A STATISTICAL TEST TO DISCERN RANDOM FROM CONSTANT AMPLITUDE HARMONICS

Guotong Zhou
School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250, USA.

ABSTRACT

Periodogram is an important tool to reveal hidden periodicities in a given time series but does not tell whether the resulting spectral lines are associated with constant or random amplitude harmonics. Applications dealing with random amplitude models include Doppler spread targets and detection in the presence of fading. We propose to estimate the variance of the harmonic amplitude and then make the decision based on whether the variance can be regarded as zero in a statistical sense. This is a viable approach because any constant has variance zero whereas any real random process has a positive variance. A rigorous statistical test is formulated and illustrated with simulations.

1. INTRODUCTION

Detection of hidden periodicities embedded in a random process has been a concern over one hundred years. Schuster in 1894 devised the periodogram as a means of searching for hidden periodicities. It has had much success in many areas ranging from seasonal and economic time series analysis, seismology, geophysics, spectroscopy, and communications to sonar and radar signal processing (see e.g., [1], [2], [4], [6], [8] and references therein).

If the periodogram of a process shows peaks at \( \pm \omega_0 \), we are led to believe that a cosine, \( \cos(\omega_0 t + \phi_0) \), of some form is present in the data. Discrete-time processes are considered in this paper and we shall discuss two possibilities:

\[
x(t) = A \cos(\omega_0 t + \phi_0) + v(t),
\]

and

\[
x(t) = s(t) \cos(\omega_0 t + \phi_0) + v(t).
\]

In (2), multiplicative noise (random amplitude) \( s(t) \) and additive noise \( v(t) \) are assumed to be real, stationary, mixing, and mutually independent. The mixing condition [3, p. 8] ensures that the cumulants of \( s(t) \) and \( v(t) \) are absolutely summable and hence the corresponding (higher-order) spectra are finite. The k-th order cumulant of \( s(t) \) at lags \( (\tau_1, \tau_2, \ldots, \tau_{k-1}) \) is defined as \( c_k(s(t), s(t + \tau_1), \ldots, s(t + \tau_{k-1})) \), and similarly for \( v(t) \). Performance analysis results of this paper require that \( \omega_0 \not\equiv 0 \mod (\pi) \) which are usually met in practice. The phase \( \phi_0 \) is assumed to be deterministic here because we only consider single record detection and estimation. Note that the constant amplitude harmonic model (1) can be regarded as a special case of (2) with \( s(t) \equiv A \). The objective of this paper is to devise statics to test the following hypotheses: \( H_0 : s(t) \text{ constant} = A \) vs. \( H_1 : s(t) \text{ random} \).

Random variations of the form (2) are known to occur in real systems. As an example, in radar processing, when a nonpoint target is fast maneuvering or scintillating, the resulting covariance due to Doppler shift carries a random amplitude [9]. In underwater acoustical applications, when the medium is dispersive or fluctuating, the sonar return also experiences the random amplitude effect [6]. The model in (2) is also appropriate for Doppler weather radar/sonar returns, where \( s(t) \) is due to the randomness of the scatterers (hydro-meteors or aerosol particles). Due to carrier modulation, (2) appears with timing and carrier synchronization of communications signals as well.

It is important to determine the correct underlying model at least for the following reasons: 1) \( s(t) \) random or not reveals partial information about the source (target) such as scattering and fading [7]; 2) The Cramér-Rao bounds on the parameter estimates are different for the two models; 3) The corresponding maximum likelihood (ML) estimates are also different.

Main contributions of this paper are: (i) methods of estimating the mean \( m_s \) and variance \( \sigma^2_s \) of \( s(t) \) and closed-form variance expressions of these estimates; (ii) formulation of a rigorous statistical test to determine the zeroeness of the \( \sigma^2_s \) estimate. Based on the test result, we then declare whether random \( s(t) \) or constant \( s(t) = A \) is present in the data. The variance expressions derived in this paper can also be used to predict the reliability of these estimates.

2. PERIODOGRAM ANALYSIS

The raw periodogram of the discrete-time \( x(t) \) is defined as

\[
I_{2x}(\lambda) = \frac{1}{T} \sum_{t=0}^{T-1} x(t) e^{-j\lambda t} = \frac{|X_T(\lambda)|^2}{T},
\]

where \( X_T(\lambda) \) is the DFT of the data. If \( x(t) = 0 \) of zero-mean stationary, then it is well known that \( I_{2x}(\lambda) \) is an asymptotically unbiased but inconsistent estimator of the power spectral density (PSD) of \( x(t) \) (see e.g., [3]).

Now suppose that the \( s(t) \) in (2) has non-zero mean. Then both (1) and (2) are cyclostationary and have nonzero and periodically time-varying mean. Previous results on periodograms of zero-mean and stationary processes do not apply. When \( m_s := E[s(t)] \neq 0 \), we can show that as long as \( s(t) \) and \( v(t) \) are absolutely summable covariance functions (or equivalently, their power spectra \( S_{2x}(\omega) \) and \( S_{2v}(\omega) \) are finite), then for \( T \) large, the expected value of \( I_{2x}(\lambda) \) of (2) has

\[
E[I_{2x}(\lambda)] \approx T \left[ \frac{m_s^2}{4} \delta(\lambda + \omega_0) + \frac{m_s^2}{4} \delta(\lambda - \omega_0) + m_n^2 \delta(\lambda) \right].
\]

Therefore when \( m_s \neq 0 \), second-order information is "lost" in the periodogram. Since (1) can be regarded as a special case of (2) with \( s(t) \equiv A \), we have for model (1),

\[
E[I_{2x}(\lambda)] \approx T \left[ \frac{A^2}{4} \delta(\lambda + \omega_0) + \frac{A^2}{4} \delta(\lambda - \omega_0) + m_n^2 \delta(\lambda) \right].
\]
Therefore, if the amplitude $A$ in (1) is the same as the mean $m_s$ in (2), the two processes will have almost identical raw periodograms and thus are indistinguishable.

Since $\sigma_s^2 > 0$ for $s(t)$ real and random and $\sigma_s^2 = 0$ for $s(t)$ constant, our approach to differentiating the two models is to estimate $\sigma_s^2$ and then check on the zeroeness of $\sigma_s^2$.

3. ESTIMATION OF $\sigma_s^2$

Our goal is to estimate $\sigma_s^2 = m_{ss} - m_s^2$. The mean $m_s = E[s(t)]$ will be retrieved from the cyclic mean of $x(t)$, and the mean square $m_{ss} = E[s^2(t)]$ will be estimated from the cyclic mean square of $x(t)$.

3.1. Cyclic mean – the estimation of $m_s$

A quantity that is closely related to the periodogram is the so-called cyclic mean. If $x(t)$ is cyclostationary, then its time-varying mean, denoted as $m_{1x}(t)$, is an almost periodic function of $t$. Hence its FS coefficient, termed the cyclic mean,

$$C_{1x}(\alpha) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} m_{1x}(t) e^{-j\alpha t}, \quad (4)$$

exhibits peaks at some $\alpha$. It is straightforward to show that the cyclic mean of (2) is

$$C_{1x}(\alpha) = m_s \frac{e^{j\phi_0}}{2} \delta(\alpha - \omega_0) + m_s \frac{e^{-j\phi_0}}{2} \delta(\alpha + \omega_0) + m_v \delta(\alpha).$$

The cyclic mean of (1) can be obtained simply by replacing $m_s$ by $A$.

The following cyclic mean estimator can be shown to be asymptotically unbiased and m.s.s. consistent [5]:

$$\bar{C}_{1x}(\alpha) = \frac{1}{T} \sum_{t=0}^{T-1} x(t) e^{-j\alpha t}, \quad (5)$$

We recognize that (5) is nothing but the normalized (by data length $T$) DFT of the data and can be computed using the FFT algorithm. Its amplitude is related to the periodogram through $S_{x}(\lambda) = T|\bar{C}_{1x}(\lambda)|^2$.

Estimates of $\omega_0$, $\phi_0$, and $m_s$ are constructed based on $C_{1x}(\alpha)$ as follows:

$$\hat{\omega}_0 = \arg \max_{\alpha > 0} |\bar{C}_{1x}(\alpha)|, \quad (6)$$

$$\hat{\phi}_0 = \arg \{\bar{C}_{1x}(\omega_0)\}, \quad (7)$$

$$\hat{m}_s = 2 \text{ Re} \left[e^{-j\hat{\omega}_0} \bar{C}_{1x}(\hat{\omega}_0)\right] = \frac{2}{T} \sum_{t=0}^{T-1} x(t) \cos(\hat{\omega}_0 t + \hat{\phi}_0). \quad (8)$$

In [11], we have shown that when the SNR is moderate to high (which requires a combination of good $m_s^2/\sigma_s^2$ and $m_v^2/\sigma_v^2$ ratios), then the estimators in (6)-(8) will be close to their true values with the following large sample variance [11]:

$$\text{var}(\hat{\omega}_0) = \frac{1}{T^2} \left[ \frac{24 S_{2x}(\omega_0)}{m_s^2} + \frac{6 S_{2x}(2\omega_0)}{m_s^2} \right], \quad (9)$$

$$\text{var}(\hat{\phi}_0) = \frac{1}{T} \left[ \frac{8 S_{2x}(\omega_0)}{m_s^2} + \frac{2 S_{2x}(2\omega_0)}{m_s^2} \right], \quad (10)$$

$$\text{var}(\hat{m}_s) = \frac{1}{T} \left[ \frac{S_{2x}(0)}{S_{2x}(\omega_0)} + \frac{1}{2} \frac{S_{2x}(2\omega_0)}{S_{2x}(\omega_0)} + 2 S_{2x}(\omega_0) \right]. \quad (11)$$

We emphasize that neither the parameter estimation algorithm nor the variance expressions depend on the distributions of $s(t)$ and $v(t)$.

Under the same SNR assumption, we have also shown in [11] that (8) is asymptotically equivalent to the following:

$$\hat{m}_s = \frac{2}{T} \sum_{t=0}^{T-1} x(t) \cos(\omega_0 t + \phi_0), \quad (12)$$

which removes the finite-sample dependence of (8) on $\omega_0$ and $\phi_0$ and makes large sample performance analysis tractable.

3.2. Cyclic mean square – the estimation of $m_{ss}$

Now let us consider the time-varying mean square of $x(t)$,

$$m_{2x}(t) = E[x^2(t)] = m_s \cos^2(\omega_0 t + \phi_0) + m_v^2 + 2 m_s m_v \cos(\omega_0 t + \phi_0). \quad (13)$$

Since $m_{2x}(t)$ is a periodic function of $t$, we consider its FS coefficients, which we term the cyclic mean square of $x(t)$,

$$M_{2x}(\alpha) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} m_{2x}(t) e^{-j\alpha t} = \left( m_{2x} + \frac{m_{2x}}{2} \right) \delta(\alpha) + m_z m_v e^{j\phi_0} \delta(\alpha - \omega_0) + m_s m_v e^{-j\phi_0} \delta(\alpha + \omega_0) + \frac{m_{2x}}{4} e^{j\phi_0} \delta(\alpha - 2\omega_0) + \frac{m_{2x}}{4} e^{-j\phi_0} \delta(\alpha + 2\omega_0). \quad (14)$$

Consistent sample estimate of $M_{2x}(\alpha)$ is given by [5]

$$\hat{M}_{2x}(\alpha) = \frac{2}{T} \sum_{t=0}^{T-1} x(t) e^{-j\alpha t}, \quad (15)$$

and hence $m_{2x}$ can be estimated via

$$\hat{m}_{2x} = \arg \max_{\alpha > 0} \left| \hat{M}_{2x}(\alpha) \right|. \quad (16)$$

Mimicking the steps used in [11], we can show that (16) can be approximated by

$$\hat{m}_{2x} = \frac{4}{T} \sum_{t=0}^{T-1} x(t) \cos(2\omega_0 t + 2\phi_0), \quad (17)$$

and its variance analysis is discussed next.

When $m_s = m_v = 0$, the variance expression of the $m_{2x} = \sigma_s^2$ estimate was derived in [11]. The variance expression of $m_{2x}$ for the general $m_s \neq 0$, $m_v \neq 0$ case is presented here. Detailed derivation can be found in [12].

Define

$$h_1(\tau) = c_{4x}(0, \tau, \tau) + 2 c_{2x}(\tau), \quad (18)$$

$$h_2(\tau) = c_{4x}(0, \tau, -\tau) + 2 c_{2x}(\tau), \quad (19)$$

$$h_3(\tau) = 4 c_{2x}(\tau) c_{2x}(\tau), \quad (20)$$

$$h_4(\tau) = c_{4x}(\tau, \tau) + c_{4x}(0, \tau, 0), \quad (21)$$

$$h_5(\tau) = c_{3x}(\tau, \tau) + c_{3x}(0, \tau, 0), \quad (22)$$

$$H_1(\lambda) = \sum_{\tau=-\infty}^{\infty} h_1(\tau) \cos(\lambda \tau) = \sum_{\tau=-\infty}^{\infty} h_1(\tau) \exp(-j\lambda \tau).$$

The large sample variance of $\hat{m}_{2x}$ can be shown to be [12]

$$\text{var}(\hat{m}_{2x}) = \frac{1}{T} \left[ H_1(0) + 2 H_1(2\omega_0) + \frac{1}{2} H_1(4\omega_0) \right].$$
5. STATISTICAL TEST FOR $\hat{\sigma}^2$

It can be shown that when $T$ is large, $\sqrt{T}(\hat{\sigma}^2 - \sigma^2)$ is asymptotically Gaussian distributed [5] with mean zero and variance given by $T$ times the r.h.s. of (27). We postulate the following two hypotheses:

$H_P_0$: $s(t)$ constant $= A$  vs.  $H_P_1$: $s(t)$ random

Under $H_P_0$, $\sigma^2 = 0$, and hence $\sqrt{T}(\hat{\sigma}^2)$ has mean zero and variance $T\text{var}(\hat{\sigma}^2)$. As a result, $T(\hat{\sigma}^2)^2/T\text{var}(\hat{\sigma}^2) = (\hat{\sigma}^2)^2/\text{var}(\hat{\sigma}^2)$ is $\chi^2(1)$ distributed. The probability of false alarms is defined as

$$P_{FA} = \Pr \left\{ \frac{(\hat{\sigma}^2)^2}{\text{var}(\hat{\sigma}^2)} > \Gamma \mid H_P_0 \right\}$$  \hspace{1cm} (29)

For a given $P_{FA}$, we first determine a threshold $\Gamma = \Gamma\text{var}(\hat{\sigma}^2)$, and upon receiving a $\hat{\sigma}^2$ estimate, we compare $(\hat{\sigma}^2)^2$ with $\Gamma$, accept $H_P_0$ if $(\hat{\sigma}^2)^2$ is below $\Gamma$ and reject $H_P_0$ if otherwise.

Under $H_P_0$, $s(t) \equiv A$; hence, except for the mean $m_s = A \neq 0$, all cumulants of $s(t)$ are zero: $h_1(\tau) = h_4(\tau) = 0$. It follows from (27) that

$$\text{var}(\hat{\sigma}^2) = \frac{8m_s^2}{T} S_{2v}(3\omega_0) + \frac{32\sigma_r^2}{T} S_{2v}(2\omega_0) + \frac{16m_r}{T} H_5(2\omega_0)$$

under $H_P_0$. If $v(t)$ is symmetrically distributed, then $H_5(\lambda) = 0$. If $v(t)$ is Gaussian, then $H_5(\lambda) = 0$ and $H_2(\lambda) = \sum 2c_{2v}(\lambda) \cos(\lambda \tau)$. In practice, we need to estimate $\text{var}(\hat{\sigma}^2)$ (and hence $\Gamma$) from the same data. Operating under $H_P_0$, we first remove $m_s \cos(\omega_0 t + \phi_0)$ from $x(t)$. The residue is regarded as an approximation of $v(t)$ (recall that $m_s = A$ here) and many available (polyspectral estimation procedures can be followed to estimate $S_{2v}(2\omega_0), S_{2v}(3\omega_0), H_3(2\omega_0)$, and $H_4(2\omega_0)$. The mean of $v(t)$ can be estimated by simply taking the running average of the data,

$$\hat{m}_v = \frac{1}{T} \sum_{t=0}^{T-1} x(t)$$  \hspace{1cm} (31)

When $v(t)$ is white Gaussian, the variance expression (30) further simplifies to

$$HP_0: \quad \text{var}(\hat{\sigma}^2) = \frac{16\sigma_r^4}{T} + \frac{8\sigma_r^2(m_s^2 + 4m_r^2)}{T}$$  \hspace{1cm} (32)

In order to estimate the above expression, we first obtain $\hat{m}_s$ and $\hat{m}_v$ as in (12) and (31), and then calculate

$$HP_0: \quad \hat{\sigma}^2 = \frac{1}{T} \sum_{t=0}^{T-1} x^2(t) - \frac{\hat{m}_s^2 + \hat{m}_v^2}{2}$$  \hspace{1cm} (33)

The overall algorithm is summarized next.

Algorithm

Step 1: For a given probability of false alarms $P_{FA}$, find the corresponding $\Gamma$ from a $\chi^2(1)$ table.

Step 2: Obtain estimates $\hat{\omega}_0$ (6), $\hat{\phi}_0$ (7), $\hat{m}_s$ (8), $\hat{m}_s$ (16), $\hat{\sigma}^2$ (24), and $\hat{m}_v$ (31).
Step 3: Subtract $\tilde{m}_s \cos(\omega_0 t + \phi_0)$ from $x(t)$ and treat the resulting process as $v(t)$. Follow existing (poly)spectral estimation procedures to estimate $S_{2e}(2\omega_0)$, $S_{v}(3\omega_0)$, $H_v(2\omega_0)$ and $H_v(2\omega_0)$ in order to calculate (30). Multiply the result by $\Gamma$ to obtain $\tau$. When it is known a priori that $v(t)$ is white Gaussian, we only need to calculate $\sigma^2_v$ (33) in order to obtain $\text{var}(\sigma^2_v)$ (32).

Step 4: If $(\hat{\sigma}^2_v)^2 < \tau$, we declare that $x(t)$ comes from the constant amplitude model (1); otherwise, it is more likely that $x(t)$ obeys the random amplitude model (2).

6. SIMULATIONS

We present here some numerical examples to verify the performance analysis results and the random amplitude detection algorithm presented in this paper.

Example 1: Verification of variance expressions

We first generated $w(t)$ which was i.i.d. one-sided exponential deviates with p.d.f. $f_W(w) = e^{-w}$. We then removed its mean and passed the mean compensated process $\tilde{w}(t) = w(t) - 1$ through a first order FIR filter with parameters $[1, 0.5]$. We then added a constant $m_s = 1$ at the output end and obtained a colored, non-Gaussian, and non-zero mean process $s(t)$. Additive noise $v(t)$ was white Gaussian with mean $m_v = -0.2$ and variance $\sigma^2_v = 0.2$. We then generated $T = 1,024$ points of $x(t)$ according to (2) with $\omega_0 = 1$, $\phi_0 = 1.8$. 200 independent realizations were used to yield the empirical asymptotic variance results. The available data were zero-padded to length $2^{14}$ when calculating the sample cyclic mean via FFT. Table I illustrates empirical vs. theoretical mean and asymptotic variance results.

Table I. Empirical and theoretical performance of estimators, colored noise case

<table>
<thead>
<tr>
<th>Formulas</th>
<th>(12), (17)</th>
<th>(8), (16)</th>
<th>Theoretical</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E[m_s]$</td>
<td>1.0038</td>
<td>1.0047</td>
<td>1.0060</td>
</tr>
<tr>
<td>$T\text{var}(m_s)$</td>
<td>1.5063</td>
<td>1.5002</td>
<td>1.4897</td>
</tr>
<tr>
<td>$E[m^2_v]$</td>
<td>2.3763</td>
<td>2.4038</td>
<td>2.3600</td>
</tr>
<tr>
<td>$T\text{var}(m^2_v)$</td>
<td>92.6108</td>
<td>92.5800</td>
<td>84.0406</td>
</tr>
<tr>
<td>$E[\tilde{\sigma}^2_v]$</td>
<td>1.3672</td>
<td>1.3929</td>
<td>1.3600</td>
</tr>
<tr>
<td>$T\text{var}(\tilde{\sigma}^2_v)$</td>
<td>66.7438</td>
<td>66.7789</td>
<td>59.2214</td>
</tr>
</tbody>
</table>

Example 2. We have available $T = 1,024$ data $x(t)$ from (1) where $v(t)$ was white Gaussian with mean $m_v = 0.5$ and variance $\sigma^2_v = 0.2$. The harmonic parameters are $A = 1$, $\omega_0 = 1$, $\phi_0 = 0.3$ is the false alarm rate chosen was $P_{FA} = 0.05$ which corresponds to $\Gamma = 3.841$. We followed the algorithm outlined in Section 5 and generated 500 independent realizations. In Fig. 1 we show the $(\tilde{\sigma}^2_v)^2$ estimates in solid lines, the true threshold $\tau$ in dashed line, and the estimated threshold $\tilde{\tau}$ in dotted line. Out of 500 $(\tilde{\sigma}^2_v)^2$ estimates, 21 exceeded $\tau$ and 24 exceeded $\tilde{\tau}$. Both numbers are close to the expected total number of false alarms $0.05 \times 500 = 25$.

7. CONCLUSIONS

Periodogram is a conventional tool to check whether a harmonic of some form is present in the given data. However, it does not tell exactly what form of harmonic is involved. Random amplitude (or multiplicative noise as is often called) appears in many important applications and its nature reflects certain characteristics of the source. In this paper, we first introduced ways of extracting information about the amplitude, such as its mean $m_s$ and variance $\sigma^2_v$. We then analyzed the performance of the $\tilde{m}_s$ and $\tilde{\sigma}^2_v$ estimates. The value of $\sigma^2_v$ can be used as a quantitative measure for target spread or source (in)coherence in Doppler applications. To make a decision as to whether the harmonic amplitude can be regarded as truly random, we compare $(\tilde{\sigma}^2_v)$ with a threshold normalized by the variance of $\sigma^2_v$ and employ a $x^2(1)$ test. The algorithms and variance expressions developed in this paper are also easy to implement as illustrated by the numerical simulations.

REFERENCES


