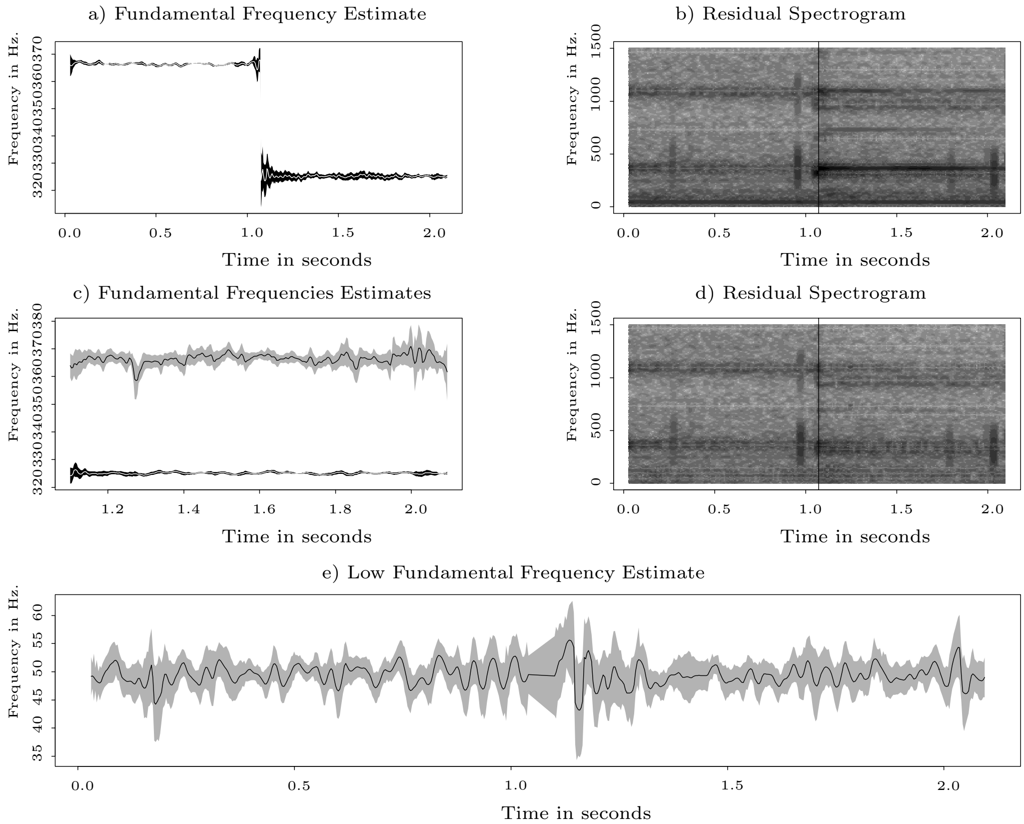

Figure 2. Estimated fundamental frequency with marginal $\pm$ 2 standard errors limits, when fitting a model with one fundamental; residual spectrogram for this fit; estimated fundamental frequencies when fitting a model with multiple fundamentals; residual spectrum for this fit; and the low frequency component fundamental frequency estimate.

The fit is greatly improved by fitting a multi-fundamental model as described in this paper. The location of the dark horizontal lines in the spectrograms seems to suggest that for the part of the signal corresponding to the first note, a model defined with two fundamentals ($J = 2$), one corresponding to the note being played (368 Hz.) and one corresponding to the low frequency component at 50 Hz., may be appropriate. For the second part of the signal corresponding to the second note, the spectrogram in Figure 1b suggests that we fit a model with 3 fundamentals ($J = 3$), one corresponding to the note being played (325 Hz.), one corresponding to the echo of the first note, and one corresponding to the low frequency at 50 Hz.

Before fitting the model we must decide on the size of the segments we consider. As mentioned in the discussion in Section 5, we need to choose segment sizes such that the harmonic parameters are approximately constant within the segments. For the analyses in this paper we chose the segment sizes in a heuristic fashion. In Figure 1 we see two segments around time 0.115 seconds, the first with 20 millisecond duration and the second with 60 millisecond duration. The harmonic parameters appear to be usefully constant in the first segment. However, when examining many consecutive segments, we see that the total amplitude seems to follow a sinusoidal pattern of about 50 Hz. This is in agreement with the appearance of a dark horizontal line in both the spectrograms of the original signal and the residuals after fitting the one fundamental frequency model. The harmonic parameters for the second segment, if we consider there to be a fundamental frequency around 368 Hz. and another at 50 Hz., appear to be usefully constant. If we consider segments of durations longer than 60 milliseconds, then the slowly varying amplitude phenomenon, seen in Figure 1a, begins to be apparent. Notice that segments of 60 millisecond contain 2650 data points. For these segments we have around 20 oscillations related to the fundamental frequency associated with the note being played and 3 oscillations associated with the lower frequency of 50 Hz. Using the asymptotic approximations described in Section 5 seems appropriate.

We use the BIC described in Section 4 to verify the choice for the number of fundamentals for the two parts of the sound and to choose the number of harmonics for each fundamental. For the segments in the first part of the sound signal we fit a harmonic model (1)–(7) with $J = 2$, $K_1 = 7$ and $K_2 = 3$. For the segments in the second part we fit a model with $J = 3$, $K_1 = 7$, $K_2 = 6$, and $K_3 = 3$. In Figure 2c we see the estimates obtained for the fundamental frequencies related to the note being played, and the echo of the first note for the second part of the signal. In Figure 2e we see the estimate of the fundamental frequency related to the low frequency component.

Using Corollary 2, the variance of the estimates may be approximated. Marginal ± 2 standard errors around the estimates are included in Figures 2a, 2c, and 2e. Notice that the difference between the two estimates appears highly significant. Furthermore, notice that the approximate standard errors are larger for the estimate related to the reverberated note ($F\sharp$). This is due to the fact that the signal-to-noise ratio is smaller for the part of the sound related to the echo or reverberation, than for the part of the sound currently being produced by the instrument (E).

By looking at the residual spectrogram in Figure 2d we see the effect of reverberation in the residuals and the low frequency component have been removed by the addition of the second and third fundamentals in the model.

7. Discussion and Extensions

We have presented a useful method for decomposing a sound signal into harmonic components produced by different fundamental frequencies and noise. Theoretical results needed to justify the approximations for the standard errors of our estimates have been presented.

Furthermore, we have introduced a criterion useful for choosing the number of fundamentals and harmonic frequencies to be included in the models being fit. However, the results presented for this criterion are asymptotic. In practice we need to choose the value of the penalty multiplier b_T somewhat arbitrarily (for example, the asymptotics still hold if we multiply b_T by a constant) and formal methods of selecting this parameter in practical situations is a subject of future work. In the present work b_T was chosen to be $\log W_0 T$ and the resulting estimates were in agreement with the spectrograms, and with what we hear when listening to the original signals, to make appropriate choices for our models.

In the example presented in this paper the segment sizes were chosen in a heuristic fashion. Much work was put into choosing window sizes that provide reasonable fits. Fitting models using different window sizes and comparing the spectrograms of the residuals may be used as a way to verify that our choice is reasonable. The residuals may also be played and heard. Residual analysis by ear is a useful tool for detecting lack of fit. For example, when listening to the residual obtained when ignoring the fundamental related to the low frequency component at 50 Hz., a sound characteristic of wind going through pipes is heard suggesting that an important component of the sound has not been included in the model. Studying the usefulness of methods for choosing the window sizes automatically is an important subject for future work.

The sounds associated with the analyses presented in this paper can be heard by visiting a demo on the author's home page at:
http://biosun01.biostat.jhsph.edu/~ririzarr/Demo/index.html.

Acknowledgements

This paper is part of the author's doctoral dissertation which was written at the University of California, Berkeley, under the supervision of Professor David R. Brillinger. Professor Brillinger's suggestions and encouragement were an essential part of this work. I would also like to thank the editor and the referee for suggestions that have improved this paper. Also, thanks to Rick Schoenberg and Jonah Goldstein for insightful comments, and to Adrian Freed for the data.

Appendix

A. Definition of constants referred to in Theorem 1

The constants are defined by:

$$\begin{aligned}
&c_0 = a_0 b_0,\ c_1 = U_0 W_0^{-2},\ c_2 = a_0 b_1 \\
&c_3 = a_0 W_1 W_0^{-2}(W_0^2 W_1 U_2 - W_1^3 U_0 - 2W_0^2 W_2 U_1 + 2W_0 W_1 W_2 U_0) \\
&c_4 = a_0(W_0 W_1 U_2 - W_1^2 U_1 - W_0 W_2 U_1 + W_1 W_2 U_0)
\end{aligned} \tag{A.1}$$

where

$$\begin{aligned}
&a_0 = (W_0 W_2 - W_1^2)^{-2},\ a_1 = (U_0 U_2 - U_1^2),\ a_2 = W_0^{-2}(W_0 U_1 - W_1 U_0)^2 \\
&b_n = W_n^2 U_2 + W_{n+1}(W_{n+1} U_0 - 2W_n U_1), n = 0, 1.
\end{aligned}$$

Here W_0, W_1, and W_2 are defined by (10) and U_0, U_1 and U_2 are defined by

$$U_n = \int_0^1 s^n w(s)^2\, ds. \tag{A.2}$$

B. Definition of matrices referred to in Corollary 1

The matrices needed to define the asymptotic variance in Theorem 2 are given in the following way:

D_j is a $2K_j \times 2K_j$ matrix:

$$D_j = \left(\sum_{k=1}^{K_j} k^2 (A_{j,k}^2 + B_{j,k}^2)/f_{\epsilon\epsilon}(k\theta_j)\right) \begin{pmatrix} D_{j,1} & \dots & O \\ \vdots & \ddots & \vdots \\ O & \dots & D_{j,K_j} \end{pmatrix}, \tag{A.3}$$

where

$$D_{j,k} = \frac{f_{\epsilon\epsilon}(k\theta_j)}{b_0(A_{j,k}^2 + B_{j,k}^2)} \begin{pmatrix} c_1 b_0 A_{j,k}^2 + a_1 B_{j,k}^2 & a_2 A_{j,k} B_{j,k} \\ a_2 A_{j,k} B_{j,k} & a_1 A_{j,k}^2 + c_1 b_0 B_{j,k}^2 \end{pmatrix} \tag{A.4}$$

and

$$E_j = c_4 \left(-B_{j,1}, A_{j,1}, \dots, -K_j B_{j,K_j}, K_j A_{j,K_j}\right)'. \tag{A.5}$$

C. Proof of Theorem 1

Before proving Theorem 1 we need to prove a few simple results. Set $\Delta_n^T(\lambda) = \sum_{t=1}^T w(t/T) t^n \exp(i\lambda t)$. We need the following result.

Lemma 1. *If $w(t)$ satisfies Condition 2 then we have for $n = 0, 1, \dots$,*

$$\lim_{T\to\infty} T^{-(n+1)} \Delta_n^T(\lambda) = W_n,\ \textit{for}\ \lambda = 0, 2\pi, \tag{A.6}$$

$$\Delta_n^T(\lambda) = O(T^n),\ \textit{for}\ 0 < \lambda < 2\pi, \tag{A.7}$$

with W_n defined by (10).

Proof of Lemma 1. Fix n. To prove (A.6) notice that for $\lambda = 0, 2\pi$ we have, from the boundedness and bounded variation of $w(s)$,

$$\lim_{T\to\infty} T^{-(n+1)}\Delta_n^T(\lambda) = \lim_{T\to\infty}\sum_{t=1}^{T}(\frac{t}{T})^n w(\frac{t}{T})(\frac{1}{T}) = \int_0^1 u^n w(u)\,du = W_n.$$

To prove (A.7), let $0 < \lambda < 2\pi$ and define $\Delta^t(\lambda) = \sum_{s=1}^{t}\exp(i\lambda s)$, with the convention that $\Delta^0(\lambda) = 0$. Letting $h(u) = u^n w(u)$ and using summation by parts we have

$$\Delta_n^T(\lambda) = T^n\left[h(1)\Delta^T(\lambda) + \sum_{t=1}^{T-1}\left\{h(\frac{t}{T}) - h(\frac{t+1}{T})\right\}\Delta^t(\lambda)\right].$$

Notice that if $w(t)$ is bounded and has bounded variation on $[0,1]$, so does $h(s)$. Let M be $\sup_s|h(s)|$ and V be the total variation of $h(s)$. Then we have

$$\left|\Delta_n^T(\lambda)\right| \le T^n\left[M|\Delta^T(\lambda)| + V\max_{1\le t\le T}|\Delta^t(\lambda)|\right].$$

We know, see for example Brillinger (1981), that $|\Delta^t(\lambda)| \le L = 1/|\sin(\frac{1}{2}\lambda)|$ for all t. Notice that L depends on λ, but given $0 < \lambda < 2\pi$ it is constant for all t, and $|\Delta_n^T(\lambda)| \le T^n L(M+V)$. This completes the proof of the Lemma.

To prove consistency and asymptotic normality for the weighted least squares, or equivalently the estimates defined by (11)–(14), we need a result concerning the behavior of the periodogram of the noise and its derivatives with respect to ω.

Lemma 2. *Let the stationary noise process* $\{\epsilon_t\}$ *satisfy Condition* 1 *and let the weight function* $w(s)$ *satisfy Condition* 2. *Then if*

$$p_T(\omega) = \left|T^{-(n+1)}\sum_{t=1}^{T} w(\frac{t}{T})t^n\epsilon_t\exp(-it\omega)\right|$$

one has for $n = 0, 1, \ldots$, $\lim_{T\to\infty}\sup_{0\le\omega\le\pi} p_T(\omega) = 0$, *in probability.*

Remark 1. Lemma 2 has been shown to be true under different assumptions for the equally weighted case, $w(s) = 1$. In most cases the result for the weighted case follows similarly. Walker (1971) proves the lemma for white noise with finite variance. Hannan (1973) proves it under ergodic and purely non-deterministic conditions. Brillinger (1986) proves a version for spatial point processes. Under Conditions 1 and 2, Lemma 2 follows directly from Theorem 4.5.1 in Brillinger (1981, p.98).

Using Lemmas 1 and 2 we prove consistency in a way similar to Walker (1971) or Hannan (1973).

Consider first the one sinusoidal case as defined by (15). We start by proving

$$\lim_{T\to\infty} T|\tilde{\omega}_T - \omega_0| = 0, \text{ in probability.} \tag{A.8}$$

This is stronger than ordinary consistency, but is needed to prove the consistency of the remaining two estimates and asymptotic normality.

Letting $D_0 = \frac{1}{2}(A_0 - iB_0)$ we have

$$\begin{aligned} q_T(\omega) = &\left|T^{-1}d_T(\omega)\right|^2 + \left|T^{-1}\{D_0\Delta_0^T(\omega_0+\omega) + \overline{D}_0\Delta_0^T(\omega_0-\omega)\}\right|^2 \\ &+ 2\Re\left(\left[T^{-1}d_T(\omega)\right]\left[T^{-1}\{D_0\Delta_0^T(\omega_0+\omega) + \overline{D}_0\Delta_0^T(\omega_0-\omega)\}\right]\right) \end{aligned} \tag{A.9}$$

with $q_T(\omega)$ defined in (13) and $d_T(\omega)$ defined by (25). By Lemma 1 we have, for $0 < \omega < \pi$, $T^{-1}\Delta_0^T(\omega_0+\omega) = o(1)$ and

$$T^{-1}\Delta_0^T(\omega_0-\omega) = \begin{cases} W_0 & : \quad \omega = \omega_0 \\ o(1) & : \quad \text{otherwise.} \end{cases}$$

Lemma 2 implies that for $0 < \omega < \pi$, $T^{-1}|d_T(\omega)| = o_p(1)$, so that

$$q_T(\omega) = \frac{1}{4}\rho_0^2\left|T^{-1}\Delta_0^T(\omega-\omega_0)\right|^2 + o_p(1) \quad \text{and} \quad q_T(\omega_0) = \frac{1}{4}\rho_0^2 W_0^2 + o_p(1).$$

To prove (A.8), for any $b > 0$, define

$$P_T(b) = \{\omega : T|\omega - \omega_0| \geq b\}. \tag{A.10}$$

Notice that

$$\begin{aligned} \Pr\left(T|\tilde{\omega}_T - \omega_0| \geq b\right) &\leq \Pr\left(\sup_{\omega\in P_T(b)} q_T(\omega) \geq q_T(\omega_0)\right) \\ &= \Pr\left(\sup_{\omega\in P_T(b)}\left|T^{-1}\Delta_0^T(\omega-\omega_0)\right| \geq W_0 + o_p(1)\right) \end{aligned}$$

and that, using a Riemann integration argument, we can show that

$$\sup_{\omega\in P_T(b)}\left|T^{-1}\Delta_0^T(\omega-\omega_0)\right| = \left|\int_0^1 w(s)\exp\{iT(\omega-\omega_0)s\}\,ds\right| + o(1).$$

Let ω^* be such that

$$\left|\int_0^1 w(s)\exp\{iT(\omega^*-\omega_0)s\}\,ds\right| = \sup_{\omega\in P_T(b)}\left|\int_0^1 w(s)\exp\{iT(\omega-\omega_0)s\}\,ds\right|.$$

Let $b^* = T|\omega^* - \omega_0| \geq b > 0$. Then, by the definition of $P_T(b)$ given at (A.10), we have

$$\lim_{T\to\infty} \Pr\left(T|\tilde{\omega}_T - \omega_0| \geq b\right) \leq \lim_{T\to\infty} \Pr\left(\left|\int_0^1 w(s)\exp(ib^*s)\,ds\right| + o(1) \geq W_0 + o_p(1)\right).$$

Since $W_0 > 0$ is a deterministic constant and $b^* > 0$,

$$W_0 = \left|\int_0^1 w(s)\,ds\right| = \int_0^1 |w(s)\exp(ib^*s)|\;ds > \left|\int_0^1 w(s)\exp(ib^*s)\,ds\right|$$

and we have (A.8).

To prove consistency for $\tilde{A}_T$ and $\tilde{B}_T$, let $r(t,\beta) = \{D_0\exp(i\omega_0 t) + \overline{D}_0\exp(-i\omega_0 t)\}$ and $L = 2\left\{\sum_{t=1}^T w(t/T)\right\}^{-1}$. By Lemma 1 and the Mean Value Theorem we have that, for some $\overline{\omega}_T$ satisfying $|\overline{\omega}_T - \omega_0| \leq |\tilde{\omega}_T - \omega_0|$,

$$|\tilde{A}_T - A_0 + i(\tilde{B}_T - B_0)| = \left|L\sum_{t=1}^T w(\frac{t}{T})r(t;\beta)it\exp(i\overline{\omega}_T t)(\tilde{\omega}_T - \omega_0)\right| + o(1).$$

The first term in the right hand side of the above equation is smaller than $L\sum_{t=1}^T w(t/T)|r(t;\beta)|t|\tilde{\omega}_T - \omega_0| \leq \rho_0\, T|\tilde{\omega}_T - \omega_0| = o_p(1)$ and thus $|(\tilde{A}_T - A_0) + i(\tilde{B}_T - B_0)| = o_p(1)$. Because both the real and imaginary parts converge in probability to 0, consistency for the one sinusoidal case is proven. The general case, for various harmonic components, follows in the same way. See Irizarry (1998) for details.

To show asymptotic normality, consider first the one sinusoidal case as defined by (15). Using Theorem 4.4.2 in Brillinger (1981, p.95) we have that the vector $\mathbf{u}$, with components

$$\begin{aligned} u_1 &= T^{-\frac{1}{2}}\sum w(\frac{t}{T})\epsilon_t\cos(\omega_0 t)\ , \quad u_2 = T^{-\frac{1}{2}}\sum w(\frac{t}{T})\epsilon_t\sin(\omega_0 t) \\ u_3 &= T^{-\frac{3}{2}}\sum w(\frac{t}{T})\epsilon_t\, t\cos(\omega_0 t)\ , \ u_4 = T^{-\frac{3}{2}}\sum w(\frac{t}{T})\epsilon_t\, t\sin(\omega_0 t) \end{aligned} \tag{A.11}$$

is asymptotically multivariate normal with zero mean and variance matrix

$$\mathbf{U} = \pi f_{\epsilon\epsilon}(\omega_0)\begin{pmatrix} U_0 & 0 & U_1 & 0 \\ 0 & U_0 & 0 & U_1 \\ U_1 & 0 & U_2 & 0 \\ 0 & U_1 & 0 & U_2 \end{pmatrix}.$$

Expanding $q_T'(\omega)$ about ω_0, we can write:

$$T^{-\frac{1}{2}}q_T'(\omega_0) = -T^{\frac{3}{2}}(\tilde{\omega}_T - \omega_0)T^{-2}q_T''(\overline{\omega}_T),\quad |\overline{\omega}_T - \omega_0| \leq |\tilde{\omega}_T - \omega_0|. \tag{A.12}$$

Notice that calculating the derivative and by repeated use of Lemmas 1 and 2 we can show

$$T^{-\frac{1}{2}}q_T'(\omega_0) = -W_1B_0u_1 + W_1A_0u_2 + W_0B_0u_3 - W_0A_0u_4 + o_p(1). \tag{A.13}$$

Since $T|\tilde{\omega}_T - \omega_0|$ converges to zero in probability, the second derivative and repeated use of Lemmas 1 and 2 yields

$$T^{-2}q_T''(\overline{\omega}_T) = \frac{1}{2}(A_0^2 + B_0^2)(W_1^2 - W_0W_2) + o_p(1). \tag{A.14}$$

Using (A.12), (A.13) and (A.14) we can express the vector of standardized estimates as a linear combination of the vector $\mathbf{u}$, defined by equation (A.11), plus a quantity converging to 0 in probability: $\{T^{\frac{1}{2}}(\tilde{A}_T - A_0), T^{\frac{1}{2}}(\tilde{B}_T - B_0), T^{\frac{3}{2}}(\tilde{\omega}_T - \omega_0)\}' = \mathbf{Au} + \mathbf{o}_p(1)$, with

$$\mathbf{A} = \begin{pmatrix} B_0^2W_2 + A_0^2(W_2 - \frac{W1^2}{W_0}) & -\frac{A_0B_0W_1^2}{W_0} & -B_0^2W_1 & A_0B_0W_1 \\ -\frac{A_0B_0W_1^2}{W_0} & A_0^2W_2 + B_0^2(W_2 - \frac{W_1^2}{W_0}) & A_0B_0W_1 & -A_0^2W_1 \\ -B_0W_1 & A_0W_1 & B_0W_0 & -A_0W_0 \end{pmatrix}.$$

By Condition 2 we know that all the denominators in the components of $\mathbf{A}$ are not 0. This implies that $\mathbf{Au}$ is asymptotically multivariate normal with variance matrix $\mathbf{AUA'}$. By computing $\mathbf{AUA'}$ we obtain the variance expression (20). This proves Theorem 1 for the one sinusoidal case.

Taking derivatives of $q_T(\omega)$ we notice the $\partial q_T(\omega)/\partial\omega_k$ does not depend on ω_l when $l \neq k$. Furthermore, under condition (19), the $\tilde{\omega}_j$'s are asymptotically independent, see for example Brillinger (1981). Theorem 1 now follows for the general case $K > 1$.

D. Proof of Corollary 1

As mentioned above, if we do not impose constraint (7) on the model, then we can rewrite it as model (8) with $K = \sum_{j=1}^J K_j$. For this model let $\tilde{\beta}_T = (\tilde{A}_{1,T}, \tilde{B}_{1,T}, \ldots, \tilde{A}_{K,T}, \tilde{B}_{K,T}, \tilde{\omega}_{1,T}, \ldots, \tilde{\omega}_{K,T})$ be the estimates of β, the parameter of the model defined by (1)–(6), as defined by (11) – (14). Notice that from Theorem 1 we know the asymptotic distribution of $\tilde{\beta}_T$.

Without loss of generality, assume the $\tilde{\omega}_{k,T}$'s are in ascending order. Now define $\hat{\omega}_{1,1,T} = \tilde{\omega}_{1,T}$ and, for each $1 < l \leq K_1$, define $\hat{\omega}_{1,l,T}$ to be the $\tilde{\omega}_{k,T}$ that minimizes $|\tilde{\omega}_{k,T} - l\hat{\omega}_{1,1,T}|$. Let $\hat{\omega}_{2,1,T}$ be the smallest of the $(K - K_1)$ terms $\tilde{\omega}_{j,T}$ that are not used to define the $\hat{\omega}_{1,k,T}$'s, and find $\hat{\omega}_{2,l,T}$ for $l = 2, \ldots, K_2$, as done for $j = 1$. Repeat this procedure for $j = 3, \ldots, J$. Now for each $j = 1, \ldots, J$ define $\hat{\beta}_{j,T} = (\hat{A}_{j,1,T}, \hat{B}_{j,1,T}, \hat{\omega}_{j,1,T}, \ldots, \hat{A}_{j,K_j,T}, \hat{B}_{j,K_j,T}, \hat{\omega}_{j,K_j,T})$ with the $\hat{A}_{j,k,T}$'s and $\hat{B}_{j,k,T}$'s defined as the corresponding estimates to the $\hat{\omega}_{j,k,T}$ using (11) and (12), respectively.

As done in Brillinger (1980), for the case of estimating a bifrequency, we notice that finding the weighted least squares estimate of β^c is asymptotically equivalent to estimating β^c via the following regression model:

$$\begin{pmatrix} \hat{\beta}_{1,T} \\ \vdots \\ \hat{\beta}_{J,T} \end{pmatrix} = X\beta^c + \delta,$$

where X is a block diagonal matrix with jth entry.

$$\begin{pmatrix} X_1 \ldots & 0 & 1X_2 \\ \vdots \ \ddots & \vdots & \vdots \\ 0 \ \ldots & X_1 & K_jX_2 \end{pmatrix}, \text{ where } X_1 = \begin{pmatrix} 1\,0 \\ 0\,1 \\ 0\,0 \end{pmatrix} \text{ and } X_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

and the vector δ has mean 0 and variance matrix V defined by a block diagonal matrix with matrices $V_{j,k}$, defined in (20), in its $(k-1+\sum_{i=1}^{j} K_i)$th diagonal entry.

Consistency and asymptotic normality follow from the fact that the estimates obtained from the regression are linear combinations of the estimates known to be consistent and jointly asymptotically normal from Theorem 1. To find the covariance matrix, we apply weighted regression and see that the new estimates have covariance matrix equal to $(X'V^{-1}X)^{-1}$. Using the result in Rao (1973, p.33), we can directly compute $(X'V^{-1}X)^{-1}$ to obtain the desired result.

E. Proof of Corollary 2

For this proof we assume all sums are over $1,\ldots,H_N$, unless otherwise specified. Without loss of generality assume that H_N is even.

Similar to the proof of Theorem 1 we let

$$\Delta_k^{H_N}(\lambda) = \sum w(\frac{n}{H_N})n^k \exp(i\lambda n). \tag{A.15}$$

Next we develop a parallel result to that of Lemma 1 for the quantity in (A.15).

Lemma 3. *If H_N is a sequence of integers such that $H_N \to \infty$, then*

$$\lim_{H_N\to\infty} H_N^{-(k+1)}\Delta_k^{H_N}(\lambda) = W_k, \ \textit{for } \lambda = 0, 2\pi, \tag{A.16}$$

$$\Delta_k^{H_N}(\lambda) = O(H_N^k), \ \textit{for } 0 < \lambda < 2\pi, \tag{A.17}$$

with W_k defined by (10) *for $k = 0, 1, 2$.*

This follows by noticing that $\Delta_k^{H_N}(\lambda) = \sum w(n/H_N)n^k \exp(i\lambda n)$ is a subsequence of $\sum_{n=1}^N w(n/N)n^k \exp(i\lambda n)$. Then from the proof of Lemma 1, (A.16) and (A.17) hold.

Notice that the equivalent result to Lemma 2, that the quantity defined by $p_N(\lambda) = |(H_N)^{-(k+1)} \sum w(n/H_N)n^k \epsilon_{n+l,N} \exp(-i\lambda n)|$ is such that

$$\lim_{N\to\infty} \sup_{0\leq\lambda\leq\pi} p_N(\lambda) = 0, \text{ in probability} \tag{A.18}$$

follows, since $\{H_N, N \geq 1\}$ is a subsequence of $\{N, N \geq 1\}$.

We first show consistency and asymptotic normality for the case $J = 1, K_1 = 1$. Because $J = 1$ and $K_1 = 1$ we suppress the indexes j and k for simplicity. We start by noticing that

$$\begin{aligned} q_N(\theta) = &\left| H_N^{-1} \sum w(\frac{n}{H_N})\, \epsilon_{n+l,N} \exp(in\theta) \right|^2 \\ &+ \left| H_N^{-1} \sum w(\frac{n}{H_N})\, s\left[\frac{n+l}{N}, \beta_N(\frac{n+l}{N})\right] \exp(in\theta) \right|^2 \\ &+ 2\Re\left[\left(H_N^{-1} \sum w(\frac{n}{H_N})\, \epsilon_{n+l,N} \exp(in\theta) \right) \times \right. \\ &\left. \left(H_N^{-1} \sum w(\frac{n}{H_N})\, s\left[\frac{n+l}{N}, \beta_N(\frac{n+l}{N})\right] \exp(in\theta) \right) \right]. \end{aligned} \tag{A.19}$$

As in Theorem 1 (A.18) implies that the first expression on the right goes to 0 in probability.

By the Mean Value Theorem we have

$$\begin{aligned} s\{t; \beta_N(t)\} = &\{A(t_0) + M_1(t - t_0)\} \cos\left[\{\theta_N(t_0) + M_3(t - t_0)\}t\right] \\ &+ \{B(t_0) + M_2(t - t_0)\} \sin\left[\{\theta_N(t_0) + M_3(t - t_0)\}t\right]. \end{aligned} \tag{A.20}$$

By Condition 4, the constants M_1, M_2 and M_3 are bounded. Since $\sin(t)$ and $\cos(t)$ are bounded functions we can write (A.20) as

$$\begin{aligned} s\{t; \beta_N(t)\} = &A(t_0) \cos\{\theta_N(t_0)t + M_3 t(t - t_0)\} + B(t_0) \sin\{\theta_N(t_0)t \\ &+ M_3 t(t - t_0)\} + M_4(t - t_0), \end{aligned} \tag{A.21}$$

where M_4 is a bounded constant. Notice that by Condition 3 and applying the Mean Value Theorem to the first term on the right of equation (A.21) we have $\cos\{\theta_N(t_0)t + M_3 t(t-t_0)\} = \cos\{N\theta(t_0)t + o(1)\, t + M_3 t(t-t_0)\} = \cos\{N\theta(t_0)t\} + M_5\{o(1)\, t + M_3 t(t-t_0)\}$ where M_5 is a bounded constant. We may find a similar expression for the second term on the right side of equation (A.21).

Let $A_0 = A(t_0)$, $B_0 = B(t_0)$, $\theta_0 = \theta(t_0)$ and, suppressing the N, $\beta_0 = \beta_N(t_0)$. Then since $|t| < 1$ we have that by the continuity of $A(t)$, $B(t)$, and $\theta(t)$,

$s\{n/N;\beta_N(n/N)\} = r(n,\beta_0) + M_6\,(n-n_0)/N + o(1)$, with M_6 bounded and $r(n,\beta_0)=A_0\cos(\theta_0 n)+B_0\sin(\theta_0 n)$. Now notice $|H_N^{-1}\sum(n-H_N/2)/N\exp(i\theta n)| \le (H_N+2)/4N=o(1)$. Since $w(t)$ is bounded and of bounded variation, a summation by parts argument like that in the proof of Lemma 1, gives $|H_N^{-1}\sum w(n/H_N)(n-H_N/2)/N\exp(i\theta n)| = o(1)$. Now we can write the second term in equation (A.19) as

$$\begin{aligned}&\left|H_N^{-1}\sum w(\frac{n}{H_N})\,s\left[\frac{n+l}{N},\beta_N(\frac{n+l}{N})\right]\exp(i\theta n)\right|\\ &=\left|H_N^{-1}\sum w(\frac{n}{H_N})\,r(n+l,\beta_0)\exp(i\theta n)+o(1)\right|\end{aligned}$$

and $r(n+l,\beta_0) = \{D_0\exp(i\theta_0 n)+\overline{D}_0\exp(-i\theta_0 n)\}\exp(i\theta_0 l)$, where $D_0=\frac{1}{2}(A_0-iB_0)$ as before. Next notice that

$$\begin{aligned}&H_N^{-1}\sum w(\frac{n}{H_N})\,r(\frac{n+l}{N},\beta_0)\exp(i\theta n)\\ &=H_N^{-1}[D_0\Delta_0^{H_N}(\theta_0+\theta)+\overline{D}_0\Delta_0^{H_N}(\theta_0-\theta)]\exp(i\theta_0 l).\end{aligned}$$

By Lemma 3 we have that, for $0<\theta<\pi$, $H_N^{-1}\Delta_0^{H_N}(\theta_0+\theta)=o(1)$ and that

$$H_N^{-1}\Delta_0^{H_N}(\theta_0-\theta)=\begin{cases}W_0 & : \quad \theta=\theta_0,\\ o(1) & : \quad \text{otherwise.}\end{cases}$$

Using this fact and (A.18), we have that the third term in (A.19) converges to 0 in probability and $q_N(\theta) = |H_N^{-1}[D_0\Delta_0^{H_N}(\theta_0+\theta)+\overline{D}_0\Delta_0^{H_N}(\theta_0-\theta)]\exp(i\theta_0 l)+o_p(1)|+o_p(1)=\frac{1}{4}(A_0^2+B_0^2)|H_N^{-1}\Delta_0^{H_N}(\theta-\theta_0)|^2+o_p(1)$.

Therefore

$$q_N(\theta_0)=\frac{1}{4}(A_0^2+B_0^2)W_0^2+o_p(1). \tag{A.22}$$

Finally, for any $b>0$, define

$$P_N(b)=\{\theta : H_N|\theta-\theta_0|\ge b\}. \tag{A.23}$$

Notice that as in the proof of Theorem 1, $\Pr(H_N|\hat\theta_N(t_0)-\theta(t_0)|\ge b)\le \Pr(\sup_{\theta\in P_N(b)}|(H_N)^{-1}\Delta_0^{H_N}(\theta_0-\theta)|\ge W_0+o_p(1))$ and that $H_N^{-1}\Delta_0^{H_N}(\theta-\theta_0)=H_N^{-1}\sum w(n/H_N)n\exp\{i(\theta-\theta_0)n\}=\sum w(n/H_N)(n/H_N)\exp\{iH_N(\theta-\theta_0)n/H_N\}$. Again, as in Theorem 1 we have

$$\sup_{\theta\in P_N(b)}|H_N^{-1}\Delta_0^{H_N}(\theta-\theta_0)|=\sup_{\theta\in P_N(b)}\left|\int_0^1 w(s)\exp\{iH_N(\theta-\theta_0)s\}\,ds\right|+o(1),$$

and thus $\lim_{N\to\infty}\Pr\left(H_N|\hat\theta_N(t_0)-\theta(t_0)|\ge b\right)=0$.

Now we will prove consistency for $\hat{A}_N(t_0)$ and $\hat{B}_N(t_0)$. As above, let $\beta_0 = \beta_N(t_0)$, $A_0 = A(t_0)$ and $B_0 = B(t_0)$. Then we have $\hat{A}_N(t_0) = 2(W_0^{H_N})^{-1}\sum_{n=l+1}^{u} w((n-l)/H_N)\,[s\{n/N, \beta(n/N)\} + \epsilon_{n,N}]\cos\{\hat{\theta}_N(t_0)\,n\}$, with $l = n_0 - H_N/2$, $u = n_0 + H_N/2$, and $W_0^{H_N} = \sum w(n/H_N)$. As before we use the Mean Value Theorem to obtain

$$(W_0^{H_N})^{-1}\sum_{n=l+1}^{u} w(\frac{n-l}{H_N})\, s\left\{\frac{n}{N}, \beta(\frac{n}{N})\right\}\cos\{\hat{\theta}_N(t_0)\,n\}$$
$$= (W_0^{H_N})^{-1}\sum_{n=l+1}^{u} w(\frac{n-l}{H_N})\, r(n,\beta_0)\cos\{\hat{\theta}_N(t_0)\,n\} \;+ o(1).$$

Then $\hat{A}_N(t_0) = (W_0^{H_N})^{-1}\sum_{n=l+1}^{u} w((n-l)/H_N)\,\{r(n,\beta_0) + \epsilon_{n,N}\}\cos\{\hat{\theta}_N(t_0)\,n\} + o(1)$. In the same way we obtain $\hat{B}_N(t_0) = (W_0^{H_N})^{-1}\sum_{n=l+1}^{u} w((n-l)/H_N)\,\{r(n,\beta_0) + \epsilon_{n,N}\}\sin\{\hat{\theta}_N(t_0)\,n\} + o(1)$. Since the parameter β_0 is constant over time the result now follows as the proof of Theorem 1.

To prove asymptotic normality, we see that expanding $q_N'(\theta)$ a Taylor series about $\theta(t_0)$, we can write $H_N^{-1/2}q_N'\{\theta(t_0)\} = -H_N^{3/2}\,\{\hat{\theta}_N(t_0) - \theta(t_0)\}\,H_N^{-2}q_N''\{\tilde{\theta}_N(t_0)\}$ for some $|\tilde{\theta}_N(t_0)|$ such that $|\tilde{\theta}_T(t_0) - \theta(t_0)| \le |\hat{\theta}_N(t_0) - \theta(t_0)|$. Using (A.22), Lemma 3, and the argument to obtain (A.18), we can proceed as in the proof of Theorems 1 and Corollary 1 to arrive at the desired result for the general case.

References

Brown, E. N. (1990). A note on the asymptotic distribution of the parameter estimates for the harmonic regression model. *Biometrika* **77**, 653-656.

Brillinger, D. R. (1980). The comparison of least squares and third–order periodogram procedures in the estimation of bifrequency. *J. Time Ser. Anal.* **1**, 95-102.

Brillinger, D. R. (1981). *Time Series Data Analysis and Theory.* Holden–Day, San Francisco.

Greenhouse, J. B., Kass, R. and Tsay, R. (1987). Fitting nonlinear models with ARMA errors to biological rhythm data. *Statist. Medicine* **6**, 167-183.

Hannan, E. J. (1971). Non-linear time series regression. *J. Appl. Probab.* **8**, 767-780.

Hannan, E. J. (1973). The estimation of frequency. *J. Appl. Probab.* **10**, 510-519.

Hannan, E. J. (1974). Time series analysis. *IEEE Trans. Auto Control* **AC-19**, 706-715.

He, S. (1984). *On Estimating the Hidden Periodicities in Harmonic Analysis.* M.S. Thesis, Beijing University, Beijing.

Irizarry, R. A. (1998). *Statistics and Music: Fitting a Local Harmonic Model to Musical Sound Signals.* Ph.D. Thesis, University of California, Berkeley.

Quinn, B. G. (1989). Estimating the number of terms in a sinusoidal regression. *J. Time Ser. Anal.* **10**, 12-31.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications.* John Wiley and Sons, New York.

Rodet, X. (1997). Musical sound signals analysis/synthesis: Sinusoidal+residual and elementary waveform models. *Proc. IEEE Time-Frequency and Time-Scale Workshop*, 1-10.

Walker, A. M. (1971). On the estimation of a harmonic component in a time series with stationary independent residuals. *Biometrika* **58**, 21-36.

Wang, X. (1991). An AIC type estimator for the number of cosinusoids. *J. Time Ser. Anal.* **14**, 433-440.

Johns Hopkins University, 615 N. Wolfe Street, Baltimore, Maryland 21205, U.S.A.

E-mail:ririzarr@jhsph.edu

(Received January 1999; accepted February 2000)