Coherently Averaged Power Spectral Estimate for Signal Detection

Hualin Lan[a,b], Paul R. White[c], Na Li[b], Jianghui Li[c], Dajun Sun[a,b,*]

[a] *The Acoustic Science and Technology Laboratory and the College of Underwater Acoustic Engineering, Harbin Engineering University, Harbin, Heilongjiang,150001, China.*
[b] *The College of Underwater Acoustic Engineering, Harbin Engineering University, Harbin, Heilongjiang, 150001,China.*
[c] *The Institute of Sound and Vibration Research, University of Southampton, Southampton, SO17 1BJ, U.K.*

Abstract

A common approach to detect sinusoidal signals buried in noise is based on spectral analysis, such as the periodogram. The fluctuations of the spectral components associated with the noise can be alleviated via incoherent averaging of the power spectral estimates of each segment, which is the basis of Welch's method. However, Welch's method only utilizes the incoherent information between segments of signals. In this paper, we propose a method of coherent averaging between segments, which enhances ratio of time-invariant sinusoidal signals relative to the level of the noise background. The gain of coherent averaged power spectral estimate has been derived in terms of time duration of the signal. The proposed method provides a flexible, computationally efficient implementation of signal detection, which can be formulated to allow for various integration times to be realised in different frequency bands. Simulation and experimental results show that the proposed method outperforms the Welch's method and the periodogram method.

Keywords: Signal detection; Power spectral estimation; Coherent averaging.

1. Introduction

Detection of sinusoidal signals in wideband noise is a problem which has been discussed in applications for decades, such as communications, radar, sonar, seismology, and speech processing [1, 2, 3, 4, 5, 6, 7]. Power Spectrum Estimation (PSE) methods, either parametric or non-parametric, are effective tools because the signal's energy is highly localized in the frequency domain [8] and the noise is widely distributed. Non-parametric methods [3, 9, 10, 11] are often used, not only because they are widely applicable and familiar, but also because they are based on the Fourier Transform allowing computational efficiency of the Fast Fourier Transform (FFT).

One of the methods for PSE is the periodogram [12, 13, 14, 15, 16, 17, 18]. The periodogram consists of simply computing the squared magnitude of a windowed time-series and applying a suitable normalization. The periodogram is an inconsistent estimator of the power spectral density [15], which is not suitable to PSE. However, it is the Bayesian optimal estimator of frequency for a single sinusoidal wave in additive white Gaussian noise [16]. In addition, the periodogram can be used as the detection statistic in a generalized likelihood ratio test [18]. The performance of the periodogram for detecting pure sinusoidal waves in white Gaussian noise is well understood [17, 18]. Assuming the sinusoidal wave is frequency stable, then the detection performance of the periodogram improves as the duration of the signal analysed increases since the energy could be accumulated coherently within this frequency bin.

To avoid this issue of the statistical inconsistency of the periodogram as PSE [15], the principle of averaging the periodogram across the data segments, using a rectangular windowing function [19, 20], was later generalized to allow the use of an arbitrary windowing function. This general technique of averaging the periodogram across segments is usually known as either segment averaging or Welch's method [9, 21, 22, 11] and is the most common form of non-parametric PSE [23, 24]. Inherent in Welch's method requires to select the segment length, which, for a fixed total length of data, means that one engages in a trade-off between bias in the spectral estimate and its variance [25, 15].

The problem we focus on is the detection of multiple sinusoidal waves in broadband noise. In terms of PSE, which relates to the problem of estimating the spectrum of a mixed process, i.e., a time-series

containing components with a discrete spectrum, representing the sinusoids, combined with a process with a continuous spectrum [26]. The goal of the PSE, is to detect the sinusoidal components in noise. In this context, despite the periodogram is inconsistent, and is a poor choice as an estimator of the power spectrum, it is an optimal detector [18]. For detecting sinusoidal signals in noise, the Welch's method is inferior to the periodogram, because the segment averaging does not exploit the coherence of the phase of the sine wave: in this work it is assumed that the sinusoid is idealized, in which its phase is coherent throughout the measurement, indicating that its frequency is stable. Wang [27] and Chen [28] have designed the detector of tonal signal by compensating the phase difference between segments. However, the procedure of phase difference estimation is complicated and is not suitable for weak signal.

In practical implementations of the periodogram, it is often the case that the data is received in an on-line mode, such that one is usually forced to select as a priori fixed window size. This needs to compute the periodogram in segments (a process which effectively forms a spectrogram [3, 29, 30, 31, 9, 10], albeit the result that might not be displayed or regarded in that manner). For optimal detection performance the window length plays an important role. It should be selected to match the length of time over which the phase of the sinusoid can be regarded as coherent, and this duration is not usually known as a priori. It may also be the case that the coherence length may not be the same for all the narrow-band signal components. Further, in practical systems the choice of segment length may not be subject to other pressing considerations, so detection performance can be sacrificed. The utility of this method will be explained in the context of an example application, specifically passive sonar. Traditional passive sonar systems employ a spectrogram-based approach, where it may be referred to as a lofargram [32]. Such systems usually employ a fixed FFT size and compute the periodogram for each windowed data block.

In an idealized scenario of stable the sinusoidal signal, the detection performance improves as the FFT size increases. However, there are often considerations that may limit the choice of FFT size for instance: long window leads to reduced updating rates, reduced temporal resolution (through the uncertainty principle) and for real-world signal a sinusoidal wave remains coherent only for a finite time (although this time is usually an unknown priori). We propose a method which intends to allow the integration of coherent information between windows. It can be regarded as a variation of Welch's method based on coherent averaging, as opposed to the incoherent. By using this method, the signal-to-noise ratio (SNR) is enhanced, and the detection performance is increased, making it suitable for the detection of weak but stable sinusoidal signals. The proposed method has been verified by using simulation and filed data. The simulation is performed in the background of white Gaussion noise, and the field data was collected from noisy underwater acoustic channel at 80 m depth. The results show that in both cases, the proposed method outperforms the Welch's method and the periodogram method.

This paper is organized as follows. Section 2 describes the coherent averaging power spectrum method, including the Welch's method, the proposed coherently averaged power spectral estimate (CAPSE), and the relationship between the CAPSE and Fourier Transform of full sequence. Section 3 compares the proposed method with the Welch's method using simulation and experimental data. Section 4 concludes the paper with remarks.

2. Coherent Averaging Power Spectrum Estimation

For convenience the following is a brief review of Welch's method. The transmitted signal $x(n)$ is expressed as

$$x(n) = s(n) + z(n), 0 \leq n \leq N - 1, \tag{1}$$

where $s(n)$ and $z(n)$ are signal and noise respectively, n represents sample number, and N is the number of samples in the signal. We initially assume that the signal $s(n)$ comprises a single sinusoidal signal with the amplitude A, angular frequency ω_0 and phase φ, i.e. $s(n) = A\cos(n\omega_0 + \varphi)$; $z(n)$ is a noise process with zero mean and variance σ_z^2.

As shown in Fig.1, for the Welch's method, the signal $x(n)$ is divided into K segments of length M with an overlap of $(M - D)$, so that the signal in the kth segment is expressed as:

$$
\begin{aligned}
x_k(m) &= x(m + kD), \\
m &= 0, 1, \ldots, M - 1, k = 0, 1, \ldots, K - 1.
\end{aligned}
\tag{2}
$$

The Fourier transform of a data segment, $X_k(\omega)$, is defined as

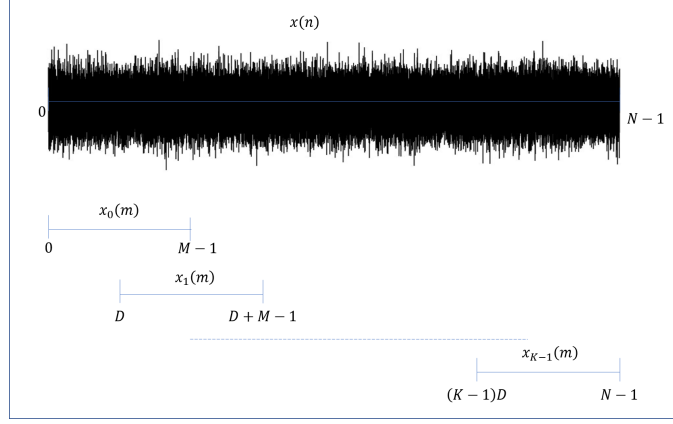$$X_k(\omega) = \sum_{m=0}^{M-1} w(m) x_k(m) e^{-jm\omega}, \tag{3}$$

2

Figure 1: Illustration of the segmentation of the signal employed within the Welch's method

where $w(m)$, $m = 0, 1, \ldots, M - 1$ is a spectral window function [30, 31, 33]. The periodogram of the kth segment is given by

$$I_k(\omega) = \frac{1}{MU}|X_k(\omega)|^2, \tag{4}$$

where U denotes the power of the window:

$$U = \frac{1}{M} \sum_{m=0}^{M-1} w^2(m). \tag{5}$$

According to the Welch's method, an estimate of the PSD is obtained through averaging the periodograms of all the segments, specifically:

$$P_{Welch}(\omega) = \frac{1}{K} \sum_{k=0}^{K-1} I_k(\omega). \tag{6}$$

2.1. Preliminary Observations

The principle of CAPSE stems from the concept of coherent addition of signals, similar to the principle that underlies delay and sum beamforming [4, 1, 5], in which the signals from different channels are summarised, or averaged, coherently after appropriate compensation for the modelled phase differences. CAPSE needs to compensate for the phase differences introduced as a result of the time shifts between different segments.

For the signal $x(n)$, as defined in Eq.(1), Welch PSE in Eq.(6) can be rewritten as

$$P_{Welch}^{xx}(\omega) = \frac{1}{KMU} \sum_{k=0}^{K-1} (|S_k(\omega)|^2 + |Z_k(\omega)|^2 + 2Re\{S_k(\omega)^* Z_k(\omega)\}), \tag{7}$$

where $S_k(\omega)$ and $Z_k(\omega)$ are the (windowed) Fourier transforms of the narrow-band signal and noise components of $x_k(m)$ respectively. The third term in the right side of Eq.(8) is a cross-term. Assume that the noise and signal are uncorrelated, the term will tend to zero as K increases. The two remaining terms are averages of the periodograms of the noise and narrow-band components.

For a single sinusoid in noise, the Fourier transform of the signal in each segment can be expressed as

$$S_k(\omega) = S_0(\omega)e^{j\phi_k}, \tag{8}$$

where

$$\phi_k = \omega_0 kD, \tag{9}$$

which represents the phase difference at the frequency ω_0 between the Fourier transforms of the kth and 1st segment.

3

2.2. Fractional phase compensation of CAPSE

The aim here is to coherently average the signal over multiple segments and increase the SNR. This can be achieved using

$$\bar{X}(\omega) = \frac{1}{K} \sum_{k=0}^{K-1} X_k(\omega) e^{-j\phi_k}$$

$$= \frac{1}{K} \sum_{k=0}^{K-1} X_0(\omega) e^{-j\omega_0 kD} \tag{10}$$

$$= S_0(\omega) + \frac{1}{K} \sum_{k=0}^{K-1} Z_k(\omega) e^{-j\omega_0 kD}.$$

Under mild constraints on the noise process, the SNR is achieved via this averaging increase by $10\log_{10}(K)$, this improves the detection performance. However, from the Eq. (10) it can be seen that the phase difference ϕ_k is highly sensitive to errors in ω_0 since the time delay kD can be significant. The implication is that the one needs a high level of precision in the estimation of frequency ω_0 to realize the anticipated SNR gains.

Considering a standard M-point Discrete Fourier Transform (DFT) [34] which represents data at a discrete set of frequencies, i.e., the frequency bins, the centre frequencies of these bins, ω_l can be expressed as

$$\omega_l = 2\pi \frac{m}{M}, m = 0, 1, \ldots, M - 1. \tag{11}$$

For the case of a single sinusoidal signal in noise, difference between the true frequency of the sinusoid ω_0, and the centre frequency of the closest DFT bin, is denoted $\delta\omega_0$ and can be expressed as

$$\delta\omega_0 = \omega_0 - \omega_m = \frac{2\pi\delta_0}{M}, |\delta_0| \leq \frac{1}{2}, \tag{12}$$

where δ_0 is fractional frequency offset of the sinusoid relative to the nearest frequency bin in the DFT. In the absence of noise the DFT of the kth segment, at frequency ω_m, can be related to that of the first window as

$$X_k(\omega_m) = X_0(\omega_m) e^{j\delta\omega_0 kD}, \tag{13}$$

where $\delta_\phi = \delta_{\omega_0}$, kD is defined as the fractional phase difference between segments.

According the Eq.(13), the time series of the DFT values associated with the frequency bin, ω_l, are themselves complex sinusoidal signals, the frequency of which corresponds to the difference between the centre frequency of the bin and the true frequency of the sinusoid. Thus, we can detect the presence of such signals by applying DFT to each frequency bin:

$$X(\omega_l, \omega_v) = \sum_{k=0}^{K-1} X_k(\omega_l) e^{-j\omega_v k}. \tag{14}$$

Eq.(14) defines a phase compensated average of the Fourier transforms. It remains the question of how to select an appropriate value for v. For every frequency bin ω_l, we can identify the frequency v in Eq.(14) which maximizes the magnitude of $X(\omega_l, \omega_v)$, and this frequency is denoted as ω_{δ_l}:

$$\omega_{\delta_l} = \arg\max_{\omega_v} |X(\omega_l, \omega_v)|^2. \tag{15}$$

The resulting CAPSE spectrum is consequently defined as

$$P_{CAPSE}^{xx}(\omega_l) = \frac{1}{UM} |X(\omega_l, \omega_{\delta_l})|^2. \tag{16}$$

The scheme of the CAPSE is summarized in Algorithm 1. The act of taking the maximum in the DFT, ensures that the large energy components, containing the information of interest, are retained. It should be noted that the in frequency bins where there is no sinusoidal component the maximum of the FFT across segments is taken.

4

Algorithm 1 Scheme of the CAPSE

1: **procedure CAPSE** (require signal $x(n)$)
2: Divide the signal into K segments $x_k(n)$ with overlap.
3: Window and Fourier transform each windowed segment.
4: Take a DFT along the segments for each frequency bin.
5: Find the maximum in the squared magnitude of the DFT, that peak defines the CAPSE spectrum in the frequency bin.
6: **end procedure**

2.3. Relationship between CAPSE and Fourier Transform of full sequence

In the previous section, the CAPSE has been derived starting from the Welch's method, but it can also be related to the Fourier transform. Considering the signal $x(n)$ in Eq.(1) of length N, such that $x_i(m) = x((i-1)L+m)$ for $m = 0, 1, \ldots, L-1$ and $i = 0, 1, \ldots, K$. It is the case of $D = L$ in Eq.(2), which means that there is no overlap and no window (or a rectangular window is being used). Further, it is assumed that there is an integer number of segments, K, in the data length, so $N = LK$.

The Fourier transform of the whole signal is:

$$X(\omega) = \sum_{n=0}^{L-1} x(n)e^{-j\omega n}, \tag{17}$$

which can be separated into contiguous blocks of L samples:

$$X(\omega) = \sum_{k=0}^{K-1} \sum_{n=kL}^{(k+1)L-1} x(n)e^{-j\omega n}. \tag{18}$$

Then we have

$$X(\omega) = \sum_{k=0}^{K-1} X_k(\omega)e^{-j\omega kL}. \tag{19}$$

Comparing Eq.(14) and (21) we can see that the Fourier transform of a large data set can be regarded as the CAPSE estimator with

$$X(\omega, \omega_r) = \sum_{k=0}^{K-1} X_{k+1}(\omega)e^{-j\omega_r kL}, \tag{20}$$

$$\omega_r = \frac{2\pi r}{KL}, r = 0, 1, \ldots, K-1,$$
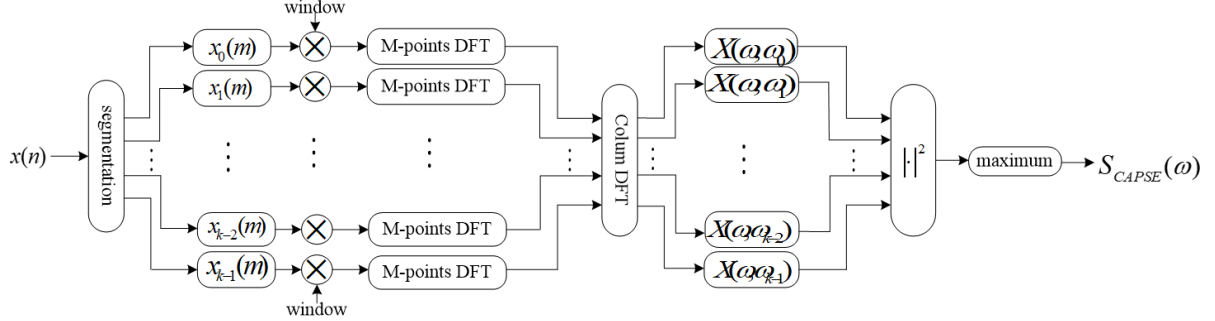
and the CAPSE is obtained as:

$$S_{CAPSE}(\omega) = max\{|X(\omega, \omega_0)|^2, |X(\omega, \omega_1)|^2, \\ \ldots, |X(\omega, \omega_{K-1})|^2\}. \tag{21}$$
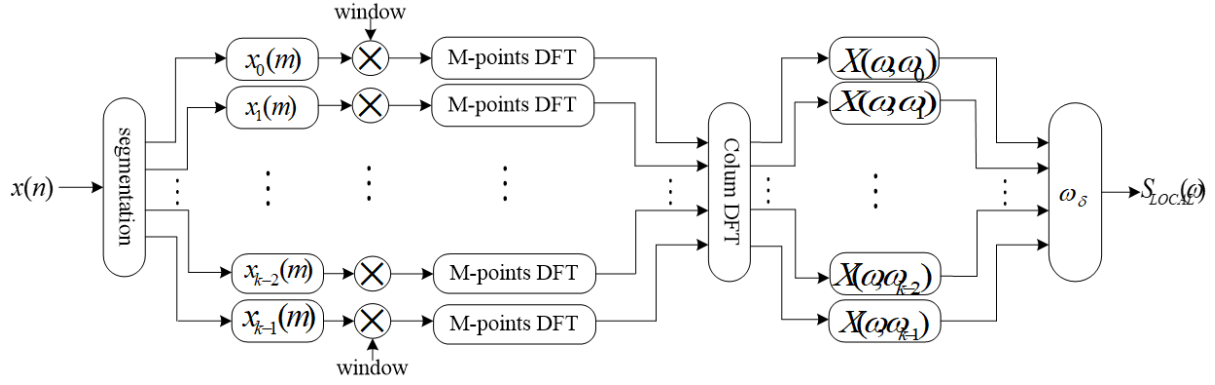
Using the CAPSE method we can maintain the accuracy associated with the FFT of the full length data, when estimating the frequency of the sine wave, although the ability for resolving two closely space sinusoids is sacrificed. The latter resolution being dictated by the choice of segment length L. For comparison, schemes of power spectral estimation of CAPSE, LOCAL, Welch and Periodogram are given in Fig.2.
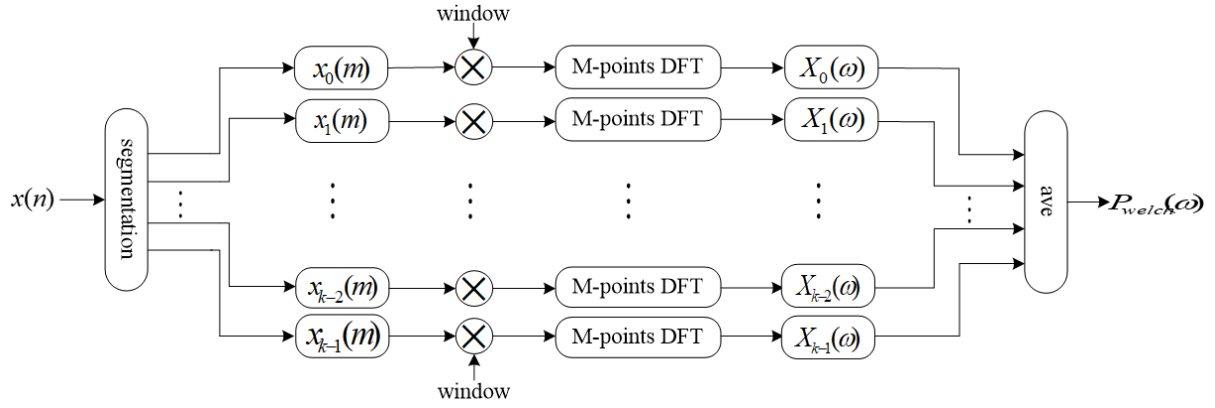
3. Results

The proposed method is verified by simulation and experiment. The simulation has been performed 100 times statistical running. The experimental data was collected in underwater acoustic channel from the South China Sea at a depth of 80 m.

window

$x(n)$ → segmentation → $x_0(m)$ ⊗ M-points DFT
$x_1(m)$ ⊗ M-points DFT
$x_{k-2}(m)$ ⊗ M-points DFT
$x_{k-1}(m)$ ⊗ M-points DFT

window

Colum DFT

$X(\omega,\omega_0)$
$X(\omega,\omega_1)$
$X(\omega,\omega_{k-2})$
$X(\omega,\omega_{k-1})$

$|\cdot|^2$ → maximum → $S_{CAPSE}(\omega)$

(a) CAPSE

window

$x(n)$ → segmentation → $x_0(m)$ ⊗ M-points DFT
$x_1(m)$ ⊗ M-points DFT
$x_{k-2}(m)$ ⊗ M-points DFT
$x_{k-1}(m)$ ⊗ M-points DFT

window

Colum DFT

$X(\omega,\omega_0)$
$X(\omega,\omega_1)$
$X(\omega,\omega_{k-2})$
$X(\omega,\omega_{k-1})$

ω_δ → $S_{LOCAL}(\omega)$

(b) LOCAL

window

$x(n)$ → segmentation → $x_0(m)$ ⊗ M-points DFT → $X_0(\omega)$
$x_1(m)$ ⊗ M-points DFT → $X_1(\omega)$
$x_{k-2}(m)$ ⊗ M-points DFT → $X_{k-2}(\omega)$
$x_{k-1}(m)$ ⊗ M-points DFT → $X_{k-1}(\omega)$

window

ave → $P_{welch}(\omega)$

(c) Welch

$x(n)$ → ⊗ → M-points DFT → $|\cdot|^2$ → $P_{Perl}(\omega)$

(d) PERIODOGRAM

Figure 2: Schemes of power spectral estimation. (a) CAPSE method, (b) LOCAL method, (c) Welch's method, (d) Periodogram method.

3.1. Simulation

We consider the mean and variance of $P^{xx}_{Welch}(\omega)$ and $P^{xx}_{CAPSE}(\omega)$ based on a rectangular window, which means $U = 1$. Assume $K = 1$ is the number of non-overlapping segments, $S_i(w) = MS_0$, and $Z_i(w) = Z_i$ is a Gaussian process because it is a linear transform of Gaussian process x_i, where $E(Z_i) = 0$ and $Var(Z_i) = MS^2_x$.

Welch [9] has shown that the mean and variance of $P^{xx}_{Welch}(\omega)$ are as follows:

$$E(P^{xx}_{Welch}(\omega)) = MS^2_0 + S^2_x. \tag{22}$$

$$Var(P^{xx}_{Welch}(\omega)) = \frac{1}{K}S^2_x. \tag{23}$$
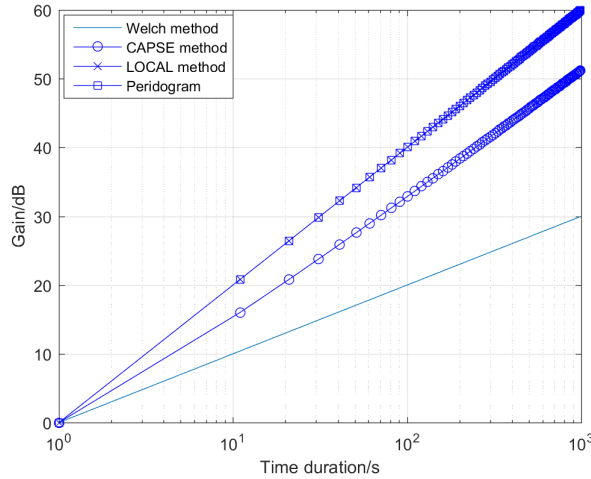
Figure 3: Comparison of gain of Welch method, the CAPSE method, the LOCAL method and Periodogram.

If $M = N$ and $K = 1$, Eq.(22) and (23) are reduced to the case of the Periodogram of full sequence of length N. It is assumed that $P^{xx}_{CAPSE}(\omega, \omega')$ is a 2-degree chi-square distribution.

To assist with the analysis, we introduce a simplified version of the algorithm. It is assumed that the frequency bin where the sinusidal wave exists is known priori and that phase compensation computed for that frequencies is applied to all other frequencies. This method refers to the LOCAL method. In this case the mean and variance of the resulting CAPSE estimate can be readily derived (see Appendix A) and are:

$$E(P^{xx}_{LOCAL}(\omega)) = MS^2_0 + \frac{1}{K}S^2_x, \tag{24}$$

$$Var(P^{xx}_{LOCAL}(\omega)) = \frac{1}{K^2}S^2_x. \tag{25}$$

Comparing Eq.(22) and (24), it can be seen that the mean of the spectrum due to the noise is reduced by $\frac{1}{K}$ in the LOCAL scheme relative to that in the Welch's method, which means the SNR is improved by K. Further, considering Eq.(23) and (25), it is evident that the variance of the CAPSE method is also reduced by a factor $\frac{1}{K}$ compared to the Welch's method.

In the implementation detailed in Section II the frequency is unknown and a different phase correction is applied to the corresponding frequency bin. This means that in frequency bins where there is no signal, the algorithm still applies a phase compensation by identifying the phase difference which yields the largest output value. This degrades the performance of the algorithm and complicates the statistical analysis, detailed in Appendix B.

It is shown that when the number of DFT is larger than the segments number K, the number of independent component used for comparison is K. The mean and variance of the CAPSE can be derived using the conclusion in Appendix B.

$$E(P^{xx}_{CAPSE}(\omega)) = MS^2_0 + \frac{1}{K}(\sum_{k=1}^{K}\frac{1}{k})S^2_x, \tag{26}$$

7

$$Var(P^{xx}_{CAPSE}(\omega)) = \frac{1}{K^2}(2K + 2K\sum_{k=1}^{K-1}(-1)^k \begin{pmatrix} K-1 \\ k \end{pmatrix}$$
$$\times \frac{1}{(k+1)^3} - (\sum_{k=1}^{K}\frac{1}{k})^2)S_x^2. \tag{27}$$

Obviously, the mean of the CAPSE method is bigger than that of the LOCAL method. If we define the gain of the processor as the difference between output and input SNR, then the gain in the Welch's method is M, in the CAPSE method is $\frac{MK}{\sum_{k=1}^{K}\frac{1}{k}}$, and in the Periodogram method is MK.

A comparison of the gain of four methods are illustrated in Fig. 3 of gain versus time duration. Here the signal segment length is 1 s. When the time duration is 1 s, the four methods are equivalent. When time duration increases, both the Periodogram and the LOCAL method have the highest gain, or lowest mean of noise spectrum level for the same noise input. When the number of segments is 100, the mean of the CAPSE method is 13 dB lower than the Welch's method, while the mean of the LOCAL method is 20 dB lower than the Welch's method.

To validate the performance of the proposed, normalized spectral level (in terms of sinusoidal signal) of the CAPSE method, the LOCAL method, the Welch's method and the Periodogram method are shown in Fig. 4 from (a) to (d) respectively. The power spectrum estimation of a signal with two sinusoids, e.g., 300.12 Hz and 312 Hz, and one Gaussian noise is given under the Welch's method and the CAPSE methods. The length of the signal is 100 s, and the signal is divided into 100 segments without overlap.

Assume that the segment interval is 1 s and the frequency resolution is 1 Hz. For the 312 Hz component, its amplitude could be estimated correctly for all the methods mentioned above, since the phase differences between the DFT the first segment and others are zero and could be added coherently no matter with compensation or not. However for the fractional frequency case, e.g, 300.12 Hz, the LOCAL method (when compensate zero fractional frequency) is unable to estimate its amplitude correctly, since spectrum of each segments are averaged coherently but the phase difference between adjacent segments of the component of 300.12 Hz is not compensated correctly. It should be noted that, if the segments are overlapped, e.g., $D = 0.8$ s, the phase difference between segments might not be 2π, even for the integer frequency, thus the compensation could not be done carefully. For the LOCAL method, uniform frequency sampling is assumed, so only the signal component with the same factional frequency can be estimated correctly.

Compare these four methods, as the spectrum are normalized with respect to the sinusoidal frequency component, the SNR determines the detection performance. The LOCAL method has the same noise level as periodogram method, about -30 dB. The CAPSE has the level of about -25 dB, which is lower than that of the Welch's method (-14 dB). Because the frequency bin of the CAPSE method and the LOCAL method are 1 Hz, whereas the frequency bin of periodogram is 0.01 Hz, the result of the other three methods (the Welch's method, the LOCAL method and the periodogram method) do not represent the 'true' spectrum density even though the level is comparable. The reason is that for the other three methods, spectral level of larger frequency bin are the signal in smaller frequency bin. In other words, these three methods are the 'down-sampling' version of periodogram, while reserving the interested signal component.

3.2. Experiment

To illustrate the capability of the proposed CAPSE method, an example of its application to real data is given as follows. Underwater acoustic channel in shallow water is often characterised as noisy and high attenuation from the acoustic source [35, 36, 37, 38]. Thus we choose the South China sea to conduct the experiment and collect data. A hydrophone has been deployed 2 m above the seabed to record the sinusoidal waves radiated from a vessel above the sea. The depth of the hydrophone was 80 m.

Fig. 4 illustrates the PSE results of ship radiated noise recorded by a hydrophone in the sea, using the CAPSE method, the LOCAL method, the Welch's method and the Periodogram method. The performance of Periodogram is best of all, with highest frequency resolution and heaviest computation load. The Welch's method in (c) is the most stable and suitable for real time processing because it segments processing. Compared with the Welch's method, both the CAPSE method and the LOCAL method can estimate spectral line with lower noise level, thus yielding higher SNR, which is critical for signal detection. The LOCAL method could only estimate one spectral line with lower noise level, while the CAPSE method could estimate all the spectral line but the noise level is fairly higher. The reason
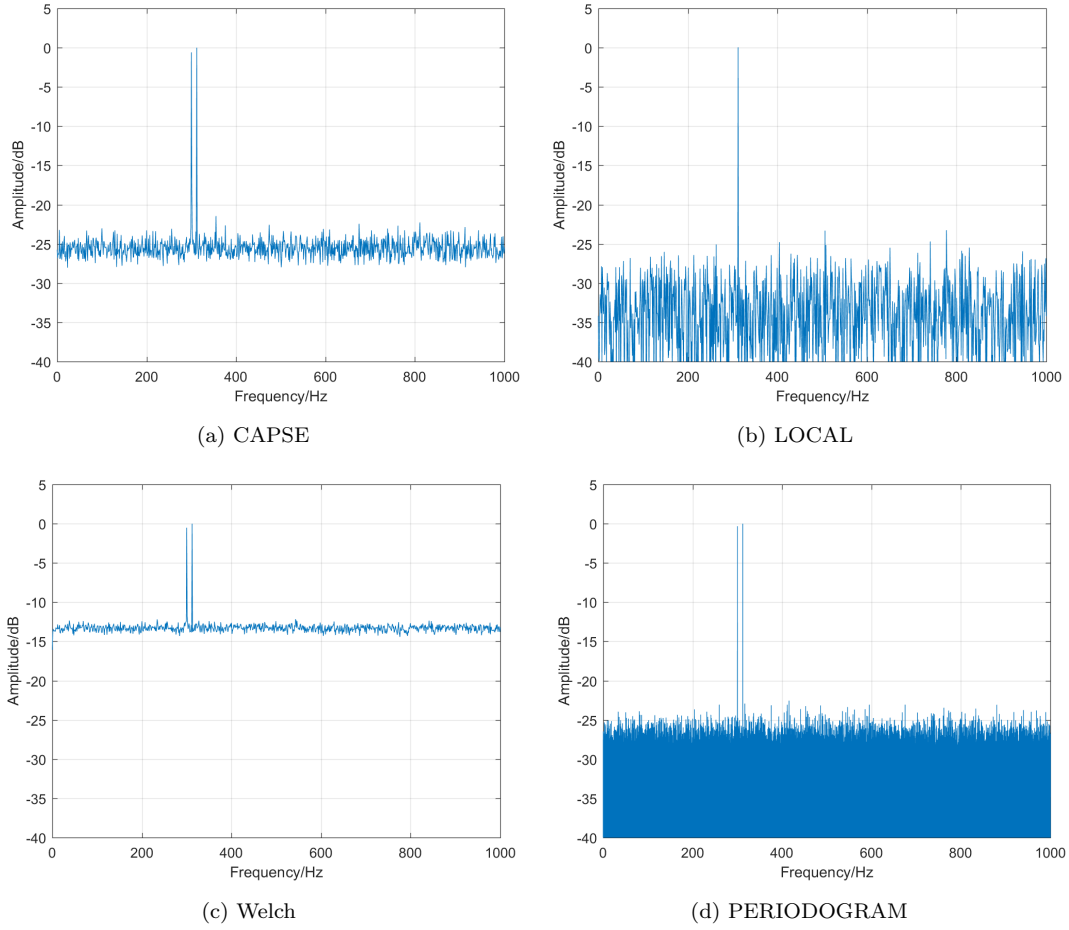
Figure 4: Simulation of spectral analysis with two sine waves buried in noise. The sinusoid frequency is 300.12 Hz and 312 Hz, time duration of the signal is 100 s, the segment(window) length is 1 s with 0.5 s overlap. (a) CAPSE method, (b) LOCAL method, (c) Welch's method, (d) Periodogram method.

of higher noise level in the CAPSE method is that, we seek the maximum for each frequency, including the frequency band with no spectral line, thus the noise is also averaged coherently. In the LOCAL method, since only one spectral line frequency is compensated, all other frequency bands are averaged incoherently, noise level is approximately to that of the periodogram.

4. Conclusion

In this paper a detection method CAPSE, has been presented, which uses coherent averaging of the spectrum of each segment. Phase difference of both the integer and fractional are compensated for inappropriate fashion. The CAPSE method can be regarded as a down-sampled version of periodogram of full sequence: with fewer frequency points (reduced resolution) but maintaining the SNR of the sinusoidal components. The proposed method can be regarded as a coherent version of the Welch's method, thus it is well-suited to real time implementation. Further, the data can be combined using different number of windows in different frequency ranges, which allows the user to match the integration time to the expected coherence of the sinusoidal components. This method can be extended to multichannel analysis, based on cross spectral estimation.
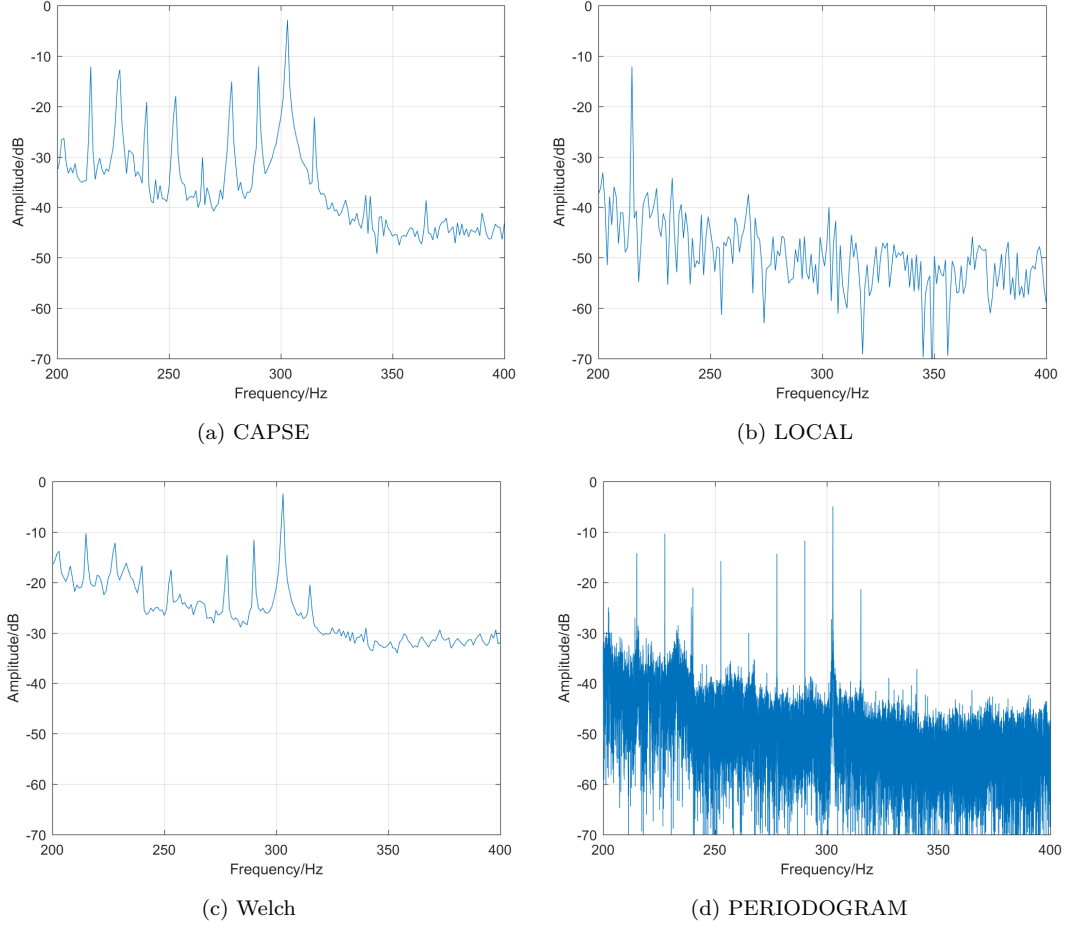
9

Figure 5: Spectral analysis of ship noise recorded by a hydrophone in the sea. (a) CAPSE method, (b) LOCAL method, (c) Welch's method, (d) Periodogram method.

Appendix A. Mean and Variance

The periodogram estimation, the Welch's estimation and the CAPSE estimation are given as follows:

$$S_W = \frac{1}{K} \sum_{k=1}^{K} I_k, \tag{Appendix A.1}$$

$$S_C = (\frac{1}{K} \sum_{k=1}^{K} Z_k)^H (\frac{1}{K} \sum_{k=1}^{K} Z_k). \tag{Appendix A.2}$$

where Z_k is the Fourier transform of Gaussian noise $z(n)$,

$$I_k = Z_k^H Z_k, \tag{Appendix A.3}$$

where H denotes complex conjugate transpose.

We assume that

$$E[I_k] = E[Z_k^H Z_k] = S_Z, \tag{Appendix A.4}$$

is the power spectral level of $z(n)$, and

$$E[Z_k^H Z_j] = 0, k \neq j. \tag{Appendix A.5}$$

The variance of I_k is

$$\begin{aligned} Var[I_k] &= E[I_k^2] - (E[I_k])^2 \\ &= 2E[Z_k^H Z_k]E[Z_k^H Z_k] - S_Z^2 \\ &= S_Z^2. \end{aligned} \tag{Appendix A.6}$$

10

The mean of the Welch's estimation is

$$E[S_W] = E[\frac{1}{K}\sum_{k=1}^{K} I_k]$$

$$= \frac{1}{K}\sum_{k=1}^{K} E[I_k] \qquad \text{(Appendix A.7)}$$

$$= S_Z.$$

The variance of the Welch's estimation is

$$Var[S_W] = E[S_W^2] - (E[S_W])^2$$

$$= E[(\frac{1}{K}\sum_{k=1}^{K} I_k)^2] - S_Z^2$$

$$= E[(\frac{1}{K}\sum_{k=1}^{K} Z_k^H Z_k)^2] - S_Z^2$$

$$= \frac{1}{K^2} E[\sum_{k=1}^{K} Z_k^H Z_k Z_k^H Z_k +$$

$$\sum_{k=1,m=1,k\neq m}^{K} Z_k^H Z_k Z_m^H Z_m] - S_Z^2 \qquad \text{(Appendix A.8)}$$

$$= \frac{1}{K^2}[\sum_{k=1}^{K} E[Z_k^H Z_k Z_k^H Z_k] +$$

$$\sum_{k=1,m=1,k\neq m}^{K} E[Z_k^H Z_k Z_m^H Z_m]] - S_Z^2$$

$$= \frac{1}{K^2}[2KS_Z^2 + K(K-1)S_Z^2] - S_Z^2$$

$$= \frac{1}{K}S_Z^2.$$

The mean of the CAPSE estimation is

$$E[S_C] = E[(\frac{1}{K}\sum_{k=1}^{K} Z_k)^H(\frac{1}{K}\sum_{k=1}^{K} Z_k)]$$

$$= \frac{1}{K^2} E[\sum_{k=1}^{K} Z_k^H Z_k + \sum_{k=1,m=1,k\neq m}^{K} Z_k^H Z_m] \qquad \text{(Appendix A.9)}$$

$$= \frac{1}{K}S_Z.$$

11

The variance of the CAPSE method is given by

$$Var[S_C] = E[S_C^2] - (E[S_C])^2$$

$$= E[(\frac{1}{K}\sum_{k=1}^{K}Z_k)^H(\frac{1}{K}\sum_{k=1}^{K}Z_k)\times$$

$$(\frac{1}{K}\sum_{k=1}^{K}Z_k)^H(\frac{1}{K}\sum_{k=1}^{K}Z_k)] - (\frac{1}{K}S_Z)^2$$

$$= \frac{1}{K^4}E[(\sum_{k=1}^{K}Z_k^H Z_k + \sum_{k=1,j=1,k\neq j}^{K}Z_k^H Z_j)\times$$

$$(\sum_{m=1}^{K}Z_m^H Z_m + \sum_{m=1,n=1,m\neq n}^{K}Z_m^H Z_n)] - \frac{1}{K^2}S_Z^2 \qquad \text{(Appendix A.10)}$$

$$= \frac{1}{K^4}E[(\sum_{k=1}^{K}Z_k^H Z_k \sum_{m=1}^{K}Z_m^H Z_m +$$

$$2\sum_{k=1}^{K}Z_k^H Z_k \sum_{m=1,n=1,m\neq n}^{K}Z_m^H Z_n +$$

$$\sum_{k=1,j=1,k\neq j}^{K}Z_k^H Z_j \sum_{m=1,n=1,m\neq n}^{K}Z_m^H Z_n)] - \frac{1}{K^2}S_Z^2.$$

Here

$$E[\sum_{k=1}^{K}Z_k^H Z_k \sum_{m=1}^{K}Z_m^H Z_m] = E[\sum_{k=1}^{K}Z_k^H Z_k Z_k^H Z_k +$$

$$\sum_{k=1,m=1,k\neq m}^{K}Z_k^H Z_k Z_m^H Z_m]$$

$$= \sum_{k=1}^{K}2E[Z_k^H Z_k]E[Z_k^H Z_k] + \qquad \text{(Appendix A.11)}$$

$$\sum_{k=1,m=1,k\neq m}^{K}E[Z_k^H Z_k]E[Z_m^H Z_m]$$

$$= 2KS_Z^2 + K(K-1)S_Z^2$$

$$= K(K+1)S_Z^2,$$

$$E[\sum_{k=1}^{K}Z_k^H Z_k \sum_{m=1,n=1,m\neq n}^{K}Z_m^H Z_n] = 0, \qquad \text{(Appendix A.12)}$$

$$E[\sum_{k=1,j=1,k\neq j}^{K}Z_k^H Z_j \sum_{m=1,n=1,m\neq n}^{K}Z_m^H Z_n]$$

$$= E[\sum_{k=1,j=1,k\neq j}^{K}Z_k^H Z_k Z_j^H Z_j] \qquad \text{(Appendix A.13)}$$

$$= K(K-1)S_Z^2.$$

The variance of CAPSE is

$$Var[S_C] = \frac{1}{K^2}S_Z^2. \qquad \text{(Appendix A.14)}$$

Specially, $K = 1$ is the case of one segment periodogram.

Appendix B.

Assume $X_1, X_2, \ldots, X_K$ are K independent 2 degrees of freedom, chi-square distribution variable, or exponential distribution. The probability density function of X_i $(i = 1, 2, \ldots, K)$ is given by

$$p(x_i) = \begin{cases} \dfrac{1}{2\sigma_X^2} e^{-\frac{x_i}{2\sigma_X^2}}, x_i \geq 0 \\ 0, x_i < 0 \end{cases}. \qquad \text{(Appendix B.1)}$$

Then the probability density function of their maximum, $X = max(X_1, X_2, \ldots, X_K)$, is

$$p(x) = Kp(x)(F(x))^{K-1}, \qquad \text{(Appendix B.2)}$$

where

$$F(x) = \int_{-\infty}^{x} p(x_i)dx_i = 1 - e^{-\frac{x}{2\sigma_X^2}} \qquad , \qquad \text{(Appendix B.3)}$$

is the cumulative distribute function of X_i.

$$p(x) = \begin{cases} \dfrac{K}{2\sigma_X^2}(1 - e^{-\frac{x}{2\sigma_X^2}})^{K-1}e^{-\frac{x}{2\sigma_X^2}}, x \geq 0 \\ 0, x < 0 \end{cases}. \qquad \text{(Appendix B.4)}$$

The expectation of X is given by

$$\begin{aligned} E[X] &= \int_{-\infty}^{+\infty} xp(x)dx \\ &= \int_{0}^{+\infty} xKp(x)(F(x))^{K-1}dx \\ &= \int_{0}^{+\infty} 1 - (F(x))^K dx \\ &= \int_{0}^{+\infty} 1 - (1 - e^{-\frac{x}{2\sigma_X^2}})^K dx. \end{aligned} \qquad \text{(Appendix B.5)}$$

Using the binomial theorem to obtain

$$\begin{aligned} E[X] &= \int_{0}^{+\infty} \sum_{k=1}^{K}(-1)^{k+1}\begin{pmatrix} K \\ k \end{pmatrix} e^{-\frac{kx}{2\sigma_X^2}} dx \\ &= 2\sigma_X^2 \sum_{k=1}^{K}(-1)^{k+1}\begin{pmatrix} K \\ k \end{pmatrix}\frac{1}{k}, \end{aligned} \qquad \text{(Appendix B.6)}$$

Lugo [39] has given and shown a more compact form of $E[X]$ as

$$E[X] = 2\sigma_X^2 \sum_{k=1}^{K}\frac{1}{k}. \qquad \text{(Appendix B.7)}$$

In some instances, it may be useful to approximate the sum of the harmonic series (Appendix B.8) using Eulers constant γ (≈ 0.577), specifically

$$E[X] \approx 2\sigma_X^2(\gamma + ln(K)) \approx 2\sigma_X^2 ln(K). \qquad \text{(Appendix B.8)}$$

When $K = 1$, from (Appendix B.5) we can obtain

$$E[X] = 2\sigma_X^2. \qquad \text{(Appendix B.9)}$$

13

From (Appendix B.5), we can obtain

$$\int_0^{+\infty} 1 - (1 - e^{-\frac{x}{2\sigma_X^2}})^{K+1} dx - \int_0^{+\infty} 1 - (1 - e^{-\frac{x}{2\sigma_X^2}})^K dx$$

$$= \int_0^{+\infty} e^{-\frac{x}{2\sigma_X^2}} (1 - e^{-\frac{x}{2\sigma_X^2}})^K dx$$

$$= -2\sigma_X^2 \int_0^{+\infty} (1 - e^{-\frac{x}{2\sigma_X^2}})^K de^{-\frac{x}{2\sigma_X^2}} \qquad \text{(Appendix B.10)}$$

$$= 2\sigma_X^2 \frac{1}{K+1} (1 - e^{-\frac{x}{2\sigma_X^2}})^K |_0^{+\infty}$$

$$= 2\sigma_X^2 \frac{1}{K+1}.$$

The variance of X is given by

$$E[X^2] = \int_{-\infty}^{+\infty} x^2 p(x) dx$$

$$= \int_0^{+\infty} x^2 K p(x) (F(x))^{K-1} dx \qquad \text{(Appendix B.11)}$$

$$= \int_0^{+\infty} x^2 \frac{K}{2\sigma_X^2} (1 - e^{-\frac{x}{2\sigma_X^2}})^{K-1} e^{-\frac{x}{2\sigma_X^2}} dx.$$

Using the binomial theorem to obtain

$$E[X^2] = \int_0^{+\infty} x^2 \frac{K}{2\sigma_X^2}$$
$$(1 + \sum_{k=1}^{K-1} (-1)^{k+1} \binom{K-1}{k} e^{-\frac{kx}{2\sigma_X^2}}) e^{-\frac{x}{2\sigma_X^2}} dx. \qquad \text{(Appendix B.12)}$$

$$\int_0^{+\infty} x^2 \frac{K}{2\sigma_X^2} e^{-\frac{x}{2\sigma_X^2}} dx = 8K\sigma_X^4. \qquad \text{(Appendix B.13)}$$

$$\int_0^{+\infty} x^2 \frac{K}{2\sigma_X^2} \sum_{k=1}^{K-1} (-1)^k \binom{K-1}{k} e^{-\frac{(k+1)x}{2\sigma_X^2}} dx$$
$$= 8K \sum_{k=1}^{K-1} (-1)^k \binom{K-1}{k} \frac{1}{(k+1)^3} \sigma_X^4. \qquad \text{(Appendix B.14)}$$

Finally,

$$E[X^2] = 8K\sigma_X^4 + 8K \sum_{k=1}^{K-1} (-1)^k \binom{K-1}{k} \frac{1}{(k+1)^3} \sigma_X^4. \qquad \text{(Appendix B.15)}$$

Acknowledgments

References

[1] H. L. Van Trees, Optimum array processing: Part IV of detection, estimation, and modulation theory, John Wiley & Sons, 2004. doi:10.1002/0471221104.

[2] R. N. McDonough, A. D. Whalen, Detection of signals in noise, Academic Press, 1995.

[3] A. V. Oppenheim, Digital signal processing, Tech. rep., MASSACHUSETTS INST OF TECH CAMBRIDGE (1981).

[4] B. D. Van Veen, K. M. Buckley, Beamforming: A versatile approach to spatial filtering, IEEE assp magazine 5 (2) (1988) 4–24. doi:10.1109/53.665.

[5] M. Brandstein, D. Ward, Microphone arrays: signal processing techniques and applications, Springer Science & Business Media, 2013. doi:10.1007/978-3-662-04619-7.

[6] A. K. Ziarani, A. Konrad, A method of extraction of nonstationary sinusoids, Signal Processing 84 (8) (2004) 1323–1346. doi:10.1016/j.sigpro.2004.05.008.

[7] H.-T. Li, P. M. Djurić, A novel approach to detection of closely spaced sinusoids, Signal processing 51 (2) (1996) 93–104. doi:10.1109/ISCAS.1995.523886.

[8] H. Akçay, Spectral estimation in frequency-domain by subspace techniques, Signal Processing 101 (2014) 204–217. doi:10.1016/j.sigpro.2014.02.015.

[9] P. Welch, The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, IEEE Transactions on audio and electroacoustics 15 (2) (1967) 70–73. doi:10.1109/TAU.1967.1161901.

[10] A. V. Oppenheim, Discrete-time signal processing, Pearson Education India, 1999.

[11] S. V. Vaseghi, Advanced digital signal processing and noise reduction, John Wiley & Sons, 2008. doi:10.1002/9780470740156.

[12] V. Katkovnik, L. Stanković, Periodogram with varying and data-driven window length, Signal Processing 67 (3) (1998) 345–358. doi:10.1016/S0165-1684(98)00049-8.

[13] P. Stoica, T. Sundin, Optimally smoothed periodogram, Signal Processing 78 (3) (1999) 253–264. doi:10.1016/S0165-1684(99)00066-3.

[14] A. Schuster, On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena, Terrestrial Magnetism 3 (1) (1898) 13–41. doi:10.1029/TM003i001p00013.

[15] J. S. Bendat, A. G. Piersol, Random data: analysis and measurement procedures, Vol. 729, John Wiley & Sons, 2011.

[16] G. L. Bretthorst, Bayesian spectrum analysis and parameter estimation, Vol. 48, Springer Science & Business Media, 2013. doi:10.1007/978-1-4684-9399-3.

[17] J. Locke, P. R. White, The performance of methods based on the fractional Fourier transform for detecting marine mammal vocalizations, The Journal of the Acoustical Society of America 130 (4) (2011) 1974–1984. doi:10.1121/1.3631664.

[18] S. Key, Fundamentals of statistical signal processing, volume ii: Detection theory (1993).

[19] M. Babtlett, Smoothing periodograms from time-series with continuous spectra, Nature 161 (4096) (1948) 686. doi:10.1038/161686a0.

[20] M. S. Bartlett, Periodogram analysis and continuous spectra, Biometrika 37 (1/2) (1950) 1–16. doi:10.2307/2332141.

[21] G. C. Carter, A. H. Nuttall, C. Yuen, On the weighted overlapped segment averaging method for power spectral estimation, Proceedings of the IEEE 68 (10) (1980) 1352–1354. doi:10.1109/PROC.1980.11866.

[22] P. E. Johnson, D. G. Long, The probability density of spectral estimates based on modified periodogram averages, IEEE transactions on signal processing 47 (5) (1999) 1255–1261. doi:10.1109/78.757213.

[23] K. Barbe, R. Pintelon, J. Schoukens, Welch method revisited: nonparametric power spectrum estimation via circular overlap, IEEE Transactions on signal processing 58 (2) (2010) 553–565. doi:10.1109/TSP.2009.2031724.

[24] R. U. Kale, P. M. Ingale, R. T. Murade, S. S. Sayyad, Comparison of quality power spectrum estimation (Bartlett, Welch, Blackman & Tukey) methods, Int. Jounal Sci. Mod. Eng 1 (2013) 28–31.

[25] D. J. Thomson, Spectrum estimation and harmonic analysis, Proceedings of the IEEE 70 (9) (1982) 1055–1096. doi:10.1109/PROC.1982.12433.

[26] M. B. Priestley, Spectral analysis and time series, Academic press, 1981.

[27] Q. Wang, C. R. Wan, A Novel CFAR Tonal Detector Using Phase Compensation, IEEE Journal of Oceanic Engineering 30 (4) (2005) 900–911. doi:10.1109/JOE.2004.836583.

[28] S. Chen, Improving coherent tonal detection with phase interpolation and compensation, in: IEEE Oceans-San Diego, 2013, 2013, pp. 1–5. doi:10.23919/OCEANS.2013.6741332.

[29] A. V. Oppenheim, A. S. Willsky, S. Nawab, Signals and systems, Massachusetts Institute Technology, Boston University.

[30] R. B. Blackman, J. W. Tukey, The measurement of power spectra, Dover Publications Inc., 1958.

[31] F. J. Harris, On the use of windows for harmonic analysis with the discrete Fourier transform, Proceedings of the IEEE 66 (1) (1978) 51–83. doi:10.1109/PROC.1978.10837.

[32] R. O. Nielsen, Sonar signal processing, Artech House, Inc., 1991.

[33] P. K. Rahi, R. Mehra, Änalysis of power spectrum estimation using welch method for various window techniques, International Journal of Emerging Technologies and Engineering 2 (6) (2014) 106–109.

[34] L. Fan, G. Qi, Frequency estimator of sinusoid based on interpolation of three DFT spectral lines, Signal Processing 144 (2018) 52–60. doi:10.1016/j.sigpro.2017.09.028.

[35] M. Stojanovic, On the relationship between capacity and distance in an underwater acoustic communication channel, ACM SIGMOBILE Mobile Computing and Communications Review 11 (4) (2007) 34–43. doi:10.1145/1161039.1161049.

[36] M. Stojanovic, Underwater acoustic communications: Design considerations on the physical layer, in: IEEE Fifth Annual Conference on Wireless on Demand Network Systems and Services., 2008, pp. 1–10. doi:10.1109/WONS.2008.4459349.

[37] J. Li, Y. V. Zakharov, Efficient use of space-time clustering for underwater acoustic communications, IEEE Journal of Oceanic Engineering 43 (1) (2018) 173–183. doi:10.1109/JOE.2017.2688558.

[38] J. Li, Y. V. Zakharov, B. Henson, Multibranch autocorrelation method for doppler estimation in underwater acoustic channels, IEEE Journal of Oceanic Engineering 43 (4) (2017) 1099 – 1113. doi:10.1109/JOE.2017.2761478.

[39] M. Lugo, The expectation of the maximum of exponentials, Stat 134.

Conflict of interests

All the authors have no conflict of interest to declare with respect to the submitted paper "Coherently Averaged Power Spectral Estimate for Signal Detection".

Authors :
Hualin Lan(E-mail: lanhualin@hrbeu.edu.cn)
Paul White(E-mail: prw@isvr.soton.ac.uk)
Na Li(E-mail: 18844188929@163.com)
Jianghui Li(E-mail: J.Li@soton.ac.uk)
Dajun Sun(E-mail: sundajunmail@163.com)