A Fractional Lower-order Bi-spectrum Estimation Method Based on Autoregressive Model

Baohai Yang

College of Physics and Electronic Engineering, Guangxi Normal University for Nationalities, Chongzuo 532200, China

ABSTRACT

The traditional fractional lower-order (FLO) spectrum estimation method cannot observe a sufficient volume of data, leading to a large variance of estimation results. To solve the problem, this paper puts forward a FLO bi-spectrum estimation method based on autoregressive (AR) model, and gives new definitions for FLO three order cumulant. The author discussed the determination of AR model parameters, and introduced how to implement the bi-spectrum estimation method based on AR model. Then, a series of tests were performed to verify the correctness of our method. The results show that our method outperformed the traditional approaches in suppressing FLO noise and identifying relevant information of signals.

Keywords: autoregressive (AR) model, bi-spectrum, fractional lower-order (FLO) statistics, three order cumulant, signal processing

1. INTRODUCTION

In the field of signal processing, the noise is mostly described by Gaussian distribution model, which has an excellent effect on two-order statistics of normal distribution [1]. When it comes to the frequency-domain analysis of noise, the existing theoretical methods include spectral feature analysis, time-frequency analysis, spectrum estimation, colored noise whitening, spatial spectrum estimation, frequency estimation, harmonic estimation and spectral line restoration. Most of these methods are based on two-order statistics like bi-spectrum and tri-spectrum [2].

Gaussian white noise and colored noise models have always occupied the leading position in signal processing, and the criterion of white noise and colored noise based on correlation function and power spectral density has been regarded as a classical rule. In practical applications, however, many noises do not conform to the normal distribution. Typical examples include low-frequency atmospheric noise and underwater noise. A viable option to describe such noises lies in setting up an α stable distribution, whose statistical features can be characterized by the relevant parameters of the feature function [3].

The best way to solve the non-Gaussian α stable distribution is the fractional lower-order (FLO) statistics. If α is greater than 2, the harmonic frequencies can be estimated accurately; otherwise, the second-order matrix does not exist, making it impossible to effectively analyze the noise [2]. To solve the problem, the FLO bi-spectrum has been conceptualized. Traditionally, the lower-order bi-spectrum is estimated by bi-spectrum and nonparametric methods. Nonetheless, these methods cannot observe a sufficient volume of data, leading to a large variance of estimation results.

In view of the above, this paper puts forward an FLO bi-spectrum estimation method based on autoregressive (AR) model [5], and compares the method with traditional approaches. The comparison shows that our method outperformed the traditional ones in spectral flatness and suppression of FLO noise.

2. α STABLE DISTRIBUTION AND FLO STATISTICS

This section briefly introduces the α stable distribution, a generalized conceptual Gaussian distribution with a wide application scope, and gives the definition to the relevant FLO statistics, which is the best way to filter the noise of non-Gaussian α stable distribution.

2.1 α stable distribution

A random variable X satisfies the α stable distribution if the parameters 0≤α≤2, γ≥0 and -1≤β≤1 have the following relationship with real numbers α:

\[ \phi(t) = \exp \left\{ j \alpha t - \gamma \right\} \left\{ 1 + j \beta \; \text{sgn}(t) \alpha(t, \alpha) \right\} \]  (1)

where

\[ \alpha(t, \alpha) = \begin{cases} \frac{\tan(\pi \alpha / 2)}{\pi} & \alpha \neq 1 \\ \frac{1}{2} & \alpha = 1 \end{cases} \]

\[ \text{sgn}(t) = \begin{cases} 1, & t > 0 \\ 0, & t = 0 \\ -1, & t < 0 \end{cases} \]

Note that α is a feature index in the interval (0, 2]. The index determines the shape of the distribution. The value of the index is negatively correlated with the trailing thickness and the pulse feature [6].

2.2 FLO statistics

(1) FLO moments. The FLO moments refer to the various moments existing at 0<α<2 (i.e. the absence of secondary moment). For the random variable \( \text{SaS} \), the FLO moments can be described by its dispersion coefficient and feature index. Meanwhile, the covariation of the two random variables \( \eta \) and \( \zeta \) can be expressed as the function of FLO moments:
The existing covariation applies to the range of $1<\alpha\leq 2$, but not defined for the range $\alpha\leq 1$ in the SaS distribution. Reference [8] proposes a more general FLO statistic that applies to the entire value range of $\alpha$. Under $0<\alpha\leq 2$, the FLO covariance of random variables $X$ and $Y$ of SaS distribution can be defined as:

$$[X, Y]_\alpha = \frac{E_X \left( \frac{Y}{|Y|^\alpha} \right)}{E_Y \left( \frac{X}{|X|^\alpha} \right)}$$

(10)

Let $\gamma_\alpha$ be the dispersion coefficient of $Y$. Under $0<\alpha\leq 2$, the covariation of random variables $X$ and $Y$ of SaS distribution can be established as:

$$[Y, Y]_\alpha = \frac{E_Y \left( \frac{Y}{|Y|^\alpha} \right)}{E_Y \left( \frac{Y}{|Y|^\alpha} \right)} = \gamma_\alpha$$

(11)

$$\lambda_{\alpha, \beta} = \frac{E_X \left( \frac{Y}{|Y|^\alpha} \right)}{E_Y \left( \frac{X}{|X|^\alpha} \right)}$$

(12)

$$[X, Y]_\alpha = \frac{E_X \left( \frac{Y}{|Y|^\alpha} \right)}{E_Y \left( \frac{X}{|X|^\alpha} \right)} = \gamma_\alpha$$

(13)

(4) FLO covariance. The existing covariation applies to the range of $1<\alpha\leq 2$, but not defined for the range $\alpha\leq 1$ in the SaS distribution. Reference [8] proposes a more general FLO statistic that applies to the entire value range of $\alpha$. Under $0<\alpha\leq 2$, the FLO covariance of random variables $X$ and $Y$ of SaS distribution can be defined as:

$$FLO(X, Y) = E(X^a Y^{a-p})$$

(14)

$$0 \leq A < \frac{\alpha}{2}, \quad 0 \leq B < \frac{\alpha}{2}$$

3. FLO BI-SPECTRUM

3.1 Definition of FLO bi-spectrum

The statistical moment of a signal tells a lot about the signal features. The spectra in statistical moments extend from low order to infinite order [9]. Only secondary moments were explored in traditional signal processing. Nevertheless, the processing methods may suffer from poor performance and high error, using variance or second-order statistics only. Recent years has seen the emergence of signal processing techniques for higher-order statistics, especially third- or fourth-order statistics. The emerging techniques not only utilize the second- or higher-order statistics, but also many fractional order statistics under the second order, i.e. the FLO statistics [10].

Both theoretical and empirical analyses prove that the FLO statistics are suitable for processing impulsive feature signals and noises. However, the FLO statistics also have two prominent defects [11]. First, there is no universal framework against algebraic smearing. Second, there is a theoretical connection between the priori knowledge of the order $p$ and the random variable $\alpha$, because the order $p$ of moments is generally limited to $(0, \alpha)$. If $p=\alpha$, the FLO statistics cannot work normally. As a result, the traditional $\alpha$ stable distribution does not contain bi-spectrum or tri-spectrum.

Based on the traditional bi-spectrum, this paper proposes the FLO bi-spectrum, and develops the nonparametric bi-spectrum estimation method for the environment of FLO noise, and verifies its performance through comparative experiments. The FLO bi-spectrum can be defined as:

$$[X, Y, Z]_\alpha = \int_S X Y Z m(ds)$$

(9)

where $S$ is an unit circle; $m(\cdot)$ is the spectral measure in SaS distribution. Then, the covariation coefficient can be defined as:

$$[X, Y]_\alpha = \int_S X Y m(ds)$$

(10)
\[ B_s(\omega_1, \omega_2) = \sum_{\tau_1}^{\infty} \sum_{\tau_2}^{\infty} C_{3s}(\tau_1, \tau_2) \exp[-j(\omega_1 \tau_1 + \omega_2 \tau_2)] \] (15)

Besides, the FLO three order cumulants were redefined as:
\[ C_{3s}(m, n) = \gamma_{3e} \sum_{i=0}^{\infty} [x(i)]^{(0)} [x(i + m)]^{(0)} [x(i + n)]^{(c)} \] (16)
where \( x(i) \) is the signal sequence; \( \gamma_{3e} \) is an adjustment coefficient; \( 0 < A + B + C \leq \alpha \) \((0 < \alpha < 2)\).

### 3.2 FLO bi-spectrum estimation with nonparametric direct estimation

The FLO bi-spectrum estimation with nonparametric direct estimation refers to the nonlinear transformation of the data \( x(n) \):
\[ g(x(t), A) = (x(t))^{-A}, 0 \leq A \leq \alpha / 3 \] (17)

After the transformation, the three order statistics of the stochastic process \( x(n) \) remain the same. Next, the transformed data can be divided into \( K \) segments, each of which has \( M \) samples, i.e. \( N = KM \). Then, the sample mean of each segment can be determined [12], allowing the overlap between two adjacent two segments. Then, the discrete Fourier transform (DFT) coefficients can be calculated as:
\[ X^{(k)}(\lambda) = \frac{1}{M} \sum_{m=1}^{M} X^{(k)}(n) e^{-j2\pi mk/M} \] (18)
where \( k = 0, 1, \ldots, M/2; k = 1, \ldots, K \). According to the DFT coefficients, the FLO bi-spectrum estimation of each segment can be obtained [13]:
\[ b_s(\lambda_k, \lambda_j) = \frac{1}{N_0} \sum_{m=1}^{M} \sum_{i=1}^{M} \sum_{j=1}^{M} X^{(k)}(\lambda_k + i)^{j} X^{(j)}(\lambda_j + i)^{k} \] (19)
where \( k = 1, \ldots, K; 0 \leq \lambda_k \leq \lambda_1; \lambda_1 + \lambda_2 \leq f_s/2; \Delta_0 = f_s/N_0 \) \((N_0 \text{ and } L \text{ satisfy } M = \frac{2L + 1}{f_s} N_0)\). The bi-spectrum estimation of the given data \( x(0), x(1), \ldots, x(N-1) \) can be determined by the mean value of the \( K \) segment bi-spectrum estimation:
\[ \hat{B}_s(\omega_1, \omega_2) = \frac{1}{K} \sum_{k=1}^{K} b_s(\omega_1, \omega_2) \] (20)
where \( \omega_1 = 2\pi \Delta_0 / N_0; \omega_2 = 2\pi \Delta_0 / \lambda_2 \).

### 3.3 FLO bi-spectrum estimation with nonparametric indirect estimation

Assuming that the observed data \( x(i)(i = 1, 2, \ldots, N) \) is a real random sequence, the three order cumulants of the observed data \( x(i) \) should be estimated, and then subjected to the DFT, completing the bi-spectrum estimation. The algorithm procedure can be described as:
1. Divide the observed data \( x(i) \) of the length \( N \) into \( K \) segments, each of which has \( M \) points, \( N = KM \), or divide the data in such a manner that the adjacent segments have a half overlap, \( 2N = KM \);
2. Remove the mean value of each segment, and make the mean value of the data to be analyzed zero;
3. Let \( x(i)(i = 1, 2, \ldots, M; j = 1, 2, \ldots, K) \) be the \( j \) segment. Then, estimate the lower-order three order cumulant of each segment:
\[ \hat{C}_{3s}(m, n) = \frac{1}{M} \sum_{i=0}^{M} [x^{(i)}(i + m)]^{(0)} [x^{(i)}(i + n)]^{(c)} \] (21)

where \( 0 < A + B + C \leq \alpha; k_j = \max\{0, -m, -n\}; k_j = \min\{M, M-n, M-m\} \).
4. Compute the statistical mean of \( \hat{C}_{3s}(m, n) \), and determine cumulative estimates of \( K \) segments:
\[ \hat{C}_s(m, n) = \frac{1}{K} \sum_{j=1}^{K} \hat{C}_{3s}(m, n) \] (22)
5. Estimate the three order cumulants through Fourier transform, and thus obtain the FLO bi-spectrum estimation:
\[ \hat{B}_s(\omega_1, \omega_2) = \sum_{m=1}^{L} \sum_{n=1}^{L} \hat{C}_s(m, n) \exp[j{(\omega_1 m + \omega_2 n)}] \] (23)
where \( L < M-1 \).

### 4. PARAMETER DETERMINATION FOR AR MODEL

If the excitation of a physical network by white noise \( \omega(n) \) is viewed as a random signal of the power spectral density under normal conditions [14], then the \( p \)-order AR model can be established as:
\[ x(n) = \sum_{k=1}^{p} a_k x(n - k) + \omega(n) \] (25)

If a \( z \) transformation is performed, the transfer function of the AR model can be expressed as:
\[ H(z) = \frac{B(z)}{A(z)} = \frac{1}{1 + \sum_{k=1}^{p} akz^{-k}} \] (26)
where the \( H(z) \) belongs to the all-pole type, i.e. it only has poles [15]. The power spectral density of the AR model can be described as:
\[ P_{xx}(\omega) = \frac{\sigma_w^2}{1 + \sum_{k=1}^{p} ak e^{-j\omega k}}^2 \] (27)
where \( \sigma_w^2 \) is the power spectral density of white noise.

The parameters of the AR model can be determined in three ways, namely, correlation function, reflection coefficient and...
normal equation. The last approach is adopted in our research.

Assuming that \( \{x(i)\} \) is a stochastic process satisfying the difference equations of the AR model, we have:

\[
\sum_{j=0}^{p} a_j x(i - j) = \sum_{j=0}^{q} b_j e(i - j')
\]  

(28)

where \( a_i \) and \( b_j \) are two key parameters in the AR model. Then, the following three conditions must be satisfied:

1. \( e(i) \) is zero-mean and a stationary, identically distributed, independent sequence; there must exist at least one nonzero pole order \( K(K>2) \) order cumulant.

2. The model is causal with a minimum phase; there must be no zero pole cancellation; the index must remain stable [16]; the transfer function can be expressed as:

\[
H(z) = \frac{b(z)}{a(z)} = \frac{1 + b(1)z^{-1} + \cdots + b(q)z^{-q}}{1 + a(1)z^{-1} + \cdots + a(p)z^{-p}} = \sum_{i=0}^{\infty} b(i)z^{-i}
\]  

(29)

3. The additive noise \( n(i) \) must be present in \( x(i) \):

\[
y(i) = x(i) + n(i)
\]

(30)

where \( n(i) \) is the FLO noise, which is independent of \( x(i) \), in the stable distribution of symmetric distributed noise.

Under these conditions, the relation between the output and the impulse response of this system can be expressed as:

\[
C_{xy}(m, n) = \gamma_{xy} \sum_{i=0}^{\infty} h^{k-2}(i) h(i + m) h(i + n)
\]

(31)

From the definition of the impulse response, we have:

\[
\sum_{i=0}^{p} a_i h(i - 1) = \sum_{j=0}^{q} b_j \delta(i - j') = b_i
\]

(32)

According to (28)–(32), we have:

\[
\sum_{i=0}^{p} a_i C_{xy}(m - 1, n) = \gamma_{xy} \sum_{i=0}^{\infty} h^{k-2}(i) h(i + n) h(i + m)
\]

(33)

If \( m>q \), \( i \geq 0 \) and \( b(i+m)=0 \), then the normal equation of the three order cumulant can be expressed as:

\[
\sum_{i=0}^{p} a_i C_{xy}(m - 1, n) = 0
\]

(34)

The parameters of the AR model can be determined by solving the equation.

5. IMPLEMENTATION OF AR-BASED BI-SPECTRUM ESTIMATION

The AR-based bi-spectrum estimation can be implemented in four steps:

Step 1: Divide \( x(i) \) into several segments with \( N \) being the segment length, \( N=KM \). Then, the three order cumulant of each data can be expressed as:

\[
C_{xy}^{(i)}(m, n) = \frac{1}{M} \sum_{i=1}^{M} [x_{(i)}^{(i)}(i) x_{(i)}^{(i)}(i + m) x_{(i)}^{(i)}(i + n)]^{(i)}
\]

(35)

where \( kl=max\{0\,\cdot\,m\,-\,n\} \); \( k2=min\{M\,M\,-\,n\,\cdot\,M\} \).

Step 2: calculate the mean value of the \( K \) segment, and estimate the three order cumulants [17] as:

\[
C_{xy}(m, n) = \frac{1}{K} \sum_{i=1}^{K} C_{xy}^{(i)}(m, n)
\]

(36)

Step 3: Determine the parameters of the AR Model, \( a\{l, 2, ...\,p\} \);

Step 4: Estimate the bi-spectrum using the parameters determined in Step 3:

\[
\hat{h}(\omega_1, \omega_2) = \sum_{i=1}^{K} \hat{h}(\omega_1, \omega_2) \hat{h}^*(\omega_1 + \omega_2)
\]

\[
\hat{h}(\omega) = [1 + \sum_{i=0}^{\infty} \exp(-j\omega n)]^{-1}, |\omega| \leq \pi
\]

(37)

6. VERIFICATION

The proposed method was verified through comparative tests on the following sequence:

\[
x(n) = s(n) + v(n)
\]

(38)

where \( n \) is an integer in the interval \( [0, N-1] \), with \( N \) being the sequence length; \( s(n) = \cos(2\pi f_1 n) + \cos(2\pi f_2 n) \), with of \( f_1=0.2 \) and \( f_2=0.4 \); \( v(n) \) is a symmetrically distributed noise.
During the tests, the mixed noise ratio was set to -20dB, and the value of a to 1.5. The proposed method, the traditional nonparametric direct estimation method and the nonparametric indirect estimation method were all applied to the tests. The test results are displayed in Figures 1 and 2. As shown in Figures 1 and 2, the proposed method outperformed the traditional methods in signal identification and background noise suppression. In addition, our method managed to preserve the amplitude and phase information of the signal excellently, and achieve the best spectral flatness, laying the basis for the optimal whitening effect.

7. CONCLUSIONS

In practice, most background noises belong to the FLO, which cannot be suppressed well by traditional methods. In this paper, a FLO bi-spectrum estimation method is proposed based on the AR model, and new definitions are given for FLO three order cumulant. The author discussed the determination of AR model parameters, and introduced how to implement the bi-spectrum estimation method based on AR model. Then, a series of tests were performed to verify the correctness of our method. The results show that our method outperformed the traditional approaches in suppressing FLO noise and identifying relevant information of signals.

ACKNOWLEDGMENT

This work is supported by the Research Foundation of health department of Jiangxi Province (Grant No.20183016).

REFERENCES


